

BROOKS-TYPE THEOREMS FOR PAIR-LIST COLORINGS AND LIST HOMOMORPHISMS*

TOMÁS FEDER[†], PAVOL HELL[‡], AND JING HUANG[§]

Abstract. Brooks proved that every connected graph other than a clique or odd cycle can be colored with Δ colors. Erdős, Rubin, and Taylor (and, independently, Vizing) generalized the theorem of Brooks to list colorings, describing all uncolorable connected graphs in which no vertex has a list smaller than its degree. Other authors have extended this to list T -colorings and their generalizations. We further extend it to model pair-list colorings. In addition to including all of the previous situations, pair-list colorings also generalize list homomorphisms (also known as list H -colorings). In the general context of pair-list colorings, we prove a Brooks-type theorem which extends many (but not all) of the existing results. Our result applies to both graphs and digraphs, with or without loops. We discuss several applications of the result, including a polynomial test for the existence of balanced list homomorphisms and retractions.

Key words. coloring, graph homomorphism, list coloring, Brooks theorem, Gallai configuration

AMS subject classifications. Primary, 05C15; Secondary, 05C85

DOI. 10.1137/06066000X

1. Background. The theorem of Brooks [5] states that a connected graph G can be colored by $\Delta(G)$ (the maximum degree) colors, except for cliques and odd cycles, which can be colored by $\Delta(G) + 1$ colors. The theorem of Erdős, Rubin, and Taylor (and, independently, Vizing) [6, 21] generalizes this result to list colorings. In a list coloring problem, each vertex v of G is equipped with a list $L(v)$ of allowed colors that can be used as a color of v , while at the same time seeking a proper coloring in the usual sense (adjacent vertices are to obtain distinct colors). Instead of a condition relating the number of colors to the maximum degree $\Delta(G)$, it is now natural to assume a condition relating the number of colors available at a vertex v to its degree $\deg_G(v)$. We say that G with lists L is *balanced* if each vertex v has $|L(v)| = \deg_G(v)$. Erdős, Rubin, and Taylor described all balanced problems that do not admit list colorings (cf. Corollary 8). Further extension of this idea to list T -colorings and their generalizations was explored in [13, 22]. Let T be a set of nonnegative integers (usually including 0). A T -coloring of G is a mapping $f : V(G) \rightarrow \mathbb{N}$ such that $|f(u) - f(v)| \notin T$ whenever uv is an edge of G . (Note that when $T = \{0\}$, a T -coloring is just a usual coloring.) Given a graph G with lists $L(v) \subseteq \mathbb{N}$, a list T -coloring of G is a T -coloring f of G with all $f(v) \in L(v)$.

List colorings and list T -colorings are naturally generalized to list homomorphisms, also known as list H -colorings. Let H be a fixed graph (or digraph), and let G be a graph (respectively, digraph) together with lists $L(v) \subseteq V(H)$, $v \in V(G)$. A *list H -coloring* of G , with respect to lists L , is a homomorphism (edge-preserving mapping of the vertices) f of G to H such that all $f(v) \in L(v)$. The *list H -coloring*

*Received by the editors May 15, 2006; accepted for publication (in revised form) August 13, 2007; published electronically February 6, 2008.

<http://www.siam.org/journals/sidma/22-1/66000.html>

[†]268 Waverley Street, Palo Alto, CA 94301.

[‡]School of Computing Science, Simon Fraser University, Burnaby, BC, V5A 1S6 Canada (pavol@cs.sfu.ca).

[§]Department of Mathematics and Statistics, University of Victoria, P.O. Box 3045, Victoria, BC, V8W 3P4 Canada (jing@math.uvic.ca).

problem asks whether or not a graph G with lists L admits a list H -coloring. The problem was investigated, and its complexity classified, in [7, 8, 9]; cf. also [17]. When all lists are equal to $V(H)$, the list H -coloring problem is known as the H -coloring problem [15]. We explicitly note that when H is a complete graph (without loops), then list H -colorings become just list colorings. On the other hand, when H is the graph with vertex set $\mathbf{N}$ and with edges ij such that $|i - j| \notin T$, then H -colorings become exactly T -colorings, and list H -colorings become exactly list T -colorings. (Note that if $0 \in T$, then the graph H has no loops.)

In [13], a more general model for list T -colorings was introduced. An instance of the *generalized list T -coloring problem* is a graph G in which each edge uv has its own constraints on the difference of colors of its endpoints, i.e., its own set $T(u, v)$ of integers, requiring that $|f(u) - f(v)| \notin T(u, v)$. The authors gave a result for these colorings which generalizes all of the earlier Brooks-type results [13].

In the same spirit, we introduce the following *pair-list coloring problem*. Let G be a graph and C a set of colors. For each edge $uv \in E(G)$, let the *forbidden set* $T(u, v)$ be a subset of $C \times C$. For consistency, we shall assume that $T(u, v) = \{(a, b) : (b, a) \in T(v, u)\}$. We shall extend the definition of $T(u, v)$ to all ordered pairs of distinct vertices, where $T(u, v) = \emptyset$ if uv is not an edge. Since edges $uv \in E(G)$ having $T(u, v) = \emptyset$ don't restrict our search for a coloring (see below), in what follows we will always assume that $T(u, v) \neq \emptyset$ if (and only if) $uv \in E(G)$. Given a graph G with forbidden sets $T(u, v)$, a *pair-list coloring* is a mapping $f : V(G) \rightarrow C$ such that $uv \in E(G)$ implies $f(u)f(v) \notin T(u, v)$. The *pair-list coloring problem* asks whether or not a given instance G, T has a pair-list coloring.

We shall also assume that each vertex v is equipped with a *list* $L(v) \subseteq C$ in the usual sense; additionally, each pair-list coloring f must have all $f(v) \in L(v)$. If these lists are not specifically given, we may set $L(v)$ equal to the set of all colors not explicitly forbidden for v by the sets $T(v, w), w \in V(G)$. Note that this is not a more general model, as the lists can be eliminated by adjusting the forbidden sets. (We can remove a list $L(v)$ and add to each $T(v, w)$ all pairs (a, b) with $a \notin L(v)$ without changing the problem.) In the rest of the paper, we shall consider an instance of the problem to be a graph G , with lists $L(v), v \in V(G)$, and forbidden sets $T(v, w), v, w \in V(G)$, such that $T(v, w) \neq \emptyset$ if and only if vw is an edge of G . We shall often refer to an instance G, L, T only by its graph G ; for example, we may say the instance is *connected* if the graph G is connected.

This concept is a common extension of (generalized) list T -colorings and list H -colorings of graphs (and digraphs). When $C = \mathbf{N}$ and each set $T(u, v)$ is described by listing the forbidden set D of differences, i.e., $T(u, v) = \{(a, b) : |a - b| \in D\}$, the pair-list coloring problem becomes the list T -coloring problem from [1, 22] (if the set D is fixed) or the generalized list T -coloring problem (if D depends on the edge uv).

Now to explain the connection of pair-list colorings to list H -colorings, let H be a graph. When the forbidden set for each edge $uv \in E(G)$ is defined as $T(u, v) = \{(a, b) : a \in L(u), b \in L(v), ab \notin E(H)\}$, the pair-list coloring problem becomes the list H -coloring problem. The same transformation works also when H is a digraph. In that case, the instance G is also a digraph; however, the directions of its arcs can be ignored by adjusting the forbidden sets so that only homomorphisms preserving the arc directions are allowed by the forbidden sets. (Note that the consistency assumption that $T(u, v) = \{(a, b) : (b, a) \in T(v, u)\}$ does not prevent us from modeling digraphs in this way.) At this point it is convenient to introduce a product graph $G \times T$ with the vertex set $V(G) \times C$, in which $(u, a)(v, b)$ is an edge whenever $(a, b) \in T(u, v)$. In fact, we may restrict our attention to vertices (v, a) with $a \in L(v)$. This product

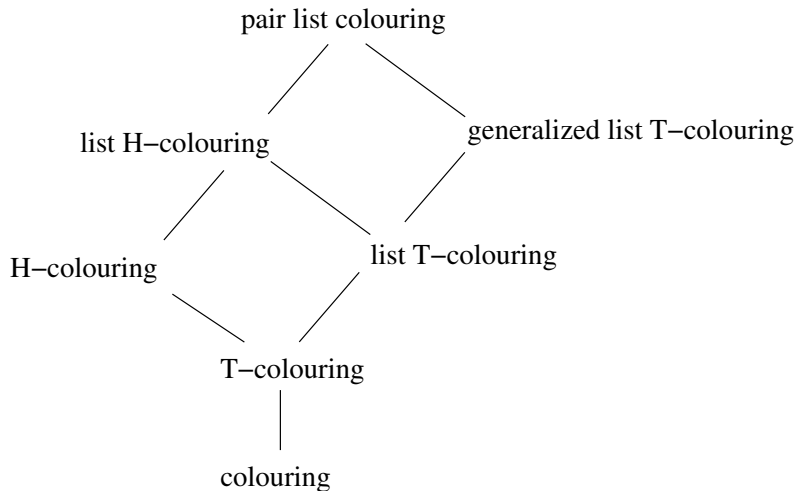


FIG. 1. Versions of coloring problems.

conveniently captures the structure of forbidden color pairs. We depict its edges by dashed lines to emphasize that these are the *forbidden* pairs. Thus a pair list coloring f of G corresponds to a set of vertices $(v, f(v)), v \in V(G)$, one from each set $\{v\} \times L(v)$, which does not include an edge of $G \times T$.

The inclusion relations among the various versions of these coloring problems are illustrated in Figure 1.

Our main result is a Brooks-like characterization of balanced instances which do not admit a pair-list coloring. It implies the theorems of Brooks and Erdős, Rubin, and Taylor, but only a weaker version of the results in [13, 22]. On the other hand, our result also applies to (generalized) list T -colorings in which $0 \notin T$, providing some answers to the questions raised in [13] concerning this case. It seems possible to extend our context to forbidden triples, and higher order tuples, to generalize also the hypergraph Brooks theorems of [18], but we will keep our focus on forbidden pairs.

We shall consider both graphs and digraphs, and allow loops, unless specifically excluded. Graphs and digraphs without loops are called *irreflexive*. Graphs and digraphs in which every vertex has a loop are called *reflexive*. An *oriented graph* is a digraph without opposite arcs uv, vu . Graphs and digraphs have different definitions, and various concepts such as degree are applied differently to them; however, if convenient, we may view graphs as symmetric digraphs (digraphs in which $uv \in E(G)$ if and only if $vu \in E(G)$), as the homomorphism properties under this definition are the same [17]. In this context, it is important that a loop contributes one to the degree of a vertex in the case of graphs, and to both the indegree and the outdegree of a vertex in the case of digraphs.

2. The general result. In this section we deal with the problem in full generality. Note that loops in G have no effect on the problem, and we shall assume that the graph G is irreflexive. Consider an instance of the pair-list coloring problem, i.e., a graph G with lists L and forbidden sets T . Let $N(v)$ denote the set of neighbors of the vertex v in G . We say that v is *pair-balanced* if it satisfies the following two assumptions:

- (i) For each $u \in N(v)$ and each assignment $f(x) \in L(x)$, $x \in N(v) - u$, there

exists a color $a \in L(v)$ with all $(a, f(x)) \notin T(v, x)$. (We say that the assignment f can be *compatibly extended* to v .)

- (ii) If an assignment $f(x) \in L(x)$ for all $x \in N(v)$ cannot be compatibly extended to v , then for each color $a \in L(v)$ there exists a unique $u \in N(v)$ with $(f(u), a) \in T(u, v)$.

Recall the product $G \times T$ defined in the introduction. The T -weight of an ordered pair (u, v) of vertices of G is the maximum number of neighbors of any one (v, b) among all $(u, a), a \in L(u)$, in $G \times T$. Note that if $uv \notin E(G)$, the weight of both pairs $(u, v), (v, u)$ is 0, and that in general the weight of (u, v) is not the same as that of (v, u) . The T -weight of (u, v) represents the maximum number of possible colors for u that can be forbidden by one color of v . The T -degree of u , denoted by $\deg_T(u)$, is the sum over all $v, v \neq u$ of the T -weights of ordered pairs (u, v) . It is easy to see that the following fact holds.

PROPOSITION 1. *A vertex v with $|L(v)| \geq \deg_T(v)$ is pair-balanced.*

A graph G with lists L and forbidden sets T is *pair-balanced* if all its vertices are pair-balanced; it is *pair-overbalanced* if all vertices are pair-balanced, and for at least one vertex v , assumption (ii) can be replaced by the following stronger assumption:

- (ii') Every assignment $f(x) \in L(x), x \in N(v)$, can be compatibly extended to v .

The easy part of Brooks' theorem states that any graph with maximum degree Δ admits a coloring with $\Delta + 1$ colors. The counterpart of this result in our context is as follows.

THEOREM 2. *Every connected pair-overbalanced instance of the pair-list problem has a solution.*

Proof. If (ii') holds for vertex v , we can color G greedily in the order of decreasing distance to v . When coloring a vertex u other than the special vertex v we can use (i); such a vertex u has at least one uncolored neighbor closer to u ; when coloring the special vertex v we can use (ii'). $\square$

There are pair-balanced instances of the pair-list problem which do not have a solution. A connected instance G, L, T of the pair-list problem is a *Gallai configuration* if

1. each block B of G is a clique or a cycle;
2. each list $L(v), v \in V$, can be partitioned into nonempty subsets $L_B(v)$, over all blocks B of G with $v \in B$, so that
3. if block B is a clique with $q + 1$ vertices, then each list $L_B(v)$ can be further partitioned into q nonempty subsets $L_B^1(v), \dots, L_B^q(v)$ such that for $u, v \in B$ with $u \neq v$ we have

$$T(u, v) = \{(a, b) : a \in L_B^i(u), b \in L_B^i(v), i = 1, 2, \dots, q\};$$

4. if block B is a cycle, then each list $L_B(v)$ can be further partitioned into two nonempty subsets $L_B^1(v), L_B^2(v)$ so that for one edge uv of the cycle B we have

$$T(u, v) = \{(a, b) : a \in L_B^i(u), b \in L_B^i(v), i = 1, 2\},$$

and for all other edges uv of B we have

$$T(u, v) = \{(a, b) : a \in L_B^i(u), b \in L_B^{3-i}(v), i = 1, 2\};$$

5. if B, B' are distinct blocks, and $u \in B, v \in B \cap B'$ are distinct vertices, then for any $a \in L_B(u), b \in L_{B'}(v)$ we have $(a, b) \notin T(u, v)$.

These conditions imply a particularly simple structure of the product $G \times T$, illustrated in Figure 2: the vertices (v, a) , $a \in L(v)$, are always grouped close together; as mentioned earlier, the edges of $G \times T$ are dashed. In both illustrations in Figure 2, all parts $L_B^i(v)$ are singletons; if a part has more vertices, the adjacencies are just repeated. As mentioned earlier, a pair-list coloring needs to choose one vertex from each group and not include a dashed edge; clearly this is impossible in both these cases.

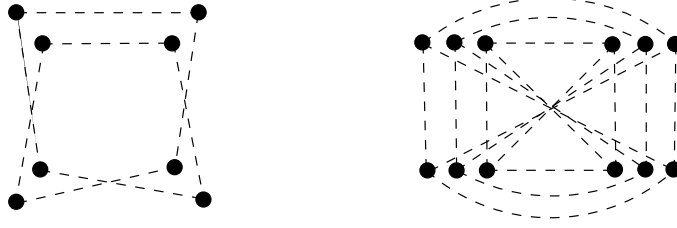


FIG. 2. The product $G \times T$ for a four-cycle and a four-clique block G .

Our main result states that Gallai configurations are the only pair-balanced instances which do not admit a pair-list coloring.

THEOREM 3. *Let G be a connected instance of the pair-list problem. Then the following statements are equivalent:*

- *The instance is pair-balanced but has no solution.*
- *The instance is a Gallai configuration.*

Proof. We first prove that a Gallai configuration is pair-balanced and has no solution. For a vertex which belongs to a unique block of a Gallai configuration, it is easy to verify assumptions (i) and (ii) by considering the above descriptions of the various types of blocks; for a vertex v which is a cutpoint, i.e., belongs to several blocks, we use condition 4 to focus on individual blocks containing v . To show that a Gallai configuration admits no solution, consider the block-cutpoint tree of G : starting with a leaf block B , it is easy to check (in each of the types of blocks described above) that no solution exists using only colors from $L_B(v)$, $v \in B$. (Compare with Figure 2.) Thus the unique cutpoint u of G contained in B must be assigned a color not from $L_B(u)$, in any solution. Since condition 5 ensures that any other color it may be assigned is compatible with other choices made within B , we may remove the entire block B , and delete from $L(u)$ the colors of $L_B(u)$, and proceed recursively. A contradiction is reached when only one block remains, since no assignment within the block is possible.

To prove the converse, we shall use the following three simple lemmas.

LEMMA 4. *Let G, L, T be a pair-balanced instance which has no pair-list coloring. Let u be a vertex of G such that $G - u$ is connected. Then for every neighbor v of u and each $a \in L(u)$ there exists some $b \in L(v)$ with $(a, b) \in T(u, v)$.*

Proof. Suppose v is a vertex of G such that $(a, b) \notin T(u, v)$ for all $b \in L(v)$. Then a pair-list coloring of G exists. We first color u by color a ; since $G - u$ is connected, we can color $G - u$ greedily in the order of decreasing distance to v , taking into account the color of u . When coloring v , we can disregard u by our assumption. $\square$

LEMMA 5. *Let G, L, T be a pair-balanced instance which has no pair-list coloring. Let u, v be two vertices of G such that $G - u - v$ is connected. Then for every vertex w adjacent to both u and v and all $a \in L(u), b \in L(v), c \in L(w)$ with $(a, c) \in T(u, w)$ and $(b, c) \in T(v, w)$, we must also have $(a, b) \in T(u, v)$.*

Proof. If $(a, b) \notin T(u, v)$, then there again is a pair-list coloring of G : color u by a and v by b , and then proceed as in Lemma 4, coloring $G - u - v$ greedily in the order of decreasing distance to w . At the end, w can be colored by (ii), since its neighbors u, v forbid the same color c . $\square$

LEMMA 6. *Let G, L, T be a pair-balanced instance which has no pair-list coloring. Let v be a vertex of G with precisely two neighbors, u and w , and assume that $G - v - w$ is connected. Then for all $a \in L(u), b, b' \in L(v)$, and $c \in L(w)$ with $(b, c) \notin T(v, w), (b', c) \notin T(v, w)$, and $(a, b) \notin T(u, v)$, we must also have $(a, b') \notin T(u, v)$.*

Proof. If $(a, b') \in T(u, v)$, then a pair-list coloring of G can be found as follows. Color w by c and v by b' , and then proceed as in Lemma 4, coloring $G - v - w$ greedily in the order of decreasing distance to u , using (i), until reaching the vertex u . If u cannot be colored, then by assumption (ii) the color a is forbidden only by the color b' on v . Recoloring v by b doesn't create a problem for w and solves the problem for u ; since the degree of v is two, this is a pair-list coloring. $\square$

We now continue proving the theorem; thus assume that we have a connected pair-balanced instance G, L, T which has no solution.

Assume first that G is a block, that is, has no cutpoint. If G has only two vertices, then it is an edge uv with $T(u, v) = L(u) \times L(v)$; this is a Gallai configuration of type 3 with $q = 1$. Thus we assume that G has at least three vertices. Assume further that G has no vertices of degree two. In this case, we first note that there must exist a vertex v such that $G - v$ is also a block. (Indeed, according to a theorem of Kaugars, every critical block other than an edge contains a vertex of degree two; cf. [2, p. 49].) We now show that v is adjacent to every vertex of $G - v$. Otherwise, there would be two vertices u, w in $G - v$ such that v is adjacent to w , w is adjacent to u , but v is not adjacent to u . Then there exist $b \in L(v)$ and $c \in L(w)$ so that $(b, c) \in T(v, w)$, since vw is an edge of G . By Lemma 4, there also exists a color $a \in L(u)$ such that $(a, c) \in T(u, w)$. Since uv is not an edge of G , $(a, b) \notin T(u, v)$, contrary to Lemma 5. This contradiction shows that v is adjacent to all other vertices of G . Since $G - v$ is a block, $G - v'$ must be a block for every other vertex v' in G , and so the preceding argument applies to v' as well; i.e., G is a clique.

We have shown that if G is a block without vertices of degree two, then G is a clique. We now prove that if G is a clique, then it is a Gallai configuration of type 3, i.e., that the lists $L(v)$ can be partitioned as claimed. Consider any three vertices u, v, w of G and $a \in L(u), b \in L(v), c \in L(w)$: if $(a, b) \in T(u, v)$ and $(b, c) \in T(v, w)$, then by Lemma 5 we also have $(a, c) \in T(u, w)$. For $a \in L(u)$, we let $S(v)$ be the set of elements $b \in L(v)$ such that $(a, b) \in T(u, v)$, and we let $S(w)$ be the set of elements $c \in L(w)$ such that $(a, c) \in T(u, w)$. Then $(b, c) \in T(v, w)$ if $b \in S(v), c \in S(w)$, and by symmetry $(b, c) \notin T(v, w)$ if $b \in S(v)$ and $c \in L(w) - S(w)$, or if $b \in L(v) - S(v)$ and $c \in S(w)$. Thus the lists $L(v)$ can each be partitioned into some number q of nonempty subsets $L^1(v), \dots, L^q(v)$ such that $(a, b) \in T(u, v)$ for $a \in L^i(u), b \in L^j(v)$ if and only if $i = j$. Since the graph has no pair-list coloring, it must have at least $q + 1$ vertices. On the other hand, G cannot have more than $q + 1$ vertices, since there is no pair-list coloring, yet all vertices are pair-balanced. Thus G has exactly $q + 1$ vertices, as claimed.

Now we claim that if G (a block pair-balanced instance without a solution) has a vertex v of degree two, then it is a Gallai configuration of type 4. This is clear if G is a three-cycle, by the above argument for cliques. (Note that in this case the partitions of lists in cases 3 and 4 can be relabeled to coincide.) Otherwise, let v be a vertex of degree two, with neighbors u, w in G . We construct a new instance G', L', T'

as follows: the graph G' is obtained from $G - v$ by adding the edge uw if it was not present (it will turn out below that it was not present); the lists L' are the lists L restricted to the vertices of G' , and the forbidden sets T' are the same as T , except for $T'(u, w)$ which is obtained from $T(u, w)$ by adding the set $T^*(u, w)$ consisting of all pairs (a, b) such that $a \in L(u)$, $b \in L(w)$ and each $c \in L(v)$ has $(a, c) \in T(u, v)$ or $(b, c) \in T(w, v)$. It is easy to observe that the instance G', L', T' has no solution, since the definition of $T'(u, w)$ would allow it to be extended to a solution of G, L, T . It is also clear that G' is a block. Moreover, it can be argued that G', L', T' is also pair-balanced. (For instance, it follows from assumption (i) that for every $c \in L(w)$ there exists a $b \in L(v)$ such that $(b, c) \notin T(v, w)$; therefore if $(a, c) \in T^*(u, w)$, then $(a, b) \in T(u, v)$. Thus when proving that u is pair-balanced in G' , we can use b instead of c and appeal to the fact that u is pair-balanced in G .)

Suppose first that G' does not contain a vertex of degree two, and so is a clique (with at least three vertices). Suppose also that some $b, b' \in L(v)$ admit a $c \in L(w)$ with $(b, c) \notin T(v, w)$, $(b', c) \notin T(v, w)$. Then, by Lemma 6, for any $a \in L(u)$, we have $(a, b) \notin T(u, v)$ if and only if $(a, b') \notin T(u, v)$. There exists an $a \in L(u)$ with $(a, b) \notin T(u, v)$ by assumption (i), and thus also $(a, b') \notin T(u, v)$. By a symmetric argument, for any $c \in L(w)$ we have $(b, c) \notin T(v, w)$ if and only if $(b', c) \notin T(v, w)$. It follows from this that there exists a $p \geq 2$ such that each of the lists $L(x)$, $x = u, v, w$, can be partitioned into nonempty disjoint subsets $L^1(x), \dots, L^p(x)$ such that for $a \in L^i(u)$, $b \in L^j(v)$, $c \in L^\ell(w)$ we have $(a, b) \in T(u, v)$ if and only if $i \neq j$, $(b, c) \in T(v, w)$ if and only if $j \neq \ell$, and hence $(a, c) \in T^*(u, w)$ if and only if $i \neq \ell$. This means that the subgraph of $G' \times T'$ induced by $\{u\} \times L(u) \cup \{w\} \times L(w)$ has at least all edges (a, c) , where $a \in L^i(u)$, $c \in L^\ell(w)$, $i \neq \ell$ (maybe more if uw was an edge of G), and hence is connected unless $p = 2$. On the other hand, since G' is a clique, $G' \times T$ has the form dictated by condition 3, in which the subgraph induced by $\{u\} \times L(u) \cup \{w\} \times L(w)$ has $q \geq 2$ components. Thus $p = 2$, and condition 3 implies that also $q = 2$, whence G' is a three-cycle. We now argue that G could not have had an edge uw ; indeed, if $T'(u, w)$ contained more pairs than $T^*(u, w)$, then the subgraph of $G' \times T'$ induced by $\{u\} \times L(u) \cup \{w\} \times L(w)$ would again be connected, contradicting the form of $G' \times T'$ from condition 3, as above. On the other hand, if $T'(u, w) = T(u, w) \cup T^*(u, w) = T^*(u, w)$, then condition (i) would imply that G' admits a pair-list coloring, which is not the case. It is now clear that G is a four-cycle, with the lists partitioned as described in condition 4.

If $G' = G_1$ does contain a vertex of degree two, we construct graphs $G_2, G_3, \dots$, where G_i is obtained from G_{i-1} by deleting a vertex of degree two and inserting an edge joining its two neighbors. (It follows as above that the two neighbors are not adjacent in G_{i-1} , unless G_{i-1} is a three-cycle.) In any event, we finally obtain a graph G_t which is a three-cycle and a Gallai configuration of type 4. Working backwards as above, we see that G_{t-1} is a four-cycle and a Gallai configuration of type 4; in fact, all G_i 's are cycles and Gallai configurations, including $G_1 = G'$, and finally G . This completes the proof in the case when G is a block.

Finally, if G is not a block, then G has a leaf block B with cutpoint $w \in B$. In B , we may assign colors $f(v) \in L(v)$ to $v \neq w$ in the order of decreasing distance from w by our assumption (i). If it is not possible to assign a color $f(w)$, then we let $L_B(w)$ be the set of colors for w forbidden by assigned colors to neighbors v of w in B , and we let $L'(w)$ be the set of colors for w forbidden by assigned colors to neighbors u of w in $G - (B - w)$. By assumption (ii), the sets $L_B(w)$ and $L'(w)$ are disjoint. Both the subproblem on B with $L(w)$ replaced by $L_B(w)$ and the subproblem on $G - (B - w)$

with $L(w)$ replaced by $L'(w)$ satisfy our assumptions (i) and (ii), and the problem on G admits an assignment if and only if either of these two subproblems on B and $G - (B - w)$ has an assignment. Proceeding inductively to decompose the subproblem on $G - (B - w)$, either the problem on G has an assignment, or it decomposes into subproblems for the various blocks B_i , with corresponding decompositions of the lists $L(v)$, so that the original problem has a solution if and only if at least one subproblem for a block B_i has a solution. This proves all the claims, except for the last one, for arbitrary G . For the last claim, if v and w are adjacent in G , with $v \in B$ and $w \in B \cap B'$ for distinct blocks B and B' , and there exist $b \in L_B(v)$ and $c \in L_{B'}(w)$ such that $(b, c) \in T(v, w)$, we proceed as follows. Consider B with lists $L_B(x)$, set $f(v) = b$, and assign $f(x) \in L_B(x)$ for $x \neq w$ in B in order of decreasing distance from w . For each block B'' containing w other than B , assign $f(x) \in L_{B''}(x)$ for $x \neq w$ from B'' in order of decreasing distance from w . The resulting assignments for blocks B'' other than B forbid at least all of $L_{B''}(w)$ for w , and the assignment for block B forbids at least all of $L_B(w)$ as well as $f(w) = c$ for w . Thus there is no possible color for w . Yet $f(w) = c$ is forbidden both by $f(v) = b$ in B and by some $f(u)$ in B' , since $c \in L_{B'}(w)$ and all of $L_{B'}(w)$ is forbidden by neighbors of w in B' . This contradicts assumption (ii). Therefore, there are no $b \in L_B(v)$ and $c \in L_{B'}(w)$ such that $(b, c) \in T(v, w)$. This completes the proof. $\square$

Since checking the block structure of an instance, and seeking the partitions of the simple kind described above, can be accomplished in polynomial time, this represents a polynomial time algorithm for solving the pair-list problem for pair-balanced instances.

3. Balanced list H -coloring problems. Recall that in the list H -coloring problem we have a fixed graph (or digraph) H , and an instance consists of a graph (respectively, digraph) G together with lists $L(v) \subseteq V(H), v \in V(G)$. A solution is a list H -coloring f of G , i.e., a homomorphism of G to H in which all $f(v) \in L(v)$. While we will allow loops in H , it will be simpler to assume that G is irreflexive. (Note that if G is not irreflexive, we can make it so by eliminating a loop at vertex v and at the same time deleting from $L(v)$ any vertices of H which do not have a loop.)

Our Theorem 3 can be applied to list H -colorings by taking each forbidden set $T(u, v)$ to be the set $L(u) \times L(v) - E(H)$. Thus there are many different applications possible. We shall single out just two very simple contexts to illustrate the applications, one for graphs and one for digraphs. In both cases we focus on instances with $\deg_G(v) = |L(v)|$, in analogy with [6]. We shall call such instances *balanced*. (For digraphs, $\deg_G(v)$ is to be interpreted as the total degree of v in G , i.e., the sum of in- and out-degrees of v in G .) A balanced instance G, L is called *overbalanced* if it contains a vertex with degree (respectively, total degree) $\deg_G(v)$, which satisfies $\deg_G(v) > |L(v)|$. The *balanced list H -coloring problem* for a graph H is the restriction of the list H -coloring problem to balanced graphs G with lists L . The *balanced list H -coloring problem* for a digraph H is the restriction of the list H -coloring problem to balanced *oriented* graphs G with lists L . (Recall that an oriented graph is a digraph without a pair of opposite arcs.) We note that our notion of balanced vertices is not the same as that of [13].

In this situation we naturally focus on graphs (respectively, digraphs) H that are *nearly complete*, in the sense that at each vertex v they have $vw \in E(H)$ for all $w \in V(H)$ except possibly one vertex w , which could be equal to v (respectively, $vw \in E(H)$ for all $w \in V(H)$ except possibly one w , and $uv \in E(H)$ for all $u \in V(H)$ except possibly one u , where $u = v = w$ is possible). Thus in a nearly complete

graph, each vertex has at most one edge missing, and that edge could be the loop, and in a nearly complete digraph, each vertex has at most one in-edge and at most one out-edge missing, and if it has the loop missing, it counts as both the missing in-edge and the missing out-edge. The *complement* of H has edges and loops precisely where H does not have edges and loops. (Normally complements are defined without reference to loops.) Note that the usual irreflexive complete graph is nearly complete in this sense, as it is missing the loop at each vertex. Some other nearly complete graphs and digraphs are depicted in Figure 3. Recall that when we model the list H -coloring problem as a pair-list coloring problem, the forbidden sets $T(u, v)$ for $uv \in E(G)$ are those pairs (a, b) , $a \in L(u)$, $b \in L(v)$ which have $ab \notin E(H)$.

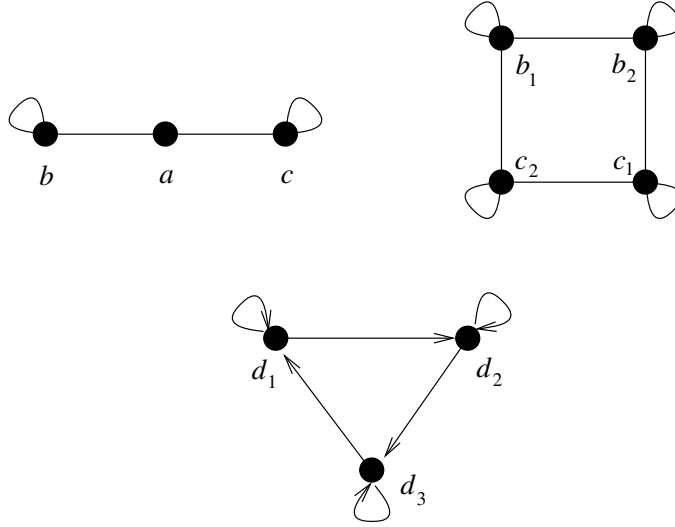


FIG. 3. Some examples of nearly complete graphs and digraphs.

Recall the definition of the T -degree of a vertex v . Since the T -weight of any pair (u, v) with $uv \in E(G)$ (or, in the case of digraphs, $uv \in E(G)$ or $vu \in E(G)$) is at most 1, the T -degree of v is at most $\deg_G(v)$. Therefore Proposition 1 ensures that a balanced graph (or oriented graph) G with lists L is pair-balanced, and hence Theorems 2 and 3 apply to guarantee that connected overbalanced instances always have a solution (list H -coloring) and to describe all connected unsolvable balanced instances as Gallai configurations. It remains to consider which of the Gallai configurations are actually possible. It is immediately obvious that all parts $L_B^i(v)$ must have a single element, since otherwise there would exist two distinct pairs $(a, b), (a', b) \in T(v, w)$.

First, we focus on the balanced list H -coloring problem for nearly complete graphs H . Note that in this situation colors $a, b \in V(H)$ either satisfy $(a, b) \notin T(u, v)$ for all edges $uv \in E(H)$ or satisfy $(a, b) \in T(u, v)$ for all edges $uv \in E(H)$. Consider a block B of a Gallai configuration, which is a clique of $q + 1 \geq 3$ vertices, each with a list of q colors. Suppose u, v, w are distinct vertices with $L_B^i(u) = \{a\}, L_B^i(v) = \{b\}, L_B^i(w) = \{c\}$. Then $(a, b) \in T(u, v), (a, c) \in T(u, w)$, i.e., $ab, ac \notin E(H)$, whence $b = c$ since H is nearly complete. The same argument shows that all sets $L_B(v), v \in B$, are the same set, L_B , of q vertices of H . Consider a block B which is a cycle $v_1, v_2, \dots, v_k$, and suppose the list of each v_i consists of $a_i \in L_B^1(v_i)$ and $b_i \in L_B^2(v_i)$, where $(a_i, a_{i+1}) \in T(v_i, v_{i+1})$ and $(b_i, b_{i+1}) \in T(v_i, v_{i+1})$ for $i = 1, 2, \dots, k-1$, and $(a_k, b_1) \in T(v_k, v_1), (b_k, a_1) \in T(v_k, v_1)$. We again apply the fact that H is nearly complete to

deduce that $a_1 = a_3 = a_5 = \dots$, $b_1 = b_3 = b_5 = \dots$, $a_2 = a_4 = \dots$, $b_2 = b_4 = \dots$. If k is even, $a_k = a_2$, $b_k = b_2$ and so both (a_1, a_2) and (a_1, b_2) are in $T(v_1, v_2)$, which is impossible in a Gallai configuration. Hence k is odd, and therefore all lists consist of the same pair of vertices in H . Thus, for nearly complete graphs H , a block of size at least three has the lists of all its vertices equal to the same set L_B , and all cycle blocks are odd. The vertices of H that appear in these blocks must be missing only their loops. On the other hand, the blocks that have two vertices could have equal lists, corresponding to missing loops in H , or they could have different lists, corresponding to missing nonloop edges in H . Summarizing this additional information, we have the following result for graphs H (with loops allowed).

COROLLARY 7. *Suppose H is a nearly complete graph, whose complement has isolated loops a_i , $i = 1, 2, \dots, k$, and isolated edges $b_i c_i$, $i = 1, 2, \dots, \ell$.*

If G is a balanced connected irreflexive graph with lists L , then either G has a list H -coloring, or it is a Gallai configuration in which

- *each block B that is a clique with $q+1 \geq 3$ vertices has all lists $L_B(v)$, $v \in B$, equal to the same set of q vertices a_i ;*
- *each block B that is a cycle has odd length and all the lists $L_B(v)$, $v \in B$, are equal to the same pair of vertices a_i, a_j ;*
- *each block B that is an edge $uv \in E(B)$ (clique with two vertices) has either $L_B(u) = L_B(v) = \{a_i\}$ for some $i = 1, 2, \dots, k$, or has $L_B(u) = \{b_j\}$, $L_B(v) = \{c_j\}$ for some $j = 1, 2, \dots, \ell$.*

Note that vertices of H that have no missing incident edges play no role in the result.

We single out two important special cases of the corollary. When H is irreflexive, we have $\ell = 0$, i.e., H is an irreflexive complete graph, and we obtain the following theorem of Erdős, Rubin, and Taylor [6]; also see Vizing [21] (for other proofs see [4, 13, 18]).

COROLLARY 8. *If G is a connected balanced irreflexive graph with lists $L(v)$, $v \in V(G)$, then either G admits a list coloring, or it is a Gallai configuration in which each block is either a clique with $q+1$ vertices, all with the same list of size q , or an odd cycle with all vertices having the same list of size two.*

The only difference from the above is that the third item on cliques of size $q+1$ also applies to $q+1 = 2$; i.e., the last item is not needed. ($L_B(u)$ and $L_B(v)$ are equal.) A Gallai configuration for this case is illustrated in Figure 4.

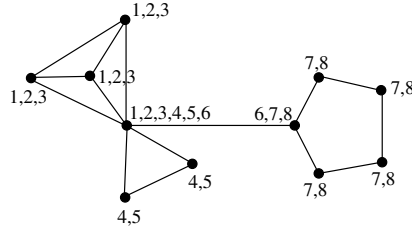


FIG. 4. A Gallai configuration for the irreflexive complete graph H .

When H is reflexive, we have $k = 0$, and we obtain the following new result.

COROLLARY 9. *Suppose H is a reflexive graph whose complement has isolated edges $b_i c_i$, $i = 1, 2, \dots, \ell$.*

If G is a balanced connected irreflexive graph with lists $L(v) \subseteq V(H)$, $v \in V(G)$, then either G admits a list H -coloring, or it is a Gallai configuration that is a tree,

in which each edge block $B = uv$ has $L_B(u), L_B(v)$ equal to $\{b_j\}, \{c_j\}$ for some j .

We illustrate a Gallai configuration for this case in Figure 5. The graph H is the reflexive four-cycle from Figure 3. Each edge uv of the tree is a block of the Gallai configuration and has $L_B(u), L_B(v)$ chosen as either b_1, c_1 or b_2, c_2 , in some order, labeled on the edge with lowercase letters. The lists of the leaves are marked, and all other vertices have the full lists $\{b_1, b_2, c_1, c_2\}$. It is instructive to see that there is no list H -coloring.

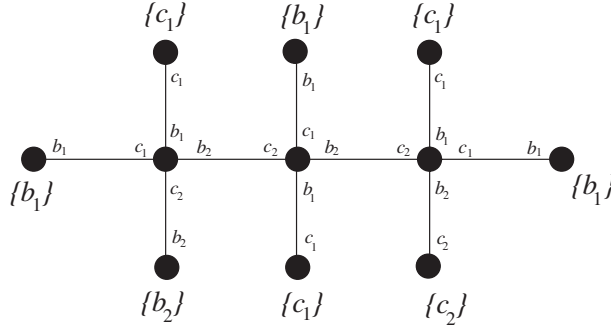


FIG. 5. A Gallai configuration for the reflexive four-cycle H .

We now briefly turn to the balanced list H -coloring problem for nearly complete digraphs H . In this situation colors $a, b \in V(H)$ may have $(a, b) \notin T(u, v)$ if uv is an edge, but not if vu is an edge, of H . However, it is still the case that (a, b) is either in $T(u, v)$ for all edges $uv \in E(H)$ or not in $T(u, v)$ for all edges $uv \in E(H)$. The instance of the balanced H -coloring problem consists of an oriented graph G with lists $L(v), v \in V(H)$, and total degrees $\deg_G(v) = |L(v)|$ for all $v \in V(H)$. Recall that to model this by a pair-list coloring problem, we remove the orientations of the edges of G and set, for each edge $uv \in E(H)$, the forbidden set $T(u, v)$ to consist of those pairs $(a, b), a \in L(u), b \in L(v)$ for which $ab \notin E(H)$. Thus the orientations of the edges of G can be ignored because they are encoded in the pair lists.

Consider a block B of a Gallai configuration containing two edges uv, uw with $v \neq w$. Since H is nearly complete, it follows as above that the lists $L_B(v)$ and $L_B(w)$ must be the same. This is easily seen to imply that, for a block B which is a clique with $q + 1 \geq 4$ vertices, each vertex has the same list q colors, and each color is a vertex of H that has no loop. A block B that is a clique with three vertices either has each vertex v with the same lists $L_B(v)$ consisting of two vertices of H without loops, or is a directed three-cycle with distinct lists corresponding to two distinct directed three-cycles in the complement of H . Consider a block B that is a cycle of G ; it may be a directed k -cycle with k even, with lists that correspond to a directed $2k$ -cycle in H ; or it may be a directed k -cycle with k odd, with lists that correspond to two disjoint k -cycles in H . The cycle B could also be oriented in more complex ways, implying certain equalities of lists; once two consecutive vertices of the cycle B are to have the same lists, all vertices of B must have the same lists. It is not hard to see that this can occur only when k is odd. Instead of stating all these necessary and sufficient conditions in such great detail, we only summarize a few necessary conditions.

COROLLARY 10. *Suppose H is a nearly complete digraph, whose complement $\overline{H}$ consists of disjoint directed cycles $D_1, D_2, \dots, D_k$ (some of which may have length one).*

If G is a balanced connected irreflexive oriented graph with lists L that does not

admit a list H -coloring, then it is a Gallai configuration in which

- each block B that is a clique with $q+1 \geq 4$ vertices has all lists $L_B(v), v \in B$, equal to the same set of q vertices lying in cycles of length one in $\overline{H}$;
- each block B that is a clique with three vertices either has all lists $L_B(v), v \in B$, equal to the same set of two vertices lying in cycles of length one in $\overline{H}$, or is a directed three-cycle of G , and its vertices have lists formed from two three-cycles of $\overline{H}$;
- each block B that is an oriented cycle has equal lists $L_B(u) = L_B(v)$ on vertices u, v that admit a $w \in B$ with $uw, vw \in E(B)$ or with $wu, wv \in E(B)$.

Thus if H is a nearly complete graph or digraph, we have solved the balanced list H -coloring problem in polynomial time.

4. Extension problems. In a future paper [11], we shall apply our results to extension problems. Let H be a fixed graph or digraph. Given a graph (respectively, oriented graph) G and a partial assignment f of some vertices of G to vertices of H , a homomorphism F of G to H is an *extension* of f if $F(v) = f(v)$ for all vertices for which $f(v)$ is defined. The *H -extension problem* asks whether or not an instance G, f has an extension. This problem is also known as the *precoloring extension problem*, the *one-or-all list H -coloring problem*, or the *H -retraction problem* [17]. Classifying the complexity of H -extension problems appears to be much harder than for list H -coloring problems [9, 12].

It is easy to see that the H -extension problem is a restriction of the list H -coloring problem to instances G, L in which each $L(v)$ either is equal to $V(H)$ or consists of a single color (vertex of H). If, moreover, the degrees of G are at most $|V(H)|$, we have basically a balanced instance, since a vertex v with $\deg_G(v) = k$ and $L(v) = \{a\}$ can be replaced by k vertices of degree one, with the same list, without changing the problem. Let us say that a graph is *k -bounded* if its maximum degree is at most k . Thus if H is a nearly complete graph or digraph with k vertices, then the H -extension problem for k -bounded graphs (respectively, oriented graphs) can be solved in polynomial time.

In [11], we shall also prove the following fact.

THEOREM 11. *Let H be a graph (with loops allowed) such that any two vertices a, b have a common neighbor c in H . Then the H -extension problem for 3-bounded irreflexive graphs can be solved in polynomial time.*

If H is a reflexive graph of diameter at most two, then it satisfies the condition. In particular, this implies that the reflexive five-cycle H has a polynomial time solvable H -extension problem for 3-bounded graphs. For the reflexive four-cycle H (which is a nearly complete graph; see Figure 3), Corollary 9 implies that the H -extension problem is polynomial time solvable for any 4-bounded graph. Finally, for the reflexive three-cycle H , any mapping of any graph G to H is a homomorphism, so the H -extension problem is trivial for all graphs. We shall prove in [11] that for all other reflexive cycles H , the H -extension problem is NP-complete, even for 3-bounded graphs; we shall also show that for the reflexive five-cycle H the H -extension problem is NP-complete. Summarizing this information, we conclude that for the reflexive k -cycle H , the H -extension problem in general is polynomial only when $k = 3$ (unless $P = NP$), but it also becomes polynomial when $k = 4$ for 4-bounded graphs and additionally when $k = 5$ for 3-bounded graphs. Similarly, it will turn out that for the irreflexive k -cycle H , the H -extension problem in general is polynomial only when $k = 4$ or $k = 6$ (unless $P = NP$), but also becomes polynomial when $k = 3$ for 3-bounded graphs [11]. This kind of sensitivity to degree constraints is unusual for

list H -coloring problems in general; in [10] we have proved that the complexity classification of list H -coloring problems (as NP-complete or polynomial time solvable) does not change when the instances are restricted to k -bounded graphs for any $k \geq 3$.

Consider now instances in which the graph G is allowed to have loops. For reflexive graphs H of diameter at most two, our polynomial time algorithm will still apply. The condition in Theorem 11 is also satisfied if H is an irreflexive graph of diameter at most two, all of whose edges belong to triangles, and the algorithm also applies to graphs G with loops allowed. For general graphs H (with loops allowed), we shall also require, in addition to having, for any two vertices a, b , a common neighbor c , that each vertex $a \in V(H)$ have a neighbor in H that has a loop. Then the H -extension problem for 3-bounded graphs G with loops allowed also has a polynomial time algorithm [11]. In all cases, the algorithms proceed by exploring some Gallai-type configurations that may or may not turn out to be uncolorable [11].

Acknowledgment. We are grateful to an anonymous referee whose uncompromising standards led to an improved presentation of several arguments.

REFERENCES

- [1] N. ALON AND A. ZAKS, *T-choosability in graphs*, Discrete Appl. Math., 82 (1998), pp. 1–13.
- [2] B. BOLLOBÁS, *Extremal Graph Theory*, Academic Press, New York, 1978.
- [3] O. V. BORODIN, *Criterion of chromaticity of a degree prescription*, in Abstracts of the 4th All-Union Conference on Theoretical Cybernetics, Novosibirsk, 1977, pp. 127–128.
- [4] M. BOROWIECKI, E. DRGAS-BURCHARDT, AND P. MIHÓK, *Generalized list colorings of graphs*, Discuss. Math. Graph Theory, 15 (1995), pp. 185–193.
- [5] R. L. BROOKS, *On coloring of nodes of a network*, Proc. Cambridge Philos. Soc., 37 (1941), pp. 194–197.
- [6] P. ERDŐS, A. L. RUBIN, AND H. TAYLOR, *Choosability in graphs*, in Proceedings of the West Coast Conference on Combinatorics, Graph Theory, and Computing, Congress. Numer. XXVI, Utilitas Math., Winnipeg, MB, 1980, pp. 125–157.
- [7] T. FEDER AND P. HELL, *List homomorphisms to reflexive graphs*, J. Combin. Theory Ser. B, 72 (1998), pp. 236–250.
- [8] T. FEDER, P. HELL, AND J. HUANG, *List homomorphisms and circular arc graphs*, Combinatorica, 19 (1999), pp. 487–505.
- [9] T. FEDER, P. HELL, AND J. HUANG, *Bi-arc graphs and the complexity of list homomorphisms*, J. Graph Theory, 42 (2003), pp. 61–80.
- [10] T. FEDER, P. HELL, AND J. HUANG, *List homomorphisms of graphs with bounded degrees*, Discrete Math., 307 (2007), pp. 386–392.
- [11] T. FEDER, P. HELL, AND J. HUANG, *Extension problems with degree bounds*, submitted.
- [12] T. FEDER AND M. Y. VARDI, *The computational structure of monotone monadic SNP and constraint satisfaction: A study through datalog and group theory*, SIAM J. Comput., 28 (1998), pp. 57–104.
- [13] J. FIALA, D. KRÁL', AND R. ŠKREKOVSKI, *A Brooks-type theorem for the generalized list T-coloring*, SIAM J. Discrete Math., 19 (2005), pp. 588–609.
- [14] A. GALLUCCIO, P. HELL, AND J. NEŠETŘIL, *The complexity of H-coloring of bounded degree graphs*, Discrete Math., 222 (2000), pp. 101–109.
- [15] P. HELL AND J. NEŠETŘIL, *On the complexity of H-coloring*, J. Combin. Theory Ser. B, 48 (1990), pp. 92–110.
- [16] P. HELL AND J. NEŠETŘIL, *Counting list homomorphisms and graphs with bounded degrees*, in Graphs, Morphisms and Statistical Physics, J. Nešetřil and P. Winkler, eds., DIMACS Ser. Discrete Math. Theoret. Comput. Sci. 63, AMS, Providence, RI, 2004, pp. 105–112.
- [17] P. HELL AND J. NEŠETŘIL, *Graphs and Homomorphisms*, Oxford University Press, Oxford, UK, 2004.
- [18] A. V. KOSTOCHKA, M. STIEBITZ, AND B. WIRTH, *The color theorems of Brooks and Gallai extended*, Discrete Math., 162 (1996), pp. 299–303.
- [19] J. KRATOCHVÍL AND A. SEBŐ, *Coloring precolored perfect graphs*, J. Graph Theory, 25 (1997), pp. 207–215.

- [20] J. KRATOCHVÍL AND ZS. TUZA, *Algorithmic complexity of list colorings*, Discrete Appl. Math., 50 (1994), pp. 297–302.
- [21] V. G. VIZING, *Coloring the vertices of a graph in prescribed colors*, Diskret. Analiz., 29 (1976), pp. 3–10 (in Russian).
- [22] A. WALLER, *An upper bound for list T -colorings*, Bull. London Math. Soc., 28 (1996), pp. 337–342.

FACES OF GENERALIZED CLUSTER COMPLEXES AND NONCROSSING PARTITIONS*

ELENI TZANAKI†

Abstract. Let Φ be a finite root system with corresponding reflection group W , and let m be a nonnegative integer. We consider the generalized cluster complex $\Delta^m(\Phi)$ defined by Fomin and Reading and the poset $NC_{(m)}(W)$ of m -divisible noncrossing partitions defined by Armstrong. We give a characterization of the faces of $\Delta^m(\Phi)$ in terms of $NC_{(m)}(W)$, generalizing that of Brady and Watt given in the case $m = 1$. Making use of this, we give a case free proof of a conjecture of Chapoton and Armstrong, which relates a certain refined face count of $\Delta^m(\Phi)$ with the Möbius function of $NC_{(m)}(W)$.

Key words. Coxeter group, cluster complex

AMS subject classifications. Primary, 20F55; Secondary, 05E99, 55U10

DOI. 10.1137/060665683

1. Introduction. Let Φ be a finite root system of rank n with associated reflection group W . Let Φ^+ be a positive system for Φ with corresponding simple system Π . Motivated by their theory of cluster algebras [16], Fomin and Zelevinsky introduced the cluster complex $\Delta(\Phi)$ in [15]. This is a pure $(n-1)$ -dimensional simplicial complex on the vertex set $\Phi^+ \cup (-\Pi)$ which is homeomorphic to a sphere [15]. Later, Fomin and Reading [14] introduced the generalized cluster complex $\Delta^m(\Phi)$, where m is any nonnegative integer (see also [20]). This is a simplicial complex on the vertex set of *colored almost positive roots* $\Phi_{\geq -1}^m$, that is, the set consisting of m (colored) copies of each positive root and one copy of each negative simple root. In the case $m = 1$, the generalized cluster complex $\Delta^m(\Phi)$ reduces to $\Delta(\Phi)$. The complex $\Delta^m(\Phi)$ has remarkable properties and surprising connections with other combinatorial objects such as m -divisible noncrossing partitions $NC_{(m)}(W)$ [1, 17, 18] and Catalan hyperplane arrangements [2, 3, 7]. For instance, it is proved in [14, Proposition 8.4] that, if Φ is irreducible, the number of facets of $\Delta^m(\Phi)$ is equal to the generalized Catalan number $N_m(\Phi) = \prod_{i=1}^n \frac{e_i + mh + 1}{e_i + 1}$ defined in [2]. Moreover, the entries of the f and h vectors of $\Delta^m(\Phi)$, as well as those of the natural subcomplex $\Delta_+^m(\Phi)$ called its *positive part*, have many interesting combinatorial interpretations [3, 7].

In his thesis [1] Armstrong defined the poset $NC_{(m)}(\gamma)$, where γ is a Coxeter element of W . There, he proved that the isomorphism type of the poset $NC_{(m)}(\gamma)$ is independent of the Coxeter element γ . We denote this poset by $NC_{(m)}(W)$ when the choice of γ is irrelevant and call it the poset of *m -divisible noncrossing partitions*. The poset $NC_{(m)}(W)$ is a graded meet-semilattice, and it reduces to the lattice of noncrossing partitions NC_W [9, 11] associated to W in the case $m = 1$. If W is irreducible, the number of elements of $NC_{(m)}(W)$ is equal to $N_m(\Phi)$ and the h -polynomial of $\Delta^m(\Phi)$ is equal to the rank-generating polynomial of $NC_{(m)}(W)$ [1, 14].

It was shown in [12, section 8] that the cluster complex $\Delta(\Phi)$ arises naturally in the context of the lattice NC_W . Our first main result is to extend this connection

*Received by the editors July 21, 2006; accepted for publication (in revised form) August 27, 2007; published electronically February 6, 2008. The present research will be part of the author's Ph.D. thesis at the University of Crete.

<http://www.siam.org/journals/sidma/22-1/66568.html>

†Department of Mathematics, University of Crete, 71409 Heraklion, Crete, Greece (etzanaki@math.uoc.gr).

by means of a new characterization of the faces of $\Delta^m(\Phi)$ in terms of m -divisible noncrossing partitions. More precisely, let Φ be an irreducible root system, and let $\Pi = \Pi_+ \cup \Pi_-$ be a partition of the set of simple roots Π into two disjoint sets such that the roots within each are pairwise orthogonal. Let γ be a bipartite Coxeter element with respect to this partition of Π (see (5) in section 2). Consider a face σ of $\Delta^m(\Phi)$, and let σ^i be the subset of σ consisting of positive roots of color i and $\sigma_\pm = \sigma \cap (-\Pi_\pm)$. If $\tau \subseteq \Phi_{\geq -1}^m$ such that either $\tau \subseteq (-\Pi)$ or τ consists of positive roots of the same color, we denote by w_τ the product of reflections through the roots in τ taken in a certain order (see (7) in section 2). We make the convention that $w_\emptyset = \mathbf{1}$, where $\mathbf{1}$ is the identity in W . The faces of $\Delta^m(\Phi)$ can be characterized by the following criterion.

THEOREM 1.1. *For $m \geq 2$, the set $\sigma \subseteq \Phi_{\geq -1}^m$ is a face of $\Delta^m(\Phi)$ if and only if the sequence $(w_{\sigma_+} w_{\sigma^m}, w_{\sigma^{m-1}}, \dots, w_{\sigma^2}, w_{\sigma^1} w_{\sigma_-})$ is an element of $NC_{(m)}(\gamma)$ of rank $|\sigma|$.*

As will be made clear by the restatement of Theorem 1.1 in section 3 (see Theorem 3.1), the above criterion generalizes the characterization of Brady and Watt [12, section 8] in the case $m = 1$. Moreover, making use of this criterion one may discover many interesting properties of $\Delta^m(\Phi)$, for instance, that $\Delta^m(\Phi)$ is shellable and $(m+1)$ -Cohen–Macaulay. This is discussed in the article [8].

Our second main result is the proof of a conjecture that relates the cluster complex $\Delta^m(\Phi)$ with the poset $NC_{(m)}(W)$. Chapoton [13] conjectured a surprising enumerative relation between a refined face count of $\Delta(\Phi)$ and the Möbius function of NC_W . This conjecture, which was proved by Athanasiadis [4], can be stated in the $m \geq 1$ case as follows [1] (see also [17, 18]). The F -triangle for $\Delta^m(\Phi)$ is defined by the generating function

$$(1) \quad F_\Phi^{(m)}(x, y) = \sum_{k=0}^n \sum_{l=0}^n f_{k,l}(\Phi, m) x^k y^l,$$

where $f_{k,l}(\Phi, m)$ is the number of faces of $\Delta^m(\Phi)$ consisting of k colored positive roots and l negative simple roots. The M -triangle for $NC_{(m)}(W)$ is defined similarly as

$$(2) \quad M_W^{(m)}(x, y) = \sum_{\substack{a \leq b \\ \text{in } NC_{(m)}(W)}} \mu(a, b) x^{rk(b) - rk(a)} y^{rk(a)},$$

where $\leq$ denotes the order relation in the poset $NC_{(m)}(W)$, μ stands for its Möbius function, and $rk(a)$ is the rank of $a \in NC_{(m)}(W)$.

The following relation was formulated by Chapoton [13, Conjecture 1] as a conjecture in the case $m = 1$ and was restated by Armstrong for any $m \geq 1$.

THEOREM 1.2. *Let Φ be a finite root system of rank n with corresponding reflection group W , and let m be a nonnegative integer. The F -triangle for $\Delta^m(\Phi)$ and the M -triangle for $NC_{(m)}(W)$ are related by the equality*

$$(3) \quad (1-y)^n F_\Phi^{(m)}\left(\frac{x+y}{1-y}, \frac{y}{1-y}\right) = M_W^{(m)}(-x, -y/x).$$

Theorem 1.2 has been proved in part by Krattenthaler [17], in a case-by-case fashion, for all finite root systems when $m = 1$ and for those that do not contain a copy of D_k for any $k \geq 4$ in the case $m \geq 2$. It was also observed by Krattenthaler that the relation (3) implies the following interesting reciprocity [18, Theorem 8]:

$$y^n M_W^{(-m)}(xy, 1/y) = M_W^{(m)}(x, y),$$

for which there is no intrinsic explanation yet.

This paper is organized as follows. After providing the necessary background, we proceed with the proof of Theorem 1.1 in section 3. In section 4 we find an EL-labeling of the poset $NC_{(m)}(\gamma)$, in which the decreasing chains are in bijection with facets of $\Delta_+^m(\Phi)$. Given the above results, we conclude with the proof of Theorem 1.2 in section 5, generalizing that of Athanasiadis [4] for the case $m = 1$.

2. Preliminaries. In this section we introduce our main objects of study and state a few lemmas and theorems necessary in later sections.

The lattice NC_W and the reflection length order on W . Let W be a finite Coxeter group of rank n , and let T be the set of all reflections in W . The group W is generated by T , and one can define a length function l_T on W so that $l_T(w)$ is the smallest k such that w can be written as a product of k reflections in T . The set W can be partially ordered by letting

$$u \leq v \quad \text{if and only if} \quad l_T(u) + l_T(u^{-1}v) = l_T(v),$$

in other words, if there exists a shortest factorization of u into reflections in T which is a prefix of such a shortest factorization of v . Note that l_T is not the usual Coxeter length function associated to a simple system for W . The order $\leq$ is called *reflection length order* or *absolute order* and turns W into a graded poset having the identity of W as its unique minimal element and rank function l_T . If $\Phi = A_{n-1}$, so that W can be realized as the symmetric group of permutations of the set $\{1, 2, \dots, n\}$, then $l_T(w) = n - c(w)$, where $c(w)$ is the number of cycles of w . Moreover, $u \leq w$ if and only if each cycle of u can be obtained from some cycle of w by deleting elements.

Note that all Coxeter elements are maximal in this partial order of W . The group W acting by conjugation gives automorphisms of this partial order, since the set of reflections is stable under conjugation. We define the *noncrossing partition lattice* NC_W to be an interval between the identity 1 and any Coxeter element γ . Since all Coxeter elements are conjugate, the isomorphism type of NC_W is independent of γ . The poset NC_W is a self-dual lattice of rank n [12]. If we have chosen some particular Coxeter element γ , then we denote the noncrossing partition lattice by $NC_W(\gamma)$. Should there be no element of confusion, we just write $NC(\gamma)$, suppressing W in the notation.

For later reference we summarize a few definitions and general facts.

DEFINITION 2.1 (see [1, Definition 3.1.1]). *The m -tuple $(w_1, \dots, w_m) \in W^m$ is a minimal factorization of $w \in W$ if $w = w_1 w_2 \dots w_m$ and $l_T(w) = \sum_{i=1}^m l_T(w_i)$.*

LEMMA 2.2.

- (i) *Conjugate elements have the same length.*
- (ii) [1, Lemma 3.1.2] *If $(w_1, w_2, \dots, w_k)$ is a minimal factorization of $w \in W$, then so is the k -tuple*

$$(w_i, w_i^{-1} w_1 w_i, \dots, w_i^{-1} w_{i-1} w_i, w_{i+1}, \dots, w_k)$$

for every $1 \leq i \leq k$. In other words, for every w_i there is a minimal factorization of w with w_i in the first place.

- (iii) *If $(w_1, w_2, \dots, w_k)$ is a minimal factorization of w , then $w_{i_1} w_{i_2} \dots w_{i_\ell} \leq w$ and $l_T(w_{i_1} \dots w_{i_\ell}) = \sum_{j=1}^\ell l_T(w_{i_j})$ for every $1 \leq i_1 < i_2 < \dots < i_\ell \leq k$.*

Proof. Part (i) follows from the fact that all reflections in W are conjugate, and then part (ii) is immediate. To prove (iii) it suffices to apply (ii) repeatedly for every w_{i_j} , $1 \leq j \leq \ell$, to get a minimal factorization of w starting with $w_{i_1} w_{i_2} \dots w_{i_\ell}$. $\square$

LEMMA 2.3 (see [4, Lemma 2.1]). *Let a, b, w be elements of W .*

- (i) *$a \leq aw \leq b$ if and only if $w \leq a^{-1}b \leq b$.*
- (ii) *$a \leq aw \leq b$ if and only if $a \leq bw^{-1} \leq b$.*
- (iii) *$a \leq b$ if and only if $a^{-1}b \leq b$, and, in that case, the interval $[a, b]$ is isomorphic to $[1, a^{-1}b]$.*

LEMMA 2.4.

- (i) [6, Lemma 2.1(iv)] *If $a, b \leq c \leq w$ for some $w \in W$ and $ab \leq w$, then $ab \leq c$.*
- (ii) [12, Relation (3)] *If $a \leq b \leq c$, then $a^{-1}b \leq a^{-1}c$ and $ba^{-1} \leq ca^{-1}$.*
- (iii) [12, Relation (7)] *For distinct reflections t_1, t_2 and $w \in W$ we have $t_1 t_2 \leq w \Leftrightarrow t_2 \leq t_1 w \Leftrightarrow t_1 \leq w t_2$.*

The following simple observation is immediate from the definition of the partial order $\leq$ of W and Lemma 2.2(i).

LEMMA 2.5. *For $u, v, w \in W$ we have $u \leq v$ if and only if $wuw^{-1} \leq wvw^{-1}$.*

Posets and EL-labelings. Let $(P, \leq)$ be a finite graded poset. We say that y covers x and write $x \rightarrow y$ if $x < y$ and $x < z \leq y$ holds only for $z = y$. Let $\mathcal{E}(P)$ be the set of covering relations in P and consider Λ a totally ordered set. An *edge labeling* of P with label set Λ is a map $\lambda : \mathcal{E}(P) \rightarrow \Lambda$. If $C : x_0 \rightarrow x_1 \rightarrow \cdots \rightarrow x_r$ is an unrefinable chain, we let $\lambda(C) = (\lambda(x_0 \rightarrow x_1), \lambda(x_1 \rightarrow x_2), \dots, \lambda(x_{r-1} \rightarrow x_r))$ be the label of C with respect to λ , and we call C *rising or decreasing* with respect to λ if the entries of $\lambda(C)$ strictly increase or weakly decrease, respectively, in the total order of Λ . Since P is graded, all maximal chains C in every interval $I = [x, y]$ have the same length, equal to $rk(y) - rk(x)$. The edge labeling allows us to order the maximal chains in I by ordering the labels $\lambda(C)$ lexicographically. That is, C is lexicographically smaller than C' if $\lambda(C)$ precedes $\lambda(C')$ in the lexicographic order induced by the total order of Λ .

DEFINITION 2.6 (see [10]). *An edge labeling λ of P is called EL-labeling if for every nonsingleton interval $[x, y]$ in P ,*

- (i) *there is a unique rising maximal chain in $[x, y]$ and*
- (ii) *this chain is lexicographically smallest among all maximal chains in $[x, y]$ with respect to λ .*

If, in addition, P has a unique minimal and maximal element, then we say that P is EL-shellable.

If $P_1, \dots, P_m$ are posets, then their direct product $P = P_1 \times \cdots \times P_m$ is the poset on the set $\{(x_1, \dots, x_m) : x_i \in P_i\}$ such that $(x_1, \dots, x_m) \leq (x'_1, \dots, x'_m)$ in P if $x_i \leq x'_i$ in P_i for every $1 \leq i \leq m$. Moreover $(x'_1, \dots, x'_m)$ covers $(x_1, \dots, x_m)$ in P if there is some $1 \leq i_0 \leq m$ such that $x_i = x'_i$ for all $1 \leq i \neq i_0 \leq m$ and x'_{i_0} covers x_{i_0} in P_{i_0} . The following lemma is a reformulation of [10, Theorem 4.3].

LEMMA 2.7 (see [10]). *If $P_1, P_2, \dots, P_m$ are graded posets that admit an EL-labeling, then $P = P_1 \times P_2 \times \cdots \times P_m$ admits an EL-labeling as well.*

The EL-labeling that P inherits from $P_1, \dots, P_m$ is the following. Assume that Λ_i is the totally ordered set of the EL-labels of P_i . If $\lambda_i \in \Lambda_i$, we totally order the m -tuples $(\emptyset, \dots, \emptyset, \lambda_i, \emptyset, \dots, \emptyset)$ with λ_i in the i th entry by letting

$$(4) \quad (\emptyset, \dots, \emptyset, \lambda_i, \emptyset, \dots, \emptyset) \leq (\emptyset, \dots, \emptyset, \lambda_j, \emptyset, \dots, \emptyset) \text{ if and only if} \\ i = j \text{ and } \lambda_i \leq \lambda_j \text{ in the total order of } \Lambda_i \\ \text{or } i > j.$$

If $x = (x_1, \dots, x_m)$ is covered by $x' = (x'_1, \dots, x'_m)$ in the poset P , then there is some $1 \leq i_0 \leq m$ such that $x_i = x'_i$ for all $1 \leq i \neq i_0 \leq m$ and $x_{i_0} \rightarrow x'_{i_0}$ in P_{i_0} .

If $\lambda_{i_0} = \lambda(x_{i_0} \rightarrow x'_{i_0})$ is the edge label from the EL-labeling of P_{i_0} , then we label the edge $x \rightarrow x'$ in P by the m -tuple $(\emptyset, \dots, \emptyset, \lambda_{i_0}, \emptyset, \dots, \emptyset)$ with λ_{i_0} in the i_0 th entry. The edge labels of P we obtain this way, totally ordered as in (4), form an EL-labeling of $P_1 \times P_2 \times \dots \times P_m$.

The EL-labeling of $NC(\gamma)$. Let Φ be an irreducible root system of rank n with positive part Φ^+ , and let $\Pi = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$ be a choice of simple system for Φ such that $\Pi_+ = \{\sigma_1, \sigma_2, \dots, \sigma_r\}$ and $\Pi_- = \{\sigma_{r+1}, \dots, \sigma_n\}$ are orthonormal sets. For each $\alpha \in \Phi$ we denote by r_α the reflection in the hyperplane in $\mathbb{R}^n$ orthogonal to α . A *bipartite Coxeter element* of the root system Φ is a Coxeter element of type $\gamma = \gamma_+ \gamma_-$, where

$$(5) \quad \gamma_\pm = \prod_{\alpha \in \Pi_\pm} r_\alpha.$$

Thus, in the present case, the bipartite Coxeter element is $\gamma = r_{\sigma_1} \dots r_{\sigma_n}$.

Let $N = nh/2$ be the number of positive roots of Φ^+ . For $1 \leq i \leq 2N$ we define $\rho_i = r_{\sigma_1} r_{\sigma_2} \dots r_{\sigma_{i-1}}(\sigma_i)$, where the simple roots σ_i are indexed cyclically modulo n and we make the convention that $\rho_{-i} = \rho_{2N-i}$ for $i \geq 0$. The following formula, which is easily verified, appears in [12, section 3]:

$$(6) \quad \rho_i = \begin{cases} \sigma_i & \text{for } i = 1, \dots, r, \\ -\gamma(\sigma_i) & \text{for } i = r+1, \dots, n, \\ \gamma(\rho_{i-n}) & \text{for } i > n. \end{cases}$$

Then we have

$$\begin{aligned} \{\rho_1, \dots, \rho_N\} &= \Phi^+, \\ \{\rho_{N+i} : 1 \leq i \leq r\} &= -\Pi_+, \\ \{\rho_{-i} : 0 \leq i < n-r\} &= \{-\rho_{N-i} : 0 \leq i < n-r\} = -\Pi_-. \end{aligned}$$

Let $\Phi_{\geq -1} = \Phi^+ \cup (-\Pi)$ be the set of *almost positive roots*. We totally order the roots in $\Phi_{\geq -1}$ as follows:

$$(7) \quad \underbrace{\rho_{-n+r+1} \prec \dots \prec \rho_0}_{-\Pi_-} \prec \underbrace{\rho_1 \prec \dots \prec \rho_N}_{\Phi^+} \prec \underbrace{\rho_{N+1} \prec \dots \prec \rho_{N+r}}_{-\Pi_+}.$$

This order induces a total order, which we again denote by $\prec$, on the set T of reflections in W as follows:

$$(8) \quad r_{\rho_1} \prec r_{\rho_2} \prec \dots \prec r_{\rho_N}.$$

Let $u, v \in NC(\gamma)$ such that v covers u . Clearly $v = ut$, where t is a reflection in W . The *natural edge labeling* of the edge $u \rightarrow v$ is the reflection $t = u^{-1}v$.

THEOREM 2.8 (see [5, Theorem 4.2]). *If the set T of reflections in W is totally ordered by (8) and $\gamma = r_{\sigma_1} r_{\sigma_2} \dots r_{\sigma_n}$, then the natural edge labeling of $NC(\gamma)$ with label set T is an EL-labeling.*

The poset $NC_{(m)}(W)$. Here we recall some general facts for the poset $NC_{(m)}(W)$ of m -divisible noncrossing partitions. We refer the reader to [1] for more details.

DEFINITION 2.9 (see [1, Definition 3.2.2]). *Let W be a reflection group, and fix a Coxeter element γ . We denote by $NC_{(m)}(\gamma)$ the set of m -tuples $(w_1, \dots, w_m) \in (NC(\gamma))^m$ for which*

- (i) $w = w_1 \dots w_m \leq \gamma$ and
- (ii) $l_T(w) = \sum_{i=1}^m l_T(w_i)$.

Recall that a subset I of a poset $(P, \leq)$ is an *order ideal* if $x, y \in P$, $x \in I$, and $y \leq x$ implies $y \in I$. The set $NC_{(m)}(\gamma)$ is a subposet of $(NC(\gamma))^m$. Moreover it is an order ideal of $(NC(\gamma))^m$ whose maximal elements are the m -tuples $(w_1, \dots, w_m)$ where $w_1 \dots w_m$ is a minimal factorization of γ [1, Lemma 3.4.3]. Thus, the set $NC_{(m)}(\gamma)$ inherits the partial order as well as the rank function of $(NC(\gamma))^m$. More specifically, for $(u_1, \dots, u_m), (w_1, \dots, w_m) \in NC_{(m)}(\gamma)$ we have

$$(u_1, u_2, \dots, u_m) \leq (w_1, w_2, \dots, w_m) \text{ if } u_i \leq w_i \text{ for all } 1 \leq i \leq m$$

and

$$rk((w_1, \dots, w_m)) = \sum_{i=1}^m l_T(w_i).$$

By the fact that all Coxeter elements are conjugate and Lemma 2.5, we deduce that the isomorphism type of $NC_{(m)}(\gamma)$ is independent of the choice of γ . Therefore, if we are not interested in the particular choice of γ , we write $NC_{(m)}(W)$, and we call this the poset of *m -divisible noncrossing partitions*. The poset $NC_{(m)}(W)$ is a ranked meet-semilattice [1, Lemma 3.4.4] with unique minimal element the m -tuple $(1, \dots, 1)$.

The cluster complex $\Delta(\Phi)$. Let Φ be an irreducible root system, and let $\Phi_{\geq -1} = \Phi^+ \cup (-\Pi)$ be the set of almost positive roots. We define the map $\mathcal{R}$ on $\Phi_{\geq -1}$ as follows:

$$(9) \quad \mathcal{R}(\alpha) = \begin{cases} \gamma^{-1}(\alpha) & \text{if } \alpha \notin \Pi_+ \cup (-\Pi_-), \\ -\alpha & \text{if } \alpha \in \Pi_+ \cup (-\Pi_-). \end{cases}$$

THEOREM 2.10 (see [15]). *There is a unique symmetric binary relation on $\Phi_{\geq -1}$ called compatibility such that*

- (i) α and β are compatible if and only if $\mathcal{R}(\alpha)$ and $\mathcal{R}(\beta)$ are compatible, and
- (ii) a negative simple root $-\sigma_i$ is compatible with a positive root β if and only if the simple root expansion of β does not involve σ_i .

The cluster complex $\Delta(\Phi)$ is defined as the simplicial complex on the vertex set $\Phi_{\geq -1}$ such that a set $\sigma \subseteq \Phi_{\geq -1}$ is a face of $\Delta(\Phi)$ if every pair of roots in σ is compatible. The maximal faces of $\Delta(\Phi)$ are called *clusters*. The simplicial complex $\Delta(\Phi)$ is pure, $(n-1)$ -dimensional, and homeomorphic to a sphere [15].

The generalized cluster complex $\Delta^m(\Phi)$. The *generalized cluster complex* $\Delta^m(\Phi)$, introduced in [14], is a simplicial complex on the vertex set $\Phi_{\geq -1}^m$ of *colored almost positive roots*. More specifically, $\Phi_{\geq -1}^m$ is the union of the set $\Phi_{>0}^m$ of m (colored) copies $\alpha^1, \dots, \alpha^m$ of each positive root $\alpha \in \Phi^+$ with the set of negative simple roots. We make the convention that each negative simple root is colored by 1. Thus,

$$\Phi_{\geq -1}^m = \Phi_{>0}^m \cup (-\Pi)^1 = \{\alpha^k : \alpha \in \Phi^+, k \in \{1, \dots, m\}\} \cup \{\alpha^1 : \alpha \in -\Pi\}.$$

The generalized cluster complex $\Delta^m(\Phi)$ is a simplicial complex on the vertex set $\Phi_{\geq -1}^m$. The faces of $\Delta^m(\Phi)$ are the sets $\sigma \subseteq \Phi_{\geq -1}^m$ whose elements are pairwise m -compatible in the sense we are going to describe just below. For every $\alpha \in \Phi_{\geq -1}^m$ we define the *degree* $d(\alpha)$ of α to be the smallest d such that $\mathcal{R}^d(\alpha)$ is a negative simple root. Clearly, $d(\alpha) = 0$ if α is a negative simple root.

DEFINITION 2.11 (see [14, Definition 3.1]). *Two colored roots $\alpha^k, \beta^l \in \Phi_{\geq -1}^m$ are called m -compatible if and only if one of the following conditions is satisfied:*

- $k > l$, $d(\alpha) \leq d(\beta)$, and the roots $\mathcal{R}(\alpha)$ and β are compatible in the original noncolored sense (Theorem 2.10),
- $k < l$, $d(\alpha) \geq d(\beta)$, and the roots α and $\mathcal{R}(\beta)$ are compatible,
- $k > l$, $d(\alpha) > d(\beta)$, and the roots α and β are compatible,
- $k < l$, $d(\alpha) < d(\beta)$, and the roots α and β are compatible, and
- $k = l$ and the roots α and β are compatible.

It is immediate from the definition that m -compatibility is a symmetric relation. There is another equivalent way to define this relation, by introducing the m -analogue of the map $\mathcal{R}$, as follows.

DEFINITION 2.12 (see [14, Definition 3.3]). *For $\alpha^k \in \Phi_{\geq -1}^m$, we set*

$$(10) \quad \mathcal{R}_m(\alpha^k) = \begin{cases} \alpha^{k+1} & \text{if } \alpha \in \Phi_{>0}^m \text{ and } k < m, \\ (\mathcal{R}(\alpha))^1 & \text{otherwise.} \end{cases}$$

THEOREM 2.13 (see [14, Theorem 3.4]). *The m -compatibility relation on $\Phi_{\geq -1}^m$ has the following properties:*

- (i) α^k is m -compatible with β^l if and only if $\mathcal{R}_m(\alpha^k)$ is m -compatible with $\mathcal{R}_m(\beta^l)$;
- (ii) $(-\sigma_i)^1$ is m -compatible with β^l if and only if the simple root expansion of β does not involve σ_i .

Furthermore, the above conditions uniquely determine this relation.

We denote by $\Delta_+^m(\Phi)$ the *positive part* of $\Delta^m(\Phi)$, which is the induced subcomplex of $\Delta^m(\Phi)$ on the vertex set $\Phi_{>0}^m$ of positive colored roots. If Φ is reducible with irreducible components $\Phi_1, \dots, \Phi_\ell$, then $\Phi_{\geq -1}^m = \bigcup_{i=1}^\ell (\Phi_i^m)_{\geq -1}$. We declare two roots in $\Phi_{\geq -1}^m$ m -compatible if they either belong to different components or belong to the same component and are m -compatible within it. Thus, the complex $\Delta^m(\Phi)$ is the simplicial join of the complexes $\Delta^m(\Phi_i)$.

3. Proof of Theorem 1.1. Throughout this section we assume that Φ is an irreducible root system and γ a bipartite Coxeter element. We will prove the following restatement of Theorem 1.1 which can be deduced immediately from Definition 2.9.

THEOREM 3.1. *The set $\sigma \subseteq \Phi_{\geq -1}^m$ is a face of $\Delta^m(\Phi)$ if and only if*

- (i) $w_{\sigma^+} w_{\sigma^m} w_{\sigma^{m-1}} \dots w_{\sigma^1} w_{\sigma^-} \leq \gamma$ and
- (ii) $l_T(w_{\sigma^+} w_{\sigma^m} w_{\sigma^{m-1}} \dots w_{\sigma^1} w_{\sigma^-}) = |\sigma|$.

The above criterion reduces to that of Brady and Watt [12, section 8] in the case $m = 1$. Before proving Theorem 3.1 we need to establish a few more facts.

Let $\mu : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the map introduced in [12], with

$$\mu(x) = 2(I - \gamma)^{-1}(x)$$

for all $x \in \mathbb{R}^n$. In what follows, we denote by $\alpha \cdot \beta$ the inner product of $\alpha, \beta \in \mathbb{R}^n$.

LEMMA 3.2.

- (i) [12, section 8] *Let $\alpha, \beta \in \Phi_{\geq -1}$, with $\alpha \prec \beta$. The roots α, β are compatible if and only if $\mathbf{1} \neq r_\beta r_\alpha \leq \gamma$.*

(ii) For $\alpha, \beta \in \Phi_{\geq -1}$ we have $\mu(\alpha) \cdot \beta = 0$ if and only if $\mathbf{1} \neq r_\alpha r_\beta \leq \gamma$.

Proof of (ii). Part (ii) is a slight modification of [6, Lemma 2.2], which states that if the roots $\alpha, \beta \in \Phi_{\geq -1}$ are nonparallel, then $\mu(\alpha) \cdot \beta = 0 \Leftrightarrow r_\alpha r_\beta \leq \gamma$. From [12, Corollary 3.3(b)] we have that $\mu(\alpha) \cdot \alpha = 1$ for all $\alpha \in \Phi_{\geq -1}$. Thus, if $\mu(\alpha) \cdot \beta = 0$, then $\alpha \neq \pm\beta$, or, equivalently, the roots α, β are nonparallel and therefore $\mathbf{1} \neq r_\alpha r_\beta \leq \gamma$. Conversely, if $\mathbf{1} \neq r_\alpha r_\beta \leq \gamma$, then $\alpha \neq \pm\beta$; that is, the roots α, β are nonparallel, and therefore $\mu(\alpha) \cdot \beta = 0$ by [6, Lemma 2.2]. $\square$

LEMMA 3.3 (see [12, Theorem 3.7]). For the total order (7) of roots in $\Phi_{\geq -1}$ we have $\mu(\rho_j) \cdot \rho_{i-n} = -\mu(\rho_i) \cdot \rho_j$ for all i, j .

LEMMA 3.4. If ρ_i is a positive root, then $\rho_{i-n} = \gamma^{-1}(\rho_i)$.

Proof. If $n+1 \leq i \leq N$, the result is clear by (6). If $1 \leq i \leq n$, then recall that $\rho_{i-n} = \rho_{2N+i-n} = \rho_{nh+i-n}$, where N is the number of positive roots. By the relation $\rho_i = \gamma(\rho_{i-n})$ holding for $i > n$ and the fact that $\gamma^h = \mathbf{1}$, we have

$$\rho_{n(h-1)+i} = \gamma(\rho_{n(h-2)+i}) = \gamma^2(\rho_{n(h-3)+i}) = \cdots = \gamma^{h-1}(\rho_i) = \gamma^{-1}(\rho_i),$$

which completes our proof. $\square$

LEMMA 3.5. Let $w \in W$ and $t_1, t_2, \dots, t_k$ be distinct reflections in W such that $t_i t_j \leq w$ for every $i < j$. Then $t_1 t_2 \dots t_k \leq w$ and $l_T(t_1 t_2 \dots t_k) = k$.

Proof. We will proceed by induction, the case $k = 2$ being trivial. We assume that the statement holds for k , and we prove it for $k+1$. To this end we consider $t_1, \dots, t_k, t_{k+1}$ distinct reflections in W such that $t_i t_j \leq w$ for every $i < j$. We have $t_1 \dots t_k \leq w$ and $t_2 \dots t_{k+1} \leq w$ by the inductive assumption. Since $t_1 \leq t_1 t_2 \dots t_k \leq w$, then $t_2 \dots t_k \leq t_1 w$ by Lemma 2.4(ii). Moreover, since $t_1 t_{k+1} \leq w$, then $t_{k+1} \leq t_1 w$ by part (iii) of the same lemma. By setting $a = t_2 \dots t_k$, $b = t_{k+1}$, $c = t_1 w$, and applying Lemma 2.4(i), we get $t_2 \dots t_{k+1} \leq t_1 w$. Thus $t_1 t_2 \dots t_{k+1} \leq w$ by Lemma 2.4(ii).

To prove the statement for the length, note that since $t_2 \dots t_{k+1} \leq t_1 w$, then $l_T(t_1 w) = l_T(t_2 \dots t_{k+1}) + l_T(t_{k+1} \dots t_2 t_1 w)$. Equivalently, $l_T(w) - 1 = k + l_T(w) - l_T(t_1 \dots t_{k+1})$ by the inductive assumption and the fact that $t_1 \dots t_{k+1} \leq w$. Thus $l_T(t_1 \dots t_{k+1}) = k + 1$, which completes our proof. $\square$

In view of (6) and the remarks following it, the total order (7) of the set of almost positive roots $\Phi_{\geq -1}$ can be reformulated as follows:

$$(11) \quad \begin{aligned} & -\Pi_- \prec \underbrace{\Pi_+ \prec \gamma(-\Pi_-)}_{d=1} \prec \underbrace{\gamma(\Pi_+) \prec \gamma^2(-\Pi_-)}_{d=2} \prec \gamma^2(\Pi_+) \prec \cdots \prec \\ & \prec \gamma^{-1}(-\Pi_+) \prec \Pi_- \prec -\Pi_+. \end{aligned}$$

LEMMA 3.6.

- (i) If $\alpha, \beta \in \Phi^+$ and $d(\alpha) < d(\beta)$, then $\alpha \prec \beta$.
- (ii) If $\alpha \in \Phi^+$, $\beta \in \Phi^+ \setminus \Pi_+$, and $d(\beta) \leq d(\alpha)$, then $\mathcal{R}(\beta) \prec \alpha$.

Proof. Part (i) follows from the fact that for the set of positive roots the degree d increases according to the total order $\prec$, as one may check from (11). To prove (ii), we have that since $\beta \in \Phi^+ \setminus \Pi_+$, then $\mathcal{R}(\beta) = \gamma^{-1}(\beta)$ by (9) so that $d(\mathcal{R}(\beta)) = d(\beta) - 1$. Thus $d(\mathcal{R}(\beta)) < d(\alpha)$, and therefore $\mathcal{R}(\beta) \prec \alpha$ by (i). $\square$

To prove Theorem 3.1 we use a criterion which determines when two colored almost positive roots in $\Phi_{\geq -1}^m$ are m -compatible. More precisely, we totally order the roots in $\Phi_{\geq -1}^m$ as follows:

$$(12) \quad (-\Pi_-)^1 \prec (\Phi^+)^1 \prec \cdots \prec (\Phi^+)^{m-1} \prec (\Phi^+)^m \prec (-\Pi_+)^1,$$

where the roots in each of the sets $(-\Pi_-)^1$, $(-\Pi_+)^1$, and $(\Phi^m)^k$, $1 \leq k \leq m$, are ordered as in (7). The following criterion generalizes Lemma 3.2(i). We make the convention that $r_{\alpha^i} = r_\alpha$ for every colored root $\alpha^i \in \Phi_{\geq -1}^m$.

PROPOSITION 3.7. *Let $\alpha^i, \beta^j \in \Phi_{\geq -1}^m$, with $\alpha^i \prec \beta^j$ in the total order (12). The roots α^i, β^j are m -compatible if and only if $\mathbf{1} \neq r_{\beta^j} r_{\alpha^i} \leq \gamma$.*

Proof. Let $\alpha^i, \beta^j \in \Phi_{\geq -1}^m$, with $\alpha^i \prec \beta^j$ in the total order (12). We distinguish cases:

(i) $i = j$. Notice that if $i = j$, then $\alpha^i \prec \beta^j$ in the total order (12) if and only if $\alpha \prec \beta$ in the total order (7). Moreover, recall by Definition 2.11 that roots of the same color are m -compatible if and only if they are compatible in the noncolored sense. Our claim then follows from Lemma 3.2(i).

(ii) $\beta^j \in \Phi_{>0}^m$, $i = 1$, and $\alpha^1 \in (-\Pi_-)^1$. Recall that the roots $\alpha^1 \in (-\Pi_-)^1$ and $\beta^j \in \Phi_{\geq -1}^m$ are m -compatible if and only if the roots $\alpha \in -\Pi, \beta \in \Phi^+$ are compatible in the noncolored sense (Theorems 2.10(ii) and 2.13(ii)). Moreover, the roots in $-\Pi_-$ precede all roots in Φ^+ in the total order (7). Thus $\alpha \in -\Pi_-$ and $\beta \in \Phi^+$ are compatible if and only if $\mathbf{1} \neq r_\beta r_\alpha \leq \gamma$ by Lemma 3.2(i).

(ii)' $\alpha^i \in \Phi_{>0}^m$, $j = 1$, and $\beta^1 \in (-\Pi_+)^1$. This case is symmetric to case (ii), bearing in mind that the roots in $-\Pi_+$ succeed all roots in Φ^+ .

(iii) $\alpha^i, \beta^j \in \Phi_{>0}^m$ and $i < j$. We first suppose that $d(\alpha) < d(\beta)$ so that $\alpha \prec \beta$ by Lemma 3.6(i). In this case, the roots α^i, β^j are m -compatible if and only if the noncolored roots α, β are compatible. By Lemma 3.2(i) we have that the roots α, β are compatible if and only if $\mathbf{1} \neq r_\beta r_\alpha \leq \gamma$, which proves our claim. Next, we suppose that $d(\alpha) \geq d(\beta)$ so that the roots α^i, β^j are m -compatible if and only if the roots $\alpha, \mathcal{R}(\beta)$ are compatible. We distinguish two cases:

Case 1: $\alpha \in \Phi^+$ and $\beta \in \Pi_+$. From (9) we have that $\mathcal{R}(\beta) = -\beta \in -\Pi_+$, and, since the roots in Φ^+ precede those in $-\Pi_+$, we have that $\alpha \prec \mathcal{R}(\beta)$. By Lemma 3.2(i), the roots $\alpha, \mathcal{R}(\beta)$ are compatible if and only if $\mathbf{1} \neq r_{\mathcal{R}(\beta)} r_\alpha \leq \gamma$. Since $r_{\mathcal{R}(\beta)} = r_{-\beta} = r_\beta$, then the roots $\alpha, \mathcal{R}(\beta)$ are compatible if and only if $\mathbf{1} \neq r_\beta r_\alpha \leq \gamma$, which is what we are after.

Case 2: $\alpha \in \Phi^+$ and $\beta \in \Phi^+ \setminus \Pi_+$. Since $d(\beta) \leq d(\alpha)$, then $\mathcal{R}(\beta) \prec \alpha$ by Lemma 3.6(ii). By Lemma 3.2(i) we have that the roots $\alpha, \mathcal{R}(\beta)$ are compatible if and only if $\mathbf{1} \neq r_\alpha r_{\mathcal{R}(\beta)} \leq \gamma$. We set $\alpha = \rho_i$, $\beta = \rho_j$, with $i \geq 0$ and $j \geq r+1$, and we notice that $\mathcal{R}(\rho_j) = \gamma^{-1}(\rho_j) = \rho_{j-n}$ by the definition of $\mathcal{R}$ and Lemma 3.4. We have $\mathbf{1} \neq r_\alpha r_{\mathcal{R}(\beta)} \leq \gamma \Leftrightarrow \mathbf{1} \neq r_{\rho_i} r_{\rho_{j-n}} \leq \gamma \Leftrightarrow \mathbf{1} \neq r_{\rho_i} r_{\rho_{j-n}} \leq \gamma \Leftrightarrow \mu(\rho_i) \cdot \rho_{j-n} = 0 \Leftrightarrow \mu(\rho_j) \cdot \rho_i = 0 \Leftrightarrow \mathbf{1} \neq r_{\rho_j} r_{\rho_i} \leq \gamma \Leftrightarrow \mathbf{1} \neq r_\beta r_\alpha \leq \gamma$ by Lemmas 3.2(ii) and 3.3. Therefore, the roots $\alpha, \mathcal{R}(\beta)$ are compatible if and only if $\mathbf{1} \neq r_\beta r_\alpha \leq \gamma$, and this completes our proof. $\square$

Proof of Theorem 3.1. Let $\sigma \subseteq \Phi_{\geq -1}^m$. Recall that the set σ is a face of $\Delta^m(\Phi)$ if and only if any two roots $\alpha^i, \beta^j \in \sigma$ are m -compatible. We totally order the elements of the set σ in decreasing order with respect to (12). That is, we partition the set σ into the sets $\sigma_+, \sigma^m, \dots, \sigma^1, \sigma_-$, and we totally order them as follows:

$$(13) \quad \sigma_+, \sigma^m, \sigma^{m-1}, \dots, \sigma^1, \sigma_-,$$

where the roots in each of the sets σ_+, σ_- , and σ^k , $1 \leq k \leq m$, are in decreasing order with respect to (7). The criterion of Proposition 3.7 implies that the roots in σ are m -compatible if and only if for any two roots $\alpha^i, \beta^j \in \sigma$, with $\alpha^i \prec \beta^j$, we have $r_{\beta^j} r_{\alpha^i} \leq \gamma$. This, in view of Lemma 3.5 and the total ordering (13) of the roots in σ , implies that σ is a face of $\Delta^m(\Phi)$ if and only if $w_{\sigma_+} w_{\sigma^m} w_{\sigma^{m-1}} \dots w_{\sigma^1} w_{\sigma_-} \leq \gamma$ and $l_T(w_{\sigma_+} w_{\sigma^m} w_{\sigma^{m-1}} \dots w_{\sigma^1} w_{\sigma_-}) = |\sigma|$, which completes our proof.

Remark 3.8. Note that the total order (7) of the set of almost positive roots $\Phi_{\geq -1}$ can be expressed in terms of the map $\mathcal{R}$ as follows:

$$(14) \quad \begin{aligned} -\Pi_- < \Pi_+ < \mathcal{R}^{-1}(-\Pi_-) < \mathcal{R}^{-1}(\Pi_+) < \mathcal{R}^{-2}(-\Pi_-) < \mathcal{R}^{-2}(-\Pi_+) < \dots \\ < \mathcal{R}(-\Pi_+) < \Pi_- < -\Pi_+. \end{aligned}$$

However, this property does not generalize when $m > 1$. That is, the total order (12) of the set of colored almost positive roots $\Phi_{\geq -1}^m$ is not the one suggested by $\mathcal{R}_m$.

4. The EL-labeling and decreasing chains in $NC_{(m)}(\gamma)$. In this section we describe the EL-labeling that $NC_{(m)}(\gamma)$ inherits from $NC(\gamma)$ and explain the relation of decreasing chains with respect to this EL-labeling to facets of $\Delta_+^m(\Phi)$. The idea for the EL-labeling was suggested by Armstrong (see [1, section 3.7]).

Since $NC(\gamma)$ admits an EL-labeling, so does $(NC(\gamma))^m$, by Proposition 2.7. Moreover, recall that $NC_{(m)}(\gamma)$ is an order ideal in $(NC(\gamma))^m$ so that we can restrict the EL-labeling of the latter to the former. The labeling that $NC_{(m)}(\gamma)$ inherits from $(NC(\gamma))^m$ is the following. If $(w_1, \dots, w_m) \rightarrow (w'_1, \dots, w'_m)$ in $NC_{(m)}(\gamma)$, then there exists some $1 \leq i_0 \leq m$ such that $w_i = w'_i$ for all $1 \leq i \neq i_0 \leq m$ and $w_{i_0} \leq w'_{i_0} = w_{i_0} t_{i_0}$ for some reflection t_{i_0} in W . We label the edge $(w_1, \dots, w_m) \rightarrow (w'_1, \dots, w'_m)$ by $(\mathbf{1}, \dots, \mathbf{1}, t_{i_0}, \mathbf{1}, \dots, \mathbf{1})$, with t_{i_0} in the i_0 th entry, and we call this the *natural edge labeling* of $NC_{(m)}(\gamma)$. Let Λ be the set of m -tuples $(\mathbf{1}, \dots, r_\alpha, \dots, \mathbf{1})$, where r_α is the reflection corresponding to the root α . Following (4), we totally order the elements of Λ by letting

$$(15) \quad (\mathbf{1}, \dots, r_\alpha, \dots, \mathbf{1}) \preceq (\mathbf{1}, \dots, r_{\alpha'}, \dots, \mathbf{1})$$

with the reflection on the i th and i' th entries, respectively, if and only if

$$i = i' \quad \text{and} \quad \alpha \preceq \alpha' \quad \text{or} \\ i > i'.$$

The preceding discussion leads us to the following result. We should point out that the following proposition is part of a stronger result [1, Theorem 3.7.2], namely, that the poset $NC_{(m)}(\gamma) \cup \{\hat{1}\}$ of m -divisible noncrossing partitions with a maximal element adjoined is EL-shellable.

PROPOSITION 4.1. *If Λ is totally ordered as in (15), then the natural edge labeling of $NC_{(m)}(\gamma)$ is an EL-labeling.*

Let $[\hat{0}, w]$ be an interval in $NC_{(m)}(\gamma)$. Our next goal is to relate the maximal decreasing chains of $[\hat{0}, w]$ to facets of certain subcomplexes of $\Delta_+^m(\Phi)$. For a fixed $w = (w_1, \dots, w_m) \in NC_{(m)}(\gamma)$ we define $\Delta_+^m(w)$ as the induced subcomplex of $\Delta_+^m(\Phi)$ on the vertex set of colored positive roots $\alpha^i \in \Phi_{>0}^m$, with $r_{\alpha^i} \leq w_{m-i+1}$. By Lemmas 3.2(i) and 3.5 and the fact that m -compatible roots of the same color are compatible, we deduce that $\Delta_+^m(w)$ is the simplicial complex whose faces are the sets $\sigma \subseteq \Phi_{>0}^m$ with $(w_{\sigma^m}, \dots, w_{\sigma^1}) \leq (w_1, \dots, w_m)$. Equivalently, the faces of $\Delta_+^m(w)$ are the sets $\sigma \subseteq \Phi_{>0}^m$ with $w_{\sigma^i} \leq w_{m-i+1}$ for all $1 \leq i \leq m$. The complex $\Delta_+^m(w)$ is the simplicial join of $\Delta_+(w_i)$, $1 \leq i \leq m$, and therefore it is pure of dimension $rk(w) - 1$.

Now consider $w = (w_1, \dots, w_m) \in NC_{(m)}(\gamma)$ with $rk(w) = k$, and write $\hat{0} = (\mathbf{1}, \dots, \mathbf{1})$. If C is a maximal decreasing chain in $[\hat{0}, w]$, then we can decompose it into m unrefinable parts C_i as follows. We define C_i to be the unrefinable subchain of C whose edge labels have a reflection on the i th entry. Clearly, in view of the total

order (15), C is the chain formed by arranging the subchains C_i one following the other, starting with C_1 and terminating with C_m . Next, consider the subsets of Φ^+ defined by

$$\begin{aligned} \sigma(C_i) &= \{\alpha \in \Phi^+ : (\mathbf{1}, \dots, r_\alpha, \dots, \mathbf{1}) \text{ is an edge label on } C_i\} \\ &= \{\alpha \in \Phi^+ : (\mathbf{1}, \dots, r_\alpha, \dots, \mathbf{1}) \text{ is an edge label on } C \\ &\quad \text{with } r_\alpha \text{ on the } i\text{th coordinate}\} \end{aligned}$$

for $1 \leq i \leq m$. Let $w_{\sigma(C_i)}$ be the product of reflections corresponding to the roots in $\sigma(C_i)$ in decreasing order with respect to (7). By (15), this is the order in which these reflections appear on the edge labels of C . Thus $w_{\sigma(C_i)} = w_i$, and therefore $(w_{\sigma(C_1)}, \dots, w_{\sigma(C_m)}) \in NC_{(m)}(\gamma)$. By Theorem 1.1, if we color the roots in each $\sigma(C_i)$ by $m - i + 1$, we obtain a face of $\Delta_+^m(\Phi)$ and in particular a facet of $\Delta_+^m(w)$.

Conversely, let σ be a facet of $\Delta_+^m(w)$, so that $w_{\sigma^{m-i+1}} = w_i$ for all $1 \leq i \leq m$. Consider the set

$$\Lambda_\sigma := \{(\mathbf{1}, \dots, r_{\alpha^i}, \dots, \mathbf{1}) : \alpha^i \in \sigma^i \text{ and } r_{\alpha^i} \text{ on the } (m - i + 1)\text{th entry}, 1 \leq i \leq m\},$$

and order its elements in decreasing order with respect to (15). Now we construct by induction a decreasing chain C_σ in the interval $[\hat{0}, w]$ of $NC_{(m)}(\gamma)$ as follows. The first element of C is $(\mathbf{1}, \dots, \mathbf{1})$. Suppose that we have found the first $k + 1$ elements of C_σ by using the first k (in the total order (15)) elements of Λ_σ . If the $(k + 1)$ th element of C_σ is $(u_1, \dots, u_m)$ and the $(k + 1)$ th element of Λ_σ is $(\mathbf{1}, \dots, r_{\alpha_j}, \dots, \mathbf{1})$, with r_{α_j} on the j th coordinate, then the $(k + 2)$ th element of the chain C_σ is $(u_1, \dots, u_{j-1}, r_{\alpha_j} u_j, u_{j+1}, \dots, u_m)$.

We summarize the above facts in the following proposition.

PROPOSITION 4.2. *Let $w = (w_1, \dots, w_m) \in NC_{(m)}(\gamma)$, and let C be a maximal decreasing chain in $[\hat{0}, w]$. The map which sends C to $\bigcup_{i=1}^m \sigma(C_i)^{m-i+1}$ is a bijection from the set of maximal decreasing chains of $[\hat{0}, w]$ to the set of facets of $\Delta_+^m(w)$.*

In the case $m = 1$ the next corollary reduces to the remark following Corollary 4.3 in [5].

COROLLARY 4.3. *The set of maximal decreasing chains in $NC_{(m)}(\gamma)$ bijects to the set of facets of $\Delta_+^m(\Phi)$.*

Proof. The set of maximal decreasing chains in $NC_{(m)}(\gamma)$ is the disjoint union of the sets of such chains within each $[\hat{0}, w]$ with $rk(w) = n$. Moreover, the set of facets of $\Delta_+^m(\Phi)$ is the disjoint union of the set of facets of each $\Delta_+^m(w)$, where $rk(w) = n$. Our claim then follows from Proposition 4.2. $\square$

The reader is invited to verify Corollary 4.3 in the following example.

Example 4.4. Consider the root system A_2 , and identify the corresponding reflection group W_{A_2} with the group of permutations on the set $\{1, 2, 3\}$. Let $\sigma_1, \sigma_2, \alpha$ be the standard choice of positive roots, where σ_1, σ_2 are simple roots and $\alpha = \sigma_1 + \sigma_2$. More precisely, the reflection in the hyperplane of $\mathbb{R}^3$ orthogonal to each of σ_1, σ_2 , and α acts on $\mathbb{R}^3$ by transposing coordinates, so that $r_{\alpha_1} = (12)$, $r_{\alpha_2} = (23)$, and $r_\alpha = (13)$. We set $\Pi_+ = \{\sigma_1\}$, $\Pi_- = \{\sigma_2\}$ so that $\sigma_1 \prec \alpha \prec \sigma_2$. The bipartite Coxeter element in this case is $\gamma = (123)$. Figure 1 illustrates the natural edge labeling of $NC_{(2)}(\gamma)$. The set of maximal decreasing chains in $NC_{(2)}(\gamma)$ bijects to the set of

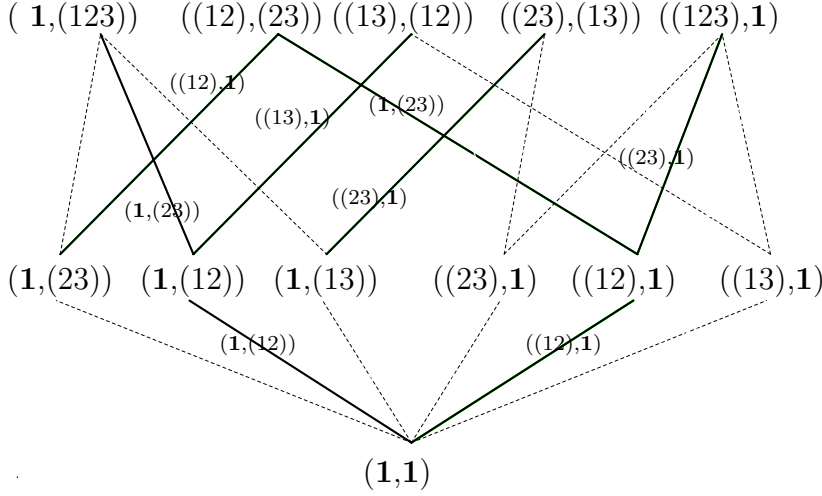


FIG. 1. The EL-labeling of $NC_2((123))$ for $\Phi = A_2$. The decreasing chains are marked with dashed lines.

positive clusters as shown in the following table.

Edge labels of decreasing chains	Positive clusters
$(1, (23)) \rightarrow (1, (13))$	σ_2^1, α^1
$(1, (13)) \rightarrow (1, (12))$	α^1, σ_1^1
$((23), 1) \rightarrow (1, (13))$	σ_2^2, α^1
$((23), 1) \rightarrow ((13), 1)$	σ_2^2, α^2
$((12), 1) \rightarrow (1, (23))$	σ_1^2, σ_2^1
$((13), 1) \rightarrow ((12), 1)$	α^2, σ_1^1
$((13), 1) \rightarrow (1, (12))$	α^2, σ_1^1

LEMMA 4.5. For $w = (w_1, \dots, w_m) \in NC_{(m)}(\gamma)$ the number $(-1)^{rk(w)} \mu(\hat{0}, w)$ is equal to the number of facets of $\Delta_+^m(w)$.

Proof. A standard fact on Möbius functions of EL-shellable posets is that the Möbius function on every interval $[a, b]$ is equal to $(-1)^{rk(b) - rk(a)}$ times the number of maximal decreasing chains in $[a, b]$ [19, section 3.13]. Applying this in our case with $a = \hat{0}$ and $b = w$, we deduce that $(-1)^{rk(w)} \mu(\hat{0}, w)$ is equal to the number of maximal decreasing chains in $[\hat{0}, w]$ which, in view of Proposition 4.2, is equal to the number of facets of $\Delta_+^m(w)$. $\square$

5. Proof of Theorem 1.2. As pointed out in [17, Proposition F(i)], for any finite root systems Φ and Φ' we have

$$F_{\Phi \times \Phi'}^{(m)}(x, y) = F_{\Phi}^{(m)}(x, y) \cdot F_{\Phi'}^{(m)}(x, y).$$

Moreover, by the multiplicativity of the Möbius function, the above relation holds for the M -triangle as well. Therefore, it suffices to prove Theorem 1.2 in the case where Φ is irreducible.

The h -polynomial of an abstract $(n - 1)$ -dimensional simplicial complex Δ is defined as

$$h(\Delta, y) = \sum_{i=0}^n f_i(\Delta) y^i (1 - y)^{n-i},$$

where $f_i(\Delta)$ is the number of faces of Δ of dimension $i - 1$. The link of a face σ of Δ is the abstract simplicial complex $\text{lk}_\Delta(\sigma) = \{\tau \setminus \sigma : \sigma \subseteq \tau \in \Delta\}$. One can check by comparing the table in [1, Figure 3.4] and [14, Theorem 10.2] that for any irreducible root system Φ the h -polynomial of $\Delta^m(\Phi)$ coincides with the rank-generating polynomial of $NC_{(m)}(\gamma)$, and thus

$$h(\Delta^m(\Phi), y) = \sum_{w \in NC_{(m)}(\gamma)} y^{rk(w)}.$$

The following lemma will be used as in [4]. For $\alpha \in \Pi$ we denote by Φ_α the standard parabolic root subsystem obtained by intersecting Φ with the linear span of $\Pi \setminus \{\alpha\}$.

LEMMA 5.1. *Let Φ be irreducible, $\alpha \in \Pi$, and $\sigma \subseteq \Phi_{>-1}^m$.*

(i) *For $\sigma \in \Delta^m(\Phi)$ we have $-\alpha \in \sigma$ if and only if $\sigma \setminus \{-\alpha\} \in \Delta^m(\Phi_\alpha)$.*

(ii) *For any $\beta^\ell \in \Phi_{>0}^m$ there exists i such that $\mathcal{R}_m^i(\beta^\ell) \in (-\Pi)$.*

(iii) *$\sigma \in \Delta^m(\Phi)$ if and only if $\mathcal{R}_m(\sigma) \in \Delta^m(\Phi)$.*

Proof. Part (i) can be verified from [14, Theorem 3.7]. Part (ii) follows from Definition 2.12 and the fact that for every $\beta \in \Phi^+$ there exists some j such that $\mathcal{R}^j(\beta) \in (-\Pi)$ [15, section 3]. Finally, part (iii) is clear from Theorem 2.13(i). $\square$

We continue with some technical lemmas required in the proof of Theorem 1.2.

LEMMA 5.2. *If σ is a face of $\Delta = \Delta^m(\Phi)$ and $w_\sigma = w_{\sigma_+} w_{\sigma^m} w_{\sigma^{m-1}} \dots w_{\sigma^1} w_{\sigma_-}$, then*

$$h(\text{lk}_\Delta(\sigma), y) = \sum_{\alpha \in NC_{(m)}(\gamma w_\sigma^{-1})} y^{rk(\alpha)}.$$

Proof. The proof is analogous to the proof of [4, Lemma 2.6], replacing $\mathcal{R}$ by $\mathcal{R}_m$ and using Lemma 5.1. $\square$

LEMMA 5.3. *If $(a_1, \dots, a_m) \leq (b_1, \dots, b_m)$ in $NC_{(m)}(\gamma)$, then $(a_1^{-1}b_1, \dots, a_m^{-1}b_m) \in NC_{(m)}(\gamma)$ and the intervals $I = [(a_1, \dots, a_m), (b_1, \dots, b_m)]$ and $I' = [(\mathbf{1}, \dots, \mathbf{1}), (a_1^{-1}b_1, \dots, a_m^{-1}b_m)]$ are isomorphic.*

Proof. Let $(a_1, \dots, a_m) \leq (b_1, \dots, b_m)$ in $NC_{(m)}(\gamma)$ so that $a_i \leq b_i$ for all $1 \leq i \leq m$. It follows from Lemma 2.3(iii) that since $a_i \leq b_i$, then $a_i^{-1}b_i \leq b_i$. Hence $(a_1, a_1^{-1}b_1, \dots, a_m, a_m^{-1}b_m)$ is a minimal factorization of $b_1 \dots b_m$, and in view of Lemma 2.2(iii) we have that $(a_1^{-1}b_1, \dots, a_m^{-1}b_m) \in NC_{(m)}(\gamma)$. This proves the first part of our statement.

Next, we define the map $\varphi : I \rightarrow I'$ with $\varphi((w_1, \dots, w_m)) = (a_1^{-1}w_1, \dots, a_m^{-1}w_m)$. To prove that the intervals I and I' are isomorphic, we claim that φ is an order-preserving bijection and that its inverse φ^{-1} is order-preserving as well. We first have to check that φ and φ^{-1} are well-defined. Indeed, if $(w_1, \dots, w_m) \in I$, then $(a_1, a_1^{-1}w_1, \dots, a_m, a_m^{-1}w_m)$ is a minimal factorization of $w_1 \dots w_m$, and therefore $(a_1^{-1}w_1, \dots, a_m^{-1}w_m) \in NC_{(m)}(\gamma)$ by Lemma 2.2(iii). Conversely, let $(w_1, \dots, w_m) \in I'$ so that $w_i \leq a_i^{-1}b_i \leq b_i$ for $1 \leq i \leq m$. In view of Lemma 2.3(i) we have $a_i \leq a_i w_i \leq b_i$, and therefore $(a_1 w_1, w_1^{-1} a_1^{-1} b_1, \dots, a_m w_m, w_m^{-1} a_m^{-1} b_m)$ is a minimal factorization of $b_1 \dots b_m$. Thus $(a_1 w_1, \dots, a_m w_m) \in NC_{(m)}(\gamma)$ as in the previous situation. It is immediate to see that φ is a bijection. Moreover, by the fact that $a_i \leq w_i \leq b_i$ if and only if $a_i^{-1}w_i \leq a_i^{-1}b_i$ (Lemma 2.3(i)), we deduce that both φ and φ^{-1} are order-preserving. This completes our proof. $\square$

LEMMA 5.4. *Let $(w_1, \dots, w_m) \in NC_{(m)}(\gamma)$. There is a rank-preserving bijection between the sets $A = \{(a_1, \dots, a_m) : (a_1, \dots, a_m) \leq (a_1 w_1, \dots, a_m w_m) \text{ in } NC_{(m)}(\gamma)\}$ and $NC_{(m)}(\gamma w_m^{-1} \dots w_1^{-1})$.*

Proof. Let $a'_i = w_1 \dots w_{i-1} a_i w_{i-1}^{-1} \dots w_1^{-1}$ for $1 \leq i \leq m$, and consider the map $\phi(a_1, \dots, a_m) = (a'_1, \dots, a'_m)$. We will prove that ϕ is a rank-preserving bijection between the sets A and $NC_{(m)}(\gamma w_m^{-1} \dots w_1^{-1})$. Clearly, the map ϕ as well as its inverse $\phi^{-1}(a'_1, \dots, a'_m) = (a_1, \dots, a_m)$, with $a_i = w_{i-1}^{-1} \dots w_1^{-1} a'_i w_1 \dots w_{i-1}$, are injective. Moreover, since a_i, a'_i are conjugate, then $l_T(a_i) = l_T(a'_i)$, and therefore ϕ is rank-preserving. So it suffices to prove that ϕ is well-defined. To this end let $(a_1, \dots, a_m) \in A$, and note that $a'_1 \dots a'_m = (a_1 w_1 \dots a_m w_m) w_m^{-1} \dots w_1^{-1}$. Since $a_i \leq a_i w_i$, then $l_T(a_i w_i) = l_T(a_i) + l_T(w_i)$, and therefore $(a_1, w_1, \dots, a_m, w_m)$ is a minimal factorization of $a_1 w_1 \dots a_m w_m$. Thus, by Lemma 2.2(iii) we have

$$w_1 \dots w_m \leq a_1 w_1 \dots a_m w_m \leq \gamma,$$

with $l_T(w_1 \dots w_m) = \sum_{i=1}^m l_T(w_i)$. In view of Lemma 2.4(ii) this implies that

$$(a_1 w_1 \dots a_m w_m) w_m^{-1} \dots w_1^{-1} \leq \gamma w_m^{-1} \dots w_1^{-1},$$

with $l_T((a_1 w_1 \dots a_m w_m) w_m^{-1} \dots w_1^{-1}) = \sum_{i=1}^m l_T(a_i)$. Equivalently, $a'_1 \dots a'_m \leq \gamma w_m^{-1} \dots w_1^{-1}$ and $l_T(a'_1 \dots a'_m) = \sum_{i=1}^m l_T(a_i) = \sum_{i=1}^m l_T(a'_i)$, which shows that indeed $(a'_1, \dots, a'_m) \in NC_{(m)}(\gamma w_m^{-1} \dots w_1^{-1})$.

For the converse, let $(a'_1, \dots, a'_m) \in NC_{(m)}(\gamma w_m^{-1} \dots w_1^{-1})$, and notice that $a_1 w_1 \dots a_m w_m = a'_1 \dots a'_m w_1 \dots w_m$. Since $a'_1 \dots a'_m \leq \gamma w_m^{-1} \dots w_1^{-1} \leq \gamma$, then by Lemma 2.3(ii) we have $a'_1 \dots a'_m w_1 \dots w_m \leq \gamma$, with $l_T(a'_1 \dots a'_m w_1 \dots w_m) = \sum_{i=1}^m l_T(a'_i) + \sum_{i=1}^m l_T(w_i)$. This forces $(a_1 w_1, \dots, a_m w_m) \in NC_{(m)}(\gamma)$. We next prove that $(a_1, \dots, a_m) \leq (a_1 w_1, \dots, a_m w_m)$. We have $a'_i \leq \gamma w_m^{-1} \dots w_1^{-1}$ or, equivalently, $w_1 \dots w_{i-1} a_i w_{i-1}^{-1} \dots w_1^{-1} \leq \gamma w_m^{-1} \dots w_1^{-1}$. In view of Lemma 2.5 this implies that $a_i \leq w_{i-1}^{-1} \dots w_1^{-1} \gamma w_m^{-1} \dots w_i^{-1}$. By Lemma 3.8 and elementary calculations one may check that $w_{i-1}^{-1} \dots w_1^{-1} \gamma w_m^{-1} \dots w_i^{-1} \leq w_{i-1}^{-1} \dots w_1^{-1} \gamma w_i^{-1} \leq \gamma w_i^{-1}$. Thus $a_i \leq \gamma w_i^{-1} \leq \gamma$, and therefore $a_i \leq a_i w_i \leq \gamma$ by Lemma 2.3(ii). This completes our proof. $\square$

Proof of Theorem 1.2. To simplify notation let us write Δ^m , Δ_+^m , $F(x, y)$, and $M(x, y)$ instead of $\Delta^m(\Phi)$, $\Delta_+^m(\Phi)$, $F_\Phi^{(m)}(x, y)$, and $M_W^{(m)}(x, y)$, respectively. We use the relation

$$(16) \quad (1-y)^n F\left(\frac{x+y}{1-y}, \frac{y}{1-y}\right) = \sum_{\sigma \in \Delta_+^m} x^{|\sigma|} h(\text{lk}_{\Delta^m}(\sigma), y),$$

which appears in the course of the proof of [4, Theorem 1.1] and can be generalized straightforwardly in the case where $m \geq 1$. We have

$$M(-x, -y/x) = \sum_{\substack{a \leq b \\ a, b \in NC_{(m)}(\gamma)}} \mu(a, b) (-x)^{rk(b) - rk(a)} y^{rk(a)}.$$

Let $\hat{0} = (\mathbf{1}, \dots, \mathbf{1})$, $w = (w_1, \dots, w_m) = (a_1^{-1} b_1, \dots, a_m^{-1} b_m)$, and note that $rk(w) = rk(b) - rk(a)$. In view of Lemmas 5.3 and 5.4 the last sum becomes

$$\sum_{w \in NC_{(m)}(\gamma)} \mu(\hat{0}, w) (-x)^{rk(w)} \sum_{a' \in NC_{(m)}(\gamma w_m^{-1} \dots w_1^{-1})} y^{rk(a')},$$

which by Lemma 4.5 is equal to

$$\sum_{w \in NC_{(m)}(\gamma)} \sum_{\substack{\sigma \in \Delta_+^m(w) \\ |\sigma| = rk(w)}} x^{rk(w)} \sum_{a' \in NC_{(m)}(\gamma w_m^{-1} \dots w_1^{-1})} y^{rk(a')}.$$

If σ is a facet of $\Delta_+^m(w)$, then $w_\sigma = w_{\sigma^m} \dots w_{\sigma^1} = w_1 \dots w_m$, and thus the last expression is equal to

$$\sum_{w \in NC_{(m)}(\gamma)} \sum_{\substack{\sigma \in \Delta_+^m(w) \\ |\sigma| = rk(w)}} x^{|\sigma|} \sum_{a' \in NC_{(m)}(\gamma w_\sigma^{-1})} y^{rk(a')}$$

or, in view of Lemma 5.2,

$$\sum_{w \in NC_{(m)}(\gamma)} \sum_{\substack{\sigma \in \Delta_+^m(w) \\ |\sigma| = rk(w)}} x^{|\sigma|} h(\text{lk}_{\Delta^m}(\sigma), y).$$

Observe that the sets of facets of the subcomplexes $\Delta_+^m(w)$ for $w \in NC_{(m)}(\gamma)$ form a partition of Δ_+^m , so that the last sum is

$$\sum_{\sigma \in \Delta_+^m} x^{|\sigma|} h(\text{lk}_{\Delta^m}(\sigma), y),$$

which, in view of (16), is equal to

$$(1 - y)^n F\left(\frac{x + y}{1 - y}, \frac{y}{1 - y}\right),$$

and this completes our proof. $\square$

Acknowledgments. I thank Drew Armstrong for making available his thesis [1] at an early stage, and Christos Athanasiadis for useful comments and discussions. I also thank the anonymous referees for their thorough comments.

REFERENCES

- [1] D. ARMSTRONG, *Generalized Noncrossing Partitions and Combinatorics of Coxeter Groups*, Ph.D. thesis, Department of Mathematics, Cornell University, Ithaca, New York, 2006.
- [2] C. A. ATHANASIADIS, *Generalized Catalan numbers, Weyl groups and arrangements of hyperplanes*, Bull. London Math. Soc., 36 (2004), pp. 294–392.
- [3] C. A. ATHANASIADIS, *On a refinement of the generalized Catalan numbers for Weyl groups*, Trans. Amer. Math. Soc., 357 (2005), pp. 179–196.
- [4] C. A. ATHANASIADIS, *On some enumerative aspects of generalized associahedra*, European J. Combin., 28 (2007), pp. 1208–1215.
- [5] C. A. ATHANASIADIS, T. BRADY, AND C. WATT, *Shellability of noncrossing partition lattices*, Proc. Amer. Math. Soc., 135 (2007), pp. 939–949.
- [6] C. A. ATHANASIADIS, T. BRADY, J. MCCAMMOND, AND C. WATT, *h-vectors of generalized associahedra and noncrossing partitions*, Int. Math. Res. Not., 2006, article 69705.
- [7] C. A. ATHANASIADIS AND E. TZANAKI, *On the enumeration of positive cells in generalized cluster complexes and Catalan hyperplane arrangements*, J. Algebraic Combin., 23 (2006), pp. 355–375.
- [8] C. A. ATHANASIADIS AND E. TZANAKI, *Shellability and higher Cohen-Macaulay connectivity of generalized cluster complexes*, Israel J. Math., to appear.
- [9] D. BESSIS, *The dual braid monoid*, Ann. Sci. École Norm. Sup. (4), 36 (2003), pp. 647–683.
- [10] A. BJÖRNER, *Shellable and Cohen-Macaulay partially ordered sets*, Trans. Amer. Math. Soc., 260 (1980), pp. 159–183.
- [11] T. BRADY AND C. WATT, *$K(\pi, 1)$'s for Artin groups of finite type*, Geom. Dedicata, 94 (2002), pp. 225–250.
- [12] T. BRADY AND C. WATT, *Lattices in finite real reflection groups*, Trans. Amer. Math. Soc., to appear.

- [13] F. CHAPOTON, *Enumerative properties of generalized associahedra*, Sémin. Lothar. Combin., 51 (2004/05), article B51b.
- [14] S. FOMIN AND N. READING, *Generalized cluster complexes and Coxeter combinatorics*, Int. Math. Res. Not., 44 (2005), pp. 2709–2757.
- [15] S. FOMIN AND A. V. ZELEVINSKY, *Y-systems and generalized associahedra*, Ann. of Math., 158 (2003), pp. 977–1018.
- [16] S. FOMIN AND A. V. ZELEVINSKY, *Cluster algebras I: Foundations*, J. Amer. Math. Soc., 15 (2002), pp. 497–529.
- [17] C. KRATTENTHALER, *The F-triangle of the generalized cluster complex*, in Topics in Discrete Mathematics, Algorithms Combin. 26, Springer, Berlin, 2006, pp. 93–126.
- [18] C. KRATTENTHALER, *The M-triangle of the generalized noncrossing partitions for the types E_7 and E_8* , Sémin. Lothar. Combin., 54 (2005/06), article B54l.
- [19] R. P. STANLEY, *Enumerative Combinatorics*, Vol. 1, Wadsworth and Brooks/Cole, Pacific Grove, CA, 1986, second printing, Cambridge University Press, Cambridge, UK, 1997.
- [20] E. TZANAKI, *Polygon dissections and some generalizations of cluster complexes*, J. Combin. Theory Ser. A, 113 (2006), pp. 1189–1198.

SWITCHING OPERATIONS FOR HADAMARD MATRICES*

WILLIAM P. ORRICK†

Abstract. We define several operations that switch substructures of Hadamard matrices, thereby producing new, generally inequivalent, Hadamard matrices. These operations have application to the enumeration and classification of Hadamard matrices. To illustrate their power, we use them to greatly improve the lower bounds on the number of equivalence classes of Hadamard matrices in orders 32 and 36 to 3,578,006 and 18,292,717.

Key words. Hadamard equivalence, Hadamard matrix, integer equivalence, Smith normal form, Hadamard code

AMS subject classifications. Primary, 05B20; Secondary, 11C20, 15A36, 94B05

DOI. 10.1137/050641727

1. Introduction. Two matrices, A and B , with entries in the set $\{-1, 1\}$ are *Hadamard equivalent* if B can be obtained from A by some sequence of

- row negations,
- column negations,
- row permutations, and
- column permutations.

Hadamard equivalence is so named because of its connection with Hadamard matrices, defined as square matrices with elements equal to ± 1 whose rows are mutually orthogonal. The listed moves all preserve the property of being a Hadamard matrix.

In this paper, we describe some additional moves, called *switching operations*, that preserve the property of being a Hadamard matrix. These operations, when applied over and over again to a seed matrix, generally produce many inequivalent Hadamard matrices.

Furthermore, adjoining the new operations to the list above gives new notions of equivalence. These weaker notions of equivalence may be useful in the classification of Hadamard matrices since they partition the set of Hadamard matrices into a much smaller number of equivalence classes than does Hadamard equivalence, but at the same time provide an effective method for enumerating the elements of these newly defined equivalence classes.

Extensive calculation indicates that the number of Hadamard equivalence classes that can be constructed using the new operations is enormous. This is a big step forward since, although complete enumerations up to order 28 suggest that the number of equivalence classes grows rapidly in higher order, until now there has been no general method for producing the vast numbers of equivalence classes that we expect to exist. The many known Hadamard matrix construction techniques typically apply only in scattered orders, or tend to produce Hadamard matrices with special features such as large automorphism groups, large Hadamard submatrices, or self-duality.

*Received by the editors October 3, 2005; accepted for publication (in revised form) October 8, 2007; published electronically February 6, 2008.

<http://www.siam.org/journals/sidma/22-1/64172.html>

†Department of Mathematics, Indiana University, Bloomington, IN 47405 (worrick@indiana.edu).

The most prolific method for constructing Hadamard matrices has been to use two Hadamard matrices A and B , both of size n , to build Hadamard matrices of size $2n$,

$$(1.1) \quad H = \begin{bmatrix} A & PB \\ A & -PB \end{bmatrix} \quad \text{and} \quad \tilde{H} = \begin{bmatrix} A & A \\ BP & -BP \end{bmatrix},$$

where P is any permutation matrix. Both A and B can be taken from any equivalence class. In order 32, Lin, Wallis, and Lie [25] produced at least 66099 inequivalent matrices from the five equivalence classes in order 16. Since the resulting matrices contain Hadamard submatrices of order 16, however, they cannot be considered generic. In contrast, the new operations produce at least 3.57 million equivalence classes, most of which do not contain Hadamard submatrices of order 16.

Lam, Lam, and Tonchev [22, 23] have exercised great ingenuity in deriving lower bounds on the number of Hadamard matrices of size $2n$ of the form (1.1) and have produced spectacularly large bounds in orders 40 and higher. If the lessons learned from order 32 are any guide, the true numbers of Hadamard equivalence classes in these orders are far greater still.

Our results are even more striking in orders congruent to 4 (mod 8) since the construction (1.1) does not apply. The previously known equivalence classes in order 36 numbered in the hundreds. By the new methods, at least 18.29 million classes can be produced.

The seed matrices used to obtain all these new equivalence classes were derived from the Hadamard matrix literature up to 2005. After this work was substantially complete, Bouyukliev, Fack, and Winne announced the classifications of 2-(31, 15, 7) and 2-(35, 17, 8) designs with automorphisms of odd prime order. From these designs, they found tens of thousands of new Hadamard equivalence classes in orders 32 and 36 [4, 5]. Most of these matrices have not yet been analyzed by our method. Compared with the analysis of the dozens of previously known Hadamard equivalence classes in order 32 (excluding the matrices from construction (1.1)), and the hundreds of previously known H-classes in order 36, analyzing these new matrices is a major undertaking and will require considerable optimization of our methods. Therefore, with one important exception, we have not used the matrices of Bouyukliev, Fack, and Winne in our enumeration, although we make a few remarks on our preliminary analysis in the next paragraph. The exception is a matrix of order 36 in Smith class 16 (defined in section 4.3), no previous example of which appears to have been known. This was used as a seed matrix to produce a new family containing at least five million Hadamard equivalence classes.

In orders 4, 8, 12, 16, 20, 24, 28 the numbers of Hadamard equivalence classes are known to be 1, 1, 1, 5, 3, 60, 487 [11, 12, 13, 18, 19]. We define a weaker notion of equivalence, which we call *Q-equivalence*, by adjoining the new operations to the operations that define Hadamard equivalence. The numbers of Q-equivalence classes are 1, 1, 1, 1, 1, 2, 2. In order 32, we find that the 3.57 million known Hadamard equivalence classes are grouped into 11 Q-equivalence classes, and that in order 36, the 18.29 million known equivalence classes are grouped into 21 Q-equivalence classes. As mentioned above, these numbers do not include Hadamard equivalence classes or Q-equivalence classes derived from the recently discovered matrices of Bouyukliev, Fack, and Winne. An analysis of their matrices should provide a good test of the ideas of this paper regarding using Q-equivalence in classifying Hadamard matrices. Preliminary analysis of a sampling of the new matrices does not turn up any new large Q-classes

but does indicate the presence of a large number of new small Q-classes (perhaps in the hundreds or more). We intend to make a complete enumeration of these and a full analysis of all the new matrices. The results will be presented in a follow-up to the present paper.

2. Overview of switching. Suppose that an $n \times n$ Hadamard matrix can be put in the form

$$(2.1) \quad \begin{bmatrix} \mathbf{1} & \cdots & \mathbf{1} & - & \cdots & - & - & \cdots & - & 1 & \cdots & 1 \\ \mathbf{1} & \cdots & \mathbf{1} & - & \cdots & - & 1 & \cdots & 1 & - & \cdots & - \\ \mathbf{1} & \cdots & \mathbf{1} & 1 & \cdots & 1 & - & \cdots & - & - & \cdots & - \\ \mathbf{1} & \cdots & \mathbf{1} & 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ & a_5 & & b_5 & & c_5 & & d_5 & & & & \\ & \vdots & & \vdots & & \vdots & & \vdots & & & & \\ & a_n & & b_n & & c_n & & d_n & & & & \end{bmatrix},$$

where a_i , b_i , c_i , and d_i are $\{-1, 1\}$ -vectors of length $n/4$. The columns of the matrix have been grouped into four sets of $n/4$ columns each. A new, generally inequivalent, Hadamard matrix can be obtained by negating the $4 \times \frac{n}{4}$ block of 1's in the upper left corner (shown in boldface). We call this operation *switching a closed quadruple*.

Suppose instead that we can put the matrix in the form

$$(2.2) \quad \begin{bmatrix} 1 & - & - & - & \mathbf{1} & \cdots & \mathbf{1} & 1 & \cdots & 1 & 1 & \cdots & 1 & - & \cdots & - \\ - & 1 & - & - & \mathbf{1} & \cdots & \mathbf{1} & - & \cdots & - & - & \cdots & - & - & \cdots & - \\ - & - & 1 & - & \mathbf{1} & \cdots & \mathbf{1} & - & \cdots & - & 1 & \cdots & 1 & 1 & \cdots & 1 \\ - & - & - & 1 & \mathbf{1} & \cdots & \mathbf{1} & 1 & \cdots & 1 & - & \cdots & - & 1 & \cdots & 1 \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & & & & & & & & & & & & \\ \vdots & \vdots & \vdots & \vdots & A_{11} & & A_{12} & & A_{13} & & A_{14} & & & & & \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & & & & & & & & & & & & \\ - & 1 & 1 & - & & & & & & & & & & & & \\ \vdots & \vdots & \vdots & \vdots & A_{21} & & A_{22} & & A_{23} & & A_{24} & & & & & \\ - & 1 & 1 & - & & & & & & & & & & & & \\ - & 1 & - & 1 & & & & & & & & & & & & \\ \vdots & \vdots & \vdots & \vdots & A_{31} & & A_{32} & & A_{33} & & A_{34} & & & & & \\ - & 1 & - & 1 & & & & & & & & & & & & \\ 1 & 1 & - & - & & & & & & & & & & & & \\ \vdots & \vdots & \vdots & \vdots & A_{41} & & A_{42} & & A_{43} & & A_{44} & & & & & \\ 1 & 1 & - & - & & & & & & & & & & & & \end{bmatrix},$$

where the A_{ij} are square matrices of size $(n-4)/4$. A new, often inequivalent, matrix can be obtained by negating the all 1 block of size $4 \times \frac{n-4}{4}$ contained in the first four rows, and the all 1 block of size $\frac{n-4}{4} \times 4$ contained in the first four columns (both shown in boldface). We call this operation *switching a Hall set*.

Justification for these claims and further elaboration are given in the subsequent sections.

3. Closed quadruples and Hall sets.

3.1. 3-normalization. Let H be a Hadamard matrix of size n . Denote its rows by h_i and its elements by h_{ij} . Define the Hadamard product of two vectors to be

$$(a_1, \dots, a_n) \circ (b_1, \dots, b_n) := (a_1 b_1, \dots, a_n b_n).$$

Let j_k be the all 1 vector of length k .

DEFINITION 3.1. A Hadamard matrix of size n is 3-normalized on rows (i, j, k) if, in every column ℓ , the set $\{h_{i\ell}, h_{j\ell}, h_{k\ell}\}$ contains an even number of -1 's, or equivalently, if $h_i \circ h_j \circ h_k = j_n$.

3-normalization is a normalization of the columns. A 3-normalized matrix remains 3-normalized if any two of the rows i, j, k , or any single row other than i, j, k , is negated. Note that 3-normalization was introduced in [31]. The definition given here is slightly weaker in that it makes no stipulation that the row sums be positive and does not impose any particular ordering on the columns. In the next paragraph we restate some needed results from [31].

The field structure (C_1, C_2, C_3, C_4) of a 3-normalized Hadamard matrix of size n is the partition of the set of columns c into four classes, C_i , accordingly as

$$(h_{jc}, h_{kc}, h_{\ell c}) = (1, 1, 1), (-1, -1, 1), (-1, 1, -1), \text{ or } (1, -1, -1).$$

The four classes are called *fields* and are all of length $n/4$. In a row $r \notin \{j, k, \ell\}$ the sum of the elements in a field is the same for each of the four fields in the row. This follows from orthogonality of row r with rows j, k, ℓ . Since the sum of the entries in a field is even if $n/4$ is even, and odd if $n/4$ is odd, the row sum of row $r \notin \{j, k, \ell\}$ must be congruent to $n \pmod{8}$.

A quadruple of rows (i, j, k, ℓ) of a Hadamard matrix H of size n is said to be of *type* r , $0 \leq r \leq n/8$, if exactly $4r$ of the entries in $h_i \circ h_j \circ h_k \circ h_\ell$ equal -1 or exactly $4r$ entries equal $+1$. This notion was introduced by Kimura [19].

DEFINITION 3.2. A quadruple of rows (i, j, k, ℓ) of a Hadamard matrix H of size n is closed if $h_i \circ h_j \circ h_k \circ h_\ell = \pm j_n$.

A closed quadruple is a quadruple of type 0. Thus if H is 3-normalized on three rows of a closed quadruple, then the fourth will consist entirely of 1's or entirely of -1 's. The field structure is independent of which three rows of the closed quadruple are chosen.

Quadruples of type 1 will also play an important role in what follows. They were used extensively by Hall in the classification of Hadamard matrices of order 20 [12] and by Kimura in the classification for order 28 [20, 19]. If H is 3-normalized on three rows of a type-1 quadruple, then the fourth row will contain one odd-sign entry in each of the fields induced by the 3-normalization. Kimura and Ohmori referred to such quadruples as *Hall sets* [21].

PROPOSITION 3.3. If a Hadamard matrix of size n has a closed quadruple, then $n = 4$ or $n \equiv 0 \pmod{8}$.

Proof. Let (i, j, k, ℓ) be the closed quadruple. 3-normalize the matrix on rows i, j, k so that $h_\ell = \pm j_n$. Orthogonality implies that all rows except for h_ℓ have row sum 0. All row sums of rows other than i, j, k must be congruent to $n \pmod{8}$. If $n > 4$, this can happen only when $n \equiv 0 \pmod{8}$. $\square$

3.2. Obtaining new Hadamard matrices by switching closed quadruples.

DEFINITION 3.4. Let H be a Hadamard matrix of size n which has a closed quadruple, Q . Let (C_1, C_2, C_3, C_4) be the partition of columns induced by 3-normalization on Q . Switching the closed quadruple Q means negating all the elements h_{rc} , where $r \in Q$ and $c \in C_i$ for some $i \in \{1, 2, 3, 4\}$.

PROPOSITION 3.5. The matrix produced by switching a closed quadruple Q in a Hadamard matrix H is a Hadamard matrix.

Proof. Any matrix containing a closed quadruple is Hadamard equivalent to one of the form (2.1). It is evident that switching preserves orthogonality of the columns in that matrix. Since column orthogonality is preserved under the operations needed to put H in the form (2.1), the conclusion holds generally. $\square$

It appears that when $n > 8$, switching always produces a Hadamard matrix that is inequivalent to the original Hadamard matrix.

Note that the equivalence class of the Hadamard matrix produced by switching Q is independent of which of the four fields C_i we choose to negate. To see this, note that negating the closed quadruple elements in C_2 is equivalent to first negating the closed quadruple elements in C_1 , then negating all four rows of the closed quadruple, and finally performing a certain permutation of the rows of Q . The same holds for C_3 and C_4 .

3.3. More general row switching operations. It was observed by Denniston [7] in connection with symmetric $(25, 9, 3)$ designs that, starting from a design, a new inequivalent design can be obtained by switching a substructure known as an oval. Denniston's switching operation can be thought of as an operation that permutes certain elements of the incidence matrix of the design. In fact, if 0's are replaced with -1 's in the incidence matrix of a $(25, 9, 3)$ design, ovals satisfy our definition of a closed quadruple, and our switching operation is equivalent to Denniston's.

We can formulate more general switching operations acting on more general structures, which we will refer to generically as "designs." Consider a set, $\mathcal{M}$, of matrices of a fixed size which represent the designs in question. If $R \in \mathcal{M}$, we suppose that the elements of R are taken from some set $\mathcal{S}$, that the rows satisfy some set of properties $\mathcal{P}$, and furthermore, that R satisfies

$$R^T R = M,$$

where M is some fixed matrix. For example, if $\mathcal{M}$ represents $(25, 9, 3)$ designs, then $\mathcal{S} = \{0, 1\}$, the set $\mathcal{P}$ contains the property that every row of $R \in \mathcal{M}$ has exactly nine 1's, and $M = 6I + 3J$, where J is the 25×25 all 1 matrix. If $\mathcal{M}$ represents Hadamard matrices of size n , then $\mathcal{S} = \{-1, 1\}$, the set $\mathcal{P}$ is empty, and $M = nI$. Many other types of matrices and designs, including certain D-optimal designs, can also be defined within this framework.

Let $R \in \mathcal{M}$ and partition the incidence matrix into two submatrices, A and X ,

$$R = \begin{bmatrix} A \\ X \end{bmatrix}.$$

Now suppose that B is a matrix of the same dimensions as A , with elements taken from the same set $\mathcal{S}$, satisfying the same properties $\mathcal{P}$, and that $B^T B = A^T A$. Then the matrix obtained from R by replacing A with B is also a matrix of the original type.

Suppose, for example, that R is an $n \times n$ Hadamard matrix with an $m \times n$ submatrix A whose columns are all identical to columns of a particular $m \times m$ Hadamard matrix H_m or to negations of such columns. Let H'_m be another $m \times m$ Hadamard matrix. Denote column j of H_m by v_j and column j of H'_m by v'_j . Define $[a, b]$ to be the set of integers i satisfying $a \leq i \leq b$. For $j = 1, \dots, n$, let column j of A be $\sigma(j)v_{a(j)}$, where $\sigma : [1, n] \rightarrow \{-1, 1\}$ and $a : [1, n] \rightarrow [1, m]$. Let B be the matrix with the same dimensions as A whose columns are $\sigma(j)v'_{a(j)}$ for $j = 1, \dots, n$. Then B will satisfy $B^T B = A^T A$, and so we may use it to obtain a new Hadamard matrix of order n .

Note that if $m = 1$ and we let $H_1 = [1]$ and $H'_1 = [-1]$, then the above operation amounts to negation of a row. Likewise, if $m = 2$ and $H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ while $H'_2 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, then the above operation amounts to swapping two rows.

Switching closed quadruples is an instance of the $m = 4$ case. Let A be a $4 \times n$ matrix whose columns, or their negations, are columns of H_4 , a 4×4 Hadamard matrix. Orthogonality of the rows of A implies that A is a closed quadruple. Now negating one column of H_4 and using the resulting matrix to construct B has the effect of switching the closed quadruple formed by the rows of A . Thus in some sense switching closed quadruples is a natural extension of the operations of row negation and row permutation.

3.4. Closed quadruples and Hadamard submatrices. There is an additional sense in which switching closed quadruples is a natural extension of the operation of row permutation. Consider the matrix H of size $2n$ defined in (1.1). One may negate or permute the columns of A or B without changing the equivalence class of H . One may also negate a row of PB (or of A) without changing the equivalence class of H . The reason is that negating row j of PB amounts to swapping rows j and $n + j$ of H .

On the other hand, changing the permutation P , for example, by performing the additional row swap (i, j) , usually does change the equivalence class of H . The additional swap will affect four rows of H , namely, $i, j, i + n, j + n$. These four rows form a closed quadruple. One of the four fields of this quadruple is the set of columns of H in which rows i and j of PB differ. We make the switch that negates the entries in rows $i, j, i + n$, and $j + n$ that lie within this field. The result is identical to the result of swapping rows i and j of PB . Therefore, in this context, switching a closed quadruple amounts to swapping a pair of rows in one of the two matrices from which H was constructed.

3.5. Properties of Hall sets. Hall sets play the role for matrices of order $n \equiv 4 \pmod{8}$ that closed quadruples play for matrices of order $n \equiv 0 \pmod{8}$.

Hall sets can be found both in Hadamard matrices of order $n \equiv 0 \pmod{8}$ and in those of order $n \equiv 4 \pmod{8}$. Four columns are singled out in the definition of a Hall set, namely, the columns whose sign in the Hadamard product differs from the sign of all the other columns. When $n \equiv 4 \pmod{8}$ these form a Hall set in the columns, as shown by Kimura and Ohmori [21]. For convenience of the reader, we reprove this here. We include the corresponding result for $n \equiv 0 \pmod{8}$ for good measure. Define the *Hall columns* to be the four distinguished columns. There is one Hall column in each field.

PROPOSITION 3.6. *Let H be a Hadamard matrix of order n . If $n \equiv 0 \pmod{8}$, then the Hall columns form a closed quadruple. If $n \equiv 4 \pmod{8}$, then the Hall columns form a Hall set.*

Proof. We assume without loss of generality that H is 3-normalized on three rows of the Hall set. Consider a row not contained in the Hall set. Let x_i denote the element of that row in the Hall column of field i . Let a_i denote the sum of the remaining elements of field i . Then orthogonality with the Hall set rows implies

$$(3.1) \quad x_1 + a_1 = x_2 + a_2 = x_3 + a_3 = x_4 + a_4,$$

$$(3.2) \quad a_1 + a_2 + a_3 + a_4 = x_1 + x_2 + x_3 + x_4,$$

which implies that the row sum, which must be congruent to $n \pmod{8}$, equals $2(x_1 + x_2 + x_3 + x_4)$. Hence the product $x_1 x_2 x_3 x_4$ is positive for $n \equiv 0 \pmod{8}$ and negative for $n \equiv 4 \pmod{8}$. In each row of the Hall set, the product of the four elements in Hall columns is always positive, so the result follows. $\square$

Remark. When $n \equiv 0 \pmod{8}$ the existence of a Hall set implies the existence of a closed quadruple in the columns, but the converse is not true. The existence of a closed quadruple does not imply the existence of a corresponding Hall set.¹

Henceforth we will consider the $n \equiv 4 \pmod{8}$ case, and when we speak of a Hall set, we will mean both the four rows of the set and the four corresponding Hall columns.

By permuting the Hall rows and columns to the top- and leftmost positions and normalizing appropriately, we obtain the form

$$(3.3) \quad H = \begin{bmatrix} H_4 & F_1 & F_2 & F_3 & F_4 \\ G_1 & A_{11} & A_{12} & A_{13} & A_{14} \\ G_2 & A_{21} & A_{22} & A_{23} & A_{24} \\ G_3 & A_{31} & A_{32} & A_{33} & A_{34} \\ G_4 & A_{41} & A_{42} & A_{43} & A_{44} \end{bmatrix},$$

where

$$(3.4) \quad H_4 = \begin{bmatrix} 1 & - & - & - \\ - & 1 & - & - \\ - & - & 1 & - \\ - & - & - & 1 \end{bmatrix}, \quad F_1 = \begin{bmatrix} 1 & \dots & 1 \\ 1 & \dots & 1 \\ 1 & \dots & 1 \\ 1 & \dots & 1 \end{bmatrix}, \quad F_2 = \begin{bmatrix} 1 & \dots & 1 \\ - & \dots & - \\ - & \dots & - \\ 1 & \dots & 1 \end{bmatrix},$$

$$F_3 = \begin{bmatrix} 1 & \dots & 1 \\ - & \dots & - \\ 1 & \dots & 1 \\ - & \dots & - \end{bmatrix}, \quad F_4 = \begin{bmatrix} - & \dots & - \\ - & \dots & - \\ 1 & \dots & 1 \\ 1 & \dots & 1 \end{bmatrix},$$

$G_1 = F_1^T$, $G_j = -F_j^T$ for $j \in \{2, 3, 4\}$, and A_{ij} are submatrices whose row and column sums equal 2 when $i = j$ and 0 when $i \neq j$.

DEFINITION 3.7. By switching a Hall set in the matrix H defined in (3.3) we mean the operation of replacing F_i by its negation and G_i by its negation for one of the choices $i = 1, 2, 3, 4$.

The four possible negations of the definition produce equivalent matrices. The proof of this is similar to the proof of the analogous property of closed quadruples

¹Sylvester matrices exhibit this in extreme form. One can show by induction on k that the Sylvester matrix of size 2^k has $\frac{1}{4}\binom{2^k}{3}$ closed quadruples and that any of its other row quadruples is of type 2^{k-3} (which means its Hadamard product has as many entries $+1$ as -1). Therefore, if $k \geq 4$, the Sylvester matrix has many closed quadruples but no Hall sets. Since Sylvester matrices are self-dual, the same is true of column quadruples.

given in the discussion following Proposition 3.5. Switching is well defined even when the Hall rows and columns do not appear in positions 1–4 or when the normalization is different from the one in (3.3). We need only apply a signed permutation to put the matrix into the form (3.3), switch as in the definition, and then apply the inverse signed permutation.

PROPOSITION 3.8. *The matrix produced by switching a Hall set in a Hadamard matrix is a Hadamard matrix.*

Proof. We will assume the form (3.3) since the conclusion is unaffected by the permutations and negations needed to convert the matrix to that form. When $i \neq j$, the rows of F_j are orthogonal to the rows of A_{ij} as the latter have row sum 0. Therefore, negating F_j does not alter the orthogonality of rows 1–4 of H with the rows of H contained in the block $[G_i \ A_{i1} \ A_{i2} \ A_{i3} \ A_{i4}]$. Row k ($k = 1, 2, 3, 4$) of F_j has inner product ± 2 with any of the rows of A_{jj} , while row k of H_4 has inner product ∓ 2 with any of the rows of G_j . Negating both F_j and G_j produces sign changes in these inner products that produce opposite contributions to any of the inner products of rows 1–4 of H with the rows of H contained in the block $[G_j \ A_{j1} \ A_{j2} \ A_{j3} \ A_{j4}]$. $\square$

Examples are known in which switching a Hall set in a Hadamard matrix H produces a matrix equivalent to H . In general, however, one obtains an inequivalent matrix.

4. Invariants. First, we prove that the number of closed quadruples in a Hadamard matrix of size $16k + 8$ is invariant under switching closed quadruples. Second, we show that the binary, doubly even, self-dual code associated to the transpose of a Hadamard matrix of size $16k + 8$ is unchanged by switching a closed quadruple of that matrix. Finally, we show that the integer equivalence class is preserved under switching of Hall sets.

4.1. A closed quadruple switching invariant for $n \equiv 8 \pmod{16}$. We will need to understand the ways that closed row quadruples may overlap within a Hadamard matrix.

PROPOSITION 4.1. *Suppose (i, j, k, ℓ) and (i', j', k', ℓ') are distinct closed quadruples with nonempty intersection. Then the number of rows common to the two quadruples is 2 if $n \equiv 8 \pmod{16}$ and 1 or 2 if $n \equiv 0 \pmod{16}$.*

Proof. The number of common rows cannot be 3 since the fourth row of a closed quadruple is determined, up to sign, by the other three, and the two quadruples are assumed distinct. Therefore the number of common rows must be either 1 or 2.

We will show that if the number of common rows is 1, then $n \equiv 0 \pmod{16}$.

Assume the number of common rows to be 1 and let $n = 8r$. Take the two quadruples to be $(1, 2, 3, 4)$ and $(1, 5, 6, 7)$, and 3-normalize the matrix on rows 2, 3, 4. Normalize row 1 to have positive entries. By suitable column permutations, the structure of the first five rows can be brought to the form

$$\begin{array}{cccccccc} 1. & 1_r & 1_r & 1_r & 1_r & 1_r & 1_r & 1_r \\ 2. & 1_r & 1_r & -1_r & -1_r & -1_r & -1_r & 1_r \\ 3. & 1_r & 1_r & -1_r & -1_r & 1_r & 1_r & -1_r \\ 4. & 1_r & 1_r & 1_r & 1_r & -1_r & -1_r & -1_r \\ 5. & 1_r & -1_r & 1_r & -1_r & 1_r & -1_r & 1_r \end{array}$$

The form of row 5 is a consequence of the fact that the sum of elements in each of the four fields must be zero. Since $(1, 5, 6, 7)$ is closed, the Hadamard product of rows 6 and 7 equals either row 5 or its negation. By normalizing row 7 appropriately we

may assume the former. Consider the two subfields that compose the first field in the above structure. They will be further subdivided as

$$\begin{array}{llllll} 5. & 1_a & 1_{r-a} & -1_b & -1_{r-b} & \dots \\ 6. & 1_a & -1_{r-a} & 1_b & -1_{r-b} & \dots \\ 7. & 1_a & -1_{r-a} & -1_b & 1_{r-b} & \dots \end{array}$$

The subfields composing the remaining three fields will be subdivided similarly. Because there are r 1's per field in rows 6 and 7, just as in row 5, we have the constraints $a + b = r$ and $a + (r - b) = r$. Therefore $a = b = r - a = r - b = r/2$, and hence r is even. Consequently $n \equiv 0 \pmod{16}$. $\square$

Note that all of the degrees of overlap between closed quadruples allowed by the proposition occur in practice.

PROPOSITION 4.2. *Let $n \equiv 8 \pmod{16}$. Let H be a Hadamard matrix of size n which has a closed row quadruple Q . Switching Q does not change the number of closed row quadruples in H .*

Proof. In the matrix obtained from H by switching Q , the rows of Q still form a closed quadruple. Also, any quadruple, whether closed or not, that does not involve any rows of Q is unaffected by switching. The only way the number of closed quadruples could change is if a closed quadruple were created or destroyed by switching Q . Such a closed quadruple would have to overlap Q (either before or after switching) and would therefore share exactly two of Q 's rows. However, the Hadamard product of any pair of rows in Q is not altered by negation of any of the fields of Q . Hence the Hadamard product of the four rows of a putative overlapping quadruple would be unchanged by such a negation. Therefore, any closed quadruple overlapping Q in two rows remains closed after switching Q . Likewise, any quadruple overlapping Q in two rows which is not closed initially will not be closed after switching Q . $\square$

It is worth pointing out that switching a closed *column* quadruple does change the number of closed row quadruples in general. Furthermore, switching closed row quadruples generally does change the number of closed row quadruples when $n \equiv 0 \pmod{16}$. For example, when $n = 16$, the five equivalence classes of Hadamard matrices have 140, 76, 44, 28, and 28 closed row quadruples. Each of these five classes can be obtained starting from any of the others and performing a series of switches of closed row quadruples.

4.2. Invariant codes. Codes can be associated with Hadamard matrices and are useful in their classification. For our purposes, codes can be thought of as collections of vectors over some finite field. The vectors in a code are called *codewords*, and the *weight* of a codeword is the number of its entries that are nonzero. The *support* of a codeword is the set of positions in which it has a nonzero entry.

One way to associate a *linear* code with a Hadamard matrix of size n is to normalize the columns of the matrix so that all entries in the first row equal 1, then to change all -1 entries to 0, and finally to take the linear span of the rows of the resulting matrix over some finite field $\mathbb{F}_p$, where p is a prime. One could equally well normalize on a row other than the first and the resulting code would be the same. The *dimension* of such a linear code is its dimension as a subspace of $\mathbb{F}_p^n$. If $\mathcal{C}$ is such a linear code, then its *dual code*, $\mathcal{C}^\perp$, is the subspace of $\mathbb{F}_p^n$ consisting of all vectors orthogonal to all codewords in $\mathcal{C}$. Basic linear algebra implies that the dimensions of a code and its dual satisfy $\dim(\mathcal{C}) + \dim(\mathcal{C}^\perp) = n$. If $\mathcal{C} \subseteq \mathcal{C}^\perp$, then $\mathcal{C}$ is said to be *self-orthogonal*. If $\mathcal{C} = \mathcal{C}^\perp$, then $\mathcal{C}$ is said to be *self-dual*.

We will consider only binary codes ($p = 2$) in this paper, but it should be noted

that codes over $\mathbb{F}_p$, p an odd prime, are closely connected with integer equivalence, which is discussed in the next section. The 2-rank of a Hadamard matrix is the same as the dimension of its associated binary code. Two binary codes are *isomorphic* if one can be converted to the other by a permutation of coordinate positions. The following result is proved (in greater generality) in many places. (For example, see [24, section 2.3].)

THEOREM 4.3. *Let H be a Hadamard matrix of size n . Let $\mathcal{C}$ be a binary code associated to H as described above. Then*

- (1) *if $n \equiv 4 \pmod{8}$, then $\mathcal{C} = \{1_n\}^\perp$ which implies $\dim(\mathcal{C}) = n - 1$;*
- (2) *if $n \equiv 0 \pmod{8}$, then $\mathcal{C}$ is self-orthogonal which implies $\dim(\mathcal{C}) \leq n/2$;*
- (3) *if $n \equiv 8 \pmod{16}$, then $\mathcal{C}$ is self-dual which implies $\dim(\mathcal{C}) = n/2$.*

Since all Hadamard matrices of a size congruent to 4 (mod 8) have the same binary code, the binary code does not help with classification (although codes over other fields may). For the present, we focus on binary codes associated with matrices of size $n \equiv 0 \pmod{8}$. It is not hard to show that such codes are *doubly even*; that is, all of their codewords have weight divisible by 4. For an illustration that various 2-ranks allowed by Theorem 4.3 do occur in practice, we consider some results discussed by Assmus and Key in [3, 2]. They note that the five inequivalent Hadamard matrices of size 16 have binary codes of dimensions 5, 6, 7, 8, and 8. Only the last two are self-dual, and they turn out not to be isomorphic. Contrast this with size 24, where the 60 nonequivalent Hadamard matrices must *all* have self-dual codes of dimension 12. Assmus and Key proved that these 60 classes of matrices are associated with six different doubly even, self-dual binary codes.

Key pointed out [17] that when H is a Hadamard matrix of size 24, the number of closed quadruples coincides with the number of codewords of weight 4 in the binary code associated with the *columns* of H . (We might also call this the code associated with H^T .) We elaborate a bit on her observation, which reflects a general phenomenon for matrices of size $n \equiv 8 \pmod{16}$.

PROPOSITION 4.4. *Let H be a Hadamard matrix of size $n \equiv 0 \pmod{8}$. Let $\mathcal{C}$ be the linear binary code constructed from the columns of H . That is, $\mathcal{C}$ is the linear span over $\mathbb{F}_2$ of the columns of a matrix A formed by normalizing H so that one of its columns consists entirely of 1's and then changing -1 's to 0's. Let $\{i, j, k, \ell\}$ be the support of a weight 4 codeword in $\mathcal{C}$. Then rows i, j, k , and ℓ of H form a closed quadruple.*

Proof. Since $\mathcal{C}$ is self-orthogonal, the weight 4 codeword with support $\{i, j, k, \ell\}$ is orthogonal to every column of A . This means that every column of A has an even number of 1's in positions i, j, k , and ℓ , which implies the result. $\square$

This result has a partial converse with self-duality of $\mathcal{C}$ being the needed additional assumption.

PROPOSITION 4.5. *Let H, A , and $\mathcal{C}$ be defined as in Proposition 4.4 and suppose in addition that $\mathcal{C}$ is self-dual. Let i, j, k , and ℓ label the rows of a closed quadruple in H . Then $\{i, j, k, \ell\}$ is the support of a weight 4 codeword in $\mathcal{C}$.*

Proof. Since one column of A consists entirely of 1's, and since i, j, k , and ℓ label a closed quadruple, every column of A has an even number of 1's among the positions i, j, k , and ℓ . Let $\mathbf{c}$ be the vector in $\mathbb{F}_2^n$ with support $\{i, j, k, \ell\}$. Then $\mathbf{c}$ is orthogonal to every column of A and therefore to every codeword in $\mathcal{C}$. Hence $\mathbf{c} \in \mathcal{C}^\perp$. Since $\mathcal{C}$ is self-dual, we also have $\mathbf{c} \in \mathcal{C}$. $\square$

We have established a one-to-one correspondence between the closed quadruples of a Hadamard matrix H of size $n \equiv 0 \pmod{8}$ and weight 4 codewords in the binary code associated to the columns of H , provided that code is self-dual. This correspon-

dence therefore holds for all Hadamard matrices of size congruent to 8 (mod 16).

We finally investigate the effect of switching a closed quadruple of H on the binary code associated to the columns of H . In this connection, we note that the closed quadruple switching operation was defined and used in the coding theory context by Phelps, Rifà, and Villanueva [32]. They were concerned with Hadamard matrices of size 2^t whose codes can range in dimension from $t + 1$ to 2^{t-1} . Starting with a code of minimal dimension, corresponding to the Sylvester matrix, they produced codes, and the corresponding matrices, of the next two higher dimensions by switching. For further details, see Lemmas 4.2 and 4.3 of [32]. Note that closed quadruples correspond to subcodes of dimension 3.

Our focus in this paper will be on codes at the opposite end of the range of possible dimensions, that is, on the self-dual codes. We have the following inclusion of codes.

PROPOSITION 4.6. *Let H , A , and $\mathcal{C}$ be defined as in Proposition 4.4 and suppose in addition that $\mathcal{C}$ is self-dual. Let H' be a Hadamard matrix obtained from H by switching a closed quadruple, and let $\mathcal{C}'$ be the code associated to the columns of H' . Then $\mathcal{C}' \subseteq \mathcal{C}$. Furthermore, if $\mathcal{C}' \neq \mathcal{C}$, then $\mathcal{C}$ is spanned by $\mathcal{C}'$ and a particular weight 4 vector.*

Proof. We may assume that H has been normalized so that all entries in its first column equal 1. The matrix A is then obtained from H simply by replacing -1 's with 0 's. The code $\mathcal{C}$ is the span of the columns of A .

Let (i, j, k, ℓ) be a closed quadruple of H . Proposition 4.5 asserts that $\mathcal{C}$ contains a codeword $\mathbf{c}$ with support $\{i, j, k, \ell\}$. Let $\{C_1, C_2, C_3, C_4\}$ be the partition into fields of the set of columns of H induced by the closed quadruple (i, j, k, ℓ) . Switching means negating all matrix elements in rows i, j, k, ℓ and in the columns of one of the C_m , $m = 1, 2, 3$, or 4 . Should column 1 be one of the affected columns, the normalization of the resulting matrix, H' , will no longer be such that column 1 contains 1 's only. To restore the normalization, we simply negate rows i, j, k, ℓ . The net result will be that all columns *but* those of C_m are affected by the switching.

At any rate, the matrix A' , obtained from H' by changing -1 's to 0 's, will differ from A only in that the elements in rows i, j, k, ℓ and in a certain subset of the columns will have been changed to their complements ($0 \rightarrow 1, 1 \rightarrow 0$). This change can be effected by adding the vector $\mathbf{c}$ to the appropriate columns of A . Therefore, $\mathcal{C}'$, which is the span of the columns of A' , is spanned by a set of linear combinations of codewords in $\mathcal{C}$. Hence $\mathcal{C}' \subseteq \mathcal{C}$. Finally, $\mathcal{C}$ is clearly the span of $\mathcal{C}' \cup \{\mathbf{c}\}$. $\square$

We note that the code $\mathcal{C}'$ obtained in the above proof depends on which of the four fields, C_m , was used in the switching. Nevertheless, the isomorphism class of the code will be independent of this choice.

As an illustration of the use of Proposition 4.6 consider the codes associated with the five equivalence classes of 16×16 Hadamard matrices. Matrices in either of the two classes associated with self-dual codes have 28 closed quadruples. Switching any of these 28 quadruples produces a matrix in the class corresponding to the code of dimension 7. We therefore conclude that the code of dimension 7 is a subspace of both of the codes of dimension 8, and that each of the latter can be obtained by augmenting the code of dimension 7 with a weight 4 vector whose support corresponds to a suitable closed quadruple.

A corollary of Proposition 4.6 is immediate.

COROLLARY 4.7. *Let H' be obtained from a Hadamard matrix H by switching a closed quadruple. If the linear binary codes associated to the columns of H and H'*

are both self-dual, then they are equal. In particular, if H is of size $n \equiv 8 \pmod{16}$, then the linear binary codes associated to the columns of H and H' are equal.

Corollary 4.7 will be important when we discuss classification of Hadamard matrices of size 24. Since in the setting of Corollary 4.7 there is a one-to-one correspondence between closed quadruples and weight 4 codewords, the corollary also provides an alternative proof of Proposition 4.2.

4.3. A Hall set switching invariant. An important notion used in the classification of Hadamard matrices is that of integer equivalence.

DEFINITION 4.8. *Two integer matrices A and B are integer equivalent if A can be converted to B by some sequence of the following row and column operations:*

- *permutation of rows (columns);*
- *negation of rows (columns);*
- *addition of an integer multiple of a row (column) to another row (column).*

Associated to the integer equivalence class of a matrix A of size n is a set of integers $s_1, \dots, s_n$ called *invariant factors* satisfying that

1. the matrix $\text{diag}(s_1, \dots, s_n)$ is integer equivalent to A .
2. there exists r such that $1 \leq r \leq n$ and $s_i | s_{i+1}$ for $1 \leq i \leq r-1$ and $s_{r+1} = \dots = s_n = 0$.
3. the product $s_1 s_2 \dots s_i$ equals the GCD of the $i \times i$ minors of A .

The matrix $\text{diag}(s_1, \dots, s_n)$ is called the *Smith normal form* of A . Two integer equivalent matrices have the same Smith normal form.

A number of properties of the Smith normal form of a Hadamard matrix have been proved [38, 30]:

1. $s_1 = 1$; $s_2 = \dots = s_{\alpha+1} = 2$, for some $\alpha \geq \lfloor \log_2 n \rfloor + 1$;
2. $s_i s_{n+1-i} = n$.

In order 36, for example, we have [6]

- $s_1 = 1$;
- $s_i = 2$ for the next α values of i ($2 \leq i \leq \alpha + 1$);
- $s_i = 6$ for the next $34 - 2\alpha$ values of i ;
- $s_i = 18$ for the next α values of i ;
- $s_{36} = 36$,

where $6 \leq \alpha \leq 17$. The single parameter α determines the integer equivalence class of a Hadamard matrix H in order 36, and we say that H is in *Smith class* α .

That the Smith class is invariant under switching Hall sets is implied by the following.

PROPOSITION 4.9. *If B is obtained from A by switching a Hall set, then B is integer equivalent to A .*

Proof. Switching a Hall set can be achieved by a sequence of integer row and column operations. Let the order of the matrix in (3.3) be $4k + 4$. Adding each of rows 1 through 4 to each of the k rows 5 through $k + 4$, and then adding each of columns 1 through 4 to each of columns 5 through $k + 4$ has the effect of negating F_1 and G_1 . $\square$

5. Equivalence relations. Hadamard equivalence, usually just called “equivalence,” was defined in the introduction. We will call Hadamard equivalence classes *H-classes*. By adjoining additional operations to the list of operations given there, we can define new equivalence relations. We already did this in the previous section when we defined integer equivalence, whose equivalence classes are the Smith classes. The considerations of the previous section also allow us to define *equivalence with respect to the associated binary linear code*, or *code equivalence* for short: Two

Hadamard matrices are code equivalent if their binary linear codes, as defined in the statement of Proposition 4.4, are isomorphic. In this section, we define further notions of equivalence.

DEFINITION 5.1. *If $n \equiv 0 \pmod{8}$, then two Hadamard matrices A and B of size n are Q-equivalent if B can be obtained from A by some sequence of the operations*

- *row or column negation;*
- *row or column permutation;*
- *switching a closed quadruple of rows;*
- *switching a closed quadruple of columns.*

If the last operation is disallowed, then A and B are said to be QR-equivalent; if the third operation is disallowed, then A and B are said to be QC-equivalent. When $n \equiv 4 \pmod{8}$, Q-equivalence is defined by replacing the last two operations with

- *switching a Hall set.*

Associated with these equivalence relations are equivalence classes, called Q-classes, QR-classes, and QC-classes.

Hadamard equivalence is stronger than Q-equivalence and therefore has a more refined equivalence class structure. In other words, there are at least as many H-classes as there are Q-classes, and each H-class is contained entirely within a particular Q-class. QR-equivalence (or QC-equivalence) is intermediate in strength between H-equivalence and Q-equivalence and will therefore have an intermediate number of equivalence classes. When $n \equiv 8 \pmod{16}$, Corollary 4.7 implies that QR-equivalence is a refinement of code equivalence: two Hadamard matrices in the same QR class have isomorphic codes; the converse does not necessarily hold, as will be seen in the case $n = 24$, which is discussed in the next section.

By Proposition 4.9, Q-equivalence is stronger than integer equivalence when $n \equiv 4 \pmod{8}$, which implies that there are at least as many Q-classes as there are Smith classes in those orders.

An equivalence class, of any type, may or may not be *self-dual*. The *dual* of a set of matrices is the set containing their transposes. A set that equals its own dual is self-dual. Many, but not all, Q-classes turn out to be self-dual. In other words, many matrices are Q-equivalent to their transposes. From the row-column symmetry in the definition of Q-equivalence it follows that if a Q-class contains at least one self-dual matrix, then that Q-class is self-dual.

We will see examples of these phenomena in the next section.

6. Application to the enumeration of inequivalent Hadamard matrices.

We remind the reader that Hadamard matrices have been completely classified up to order 28. There are five H-classes in order 16 [11], three in order 20 [12], 60 in order 24 [13, 18], and 487 in order 28 [19]. Using the available lists of H-classes, which can be obtained from a number of sources [33, 34, 36], we will be able to determine the structure of the Q-classes in these orders. The classification of H-classes in orders 32 and higher appears to be very difficult. We will content ourselves with identifying the Q-classes of all Hadamard matrices in orders 32 and 36 that were known before the recent work of Bouyukliev, Fack, and Winne (see the introduction) and completely enumerating those Q-classes that are small enough for this to be feasible.

Our procedure requires that we maintain a database of inequivalent matrices. As new matrices are generated, they are put in a canonical form and compared with known matrices to prevent duplication in the database. To put the matrices in canonical form, we followed the suggestion of Brendan McKay [26], converting $n \times n$ matrices to graphs on $4n$ vertices and then using the graph isomorphism program *nauty* that

he developed [27]. The canonical form of the graph computed by *nauty* was then converted back into a matrix. As suggested in the *nauty* User's Guide [28], we used the vertex invariant *cellquads* at level 2, which improves the efficiency in processing this type of graph.

To generate lists of inequivalent Hadamard matrices of order n we carried out the following procedure, which requires a seed Hadamard matrix of order n as input:

1. Initialize **hadList** to null list.
2. Compute canonical form of seed matrix using *nauty*. Append it to **hadList**.
3. Compute canonical form of transpose of seed matrix. If it differs from canonical form of seed matrix, append it to **hadList**.
4. Initialize **ctr** to 1.
5. Let H be matrix number **ctr** on **hadList**. If $n \equiv 4 \pmod{8}$ and this matrix is in the H-class of the transpose of the previous one, skip to step 7.
6. For each closed row quadruple ($n \equiv 0 \pmod{8}$) or Hall set ($n \equiv 4 \pmod{8}$) in H ,
 - (a) switch the quadruple (Hall set) and compute the canonical form of the resulting matrix to obtain H' .
 - (b) if H' differs from all matrices on **hadList**, append it to **hadList**. Then if the canonical form of the transpose of H' differs from H' , append it to **hadList** as well.
7. Increment **ctr**. If **hadList** is not exhausted, return to step 5.

Note that this procedure generates the Q-class of the seed matrix unless the Q-class happens to be non-self-dual, in which case it generates the union of the Q-class and its dual. This is due to the use of the transposition operation in step 6(b). Non-self-dual Q-classes always turn out to be small, and when the situation arises, we partition the union into two Q-classes by hand. (We could use column quadruple switching in the $n \equiv 0 \pmod{8}$ case and dispense with transposition in both cases, thereby avoiding this issue, but we found it convenient to use transposition to keep track of duality.) We can also modify the procedure by simply eliminating the transposition step, in which case the procedure generates the QR-class of the seed matrix in the $n \equiv 0 \pmod{8}$ case.

Here are the results on the Q-classes and QR-classes for orders 16 and 24:

- $n = 16$: The five H-classes are all Q-equivalent. More strikingly, they are all QR-equivalent.
- $n = 24$: Of the 60 H-classes, 59 are Q-equivalent. The H-class missing from the main Q-class is that of the Paley matrix which has no closed quadruples and is self-dual. It forms a Q-class all by itself.

As stated in section 4.2, Assmus and Key classified the 60 H-classes according to the doubly even binary codes associated to the columns of the matrices. (See Table 1 in [3] or Table 7.1 in [2], but beware that, in both tables, matrix 42_D^{32} , listed with the code D , should be listed with the code C , and that 32_D^{42} in line 3 should be changed to 32_C^{42} .) We use QR-equivalence to refine this classification.

Assmus and Key found that six codes, labeled A , C , D , E , F , and G , occur. They are distinguished by the number of codewords of weight 4 and therefore, according to Propositions 4.4 and 4.5, by the number of closed row quadruples in the associated matrices. We now see Corollary 4.7, on the invariance of codes under switching, in action. For example, the matrices associated with the code D all have 12 closed row quadruples. Switching any of these

TABLE 1
The six codes associated with the 60 Hadamard matrices of order 24.

Code	# weight 4 codewords	Size of code class	Sizes of QR-classes
A	30	8	8
C	18	17	17
D	12	15	5, 10
E	66	8	8
F	6	10	5, 5
G	0	2	1, 1

quadruples produces another matrix with code D . Depending on which of these matrices one starts with, switching row quadruples produces a QR-class of size 5 or of size 10. These two QR-classes together account for all 15 H-classes associated with the code D . Results for all the codes appear in Table 1.

Note that the matrices associated with the $[24, 12]$ extended Golay code G do not contain closed row quadruples. One class of such matrices must be that of the Paley Hadamard matrix, as we have already stated that it has no closed quadruples. There is a second class of matrices with no closed row quadruples. The matrices in this class, however, do each have 66 closed column quadruples, since their duals turn out to be in the class of the code E .

The results on Q-classes in orders 20 and 28 are

- $n = 20$: The 3 H-classes are Q-equivalent.
- $n = 28$: Of the 487 H-classes, 486 of them (the ones containing Hall sets [20]) are Q-equivalent. The Paley matrix (generated from quadratic residues in $\text{GF}(3^3)$) contains no Hall set, and therefore its H-class forms a Q-class by itself.

Before presenting our results in orders 32 and 36, we ask what might the results so far lead us to expect in higher order? It is striking that except for a small number of exceptions (the H-classes of the Paley matrices in orders 24 and 28), all Hadamard matrices of given order are Q-equivalent. Could this be a general phenomenon?

In order 36, a difficulty arises. By Proposition 4.9 the Smith class is invariant under the defining operations of Q-equivalence. We will see that at least six different Smith classes occur, and so there must be at least six Q-classes, each possibly containing many H-classes. The reason the multiplicity of Smith classes was not an issue in order 28 is that $7 = 28/4$ is an odd square-free number. By a result in [39] this implies that all Hadamard matrices in order 28 lie in a single Smith class. From the foregoing discussion, the best we can hope for, for general $n \equiv 4 \pmod{8}$, is that within each Smith class there will be a single dominant Q-class, and that the total number of Q-classes will still be small.

The results we have obtained so far appear to support the idea of a single dominant Q-class in order 32, and of a single dominant Q-class within each Smith class in order 36. The total number of Q-classes also appears to be very small relative to the number of H-classes. We found only a few tens of Q-classes in our analysis of the matrices known prior to the work of Bouyukliev, Fack, and Winne, but a preliminary analysis of their matrices suggests that the number will rise into the hundreds, if not higher. Our method was to collect as many Hadamard matrices as possible from the literature or using known construction techniques, and then to apply our algorithm to each of these

matrices in order to obtain its Q-class. In fortunate cases our program terminated in a reasonable time, giving us a complete enumeration of the elements of the Q-class of the given seed matrix. In less fortunate cases—and if our speculations are correct, this is expected to be the usual situation—the Q-class was too big to enumerate completely. Instead, we compared partially constructed Q-classes with each other and looked for overlaps. By so doing, we managed to identify unambiguously the Q-class of every Hadamard matrix in orders 32 and 36 known to us prior to the work of Bouyukliev, Fack, and Winne, to enumerate the smaller of these Q-classes, and to obtain lower bounds on the sizes of the larger Q-classes.

6.1. Order 32.

PROPOSITION 6.1. *All Hadamard matrices of either of the forms*

$$H = \begin{bmatrix} A & B \\ A & -B \end{bmatrix}, \quad \tilde{H} = \begin{bmatrix} A & A \\ B & -B \end{bmatrix},$$

where A and B are any Hadamard matrices of order 16, are Q-equivalent.

Proof. From the discussion in section 3.4 it follows that, from the matrix $\begin{bmatrix} A & B \\ A & -B \end{bmatrix}$, with A and B fixed, we may obtain any matrix of the form $\begin{bmatrix} A & PB \\ A & -PB \end{bmatrix}$ by switching closed row quadruples.

To show that all matrices of the form H are Q-equivalent we need only show that we can change the H-class of A or of B to any of the five classes in order 16 by switching closed quadruples. Since all Hadamard matrices of order 16 are QR- and QC-equivalent, we can achieve this by switching closed column quadruples in the A columns of H only or in the B columns of H only. (Closed column quadruples of A or of B extend to closed column quadruples of H , and switching transforms the top and bottom halves of a column the same way.)

Analogous arguments, with rows and columns interchanged, show that all matrices of the form $\tilde{H}$ are Q-equivalent. To show that matrices of the form H and of the form $\tilde{H}$ are Q-equivalent to each other, simply note that both sets contain the Sylvester Hadamard matrix. $\square$

Thus the 66099 H-classes identified in [25] are Q-equivalent. We call the Q-class of these matrices the *Sylvester Q-class*. We now turn to other known Hadamard matrices in order 32:

- the Paley matrix,
- 13 matrices from generalized Legendre (GL) pairs [9],
- four matrices listed in [1] and their transposes,
- the maximal excess matrix in [8],
- four matrices from Construction II in [25],
- two Williamson matrices,
- eight Goethals–Seidel matrices constructed from circulant blocks,
- 18 Goethals–Seidel matrices constructed from negacyclic blocks,
- 10 matrices constructed from two circulants,
- 17 matrices constructed from two negacyclic matrices,
- a matrix from the appendix of [23] and its transpose.

Some of these matrices were provided by Hadi Kharaghani. Discarding duplicates (which occur due to accidental equivalences) and matrices that happen to have one of the forms in Proposition 6.1, we are left with a list of 59 matrices. Of these, 49 are in the Sylvester Q-class. Using these matrices, and some matrices from Proposition 6.1 as seeds, we have managed to generate 3,577,996 H-classes in the Sylvester Q-class by using our program and then piecing together the results. This is certainly a gross

underestimate of the actual number.

The ten exceptional matrices among the 59 all lack closed quadruples both in rows and in columns, and therefore form Q-classes by themselves. Of the ten exceptional matrices, six are constructed from GL pairs, and four are constructed from two negacyclic blocks. The matrices from GL pairs are listed on the Web page [33] as P12–P19 (with transposes of nonself-dual matrices omitted). The exceptional GL pair matrices are P13, P15 and its transpose, P17, and P19 and its transpose. Matrix P17 is Hadamard equivalent to the Paley matrix. Of the matrices constructed from two negacyclic blocks, the exceptional ones come in two dual pairs.

The Sylvester Q-class and the 10 singleton Q-classes give a total of 11 known Q-classes in order 32, containing at least 3,578,006 Hadamard equivalence classes.

6.2. Order 36. As noted above, in order 36 we must consider each Smith class separately. Although Smith classes $\alpha = 6, 7, \dots, 17$ are allowed, the only Smith classes known to be nonempty are $\alpha = 11, 12, 13, 14, 15, 16, 17$.

A complete summary of the seed matrices we compiled in order 36 follows:

- Ted Spence’s 180 matrices related to regular 2-graphs (S1–S180) [35, 29, 36];
- the 24 matrices of Goethals–Seidel type classified by Spence and Turyn (GS1–GS24) [36];
- the 11 matrices with automorphism of size 17 classified by Tonchev (T1–T11) [37];
- the Bush-type Hadamard matrix found by Janko (B1) [14];
- a regular Hadamard matrix found by Seberry and listed on her Web page (R1) [33] (she actually lists four, but two are duplicates, and two are of Goethals–Seidel type);
- four Williamson Hadamard matrices (W1–W4) (there is a fifth, but it is equivalent to one of Tonchev’s);
- the (35, 17, 8)-difference set construction (D1);
- seven matrices of the type defined by Whiteman (a Goethals–Seidel array bordered by a Hall set) (Wh1–Wh7) [40];
- two block negacyclic Bush-type Hadamard matrices, the first given in the paper of Janko and Kharaghani (NB1, NB2) [15];
- a matrix in Smith class 11, found in the course of a (fruitless) search for block circulant Bush-type matrices (O1);
- a skew Bush-type Hadamard matrix found by Jørgensen and its transpose (J1, J2) [16];
- a matrix listed in the appendix of [23] (LLT1);
- the first matrix known (to us) in Smith class 16, found by Bouyukliev, Fack, and Winne [4] (BFW1).

Reference [6] was helpful in assembling the above list, but the reader should note that the 80 matrices from Steiner triple systems, which are a subset of Spence’s 180 matrices, are in Smith class 13, not 12 as stated there. We have not made a serious effort to credit the original author of every matrix on our list, as we were more concerned with compiling as complete a list as possible from readily obtainable sources. We should note, however, that many of these matrices derive from the important work of Goethals and Seidel [10], including the 80 matrices from Steiner triple systems mentioned above, and 11 matrices derived from Latin squares of order 6, which are also a subset of Spence’s list.

The structure we have uncovered in Smith class 13 is interesting, so we describe it in detail. Hadamard matrices in this class include 179 of Spence’s 180 matrices.

TABLE 2

Sizes of known Q-classes in order 36 for the six known Smith classes, α . A representative matrix is listed for each Q-class.

α	Q-classes
11	1 (S137), 1 (O1)
12	1 (D1)
13	1 (S179), 1 (S180), 1 (NB1), 6 (S172), 3425 (S1)
14	954,254 (B1)
15	5 (W3), 5 (W4), ≥ 5 , 520, 880 (GS1)
16	≥ 5 , 814, 129 (BFW1)
17	1 (GS11), 1 (GS12), 1 (T1), 1 (T2), 1 (T5), 1 (T6), 1 (T7), $\geq 6,000,000$ (GS4)

(His matrix 137 is in Smith class 11.) Two other matrices in Smith class 13 were previously known: the regular Hadamard matrix constructed by Seberry, and the block negacyclic Bush-type Hadamard matrix constructed by Janko and Kharaghani. Seberry's matrix and 172 of Spence's fall into the same Q-class which we found has size 3425. Two of Spence's matrices (179 and 180) and the Bush-type matrix form singleton Q-classes. They have no Hall sets. The remaining five of Spence's matrices lie in a Q-class of size 6.

Spence's matrix 137, which is in Smith class 11 and is one of the matrices derived from a Latin square of order 6, is intriguing. It has nine Hall sets, but switching any of these produces a matrix H-equivalent to the original.

Only two matrices on our list are in Smith class 14, B1 and LLT1. They are Q-equivalent. A major success of our program has been the complete enumeration of their Q-class, which has 954,254 elements. In each of Smith classes 15, 16, and 17 there is one known Q-class of size above 5 million, while all other known Q-classes are of size no greater than 5. The three large Q-classes have not yet been completely enumerated. At present, there is no evidence for more than one large Q-class in any Smith class.

The 236 matrices we compiled represent seven different Smith classes and lie in 21 different Q-classes. Some details are given in Table 2. The union of the known Q-classes contains at least 18,292,717 Hadamard equivalence classes of order 36.

Acknowledgments. I thank Bruce Solomon for numerous interesting discussions and for careful comments on the manuscript. I thank Hadi Kharaghani for extensive correspondence and for providing many unpublished Hadamard matrices in orders 32 and 36. I thank Jennifer Key for recomputing a table that appears in [3, 2], and for pointing out the connection between closed quadruples and codewords of weight 4. I am indebted to Robin Chapman who pointed out an error in the proof of Proposition 3.8 in an earlier version of the manuscript and who suggested the simple proof of Proposition 3.5. Iliya Bouyukliev, Veerle Fack, and Joost Winne kindly provided me with a list of the matrices they obtained in their work on classification of 2-(35, 17, 8)-designs. Jayoung Nam made numerous helpful suggestions about presentation. I thank the editor, Kevin Phelps, and the anonymous referees for suggestions that improved the paper. This project relied on the High Performance Computing facilities of Indiana University, in particular the IBM RS/6000 SP system.

REFERENCES

- [1] M. ARAYA, M. HARADA, AND H. KHARAGHANI, *Some Hadamard matrices of order 32 and their binary codes*, J. Combin. Des., 12 (2004), pp. 142–146.

- [2] E. F. ASSMUS, JR., AND J. D. KEY, *Designs and Their Codes*, Cambridge Tracts in Math. 103, Cambridge University Press, Cambridge, UK, 1992.
- [3] E. F. ASSMUS, JR., AND J. D. KEY, *Hadamard matrices and their designs: A coding-theoretic approach*, Trans. Amer. Math. Soc., 330 (1992), pp. 269–293.
- [4] I. BOUYUKLIEV, V. FACK, AND J. WINNE, *Hadamard matrices of order 36 and double-even self-dual $[72, 36, 12]$ codes*, in Proceedings of the European Conference on Combinatorics, Graph Theory and Applications 2005, Stefan Felsner, ed., vol. AE of DMTCS Proceedings, Discrete Mathematics and Theoretical Computer Science, 2005, pp. 93–98. Available online at <http://www.dmtcs.org/dmtcs-ojs/index.php/proceedings/article/viewFile/dmAE0119/988>
- [5] I. BOUYUKLIEV, V. FACK, AND J. WINNE, *2-(31, 15, 7), 2-(35, 17, 8) and 2-(36, 15, 6) Designs with Automorphisms of Odd Prime Order, and Their Related Hadamard Matrices and Codes*, preprint, 2006.
- [6] J. COOPER, J. MILAS, AND W. D. WALLIS, *Hadamard equivalence*, in Combinatorial Mathematics (Proc. Internat. Conf. Combinatorial Theory, Australian Nat. Univ., Canberra, 1977), Lecture Notes in Math. 686, Springer, Berlin, 1978, pp. 126–135.
- [7] R. H. F. DENNISTON, *Enumeration of symmetric designs* (25, 9, 3), in Algebraic and Geometric Combinatorics, North-Holland Math. Stud. 65, North-Holland, Amsterdam, 1982, pp. 111–127.
- [8] H. ENOMOTO AND M. MIYAMOTO, *On maximal weights of Hadamard matrices*, J. Combin. Theory Ser. A, 29 (1980), pp. 94–100.
- [9] R. J. FLETCHER, M. GYSIN, AND J. SEBERRY, *Application of the discrete Fourier transform to the search for generalised Legendre pairs and Hadamard matrices*, Australas. J. Combin., 23 (2001), pp. 75–86.
- [10] J.-M. GOETHALS AND J. J. SEIDEL, *Strongly regular graphs derived from combinatorial designs*, Canad. J. Math., 22 (1970), pp. 597–614.
- [11] M. HALL, JR., *Hadamard matrices of order 16*, in Research Summary No. 36-10, Volume I, Jet Propulsion Laboratory, Pasadena, 1961, pp. 21–26.
- [12] M. HALL, JR., *Hadamard Matrices of Order 20*, Tech. report 32-761, Jet Propulsion Laboratory, Pasadena, 1965.
- [13] N. ITO, J. S. LEON, AND J. Q. LONGYEAR, *Classification of 3-(24, 12, 5) designs and 24-dimensional Hadamard matrices*, J. Combin. Theory Ser. A, 31 (1981), pp. 66–93.
- [14] Z. JANKO, *The existence of a Bush-type Hadamard matrix of order 36 and two new infinite classes of symmetric designs*, J. Combin. Theory Ser. A, 95 (2001), pp. 360–364.
- [15] Z. JANKO AND H. KHARAGHANI, *A block negacyclic Bush-type Hadamard matrix and two strongly regular graphs*, J. Combin. Theory Ser. A, 98 (2002), pp. 118–126.
- [16] L. K. JØRGENSEN, *Nonsymmetric 3-Class Association Schemes*, Tech. report R-2005-13, Department of Mathematical Sciences, Aalborg University, Aalborg, Denmark, 2005. Available online at <http://www.math.aau.dk/research/reports/R-2005-13.pdf>
- [17] J. D. KEY, *Personal communication*, 2005.
- [18] H. KIMURA, *New Hadamard matrix of order 24*, Graphs Combin., 5 (1989), pp. 235–242.
- [19] H. KIMURA, *Classification of Hadamard matrices of order 28*, Discrete Math., 133 (1994), pp. 171–180.
- [20] H. KIMURA, *Classification of Hadamard matrices of order 28 with Hall sets*, Discrete Math., 128 (1994), pp. 257–268.
- [21] H. KIMURA AND H. OHMORI, *Construction of Hadamard matrices of order 28*, Graphs Combin., 2 (1986), pp. 247–257.
- [22] C. LAM, S. LAM, AND V. D. TONCHEV, *Bounds on the number of affine, symmetric, and Hadamard designs and matrices*, J. Combin. Theory Ser. A, 92 (2000), pp. 186–196.
- [23] C. LAM, S. LAM, AND V. D. TONCHEV, *Bounds on the number of Hadamard designs of even order*, J. Combin. Des., 9 (2001), pp. 363–378.
- [24] E. S. LANDER, *Symmetric Designs: An Algebraic Approach*, London Math. Soc. Lecture Note Ser. 74, Cambridge University Press, Cambridge, UK, 1983.
- [25] C. LIN, W. D. WALLIS, AND Z. LIE, *Equivalence classes of Hadamard matrices of order 32*, in Proceedings of the Twenty-fourth Southeastern International Conference on Combinatorics, Graph Theory, and Computing (Boca Raton, FL, 1993), Congr. Numer., 95 (1993), pp. 179–182.
- [26] B. D. MCKAY, *Hadamard equivalence via graph isomorphism*, Discrete Math., 27 (1979), pp. 213–214.
- [27] B. D. MCKAY, *Practical graph isomorphism*, in Proceedings of the Tenth Manitoba Conference on Numerical Mathematics and Computing, Vol. I (Winnipeg, Man., 1980), Congr. Numer., 30 (1981), pp. 45–87.

- [28] B. D. MCKAY, *Nauty User's Guide (Version 2.2)*, Computer Science Department, Australian National University, Canberra, 2004.
- [29] B. D. MCKAY AND E. SPENCE, *Classification of regular two-graphs on 36 and 38 vertices*, Australas. J. Combin., 24 (2001), pp. 293–300.
- [30] M. NEWMAN, *Invariant factors of combinatorial matrices*, Israel J. Math., 10 (1971), pp. 126–130.
- [31] W. P. ORRICK AND B. SOLOMON, *Large-determinant sign matrices of order $4k + 1$* , Discrete Math., 307 (2007), pp. 226–236.
- [32] K. T. PHELPS, J. RIFÀ, AND M. VILLANUEVA, *Rank and kernel of binary Hadamard codes*, IEEE Trans. Inform. Theory, 51 (2005), pp. 3931–3937.
- [33] J. SEBERRY, *Library of Hadamard Matrices*, available online at <http://www.uow.edu.au/~jennie/hadamard.html> (2001).
- [34] N. J. A. SLOANE, *A Library of Hadamard Matrices*, available online at <http://www.research.att.com/~njas/hadamard/> (1999).
- [35] E. SPENCE, *Regular two-graphs on 36 vertices*, Linear Algebra Appl., 226/228 (1995), pp. 459–497.
- [36] E. SPENCE, *Ted Spence's Home Page*, <http://www.maths.gla.ac.uk/~es/> (2005).
- [37] V. D. TONCHEV, *Hadamard matrices of order 36 with automorphisms of order 17*, Nagoya Math. J., 104 (1986), pp. 163–174.
- [38] W. D. WALLIS, *Integral equivalence of Hadamard matrices*, Israel J. Math., 10 (1971), pp. 349–358.
- [39] W. D. WALLIS AND J. WALLIS, *Equivalence of Hadamard matrices*, Israel J. Math., 7 (1969), pp. 122–128.
- [40] A. L. WHITEMAN, *Hadamard matrices of order $4(2p+1)$* , J. Number Theory, 8 (1976), pp. 1–11.

ALGORITHMS FOR SQUARE-3PC($\cdot, \cdot$)-FREE BERGE GRAPHS*

FRÉDÉRIC MAFFRAY[†], NICOLAS TROTIGNON[‡], AND KRISTINA VUŠKOVIĆ[§]

Abstract. We consider the class of graphs containing no odd hole, no odd antihole, and no configuration consisting of three paths between two nodes such that any two of the paths induce a hole, and at least two of the paths are of length 2. This class generalizes claw-free Berge graphs and square-free Berge graphs. We give a combinatorial algorithm of complexity $O(n^7)$ to find a clique of maximum weight in such a graph. We also consider several subgraph-detection problems related to this class.

Key words. recognition algorithm, maximum weight clique algorithm, combinatorial algorithms, perfect graphs, star decomposition

AMS subject classifications. 68R10, 68Q25, 05C85, 05C17, 90C27

DOI. 10.1137/050628520

1. Introduction. A graph G is *perfect* if every induced subgraph G' of G satisfies $\chi(G') = \omega(G')$, where χ denotes the chromatic number and ω denotes the size of a maximum clique. We say that a graph G *contains* a graph H if H is isomorphic to an induced subgraph of G . A graph G is *H-free* if it does not contain H . A *hole* is a chordless cycle of length at least 4. A *square* is a hole of length 4. A graph is said to be *Berge* if it does not contain an odd hole or the complement of an odd hole.

Berge conjectured in 1960 that a graph is Berge if and only if it is perfect. This was proved by Chudnovsky et al. in 2002 [8]. Later, Chudnovsky et al. [7] gave a polynomial time algorithm that recognizes Berge graphs. In the 1980s, Grötschel, Lovász, and Schrijver [16], [17] gave a polynomial time algorithm that for any perfect graph computes an optimal coloring and a clique of maximum size. This algorithm uses the ellipsoid method and a polynomial time separation algorithm for a certain class of positive semidefinite matrices related to Lovász's upper bound on the Shannon capacity of a graph [20]. The question remains whether these optimization problems can be solved by purely combinatorial polynomial time algorithms, and thus avoiding the numerical instability of the ellipsoid method. The aim of this paper is to give such an algorithm for finding a clique of maximum weight in a subclass of perfect graphs that generalizes claw-free perfect graphs and square-free perfect graphs.

3-path-configurations ($\cdot, \cdot$)'s: A $3PC(x, y)$ is a graph induced by three chordless paths that have the same endnodes x and y and such that the union of any two of them induces a hole. We say that a graph G contains a $3PC(\cdot, \cdot)$ if it contains a $3PC(x, y)$ for some $x, y \in V(G)$. It is easy to see that in a $3PC(x, y)$, each of the three paths must have length at least 2. In the literature, $3PC(\cdot, \cdot)$'s are also known as *thetas* in [9]. A *square-3PC($\cdot, \cdot$)* is a $3PC(\cdot, \cdot)$ that has at least two paths of length 2.

*Received by the editors April 5, 2005; accepted for publication (in revised form) July 5, 2007; published electronically February 8, 2008.

<http://www.siam.org/journals/sidma/22-1/62852.html>

[†]C.N.R.S., Laboratoire G-SCOP, 46 Avenue Félix Viallet, 38031 Grenoble Cedex, France (frederic.maffray@g-scop.inpg.fr).

[‡]Université Paris I, Centre d'Economie de la Sorbonne, 106–112 Boulevard de l'Hôpital, 75647 Paris Cedex 13, France (nicolas.trotignon@univ-paris1.fr). This author was partially supported by the ADONET network, a Marie Curie training network of the European Community.

[§]School of Computing, University of Leeds, Leeds LS2 9JT, UK (vuskovi@comp.leeds.ac.uk). This author was partially supported by EPSRC grant EP/C518225/1.

In this paper we give a combinatorial algorithm, with time complexity $O(n^7)$, that computes a maximum weight clique in every square-3PC($\cdot, \cdot$)-free Berge graph. We will show that every square-3PC($\cdot, \cdot$)-free Berge graph has a node whose neighborhood has no long hole (where a *long hole* is a hole of length greater than 4). This yields a linear-size decomposition tree into square-3PC($\cdot, \cdot$)-free Berge graphs that have no long hole, and then these graphs are further decomposed into cobipartite graphs, resulting in a decomposition tree of total size $O(n^4)$.

Recall that there is an $O(n^9)$ recognition algorithm for the class of Berge graphs [7]. Detecting square-3PC($\cdot, \cdot$)-s in a graph G can be done easily: it suffices to check, for every square a_1, a_2, a_3, a_4, a_1 , whether a_1 and a_3 are in the same connected component of $G \setminus ((N(a_2) \cup N(a_4)) \setminus \{a_1, a_3\})$. This takes time $O(n^6)$. In section 4, we deal with the complexity of several subgraph-detection problems related to this class.

A *claw* is a graph on nodes u, a, b, c with three edges ua, ub, uc . It is easy to see that every 3PC($\cdot, \cdot$) contains a claw. So 3PC($\cdot, \cdot$)-free graphs generalize claw-free graphs. 3PC($\cdot, \cdot$)-free Berge graphs were first studied by Aossey and Vušković [2], [3] in the context of proving the strong perfect graph conjecture for this class. The conjecture was proved by decomposing 3PC($\cdot, \cdot$)-free Berge graphs into claw-free graphs using star cutsets, homogeneous pairs, and 6-joins (a new edge cutset introduced in [3]).

Clearly, square-3PC($\cdot, \cdot$) graphs generalize square-free graphs. In [12] square-free Berge graphs are decomposed by 2-joins and star cutsets into bipartite graphs and line graphs of bipartite graphs (hence proving the strong perfect graph conjecture for this class).

Square-3PC($\cdot, \cdot$)-free Berge graphs contain both claw-free Berge graphs and square-free Berge graphs, and it is likely that one might be able to obtain a similar decomposition theorem that uses star cutsets and some of the other mentioned cutsets. Of course all Berge graphs have been decomposed in [8] (thus proving the strong perfect graph conjecture) into basic classes by skew cutsets, 2-joins, and their complements; see also [6].

Our initial idea was to try to use the above mentioned types of decomposition theorems to develop an algorithm for finding a maximum weight clique in a square-3PC($\cdot, \cdot$)-free Berge graph. Interestingly, we did end up developing a *decomposition-based algorithm* for finding a maximum weight clique, but it does not use any of the types of decomposition theorems mentioned above.

Finding a maximum weight clique in a claw-free Berge graph is not difficult. Indeed, in such a graph G the neighborhood of every node induces a cobipartite graph (first observed in [18]; see also [19]), so the problem reduces to n instances of the maximum weight stable set problem in a bipartite graph, which can be reduced to a maximum flow problem and can be done in time $O(n^3)$; see [19].

For a graph G let k denote the number of maximal cliques in G , n the number of nodes in G , and m the number of edges of G . Farber [14], and independently Alekseev [1], showed that there are $O(n^2)$ maximal cliques in any square-free graph. Tsukiyama et al. [27] gave an $O(nmk)$ algorithm for generating all maximal cliques of a graph, and Chiba and Nishizeki [5] improved this complexity to $O(\sqrt{m+n}mk)$. So one can generate all the maximal cliques of a square-free graph in time $O(\sqrt{m+n}mn^2)$.

For square-free Berge graphs, one can obtain a slightly better algorithm by using the following characterization obtained by Parfenoff, Roussel, and Rusu [24]: *Every square-free Berge graph has a node whose neighborhood is triangulated.*

We conclude this section by defining two more types of 3PC's and wheels.

3PC(Δ, Δ)'s. Let $x_1, x_2, x_3, y_1, y_2, y_3$ be six distinct nodes of G such that $\{x_1, x_2, x_3\}$ and $\{y_1, y_2, y_3\}$ both induce triangles. A $3PC(x_1x_2x_3, y_1y_2y_3)$ is a graph induced by three chordless paths $P_1 = x_1 \cdots y_1$, $P_2 = x_2 \cdots y_2$, and $P_3 = x_3 \cdots y_3$ such that any two of them induce a hole. We say that a graph G contains a $3PC(\Delta, \Delta)$ if it contains a $3PC(x_1x_2x_3, y_1y_2y_3)$ for some $x_1, x_2, x_3, y_1, y_2, y_3 \in V(G)$. Such graphs are also known as prisms in [8] and stretchers in [13].

3PC($\Delta, \cdot$)'s. Let x_1, x_2, x_3, y be four distinct nodes of G such that $\{x_1, x_2, x_3\}$ induces a triangle. We call $3PC(x_1x_2x_3, y)$ any graph induced by three chordless paths $P_1 = x_1 \cdots y$, $P_2 = x_2 \cdots y$, and $P_3 = x_3 \cdots y$ such that the union of any two of them induces a hole. Note that at least two of the three paths must have length at least 2. We say that a graph G contains a $3PC(\Delta, \cdot)$ if it contains a $3PC(x_1x_2x_3, y)$ for some $x_1, x_2, x_3, y \in V(G)$. Such graphs are called pyramids in [8].

Wheels. A *wheel* (H, x) is a graph induced by a hole H and a node $x \notin V(H)$ that has at least three neighbors in H . Node x is the *center* of the wheel. A subpath of H of length at least 1, whose endnodes are adjacent to x and where no intermediate node is adjacent to x , is called a *sector* of (H, x) . A *short sector* is a sector of length 1, and a *long sector* is a sector of length greater than 1. A wheel is *odd* if it contains an odd number of short sectors.

It is easy to see that every odd wheel and every $3PC(\Delta, \cdot)$ contains an odd hole, so Berge graphs cannot contain these two structures. These facts will be used repeatedly in the proofs.

2. Finding a maximum weight clique in a square-3PC($\cdot, \cdot$)-free Berge graph. We assume that we are given a graph G with a weight $f(x)$ associated with every node x . The problem is to find a clique of G of maximum weight, where the weight of a subset of nodes is the sum of the weights of its elements. The maximum weight of a clique is denoted by $\omega_f(G)$. The next theorem will help us solve this problem.

For $x \in V(G)$, $N(x)$ denotes the set of nodes of G that are adjacent to x , and $N[x] = N(x) \cup \{x\}$. For $A \subseteq V(G)$, $G[A]$ denotes the subgraph of G induced by the node set A . $G \setminus A$ denotes the subgraph of G obtained by removing the node set A , i.e., $G \setminus A = G[V(G) \setminus A]$.

THEOREM 2.1. *Let G be a square-3PC($\cdot, \cdot$)-free Berge graph. Let x be a node of G such that $N(x)$ contains a long hole H , and let C be any connected component of $G \setminus N[x]$. Then some node of H has no neighbor in C .*

The proof of this theorem is long and technical, and we leave it for section 3. Here we give a corollary of Theorem 2.1 and show how to use it in an algorithm for the maximum clique problem.

Let $\mathcal{F}$ be a class of graphs. We say that a graph G is $\mathcal{F}$ -free if G does not contain any of the graphs from $\mathcal{F}$.

A class $\mathcal{F}$ of graphs satisfies *property (*)* w.r.t. a graph G if the following holds: For every node x of G such that $G \setminus N[x] \neq \emptyset$, and for every connected component C of $G \setminus N[x]$, if $F \in \mathcal{F}$ is contained in $N(x)$, then there exists a node of F that has no neighbor in C .

In a graph G , for any node x , let $C_1, \dots, C_k$ be the components of $G \setminus N[x]$, with $|C_1| \geq |C_2| \geq \dots \geq |C_k|$, and let the numerical vector $(|C_1|, \dots, |C_k|)$ be associated with x . The nodes of G can thus be ordered according to the lexicographical ordering of the numerical vectors associated with them. Say that a node x is *lex-maximal* if the associated numerical vector is lexicographically maximal over all nodes of G .

THEOREM 2.2. *Let $\mathcal{F}$ be a class of graphs such that for every $F \in \mathcal{F}$, no node of F is adjacent to all the other nodes of F . If $\mathcal{F}$ satisfies property (*) w.r.t. a graph G and x is a lex-maximal node of G , then $N(x)$ is $\mathcal{F}$ -free.*

Proof. Let $\mathcal{F}$ be a class of graphs such that for every $F \in \mathcal{F}$, no node of F is adjacent to all the other nodes of F . Assume that $\mathcal{F}$ satisfies property (*) w.r.t. G .

Let x be a lex-maximal node of G and suppose that $N(x)$ is not $\mathcal{F}$ -free. Then G is not a clique, and hence, since x is lex-maximal, $G \setminus N[x] \neq \emptyset$.

Let $C_1, \dots, C_k$ be the connected components of $G \setminus N[x]$, with $|C_1| \geq |C_2| \geq \dots \geq |C_k|$. Let $N = N(x)$, and for $i = 1, \dots, k$, let $N_i = N(x) \cap N(C_i)$.

CLAIM 2.1. *$N_1 \subseteq N_2 \subseteq \dots \subseteq N_k$, and for every $i = 1, \dots, k$, every node of $(N \setminus N_i) \cup (C_{i+1} \cup \dots \cup C_k)$ is adjacent to every node of N_i .*

Proof of Claim 2.1. We argue by induction. First we show that every node of $(N \setminus N_1) \cup (C_2 \cup \dots \cup C_k)$ is adjacent to every node of N_1 . Assume the contrary and let $y \in (N \setminus N_1) \cup (C_2 \cup \dots \cup C_k)$ be such that it is not adjacent to $z \in N_1$. Clearly y has no neighbor in C_1 , but z does. So $G \setminus N[y]$ contains a connected component that contains $C_1 \cup z$, contradicting the choice of x .

Now let $i > 1$ and assume that $N_1 \subseteq \dots \subseteq N_{i-1}$ and that every node of $(N \setminus N_{i-1}) \cup (C_i \cup \dots \cup C_k)$ is adjacent to every node of N_{i-1} . Since every node of C_i is adjacent to every node of N_{i-1} , it follows that $N_{i-1} \subseteq N_i$. Suppose that there exists a node $y \in (N \setminus N_i) \cup (C_{i+1} \cup \dots \cup C_k)$ that is not adjacent to a node $z \in N_i$. Then $z \in N_i \setminus N_{i-1}$, and z has a neighbor in C_i . Also y is adjacent to all nodes in N_{i-1} and no node of $C_1 \cup \dots \cup C_i$. So there exist connected components of $G \setminus N[y]$, $C_1^y, \dots, C_i^y$ such that $C_1 = C_1^y, \dots, C_{i-1} = C_{i-1}^y$ and $C_i \cup z$ is contained in C_i^y . This contradicts the choice of x . This completes the proof of Claim 2.1.

Since $G[N]$ is not $\mathcal{F}$ -free, it contains $F \in \mathcal{F}$. By property (*), a node y of F has no neighbor in C_k . By Claim 2.1, y is adjacent to every node of N_k , and no node of $N \setminus N_k$ has a neighbor in $C = C_1 \cup \dots \cup C_k$. So (since every node of F has a nonneighbor in F) F must contain another node $z \in N \setminus N_k$, nonadjacent to y . But then $C_1, \dots, C_k$ are connected components of $G \setminus N[y]$ and z is contained in $(G \setminus N[y]) \setminus C$, so y contradicts the choice of x . $\square$

THEOREM 2.3. *Let G be a square-3PC($\cdot, \cdot$)-free Berge graph. Let x be a lex-maximal node in G . Then the neighborhood of x in G contains no long hole.*

Proof. Let G be a square-3PC($\cdot, \cdot$)-free Berge graph and let x be a lex-maximal node of G . Let $\mathcal{F}$ be the set of all long holes of G . By Theorem 2.1, $\mathcal{F}$ satisfies property (*) w.r.t. G . So by Theorem 2.2, $N(x)$ is $\mathcal{F}$ -free, i.e., long-hole-free. $\square$

Let $\mathcal{F}$ be the class of square-3PC($\cdot, \cdot$)-free Berge graphs which contain no long hole. Suppose that we have an algorithm A that computes a clique of maximum weight for every graph in $\mathcal{F}$ in time $O(n^t)$. Then we can compute a clique of maximum weight for every square-3PC($\cdot, \cdot$)-free Berge graph G as follows. By Theorem 2.3, G has a node x whose neighborhood contains no long hole. Let G_0 be the subgraph of G induced by $N(x)$. So G_0 is in $\mathcal{F}$. Clearly, since every clique of G either contains x or does not, we have $\omega_f(G) = \max\{f(x) + \omega_f(G_0), \omega_f(G \setminus \{x\})\}$. Thus, in order to compute $\omega_f(G)$, we need only compute $\omega_f(G_0)$ and $\omega_f(G \setminus \{x\})$. The former can be done by algorithm A , and the latter can be done recursively. Note that computing the numerical vector associated with a node takes time $O(n^2)$, and so we can find a lex-maximal node in time $O(n^3)$. So we can find in time $O(n^4)$ an ordering $x_1, \dots, x_n$ of the nodes of G such that, for each $i = 1, \dots, n$, node x_i is lex-maximal in the subgraph induced by $x_i, \dots, x_n$. Thus we can compute $\omega_f(G)$ for every square-3PC($\cdot, \cdot$)-free Berge graph G in time $O(n^4 + n^{t+1})$. Now we describe such an algorithm A . For this purpose we

will use the following definition.

Full star decomposition. For $x \in V(G)$ such that x is not adjacent to every node of $G \setminus \{x\}$, let $C_1, \dots, C_m$ be the connected components of $G \setminus N[x]$. Note that $N[x]$ need not be a cutset, i.e., possibly $m = 1$. The blocks of the *full star decomposition* at x are the graphs $G_0, G_1, \dots, G_m$ defined as follows: G_0 is the subgraph of G induced by $N(x)$ and, for $i = 1, \dots, m$, G_i is the subgraph of G induced by $C_i \cup N_i$, where N_i is the set of nodes of $N(x)$ which have a neighbor in C_i .

Remark 2.4. From the construction of the blocks of full star decomposition of G it follows easily that $\omega_f(G) = \max\{f(x) + \omega_f(G_0), \omega_f(G_1), \dots, \omega_f(G_m)\}$.

This remark shows that the problem of finding a maximum weight clique in G can be reduced to finding a maximum weight clique in some subgraphs of G . Our algorithm for finding a maximum weight clique in a square-3PC($\cdot, \cdot$)-free Berge graph G with no long hole consists of the following two stages:

- Stage 1: A decomposition tree T is constructed, where each leaf node is cobipartite (i.e., complement of a bipartite graph), and for each nonleaf node F , the children of F in T represent the blocks of a full star decomposition of F .
- Stage 2: A maximum weight clique is computed for each of the leaf nodes, and then the algorithm backtracks along T to find a maximum weight clique for G using Remark 2.4.

Finding a maximum weight clique in a cobipartite graph is equivalent to finding a maximum weight stable set in a bipartite graph, and it is well known that this problem can be reduced to a maximum flow in a directed network associated with the bipartite graph; see [19]. From that it is easy to deduce an algorithm (henceforth the “cobipartite maximum weight clique algorithm”) that computes a maximum weight clique in a cobipartite graph G in time $O(n^3)$. So if the size of the decomposition tree is polynomial, then Stage 2 of the algorithm can be performed in polynomial time. The key difficulty is to construct T in polynomial time. To perform Stage 1, we decompose using a full star centered at a node contained in an independent set of size 3. Note that when a Berge graph does not contain any independent set of size 3, then it is cobipartite. The following lemma implies that the decomposition tree so constructed has polynomial size.

LEMMA 2.5. *Let G be a square-3PC($\cdot, \cdot$)-free Berge graph that contains no long hole, and let $G_0, G_1, \dots, G_m$ be the blocks of a full star decomposition at some node x of G . Then the following hold:*

- (1) *No independent set of G of size 2 is contained in both G_i and G_j for any $1 \leq i \neq j \leq m$.*
- (2) *No independent set of G of size 3 is contained in both G_0 and G_i for any $1 \leq i \leq m$.*

Proof. Note that in a long-hole-free graph every 3PC($\cdot, \cdot$) is a square-3PC($\cdot, \cdot$). To prove (1), suppose that $\{a, b\}$ is an independent set of G contained in both G_i and G_j with $1 \leq i \neq j \leq m$. Then $\{a, b\} \subseteq N(x)$, and both a and b have neighbors in both C_i and C_j . But then there exists a chordless a, b -path P_i (resp., P_j) whose intermediate nodes are in C_i (resp., C_j), and hence $P_i \cup P_j \cup x$ induces a 3PC(a, b), which is a contradiction.

To prove (2), suppose that there exists an independent set $\{a, b, c\}$ which is contained in both G_0 and G_1 . Then $\{a, b, c\} \subseteq N(x)$ and every node of $\{a, b, c\}$ has a neighbor in C_1 . Let u, v, t be neighbors of a, b, c , respectively, in C_1 . Then there is a path P in C_1 from u to v . W.l.o.g., u, v, P are chosen so that P is minimal. Then

$P \cup \{x, a, b\}$ induces a hole, and since this hole cannot be long, $u = v$. If $t = u$, then $\{x, a, b, c, u\}$ induces a $3PC(x, u)$. So $t \neq u$. Let Q be a shortest path from t to u in C_1 . W.l.o.g., no node of $Q \setminus \{t\}$ is adjacent to c . Since the graph induced by $Q \cup \{x, a, c\}$ cannot contain a long hole, a is adjacent to t . By symmetry, b is also adjacent to t , and hence $\{x, a, b, c, t\}$ induces a $3PC(x, t)$. $\square$

ALGORITHM 2.6.

INPUT: A square- $3PC(\cdot, \cdot)$ -free Berge graph G with no long hole, and a weight function f on $V(G)$.

OUTPUT: A maximum weight clique of G .

METHOD: Step 1. Let $\mathcal{L} = \{G\}$, $\mathcal{L}' = \emptyset$, and let T be a tree that consist of a single node G .

Step 2. If $\mathcal{L} = \emptyset$, then go to Step 3. Otherwise, remove a graph F from $\mathcal{L}$. If F does not contain an independent set of size 3, then place F in $\mathcal{L}'$ and return to Step 2. Otherwise, let $\{x, y, z\}$ be an independent set of F . Decompose F using the full star decomposition centered at x . Place the blocks of the decomposition in $\mathcal{L}$, add the blocks of the decomposition to T as children of F , and return to Step 2.

Step 3. For every $F \in \mathcal{L}'$, find a maximum weight clique of F using the cobipartite maximum weight clique algorithm. Note that the leaves of T are precisely the graphs in $\mathcal{L}'$. Using Remark 2.4, backtrack along T to a maximum weight clique of G .

COMPLEXITY: $O(n^6)$.

Proof. By the definition of Step 2, the graphs in $\mathcal{L}'$ (that represent the leaves of T) do not contain any independent set of size 3. Since G is Berge, these graphs are cobipartite. So Step 3 correctly finds a maximum weight clique.

Now we determine the complexity of the algorithm. Consider the tree T obtained at the end of Step 2. We show that the number of nonleaf nodes in T is $O(n^3)$. Let F be a nonleaf node of T . Let $\{x, y, z\}$ be an independent set of F from Step 2 such that F is decomposed by the full star centered at x . We view $\{x, y, z\}$ as the label of F . By Lemma 2.5 and the fact that x is not contained in any of the blocks of decomposition of F , no two nonleaf nodes of T have the same label. So the number of nonleaf nodes of T is at most n^3 .

We now show that the number of leaf nodes of T is also $O(n^3)$. For a node F of T , define a measure $\tau(F)$ as follows: If F is a nonleaf node of T , then let $\tau(F)$ be the number of independent sets of size 3 in F ; if F is a leaf node and F has at least three siblings in the decomposition tree, then $\tau(F) = 1$; otherwise, $\tau(F) = 0$. Let F be a nonleaf node of T . Suppose that F is decomposed by a full star centered at x . Let $C_1, \dots, C_m$ be the connected components of $F \setminus N[x]$, and let $F_0, \dots, F_m$ be the blocks of decomposition. We claim that the following inequality holds:

$$(1) \quad \tau(F) \geq \tau(F_0) + \tau(F_1) + \dots + \tau(F_m).$$

Indeed, by Lemma 2.5, no independent set of F of size 3 is contained in more than one block of the decomposition. So if $m < 3$, then (1) clearly holds. Suppose that $m \geq 3$. To show (1), it is enough to show that the number of independent sets of F of size 3 that are not contained in any of the blocks is $\geq m + 1$. For $i = 1, \dots, m$, let c_i be a node of C_i . The number of sets that contain x and two nodes from $\{c_1, \dots, c_m\}$ is $\binom{m}{2}$. The number of sets that contain three nodes from $\{c_1, \dots, c_m\}$ is $\binom{m}{3}$. Note that all these sets of size 3 are independent sets of F that are not contained in any of the blocks of the decomposition. So the number of independent sets of size 3 of

F that are not contained in any of the blocks is at least $\binom{m}{2} + \binom{m}{3}$. Since $m \geq 3$, $\binom{m}{2} + \binom{m}{3} \geq m + 1$. Therefore, (1) holds.

By repeated applications of (1) we get the inequality $\tau(G) \geq \sum \{\tau(F) \mid F \text{ leaf of } T\}$. So the number of leaves F of T such that $\tau(F) = 1$ is $O(n^3)$. By the definition of τ , the number of leaves F of T such that $\tau(F) = 0$ is at most 3 times the number of internal nodes of T . Hence, the number of leaves F of T such that $\tau(F) = 0$ is $O(n^3)$. Therefore the number of leaves of T is $O(n^3)$.

So the size of T is $O(n^3)$. When the algorithm examines a nonleaf node of T , it looks for an independent set of size 3. Since F is Berge, it suffices to check whether its complement $\bar{F}$ is bipartite, which can be done in time $O(n^2)$ with standard breadth-first search (and this method will produce an independent set of size 3 whenever $\bar{F}$ is not bipartite). So the complexity of constructing T is $O(n^5)$, and the total complexity of Step 3 is $O(n^6)$. Therefore, the overall complexity of the algorithm is $O(n^6)$. $\square$

This implies that we can compute $\omega_f(G)$ for every square-3PC($\cdot, \cdot$)-free Berge graph G in time $\mathcal{O}(n^7)$. In fact this algorithm can be turned into a robust algorithm (in Spinrad's sense [25]). We would not need to know that the input graph G to the algorithm is a square-3PC($\cdot, \cdot$)-free Berge graph. The algorithm would then either correctly compute $\omega_f(G)$ (in all cases when G is a square-3PC($\cdot, \cdot$)-free Berge graph, and in some cases when it is not) or it would identify G as not being a square-3PC($\cdot, \cdot$)-free Berge graph. As in the algorithm we have just given, we would start by looking for a lex-maximal vertex x . We would then check whether $N(x)$ is long-hole-free. If it is not, we would terminate the algorithm, outputting that G is not a square-3PC($\cdot, \cdot$)-free Berge graph (by Theorem 2.3). In the second part of the algorithm, where we decompose graphs with no long hole, at each decomposition we would verify that (1) and (2) of Lemma 2.5 hold (which can easily be done). If one of those conditions fails, we would again terminate the algorithm, outputting that G is not a square-3PC($\cdot, \cdot$)-free Berge graph. Otherwise, we would end up with an $\mathcal{O}(n^3)$ decomposition tree as before. We would now just have to check whether the leaves are cobipartite. If they are not, we would again terminate, outputting that G is not a square-3PC($\cdot, \cdot$)-free Berge graph.

3. Proof of Theorem 2.1. To prove Theorem 2.1 we prove the following stronger result. We *sign* a graph by assigning 0, 1 weights to its edges in such a way that, for every triangle in the graph, the sum of the weights of its edges is odd. A graph G is *even-signable* if there is a signing of its edges so that for every hole in G , the sum of the weights of its edges is even. Clearly, every odd-hole-free graph is even-signable (assign weight 1 to all its edges). The following theorem is an easy consequence of a theorem of Truemper [26].

THEOREM 3.1 (Conforti et al. [10], [11]). *A graph is even-signable if and only if it does not contain an odd wheel or a 3PC($\Delta, \cdot$).*

The fact that even-signable graphs do not contain odd wheels and 3PC($\Delta, \cdot$)'s will be used throughout the proof of the next theorem.

Remark. Even though Theorem 3.2 below implies that every square-3PC($\cdot, \cdot$) even-signable graph has a node whose neighborhood is long-hole-free, finding a largest clique in a claw-free odd-hole-free graph (and hence in a square-3PC($\cdot, \cdot$)-free even-signable graph) is NP-hard. Indeed, it is proved in [22] that it is NP-hard to find a largest independent set in a graph with no cycle of length 3, 4, or 5, and so it is NP-hard to find a largest clique in a graph with no stable set of size 3 and no hole of length 5.

THEOREM 3.2. *Let G be a square-3PC($\cdot, \cdot$)-free even-signable graph. Let x be a node of G such that $N(x)$ contains a long hole H , and let C be any connected component of $G \setminus N[x]$. Then some node of H has no neighbor in C .*

Proof. Assume that every node of H has a neighbor in C . We will show that this leads to a contradiction. Let n be the length of H , and let $H = h_1 h_2 \cdots h_n h_1$. Note that since (H, x) cannot be an odd wheel, H must be of even length, so $n \geq 6$. For any node u , we denote by $N_H(u)$ the set $N(u) \cap V(H)$.

CLAIM 3.1. *Every node u of C that has a neighbor in H is one of the following five types:*

- *Type 1, $i = 1, 2, 3$: u has exactly i neighbors in H , and they are consecutive along H .*
- *Type 4: u has exactly four neighbors $h_i, h_{i+1}, h_j, h_{j+1}$ in H (that appear in this order when traversing H clockwise), where i, j have different parities.*
- *Type 5: u has exactly two neighbors h_i and h_j in H , i and j are of the same parity, and the two subpaths of H from h_i to h_j are of length greater than 2.*

Proof. Let $s = |N_H(u)|$. Note that if u has three pairwise nonadjacent neighbors a, b, c in H , then $\{x, u, a, b, c\}$ induces a square-3PC(x, u), a contradiction. Therefore $N_H(u)$ is covered by at most two cliques of H and $s \leq 4$. If $s = 1$, then u is of Type 1. Suppose $s = 2$ and u is not of Type 2. Let h_i and h_j be the two neighbors of u in H . Let H' be a subpath of H whose 1 endnode is h_i and the other is h_j . If i and j are not of the same parity, then H' is of odd length greater than 1 and $H' \cup \{u, x\}$ induces an odd wheel with center x . If H' is of length 2, then $H \cup u$ induces a square-3PC(h_i, h_j). So u must be of Type 5. If $s = 3$ and u is not of Type 3, then (H, u) is an odd wheel. Suppose $s = 4$ and u is not of Type 4. Then u has four neighbors $h_i, h_{i+1}, h_j, h_{j+1}$ in H , where i, j have the same parity. Let H' be a subpath of H from h_{i+1} to h_j . H' cannot be of length 1, since then (H, u) is an odd wheel. So H' is of odd length greater than 1, and hence $H' \cup \{u, x\}$ induces an odd wheel with center x . This proves the claim. $\square$

CLAIM 3.2. *Let $P = p_1 \cdots p_k$ be a chordless path in C such that $k \geq 2$, nodes p_1 and p_k both have neighbors in H , and no node of $P \setminus \{p_1, p_k\}$ has a neighbor in H . If $N_H(p_1) \not\subseteq N_H(p_k)$ and $N_H(p_k) \not\subseteq N_H(p_1)$, then one of the following holds:*

- (i) $N_H(p_1) = \{a\}$, $N_H(p_k) = \{b\}$, and either ab is an edge or the two subpaths of $H \setminus \{a, b\}$ are both of even length.
- (ii) $N_H(p_1) = \{a, b\}$, $N_H(p_k) = \{c, d\}$, ab and cd are edges, and the subpaths of $H \setminus \{a, b, c, d\}$ are of even length.
- (iii) For some $i \in \{1, \dots, n\}$, $N_H(p_1) = \{h_i, h_{i+1}, h_{i+2}\}$, indices taken modulo n , $N_H(p_k) \cap \{h_i, h_{i+1}, h_{i+2}\} = h_i$, and p_k is of Type 3 or 5.
- (iv) Nodes p_1 and p_k are both of Type 5 and have a common neighbor in H .

Proof. Consider the following property S_3 : There are three pairwise nonadjacent nodes h_r, h_s, h_t of H such that p_1 is adjacent to h_r and h_s (and thus not to h_t) and p_k is adjacent to h_t and not to h_r, h_s . Note that S_3 does not hold, for otherwise $P \cup \{h_r, h_s, h_t, x\}$ induces a square-3PC(p_1, x). By Claim 3.1, p_1 and p_k are of Type 1, 2, 3, 4, or 5. This leads, up to symmetry, to the following case analysis.

First suppose that p_1 is of Type 4, with $N_H(p_1) = \{h_1, h_2, h_t, h_{t+1}\}$ with t even, $4 \leq t \leq n - 2$. If p_k is of Type 1, 2, or 5, then S_3 holds. If p_k is of Type 3, then either S_3 holds or $N_H(p_k) = \{h_1, h_2, h_3\}$, and in this case either $k = 2$ and $\{p_1, p_2, h_1, h_3, h_{t+1}, x\}$ induces a 3PC($h_1 p_1 p_2, x$), or $k > 2$ and $(H \setminus \{h_2\}) \cup \{p_1, p_k\}$ induces an odd wheel with center p_1 . If p_k is of Type 4, then either S_3 holds; or $N_H(p_k) = \{h_1, h_2, h_s, h_{s+1}\}$ with s even and $4 \leq s \leq t - 2 < t \leq n - 2$, and in this

case either $k = 2$ and $\{p_1, p_2, h_1, h_{s+1}, h_{t+1}, x\}$ induces a $3PC(h_1 p_1 p_2, x)$, or $k > 2$ and $\{h_{s+1}, \dots, h_n, h_1, p_1, p_k\}$ induces an odd wheel with center p_1 ; or $N_H(p_k) = \{h_2, h_3, h_{t+1}, h_{t+2}\}$, and in this case either $k = 2$ and $\{x, h_1, h_2, h_3, p_1, p_2\}$ induces an odd wheel with center h_2 , or $k > 2$ and $\{x, h_2, h_{t+1}, p_1, p_k\}$ induces a square- $3PC(h_2, h_{t+1})$.

Now suppose that p_1 is of Type 3, with $N_H(p_1) = \{h_1, h_2, h_3\}$. If p_k is of Type 1, 2 or 3, then either property S_3 holds, or we have outcome (iii), or p_k is adjacent to h_4 and $N_H(p_k) \subseteq \{h_2, h_3, h_4\}$, and in this case $(H \setminus \{h_2, h_3\}) \cup P \cup x$ induces an odd wheel with center x . If p_k is of Type 5, then either S_3 holds, or we have outcome (iii).

Now suppose that p_1 is of Type 5, with $N_H(p_1) = \{h_1, h_t\}$ and t odd, $5 \leq t \leq n-3$. If p_k is of Type 1 or 2, then either S_3 holds, or p_k is adjacent to h_2 and $N_H(p_k) \subseteq \{h_1, h_2\}$, and in this case $P \cup \{x, h_2, \dots, h_t\}$ induces an odd wheel with center x . If p_k is of Type 5, then either S_3 holds, or we have outcome (iv), or $N_H(p_k) = \{h_2, h_s\}$ with $s = t \pm 1$, and in this case $P \cup \{x, h_1, h_2, h_t\}$ induces a $3PC(x h_1 h_2, p_1)$.

Finally, if p_1, p_k are both of Type 1, say $N_H(p_1) = h_i$ and $N_H(p_k) = h_j$ with $i < j$, then either we have outcome (i) or $P \cup \{h_i, \dots, h_j, x\}$ induces an odd wheel with center x . If one of p_1, p_k is of Type 1 and the other is of Type 2, then $H \cup P$ induces a $3PC(\Delta, \cdot)$. If p_1, p_k are both of Type 2, then either we have outcome (ii) or there is a subpath $H' = h_i \dots h_j$ of H such that $N(p_1) \cap H' = \{h_i\}$, $N(p_k) \cap H' = \{h_j\}$, and H' is of odd length, and then $P \cup H' \cup \{x\}$ induces an odd wheel with center x . This proves the claim. $\square$

CLAIM 3.3. *C does not contain a path $P = p_1 \dots p_k$ such that either $k = 1$ and p_1 is of Type 4 or $k \geq 2$ and $H \cup P$ induces a $3PC(\Delta, \Delta)$.*

Proof. Suppose that there is such a path. W.l.o.g., we have either $k = 1$ and $N_H(p_1) = \{h_1, h_2, h_t, h_{t+1}\}$, or $k \geq 2$, $N_H(p_1) = \{h_1, h_2\}$, and $N_H(p_k) = \{h_t, h_{t+1}\}$. Then by Claims 3.1 and 3.2, t is even. In particular, $h_2 h_t$ and $h_{t+1} h_1$ are not edges. Since every node of H has a neighbor in C , there exists a chordless path $Q = q_1 \dots q_l$ in C such that q_1 is adjacent to a node of $H \setminus \{h_1, h_2, h_t, h_{t+1}\}$ and q_l is adjacent to a node of P . We may assume that such paths P and Q are chosen so that $|V(P) \cup V(Q)|$ is minimized. Thus no node of $Q \setminus \{q_l\}$ has a neighbor in P , and the only nodes of H that can have a neighbor in $Q \setminus \{q_1\}$ are h_1, h_2, h_t, h_{t+1} .

First suppose that $k = 1$. W.l.o.g., q_1 has a neighbor h_i in $h_3 \dots h_{t-1}$. Some node q_j of Q must be adjacent to one of h_1, h_{t+1} , else $Q \cup \{p_1, h_1, h_{t+1}, h_i, x\}$ induces a square- $3PC(x, p_1)$. Let j be the smallest such index; say q_j is adjacent to h_1 . Then q_j is not adjacent to h_{t+1} , else $\{q_1, \dots, q_j, x, h_1, h_i, h_{t+1}\}$ induces a square- $3PC(x, q_j)$. By the choice of j , node q_1 is not adjacent to h_{t+1} . Then $j < l$, else $\{q_1, \dots, q_j, p_1, h_1, h_i, h_{t+1}, x\}$ induces a $3PC(h_1 p_1 q_j, x)$. Then q_1 has a neighbor in $h_{t+2} \dots h_n$, else $\{p_1, q_1, \dots, q_j\} \cup H \setminus \{h_2, \dots, h_{i-1}\}$ contains a $3PC(h_t h_{t+1} p_1, h_1)$. It follows that q_1 is either of Type 4, or of Type 5 not adjacent to any of h_1, h_2, h_t, h_{t+1} . By Claim 3.2 applied to $Q \cup \{p_1\}$, some node q_s of $Q \setminus \{q_1\}$ has a neighbor in H , and we choose the smallest such s . By Claim 3.2 applied to $q_1 \dots q_s$, we have $N_H(q_s) \subseteq N_H(q_1)$, which is possible only if q_1 is of Type 4 with a neighbor in $\{h_1, h_2, h_t, h_{t+1}\}$; so, up to symmetry, $N_H(q_1) = \{h_n, h_1, h_i, h_{i+1}\}$, $2 < i < t$. Let R be the shortest path from q_1 to h_{t+1} in the subgraph induced by $Q \cup \{p_1, h_{t+1}\}$. If $n > t + 2$, then $R \cup \{h_{t+2}, \dots, h_n, x\}$ induces an odd wheel with center x , while if $n = t + 2$, then $R \cup \{h_n, x, h_i\}$ induces a $3PC(x h_{t+1} h_{t+2}, q_1)$, which is a contradiction. Therefore we must have $k \geq 2$, and p_1 and p_k are of Type 2.

Now we show that either p_1 or p_k is the only neighbor of q_l in P . Suppose the contrary. Let a, b be, respectively, the smallest and largest integers such that q_l is

adjacent to nodes p_a and p_b of P . So either $a \neq b$ or $1 < a = b < k$. Let j be the largest integer such that q_j has a neighbor in H . We can apply Claim 3.2 to paths $P_a = p_1 \cdots p_a q_l \cdots q_j$ and $P_b = p_k \cdots p_b q_l \cdots q_j$. Since $N_H(q_j)$ cannot be a subset of both $N_H(p_1)$ and $N_H(p_k)$, this implies that either $N_H(q_j) = \{h_1, h_2, h_t, h_{t+1}\}$ or $N_H(q_j) = \{h_i, h_{i+1}\}$ for some i . In the first case, $\{q_1, \dots, q_j\}$ contradicts the minimality of $P \cup Q$. So we have the second case. If $i = 1$, then $\{p_b, \dots, p_k\} \cup Q$ contradicts the minimality of $P \cup Q$. If $i = t$, we have a similar contradiction. If $i \notin \{1, t\}$, then the parity condition of Claim 3.2(ii) is violated by one of P_a, P_b . Therefore, and up to symmetry, we may assume that p_k is the only neighbor of q_l in P .

Put $p_k = q_{l+1}$. Let r be the largest index such that a node q_r of Q has a neighbor in $H \setminus \{h_t, h_{t+1}\}$. Along the path $q_{r+1} \cdots q_{l+1}$, let s be the smallest index such that q_s has a neighbor in H . By the choice of q_r , we have $N_H(q_s) \subseteq \{h_t, h_{t+1}\}$, and q_s is of Type 1 or 2. W.l.o.g., q_s is adjacent to h_t . By Claim 3.2 applied to path $q_r \cdots q_s$, we have either case (a) nodes q_r, q_s are both of Type 1, or (b) nodes q_r, q_s are both of Type 2, or (c) $N_H(q_s) \subseteq N_H(q_r)$. More precisely, we have the following.

In case (a), we have $N_H(q_s) = \{h_t\}$ and $N_H(q_r) = \{h_i\}$ for some even $i \neq t$. Suppose $t + 2 \leq i \leq n$. So $r = 1$. If $s = l$, then $P \cup Q \cup \{x, h_2, h_t, h_i\}$ induces a 3PC($p_k q_l h_t, x$). If $s < l$, then $P \cup \{q_1, \dots, q_s\} \cup \{h_t, \dots, h_n, h_1\}$ induces a 3PC($p_k h_t h_{t+1}, h_i$). Thus $2 \leq i \leq t - 2$.

In case (b), we have $N_H(q_s) = \{h_t, h_{t+1}\}$ and $N_H(q_r) = \{h_i, h_{i+1}\}$ for some odd i . If $i = 1$, then $r > 1$, and $\{q_1, q_2, \dots, q_s\}$ contradicts the minimality of $P \cup Q$. Thus we may assume w.l.o.g. that $3 \leq i \leq t - 1$.

In case (c), we have either case (c1) node q_r is of Type 2 or 3 adjacent to h_{t-1} and h_t , or (c2) node q_r is of Type 4 with $N_H(q_r) = \{h_i, h_{i+1}, h_t, h_{t+1}\}$ for some odd i with $3 \leq i \leq t - 3$, or (c3) node q_r is of Type 4 and not adjacent to one of h_t, h_{t+1} , or (c4) node q_r is of Type 5 adjacent to h_t and h_i for some even $i \neq t - 2, t, t + 2$. In case (c3), $\{q_r, \dots, q_l, p_k\}$ contradicts the minimality of $P \cup Q$. In case (c4), suppose $t + 2 \leq i \leq n$. If $r = l$, then $P \cup \{q_r, x, h_2, h_t, h_i\}$ induces a 3PC($p_k q_l h_t, x$), while if $r < l$, then $P \cup \{q_r\} \cup \{h_t, \dots, h_n, h_1\}$ induces a 3PC($p_k h_t h_{t+1}, h_i$). Thus $2 \leq i \leq t - 2$. So we have case (a), (b), (c1), (c2), or (c4), and in either case there is an index i , with $2 \leq i \leq t - 1$, such that q_r is adjacent to h_i and $N_H(q_r) \subseteq \{h_i, \dots, h_{t+1}\}$. Now, if h_{t+1} has no neighbor in $q_r \cdots q_l$, then $P \cup (H \setminus \{h_{i+1}, \dots, h_t\}) \cup \{q_r, \dots, q_l\}$ induces a 3PC($h_1 h_2 p_1, p_k$). If h_{t+1} has a neighbor in $q_r \cdots q_{l-1}$, then $P \cup (H \setminus \{h_{i+1}, \dots, h_t\}) \cup \{q_r, \dots, q_{l-1}\}$ contains a 3PC($h_1 h_2 p_1, h_{t+1}$). So q_l is the unique neighbor of h_{t+1} in $q_r \cdots q_l$. If $i = 2$, then $P \cup \{q_r, \dots, q_l, x, h_2, h_{t+1}\}$ induces a 3PC($h_{t+1} p_k q_l, h_2$). If $i > 2$, then $P \cup \{q_r, \dots, q_l, x, h_1, h_i, h_{t+1}\}$ induces a 3PC($h_{t+1} p_k q_l, x$). This proves the claim. $\square$

For $i = 1, \dots, n$, let H_i be the set of Type 3 nodes of C adjacent to h_i, h_{i+1}, h_{i+2} (indices taken modulo n). Note that H_i induces a clique, for if $a, b \in H_i$ are not adjacent, then $\{a, b, x, h_i, h_{i+2}\}$ induces a square-3PC(h_i, h_{i+2}).

CLAIM 3.4. *Let h_s, h_t be nonadjacent nodes of H and $P = y \cdots z$ be a chordless path in C such that y is the only neighbor of h_s in P and z is the only neighbor of h_t in P . If both h_{s-1}, h_{s+1} have a neighbor in P , then $N_H(y) = \{h_{s-1}, h_s, h_{s+1}\}$.*

Proof. We may assume up to symmetry that $s = 2$ and $t \neq 4$. Let H' be the hole induced by $P \cup \{x, h_2, h_t\}$. Since (H', h_3) cannot be an odd wheel, $(H' \setminus \{x\}) \cup \{h_3\}$ must contain an odd number of triangles. Let R be any long sector of (H', h_1) . If $R \cup \{h_3\}$ contains an odd number of triangles and R contains at least three neighbors of h_3 , then $R \cup \{h_1, h_3\}$ induces an odd wheel with center h_3 . If R contains only two

adjacent neighbors a, b of h_3 , then R cannot contain x and h_t ; and then, if R does not contain h_2 , then $R \cup \{h_1, h_3, x\}$ induces a $3PC(h_3ab, h_1)$, while if R contains h_2 , then $\{a, b\} = \{h_2, y\}$, and $R \cup \{h_1, h_3, x\}$ induces an odd wheel with center h_2 . Thus $R \cup \{h_3\}$ contains an even number of triangles for every long sector R of (H', h_1) . It follows that some edge of $H' \setminus \{x\}$ is a short sector of both (H', h_1) and (H', h_3) , and thus some node of P is adjacent to h_1 and h_3 ; by Claims 3.1 and 3.3, such a node is of Type 3 adjacent to h_1, h_2, h_3 and so it can only be y . This proves the claim. $\square$

CLAIM 3.5. *Let u be a node of C that has a neighbor in $h_4 \cdots h_n$. Let $P = p_1 \cdots p_k$ be a path of C such that $p_k = u$, $p_1 \in H_1$, and no node of $P \setminus \{p_1\}$ belongs to H_1 . Then exactly one node of h_1, h_3 has a neighbor in $P \setminus \{p_1\}$, say h_1 . Node h_1 must in fact have a neighbor in $P \setminus \{p_1, p_2\}$, and so $k \geq 3$. Furthermore, if $Q = q_1 \cdots q_l$ is any other path of C such that $q_l = u$, $q_1 \in H_1$, and no node of $Q \setminus \{q_1\}$ belongs to H_1 , then h_1 has a neighbor in $Q \setminus \{q_1\}$.*

Proof. Let h_i be any neighbor of u in $h_4 \cdots h_n$, and let j be the smallest index such that h_i is adjacent to p_j . Suppose that neither h_1 nor h_3 has a neighbor in $p_2 \cdots p_j$. If $i \notin \{4, n\}$, then $\{p_1, \dots, p_j, h_1, h_3, h_i, x\}$ induces a square-3PC(p_1, x). Otherwise, w.l.o.g. $i = 4$, and hence the same node set induces a $3PC(h_3h_4x, p_1)$. Therefore h_1 or h_3 must have a neighbor in $P \setminus \{p_1\}$.

Suppose that both h_1 and h_3 have a neighbor in $P \setminus \{p_1\}$. If they both have a neighbor in $P \setminus \{p_1, p_2\}$, then there is a shortest subpath P' of $P \setminus \{p_1, p_2\}$ whose one endnode is adjacent to h_1 and the other to h_3 , and hence $P' \cup \{p_1, x, h_1, h_3\}$ induces a square-3PC(h_1, h_3). So we may assume w.l.o.g. that p_2 is the unique neighbor of h_1 in $P \setminus \{p_1\}$. Let p_t be the node of P with lowest index adjacent to h_3 . If $t = 2$, then Claim 3.1 implies $p_2 \in H_1$, a contradiction. So $t > 2$, and hence $\{p_1, \dots, p_t, h_1, h_3, x\}$ induces a $3PC(p_1p_2h_1, h_3)$. Therefore, both h_1 and h_3 cannot have a neighbor in $P \setminus \{p_1\}$.

Assume w.l.o.g. that h_1 has a neighbor in $P \setminus \{p_1\}$. Suppose that p_2 is the unique neighbor of h_1 in $P \setminus \{p_1\}$. If a node h_t , $4 < t < n$, has a neighbor in P , then $P \cup \{h_1, h_3, h_t, x\}$ contains a $3PC(h_1p_1p_2, x)$. So no node of $h_5 \cdots h_{n-1}$ has a neighbor in P . If h_n has no neighbor in $p_1 \cdots p_j$, then $i = 4$, and hence $\{p_1, \dots, p_j, x\} \cup (H \setminus \{h_2, h_3\})$ induces an odd wheel with center x . So h_n has a neighbor in $p_1 \cdots p_j$. Let p_t be such a neighbor with smallest index. If $t < j$ or $t = j$ and p_j is not adjacent to h_4 , then $\{p_1, \dots, p_t, x\} \cup (H \setminus \{h_1, h_2\})$ induces an odd wheel with center x . So $t = j$, and p_j is adjacent to h_4 . Hence $\{p_1, \dots, p_j, h_3, h_4, h_n, x\}$ induces a $3PC(h_3h_4x, p_j)$. Therefore, h_1 has a neighbor in $P \setminus \{p_1, p_2\}$.

Now suppose that h_1 does not have a neighbor in $Q \setminus \{q_1\}$. Then h_3 must have a neighbor in $Q \setminus \{q_1, q_2\}$, and hence $(P \setminus \{p_1, p_2\}) \cup (Q \setminus \{q_1, q_2\}) \cup \{h_1, h_3\}$ contains a chordless path R from h_1 to h_3 . If p_1 does not have a neighbor in $Q \setminus \{q_1, q_2\}$, then $R \cup \{p_1, x\}$ induces a square-3PC(h_1, h_3). So p_1 has a neighbor in $Q \setminus \{q_1, q_2\}$. Let Q' be the shortest path from p_1 to u whose vertices are contained in $(Q \setminus \{q_1, q_2\}) \cup \{p_1\}$. Now apply the same argument to P and Q' . This proves the claim. $\square$

Suppose that $H_1 \neq \emptyset$. Let u be any node of C that has a neighbor in $h_4 \cdots h_n$. Then there exists a path $P = p_1 \cdots p_k$ in C such that $p_k = u$, $p_1 \in H_1$, and no node of $P \setminus \{p_1\}$ belongs to H_1 . By Claim 3.5, exactly one of h_1, h_3 has a neighbor in $P \setminus \{p_1\}$. If h_1 does, then we say that u is *labeled 1* w.r.t. H_1 , and otherwise u is *labeled 3* w.r.t. H_1 . Note that by Claim 3.5 this label is unique.

CLAIM 3.6. *If $H_1 \neq \emptyset$, every node of C adjacent to h_4 must be labeled 3 w.r.t. H_1 , and every node of C adjacent to h_n must be labeled 1 w.r.t. H_1 .*

Proof. Suppose, up to symmetry, that some node u of C adjacent to h_4 is labeled

1 w.r.t. H_1 . Let $P = p_1 \cdots p_k$ be a path of C such that $p_1 \in H_1$, $p_k = u$, and no node of $P \setminus \{p_1\}$ belongs to H_1 . Then h_3 has no neighbor in $P \setminus \{p_1\}$. Let p_j be the node of P with lowest index adjacent to h_4 . By Claim 3.5, we have $j \geq 3$ and h_1 has a neighbor in $p_3 \cdots p_j$. But then $\{p_1, p_3, \dots, p_j, h_1, h_3, h_4, x\}$ contains a $3PC(h_3 h_4 x, h_1)$. This proves the claim. $\square$

CLAIM 3.7. *No node of C is of Type 3.*

Proof. Assume the contrary. W.l.o.g. $H_1 \neq \emptyset$.

Suppose that $H_3 = \emptyset$. Let $P = p_1 \cdots p_k$ be a shortest path in C from a node p_1 of H_1 to a node p_k adjacent to h_4 . So no node of $P \setminus \{p_1\}$ belongs to H_1 , and no node of $P \setminus \{p_k\}$ is adjacent to h_4 . By Claims 3.5 and 3.6, we have $k \geq 3$, node h_3 has a neighbor in $P \setminus \{p_1, p_2\}$, node h_1 has no neighbor in $P \setminus \{p_1\}$, and so, by Claim 3.6 again, h_n has no neighbor in P . Then h_5 has no neighbor in P , else by Claim 3.4, applied to h_4 and P , we should have $p_k \in H_3$. Some node of $h_6, \dots, h_{n-1}$ must have a neighbor in P , else $P \cup (H \setminus \{h_2, h_3\}) \cup \{x\}$ induces an odd wheel with center x . Let h_i be such a node with smallest index. If i is odd, then $P \cup \{h_4, \dots, h_i, x\}$ contains an odd wheel with center x . So i is even. If h_i has a neighbor in $P \setminus \{p_k\}$, then $(P \setminus \{p_k\}) \cup \{h_3, \dots, h_i, x\}$ contains an odd wheel with center x . So p_k is the only neighbor of h_i in P . By Claims 3.1 and 3.3, node h_3 cannot be adjacent to p_k . But then $P \cup \{h_3, h_4, h_i, x\}$ contains a $3PC(xh_3h_4, p_k)$, a contradiction. So $H_3 \neq \emptyset$.

Repeating this argument, we obtain that $H_i \neq \emptyset$ for each odd i .

Let $y \cdots z$ be any shortest path in C from a node y of H_1 to a node z of H_3 . By Claim 3.6, z is labeled 3 w.r.t. H_1 , and y is labeled 3 w.r.t. H_3 . So there exists a largest odd integer i such that C contains a chordless path $P = p_1 \cdots p_k$ from a node p_1 of H_1 to a node p_k of H_i , with no intermediate nodes in $H_1 \cup H_i$, such that p_k is labeled 3 w.r.t. H_1 and p_1 is labeled i w.r.t. H_i . In particular, by Claim 3.5, we have $k \geq 3$, node h_i has a neighbor in $P \setminus \{p_k, p_{k-1}\}$, and h_{i+2} has no neighbor in $P \setminus \{p_k\}$. Also since p_k is labeled 3 w.r.t. H_1 , and since by Claim 3.6 all nodes of H_{n-1} are labeled 1 w.r.t. H_1 , it follows that $i < n - 1$. Since $H_{i+2} \neq \emptyset$, there exists a shortest path $Q = q_1 \cdots q_l$ in C such that $q_1 \in H_{i+2}$ and q_l has a neighbor in P . Node q_1 is not adjacent to p_k , for otherwise $\{x, h_{i+1}, h_{i+2}, h_{i+3}, p_k, q_1\}$ induces an odd wheel.

Suppose that q_l is not adjacent to p_k . If q_l has a neighbor in $P \setminus \{p_k, p_{k-1}\}$, then $(P \setminus \{p_{k-1}\}) \cup Q \cup \{h_i, h_{i+2}, x\}$ contains a square- $3PC(h_i, h_{i+2})$. So $N(q_l) \cap P = \{p_{k-1}\}$. Let $Q' = q'_1 \cdots q'_t$ be the path induced by $(P \setminus \{p_k\}) \cup Q$, where $q'_1 = q_1$ and $q'_t = p_1$. Suppose that q_1 is labeled 1 w.r.t. H_1 . By Claim 3.5 applied to Q' , this is possible only if p_k is the only neighbor of h_3 in P , so $i = 3$, and h_1 has a neighbor in Q . Then $k = 3$, for otherwise $Q \cup \{x, h_1, h_3, p_1, p_{k-1}, p_k\}$ contains a square- $3PC(h_1, h_3)$. But then $\{x, h_3, h_6\} \cup P \cup Q$ contains a square- $3PC(h_3, p_2)$. So q_1 is labeled 3 w.r.t. H_1 . If p_1 is labeled $i + 2$ w.r.t. H_{i+2} , the maximality of i is contradicted. Hence p_1 is labeled $i + 4$ w.r.t. H_{i+2} , and by Claim 3.5, we have $t \geq 3$, node h_{i+4} has a neighbor in $Q' \setminus \{q'_1, q'_2\}$, and h_{i+2} has no neighbor in $Q' \setminus \{q'_1\}$. If $l > 1$, then $(Q' \setminus \{q'_2\}) \cup \{p_k, h_{i+2}, h_{i+4}, x\}$ contains a square- $3PC(h_{i+2}, h_{i+4})$. So $l = 1$. But then $P \cup \{q_1, h_2, h_{i+2}, x\}$ contains a square- $3PC(p_{k-1}, h_{i+2})$. Therefore, q_l must be adjacent to p_k . Thus $l \geq 2$.

Let p_j be the node of P with smallest index adjacent to q_l . Let $Q' = q'_1 \cdots q'_t$ be the path induced by $Q \cup \{p_1, \dots, p_j\}$, where $q'_1 = q_1$ and $q'_t = p_1$. Note that h_1 cannot have a neighbor in Q , since otherwise $(P \setminus \{p_2\}) \cup Q \cup \{h_1, h_3, x\}$ contains a square- $3PC(h_1, h_3)$. So node q'_1 must be labeled 3 w.r.t. H_1 . Node q'_t must be labeled $i + 4$ w.r.t. H_{i+2} , else the maximality of i is contradicted. So, by Claim 3.5, we have $t \geq 3$, node h_{i+4} has a neighbor in $Q' \setminus \{q'_1, q'_2\}$, and h_{i+2} has no neighbor in $Q' \setminus \{q'_1\}$.

But then $P \cup (Q \setminus \{q_2\}) \cup \{h_{i+2}, h_{i+4}, x\}$ contains a square-3PC(h_{i+2}, h_{i+4}). This proves the claim. $\square$

By Claims 3.1, 3.3, and 3.7, the nodes of C that have a neighbor in H are of Type 1, 2, or 5. Let C' be a minimal connected induced subgraph of C such that for some $s, t \in \{1, \dots, n\}$ node h_t is not adjacent to h_s or h_{s+1} , and each of h_s, h_{s+1}, h_t has a neighbor in C' . W.l.o.g. $s = 1$. Let $P = p_1 \cdots p_k$ be a shortest path in C' such that h_1 is adjacent to p_1 and h_2 to p_k .

Suppose that $k = 1$. So p_1 is of Type 2. Let $Q = q_1 \cdots q_l$ be a path in C' such that q_1 has a neighbor in $H \setminus \{h_n, h_1, h_2, h_3\}$ and q_l is adjacent to p_1 . Thus $C' = Q \cup \{p_1\}$, and no node of $Q \setminus \{q_1\}$ has a neighbor in $H \setminus \{h_n, h_1, h_2, h_3\}$. If both h_1, h_2 have a neighbor in Q , then Q contradicts the minimality of C' . So we may assume that h_1 has no neighbor in Q . Then h_n has no neighbor in $Q \cup \{p_1\}$, for otherwise a subpath of $Q \cup \{p_1\}$ violates Claim 3.2 or 3.3. If h_2 too has no neighbor in Q , then similarly h_3 has no neighbor in $Q \cup \{p_1\}$, and then by Claim 3.2, $H \cup Q \cup p_1$ induces a 3PC(Δ, Δ), contradicting Claim 3.3. So h_2 has a neighbor in Q . Let h_t be the node of $H \setminus \{h_n, h_1, h_2, h_3\}$ with highest index adjacent to q_1 . Then $Q \cup \{h_t, \dots, h_n, h_1, h_2\}$ and $Q \cup \{h_1, h_2, h_t, x\}$ induce two wheels with center h_2 , one of which is odd, a contradiction. So $k > 1$.

Let $Q = q_0 \cdots q_l$ be a shortest path such that $q_0 \in H \setminus \{h_n, h_1, h_2, h_3\}$, $Q \setminus \{q_0\} \subseteq C'$, and q_l has a neighbor in P (possibly $l = 0$). So if $l > 0$, no node of $P \cup Q \setminus \{q_0, q_1\}$ has a neighbor in $H \setminus \{h_n, h_1, h_2, h_3\}$. Note that $C' = P \cup Q \setminus \{q_0\}$.

Suppose that h_1 has a neighbor in Q . So $l > 0$. Let h_t be the node of $h_4 \cdots h_{n-1}$ with smallest index adjacent to q_1 . Then, by the minimality of C' , $N(q_l) \cap P = \{p_1\}$ and h_2 has no neighbor in Q . If h_3 has no neighbor in $P \cup Q$, then $P \cup (Q \setminus \{q_0\}) \cup \{h_1, h_2, \dots, h_t\}$ or $P \cup (Q \setminus \{q_0\}) \cup \{x, h_1, h_2, h_t\}$ induces an odd wheel with center h_1 . So h_3 has a neighbor in $P \cup Q$. Then Claim 3.4, applied to h_2 and $P \cup Q$, implies that p_k is of Type 3, a contradiction. Therefore h_1 cannot have a neighbor in Q , and similarly neither can h_2 .

If q_l has exactly one neighbor p_a in P , then $P \cup Q \cup \{h_1, h_2, x\}$ induces a 3PC($h_1 h_2 x, p_a$). If q_l has two nonconsecutive neighbors in P , then the same node set contains a 3PC($h_1 h_2 x, q_l$). So $N(q_l) \cap P = \{p_a, p_{a+1}\}$. Node h_3 must have a neighbor in P , else $P \cup Q \cup \{h_1, h_2, \dots, q_0\}$ contains a 3PC($q_l p_a p_{a+1}, h_2$). By Claim 3.1, h_3 cannot be adjacent to p_1 . Now h_2, h_3, q_0 all have a neighbor in $C' \setminus \{p_1\}$, which is connected, and so the minimality of C' implies $q_0 = h_4$. Then $N(h_3) \cap P = \{p_k\}$, else h_1, h_3, h_4 all have a neighbor in $C' \setminus \{p_k\}$, contradicting the minimality of C' . But then $P \cup \{h_1, h_2, h_3, x\}$ induces an odd wheel with center h_2 . This completes the proof of the theorem. $\square$

4. On the complexity of several detection problems.

4.1. Detecting a 3PC($\cdot, \cdot$) or a 3PC($\Delta, \cdot$). Chudnovsky and Seymour gave an $O(n^{11})$ to decide if a graph contains a 3PC($\cdot, \cdot$) [9] and an $O(n^9)$ algorithm to decide if a graph contains a 3PC($\Delta, \cdot$) [7]. Here we give an $O(n^7)$ algorithm that decides whether a given graph has either a 3PC($\cdot, \cdot$) or a 3PC($\Delta, \cdot$). Say that a 3PC($\Delta, \cdot$) is *long* if its three paths have length at least 2. Otherwise, exactly one of its paths has length 1, and we say it is *short*. Here is a sufficient condition for a graph to have a 3PC($\cdot, \cdot$) or a long 3PC($\Delta, \cdot$).

LEMMA 4.1. *Let G be a graph with four nodes u, a, b, c and a set $W \subseteq V(G) \setminus \{u, a, b, c\}$ such that*

- $\{u, a, b, c\}$ induces a claw centered at u ;
- W induces a connected subgraph of G ;

- u has no neighbor in W ;
- Every node in $\{a, b, c\}$ has exactly one neighbor in W .

Then, G contains a $3PC(\cdot, \cdot)$ or a long $3PC(\Delta, \cdot)$.

Proof. Let a', b', c' be the neighbors of a, b, c in W . If $a' = b' = c'$, then $\{u, a, b, c, a'\}$ induces a $3PC(u, a')$. Now we may assume $a' \neq b'$. Then $G[W]$ contains a path with ends a' and b' , and we let P be a shortest such path. Let $Q = c', \dots, v$ be a path in $G[W]$ such that v has a neighbor in P and no node of $Q \setminus v$ has a neighbor in P . Such a Q exists because W is connected (possibly $c' = v$). Let d (resp., e) be the neighbor of v in P closest to a' (resp., to b'). Let T be the graph induced by $P \cup Q \cup \{u, a, b, c\}$. If $d = e$, then T is a $3PC(d, u)$. If d, e are distinct and adjacent, then T is a long $3PC(vde, u)$. If d, e are distinct and nonadjacent, then T contains a $3PC(v, u)$. $\square$

Now we can give an $O(n^6)$ algorithm for the detection of nonsquare- $3PC(\cdot, \cdot)$'s or long $3PC(\Delta, \cdot)$'s that is very similar to an $O(n^5)$ algorithm by Maffray and Trotignon [21] that detects $3PC(\Delta, \Delta)$'s or $3PC(\Delta, \cdot)$'s.

ALGORITHM.

INPUT: A graph G .

OUTPUT: The positive answer “ G contains a nonsquare- $3PC(\cdot, \cdot)$ or a long $3PC(\Delta, \cdot)$ ” if it does; else the negative answer “ G contains no nonsquare- $3PC(\cdot, \cdot)$ and no long $3PC(\Delta, \cdot)$.”

METHOD: For every claw $\{u, b_1, b_2, b_3\}$ centered at u do:

Step 1. Compute the set X_1 of those nodes of $V(G)$ that are adjacent to b_1 and not adjacent to u, b_2 , or b_3 , and the similar sets X_2, X_3 , and compute the set X of those nodes of $V(G)$ that are not adjacent to any of u, b_1, b_2, b_3 . Compute the connected components of X in G . For each component H of X , and for $i = 1, 2, 3$, if some node of H has a neighbor in X_i then mark H with label i .

Step 2. For every component H of X that has received label $i \in \{1, 2, 3\}$, and for every node x of X_i that has a neighbor in H , assign to x the other labels of H (if any). For each $i = 1, 2, 3$ and for every node x of X_i that has a neighbor in X_j with $j \in \{1, 2, 3\}$ and $j \neq i$, assign label j to x .

Step 3. If some node of $X_1 \cup X_2 \cup X_3$ gets two different labels, return the positive answer and stop.

If the positive answer has not been returned at Step 3, return the negative answer.

COMPLEXITY: $O(n^6)$.

Proof of correctness. Suppose that G contains a nonsquare- $3PC(\cdot, \cdot)$ or a long $3PC(\Delta, \cdot)$, say K . Let u, b_1, b_2, b_3 be the four nodes of a u -centered claw of K , and for $i = 1, 2, 3$ let c_i be the neighbor of b_i in $K \setminus \{u, b_1, b_2, b_3\}$. Let us observe what the algorithm will do when it examines the 4-tuple $\{u, b_1, b_2, b_3\}$. The algorithm will place the three nodes c_1, c_2, c_3 in the sets X_1, X_2, X_3 , respectively.

First suppose that K is neither a $3PC(\cdot, \cdot)$ with one of the paths of length 2 nor a $3PC(\Delta, \cdot)$ with all three paths of length 2. Then $K \setminus \{u, b_1, b_2, b_3, c_1, c_2, c_3\}$ is contained in a connected component H of X , and all three nodes c_1, c_2, c_3 have a neighbor in H , i.e., H gets assigned all three labels 1, 2, and 3. So c_1 gets labels 2 and 3, and hence Step 3 returns the positive answer. If K is a $3PC(\Delta, \cdot)$ with all three paths of length 2, then K is a $3PC(c_1 c_2 c_3, u)$, so by Step 2 c_1 gets labels 2 and 3, and hence Step 3 returns the positive answer. Finally, assume that K is a $3PC(\cdot, \cdot)$ with one of the paths of length 2. W.l.o.g. K is a $3PC(u, c_1)$. Let H_2 (resp., H_3) be the

connected component of $K \setminus \{u, b_1, b_2, b_3, c_1, c_2, c_3\}$ in which both c_1 and c_2 (resp., c_1 and c_3) have a neighbor. Since c_1 has a neighbor in both H_2 and H_3 , in Step 2 c_1 gets labels 2 and 3, and hence Step 3 returns the positive answer.

Conversely, suppose that the algorithm returns the positive answer when it is examining a u -centered claw $\{x, b_1, b_2, b_3\}$. So (up to symmetry) some node $c_1 \in X_1$ gets labels 2 and 3 at Step 2. This means that for $j = 2, 3$, there exists a path R_j from c_1 to a node of X_j such that the interior nodes of R_j (if any) lie in X . We can apply Lemma 4.1 to the claw $\{u, b_1, b_2, b_3\}$ and the set $W = V(R_2) \cup V(R_3)$, which implies that this subgraph (and thus G itself) contains a long 3PC($\Delta, \cdot$) or a 3PC($\cdot, \cdot$) (that is nonsquare because the c_i 's are pairwise distinct). This completes the proof of correctness.

Finding all 4-tuples takes time $O(n^4)$. For each 4-tuple, computing the sets X_1, X_2, X_3, X takes time $O(n^2)$. Finding the components of X takes time $O(n^2)$. Marking the components at the end of Step 1 can be done as follows: For each edge uv of G , if u is in a component H of X and v is in some X_i then mark H with label i . This takes time $O(n^2)$. Marking the nodes of $X_1 \cup X_2 \cup X_3$ at Step 2 can be done similarly. Thus the overall complexity is $O(n^6)$. $\square$

Detecting square-3PC($\cdot, \cdot$)'s is easy to do in time $O(n^6)$, as noted in the introduction. So, in order to detect 3PC($\cdot, \cdot$)'s or 3PC($\Delta, \cdot$)'s, we are left with the problem of deciding whether a graph has a short 3PC($\Delta, \cdot$). This is an NP-complete problem (proved in the next section) but we can solve it assuming that the graph has no 3PC($\cdot, \cdot$) and no long 3PC($\Delta, \cdot$). This could be done in time $O(n^9)$ using the algorithm of Chudnovsky et al. [7] that detects a 3PC($\Delta, \cdot$). We propose here something faster and simpler but based on the same ideas.

If K is a short 3PC(bde, u) such that u sees b , a is the neighbor of u along the path of K from u to d , and c is the neighbor of u along the path of K from u to e , then we say that (u, a, b, c, d, e) is a *frame* of K .

LEMMA 4.2. *Let G be a graph with no 3PC($\cdot, \cdot$) and no long 3PC($\Delta, \cdot$). Let K be a smallest short 3PC($\Delta, \cdot$) of G with frame (u, a, b, c, d, e) . Let P be the path of $K \setminus \{u\}$ between a and d . Let R be a shortest path of G between a and d such that the interior nodes of R are not adjacent to b, c, e . Then the graph induced by $(K \setminus P) \cup R$ induces a smallest short 3PC($\Delta, \cdot$) of G .*

Proof. Note that G has no 3PC($\Delta, \cdot$) smaller than K , because it has no long 3PC($\Delta, \cdot$) at all, and because K is a smallest short 3PC($\Delta, \cdot$). Let us denote by $r_1 = a, \dots, r_k = d$ the nodes of R . Let r_s be the neighbor of u in R with greatest index. Note that r_s exists because r_1 is a neighbor of u .

We claim that the graph induced by $(K \setminus P) \cup \{r_s, \dots, r_k\}$ induces a short 3PC($\Delta, \cdot$). Indeed, let Q be the path of $K \setminus \{u\}$ with endnodes c and e . Let us denote by $q_1 = c, \dots, q_l = e$ the nodes of Q . If no node of $r_s \dots r_{k-1}$ has neighbors in Q , then the claim holds. So, we may assume that there is a node r_t in $r_s \dots r_{k-1}$ that has a neighbor in the interior of Q , and we choose t maximum. Note that $t > 1$. Let i be the smallest index and j be the greatest index such that q_i, q_j are neighbors of r_t . If $t = s$, then $\{r_t, \dots, r_k, q_j, \dots, q_l, u, b\}$ induces a 3PC(bde, r_t) that is smaller than K , a contradiction. So $t > s$. If $i = j$, then $Q \cup \{r_t, \dots, r_k, u, b\}$ induces a 3PC(bde, q_j) that is smaller than K , a contradiction. If $j > i + 1$, then $\{r_t, \dots, r_k, q_1, \dots, q_i, q_j, \dots, q_l, u, b\}$ induces a 3PC(bde, r_t) that is smaller than K , a contradiction. So $j = i + 1$. There is a shortest path S with endnodes r_t and a in the graph induced by $P \cup \{r_t \dots, r_k\}$. If $d \notin V(S)$, then $Q \cup S \cup \{u, b\}$ induces a long 3PC($r_t q_i q_j, u$), a contradiction. If $d \in V(S)$, then $S \cup \{q_1, \dots, q_i, b, u\}$ induces a

$3PC(d, u)$, a contradiction. This proves the claim.

We proved that $K' = (K \setminus P) \cup \{r_s, \dots, r_k\}$ induces a $3PC(\Delta, \cdot)$. If $s > 1$, then K' is a $3PC(\Delta, \cdot)$ smaller than K , a contradiction. So $s = 1$, proving the lemma. $\square$

Now we can give an algorithm for the detection of short $3PC(\Delta, \cdot)$'s in graphs with no $3PC(\cdot, \cdot)$ and no long $3PC(\Delta, \cdot)$.

ALGORITHM.

INPUT: A graph G with no $3PC(\cdot, \cdot)$ and no long $3PC(\Delta, \cdot)$.

OUTPUT: The positive answer “ G contains a short $3PC(\Delta, \cdot)$ ” if it does; else the negative answer “ G contains no short $3PC(\Delta, \cdot)$.”

METHOD: For every 5-tuple (a, b, c, d, e) of nodes do:

Step 1. Compute in $V(G) \setminus (N(b) \cup N(c) \cup N(e))$ a shortest path P from a to d (if any). Compute in $V(G) \setminus (N(b) \cup N(a) \cup N(d))$ a shortest path Q from c to e (if any). If at least one of the paths does not exist, go to the next 5-tuple.

Step 2. Check if the edge-set of $G[V(P) \cup V(Q) \cup \{a, b, c, d, e\}]$ is exactly $E(P) \cup E(Q) \cup \{bd, be, de\}$. If it is not, go to the next 5-tuple.

Step 3. For every node u of G , check if ua , ub , and uc are the only edges from u to $V(P) \cup V(Q) \cup \{a, b, c, d, e\}$. If it is the case, return the positive answer and stop. Else, go to the next 5-tuple.

If, after checking every 5-tuple, the positive answer has not been returned, return the negative answer.

COMPLEXITY: $O(n^7)$.

Proof. If the algorithm gives the positive answer, let us consider the 5-tuple (a, b, c, d, e) , the paths P, Q , and the node u that make the algorithm stop. It is clear by the method that $V(P) \cup V(Q) \cup \{a, b, c, d, e\}$ induces a short $3PC(\Delta, \cdot)$. Conversely, if G has a short $3PC(\Delta, \cdot)$ K with frame (u, a, b, c, d, e) , then let us examine what the algorithm will do when checking the 5-tuple (a, b, c, d, e) . By two applications of Lemma 4.2, we see that the two paths computed by the algorithm can take the place of the corresponding paths of K to give another short $3PC(\Delta, \cdot)$ K' (possibly not K) with apex u . So, since u exists, the algorithm will find a node that has the same neighborhood as u (in K') and give the positive answer. This proves the correctness of the algorithm. Checking every 5-tuple takes $O(n^5)$, computing shortest paths takes $O(n^2)$, checking every possible edge at Step 2 takes $O(n)^2$, checking every u at Step 3 takes $O(n)$, and checking every neighbor of u takes $O(n)$. So the overall complexity is $O(n^7)$. $\square$

By the algorithms above, we obtain the following.

THEOREM 4.3. *There is an $O(n^7)$ -time algorithm that decides whether a graph has a $3PC(\cdot, \cdot)$ or a $3PC(\Delta, \cdot)$.*

4.2. NP-completeness results. Let us denote by “problem II” the decision problem whose input is a graph G , whose two nonadjacent nodes a, b of G are of degree 2, and whose question is the following: “Does G have a hole that contains both a, b ?” Bienstock [4] proved that this problem is NP-complete. Adapting Bienstock’s proof, Maffray and Trotignon [21] remarked that the problem remains NP-complete for triangle-free graphs. Here is an easy consequence.

THEOREM 4.4. *The problem of deciding whether a graph has an odd wheel is NP-complete. The problem of deciding whether a graph has a short $3PC(\Delta, \cdot)$ is NP-complete.*

Proof. Suppose there is a polynomial time algorithm $\mathcal{A}$ for the detection of short $3PC(\Delta, \cdot)$'s or an algorithm $\mathcal{A}'$ for the detection of odd wheels. Let G, a, b be an

instance of problem Π . Let b', b'' be the neighbors of b in G . Build a graph H by adding to G nodes c_1, c_2, c_3, c_4, c_5 and edges $c_1a, c_1c_2, c_1c_3, c_2c_3, c_2c_4, c_4b', c_3c_5, c_5b''$. Since G has no triangle, every short $3PC(\Delta, \cdot)$ in H has apex a . So there is a short $3PC(\Delta, \cdot)$ in H if and only if there is a hole passing through a and b in G . Similarly, there is an odd wheel in H if and only if there is a hole passing through a and b in G . Thus, algorithm $\mathcal{A}$ (or $\mathcal{A}'$) yields a polynomial time algorithm that solves the NP-complete problem Π . $\square$

When $k \geq 2$, a $kPC(\Delta, \cdot)$ is a graph induced by k chordless paths $P_1, \dots, P_k$, such that each path P_i is from a node x to a node $y_i \neq x$, the nodes $y_1, \dots, y_k$ are distinct and pairwise adjacent, and the union of any two paths P_i, P_j induces a hole. Note that this latter condition implies that at most one of the paths $P_1, \dots, P_k$ can have length 1.

When $k \geq 2$, a $kPC(\cdot, \cdot)$ is a graph induced by k chordless paths $P_1, \dots, P_k$, such that they all have the same endnodes, and the union of any two paths P_i, P_j induces a hole.

For any integer $n \geq 1$, let $K_{1,n}$ denote the graph on $n + 1$ nodes with n edges and a node of degree n . Adapting the proof of Bienstock, we prove that problem Π remains NP-complete for $K_{1,4}$ -free graphs. Before presenting the proof of this result, we point out that problem Π is polynomial for $K_{1,3}$ -free graphs. To see this, consider an instance (G, a, b) of problem Π , where G is $K_{1,3}$ -free. Consider the graph G' obtained from G by adding a node c of degree 2 adjacent to a and b . It is easy to see that G contains a hole going through a and b if and only if G' contains a $3PC(\cdot, \cdot)$. Since this last problem is polynomial [9], problem Π is polynomial when restricted to $K_{1,3}$ -free graphs.

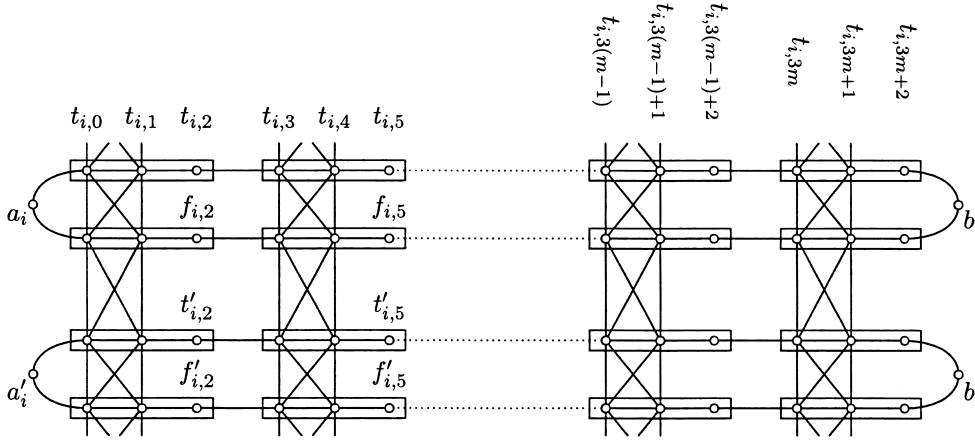
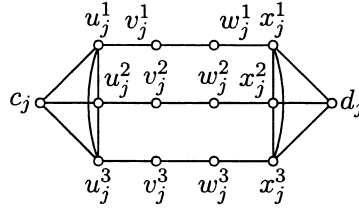
THEOREM 4.5. *Problem Π is NP-complete for $K_{1,4}$ -free graphs.*

Proof. Let us give a polynomial reduction from the problem 3-SATISFIABILITY of Boolean functions to problem Π restricted to $K_{1,4}$ -free graphs. Recall that a Boolean function with n variables is a mapping f from $\{0, 1\}^n$ to $\{0, 1\}$. A Boolean vector $\xi \in \{0, 1\}^n$ is a *truth assignment* for f if $f(\xi) = 1$. For any Boolean variable z on $\{0, 1\}$, we write $\bar{z} := 1 - z$, and each of $z, \bar{z}$ is called a *literal*. An instance of 3-SATISFIABILITY is a Boolean function f given as a product of clauses, each clause being the Boolean sum $\vee$ of three literals; the question is whether f admits a truth assignment. The NP-completeness of 3-SATISFIABILITY is a fundamental result in complexity theory; see [15].

Let f be an instance of 3-SATISFIABILITY consisting of m clauses $C_1, \dots, C_m$ on n variables $z_1, \dots, z_n$. Let us build a graph G_f with two specialized nodes a, b such that there will be a hole containing a and b in G if and only if there exists a truth assignment for f .

For each variable z_i ($i = 1, \dots, n$), make a graph $G(z_i)$ with four nodes a_i, b_i, a'_i, b'_i , and $4(3(m + 1))$ nodes $t_{i,j}, f_{i,j}, t'_{i,j}, f'_{i,j}$, with $j \in \{0, \dots, 3m + 2\}$. Add edges so that the four sets $\{a_i, t_{i,0}, t_{i,1}, \dots, t_{i,3m+2}, b_i\}$, $\{a_i, f_{i,0}, f_{i,1}, \dots, f_{i,3m+2}, b_i\}$, $\{a'_i, t'_{i,0}, t'_{i,1}, \dots, t'_{i,3m+2}, b'_i\}$, $\{a'_i, f'_{i,0}, f'_{i,1}, \dots, f'_{i,3m+2}, b'_i\}$ all induce paths (and the nodes appear in this order along these paths). For $k = 0, \dots, m$, add every possible edge between $\{t_{i,3k}, t_{i,3k+1}\}$ and $\{f_{i,3k}, f_{i,3k+1}\}$, between $\{f_{i,3k}, f_{i,3k+1}\}$ and $\{t'_{i,3k}, t'_{i,3k+1}\}$, between $\{t'_{i,3k}, t'_{i,3k+1}\}$ and $\{f'_{i,3k}, f'_{i,3k+1}\}$, and between $\{f'_{i,3k}, f'_{i,3k+1}\}$ and $\{t_{i,3k}, t_{i,3k+1}\}$. See Figure 1.

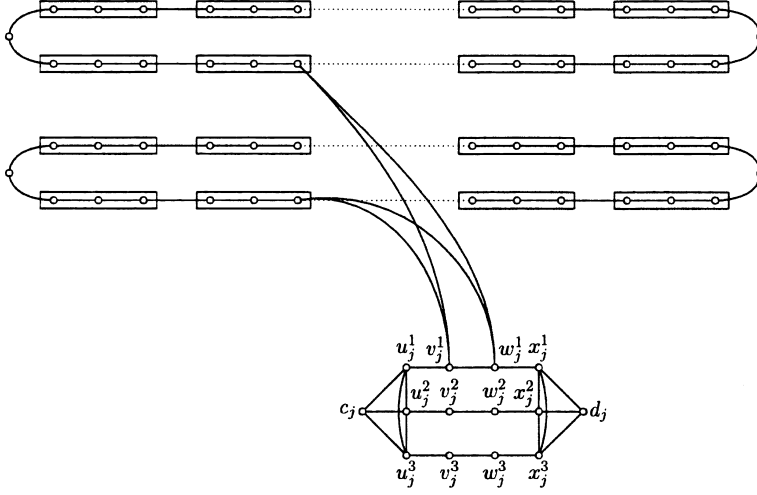
For each clause C_j ($j = 1, \dots, m$), with $C_j = y_j^1 \vee y_j^2 \vee y_j^3$, where each y_j^p ($p = 1, 2, 3$) is a literal from $\{z_1, \dots, z_n, \bar{z}_1, \dots, \bar{z}_n\}$, make a graph $G(C_j)$ with fourteen nodes $c_j, d_j, u_j^1, v_j^1, w_j^1, x_j^1, u_j^2, v_j^2, w_j^2, x_j^2, u_j^3, v_j^3, w_j^3, x_j^3$. Add edges so that the

FIG. 1. Graph $G(z_i)$.FIG. 2. Graph $G(C_j)$.

three sets $\{c_j, u_j^1, v_j^1, w_j^1, x_j^1, d_j\}$, $\{c_j, u_j^2, v_j^2, w_j^2, x_j^2, d_j\}$, $\{c_j, u_j^3, v_j^3, w_j^3, x_j^3, d_j\}$ all induce paths (and the nodes appear in this order along these paths). Add six more edges so that $\{u_j^1, u_j^2, u_j^3\}$, $\{x_j^1, x_j^2, x_j^3\}$ induce two triangles. See Figure 2.

The graph G_f is obtained from the disjoint union of the $G(z_i)$'s and the $G(C_j)$'s as follows. For $i = 1, \dots, n-1$, add edges $b_i a_{i+1}$ and $b'_i a'_{i+1}$. Add an edge $b'_n c_1$. For $j = 1, \dots, m-1$, add an edge $d_j c_{j+1}$. Introduce the two special nodes a, b and add edges aa_1, aa'_1 and bd_m, bb_n . See Figure 4. For $p = 1, 2, 3$, if $y_j^p = z_i$, then add four edges $v_j^p f_{i,3j-1}, v_j^p f'_{i,3j-1}, w_j^p f_{i,3j-1}, w_j^p f'_{i,3j-1}$, while if $y_j^p = \bar{z}_i$, then add four edges $v_j^p t_{i,3j-1}, v_j^p t'_{i,3j-1}, w_j^p t_{i,3j-1}, w_j^p t'_{i,3j-1}$. See Figure 3. Clearly the size of G_f is polynomial (actually quadratic) in the size $n+m$ of f . Moreover, it is a routine matter to check that G_f contains no $K_{1,4}$, and that a, b are nonadjacent and both have degree 2.

Suppose that f admits a truth assignment $\xi \in \{0, 1\}^n$. We can build a hole in G by selecting nodes as follows. Select a, b . For $i = 1, \dots, n$, select a_i, b_i, a'_i, b'_i . If $\xi_i = 1$, select $t_{i,j}, t'_{i,j}$ where $j \in \{0, \dots, 3m+2\}$. If $\xi_i = 0$, select $f_{i,j}, f'_{i,j}$ where $j \in \{0, \dots, 3m+2\}$. For $j = 1, \dots, m$, since ξ is a truth assignment for f , at least one of the three literals of C_j is equal to 1, say $y_j^p = 1$ for some $p \in \{1, 2, 3\}$. Then select c_j, d_j and $u_j^p, v_j^p, w_j^p, x_j^p$. Now it is a routine matter to check that the selected nodes induce a cycle Z that contains a, b , and that Z is chordless, so it is a hole. The main point is that there is no chord in Z between some subgraph $G(C_j)$ and some subgraph $G(z_i)$, for that would be either an edge $t_{i,3j-1} v_j^p$ (or similarly $t'_{i,3j-1} v_j^p$,

FIG. 3. The four edges added to G_f in the case $y_j^1 = z_i$.FIG. 4. Graph G_f .

$t_{i,3j-1}w_j^p, t'_{i,3j-1}w_j^p$) with $y_j^p = z_i$ and $\xi_i = 1$, or, symmetrically, an edge $f_{i,3j-1}v_j^p$ (or similarly $f'_{i,3j-1}v_j^p, f_{i,3j-1}w_j^p, f'_{i,3j-1}w_j^p$) with $y_j^p = \bar{z}_i$ and $\xi_i = 0$, and in either case this would contradict the way the nodes of Z were selected.

Conversely, suppose that G_f admits a hole Z that contains a, b . Clearly Z contains a_1, a'_1 since these are the only neighbors of a in G_f .

CLAIM 4.1. *For $i = 1, \dots, n$, Z contains exactly $6m + 10$ nodes of $G(z_i)$: four of these are a_i, a'_i, b_i, b'_i , and the others are either the $t_{i,q}, t'_{i,q}$'s or the $f_{i,q}, f'_{i,q}$'s, where $q \in \{0, \dots, 3m + 2\}$.*

Proof. First we prove the claim for $i = 1$. Since a, a_1 are in Z and a_1 has only three neighbors $a, t_{1,0}, f_{1,0}$, exactly one of $t_{1,0}, f_{1,0}$ is in Z . Likewise exactly one of $t'_{1,0}, f'_{1,0}$ is in Z . If $t_{1,0}, f'_{1,0}$ are in Z , then the nodes $a, a_1, a'_1, t_{1,0}, f'_{1,0}$ are all in Z and they induce a hole that does not contain b , a contradiction. Likewise we do not have both $t'_{1,0}, f_{1,0}$ in Z . Therefore, up to symmetry we may assume that $t_{1,0}, t'_{1,0}$ are in Z and $f_{1,0}, f'_{1,0}$ are not. This implies $t_{1,1}, t'_{1,1} \in Z$.

Suppose that for some $j \in \{1, \dots, m + 1\}$ and $k \in \{0, 1, 2\}$ one of $t_{1,3(j-1)+k}, t'_{1,3(j-1)+k}$ is not in Z . Let $3(j-1) + k$ be minimum with that property and assume up to a symmetry that $t_{1,3(j-1)+k}$ is not in Z . If $k = 0$ or $k = 1$, then Z contains up to a symmetry $f_{1,3(j-1)+k}$ that is adjacent to $t'_{1,3(j-1)+k}$, so Z cannot contain a, b , a contradiction. So, $k = 2$, $t_{1,3(j-1)+k}$ is in Z and one of v_j^p, w_j^p , $p \in \{1, 2, 3\}$, say v_j^1 up to a symmetry, must be in Z . But then, by the definition of G_f , v_j^1 is adjacent to $t'_{1,3(j-1)+2}$. So, Z has node set $\{a, t_{1,0}, t'_{1,0}, \dots, t_{1,3(j-1)+2}, t'_{1,3(j-1)+2}, v_j^1\}$ and b is not in Z , a contradiction. So, for $q = 0, \dots, 3m + 2$, we have $t_{1,j}, t'_{1,j} \in Z$ and $b_1, b'_1 \in Z$.

Now, for $j = 0, \dots, m$, none of $f_{i,3j}, f_{i,3j+1}$ can be in Z . So, $f_{i,3j+2}$ cannot be in Z because only one of its neighbor can be in Z . In particular, $a_2 \in Z$ and similarly $a'_2 \in Z$.

This proves our claim for $i = 1$. The proof of the claim for $i = 2$ is essentially the same as for $i = 1$, and by induction the claim holds up to $i = n$. $\square$

CLAIM 4.2. *For $j = 1, \dots, m$, Z contains c_j, d_j and exactly one of $\{u_j^1, v_j^1, w_j^1, x_j^1\}, \{u_j^2, v_j^2, w_j^2, x_j^2\}, \{u_j^3, v_j^3, w_j^3, x_j^3\}$.*

Proof. First we prove this claim for $j = 1$. By Claim 4.1, b'_n is in Z and exactly one of $t'_{n,3m+2}, f'_{n,3m+2}$ is in Z , so (since b'_n has degree 3 in G_f) c_1 is in Z . Consequently exactly one of u_1^1, u_1^2, u_1^3 is in Z , say u_1^1 . The neighbor of u_1^1 in $Z \setminus c_1$ cannot be a node among u_1^2, u_1^3 , for this would imply that Z contains a triangle. Hence $v_1^1 \in Z$. The neighbor of v_1^1 in $Z \setminus u_1^1$ cannot be in some $G(z_i)$ ($1 \leq i \leq n$), for in that case that neighbor would be either $t_{i,2}$ (or $f_{i,2}$) and thus, by Claim 4.1, node $t'_{i,2}$ (or $f'_{i,2}$) would be a third neighbor of v_1^1 in Z , a contradiction. Thus the other neighbor of v_1^1 in Z is w_1^1 . Similarly, we prove that w_1^1, x_1^1, d_1 are in Z , and so the claim holds for $j = 1$. Since d_1 has degree 4 in G_f and exactly one of x_1^1, x_1^2, x_1^3 is in Z , it follows that the fourth neighbor c_2 of d_1 is in Z . Now the proof of the claim for $j = 2$ is the same as for $j = 1$, and by induction the claim holds up to $j = m$. $\square$

We can now make a Boolean vector ξ as follows. For $i = 1, \dots, n$, if Z contains $t_{i,0}, t'_{i,0}$, then set $\xi_i = 1$; if Z contains $f_{i,0}, f'_{i,0}$, then set $\xi_i = 0$. By Claim 4.1 this is consistent. Consider any clause C_j ($1 \leq j \leq m$). By Claim 4.2 and up to symmetry we may assume that v_j^1 is in Z . If $y_j^1 = z_i$ for some $i \in \{1, \dots, n\}$, then the construction of G_f implies that $f_{i,3j-1}, f'_{i,3j-1}$ are not in Z , so $t_{i,3j-1}, t'_{i,3j-1}$ are in Z , so $\xi_i = 1$, so clause C_j is satisfied by x_i . If $y_j^1 = \bar{z}_i$ for some $i \in \{1, \dots, n\}$, then the construction of G_f implies that $t_{i,3j-1}, t'_{i,3j-1}$ are not in Z , so $f_{i,3j-1}, f'_{i,3j-1}$ are in Z , so $\xi_i = 0$, and so clause C_j is satisfied by $\bar{z}_i$. Thus ξ is a truth assignment for f . This completes the proof of the lemma. $\square$

THEOREM 4.6. *For each integer $k \geq 4$, the problem of deciding whether a graph contains a $kPC(\cdot, \cdot)$ is NP-complete, and the problem of deciding whether a graph contains a $kPC(\Delta, \cdot)$ is NP-complete.*

Proof. Let $k \geq 4$ be an integer. We give a reduction from problem Π to the problems whose NP-completeness is claimed. So let (G, a, b) be any instance of problem Π , where G is a $K_{1,4}$ -free graph and a, b are nonadjacent nodes of G of degree 2. Let us call a', a'' the two neighbors of a and b', b'' the two neighbors of b in G .

Reduction to the detection of a $kPC(\cdot, \cdot)$. Starting from G , build a graph G' as follows: Add nodes $y_1, \dots, y_{k-2}$. Add edges $ay_1, \dots, ay_{k-2}, by_1, \dots, by_{k-2}$. We see that G' contains a $kPC(\cdot, \cdot)$ if and only if G contains a hole that contains a and b . So every instance of Π can be reduced polynomially to an instance of the detection of a $kPC(\cdot, \cdot)$, which proves that this problem is NP-complete.

Reduction to the detection of a $kPC(\Delta, \cdot)$. Starting from G , build the same graph G' as above. Subdivide every edge ay_i , $i \in 1, \dots, k-2$, by adding a node z_i of degree 2. Subdivide edge aa' by adding a node a''' of degree 2. Add every possible edge between the nodes of $\{b', b'', y_1, \dots, y_{k-2}\}$. We see that G' contains a $kPC(\Delta, \cdot)$ if and only if G contains a hole that contains a and b . So every instance of Π can be reduced polynomially to an instance of the detection of a $kPC(\Delta, \cdot)$, which proves that this problem is NP-complete. $\square$

REFERENCES

- [1] V. E. ALEKSEEV, *On the number of maximal independence sets in graphs from hereditary classes*, in Combinatorial-Algebraic Methods in Applied Mathematics, V. N. Shevchenko, ed., Gorky University Press, Gorky, 1991, pp. 5–8 (in Russian).
- [2] C. AOSSEY, *3PC($\cdot, \cdot$)-Free Berge Graphs Are Perfect*, Dissertation, University of Kentucky, Lexington, KY, 2000.
- [3] C. AOSSEY AND K. VUŠKOVIĆ, *3PC($\cdot, \cdot$)-Free Berge Graphs Are Perfect*, manuscript, 1999.
- [4] D. BIENSTOCK, *On the complexity of testing for odd holes and induced odd paths*, Discrete Math., 90 (1991), pp. 85–92; see also B. Reed, *Corrigendum*, Discrete Math., 102 (1992), p. 109.
- [5] N. CHIBA AND T. NISHIZEKI, *Arboricity and subgraph listing algorithms*, SIAM J. Comput., 14 (1985), pp. 210–223.
- [6] M. CHUDNOVSKY, *Berge Trigraphs and Their Applications*, Ph.D. dissertation, Princeton University, Princeton, NJ, 2003.
- [7] M. CHUDNOVSKY, G. CORNUÉJOLS, X. LIU, P. SEYMOUR, AND K. VUŠKOVIĆ, *Recognizing Berge graphs*, Combinatorica, 25 (2005), pp. 143–187.
- [8] M. CHUDNOVSKY, N. ROBERTSON, P. SEYMOUR, AND R. THOMAS, *The strong perfect graph theorem*, Ann. of Math. (2), 164 (2006), pp. 51–229.
- [9] M. CHUDNOVSKY AND P. SEYMOUR, *The Three-in-a-Tree Problem*, preprint, 2006, submitted.
- [10] M. CONFORTI, G. CORNUÉJOLS, A. KAPOOR, AND K. VUŠKOVIĆ, *A Mickey Mouse decomposition theorem*, in Proceedings of the Fourth International Integer Programming and Combinatorial Optimization Conference (Copenhagen, Denmark), E. Balas and J. Clausen, eds., Lecture Notes in Comput. Sci. 920, Springer, Berlin, 1995, pp. 321–328.
- [11] M. CONFORTI, G. CORNUÉJOLS, A. KAPOOR, AND K. VUŠKOVIĆ, *Universally signable graphs*, Combinatorica, 17 (1997), pp. 67–77.
- [12] M. CONFORTI, G. CORNUÉJOLS, AND K. VUŠKOVIĆ, *Square-free perfect graphs*, J. Combin. Theory Ser. B, 90 (2004), pp. 257–307.
- [13] H. EVERETT, C. M. H. DE FIGUEIREDO, C. LINHARES SALES, F. MAFFRAY, O. PORTO, AND B. A. REED, *Even pairs*, in Perfect Graphs, J. L. Ramírez-Alfonsín and B. A. Reed, eds., Wiley Interscience, Chichester, 2001, pp. 67–92.
- [14] M. FARBER, *On diameters and radii of bridged graphs*, Discrete Math., 73 (1989), pp. 249–260.
- [15] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, San Francisco, 1979.
- [16] M. GRÖSTCHEL, L. LOVÁSZ, AND A. SCHRIJVER, *The ellipsoid method and its consequences in combinatorial optimization*, Combinatorica, 1 (1981), pp. 169–197.
- [17] M. GRÖSTCHEL, L. LOVÁSZ, AND A. SCHRIJVER, *Geometric Algorithms and Combinatorial Optimization*, 2nd ed., Springer-Verlag, Berlin, 1993.
- [18] W.-L. HSU AND G. L. NEMHAUSER, *Algorithms for minimum covering by cliques and maximum clique in claw-free perfect graphs*, Discrete Math., 37 (1981), pp. 181–191.
- [19] W.-L. HSU AND G. L. NEMHAUSER, *Algorithms for maximum weight clique, minimum weighted clique covers and minimum colorings of claw-free perfect graphs*, in Topics on Perfect Graphs, C. Berge and V. Chvátal, eds., North-Holland, Amsterdam, 1984, pp. 357–369.
- [20] L. LOVÁSZ, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory, 25 (1979), pp. 1–7.
- [21] F. MAFFRAY AND N. TROTIGNON, *Algorithms for perfectly contractile graphs*, SIAM J. Discrete Math., 19 (2005), pp. 553–574.
- [22] O. J. MURPHY, *Computing independent sets in graphs with large girth*, Discrete Appl. Math., 35 (1992), pp. 167–170.
- [23] S. D. NIKOLOPOULOS AND L. PALIOS, *Hole and antihole detection in graphs*, in Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms (New Orleans, LA), SIAM, Philadelphia, 2004, pp. 850–859.
- [24] I. PARFENOFF, F. ROUSSEL, AND I. RUSU, *Triangulated neighborhoods in C_4 -free Berge graphs*, in Proceedings of Graph-Theoretic Concepts in Computer Science (WG'99), Lecture Notes in Comput. Sci. 1665, Springer, Berlin, 1999, pp. 402–412.
- [25] J. SPINRAD, *Efficient Graph Representations*, Fields Inst. Monogr. 19, American Mathematical Society, Providence, RI, 2003.
- [26] K. TRUEMPER, *Alpha-balanced graphs and matrices and $GF(3)$ -representability of matroids*, J. Combin. Theory Ser. B, 32 (1982), pp. 112–139.
- [27] S. TSUKIYAMA, M. IDE, H. ARIYOSHI, AND I. SHIRAKAWA, *A new algorithm for generating all the maximal independent sets*, SIAM J. Comput., 6 (1977), pp. 505–517.

A NEW ALGORITHM FOR ON-LINE COLORING BIPARTITE GRAPHS*

HAJO J. BROERSMA[†], AGOSTINO CAPPONI[‡], AND DANIEL PAULUSMA[†]

Abstract. We first show that for any bipartite graph H with at most five vertices there exists an on-line competitive algorithm for the class of H -free bipartite graphs. We then analyze the performance of an on-line algorithm for coloring bipartite graphs on various subfamilies. The algorithm yields new upper bounds for the on-line chromatic number of bipartite graphs. We prove that the algorithm is on-line competitive for P_7 -free bipartite graphs, i.e., that do not contain an induced path on seven vertices. The number of colors used by the on-line algorithm for P_6 -free and P_7 -free bipartite graphs is, respectively, bounded by roughly twice and roughly eight times the on-line chromatic number. In contrast, it is known that there exists no competitive on-line algorithm to color P_6 -free (or P_7 -free) bipartite graphs, i.e., for which the number of colors is bounded by any function depending only on the chromatic number.

Key words. on-line coloring, bipartite graph, (on-line) competitive

AMS subject classifications. 05C15, 05C85

DOI. 10.1137/060668675

1. Introduction. In static optimization problems one is often faced with the challenge of determining efficient algorithms that solve a particular problem (nearly) optimally for any given instance of the problem. This task is usually facilitated if the structure of the instances is pretty straightforward. As an example, it is a trivial exercise to determine an algorithm for finding a 2-coloring of a given bipartite graph.

In the area of dynamic optimization the situation gets more complicated. There, one often lacks the knowledge of the complete instances of the problems. As an illustration, compare the previous problem with the slightly changed situation in which the bipartite graph is presented on-line, i.e., vertex by vertex, and the algorithm has to assign a color irrevocably to a vertex as it comes in, i.e., only based on the knowledge of the subgraph that has been revealed so far. This slight change of the problem formulation makes it a lot more difficult: whereas the static problem was trivial, no algorithm for the dynamic problem can guarantee an optimal solution for every instance. In [12] it has been shown that the worst-case performance ratio between on-line and off-line coloring of a known input graph on n vertices is at least $2n/(\log_2 n)^2$. It is even questionable whether one can expect to determine an on-line algorithm that does reasonably well, in the sense that the number of colors used is bounded in some other reasonable way, e.g., as a function of some invariant of the input instances.

In this paper we will focus on particular questions of this type related to coloring bipartite graphs. These types of questions in a more general setting are at the heart of the areas of on-line algorithms and of approximation algorithms.

*Received by the editors August 24, 2006; accepted for publication (in revised form) July 17, 2007; published electronically February 8, 2008. An extended abstract of some of the results in this paper was presented at the 6th Italian Conference on Algorithms and Complexity 2006 [3].

<http://www.siam.org/journals/sidma/22-1/66867.html>

[†]Department of Computer Science, Durham University, South Road, Durham, DH1 3LE, UK (hajo.broersma@durham.ac.uk, daniel.paulusma@durham.ac.uk).

[‡]Computer Science, Division of Engineering and Applied Sciences, California Institute of Technology, Pasadena, CA 91125 (acaponi@cs.caltech.edu).

We first give a short historical excursion starting with a benchmark paper from Gyárfás and Lehel [9]. They introduced the concept of on-line coloring as a general approach. This was motivated by their translation of a rectangle packing problem related to dynamical storage allocation appearing in [4] into an on-line coloring problem. The latter problem was to decide whether the on-line coloring algorithm known as *First-Fit* (*FF*) has a constant worst-case performance ratio on the family of interval graphs. We note that since [9] many papers on on-line (coloring) problems have appeared. We refer to [14] for a survey on on-line coloring and to [2] for more background on the general area of on-line algorithms.

In order to have some measure of the performance of on-line algorithms, the notion of competitive algorithms has been introduced in [20] and specifically for coloring in [9]. Intuitively, an on-line coloring algorithm is said to be competitive for a family of graphs $\mathcal{G}$ if, for any graph $G \in \mathcal{G}$, the number of colors used by the algorithm on G is bounded from above by a function depending only on the chromatic number of G . The chromatic number of G is the smallest number of colors that is necessary to properly color the vertices of G (off-line), i.e., such that adjacent vertices receive different colors. In [13] it is shown that *FF* is competitive for interval graphs, with a bounding function that is linear in the chromatic number, and in [6] competitiveness of *FF* for geometric intersection graphs has been proven. In [5] it is shown that *FF* is competitive for graphs with a bounded independence number, including co-planar graphs.

It is well known that *FF* is not competitive for P_6 -free bipartite graphs, i.e., bipartite graphs that do not contain an induced path on six vertices: if the vertices of a complete bipartite graph $K_{m,m}$ minus a perfect matching $\{u_1, v_1\}, \{u_2, v_2\}, \dots, \{u_m, v_m\}$ are presented in the ordering $u_1, v_1, u_2, v_2, \dots, u_m, v_m$, then *FF* uses m colors. In fact, there are many families of graphs for which it has been proven that no competitive algorithms exist: two examples given in [9] are the family of trees and the family of P_6 -free bipartite graphs. These negative results have led to the definition of a weaker form of competitiveness in [7], namely on-line competitiveness, although results of this type have been obtained before the term was formally introduced.

We can explain the notion of on-line competitiveness as follows. It is perhaps more natural to compare the number of colors that is needed to color a graph G by an on-line algorithm to the on-line chromatic number instead of the chromatic number. We will define the on-line chromatic number formally in section 3. Intuitively, it is the number of colors used by the best performing on-line algorithm for G , i.e., that gives the smallest number of colors in the worst-case ordering of the presented vertices of G . An on-line coloring algorithm is said to be on-line competitive if the number of colors is bounded from above by a function only depending on the on-line chromatic number of G . The main open problem is whether on-line competitive coloring algorithms exist for all graphs. A subproblem is to establish such algorithms for certain classes of graphs for which competitive algorithms do not exist. Solving this subproblem can be particularly useful if somebody must design an on-line coloring algorithm in the case that the input graphs are only known to be in a specified class of graphs. This also motivated our choice for considering special graph classes.

It is shown in [10] that *FF* is on-line competitive for trees; it is even optimal for trees, in the sense that if *FF* uses k colors, then the on-line chromatic number of the tree is also k . In [7] it is shown that *FF* is on-line competitive with an exponential bounding function for graphs with girth of at least five.

In the context of algorithmic graph theory it has become rather fashionable to

consider forbidden subgraph conditions. For instance, many NP-hard problems turn out to be solvable in polynomial time when restricted to H -free graphs for particular choices of H . Therefore, these graph classes are well studied throughout a range of NP-hard problems. In the context of coloring, e.g., 3-colorability is polynomially solvable for P_6 -free graphs, while 4-colorability remains NP-hard for P_{12} -free graphs, and 5-colorability remains NP-hard for P_8 -free graphs. We refer the reader to the survey paper [19] for more details. Note also that well-studied graph classes like chordal graphs (or, more generally, perfect graphs) and line graphs (or, more generally, claw-free graphs) can be characterized by forbidden subgraph conditions.

2. Results of this paper. One of the main open problems concerning on-line competitive coloring algorithms [7] is to decide whether for every k there exists an on-line competitive coloring algorithm for the family of graphs with on-line chromatic number k . Perhaps surprisingly, this is even open for bipartite graphs for $k = 4$, whereas it has been solved for general graphs for $k \leq 3$: in both [8] and [17] it is proven that for the family of graphs with on-line chromatic number 3 at most four colors are needed. The open problem on bipartite graphs seems to be very hard and emphasizes how much on-line coloring differs from off-line coloring.

Our results are motivated by a number of open problems, but most strongly by the above open problem for bipartite graphs. We solve the problem for several subclasses of bipartite graphs which are defined by forbidding a certain fixed bipartite graph H as an induced subgraph. For a relatively small graph H this is an easy exercise, but for larger graphs this gets difficult, in correspondence with the fact that the class of H -free graphs contains the class of H' -free graphs if H' is a subgraph of H . By combining known results and dealing with a few cases ourselves, we show that for every bipartite graph H with at most five vertices there exists an on-line competitive coloring algorithm for the class of H -free bipartite graphs. Since for P_4 -free and P_5 -free graphs there even exists a competitive algorithm [9, 11], and since P_6 -free bipartite graphs do not admit a competitive algorithm [9], our natural starting point from there is the latter class. In [3] we proved that there exists an on-line competitive algorithm for the class of P_6 -free bipartite graphs; its bounding function is *linear* in the on-line chromatic number, namely roughly twice the on-line chromatic number. In fact, this gives a 2-approximation algorithm for on-line coloring P_6 -free bipartite graphs. In this paper we prove a similar result for the larger class of P_7 -free bipartite graphs with a bounding function that is roughly eight times the on-line chromatic number. Note that the on-line chromatic number for both these graph classes can be arbitrarily high, so these classes are definitely no subclasses of the class of bipartite graphs with on-line chromatic number 4. In this sense, our results have a broader appeal than just solving the aforementioned problem with $k = 4$ for the restricted classes of P_6 -free and P_7 -free bipartite graphs. It might be possible that our algorithm or variations on it can be used to prove similar results for larger subclasses of bipartite graphs, although we have not been able to do so yet. We also note that our algorithm uses at most four colors (so it is competitive) for the class of P_5 -free bipartite graphs.

The rest of the paper is organized as follows. Section 3 contains the basic notation and definitions. In section 4 we start our exposition by proving the result on H -free bipartite graphs with $|V(H)| \leq 5$. Next we introduce the key algorithm of this paper called *BicolorMax*. We prove in section 6 that it is on-line competitive for P_7 -free bipartite graphs. In section 7 we deduce new upper bounds for the on-line chromatic number of bipartite graphs. We give some final remarks in section 8 and conclusions in section 9.

3. Preliminaries. Throughout we consider simple graphs $G = (V(G), E(G))$, where $V(G)$ is a set of vertices and $E(G)$ is a set of unordered pairs of vertices, called edges. For graph terminology not defined below we refer to [1].

A graph is called *bipartite* if its vertex set can be partitioned into two disjoint sets so that all edges are incident with vertices from both sets. If H is a *subgraph* of G , i.e., if H is a graph and $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$, then we write $H \subseteq G$. If $S \subseteq V(G)$, then $G[S]$ denotes the subgraph of G with vertex set S and edge set $\{\{x, y\} \mid x \in S, y \in S\}$. A graph is an *induced subgraph* of G if it is isomorphic to $G[S]$ for some nonempty $S \subseteq V(G)$. A graph is *H-free* if it does not contain the graph H as an induced subgraph. We call two vertex-disjoint graphs *remote* if there are no edges joining them. A maximal connected subgraph of a graph G is called a *component* of G . For any two vertices x, y of a connected graph G we denote by P_{xy} a path between x and y in G . We define the *distance* $d(x, y, G)$ between x and y in G as the number of edges of a shortest path between x and y . We use K_n , C_n , and P_n to denote, respectively, the complete graph, the cycle, and the path on n vertices, and we use $K_{m,n}$ to denote the complete bipartite graph with m vertices in one bipartition class and n vertices in the other. A *coloring* of a graph G is a function $c : V(G) \rightarrow \{1, 2, \dots\}$ such that $c(v) \neq c(w)$ whenever $\{v, w\} \in E(G)$. The smallest number of colors in a coloring of G is the *chromatic number* of G and denoted by $\chi(G)$. Clearly, a graph G is bipartite if and only if $\chi(G) \leq 2$.

We assume that the reader is familiar with the basic concept of an on-line coloring algorithm. For details we refer to [14]. Intuitively, an on-line coloring algorithm properly colors the vertices of a graph one by one, consistently using a fixed strategy, depending only on the subgraph induced by the revealed vertices and the colors that have been assigned to them by the algorithm, according to an externally determined ordering of the presented vertices.

A popular informative way of looking at on-line coloring is as a two-person game. In the two-person game, the *Drawer* reveals the vertices of G one by one together with their adjacencies to vertices already revealed. The *Painter* irrevocably assigns an admissible color to the currently revealed vertex, based on a strategy involving the graph induced by the already revealed vertices and their colors, and the adjacencies of the currently revealed vertex. The aim of Drawer is to reveal the vertices in an order that forces Painter to use many colors. The aim of Painter is to invent a strategy that uses as few colors as possible. The order in which Drawer reveals the vertices of G is an ordering of the vertices of $V(G)$. Any strategy Painter chooses can be extended to some on-line coloring algorithm in the above sense that assigns Painter's coloring to G . The reverse is also true: a coloring prescribed by an on-line coloring algorithm can be simulated by a two-person game with a Drawer and Painter as described above.

We denote the (finite) set of all on-line coloring algorithms for a graph G by $AOL(G)$. Let $\Pi(G)$ denote the set of all permutations of the vertices of G . If $A \in AOL(G)$ and $\pi \in \Pi(G)$, then we denote by $\chi_A(G, \pi)$ the number of colors used by A when the vertices of G are presented according to π . The largest number of colors used by the on-line algorithm A for G is called the *A-chromatic number* of G and denoted by $\chi_A(G)$. Hence

$$\chi_A(G) = \max_{\pi \in \Pi(G)} \chi_A(G, \pi).$$

The smallest number of colors used by an on-line algorithm for G is the *on-line chromatic number* of G and denoted by $\chi_{OL}(G)$ [9]. Hence

$$\chi_{OL}(G) = \min_{A \in AOL(G)} \chi_A(G).$$

Let $\mathcal{G}$ denote a (possibly infinite) family of graphs. If $A \in AOL(G)$ for every $G \in \mathcal{G}$, then we say that A is an on-line coloring algorithm for $\mathcal{G}$ and write $A \in AOL(\mathcal{G})$. An algorithm $A \in AOL(\mathcal{G})$ is said to be *competitive* for $\mathcal{G}$ if there exists a function f such that $\chi_A(G) \leq f(\chi(G))$ for every $G \in \mathcal{G}$; it is *on-line competitive* if $\chi_A(G) \leq f(\chi_{OL}(G))$ for every $G \in \mathcal{G}$.

4. Small forbidden subgraphs. As stated before, it has been shown that there does not exist a competitive on-line coloring algorithm for P_6 -free bipartite graphs, but there exists a competitive on-line coloring algorithm for P_5 -free bipartite graphs. In fact, combining results from [7, 11, 15, 16], and analyzing a few cases ourselves, we can show that there exists an on-line coloring algorithm that is on-line competitive for the class of H -free bipartite graphs for any fixed bipartite graph H on at most five vertices. We are not able to show such a result for any fixed graph H that is not bipartite. Since a bipartite graph does not contain a nonbipartite subgraph, this would come down to proving that there exists an on-line coloring algorithm that is on-line competitive for the class of bipartite graphs. This, however, is still a major open problem.

Before stating and proving our proposition below, let us first note that we are not aiming at obtaining nice performance ratios in our proof of this proposition. As we are primarily interested in the existence result, here we took the freedom to use rather unsophisticated methods for reaching our goal.

PROPOSITION 4.1. *Let H be a bipartite graph on at most five vertices. Then there exists an on-line coloring algorithm that is on-line competitive for the class of H -free bipartite graphs.*

Proof. We may restrict ourselves to bipartite graphs on exactly five vertices, noting that an F -free bipartite graph with F bipartite on at most four vertices is also H -free for some bipartite graph H on five vertices. We use $H + H'$ to denote the disjoint union of two graphs H and H' , and pH to denote the disjoint union of $p \geq 2$ copies of H . Before we make a case distinction we first make the following easy observation:

- (1) Let F be a graph and A an on-line coloring algorithm that is on-line competitive for the class of F -free bipartite graphs. Then there exists an on-line coloring algorithm A' that is on-line competitive for the class of $F + K_1$ -free bipartite graphs.

This claim can be seen as follows. Initially we use algorithm A to color the vertices of an $F + K_1$ -free bipartite graph G . If G does not contain an induced F , then we do not need any extra colors. Suppose G contains an induced subgraph G' that is isomorphic to F . We assume that G' is the first occurrence of F if G is revealed to A . Let v be the last vertex of G' presented to A . We color v with a new color c^* . Since G is $F + K_1$ -free, any vertex w that is presented after v must have a neighbor in G' . Let $Q_1, \dots, Q_p$ be the components of G' . We note that these components are bipartite, since G is bipartite. For each component Q_i ($i = 1, \dots, p$) we define a set of two new colors $\{c_i, d_i\}$, each of which corresponds to one of the classes in the bipartition of Q_i . Now suppose Q_h is the component that has the smallest index h of all the components of G' , which w is adjacent to. Then we color w with c_h or d_h , depending on which bipartite class of Q_h vertex w belongs to. This way we can finish the coloring of G with at most $2p + 1$ extra colors.

We now distinguish a number of cases depending on the value of $|E(H)| = m$.

Case I. $m = 0$. Then $H = 5K_1$. Since G is bipartite, we obtain $\chi_{FF} \leq \Delta(G) + 1 \leq 5$.

Case II. $m = 1$. Then $H = K_2 + 3K_1$. It is trivial to see that FF is on-line competitive for the class of K_2 -free graphs, i.e., graphs with only isolated vertices. After applying (1) three times we get the desired result.

Case III. $m = 2$. Then $H = P_3 + 2K_1$ or $2K_2 + K_1$. For the first subcase we can proceed similarly as in Case II. For the second subcase we use the following result from [11]:

(2) If G is a P_5 -free graph without triangles, then $\chi_{FF}(G) \leq 3$.

Noting that $2K_2$ -free bipartite graphs are both P_5 -free and triangle-free, and combining (1) and (2), yields the result.

Case IV. $m = 3$. Then $H = P_4 + K_1$, $K_{1,3} + K_1$, or $P_3 + K_2$. Noting that P_4 -free bipartite graphs are both P_5 -free and triangle-free, and combining (1) and (2), yields the desired result for the first subcase. For the second subcase we first observe that $\chi_{FF}(G) \leq \Delta(G) + 1 \leq 3$ for any $K_{1,3}$ -free bipartite graph G and then we apply (1) to get the result. Since a $P_3 + K_2$ -free bipartite graph is a P_7 -free bipartite graph, we can of course immediately apply Theorem 6.8 (which will be presented later) for the third subcase. It is also not difficult to give a direct proof that our algorithm *BicolorMax* (which will be presented in the next section) is on-line competitive for this class of graphs.

Case V. $m = 4$. Then $H = K_{1,4}$, $C_4 + K_1$, P_5 , or the unique graph with degree sequence 3,2,1,1,1 which we denote by $K_{1,3}^+$. For the first subcase we easily get that $\chi_{FF}(G) \leq \Delta(G) + 1 \leq 4$. The *girth* of a graph G is the number of edges of a smallest cycle in G . For the second subcase we combine (1) with the following result from [7]:

(3) If G has a girth of at least five, then $\chi_{FF}(G) \leq \binom{2^{\chi_{OL}(G)}}{2}$.

For the third subcase we use (2). The *radius* of a graph G is defined as the minimum of $\max_v d(u, v, G)$ over all vertices u in G . Since $K_{1,3}^+$ has a radius of 2, we can use the following result from [16]:

(4) For every tree T with a radius of 2, there is an on-line coloring algorithm A that is on-line competitive for the class of T -free graphs.

Case VI. $m = 5$. Then $H = K_{2,3} - e$ for an edge e of $K_{2,3}$. We need a separate proof for this case and first prove the following claim.

Claim. Let G be bipartite and H -free and let C be a component of G such that C_4 is an induced subgraph of C . Then $C = K_{s,t}$ for some integers $s, t \geq 2$.

We prove this claim as follows. If $C = C_4 = K_{2,2}$, then the claim trivially holds. If not, let $C_4 = uvwxu$, and let $N(p)$ denote the neighbors of vertex p in C . If $N(u) \not\subseteq N(w)$, then G contains H as an induced subgraph. So, by symmetry, $N(u) = N(w)$, and similarly $N(v) = N(x)$. Let $y \in N(u) \cap N(w)$. Then $uvwxyu$ is an induced C_4 , so as before $N(y) = N(v) = N(x)$. Hence all neighbors of u and w are adjacent to all neighbors of v and x and vice versa. By repeating the arguments for all induced C_4 's, we obtain that $C = K_{s,t}$ for some $s, t \geq 2$.

A component in a bipartite graph that does not contain an induced C_4 has a girth of at least six. Since $\chi_{FF}(K_{s,t}) = 2$, the above claim together with (3) then implies that $\chi_{FF}(G) \leq \max\{\binom{2^{\chi_{OL}(G)}}{2}, 2\}$.

Case VII. $m = 6$. Then $H = K_{2,3}$. Kierstead and Penrice [15] showed that FF is on-line competitive for the class of H -free graphs. $\square$

We conclude that the first open question with respect to the (non)existence of on-line competitive coloring algorithms for H -free bipartite graphs concerns bipartite graphs H on six vertices, in particular $H = P_6$, which has radius 3. In the next section we present an on-line algorithm for coloring general bipartite graphs. We will show that it is on-line competitive for (P_6 -free and) P_7 -free bipartite graphs.

5. The algorithm BicolorMax. Let G be a bipartite graph on n vertices denoted by $1, 2, \dots, n$. Let $A = \{a_1, a_2, \dots, a_p\}$ and $B = \{b_1, b_2, \dots, b_p\}$ be two disjoint ordered sets of colors. For a fixed positive integer $k \leq p$, let $A(k) = \{a_1, a_2, \dots, a_k\}$ and $B(k) = \{b_1, b_2, \dots, b_k\}$.

We first give the general idea of our on-line algorithm called *BicolorMax*. Suppose that G is presented to the algorithm. At some stage a new uncolored vertex v of G is revealed, together with its adjacencies to the set S of already colored vertices of G . If v is not adjacent to any previously revealed vertex of G , then v receives color a_1 . Otherwise, the choice of the color for v is based on the present colors in the bipartition classes of the component containing v of the subgraph of G induced by v and the vertices of S with colors in $A(k) \cup B(k)$ for some suitable $k \geq 1$. To make this choice explicit we first need to introduce some additional terminology.

If $F \subseteq V(G)$, then the *hue* of F , denoted by $H(F)$, is the set of all colors used on vertices in F . Let $\pi \in \Pi(G)$ for a bipartite graph G , and assume that $v = \pi(j)$. Let G_j^k denote the subgraph of $G[\{\pi(1), \dots, \pi(j)\}]$ induced by $v = \pi(j)$ and all the vertices in $\{\pi(1), \dots, \pi(j-1)\}$ that have been assigned colors from $A(k) \cup B(k)$. We denote by C_j^k the component of G_j^k containing v , and we write $C_j^k := (I_1, I_2)$ to indicate the bipartition of its vertex set. Note that (I_1, I_2) is the unique bipartition of C_j^k (up to renaming), because C_j^k is connected. We say that color a_k is *mixed* on $C_j^k = (I_1, I_2)$ if there exist at least two vertices, $v \in I_1$ and $w \in I_2$, that have been colored with a_k . We then call (v, w) a *k-mixed pair*.

We are now ready to explain the choice of the color of a new vertex. The key idea is to prevent the mixing of colors as much as possible. Recall that in the worst case FF uses m colors for a complete bipartite graph $K_{m,m}$ minus a perfect matching, because the worst-case ordering forces FF to assign the same color to each pair of the perfect matching; hence, all m colors are used on both bipartite classes. In order to reduce the total number of colors we somehow need to avoid this as much as possible by changing the coloring strategy with respect to mixing colors.

The algorithm *BicolorMax* has been designed with this intuitive idea in mind. It is defined inductively. Vertex $\pi(1)$ is colored with a_1 . Suppose that vertices $\pi(1), \dots, \pi(j-1)$ have already been colored. Let $v = \pi(j)$ be the next vertex presented to the algorithm. The algorithm first computes the highest value of k such that a_k is mixed on C_j^k . We denote this value by m . In order to stop the mixing of colors, the algorithm colors v with a_{m+1} or with b_{m+1} . By definition of m , the color a_{m+1} is not mixed on C_j^{m+1} . The choice will fall on a_{m+1} if v is in the same bipartition class as a vertex with color a_{m+1} or if a_{m+1} has not been used before on C_j^{m+1} . Otherwise v receives color b_{m+1} . This way the mixing of color a_{m+1} within a component is prevented. We note that a_{m+1} might get mixed later on: if at a certain stage two components that both contain a vertex with color a_{m+1} are joined by a new vertex, then a_{m+1} becomes mixed if at least two vertices with color a_{m+1} belong to different bipartition classes.

Before we give a formal description of our algorithm we would like to make the following remark. The algorithm FF uses at most three colors on any path P_n . However, *BicolorMax* does not try to use the colors a_k or b_k for any $k < m$ if a_m is mixed on C_j^m . Hence, it is easy to see that *BicolorMax* is not on-line competitive for the family of paths. In section 8 we return to this observation and show how we can combine FF and *BicolorMax* to obtain an on-line competitive algorithm for families of graphs for which FF is on-line competitive as well.

```

BicolorMax( $G[\{\pi(1), \dots, \pi(j-1)\}], v$ )

 $m := \max(\{0\} \cup \{k : a_k \text{ is mixed on } C_j^k\})$ .
if  $a_{m+1} \notin H(V(C_j^{m+1}))$ 
     $C_j^{m+1} := (I_1, I_2)$  such that  $v \in I_1$ 
else
     $C_j^{m+1} := (I_1, I_2)$  such that  $a_{m+1} \in H(I_1)$ .
if  $v \in I_1$ 
    assign color  $a_{m+1}$  to  $v$ 
else
    assign color  $b_{m+1}$  to  $v$ .
```

It is immediately clear that *BicolorMax* is not able to color a nonbipartite graph G . For instance, take $G = K_3$. Then $C_3^1 = C_3^2 = K_3$ cannot be partitioned into two independent sets I_1 and I_2 . For bipartite graphs, *BicolorMax* does assign a coloring for each vertex permutation.

OBSERVATION 5.1. *BicolorMax is an on-line coloring algorithm for bipartite graphs.*

Proof. Let $\pi \in \Pi(G)$ for a bipartite graph G . Let $v = \pi(j)$ be the j th vertex presented to *BicolorMax*. Let $m := \max(\{0\} \cup \{k : a_k \text{ is mixed on } C_j^k\})$.

Suppose $a_{m+1} \notin H(V(C_j^{m+1}))$. This means that v is not adjacent to a previously presented vertex with color a_{m+1} . Hence, *BicolorMax* may color v with color a_{m+1} in a proper vertex coloring.

Suppose $a_{m+1} \in H(V(C_j^{m+1}))$. By maximality of m , we find that a_{m+1} is not mixed on C_j^{m+1} . Then, a_{m+1} is only assigned to vertices in one bipartite class of C_j^{m+1} , and we can indeed define $C_j^{m+1} = (I_1, I_2)$ with $a_{m+1} \in H(I_1)$ and $a_{m+1} \notin H(I_2)$. If v is in I_1 , then *BicolorMax* colors v with a_{m+1} . Since $a_{m+1} \notin H(I_2)$, vertex v is not adjacent to any previously presented vertex with color a_{m+1} . If v is in I_2 , then *BicolorMax* colors v with b_{m+1} . Suppose v is adjacent to a vertex $u = \pi(i)$ for some $i \leq j-1$ with color b_{m+1} . By definition of the algorithm, $C_i^{m+1} = (I'_1, I'_2)$ with $a_{m+1} \in H(I'_1)$ and $u \in I'_2$. Since $C_i^{m+1} \subseteq C_j^{m+1}$, we find that $I'_1 \subseteq I_1$ and $I'_2 \subseteq I_2$. This leads to a contradiction because the vertices u and v cannot be adjacent if they are both in I_2 . $\square$

To illustrate how the choice of the color of $v = \pi(j)$ depends on the components of G_j^k we apply the algorithm to the following example.

Example. Let G be a $K_{4,4}$ without a perfect matching, i.e., with $V(G) = \{1, 2, 3, 4, 5, 6, 7, 8\}$, bipartition in $\{1, 3, 5, 7\}$ and $\{2, 4, 6, 8\}$, and only edges $\{1, 2\}$, $\{3, 4\}$, $\{5, 6\}$, and $\{7, 8\}$ omitted. If the vertices are revealed in the order of increasing numbers, the algorithm assigns colors $a_1, a_1, b_1, b_1, a_2, b_2, a_2, b_2$, respectively. The color b_2 is assigned to vertex 8 for the following three reasons. First, a_1 is mixed on C_8^1 , i.e., the subgraph of G induced by $\{1, 2, 3, 4, 8\}$. Second, a_2 is not mixed on $C_8^2 = G$. Third, a_2 has already been assigned to a vertex in $\{1, 3, 5, 7\}$, i.e., the bipartition class of C_8^2 not containing the vertex 8.

Suppose that G is extended and a new vertex 9 is revealed to the algorithm after presenting vertices $1, \dots, 8$. Suppose 9 is only adjacent to 7. Then a_1 is not mixed on $C_9^1 = (\{9\}, \emptyset)$ and a_2 is not mixed on $C_9^2 = G$. Hence, *BicolorMax* colors 9 with a_1 . It is left to the reader to check that 9 is assigned color b_1 if 9 is adjacent to 1 and 7, color b_2 if 9 is adjacent to 1, 3, and 7, and color a_2 if 9 is adjacent to 2, 4, and 6.

For a $K_{n,n}$ without a perfect matching with $n \geq 5$, the algorithm will continue assigning a_2 and b_2 if the vertices are presented in an order alternating between the two classes of the bipartition, as in the previous example for $n = 4$. In contrast, recall that *FF* uses n colors in this case.

Remark. In [18], the authors leave it as an “easy exercise” to construct an on-line coloring algorithm A such that $\chi_A(G) = O(\log_2 n)$ on the class of bipartite graphs (where G is a bipartite graph on n vertices). There are several choices for A to solve this exercise. One of these choices is the algorithm *BicolorMax* that we presented previously and that will be analyzed in this paper. We are not aware of any publications that contain this algorithm *BicolorMax* (except for [3], of course). Clearly, our main purpose in this paper is to show that *BicolorMax* works particularly well in the sense of on-line competitiveness for P_7 -free bipartite graphs. As can be concluded from the next section this is not an easy exercise.

One might be inclined to think that *BicolorMax* coincides with other on-line algorithms that have been proposed to solve the “easy exercise” from [18], e.g., the algorithm from the survey paper of Kierstead (see [14, Theorem 2, p. 286]) that is also known under the acronym *BFF* (Bipartite First-Fit). The formulation of *BFF* is very simple: when vertex v_i is presented, there is a unique partition (I_1, I_2) of the component to which v_i belongs such that $v_i \in I_1$. The algorithm *BFF* assigns to v_i the smallest color (in some ordering of the colors) that has not already been assigned to a vertex of I_2 . We do something more advanced, as can be seen from the above example. In this example, as we have argued, if the vertices 1 to 8 are presented in the order of increasing numbers, the algorithm *BicolorMax* assigns colors $a_1, a_1, b_1, b_1, a_2, b_2, a_2, b_2$, respectively. It can be checked that *BFF* also assigns these colors in the same order. Now suppose a new vertex 9 is presented which is only adjacent to vertex 7. Then *BicolorMax* assigns color a_1 to 9, as explained in the example, whereas *BFF* assigns color b_2 to 9.

Although at first sight *BicolorMax* looks very similar to *BFF*, the subtle difference in the choice of colors to avoid mixing colors enables us to prove the on-line competitiveness of our algorithm for P_7 -free bipartite graphs. Our earlier attempts to use a simpler strategy, like in *BFF*, for such a competitiveness proof failed even for the more restricted class of P_6 -free bipartite graphs. In fact, finding a suitable variant for our competitiveness proof turned out to be rather tricky.

6. P_7 -free bipartite graphs. In this section we prove our main theorem showing that *BicolorMax* is a linear on-line competitive algorithm for the class of P_7 -free bipartite graphs. The proof is modelled along the following lines. We will define a class of tree-like bipartite graphs, show that the on-line competitive factor of any on-line coloring algorithm for these graphs is high, and show that whenever *BicolorMax* uses many colors on a P_7 -free bipartite input graph G , then G contains a large member of the class of tree-like graphs as an induced subgraph. More specifically, denote by $\chi_{BM}(G)$ the maximum number of colors used by *BicolorMax* for coloring a graph G .

By a series of lemmas and propositions we prove that $\chi_{BM}(G) \leq 8\chi_{OL}(G) + 8$ for any P_7 -free bipartite graph G . In the next two lemmas we first show that *BicolorMax* controls the mixing of color pairs. We note that these lemmas are valid for any bipartite graph.

LEMMA 6.1. *Let G be a bipartite graph. Let *BicolorMax* color vertex $v = \pi(j)$ with a_m or b_m , $m \geq 2$. Let (x, y) be a k -mixed pair in C_j^k with $k \leq m - 1$. Then any path between x and y in C_j^k must pass through v .*

Proof. Let $x = \pi(r)$, and let $y = \pi(s)$ for some $r, s \leq j - 1$. We assume without loss of generality that y has been presented to *BicolorMax* after x , i.e., $s > r$. Suppose x belongs to C_s^k , implying that $a_k \in H(V(C_s^k))$. Since y is colored with a_k , color a_k is not mixed on C_s^k . Since (x, y) is a k -mixed pair, x (together with all the other vertices in C_s^k that have color a_k) and y are in different bipartite classes. So *BicolorMax* would have colored y with color b_k . Hence, x does not belong to C_s^k .

Suppose there exists an index i with $s < i < j$ such that x and y belong to C_i^k . This means that a_k is mixed on C_i^k . Then *BicolorMax* would never use a color a_h or b_h with $h \leq k$ to color $\pi(i)$. We conclude that every path between x and y in C_j^k must pass through v . $\square$

LEMMA 6.2. *Let G be a bipartite graph. Let *BicolorMax* color vertex $v = \pi(j)$ with a_m or b_m , $m \geq 2$. For some $k \leq m - 1$ let x and y be two vertices in C_j^k colored with a_k and b_k , respectively, such that $d(x, y, C_j^k)$ is even. Then any path between x and y in C_j^k must pass through v .*

Proof. Let $x = \pi(r)$, and let $y = \pi(s)$ for some $r, s \leq j - 1$. Suppose x belongs to C_s^k implying that y appeared after x . Since y is colored with b_k , color a_k is not mixed on C_s^k . Because G is bipartite and the distance $d(x, y, C_j^k)$ is even, the distance between x (or any other vertex in C_s^k with color a_k) and y in C_s^k is also even. Then *BicolorMax* would have colored y with color a_k . Hence, x does not belong to C_s^k .

Suppose y belongs to C_r^k . Since x is colored with a_k , color a_k is not mixed on C_r^k . Then the distance between x and any vertex in C_r^k with color a_k is even. Since y has received color b_k , by definition of *BicolorMax*, there exists at least one vertex z in C_r^k with color a_k , which is in a different bipartite class than y . Since the distance between x and y in C_j^k , and consequently in C_r^k , is even, this implies that the distance between x and z in C_r^k is odd. This is a contradiction. Hence, y does not belong to C_r^k .

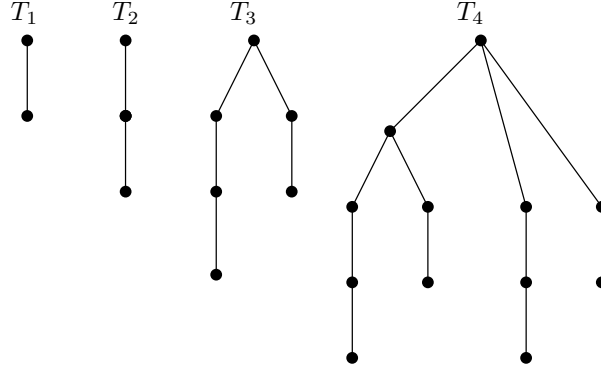
In the remaining case there exists an index i with $\max(\{r, s\}) < i < j$ such that x and y belong to C_i^k . Since y is colored with b_k , by definition of *BicolorMax*, the component C_s^k , which is a subgraph of C_i^k , must contain a vertex w with color a_k and with odd distance to y . Since the distance between x and y is even, the distance between w and x is odd, and we find that a_k is mixed on C_i^k . Then *BicolorMax* would never use a color a_h or b_h with $h \leq k$ to color $\pi(i)$. We conclude that every path between x and y in C_j^k must pass through v . $\square$

In the following lemma we show that *BicolorMax* ensures that vertices of higher colors are very “close” to at least some vertices of lower colors.

LEMMA 6.3. *Let G be a P_7 -free bipartite graph. Let *BicolorMax* color vertex $v = \pi(j)$ with a_m or b_m , $m \geq 3$. There exists an $(m - 1)$ -mixed pair (z^*, z) in C_j^{m-1} with $d(v, z^*, C_j^{m-1}) = 1$ and $d(v, z, C_j^{m-1}) = 2$.*

Proof. By definition of *BicolorMax*, color a_{m-1} is mixed on C_j^{m-1} . So C_j^{m-1} contains at least one $(m - 1)$ -mixed pair (z, z') . We assume without loss of generality that the distance between z and v is even and the distance between z' and v is odd. Since (z, z') is an $(m - 1)$ -mixed pair, by Lemma 6.1, a shortest path $P_{zz'}$ from z to z' in C_j^{m-1} must be formed by joining shortest paths P_{zv} from z to v and $P_{vz'}$ from v to z' . First, we show that $d(v, z, C_j^{m-1}) = 2$.

Suppose $d(v, z, C_j^{m-1}) \geq 4$. Then $d(v, z', C_j^{m-1}) = 1$; i.e., z' and v are adjacent. Otherwise, since z' has odd distance to v , the path $P_{zz'}$ contains an induced P_7 . Let $z' = \pi(s)$ for some $s < j$. Since *BicolorMax* has used color $a_{m-1} \neq a_1$ (due to our assumption that $m \geq 3$) on vertex z' , the component C_s^{m-2} contains an $(m - 2)$ -mixed

FIG. 6.1. The trees T_1, T_2, T_3, T_4 .

pair. This means that z' has a neighbor $w \neq v$ in $C_s^{m-2} \subseteq C_j^{m-1}$. Bipartiteness of G and Lemma 6.1 imply that w is not adjacent to any vertex in P_{zv} . This implies that G contains an induced P_7 , which is a contradiction. Hence, $d(v, z, C_j^{m-1}) = 2$.

We now show that $d(v, z', C_j^{m-1}) \leq 3$. Suppose $d(v, z', C_j^{m-1}) \geq 5$. Then $P_{zz'}$ would contain an induced P_7 . Hence, $d(v, z', C_j^{m-1}) = 1$ or $d(v, z', C_j^{m-1}) = 3$. In the first case we are done. In the second case, i.e., if $d(v, z', C_j^{m-1}) = 3$, we will show that there exists a vertex z^* on $P_{z'v}$ with color a_{m-1} that is adjacent to v .

Let y and z^* be vertices of G such that $P_{z'v} = z'y z^* v$. Note that the distance between z^* and z' is even. Then, due to Lemma 6.2, vertex z^* has not been colored with b_{m-1} . We will show that z^* has not been colored with any color from $A(m-2) \cup B(m-2)$ either. First, note that any neighbor of z' in C_j^{m-1} is adjacent to z^* . Otherwise, we could extend the path $P_{zz'}$ on six vertices with one extra vertex, and C_j^{m-1} would contain an induced P_7 .

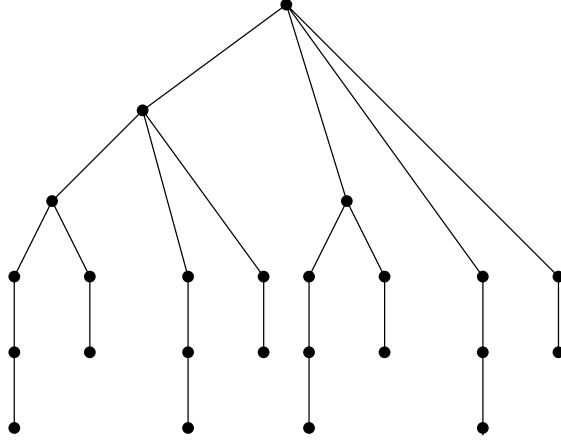
Let $r < j$ be chosen such that $z^* = \pi(r)$. Recall that $z' = \pi(s)$. We first consider the case $s > r$, i.e., vertex z' has appeared after z^* . Since z^* is adjacent to every neighbor of z' in $C_s^{m-2} \subseteq C_j^{m-1}$ and a_{m-2} is mixed on C_s^{m-2} , Lemma 6.1 prevents z^* from being in C_s^{m-2} . Otherwise, C_s^{m-2} would contain a path (using z^*) between two vertices u_1 and u_2 of an $(m-2)$ -mixed pair (u_1, u_2) in C_s^{m-2} not going through z' . We already noted that z^* has not received color b_{m-1} . Then, since z^* is in C_j^{m-1} , vertex z^* must have been colored with a_{m-1} .

Now assume $s < r$, i.e., vertex z^* has appeared after z' . Every neighbor of z' in C_r^{m-2} is adjacent to z^* . Hence, a_{m-2} is mixed not only on C_s^{m-2} but also on C_r^{m-2} . Since b_{m-1} was not allowed and z^* is in C_j^{m-1} , *BicolorMax* must have colored z^* with a_{m-1} . $\square$

We inductively define a class of trees (see Figures 6.1 and 6.2). Each tree T_k of the class has a root vertex $r(T_k)$, and the following hold:

- T_1 is a tree consisting of an edge, one of whose end vertices is the root vertex $r(T_1)$.
- T_2 is a path on three vertices, one of whose end vertices is the root vertex $r(T_2)$.
- T_k , $k \geq 3$, consists of a root vertex $r(T_k)$ that is adjacent to the root vertices of mutually disjoint copies of $T_1, T_2, \dots, T_{k-1}$ (one copy of each of these trees). These copies are then called the *child trees* of T_k .

We denote a copy of a tree T_k with root vertex $v = r(T_k)$ by $T_k(v)$. The child trees of

FIG. 6.2. The tree T_5 .

$T_k(v)$ are denoted by $T_1^v, T_2^v, \dots, T_{k-1}^v$. The following lemma turns out to be useful.

LEMMA 6.4. *Let G be a P_7 -free bipartite graph. If $BicolorMax$ uses color a_k or b_k on vertex $v = \pi(j)$ with $k \geq 2$, then C_j^{k-1} contains the tree $T_{k-1}(v)$ as a (not necessarily induced) subgraph in such a way that the following hold:*

- (i) *If there exists an edge in G between any two vertices x, y in $T_{k-1}(v)$ with $d(v, x, T_{k-1}(v)) \leq d(v, y, T_{k-1}(v))$, then x lies on the path from y to v in $T_{k-1}(v)$.*
- (ii) *The root of child tree T_i^v is colored with a_{i+1} or b_{i+1} for all $1 \leq i \leq k-2$.*

Proof. The proof proceeds by induction on k . Let $k = 2$; i.e., $BicolorMax$ uses color a_2 or b_2 on vertex v . Then C_j^1 contains a 1-mixed pair. This implies that v has a neighbor in C_j^1 , and the conditions of the lemma are trivially satisfied.

Let $k = 3$; i.e., $BicolorMax$ uses color a_3 or b_3 on vertex v . By Lemma 6.3, vertex v has a neighbor z^* in C_j^2 with color a_2 . Let $z^* = \pi(q)$ for some $q < j$. Then C_q^1 contains a 1-mixed pair. This implies that z^* has a neighbor not equal to v in C_j^2 . We conclude that the conditions of the lemma are satisfied.

Let $k \geq 4$. Since $BicolorMax$ uses color a_k or b_k on vertex v , there exists a $(k-1)$ -mixed pair (x, y) in C_j^{k-1} . By Lemma 6.3, we may without loss of generality assume that $d(v, x, C_j^{k-1}) = 2$ and $d(v, y, C_j^{k-1}) = 1$. Assume $x = \pi(h)$ for some $h < j$ and $y = \pi(i)$ for some $i < j$. By the induction hypothesis, C_h^{k-2} contains the tree $T_{k-2}(x)$, and C_i^{k-2} contains the tree $T_{k-2}(y)$. Since (x, y) is a $(k-1)$ -mixed pair in C_j^{k-1} , every path from x to y in C_j^{k-1} must go through v due to Lemma 6.1. This implies that every path in C_j^{k-1} from a vertex in $C_h^{k-2} \subseteq C_j^{k-1}$ to a vertex in $C_i^{k-2} \subseteq C_j^{k-1}$ must go through v . Then we have also found that every path in C_j^{k-1} from a vertex in $T_{k-2}(x) \subseteq C_h^{k-2}$ to a vertex in $T_{k-2}(y) \subseteq C_i^{k-2}$ must go through v . We distinguish two cases: either C_h^{k-2} contains a common neighbor of v and x , or C_h^{k-2} does not contain any common neighbors of v and x .

Case 1. Component C_h^{k-2} contains a common neighbor w of x and v . Again we need to distinguish two cases: either w is in $T_{k-2}(x)$, or w is not in $T_{k-2}(x)$.

Case 1a. Vertex w is in $T_{k-2}(x)$. Then w is in a child tree T_p^x of $T_{k-2}(x)$ for some $1 \leq p \leq k-3$. If $k = 4$, then $T_{k-2}(x) = T_2(x)$, and w must be the root of T_1^x . By the induction hypothesis w is colored with a_2 or b_2 . Recall that every path in C_j^3 from a vertex in $T_2(x)$ to a vertex in $T_2(y)$ goes through v , and that y is colored with

$a_{k-1} = a_3$. This implies that v is the root of a copy of T_3 satisfying (i) and (ii).

Suppose $k \geq 5$. Then $T_{k-2}(x)$ has a child tree T_q^x with root u for some $q \neq p$. Note that for all $k \geq 2$ any child tree of a tree T_k consists of at least two vertices. Let u' be a neighbor of u in T_q^x . Let y' be a neighbor of y in $T_{k-2}(y)$. Note that u', u, x, w, v, y, y' are seven different vertices of G . This implies that $P_{u'y'} = u'uxwvyy'$ is a path on seven vertices in C_j^{k-1} . Recall that every path from a vertex of $T_{k-2}(x)$ to a vertex of $T_{k-2}(y)$ goes through v . Then there are no edges between $\{u', u, x, w\}$ and $\{y, y'\}$. By the induction hypothesis, T_p^x and T_q^x are remote. This implies that w is adjacent neither to u nor to u' . Then there must be an edge between u and v ; otherwise $P_{u'y'}$ is an induced P_7 in G . Hence, we have found that v is adjacent to the root of all child trees of $T_{k-2}(x)$ that are not equal to T_p^x . However, the root of T_p^x must also be joined to v by an edge. This can be shown by using exactly the same arguments (in which vertex u takes over the role of vertex w).

From the above we conclude that v is adjacent to the roots of all child trees of $T_{k-2}(x)$. These trees together with tree $T_{k-2}(y)$ form the child trees of $T_{k-1}(v)$. Due to the fact that any path from a vertex in $T_{k-2}(x)$ to a vertex in $T_{k-2}(y)$ goes through v and our induction hypothesis, the child trees of $T_{k-1}(v)$ satisfy (i). Recall that the root vertex y of $T_{k-2}^v = T_{k-2}(y)$ is colored with a_{k-1} . The root vertices of the other child trees of $T_{k-1}(v)$ are colored with the desired colors due to the induction hypothesis. Hence, condition (ii) of the lemma is also satisfied.

Case 1b. Vertex w is not in $T_{k-2}(x)$. Since $k \geq 4$, there exists a vertex s with color a_{k-2} in C_h^{k-2} with $d(x, s, C_h^{k-2}) = 2$ due to Lemma 6.1. Since $d(x, w, C_h^{k-2}) = 1$, vertex s is not equal to vertex w . Again, let y' be a neighbor of y in $T_{k-2}(y)$. We first show that we may without loss of generality assume that w is adjacent to s .

Suppose w is not adjacent to s . Since $d(x, s, C_h^{k-2}) = 2$, vertices s and x have a common neighbor t in C_h^{k-2} . Note that s, t, x, w, v, y, y' are seven different vertices of G . This implies that $P_{sy'} = stxwvyy'$ is a path on seven vertices in C_j^{k-1} . Recall that every path from a vertex in C_h^{k-2} to a vertex in $T_{k-2}(y)$ goes through v . Then there are no edges between $\{s, t, x, w\}$ and $\{y, y'\}$. This together with $\{w, s\} \notin E(G)$ implies that v must be adjacent to t ; otherwise $P_{sy'}$ is an induced P_7 in G . Then, we can pick vertex t instead of vertex w . So from now on we assume that w is a common neighbor of s and x in C_h^{k-2} .

Let r be the root of child tree T_{k-3}^x of $T_{k-2}(x)$. Let r' be a neighbor of r in T_{k-3}^x . Note that r', r, x, w, v, y, y' are seven different vertices of G . This implies that $P_{r'y'} = r'rxwvyy'$ is a path on seven vertices in C_j^{k-1} . By the induction hypothesis, r is colored with a_{k-2} or b_{k-2} .

Suppose r is colored with a_{k-2} . Since $d(r, s, C_h^{k-2})$ is odd and s is colored with a_{k-2} , the pair (r, s) is a $(k-2)$ -mixed pair in C_h^{k-2} . Lemma 6.1 implies that every path from s in C_h^{k-2} , and hence every path from w in C_h^{k-2} , to a vertex in $T_{k-3}^x \subseteq C_h^{k-2}$ goes through x . Then there are no edges between $\{r, r'\}$ and w . Furthermore, recall that every path in C_j^{k-1} from a vertex in C_h^{k-2} to a vertex in $T_{k-2}(y) \subseteq C_i^{k-2}$ goes through v . Then there are no edges between $\{r', r, x, w\}$ and $\{y, y'\}$ either. Since $P_{r'y'}$ may not be an induced P_7 , these observations imply that there must be an edge between v and r . Hence, using vertex r instead of w brings us back to Case 1a.

Suppose r is colored with b_{k-2} . By Lemma 6.3, C_h^{k-2} contains a vertex $\hat{s}$ that has received color a_{k-2} and is adjacent to x . Let $\hat{s} = \pi(\ell)$ for some $\ell < j$. By our induction hypothesis, vertex $\hat{s}$ is the root of a tree $T_{k-3}(\hat{s})$ in C_ℓ^{k-3} .

Suppose every path in C_h^{k-2} from $\hat{s}$ to a vertex in any child tree T_p^x for $1 \leq p \leq k-4$ goes through x . Then in $T_{k-2}(x)$ we can replace the child tree $T_{k-3}(r)$ by the

child tree $T_{k-3}(\hat{s})$. Then we can repeat the argument above (replace r by $\hat{s}$, which is colored by a_{k-2}) in order to find that v is adjacent to $\hat{s}$, and we return to Case 1a.

Suppose C_h^{k-2} contains a path from $\hat{s}$ to a child tree T_m^x for some $1 \leq m \leq k-4$ that does not go through x . Let z be the root of T_m^x , and let z' be a neighbor of z in T_m^x . Note that z', z, x, w, v, y, y' are seven different vertices of G . This implies that $P_{z'y'} = z'zxwvyy'$ is a path on seven vertices in C_j^{k-1} . Recall that w is a neighbor of s . We then find that z' and w are not adjacent. Otherwise, C_h^{k-2} would contain a path between s and $\hat{s}$, which are the vertices of a $(k-2)$ -mixed pair, not going through x (a contradiction due to Lemma 6.1). Then as before, in order not to have an induced P_7 , vertex v must be adjacent to z . We return to Case 1a.

Case 2. Component C_h^{k-2} does not contain a common neighbor of x and v . Since x has distance 2 from v in C_j^{k-1} , there exists a common neighbor v' of v and x in C_j^{k-1} . We first prove that v' has received color b_{k-1} .

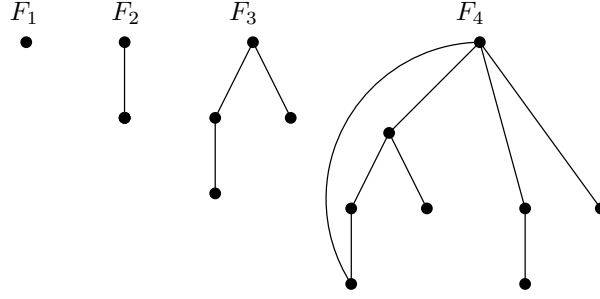
Since $k \geq 4$, Lemma 6.3 implies that C_h^{k-2} contains a vertex s with color a_{k-2} at distance $d(x, s, C_h^{k-2}) = 2$ from x , and C_h^{k-2} contains a vertex t with color a_{k-2} that is a neighbor of x . Since $d(x, s, C_h^{k-2}) = 2$, vertices s and x have a common neighbor s' in C_h^{k-2} . Since $k \geq 4$, vertex t has a neighbor t' in C_h^{k-2} that is not equal to x . Let y' be a neighbor of y in $T_{k-2}(y)$. Note that $s, s', t, t', x, v', v, y, y'$ are nine different vertices in G . This implies that both $P_{sy'} = ss'xv'vy'y'$ and $P_{t'y'} = t'txv'vy'y'$ are paths on seven vertices in C_j^{k-1} . Any path in C_j^{k-1} from a vertex in $T_{k-2}(y)$ to a vertex in C_h^{k-2} or to v' goes through v . Otherwise there exists a path in C_j^{k-1} from y to x that does not use v , which is not possible due to Lemma 6.1. Hence, there are no edges between $\{s, s', x, v'\}$ and $\{y, y'\}$, and there are no edges between $\{t, t', x, v'\}$ and $\{y, y'\}$ either. Since we assumed that C_h^{k-2} does not contain any common neighbor of v and x , there are no edges between v and $\{s', t'\}$. Then v' must be adjacent to both s and t' ; otherwise, G contains an induced P_7 . Since v' is in C_j^{k-1} and adjacent to vertex x with color a_{k-1} , the color of v' is in $A(k-2) \cup B(k-1)$.

Suppose v' has not received color b_{k-1} but some color from $A(k-2) \cup B(k-2)$. Then v' must have appeared after x , due to our assumption that v' is not in C_h^{k-2} . However, in that case, v' has also appeared after s and t . Then (s, t) is a $(k-2)$ -mixed pair of $C_{\pi^{-1}(v')}^{k-2}$ implying that *BicolorMax* would never color v' with a color from $A(k-2) \cup B(k-2)$. Hence, v' must have received color b_{k-1} .

Recall that $y = \pi(i)$ is assigned color a_{k-1} . Then, by Lemma 6.3, component C_i^{k-2} contains a $(k-2)$ -mixed pair (z^*, z) such that $d(y, z^*, C_i^{k-2}) = 1$ and $d(y, z, C_i^{k-2}) = 2$. We first show that v is adjacent to z and every neighbor of z^* in C_i^{k-2} .

Since $d(y, z, C_i^{k-2}) = 2$, component C_i^{k-2} contains a common neighbor z' of y and z . Let x' be a neighbor of x in C_h^{k-2} . Since v' with color b_{k-1} and neighbor v with color a_{k-1} are neither in C_h^{k-2} nor in C_i^{k-2} , we find that z, z', y, v, v', x, x' are seven different vertices. This implies that $P_{zx'} = zz'yvv'xx'$ is a path on seven vertices in C_j^{k-1} . By Lemma 6.1, any path from y to x in C_j^{k-1} must go through v . This implies that there are no edges between vertices from $\{z, z', y\}$ and $\{v', x, x'\}$. Since we assume that v and x do not have a common neighbor in C_h^{k-2} , vertices x' and v are not adjacent. Then v must be adjacent to z . By the same arguments we find that v is adjacent to every neighbor of z^* in C_i^{k-2} .

Let $z^* = \pi(\ell^*)$ for some $\ell^* < j$. By our induction hypothesis, $C_{\ell^*}^{k-3}$ contains the tree $T_{k-3}(z^*)$ satisfying conditions (i) and (ii) of the lemma. Since v is adjacent to every neighbor of z^* in $C_i^{k-2} \supseteq C_{\ell^*}^{k-3}$, vertex v is adjacent to the roots of all child trees $T_{k-4}^{z^*}, T_{k-5}^{z^*}, \dots, T_1^{z^*}$.

FIG. 6.3. The graphs F_1, F_2, F_3, F_4 .

We are also going to use our induction hypothesis with respect to vertex z , which has been assigned color a_{k-2} . Let $z = \pi(\ell)$ for some $\ell < j$. Then C_ℓ^{k-3} contains the tree $T_{k-3}(z)$ satisfying conditions (i) and (ii) of the lemma. Recall that z is adjacent to v . So v is adjacent to the root of a copy of a tree T_{k-3} . Due to Lemma 6.1, the trees $T_{k-3}(z)$ and $T_{k-3}(z^*)$ do not have a vertex in common.

Finally, we consider vertex v' with color b_{k-1} . Let $v' = \pi(\ell')$ for some $\ell' < j$. Again by our induction hypothesis, $C_{\ell'}^{k-2}$ contains the tree $T_{k-2}(v')$ satisfying conditions (i) and (ii) of the lemma. Recall that v is adjacent to v' . So v is adjacent to a copy of T_{k-2} that does not have any vertex in common with the trees $T_{k-3}(z)$ and $T_{k-3}(z^*)$ due to Lemma 6.1.

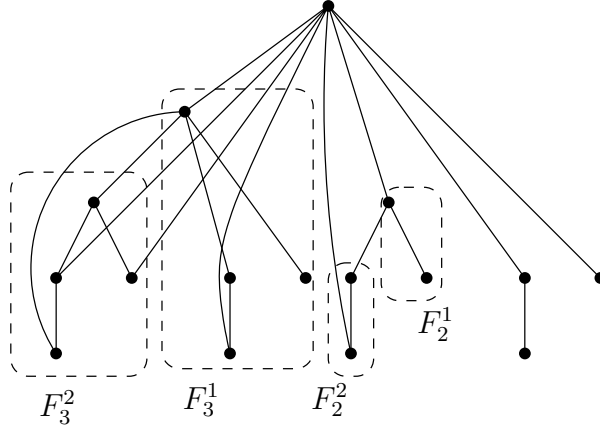
We conclude that C_j^{k-1} contains the tree $T_{k-1}(v)$ as a subgraph such that condition (ii) has been satisfied and that condition (i) will be satisfied as well if we can show the following statement: there are no edges between $T_{k-3}^v = T_{k-3}(z)$ and a child tree $T_i^v = T_i^{z^*}$ for all $1 \leq i \leq k-4$, and there are no edges between $T_{k-2}^v = T_{k-2}(v')$ and a child tree T_i^v for all $1 \leq i \leq k-3$. This claim can be seen as follows. If there is an edge between T_{k-2}^v and a child tree T_i^v for $1 \leq i \leq k-3$, then C_j^{k-1} contains a path from x to y that does not go through v . Since (x, y) is a $(k-1)$ -mixed pair in C_j^{k-1} , this is not possible due to Lemma 6.1. If there is an edge between T_{k-3}^v and a child tree T_i^v for $1 \leq i \leq k-4$, then C_i^{k-2} contains a path from z to z^* that does not go through y . Since (z^*, z) is a $(k-2)$ -mixed pair in C_i^{k-2} , this is not possible, again due to Lemma 6.1. $\square$

We also inductively define a class of P_6 -free bipartite graphs F_i (see Figures 6.3 and 6.4). Its purpose and its relation to the class of trees T_i will be made clear later on. Each graph F_i of the class has a root vertex $r(F_i)$, and the following hold:

- F_1 is a graph consisting of a single root vertex.
- F_2 is a graph consisting of an edge, one of whose end vertices is the root vertex.
- F_k , $k \geq 3$, consists of a root vertex $r(F_k)$ that is adjacent to the root vertices of disjoint copies of $F_1, F_2, \dots, F_{k-1}$ (one copy of each of these graphs). These copies are then called the *child graphs* of F_k . For all $1 \leq j \leq k-1$, we join vertex $r(F_k)$ also to every vertex in F_j that has distance 2 to $r(F_j)$. This implies that every vertex of F_k is at distance at most 2 from $r(F_k)$. Hence, the maximum distance between two vertices in F_k is at most 4, and F_k is P_6 -free.

A graph F_k has the following useful properties (see also Figure 6.4).

LEMMA 6.5. *Any graph F_k with $k \geq 4$ contains copies F_t^1 and F_t^2 of F_t for all*

FIG. 6.4. The graph F_5 which contains copies $F_2^1, F_2^2, F_3^1, F_3^2$.

$2 \leq t \leq k-2$ such that the following hold:

- (i) For all $1 \leq i, j \leq 2$ and all $2 \leq s < t \leq k-2$, the graphs F_s^i and F_t^j are remote.
- (ii) For all $2 \leq t \leq k-2$, the vertices of the graph F_t^1 in the bipartite class containing the root vertex of F_t^1 are adjacent to $r(F_k)$.
- (iii) For all $2 \leq t \leq k-2$, the vertices of the graph F_t^2 in the bipartite class not containing the root vertex of F_t^2 are adjacent to $r(F_k)$.

Proof. One easily checks that child graph F_{t+1} of graph F_k contains the desired copies F_t^1 and F_t^2 for $2 \leq t \leq k-2$. $\square$

The following proposition is a key ingredient for the proof of our main result. It also implies that there exists no competitive on-line coloring algorithm for the family of P_6 -free bipartite graphs (cf. [9], where this has been proven using a different subfamily of P_6 -free bipartite graphs).

PROPOSITION 6.6. For any $k \geq 1$, $\chi_{OL}(F_{2k}) \geq k$.

Proof. By induction on k . The case $k = 1$ is trivial. Let $k \geq 2$. Consider F_{2k} and an on-line algorithm A for coloring F_{2k} . The first time the i th color is used by A we identify it as color i . We choose an ordering on $V(F_{2k})$ such that the vertices of remote copies of $F_2, F_4, \dots, F_{2k-2}$ are presented until color i is used on F_{2i} ($i = 1, \dots, k-1$); i.e., as soon as color 1 is used on F_2 we start presenting vertices of F_4 , as soon as color 2 is used on F_4 we start presenting vertices of F_6 and so on, until color $k-1$ is used on F_{2k-2} . By the adjacency relations from the definition of F_{2k} and the properties of Lemma 6.5, the ordering of the presented vertices of $F_2, F_4, \dots, F_{2k-2}$ can be chosen in such a way that $r(F_{2k})$ is adjacent to the (not necessarily root) vertices that received colors $1, \dots, k-1$. Hence, a new color k is forced upon A . $\square$

Below we denote a copy of a graph F_k with root vertex v by $F_k(v)$. The child graphs of $F_k(v)$ are denoted by $F_1^v, F_2^v, \dots, F_{k-1}^v$.

LEMMA 6.7. Let G be a P_7 -free bipartite graph. If *BicolorMax* uses color a_k or b_k with $k \geq 3$ on vertex $v = \pi(j)$, then C_j^{k-1} contains the graph $F_{\lfloor \frac{k-1}{2} \rfloor}(v)$ as an induced subgraph.

Proof. Due to Lemma 6.4 the component C_j^{k-1} contains the tree $T_{k-1}(v)$ as a subgraph in such a way that the following hold:

- (i) If there exists an edge in G between any two vertices x, y in $T_{k-1}(v)$ with $d(v, x, T_{k-1}(v)) \leq d(v, y, T_{k-1}(v))$, then x lies on the path from y to v in

$T_{k-1}(v)$.

(ii) The root of child tree T_i^v is colored with a_{i+1} or b_{i+1} for all $1 \leq i \leq k-2$.

By induction on k we will show that the subgraph of G induced by $V(T_{k-1}(v))$ contains the graph $F_{\lfloor \frac{k-1}{2} \rfloor}(v)$ as an induced subgraph. The case $k = 3$ is trivial. Let $k \geq 4$.

Claim. The following is true for at least $k-3$ child trees T_i^v : vertex v is adjacent to every vertex in T_i^v that has distance 2 to $r(T_i^v)$.

We prove this claim by contradiction. Suppose there exist two child trees of $T_{k-1}(v)$ not satisfying the claim. Let T_h^v be a child tree of $T_{k-1}(v)$ for some $1 \leq h \leq k-2$ that contains a vertex x with $d(x, r(T_h^v), T_h^v) = 2$ and with $\{x, v\} \notin E(G)$. Let T_j^v be a child tree of $T_{k-1}(v)$ for some $1 \leq j \neq h \leq k-2$ that contains a vertex y with $d(y, r(T_j^v), T_j^v) = 2$ and with $\{y, v\} \notin E(G)$. Let y' be the common neighbor of y and $r(T_j^v)$ in T_j^v . Let x' be the common neighbor of x and $r(T_h^v)$ in T_h^v . Since $T_{k-1}(v)$ satisfies (i), G contains an induced $P_7 = xx'r(T_h^v)vr(T_j^v)y'y$, which is a contradiction. Hence, the above claim has been proven.

By the induction hypothesis and condition (ii), for $2 \leq i \leq k-2$, the subgraph induced by $V(T_i^v)$ contains the graph $F_{\lfloor \frac{i}{2} \rfloor}(r(T_i^v))$ as an induced subgraph. This, together with the above claim and the fact that $T_{k-1}(v)$ satisfies condition (i), implies that C_j^{k-1} contains the graph $F_{\lfloor \frac{k-1}{2} \rfloor}(v)$ as an induced subgraph. $\square$

THEOREM 6.8. *If G is a P_7 -free bipartite graph, then $\chi_{BM}(G) \leq 8\chi_{OL}(G) + 8$.*

Proof. Let k be the highest index such that *BicolorMax* uses color a_{4k+1} on a vertex in G . Note that it is possible that *BicolorMax* uses colors $a_{4k+2}, a_{4k+3}, a_{4k+4}$ or $b_{4k+2}, b_{4k+3}, b_{4k+4}$ to color G . Hence, every color used on a vertex of G is from $A(4k+4) \cup B(4k+4)$. Since *BicolorMax* uses b_i only if a_i has been used before, then $\chi_{B<}(G) \leq 2(4k+4) = 8k+8$. For $k = 0$ the statement obviously holds. Suppose $k \geq 1$. Due to Lemma 6.7, G contains a copy of F_{2k} as an induced subgraph. Proposition 6.6 implies that $\chi_{OL}(G) \geq \chi_{OL}(F_{2k}) \geq k$. $\square$

7. New upper bounds on χ_{OL} for bipartite graphs. As we noted in section 5, in [18], the authors leave it as an “easy exercise” to construct an on-line coloring algorithm A such that $\chi_A(G) = O(\log_2 n)$ on the class of bipartite graphs (where G is a bipartite graph on n vertices). There are several choices for A to solve this exercise. One of these choices is the algorithm *BicolorMax* that we presented and analyzed in this paper.

As a byproduct, using the derived properties of *BicolorMax*, below we give an upper bound on $\chi_A(G)$ for a bipartite graph G in terms of the independence number $\alpha(G)$ and in terms of the number of remote induced P_5 's in G . It is not possible to prove an upper bound in terms of induced subgraphs isomorphic to P_6 , since it has been proven in [9] (and also follows from Proposition 6.6) that no competitive algorithm exists for the family of P_6 -free bipartite graphs.

THEOREM 7.1. *Let G be a bipartite graph in which each component has at most s remote induced subgraphs isomorphic to P_5 . If $s = 0$, then $\chi_{BM}(G) \leq 4$. If $s > 0$, then $\chi_{BM}(G) \leq 2\log_2(s) + 6$.*

Proof. We prove the theorem by showing that a component C of G contains at least 2^{k-3} remote induced subgraphs isomorphic to P_5 if *BicolorMax* uses color a_k on C with $k \geq 3$. We use induction on k .

Let $k = 3$. It is easy to check that a component of G contains an induced P_5 if *BicolorMax* uses color a_3 on a vertex of that component. Let $k \geq 4$. Suppose $v = \pi(j)$ is colored by a_k or b_k . Then there exists a $(k-1)$ -mixed pair (x, y) in C_j^{k-1} . By Lemma 6.1, x and y belong to two different components in $G_j^{k-1} - v$, both

containing 2^{k-4} remote induced subgraphs isomorphic to P_5 . $\square$

Our next result gives an upper bound on the number of colors used by *BicolorMax* on a bipartite graph G in terms of its independence number $\alpha(G)$ (i.e., the largest number of vertices of G that are mutually nonadjacent in G).

THEOREM 7.2. *Let G be a bipartite graph with independence number α . Then*

$$\chi_{BM}(G) \leq \begin{cases} \alpha + 1 & \text{if } 1 \leq \alpha \leq 3, \\ 4 & \text{if } \alpha = 4, \\ 2p & \text{if } 5 \cdot 2^{p-3} \leq \alpha \leq 5 \cdot 2^{p-2} - 1 \text{ for some } p \geq 3. \end{cases}$$

Proof. Let k be the highest index such that *BicolorMax* uses color a_k on a vertex v in G . If $\alpha = 1$, then $k = 1$; hence, $\chi_{BM}(G) \leq 2$. If $\alpha = 2$, then $k \leq 2$ but color b_2 is never used; hence, $\chi_{BM}(G) \leq 3$. Now assume that $\alpha \geq 3$.

Let $v = \pi(j)$. We write $C_j^{k-1} = (I_1^k, I_2^k)$ for the bipartition with $v \in I_1^k$. Clearly, $|I_1^2| \geq 2$ and $|I_2^2| \geq 2$. We will use induction to show that $|I_1^k| \geq 5 \cdot 2^{k-3}$ and $|I_2^k| \geq 5 \cdot 2^{k-3} - 1$ for $k \geq 3$.

For a fixed $k \geq 3$, by definition of *BicolorMax*, there exists a $(k-1)$ -mixed pair (x, y) in C_j^{k-1} with $x \in I_1^k$ and $y \in I_2^k$. By Lemma 6.1, x and y belong to two different components $D = (J_1, J_2)$ and $D' = (J'_1, J'_2)$ in $G_j^{k-1} - v$, say $x \in J_1 \subseteq I_1^k$ and $y \in J'_1 \subseteq I_2^k$. For $k = 3$, we get $|I_1^3| \geq |J_1| + |J'_2| + 1 = 2 + 2 + 1 = 5$ and $|I_2^3| \geq |J_2| + |J'_1| \geq 2 + 2 = 4$.

Now let $k \geq 4$. Then $|I_1^k| \geq |J_1| + |J'_2| + 1 = 5 \cdot 2^{k-4} + 5 \cdot 2^{k-4} - 1 + 1 = 5 \cdot 2^{k-3}$ and $|I_2^k| \geq |J_2| + |J'_1| \geq 5 \cdot 2^{k-4} - 1 + 5 \cdot 2^{k-4} = 5 \cdot 2^{k-3} - 1$. We conclude that $|I_1^k| \geq 5 \cdot 2^{k-3}$ and $|I_2^k| \geq 5 \cdot 2^{k-3} - 1$ for $k \geq 3$.

Assume $3 \leq \alpha \leq 4$. If $k = 3$, then $\alpha \geq |I_1^3| \geq 5$. Hence $k \leq 2$. Then $\chi_{BM} \leq 2k \leq 4$.

Finally assume $5 \cdot 2^{p-3} \leq \alpha \leq 5 \cdot 2^{p-2} - 1$ for some $p \geq 3$. If $k \geq p + 1$, then $\alpha \geq |I_1^k| \geq 5 \cdot 2^{p-2}$. Hence $k \leq p$. Then $\chi_{BM} \leq 2p$. This completes the proof of Theorem 7.2. $\square$

The bounds of Theorem 7.2 improve a bound in [5] for bipartite graphs: the main result of [5] applied to bipartite graphs shows that for any bipartite graph G any greedy on-line coloring algorithm uses at most $\alpha(G) + 1$ colors.

8. Final remarks. As we noted in section 5, *BicolorMax* is not on-line competitive for the family of paths. This is in contrast to *FF*, which uses at most three colors on any path P_n . Below we show in a more general way how to combine two on-line competitive algorithms satisfying certain conditions.

For this purpose we need the following terminology. We call a family of graphs $\mathcal{G}$ *decidable* if there exists a (finite) procedure to decide whether or not a graph G belongs to $\mathcal{G}$. Note that the family of H -free graphs is decidable for any fixed graph H . Also the family of paths is decidable. We call an on-line coloring algorithm A for a graph G *component-independent* if the color assigned by A to every revealed vertex $v = \pi(j)$ depends only on the component of $G[\{\pi(1), \dots, \pi(j)\}]$ that vertex v belongs to. For instance, *FF* and *BicolorMax* are component-independent. For two families of graphs $\mathcal{F}$ and $\mathcal{G}$ we define $\mathcal{F} + \mathcal{G} = \{L \mid L \in \mathcal{F} \cup \mathcal{G}, \text{ or } L = F + G \text{ with } F \in \mathcal{F} \text{ and } G \in \mathcal{G}\}$.

THEOREM 8.1. *Let $\mathcal{F}$ be a decidable family of graphs for which an on-line competitive coloring algorithm A_1 exists, and let $\mathcal{G}$ be a decidable family of graphs for which an on-line competitive component-independent coloring algorithm A_2 exists. Then there exists an on-line coloring algorithm A_3 that is on-line competitive for $\mathcal{F} + \mathcal{G}$.*

Proof. We assume A_1 and A_2 use disjoint color sets $\{c_1, c_2, \dots\}$ and $\{d_1, d_2, \dots\}$, respectively. Let $\pi \in \Pi(L)$ for some graph $L \in \mathcal{F} + \mathcal{G}$.

We inductively define A_3 as follows. The vertex $\pi(1)$ is colored with d_1 . Suppose that vertices $\pi(1), \dots, \pi(j-1)$ have already been colored and let $v = \pi(j)$ be the next vertex of L presented to A_3 . For $i = 1, 2$, we let $L_j(A_i)$ be the subgraph of L that consists of v and all vertices in $\{\pi(1), \dots, \pi(j-1)\}$ colored by A_i . We let $C_j(A_2)$ be the component of $L_j(A_2)$ that contains v . If $C_j(A_2)$ is an induced subgraph of some graph in $\mathcal{G}$, then color v with the color that A_2 would have used on v after the other vertices of $C_j(A_2)$ had been presented to A_2 in the suborder defined by π . Otherwise, $L_j(A_1)$ is a subgraph of a graph in $\mathcal{F}$, and we color v with the color that A_1 would have used on v after the other vertices of $L_j(A_1)$ had been presented to A_1 in the suborder defined by π . It is easy to see that A_3 is on-line competitive on $\mathcal{F} + \mathcal{G}$. $\square$

9. Conclusions and future work. We have presented an on-line coloring algorithm *BicolorMax* for bipartite graphs. We have shown that the number of colors used by this algorithm on a bipartite graph G is bounded from above by the number of remote induced subgraphs of G isomorphic to P_5 , and we gave a similar upper bound in terms of the independence number of G . As a consequence we improved known upper bounds for the on-line chromatic number of bipartite graphs given in [5, 18]. We showed in [3] that for any P_6 -free bipartite graph G , *BicolorMax* uses at most twice as many colors as any optimal on-line coloring algorithm for G . Here we showed that for any P_7 -free bipartite graph G , *BicolorMax* uses at most eight times as many colors as any optimal on-line coloring algorithm for G .

In a future continuation of this work, we would like to face the problem of deciding whether for any $n \geq 8$ a linear on-line competitive algorithm can be defined for the class of P_n -free bipartite graphs. We also consider analyzing *BicolorMax* and related algorithms for other classes of H -free bipartite graphs, in particular for bipartite graphs H with six vertices. A seemingly difficult and interesting open case is the (non)existence of an on-line competitive algorithm for the class of C_6 -free bipartite graphs.

Acknowledgments. We thank the anonymous referees for their useful comments on an earlier version of this paper.

REFERENCES

- [1] J. A. BONDY AND U. S. R. MURTY, *Graph Theory with Applications*, Elsevier, New York, 1976.
- [2] A. BORODIN AND R. EL-YANIV, *Online Computation and Competitive Analysis*, Cambridge University Press, New York, 1998.
- [3] H. J. BROERSMA, A. CAPPONI, AND D. PAULUSMA, *On-line coloring of H -free bipartite graphs*, in Proceedings of the 6th Annual Italian Conference on Algorithms and Complexity (CIAC 2006), Rome, Italy, Lecture Notes in Comput. Sci. 3998, Springer-Verlag, Berlin, 2006, pp. 284–295.
- [4] M. CHROBAK AND M. ŚLUSAREK, *Problem 84–23*, J. Algorithms, 5 (1984), p. 588.
- [5] I. CIEŚLIK, M. KOZIK, AND P. MICEK, *On-line coloring of I_s -free graphs and co-planar graphs*, in Proceedings of the Annual Discrete Mathematics and Theoretical Computer Science AF, Nancy, France, 2006, pp. 61–68.
- [6] T. ERLEBACH AND J. FIALA, *On-line coloring of geometric intersection graphs*, Comput. Geom., 23 (2002), pp. 243–255.
- [7] A. GYÁRFÁS, Z. KIRÁLY, AND J. LEHEL, *Online Competitive Coloring Algorithms*, Technical report TR-9703-1, 1997; available online at <http://www.cs.elte.hu/tr97/tr9703-1.ps>.
- [8] A. GYÁRFÁS, Z. KIRÁLY, AND J. LEHEL, *On-line 3-chromatic graphs. II. Critical graphs*, Discrete Math., 177 (1997), pp. 99–122.
- [9] A. GYÁRFÁS AND J. LEHEL, *On-line and first-fit colorings of graphs*, J. Graph Theory, 12 (1988), pp. 217–227.
- [10] A. GYÁRFÁS AND J. LEHEL, *First fit and on-line chromatic number of families of graphs*, Ars Combin., 29 (1990), pp. 168–176.

- [11] A. GYÁRFÁS AND J. LEHEL, *Effective on-line coloring of P_5 -free graphs*, *Combinatorica*, 11 (1991), pp. 181–184.
- [12] M. M. HALLDÓRSSON AND M. SZEGEDY, *Lower bounds for on-line graph coloring*, *Theoret. Comput. Sci.*, 130 (1994), pp. 163–174.
- [13] H. A. KIERSTEAD, *The linearity of first-fit coloring of interval graphs*, *SIAM J. Discrete Math.*, 1 (1988), pp. 526–530.
- [14] H. A. KIERSTEAD, *Coloring graphs on-line*, in *Online Algorithms: The State of the Art*, Lecture Notes in Comput. Sci. 1442, Springer-Verlag, Berlin, 1998, pp. 281–305.
- [15] H. KIERSTEAD AND S. G. PENRICE, *Radius two trees specify χ -bounded classes*, *J. Graph Theory*, 18 (1994), pp. 119–129.
- [16] H. KIERSTEAD, S. G. PENRICE, AND W. TROTTER, *On-line coloring and recursive graph theory*, *SIAM J. Discrete Math.*, 7 (1994), pp. 72–89.
- [17] K. KOLOSSA, *On the on-line chromatic number of the family of on-line 3-chromatic graphs*, *Discrete Math.*, 150 (1996), pp. 205–230.
- [18] L. LOVÁSZ, M. SAKS, AND W. T. TROTTER, *An on-line graph coloring algorithm with sublinear performance ratio*, *Discrete Math.*, 75 (1989), pp. 319–325.
- [19] B. RANDERATH AND I. SCHIERMEYER, *Vertex colouring and forbidden subgraphs—A survey*, *Graphs Combin.*, 20 (2004), pp. 1–40.
- [20] D. D. SLEATOR AND R. E. TARJAN, *Amortized efficiency of list updates and paging rules*, *Comm. ACM*, 28 (1985), pp. 202–208.

AVERAGE DISTANCE AND EDGE-CONNECTIVITY I*

PETER DANKELMANN[†], SIMON MUKWEMBI[†], AND HENDA C. SWART[†]

Abstract. The average distance $\mu(G)$ of a connected graph G of order n is the average of the distances between all pairs of vertices of G . We prove that if G is a λ -edge-connected graph of order n , then the bounds $\mu(G) \leq 2n/15 + 9$ if $\lambda = 5, 6$, $\mu(G) \leq n/9 + 10$ if $\lambda = 7$, and $\mu(G) \leq n/(\lambda + 1) + 5$ if $\lambda \geq 8$ hold. Our bounds are shown to be best possible, and our results solve a problem of Plesník [*J. Graph Theory*, 8 (1984), pp. 1–24].

Key words. distance, average distance, edge-connectivity

AMS subject classifications. 05C12

DOI. 10.1137/060656516

1. Introduction. Let $G = (V, E)$ be a connected graph of order n . The *average distance* $\mu(G)$ of G is defined as $\mu(G) = \binom{n}{2}^{-1} \sum_{\{x,y\} \subset V} d_G(x, y)$, where $d_G(x, y)$ denotes the distance between the vertices x and y in G . The average distance has been investigated by several authors and under various names; for instance, Kouider and Winkler [8], Favaron, Kouider, and Mahéo [6], Doyle and Graver [4] depict average distance as the *mean distance*; Wuchty and Stadler [16] use the term *average path length*; Bermond, Liu, and Syska [1] use the term *mean eccentricity*; whilst Plesník [12], Soltés [14] work with the total distance of a graph and call it the *transmission*. The transmission differs from the average distance by only a factor of $n(n-1)$. The study of average distance seems to have been started by the chemist Wiener [15] in 1947, when he introduced the Wiener index $W(G)$ of a graph G which differs from the average distance by a factor of $\binom{n}{2}$, i.e., $W(G) = \binom{n}{2} \mu(G)$. The applications of the Wiener index in physical chemistry are legion (see, for example, [7, 10, 13] for references); it has been used in the characterization of many different types of chemical species, including alkanes, alkenes, and arenes. The Wiener index is also a tool for designing quantitative structure-property relations in organic chemistry.

Several upper and lower bounds on the average distance in terms of other graph parameters are known. The *edge-connectivity* $\lambda(G)$ of G is the minimum number of edges whose deletion from G results in a disconnected or trivial graph. We say G is *k-edge-connected* if $\lambda(G) \geq k$. It has been proved (see, for example, [9, 4, 5]) that the average distance of a connected graph with n vertices is at most $(n+1)/3$ and that this bound is attained only by the path of order n . Plesník [12] showed that this bound can be improved for a 2-edge-connected (2-vertex-connected) graph G of order n to $\mu(G) \leq \lfloor \frac{n^2}{4} \rfloor / (n-1)$ and that this bound is achieved if and only if G is a cycle. He posed the problem of finding upper bounds on the average distance of a graph of given order and edge-connectivity. The goal of this paper is to find asymptotically sharp upper bounds on $\mu(G)$ for a graph G of given order and edge-connectivity λ . For $\lambda \geq 8$, we obtain bounds as an immediate consequence of known bounds on

*Received by the editors April 6, 2006; accepted for publication (in revised form) May 11, 2007; published electronically February 15, 2008. The results in this paper are part of the second author's Ph.D. thesis. Financial support by the National Research Foundation and the University of KwaZulu-Natal is gratefully acknowledged.

<http://www.siam.org/journals/sidma/22-1/65651.html>

[†]School of Mathematical Sciences, University of KwaZulu-Natal, Durban, 4041, South Africa (dankelma@ukzn.ac.za, mukwembi@ukzn.ac.za, swarth@ukzn.ac.za).

the average distance in terms of order and minimum degree. For small values of λ , $\lambda \in \{3, \dots, 7\}$, the problem requires a more intricate analysis. In this paper we give asymptotically sharp bounds on μ for $\lambda \in \{5, 6, 7\}$. The cases $\lambda \in \{3, 4\}$ will be considered in a subsequent paper [3].

We use the following notation. Let $G = (V, E)$ be a connected simple graph. For a subset $S \subseteq V$, $G[S]$ denotes the subgraph induced by S in G , and $\text{diam}(S)$ is the maximum value of $d_G(x, y)$, $x, y \in S$. Let $v \in V$. The *eccentricity* $ex_G(v)$ of v is the maximum value of $d_G(v, x)$, $x \in V$. The *distance* $\sigma_G(v)$ of v is the sum $\sum_{x \in V} d_G(v, x)$, whereas the *distance* $\sigma(G)$ of G is $\sum_{u \in V} \sigma_G(u) = \sum_{(x, y) \in V \times V} d_G(x, y)$. Thus $\mu(G) = \sigma(G)/n(n-1)$, where n is the order of G . For subsets $M_1, M_2 \subseteq V$, the distance $\sigma_{M_1}(v)$ of v with respect to M_1 is the sum $\sum_{x \in M_1} d_G(v, x)$, and the distance $\sigma_{M_1}(M_2)$ of M_2 with respect to M_1 is $\sum_{x \in M_2} \sigma_{M_1}(x)$. $E(M_1, M_2)$ denotes the set of edges $\{ab \in E \mid a \in M_1, b \in M_2\}$. $N_G(v) = N(v)$ denotes the set of all vertices adjacent to v in G , and its cardinality is the degree of v in G and is denoted by $\deg_G(v)$. We denote the closed neighbourhood of v by $N[v]$, i.e., $N[v] = N(v) \cup \{v\}$. For a positive integer i , the i th *distance layer* $N_i(v)$ of v is the set of vertices at distance i from v , i.e., $N_i(v) = \{x \in V \mid d_G(v, x) = i\}$. We denote the cardinality of $N_i(v)$ by $k_i(v)$, and if the vertex v is understood we omit the argument v and simply write N_i, k_i . We denote the set $\bigcup_{0 \leq j \leq i} N_j(v)$ by $N_{\leq i}(v)$ and the set $\bigcup_{j \geq i} N_j(v)$ by $N_{\geq i}(v)$.

A *disconnecting (separating) set* of G is a set of edges (vertices) of G whose removal increases the number of components of G . A *vertex cutset* of G is a separating set, S say, such that no proper subset of S is a separating set of G . A vertex v of G is a *cutvertex* if $\{v\}$ is a separating set of G . A maximal subgraph of G with no cutvertex is called a *block*, and an *end block* of G is a block of G that contains exactly one cutvertex of G . For vertex disjoint graphs $G_1, G_2, \dots, G_k$, the *sequential join* $G_1 + G_2 + \dots + G_k$ is the graph obtained from the union of $G_1, \dots, G_k$ by joining every vertex of G_i to every vertex of G_{i+1} for $i = 1, 2, \dots, k-1$.

2. Results on 8-edge-connected graphs. Dankelmann and Entringer [2] proved that if G is a connected graph of order n and minimum degree δ , then $\mu(G) \leq n/(\delta+1) + 5$. An application of the well-known inequality $\lambda \leq \delta$ to this bound yields part (a) of the following proposition.

PROPOSITION 1. *Let G be a λ -edge-connected graph of order n .*

(a) *Then $\mu(G) \leq \frac{n}{\lambda+1} + 5$.*

(b) *The bound in (a), apart from the additive constant, is best possible for $\lambda \geq 8$.*

Proof. (b) Assume that $\lambda \geq 8$. We show that the bound is sharp apart from the exact value of the additive constant. Given integers n, λ , and k with $n = \frac{1}{3}(\lambda+1)(k+2)$, $k \equiv 1 \pmod{3}$, $k \geq 7$, let

$$G_{n,\lambda} = G_1 + G_2 + G_3 + \dots + G_k,$$

where $G_1 = K_1 = G_k$, $G_2 = K_\lambda = G_{k-1}$, and for $3 \leq i \leq k-2$, G_i

$$= \begin{cases} K_{\frac{\lambda+1}{3}} & \text{if } \lambda \equiv 2 \pmod{3}, \\ K_{\frac{\lambda}{3}} \text{ for } i \equiv 0, 2 \pmod{3} \text{ and } K_{\frac{\lambda}{3}+1} \text{ for } i \equiv 1 \pmod{3} & \text{if } \lambda \equiv 0 \pmod{3}, \\ K_{\frac{\lambda+2}{3}} \text{ for } i \equiv 0, 2 \pmod{3} \text{ and } K_{\frac{\lambda-1}{3}} \text{ for } i \equiv 1 \pmod{3} & \text{if } \lambda \equiv 1 \pmod{3}. \end{cases}$$

Clearly, since $\lambda \geq 8$, $G_{n,\lambda}$ is λ -edge-connected. It is easy to show that, for constant λ and large n , $\mu(G_{n,\lambda}) \geq \frac{n}{\lambda+1} + O(1)$. $\square$

3. Results on 5-edge-connected graphs.

LEMMA 2. *Let G be a 5-edge-connected graph with n vertices, and let v be any vertex of G . Then $\sigma_G(v) \leq \frac{1}{5}(n^2 - 4n + 14)$. Moreover, $\sigma(G) \leq \frac{n}{5}(n^2 - 4n + 14)$.*

Proof. Let e be the eccentricity of v , and note that

$$(1) \quad \sigma_G(v) = 1k_1 + 2k_2 + \cdots + ek_e.$$

The lemma can easily be verified for $e \leq 4$. Assume that $e \geq 5$. Since G is 5-edge-connected, $k_i k_{i+1} \geq |E(N_i, N_{i+1})| \geq \lambda(G) \geq 5$. From $k_i k_{i+1} \geq 5$ we deduce that $k_i + k_{i+1} \geq 5$. Hence it is easy to see that

$$(*) \quad k_i k_{i+1} \geq 5, k_i + k_{i+1} \geq 5 \text{ for } i = 0, 1, \dots, e-1, k_{e-1} + k_e \geq 6, \sum_{i=0}^e k_i = n.$$

Let $(a_0, a_1, \dots, a_e)$ be a sequence that maximizes (1) subject to (*). We first show that we can assume that $a_i > 1$ for $i = 2, \dots, e-1$. If $a_2 = 1$, then since $e \geq 5$ and $a_3 \geq 5$, the sequence $(b_0, b_1, \dots, b_e)$ given by $b_2 = 2$, $b_3 = a_3 - 2$, $b_4 = a_4 + 1$, and $b_i = a_i$ otherwise, also maximizes (1) subject to (*). Hence we can assume that $a_2 > 1$. Now we show that $a_i > 1$ for $i = 3, \dots, e-1$. If this were false, let s , $3 \leq s \leq e-1$, be the smallest integer for which $a_s = 1$. Let $(b_0, b_1, \dots, b_e)$ be the sequence given by (i) $b_i = 3$ if $i = s-1$, (ii) $b_i = 2$ if $i = s$, (iii) $b_i = a_{s-1} + a_{s+1} - 4$ if $i = s+1$, and (vi) $b_i = a_i$ otherwise. By (*), $a_{s-1}, a_{s+1} \geq 5$. This, in conjunction with $b_{s-2} = a_{s-2} \geq 2$, yields that the sequence (b_i) satisfies (*), and, moreover, $\sum_{i=0}^e i a_i = \sum_{i=0}^e i b_i + 7 - 2a_{s-1} < \sum_{i=0}^e i b_i$, contradicting our choice of the sequence (a_i) . Therefore, we can assume that $a_i \geq 2$ for all $i = 2, 3, \dots, e-1$.

Hence we maximize (1) subject to the conditions

$$k_1 \geq 5, k_i \geq 2, k_i + k_{i+1} \geq 5, \text{ for } i = 1, 2, \dots, e-1, k_{e-1} + k_e \geq 6, \sum_{i=0}^e k_i = n.$$

Clearly for fixed e , (1) is maximized, subject to these conditions, for

$$(k_0, k_1, \dots, k_{e-1}) = \begin{cases} (1, 5, 2, 3, 2, 3, \dots, 2, 3) & \text{if } e \text{ is even,} \\ (1, 5, 2, 3, 2, 3, \dots, 2, 3, 2) & \text{if } e \text{ is odd,} \end{cases}$$

and $k_e = n - \sum_{i=0}^{e-1} k_i$. Then

$$\begin{aligned} \sigma_G(v) &= 1k_1 + 2k_2 + 3k_3 + 4k_4 + 5k_5 + \cdots + ek_e \\ &\leq \begin{cases} 1 \cdot 5 + [2 \cdot 2 + 3 \cdot 3] + \cdots + [(e-2) \cdot 2 \\ \quad + (e-1) \cdot 3] + e(n - \frac{5}{2}e - 1) & \text{if } e \text{ is even,} \\ 1 \cdot 5 + [2 \cdot 2 + 3 \cdot 3] + \cdots + [(e-3) \cdot 2 \\ \quad + (e-2) \cdot 3] + (e-1) \cdot 2 + e(n - \frac{5}{2}e - \frac{1}{2}) & \text{if } e \text{ is odd} \end{cases} \\ (2) \quad &\leq en - \frac{5}{4}e^2 - 2e + \frac{9}{4}. \end{aligned}$$

By (*),

$$n = \sum_{i=0}^e k_i \geq \begin{cases} \frac{5e+8}{2} & \text{if } e \text{ is even,} \\ \frac{5e+9}{2} & \text{if } e \text{ is odd.} \end{cases}$$

(Note that for even e we use $k_2 + k_3 + k_4 \geq 7$.) Hence $e \leq \frac{2n-8}{5}$. A simple differentiation shows that (2) is maximized, subject to the constraint $e \leq \frac{2n-8}{5}$, for $e = \frac{2n-8}{5}$, and we obtain $\sigma_G(v) \leq \left(\frac{2n-8}{5}\right)n - \frac{5}{4}\left(\frac{2n-8}{5}\right)^2 - 2\left(\frac{2n-8}{5}\right) + \frac{9}{4} \leq \frac{1}{5}(n^2 - 4n + 14)$, as desired.

Summing $\sigma_G(v)$ for all v , the bound in the second part of the lemma follows from the first part. $\square$

PROPOSITION 3. *Let G be a 5-edge-connected block and v a vertex of G . If $k_l + k_{l+1} + k_{l+2} \leq 11$ for some l , $1 \leq l \leq \text{ex}_G(v) - 3$, then $\text{diam}(N_{l+1}) \leq 2$ and $2 \leq k_{l+1} \leq 7$.*

Proof. Suppose to the contrary that for some $a, b \in N_{l+1}$, $d(a, b) > 2$. Then $N[a] \cap N[b] = \emptyset$. Thus, since $N[a], N[b] \subseteq N_l \cup N_{l+1} \cup N_{l+2}$ and $\delta(G) = \delta \geq 5$, we have $k_l + k_{l+1} + k_{l+2} \geq |N[a]| + |N[b]| \geq (\delta + 1) + (\delta + 1) \geq 12$, a contradiction. Therefore, for all $a, b \in N_{l+1}$, $d(a, b) \leq 2$ as desired.

Since G is a block, $k_l, k_{l+1}, k_{l+2} \geq 2$. Therefore, $2 \leq k_{l+1} \leq 11 - (k_l + k_{l+2}) \leq 11 - 4 = 7$. $\square$

DEFINITION 4. *Let G be a 5-edge-connected graph and v a vertex of G . We say that (N_l, N_{l+1}, N_{l+2}) , where $22 \leq l \leq \text{ex}_G(v) - 24$, is a forbidden separating triple of v if $k_l + k_{l+1} + k_{l+2} \leq 11$.*

Clearly, if a vertex v with $\text{ex}_G(v) \geq 46$ has no forbidden separating triple, then for all $i \in \{22, \dots, \text{ex}_G(v) - 24\}$ we have that $k_i + k_{i+1} + k_{i+2} \geq 12$. This will allow us to show that the distance of a vertex v with no forbidden separating triple is small and bounded by a term of the form $\frac{n^2}{8} + O(n)$.

LEMMA 5. *Let G be a 5-edge-connected graph with n vertices, and let v be a vertex of G with $\text{ex}_G(v) \geq 46$. If v has no forbidden separating triple, then*

$$\sigma_G(v) \leq \frac{1}{8}n^2 + \frac{33}{4}n + \frac{1361}{8}.$$

Proof. Denote the eccentricity of v by e , where e is fixed. We first show that

$$(3) \quad \sigma_G(v) \leq en - 2e^2 + 33e + 34.$$

As above,

$$(4) \quad \sigma_G(v) = 1k_1 + 2k_2 + 3k_3 + \dots + ek_e.$$

Since G is a 5-edge-connected graph and v has no forbidden separating triple, the sequence (k_i) satisfies the conditions: $k_1 \geq 5$; $k_{e-1} + k_e \geq 6$; $k_i \geq 1$ for $i = 0, 1, \dots, e$; $k_i + k_{i+1} \geq 5$ for $i = 0, 1, \dots, 20$, $i = e - 21, \dots, e - 1$; $k_i + k_{i+1} + k_{i+2} \geq 12$ for $i = 22, 23, \dots, e - 24$ and $\sum_{i=0}^e k_i = n$. If $e \equiv 0 \pmod{3}$, then clearly (4) is maximized, subject to these conditions, for $(k_0, k_1, \dots, k_{21}) = (1, 5, 1, 4, \dots, 1, 4)$, $(k_{22}, k_{23}, \dots, k_{e-22}) = (1, 1, 10, 1, 1, 10, \dots, 1, 1, 10, 1, 1)$, $(k_{e-21}, k_{e-20}, \dots, k_{e-1}) = (1, 4, 1, 4, \dots, 1, 4, 1)$, and k_e as large as possible. Then

$$\begin{aligned} \sigma_G(v) &= 1k_1 + 2k_2 + 3k_3 + \dots + ek_e \\ &\leq 5 + [2 \cdot 1 + 3 \cdot 4] + \dots + [20 \cdot 1 + 21 \cdot 4] + [22 \cdot 1 + 23 \cdot 1 + 24 \cdot 10] \\ &\quad + \dots + [(e - 26) \cdot 1 + (e - 25) \cdot 1 + (e - 24) \cdot 10] + (e - 23) \cdot 1 \\ &\quad + (e - 22) \cdot 1 + [(e - 21) \cdot 1 + (e - 20) \cdot 4] + \dots \\ &\quad + [(e - 3) \cdot 1 + (e - 2) \cdot 4] + (e - 1) \cdot 1 + e(n - 4e + 71) \\ &= en - 2e^2 + 33e + 34. \end{aligned}$$

By a similar argument, for $e \equiv 1 \pmod 3$ and for $e \equiv 2 \pmod 3$, we have $\sigma_G(v) \leq en - 2e^2 + 33e + 34$.

Note that by the conditions satisfied by the sequence (k_i) , we have

$$n = \sum_{i=0}^e k_i \geq \begin{cases} 4e - 66 & \text{if } e \equiv 0 \pmod 3, \\ 4e - 63 & \text{if } e \equiv 2 \pmod 3, \\ 4e - 60 & \text{if } e \equiv 1 \pmod 3. \end{cases}$$

Hence $e \leq \frac{n+66}{4}$. A simple differentiation shows that (3) is maximized, subject to the constraint $e \leq \frac{n+66}{4}$, for $e = \frac{n+33}{4}$, and we obtain

$$\begin{aligned} \sigma_G(v) &\leq \left(\frac{n+33}{4}\right)n - 2\left(\frac{n+33}{4}\right)^2 + 33\left(\frac{n+33}{4}\right) + 34 \\ &= \frac{1}{8}n^2 + \frac{33}{4}n + \frac{1361}{8}, \end{aligned}$$

as desired. $\square$

We first establish an upper bound on distance for blocks.

THEOREM 6. *Let G be a 5-edge-connected block of order n . Then*

$$\sigma(G) \leq n(n-1) \left(\frac{2}{15}n + 9 \right).$$

Proof. Let v be a vertex of G of largest distance, and let e be its eccentricity.

For $e \leq 45$, the result is obtained as follows. If, on one hand, $n \leq 116$, then the second part of Lemma 2 yields the desired bound by a simple calculation. If, on the other hand, $n \geq 117$, then we have

$$\begin{aligned} \sigma_G(v) &\leq 5 + [2(2) + 3(3)] + \cdots + [42(2) + 43(3)] + 44(2) \\ &\quad + 45(n - 6 - 5 \cdot 21 - 2) \\ &= 45n - 2619. \end{aligned}$$

Since $45n - 2619 \leq (n-1)(\frac{2}{15}n + 9)$ for $n \geq 117$, it follows that $\sigma(G) \leq n\sigma_G(v) \leq n(45n - 2619) \leq n(n-1)(\frac{2}{15}n + 9)$ for all n . Thus the result holds.

So assume that $e \geq 46$ and hence $n \geq 119$. If v has no forbidden separating triple, then summing $\sigma_G(v)$ for all v , we obtain the result by Lemma 5.

If v has a forbidden separating triple, we prove the theorem by induction on the order n of G . For $n \leq 143$, the result follows by the second part of Lemma 2. So assume that $n \geq 144$ and that the result holds for any 5-edge-connected block with less than n vertices. Since v has a forbidden separating triple, let N_α , $23 \leq \alpha \leq e-23$, be chosen in accordance with Proposition 3, that is, $\text{diam}(N_\alpha) \leq 2$ and $2 \leq k_\alpha \leq 7$. Let K_5 be a copy of the complete graph with five vertices such that $V(G) \cap V(K_5) = \emptyset$. Denote the vertex set of K_5 by $A = \{a_1, \dots, a_5\}$. We form a new graph H from $G \cup K_5$ by attaching K_5 to N_α , that is, by joining every vertex in N_α to every vertex in K_5 .

Let $G_1 = H[N_{\leq \alpha} \cup A]$ and $G_2 = H[N_{\geq \alpha} \cup A]$. Since $\text{diam}(N_\alpha) \leq 2$, no distances in G have been reduced; hence $d_G(x, y) = d_H(x, y)$ for all $x, y \in V(G)$. The proof of the following claim is elementary.

Claim 1. Let $X = \sum_{i=1}^5 \sigma_{G_1}(a_i)$ and $Y = \sum_{i=1}^5 \sigma_{G_2}(a_i)$. Then

$$\begin{aligned} \sigma(G) &= \sigma(G_1) + \sigma(G_2) - 2(X + Y) + 2\sigma_A(A) \\ &\quad + 2 \sum_{x \in N_{< \alpha}} \sum_{y \in N_{> \alpha}} d_H(x, y) - \sum_{(x, y) \in N_\alpha \times N_\alpha} d_H(x, y). \end{aligned}$$

Note that G_1 and G_2 have $N_\alpha \cup A$ in common. Hence, letting $|V(G_i)| = n_i$, $i = 1, 2$, we obtain

$$(5) \quad n_1 + n_2 - k_\alpha - 10 = n.$$

Note also that $n_1 \geq |N_{\leq \alpha}| + |A| \geq (k_0 + k_1) + \dots + (k_{20} + k_{21}) + k_{22} + k_\alpha + 5 \geq 63 + k_\alpha$. Analogously, $n_2 \geq |N_{\geq \alpha}| + |A| \geq 63 + k_\alpha$. It follows from (5) that

$$63 + k_\alpha \leq n_1, \quad n_2 \leq n - 53.$$

We bound $\sigma(G_1)$ and $\sigma(G_2)$ using induction. It is easy to verify the following claim.

Claim 2. G_1 and G_2 are 5-edge-connected blocks.

We now find bounds on the terms in Claim 1. First, note that

$$(6) \quad \sigma_A(A) \leq \sigma(K_5) = 20 \quad \text{and} \quad \sum_{(x,y) \in N_\alpha \times N_\alpha} d_H(x, y) \geq k_\alpha(k_\alpha - 1).$$

Now, let a be a fixed vertex in N_α . Then

$$\begin{aligned} 2 \sum_{x \in N_{< \alpha}, y \in N_{> \alpha}} d_H(x, y) &\leq 2 \sum_{x \in N_{< \alpha}} \sum_{y \in N_{> \alpha}} [d_H(x, a) + d_H(a, y)] \\ &= 2 \sum_{x \in N_{< \alpha}} [|N_{> \alpha}| d_H(x, a) + \sigma_{G_2}(a) - \sigma_{N_\alpha \cup A}(a)] \\ &= 2[|N_{> \alpha}|(\sigma_{G_1}(a) - \sigma_{N_\alpha \cup A}(a)) \\ &\quad + |N_{< \alpha}|(\sigma_{G_2}(a) - \sigma_{N_\alpha \cup A}(a))]. \end{aligned}$$

Note that $\sigma_{N_\alpha \cup A}(a) \geq |(N_\alpha \cup A) - \{a\}| = k_\alpha + 4$. Thus,

$$\begin{aligned} (7) \quad 2 \sum_{x \in N_{< \alpha}, y \in N_{> \alpha}} d_H(x, y) &\leq 2[(n_2 - k_\alpha - 5)(\sigma_{G_1}(a) - k_\alpha - 4) + (n_1 - k_\alpha - 5)(\sigma_{G_2}(a) - k_\alpha - 4)]. \end{aligned}$$

We find lower bounds on X and Y . Note that for $i = 1, 2, \dots, 5$,

$$\begin{aligned} \sigma_{G_1}(a) &= \sum_{x \in V(G_1)} d(a, x) \leq \sum_{x \in V(G_1)} [d(a, a_i) + d(a_i, x)] \\ &= \sum_{x \in V(G_1)} [1 + d(a_i, x)] = n_1 + \sigma_{G_1}(a_i). \end{aligned}$$

Therefore, $\sigma_{G_1}(a_i) \geq \sigma_{G_1}(a) - n_1$ and

$$(8) \quad X \geq 5\sigma_{G_1}(a) - 5n_1.$$

Analogously,

$$(9) \quad Y \geq 5\sigma_{G_2}(a) - 5n_2.$$

Using induction, we have $\sigma(G_i) \leq n_i(n_i - 1) \left(\frac{2}{15}n_i + 9\right)$, whereas by Lemma 2 we have $\sigma_{G_i}(a) \leq \frac{1}{5}(n_i^2 - 4n_i + 14)$ for $i = 1, 2$.

This, together with (5), Claim 1, and (6)–(9), yields

$$\begin{aligned} \sigma(G) \leq n(n-1) \left(\frac{2}{15}n + 9 \right) &+ \left\{ (n_1 + n_2) \left(\frac{28}{3}k_\alpha + \frac{2414}{15} - \frac{2}{5}k_\alpha^2 \right) \right. \\ &\left. - n_1n_2 \left(\frac{194}{15} - \frac{4}{5}k_\alpha \right) + \frac{2}{15}k_\alpha^3 - \frac{28}{15}k_\alpha^2 - \frac{1808}{15}k_\alpha - \frac{2506}{3} \right\}. \end{aligned}$$

Denote the term in curly brackets of the right-hand side of the last inequality by $f(n_1, n_2)$. It is sufficient to show that, subject to $63 + k_\alpha \leq n_1, n_2$, we have $f(n_1, n_2) \leq 0$. Clearly, $(n_1 - (63 + k_\alpha))(n_2 - (63 + k_\alpha)) \geq 0$; hence $n_1n_2 \geq (63 + k_\alpha)(n_1 + n_2) - (63 + k_\alpha)^2$. This, in conjunction with (5), yields $f(n_1, n_2) \leq n(\frac{2}{5}k_\alpha^2 + \frac{234}{5}k_\alpha - \frac{9808}{15}) - \frac{4}{15}k_\alpha^3 - \frac{584}{15}k_\alpha^2 - 1852k_\alpha + \frac{219792}{5}$, and thus from $2 \leq k_\alpha \leq 7$ we obtain $f(n_1, n_2) \leq \frac{601448}{15} - \frac{920}{3}n \leq 0$ for all $n \geq 131$, as desired. $\square$

We now show that the upper bound on the distance proved above for 5-edge-connected blocks holds for all 5-edge-connected graphs.

THEOREM 7. *Let G be a 5-edge-connected graph of order n . Then*

$$\mu(G) \leq \frac{2}{15}n + 9.$$

Apart from the additive constant, this inequality is best possible.

Proof. We prove this result by induction on the number of blocks in G . If G is a block, then the result follows by Theorem 6. So assume that G has at least two blocks. We prove the equivalent statement

$$\sigma(G) \leq n(n-1) \left(\frac{2}{15}n + 9 \right).$$

Let G_1 be an end block, G_2 be the union of the other blocks, and u be the unique common vertex of G_1 and G_2 . For $i = 1, 2$, let $n_i = |V(G_i)|$. Then the relation $n_1 + n_2 - 1 = n$ holds. By definition,

$$\begin{aligned} \sigma(G) &= \sum_{(x,y) \in V(G) \times V(G)} d(x, y) \\ &= \sigma(G_1) + 2 \sum_{x \in V(G_1) - \{u\}} \sum_{y \in V(G_2) - \{u\}} d(x, y) + \sigma(G_2) \\ (10) \quad &= \sigma(G_1) + 2[(n_2 - 1)\sigma_{G_1}(u) + (n_1 - 1)\sigma_{G_2}(u)] + \sigma(G_2). \end{aligned}$$

Clearly, G_1 and G_2 are 5-edge-connected. Thus, by induction, $\sigma(G_i) \leq n_i(n_i - 1)(\frac{2}{15}n_i + 9)$ for $i = 1, 2$. By Lemma 2, we have $\sigma_{G_i}(u) \leq \frac{1}{5}(n_i^2 - 4n_i + 14)$ for $i = 1, 2$. This, in conjunction with (10) and $n_1 + n_2 - 1 = n$, yields $\sigma(G) \leq n(n-1)(\frac{2}{15}n + 9) + \{\frac{368}{15}(n_1 + n_2) - \frac{302}{15}n_1n_2 - \frac{434}{15}\}$. Denote the term in curly brackets of the right-hand side of the last inequality by $f(n_1, n_2)$. We show that $f(n_1, n_2) \leq 0$ for all $n_1, n_2 \geq 6$. Clearly, $(n_1 - 6)(n_2 - 6) \geq 0$; hence $n_1n_2 \geq 6(n_1 + n_2) - 36$. This, in conjunction with $n_1 + n_2 - 1 = n$, yields $f(n_1, n_2) \leq \frac{2998}{5} - \frac{1444}{15}n \leq 0$ for all $n \geq 7$. Note that if $n = 6$, then $G = K_6$, and the result holds.

We now show that, apart from the additive constant, the bound is best possible. Given integers n, k with $n = \frac{1}{2}(5k + 9)$, $k \geq 5$ odd, let

$$G_n = G_1 + G_2 + \cdots + G_k,$$

where $G_1 = K_5 = G_k$ and for $2 \leq i \leq k-1$,

$$G_i = \begin{cases} K_2 & \text{if } i \text{ is even,} \\ K_3 & \text{if } i \text{ is odd.} \end{cases}$$

Clearly, G_n is 6-edge-connected (and thus 5-edge-connected) of order n . Straight-forward calculations show that

$$\mu(G_n) \geq \frac{2}{n(n-1)} \left(\frac{1}{15}(n^3 - 15n^2 + 20n + 57) \right) = \frac{2}{15}n - \frac{28}{15} + \frac{2(6n+57)}{15n(n-1)},$$

as desired. $\square$

Since every 6-edge-connected graph is also 5-edge-connected and since the graph G_n defined above is 6-edge-connected, the following theorem becomes immediate.

THEOREM 8. *Let G be a 6-edge-connected graph of order n . Then*

$$\mu(G) \leq \frac{2}{15}n + 9.$$

Apart from the additive constant, this inequality is best possible.

4. Results on 7-edge-connected graphs. We now focus our attention on 7-edge-connected graphs.

LEMMA 9. *Let G be a 7-edge-connected graph with n vertices, and let v be an arbitrary vertex of G . Then $\sigma_G(v) \leq \frac{1}{6}(n^2 - 6n + 27)$. Moreover, $\sigma(G) \leq \frac{n}{6}(n^2 - 6n + 27)$.*

Proof. Let e be the eccentricity of v . Then

$$(11) \quad \sigma_G(v) = 1k_1 + 2k_2 + \cdots + ek_e.$$

The lemma can easily be verified for $e \leq 4$. Assume that $e \geq 5$. Since G is 7-edge-connected and as in the proof of Lemma 2, we have

$$(*) \quad k_1, k_i k_{i+1} \geq 7, k_i + k_{i+1} \geq 6 \text{ for } i = 1, \dots, e-1, k_{e-1} + k_e \geq 8, n = \sum_{i=0}^e k_i.$$

Let $(a_0, a_1, \dots, a_e)$ be a sequence that maximizes (11) subject to (*). We first show that $a_i > 1$ for $i = 2, \dots, e-1$. If this were not true, let s , $2 \leq s \leq e-1$, be the smallest integer for which $a_s = 1$. First assume that $s \geq 3$. Let $(b_0, b_1, \dots, b_e)$ be the sequence defined by (i) $b_i = 4$ if $i = s-1$, (ii) $b_i = 2$ if $i = s$, (iii) $b_i = a_{s-1} + a_{s+1} - 5$ if $i = s+1$, and (vi) $b_i = a_i$ otherwise. Then from $a_{s-2} \geq 2$, $a_{s-1}, a_{s+1} \geq 7$, we have that the sequence (b_i) satisfies (*) and $\sum_{i=0}^e ia_i = \sum_{i=0}^e ib_i - (2a_{s-1} - 9) < \sum_{i=0}^e ib_i$, contradicting our choice of the sequence (a_i) . If $s = 2$, then let $(b_0, b_1, \dots, b_e)$ be defined by $b_2 = 2$, $b_3 = 4$, $b_4 = a_3 + a_4 - 5$, and $b_i = a_i$ otherwise. Since $a_3 \geq 7$ and $e \geq 5$, the sequence (b_i) satisfies (*) and $\sum_{i=0}^e ia_i = \sum_{i=0}^e ib_i - (a_3 - 6) < \sum_{i=0}^e ib_i$, again contradicting our choice of the sequence (a_i) . Therefore, $a_i \geq 2$ for all $i = 2, \dots, e-1$. Hence we maximize (11) subject to the conditions

$$k_1 \geq 7, k_i \geq 2, k_i + k_{i+1} \geq 6 \text{ for } i = 1, 2, \dots, e-1, k_{e-1} + k_e \geq 8, \sum_{i=0}^e k_i = n.$$

Clearly for fixed e , (11) is maximized, subject to these constraints, for

$$(k_0, k_1, \dots, k_{e-1}) = \begin{cases} (1, 7, 2, 4, 2, 4, \dots, 2, 4) & \text{if } e \text{ is even,} \\ (1, 7, 2, 4, 2, 4, \dots, 2, 4, 2) & \text{if } e \text{ is odd.} \end{cases}$$

As in the proof of Lemma 2, we obtain $\sigma_G(v) \leq \frac{1}{6}(n^2 - 6n + 27)$, as desired. The second part also follows similarly. $\square$

The proposition below is an analogue to Proposition 3. Its proof is similar, so we omit it.

PROPOSITION 10. *Let G be a 7-edge-connected block and v a vertex of G . If for some l , $1 \leq l \leq \text{ex}_G(v) - 3$, we have $k_l + k_{l+1} + k_{l+2} \leq 15$, then $\text{diam}(N_{l+1}) \leq 2$ and $2 \leq k_{l+1} \leq 11$.*

DEFINITION 11. *Let G be a 7-edge-connected graph and v a vertex of G . We say that (N_l, N_{l+1}, N_{l+2}) , where $22 \leq l \leq \text{ex}_G(v) - 24$, is a forbidden separating triple of v if $k_l + k_{l+1} + k_{l+2} \leq 15$.*

The bound on the distance of a vertex of G with no forbidden separating triple is given below. Its proof is analogous to the proof of Lemma 5.

LEMMA 12. *Let G be a 7-edge-connected graph of order n and v a vertex of G with $\text{ex}_G \geq 46$. If v has no forbidden separating triple, then*

$$\sigma_G(v) \leq \frac{3}{32}n^2 + \frac{153}{16}n + \frac{9275}{32}.$$

THEOREM 13. *Let G be a 7-edge-connected block of order n . Then*

$$\mu(G) \leq \frac{1}{9}n + 10.$$

Proof. Let v be a vertex of G of largest distance and let e be its eccentricity.

If $e \leq 45$, then the result is obtained as follows. If, on one hand, $n \leq 142$, then, by the second part of Lemma 9, $\sigma(G) \leq \frac{1}{6}n(n^2 - 6n + 27) \leq n(n-1)(\frac{1}{9}n + 10)$ for all $n \leq 142$ and the theorem holds. If, on the other hand, $n \geq 143$, then, as in the proof of Lemma 9, we have

$$\begin{aligned} \sigma_G(v) &\leq 7 + [2(2) + 3(4)] + \cdots + [42(2) + 43(4)] + 44(2) + 45(n - 136) \\ &= 45n - 3169. \end{aligned}$$

Since $45n - 3169 \leq (n-1)(\frac{1}{9}n + 10)$ for $n \geq 143$, it follows that $\sigma(G) \leq n\sigma_G(v) \leq n(n-1)(\frac{1}{9}n + 10)$ for all n . Thus the theorem holds for $e \leq 45$.

So assume that $e \geq 46$ and hence $n \geq 144$. If v has no forbidden separating triple, then summing $\sigma_G(v)$ for all v , we obtain the result by Lemma 12. If v has a forbidden separating triple, then we proceed with our proof as follows. We prove the equivalent statement

$$\sigma(G) \leq n(n-1) \left(\frac{1}{9}n + 10 \right)$$

by induction on the order n of G . For $n \leq 194$, the result follows by the second part of Lemma 9. So assume that $n \geq 195$ and that the result holds for any 7-edge-connected block with fewer than n vertices. Since v has a forbidden separating triple, let N_α , $23 \leq \alpha \leq e - 23$, be chosen in accordance with Proposition 10, that is, $\text{diam}(N_\alpha) \leq 2$ and $2 \leq k_\alpha \leq 11$. Let K_7 be a copy of the complete graph with seven vertices such that $V(G) \cap V(K_7) = \emptyset$. Denote the vertex set of K_7 by $A = \{a_1, \dots, a_7\}$. We form a new graph H from $G \cup K_7$ by attaching K_7 to N_α , that is, by joining each vertex in N_α to every vertex in K_7 .

Let $G_1 = H[N_{\leq \alpha} \cup A]$ and $G_2 = H[N_{\geq \alpha} \cup A]$. Since $\text{diam}(N_\alpha) \leq 2$, no distances in G have been reduced; hence $d_G(x, y) = d_H(x, y)$ for all $x, y \in V(G)$. The claim below is easy to establish.

Claim 3. Let $X = \sum_{i=1}^7 \sigma_{G_1}(a_i)$ and $Y = \sum_{i=1}^7 \sigma_{G_2}(a_i)$. Then

$$\begin{aligned} \sigma(G) &= \sigma(G_1) + \sigma(G_2) - 2(X + Y) + 2\sigma_A(A) \\ &\quad + 2 \sum_{x \in N_{<\alpha}} \sum_{y \in N_{>\alpha}} d_H(x, y) - \sum_{(x,y) \in N_\alpha \times N_\alpha} d_H(x, y). \end{aligned}$$

By bounding the terms on the right-hand side of the equation in Claim 3 in a manner similar to that done in Theorem 6, we obtain $\sigma(G) \leq n(n-1)(\frac{1}{9}n + 10)$, as desired. $\square$

THEOREM 14. *Let G be a 7-edge-connected graph of order n . Then*

$$\mu(G) \leq \frac{1}{9}n + 10.$$

Apart from the additive constant, this inequality is best possible.

Proof. The proof of this result is treated similarly to the proof of Theorem 7. To show that the bound is best possible apart from the exact value of the additive constant, let n and k be integers satisfying $n = 3k + 4$, $k \geq 6$, and let $G_{n,7} = G_1 + G_2 + \cdots + G_k$, where $G_1 = K_1 = G_k$, $G_2 = K_7 = G_{k-1}$, and $G_i = K_3$ for $i = 3, \dots, k-2$. Clearly, $G_{n,7}$ is 7-edge-connected and $\mu(G_{n,7}) = \frac{1}{9}n + \frac{1}{9} + \frac{2(736-97n)}{9n(n-1)}$, as desired. $\square$

REFERENCES

- [1] J. C. BERMOND, Z. LIU, AND M. SYSKA, *Mean eccentricities of DeBruijn networks*, Networks, 30 (1997), pp. 187–203.
- [2] P. DANKELMANN AND R. C. ENTRINGER, *Average distance, minimum degree, and spanning trees*, J. Graph Theory, 33 (2000), pp. 1–13.
- [3] P. DANKELMANN, S. MUKWEMBI, AND H. C. SWART, *Average distance and edge-connectivity II*, SIAM J. Discrete Math., 21 (2007), pp. 1035–1052.
- [4] J. K. DOYLE AND J. E. GRAVER, *Mean distance in a graph*, Discrete Math., 7 (1977), pp. 147–154.
- [5] R. C. ENTRINGER, D. E. JACKSON, AND D. A. SNYDER, *Distance in graphs*, Czechoslovak Math. J., 26 (1976), pp. 283–296.
- [6] O. FAVARON, M. KOUIDER, AND M. MAHÉO, *Edge-vulnerability and mean distance*, Networks, 19 (1989), pp. 493–504.
- [7] I. GUTMAN AND O. E. POLANSKY, *Mathematical Concepts in Organic Chemistry*, Springer-Verlag, Berlin, 1986.
- [8] M. KOUIDER AND P. WINKLER, *Mean distance and minimum degree*, J. Graph Theory, 25 (1997), pp. 95–99.
- [9] L. LOVÁSZ, *Combinatorial Problems and Exercises*, North-Holland, Amsterdam, 1979.
- [10] Z. MIHALIC, D. VELJAN, D. AMIC, S. NIKOLIC, D. PLAVSIC, AND N. TRINAJSTIC, *The distance matrix in chemistry*, J. Math. Chem., 11 (1992), pp. 233–258.
- [11] S. MUKWEMBI, *Bounds on Distances in Graphs*, Ph.D. thesis, University of KwaZulu-Natal, Durban, South Africa, 2006.
- [12] J. PLESNÍK, *On the sum of all distances in a graph or digraph*, J. Graph Theory, 8 (1984), pp. 1–24.
- [13] D. H. ROUVRAY, *The modeling of chemical phenomena using topological indexes*, J. Comput. Chem., 8 (1987), pp. 470–480.
- [14] L. SOLTÉS, *Transmission in graphs: A bound and vertex removing*, Math. Slovaca, 41 (1991), pp. 11–16.
- [15] H. WIENER, *Structural determination of paraffin boiling points*, J. Amer. Chem. Soc., 69 (1947), pp. 17–20.
- [16] S. WUCHTY AND P. F. STADLER, *Centers of complex networks*, J. Theoret. Biol., 223 (2003), pp. 45–53.

TREE-RELATED WIDTHS OF GRAPHS AND HYPERGRAPHS*

ISOLDE ADLER[†]

Abstract. A *hypergraph pair* is a pair (G, H) where G and H are hypergraphs on the same set of vertices. We extend the definitions of *hypertree-width* [G. Gottlob, N. Leone, and F. Scarcello, *J. Comput. System Sci.*, 64 (2002), pp. 579–627] and *generalized hypertree-width* [G. Gottlob, N. Leone, and F. Scarcello, *J. Comput. System Sci.*, 66 (2003), pp. 775–808] from hypergraphs to hypergraph pairs. We show that for constant k the problem of deciding whether a hypergraph pair has generalized hypertree-width $\leq k$, is equivalent to the *hypergraph sandwich problem* (HSP) [A. Lustig and O. Shmueli, *J. Algorithms*, 30 (1999), pp. 400–422]. It was recently proved in [G. Gottlob, Z. Miklós, and Th. Schwentick, *Proceedings of the Symposium on Principles of Database Systems (PODS07)*] that the HSP is NP-complete. For constant k there is a polynomial time algorithm that decides whether a given hypergraph pair has hypertree-width $\leq k$. (For hypertree-width of hypergraphs, this was shown in [G. Gottlob, N. Leone, and F. Scarcello, *J. Comput. System Sci.*, 64 (2002), pp. 579–627].) It follows that the HSP is solvable in polynomial time for inputs (G, H) satisfying: $\text{ghw}(G, H) \leq 1$ if and only if $\text{hw}(G, H) \leq 1$. Besides this practical interest, hypergraph pairs serve as a tool for giving a common proof for the game theoretic characterizations of tree-width [P. D. Seymour and R. Thomas, *J. Combin. Theory Ser. B*, 58 (1993), pp. 22–33] and hypertree-width [G. Gottlob, N. Leone, and F. Scarcello, *J. Comput. System Sci.*, 66 (2003), pp. 775–808]. Furthermore, they enable us to show a *compactness property* of generalized hypertree-width for a large class of hypergraphs, the hypergraphs with *finite character*. Finally, we present two examples showing that neither hypertree-width of hypergraph pairs nor hypertree-width of hypergraphs has the *compactness property*.

Key words. hypergraphs, hypertree-width, hypergraph sandwich problem, tree-width, infinite hypergraphs

AMS subject classifications. 05C65, 05C85, 05C75

DOI. 10.1137/050623395

1. Introduction. The well-studied notion of tree-width of a graph (or a hypergraph) is important for various reasons, one of which is that problems that are NP-complete¹ in general become polynomially solvable when restricted to graphs of bounded tree-width (see [5]). Similarly, restriction to hypergraphs with bounded hypertree-width (hw) or bounded generalized hypertree-width (ghw) yields large classes of polynomially solvable instances of problems that are NP-complete in general (see [7]). There are classes of hypergraphs with bounded (generalized) hypertree-width, whose tree-width is unbounded, and where the tree-width of the incidence graphs is also unbounded (see [3, Example 2]).

This paper is organized as follows: section 2 contains the basic definitions, some observations, and an example of a hypergraph pair with hypertree-width 2 and generalized hypertree-width 1.

In section 3, we reformulate the hypergraph sandwich problem (HSP) in terms of generalized hypertree-width of a hypergraph pair. Gottlob, Miklós, and Schwentick [9] recently proved that the HSP is NP-complete. Therefore it is of interest to find tractable restrictions of the HSP. For constant k there is a polynomial time algorithm that decides whether a given hypergraph pair has hypertree-width $\leq k$. (For

*Received by the editors January 27, 2005; accepted for publication (in revised form) June 12, 2007; published electronically February 15, 2008.

<http://www.siam.org/journals/sidma/22-1/62339.html>

[†]Mathematisches Institut, Eckerstr. 1, D-79102 Freiburg, Germany (isolde.adler@math.uni-freiburg.de).

¹See [6] for definitions.

hypertree-width of *hypergraphs*, this was shown in [7].) It follows that the HSP is solvable in polynomial time for inputs (G, H) satisfying: $\text{ghw}(G, H) \leq 1$ if and only if $\text{hw}(G, H) \leq 1$.

Section 4 gives a characterization of the hypertree-width of a (possibly infinite) hypergraph pair (G, H) by the number of marshals necessary to catch the robber in the monotone robber and marshals game played on (G, H) . As a corollary, we obtain the game theoretic characterizations of the well-studied notion of tree-width (tw) [14] and of hypertree-width of hypergraphs [8]. We use the game theoretic characterization in section 6 for constructing examples with special properties.

In section 5 we define the notion of *finite character*, and we show that for hypergraph pairs with finite character generalized hypertree-width is *compact*, i.e.,: a hypergraph pair (G, H) with finite character has generalized hypertree-width $\leq k$ if and only if every finite induced subhypergraph pair (G_0, H_0) of (G, H) has generalized hypertree-width $\leq k$. As special cases we obtain compactness of tree-width of graphs² and compactness of generalized hypertree-width of hypergraphs with finite character.

Finally, section 6 contains two examples, one showing that hypertree-width of hypergraph pairs is not compact, and the other showing that hypertree-width of hypergraphs is not compact. Both examples have hyperedges of size at most 3, which implies that they have finite character.

2. Preliminaries.

2.1. Graphs, hypergraphs, and hypergraph pairs. A *graph* is a pair $G = (V(G), E(G))$ where $V(G)$ is a nonempty set of *vertices*, and $E(G) \subseteq \mathcal{P}_{=2}(V(G))$ is the set of *edges* of G . Thus the edges of G are two-element subsets of $V(G)$. A *clique* is a complete graph. A *forest* is a graph without cycles, and a *tree* is a connected forest. A *directed tree* is a rooted tree where all the edges are directed away from the root.

A *hypergraph* is a pair $H = (V(H), E(H))$, consisting of a nonempty set $V(H)$ of *vertices*, and a set $E(H) \subseteq \mathcal{P}_{<\omega}(V(H))$ of finite subsets of $V(H)$, the *hyperedges* of H . Let $X \subseteq V(H)$. Then $H[X] := (X, \{e \cap X \mid e \in E(H)\})$ is the *subhypergraph of H induced by X* . H_0 is an *induced subhypergraph of H* if $H_0 = H[X]$ for some $X \subseteq V(H)$. For a hypergraph H the *underlying graph of H* is the graph $\underline{H} = (V(H), \{\{v, w\} \mid \text{there exists an } e \in E(H) \text{ such that } \{v, w\} \subseteq e\})$.

A *hypergraph pair* is a pair (G, H) , where G and H are hypergraphs and $V(G) = V(H)$. A hypergraph pair (G_0, H_0) is an *induced subhypergraph pair of (G, H)* , if G_0 is an induced subhypergraph of G and H_0 is an induced subhypergraph of H .

In this paper all graphs are connected and all hypergraphs H satisfy $\bigcup E(H) = V(H)$. Furthermore we only consider hypergraph pairs (G, H) , where G is connected (i.e., $\underline{G}$ is connected).

2.2. Decompositions and acyclicity. Let H be a hypergraph. A *tree-decomposition of H* is a tuple (T, B) , consisting of a directed tree T and a family $B := (B_t)_{t \in V(T)}$ of finite subsets of $V(T)$, the *pieces of T* , satisfying

1. For each edge $e \in E(H)$ there exists $t \in V(T)$, such that $e \subseteq B_t$. (t covers e .)
2. For each $v \in V(H)$ the set $\{t \in V(T) \mid v \in B_t\}$ is connected in T .

²This was first proved by Thomas in [15]. See [16] for a shorter proof.

The *width* of a tree-decomposition of H is

$$w(T, B) := \sup \{ |B_t| - 1 \mid t \in V(T) \} \in \mathbb{N} \cup \{\infty\}.$$

The *tree-width* of H is

$$\text{tw}(H) := \min \{ w(T, B) \mid (T, B) \text{ is a tree-decomposition of } H \} \in \mathbb{N} \cup \{\infty\}.$$

Condition 2 will be referred to as the “connectedness condition.” Since $V(H) \neq \emptyset$ and $\bigcup E(H) = V(H)$, condition 1 implies that a tree-decomposition (T, B) of H always satisfies $w(T, B) \geq 0$. It is not until the definition of a hypertree decomposition that we need T to be a *directed* tree. Nevertheless, it is convenient to require T to be directed.

REMARK 1. Let $C \subseteq \underline{H}$ be a finite clique, and let (T, B) be a tree-decomposition of H . Then there exists a node $t \in V(T)$ with $V(C) \subseteq B_t$.

For a directed tree T and $t \in V(T)$, the maximal subtree of T with root t is denoted by T_t . For a tree-decomposition (T, B) we set $B_{T_t} := \bigcup_{s \in V(T_t)} B_s$.

Proof. Let (T, B) be a tree-decomposition of H . Starting with the root r of T we find a path leading us to a node $t \in V(T)$ with $V(C) \subseteq B_t$ as follows:

- Set $p_0 := r$. If $V(C) \subseteq B_r$ we are finished. Otherwise, connectedness shows that for each $v \in V(C) \setminus B_r$ there exists a unique successor s_v of r such that $v \in B_{T_{s_v}}$.
- If $u \in V(C) \setminus B_r$ and $u \neq v$, then the edge $\{u, v\}$ must be covered by some $s' \in V(T)$.
- Connectedness implies that $s' \in V(T_{s_v})$ and hence $s_u = s_v$. We set $p_1 := s_v$. By connectedness, $(B_r \cap V(C)) \subseteq (B_{p_1} \cap V(C))$.

Suppose the path $p_0, p_1, \dots, p_n$ is already defined. If $V(C) \subseteq B_{p_n}$, then we are finished. Otherwise, by the same arguments as above, there exists a unique successor s of p_n with $V(C) \setminus B_{p_n} \subseteq B_{T_s}$. We set $p_{n+1} := s$. By connectedness, $(B_{p_n} \cap V(C)) \subseteq (B_s \cap V(C))$.

Since $|V(C)| < \omega$, after finitely many steps the path ends at a node p_m with $V(C) \subseteq B_{p_m}$. $\square$

Since every hyperedge of a hypergraph H induces a clique in $\underline{H}$, with Remark 1 it is easy to see that every hypergraph H satisfies $\text{tw}(H) = \text{tw}(\underline{H})$. Hence the tree-width of H only depends on $\underline{H}$.

Let (G, H) be a hypergraph pair. A *hypertree decomposition* of (G, H) is a triple (T, B, C) , where

1. (T, B) is a tree-decomposition of G , and $C := (C_t)_{t \in V(T)}$ is a family of subsets of $E(H)$, such that:
2. Each $t \in V(T)$ satisfies $B_t \subseteq \bigcup C_t$.
3. Each $t \in V(T)$ satisfies $(\bigcup C_t) \cap B_{T_t} \subseteq B_t$.

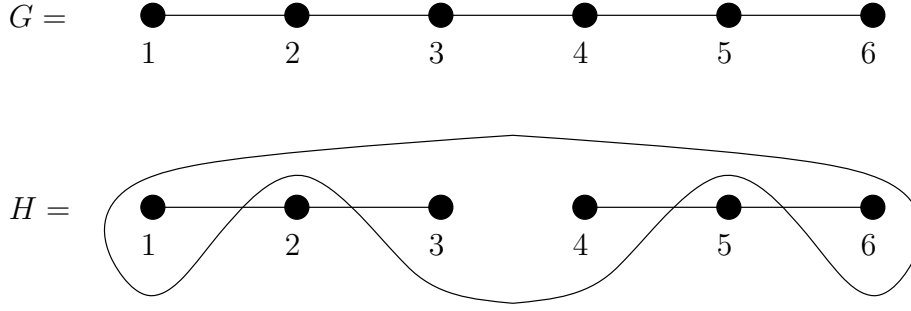
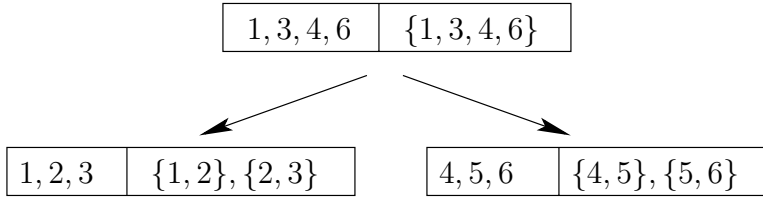
The *width* of a hypertree decomposition of (G, H) is

$$w(T, B, C) := \sup \{ |C_t| \mid t \in V(T) \} \in \mathbb{N} \cup \{\infty\}.$$

The *hypertree-width* of (G, H) is

$$\text{hw}(G, H) := \min \{ w(T, B, C) \mid (T, B, C) \text{ is a hypertree decomposition of } (G, H) \}.$$

By dropping condition 3 from the definition of hypertree decomposition of (G, H) , we obtain a *generalized hypertree decomposition* $(T, B, C,)$ of (G, H) and the *generalized hypertree-width* of (G, H) , $\text{ghw}(G, H)$.


 FIG. 2.1. A hypergraph pair (G, H) .

 FIG. 2.2. A width 2 hypertree decomposition of (G, H) from Figure 2.1.

Note that by definition, $\text{ghw}(G, H) \leq \text{hw}(G, H)$. Furthermore, every hypergraph pair (G, H) satisfies $\text{hw}(G, H) = \text{hw}(\underline{G}, H)$ and $\text{ghw}(G, H) = \text{ghw}(\underline{G}, H)$.

Figure 2.1 shows a hypergraph pair (G, H) with

$$V(G) = V(H) = \{1, 2, 3, 4, 5, 6\},$$

$$E(G) = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}, \{5, 6\}\}, \text{ and}$$

$$E(H) = \{\{1, 2\}, \{2, 3\}, \{1, 3, 4, 6\}, \{4, 5\}, \{5, 6\}\}.$$

For convenience, the hyperedges containing exactly two vertices are shown as graph edges. Figure 2.2 shows a width 2 hypertree decomposition of (G, H) . Figure 2.3 shows a width 1 generalized hypertree decomposition of (G, H) . Each tree node t is depicted with B_t on the left-hand side and C_t on the right-hand side. Actually, these decompositions are optimal in the sense that $\text{hw}(G, H) = 2$ and $\text{ghw}(G, H) = 1$. This is left to the reader as an exercise.

For a hypergraph H , define the *hypertree-width* of H by $\text{hw}(H) := \text{hw}(H, H)$ and the *generalized hypertree-width* of H by $\text{ghw}(H) := \text{ghw}(H, H)$. In this way we get exactly the definitions from [8] of (generalized) hypertree-width of a hypergraph. In particular, $\text{hw}(H) \leq k$ implies $\text{ghw}(H) \leq k$. In general the converse does not hold.³

It is easy to see that we also obtain tree-width as a special case.

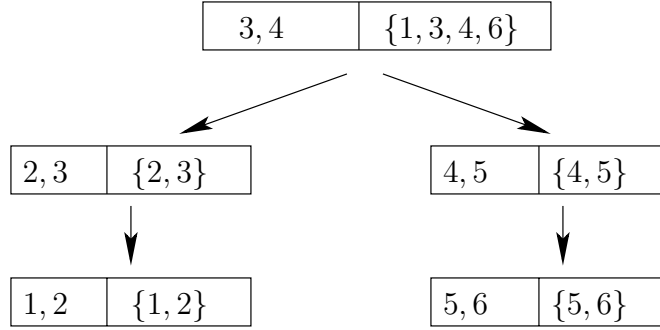
REMARK 2. Let (G, H) be a hypergraph pair with $H = \{\{v\} \mid v \in V(G)\}$, then

$$\text{hw}(G, H) = \text{tw}(G) + 1.$$

A tree-decomposition (T, B) of H is *small*, if all $s, t \in V(T)$ with $s \neq t$ satisfy $B_s \not\subseteq B_t$.

REMARK 3. If a finite hypergraph pair (G, H) has a generalized hypertree decomposition of width $k \in \mathbb{N}$, then (G, H) also has a generalized hypertree decomposition (T, B, C) of width k , such that (T, B) is small.

³A counterexample can be found in [1].

FIG. 2.3. A width 1 generalized hypertree decomposition of (G, H) from Figure 2.1.

Proof. Let (T', B', C') be a generalized hypertree decomposition of width k for the finite hypergraph pair (G, H) . We may assume that T is finite. If $\{s, t\} \in E(T)$ with $B_s \subseteq B_t$, remove s , B_s , and C_s from (T, B, C) and connect each neighbor $t' \neq t$ of s by an edge to t . It is easy to see that thus we obtain a new generalized hypertree decomposition (T'', B'', C'') of width k for (G, H) . Since T is finite, we can repeat this procedure finitely often until we obtain a generalized hypertree decomposition (T, B, C) of width k for (G, H) s.t. (T, B) is small. $\square$

Note that this is not true for infinite hypergraph pairs: Let $H = (\omega, \mathcal{P}_{<\omega}(\omega))$. Then $\text{ghw}(\underline{H}, H) = 1$ (take a one way infinite path with an increasing sequence of pieces $B_t = \bigcup C_t$), but $(\underline{H}, H)$ has no small generalized hypertree decomposition of width 1: Let (T, B, C) be a generalized hypertree decomposition of $(\underline{H}, H)$. Since all B_t are finite, for each B_t there is an edge $e \in E(H)$ s.t. $B_t \subsetneq e$ and e is covered by some $t' \in V(T)$.

A hypergraph H is *acyclic* if its hyperedges can be arranged as nodes of a tree T so that for every vertex $v \in V(H)$, the subtree of T defined by the nodes containing v is connected (cf. [13]).

LEMMA 4. *Let H be a finite hypergraph. The following statements are equivalent:*

- (i) $\text{ghw}(H) \leq 1$.
- (ii) $\text{hw}(H) \leq 1$.
- (iii) H is *acyclic*.

The equivalence of (ii) and (iii) was already observed in [7].

Proof of Lemma 4. It is easy to see that (iii) implies (ii), and that (ii) implies (i).

(i) $\Rightarrow$ (iii): Let (T, B, C) be a generalized hypertree decomposition for H of width at most 1. By Remark 3 we may assume that (T, B) is small. Then $B_t = \bigcup C_t$ for all $t \in T$. (Otherwise $B_t \subsetneq \bigcup C_t$ and $C_t = \{e\}$ for some $e \in E$. But there must be a node $s \in T$ such that $e \subseteq B_s$. Hence $B_t \subsetneq B_s$, in contradiction to (T, B) being small.) We may assume that $t = B_t$ for every $t \in T$. Now we can extend T to a tree T' with $V(T') = E(H)$ and such that each $e \in E(H) \setminus V(T)$ is attached to some $t \in T$ with $e \in t = B_t$. $\square$

3. The hypergraph sandwich problem. In this section, all graphs and hypergraphs are finite. We relate generalized hypertree-width of hypergraph pairs to the HSP.

For an integer $k \geq 0$ the problem k -GHW is defined as follows: Given a finite hypergraph H , decide whether $\text{ghw}(H) \leq k$. In [9], Gottlob, Miklós, and Schwentick proved that k -GHW is NP-complete. In [7] it was shown that k -HW, the correspond-

ing problem of deciding whether $\text{hw}(H) \leq k$, is in P (for fixed k). Hence, by Lemma 4, 1-GHW is in P.

Let $k\text{-GHW}^{\text{pair}}$ ($k\text{-HW}^{\text{pair}}$) be the extension of $k\text{-GHW}$ ($k\text{-HW}$) to hypergraph pairs: Given a finite hypergraph pair (G, H) , decide whether $\text{ghw}(G, H) \leq k$ ($\text{hw}(G, H) \leq k$). For fixed k , the problem $k\text{-HW}^{\text{pair}}$ is also solvable in polynomial time: The proof of $k\text{-HW} \in \text{P}$ from [7] generalizes to hypergraph pairs. We will give a different proof in the next section (Corollary 16). Both algorithms can easily be modified in such a way that if $\text{hw}(G, H) \leq k$, then they return a hypertree decomposition of width at most k .

For a hypergraph H and $k \in \mathbb{N}$, we define the hypergraph

$$H^{\leq k} = (V(H), \{e_1 \cup \dots \cup e_i \mid i \leq k, e_j \in E(H) \text{ for all } j = 1, \dots, i\}).$$

Note that for fixed k and input H , we can compute $H^{\leq k}$ in polynomial time.

LEMMA 5. *Let (G, H) be a hypergraph pair. Then:*

- (i) $\text{ghw}(G, H) \leq k$ if and only if $\text{ghw}(G, H^{\leq k}) \leq 1$.
- (ii) $\text{hw}(G, H) \leq k$ if and only if $\text{hw}(G, H^{\leq k}) \leq 1$.

Proof. Let (T, B, C) be a (generalized) hypertree decomposition of width $\leq k$ for (G, H) . For every $t \in V(T)$ replace the hyperedges in C_t by their union. Conversely, if (T, B, C) is a (generalized) hypertree decomposition of width ≤ 1 for $(G, H^{\leq k})$, replace each edge $e_1 \cup \dots \cup e_i$, in C_t by the edges $e_1, \dots, e_i$ ($i \leq k$). $\square$

Let (G, H) be a hypergraph pair. We write $G \leq H$, if for every hyperedge $e_1 \in E(G)$ there exists a hyperedge $e_2 \in E(H)$ such that $e_1 \subseteq e_2$. For example, $\underline{H} \leq H$ for every hypergraph H .

The *hypergraph sandwich problem* is the following problem [13]: Given a hypergraph pair (G, H) , decide whether an acyclic hypergraph A exists such that $G \leq A \leq H$. (So, A is “sandwiched” between G and H .) Clearly, the HSP is in NP. It was proved in [9] that it is actually NP-complete.

LEMMA 6. *Let (G, H) be a hypergraph pair. The following are equivalent:*

- (i) *There is an acyclic hypergraph A such that $G \leq A \leq H$.*
- (ii) $\text{ghw}(G, H) \leq 1$.

Proof. Let A be acyclic with $G \leq A \leq H$. By the obvious part of Lemma 4 (which holds even in the infinite case), A has a generalized hypertree decomposition (T, B, C) of width ≤ 1 . For every $t \in V(T)$ we set $C'_t := \{e'\}$, where $C_t = \{e\}$ and $e \subseteq e' \in E(H)$. Thus we obtain a generalized hypertree decomposition (T, B, C') of (G, H) of width ≤ 1 .

Conversely, let (T, B, C) be a generalized hypertree decomposition of (G, H) of width ≤ 1 . We may assume that $B_t \neq \emptyset$ for all $t \in V(T)$. Define a hypergraph A with $V(A) = V(G)$ and $E(A) = \{B_t \mid t \in V(T)\}$. It is easy to see that A is acyclic. Since every edge $e \in E(G)$ is covered by some B_t we have $G \leq A$. On the other hand, by condition 2 of the definition of a hypertree decomposition of (G, H) , all $t \in V(T)$ satisfy $B_t \subseteq \bigcup C_t = e'$ for some $e' \in E(H)$. Hence $A \leq H$. $\square$

Gottlob, Scarcello, and Leone pinpointed in [8] that there is a polynomial time reduction from $k\text{-GHW}$ to the HSP. We actually have the following theorem.

THEOREM 7. *The following problems are NP-complete:*

- (i) HSP,
- (ii) $1\text{-GHW}^{\text{pair}}$, and
- (iii) $k\text{-GHW}^{\text{pair}}$.

Proof. (i) is proved in [9]. By Lemma 6, (i) and (ii) are polynomial time equivalent. (ii) and (iii) are also polynomial time equivalent: (ii) $\Rightarrow$ (iii) is easy with Lemma 5. $\square$

Instead of solving the HSP with input (G, H) , or, equivalently, $1\text{-GHW}^{\text{pair}}$, we can solve $1\text{-HW}^{\text{pair}}$ with input (G, H) in polynomial time. If $\text{hw}(G, H) \leq 1$, then we get a hypertree decomposition for (G, H) of width at most 1. As in Lemma 6, the hypertree decomposition gives rise to a solution for the HSP. Thus the HSP is solvable in polynomial time for inputs (G, H) satisfying $\text{ghw}(G, H) \leq 1$ if and only if $\text{hw}(G, H) \leq 1$.

Unfortunately, if $\text{hw}(G, H) > 1$, we know nothing about the HSP: By [2, Theorem 3.3.18], for each positive integer k there is a hypergraph pair (G_k, H_k) with $\text{ghw}(G_k, H_k) = 1$ and $\text{hw}(G_k, H_k) = k$.

4. Game theoretic characterizations. This section contains a generalization of the robber and marshals game of [8] to (possibly infinite) hypergraph pairs. The main result of this section (Theorem 9) is that even for infinite hypergraph pairs (G, H) , the minimum number of marshals necessary to catch the robber on (G, H) equals $\text{hw}(G, H)$. This generalizes both the game theoretic characterization of tree-width [14] and the game theoretic characterization of hypertree-width [8].

4.1. Games. Let (G, H) be a hypergraph pair, and let $k \geq 0$ be an integer. The *monotone robber and marshals game on (G, H)* , $\text{mon-mars}((G, H), k)$, is played by two players: The *marshal player* plays k marshals on H , and the *robber player* plays the robber on G . A *position* of $\text{mon-mars}((G, H), k)$ is a pair (X, r) , where $X \in \mathcal{P}_{\leq k}(E(H))$ and $r \in V(G) \setminus \bigcup X$. At the beginning of the game, the marshal player chooses an $X \in \mathcal{P}_{\leq k}(E(H))$. The robber player responds by choosing an $r \in V(G) \setminus \bigcup X$. In each step of the game, say in position (X, r) , the marshal player chooses a set $Y \in \mathcal{P}_{\leq k}(E(H))$. Then the robber player chooses an $r' \in V(G) \setminus \bigcup Y$ that is connected to r by a path in $G \setminus (\bigcup X \cap \bigcup Y)$.

Let (X, u) and (X, v) be game positions with $u \in R$ and $v \in R$, where R is a connected component of $G \setminus \bigcup X$. Now suppose the marshals choose the new position Y . Let (Y, r) be a position which the robber can reach from (X, u) . Then he can reach (Y, r) from (X, v) as well (and vice versa). For this reason a game position will be represented by (X, R) instead of (X, u) , and we define the following.

For $W \subseteq V(G)$ a connected component $R \subseteq V(G)$ of $G \setminus W$ is called an *escape space (with respect to W)*. For $R \subseteq V(G)$ we call

$$\mathcal{R}(W'; W, R) = \{R' \mid R' \text{ is an escape space w.r.t. } W' \text{ and there is a path from } R \text{ to } R' \text{ in } G \setminus (W \cap W')\}$$

the *set of possible escape spaces w.r.t. W'* .

With these definitions, suppose the game is at position (X, R) and the marshals choose a set $Y \in \mathcal{P}_{\leq k}(E(H))$. Then $\mathcal{R}(\bigcup Y; \bigcup X, R)$ is the set of escape spaces w.r.t. $\bigcup Y$ that can be reached by the robber.

Suppose the game is at position (X, R) , where R is an escape space in G w.r.t. $\bigcup X$. The marshal player makes a *monotone move*, if she chooses a set $Y \in \mathcal{P}_{\leq k}(E(H))$ such that each $R' \in \mathcal{R}(\bigcup Y; \bigcup X, R)$ satisfies $R' \subsetneq R$.

A *game* is a sequence of moves of the marshal player and the robber player moving alternately. The marshal player is only allowed to make monotone moves. If there is a step in which the marshal player cannot make a monotone move, then the robber player wins.

The marshal player wins if she can make sure that for every $v \in V(G)$ there is a move in the game such that after this move the robber can never visit v again. Note that this includes the case that the robber cannot move at all.

In the case that (G, H) is finite, we can equivalently define that the marshal player wins, if at some point of the game the robber player cannot move any more. Otherwise, the robber player wins. The marshal player has a *winning strategy* if she can win every possible game.

Let (G, H) be a hypergraph pair. A *marshal strategy* in $\text{mon-mars}((G, H), k)$ is a triple (T, σ, ρ) satisfying:

- (a) T is a directed tree.
- (b) $\sigma : V(T) \rightarrow \mathcal{P}_{\leq k}(E(H))$.
- (c) $\rho : V(T) \rightarrow \mathcal{P}(V(G))$.
- (d) The root r of T satisfies $\rho_r = V(G)$. Every node s of T which is not the root satisfies $\rho_s \in \mathcal{R}(\bigcup \sigma_t; \bigcup \sigma_t^{\text{pred}}, \rho_t)$, where $t = \text{pred}(s)$ and $\text{pred}(s)$ denotes the unique predecessor of s in T .

Here we have made use of the abbreviation

$$\sigma_t^{\text{pred}} = \begin{cases} \emptyset & \text{if } t \text{ is the root of } T, \\ \sigma_{\text{pred}(t)} & \text{otherwise.} \end{cases}$$

- (e) For each $t \in V(T)$ and each $R \in \mathcal{R}(\bigcup \sigma_t; \bigcup \sigma_t^{\text{pred}}, \rho_t)$ there is exactly one successor $s \in V(T)$ with $\rho_s = R$.

- (f) (T, σ, ρ) is *monotone*, i.e. for all $t \in V(T) \setminus \{r\}$ the condition $\rho_t \subsetneq \rho_{\text{pred}(t)}$ holds.

A *branch* D of T is the vertex set $D \subseteq V(T)$ of a maximal path in T that begins with the root. (T, σ, ρ) is a *winning strategy* for $\text{mon-mars}((G, H), k)$, if all infinite branches D of T satisfy $\bigcap_{t \in D} \rho_t = \emptyset$. In the case that (G, H) is finite, this condition is always satisfied.

Let (G, H) be a hypergraph pair. The *monotone marshal-width* of (G, H) is

$$\text{mon-mw}(G, H) := \min \{k \mid \text{the marshals have a winning strategy in } \text{mon-mars}((G, H), k)\} \in \mathbb{N} \cup \{\infty\}.$$

Figure 4.1 shows a monotone winning strategy for two marshals and a nonmonotone winning strategy for one marshal on the hypergraph pair (G, H) of Figure 2.1. The nodes of a strategy tree are the marshals' moves and the edges are labeled by the corresponding robber's escape spaces.

The strategy on the right-hand side is not monotone: after the marshals' first move, the vertex 1 is no longer in the robber's escape space. If the robber now moves to the vertex 2, the marshals move to $\{2, 3\}$ and the robber can reach 1 again.

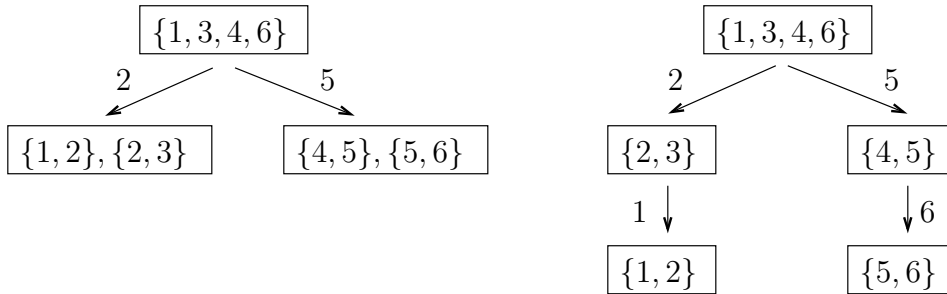


FIG. 4.1. A monotone winning strategy for two marshals and a nonmonotone winning strategy for one marshal on the hypergraph pair (G, H) of Figure 2.1.

For the monotone robber and cops game as defined⁴ in [14] we set

$$\text{mon-cw}(H) := \min\{k \in \mathbb{N} \mid k \text{ cops can win the monotone} \\ \text{robber and cops game played on } H\} \in \mathbb{N} \cup \{\infty\}.$$

For the monotone robber and marshals game as defined in [8] we set

$$\text{mon-mw}(H) := \min\{k \mid k \text{ marshals can win the monotone} \\ \text{robber and marshals game played on } H\} \in \mathbb{N} \cup \{\infty\}.$$

It is easy to see that the monotone robber and cops game and the monotone robber and marshals game on hypergraphs are special cases of the monotone robber and marshals game played on a hypergraph pair (G, H) :

REMARK 8. (a) Let (G, H) be a hypergraph pair where $E(H) = \{\{v\} \mid v \in V(G)\}$. Then $\text{mon-mars}((G, H), k)$ is the monotone robber and cops game played on G , and $\text{mon-mw}(G, H) = \text{mon-cw}(G)$.

(b) Let H be a connected hypergraph. Then $\text{mon-mars}((H, H), k)$ is the monotone robber and marshals Game played on H , and $\text{mon-mw}(H, H) = \text{mon-mw}(H)$.

4.2. Equivalence of decompositions and strategies. In this section we will prove the following theorem.

THEOREM 9. Let (G, H) be a hypergraph pair, and let $k \geq 0$ be an integer. Then $\text{hw}(G, H) \leq k$ if and only if there is a winning strategy⁵ for $\text{mon-mars}((G, H), k)$.

Together with Remarks 8 and 2 this implies the following corollary.

COROLLARY 10. Let H be a hypergraph, and let $k \geq 0$ be an integer.

(a) (see Seymour and Thomas [14]) $\text{tw}(H) + 1 \leq k$ if and only if k cops can win the monotone robber and cops game on H .

(b) (see Gottlob, Scarcello and Leone [8]) $\text{hw}(H) \leq k$ if and only if k marshals can win the monotone robber and marshals game on H .

To prove the “if” part of Theorem 9, we make use of the following lemma.

LEMMA 11. Let (T, σ, ρ) be a winning strategy for $\text{mon-mars}((G, H), k)$.

(i) For every $v \in V(G)$ there is one and only one $t_v \in V(T)$ with

$$v \in \bigcup \sigma_{t_v} \cap \rho_{t_v}.$$

(ii) For every $e \in E(G)$ there is a node $t_e \in V(T)$ with

$$e \subseteq \bigcup \sigma_{t_e} \text{ and } e \cap \rho_{t_e} \neq \emptyset.$$

Proof. The proof is left to the reader. $\square$

Proof of the “if” part of Theorem 9. Let (T, σ, ρ) be a winning strategy for $\text{mon-mars}((G, H), k)$.

Set $C := \sigma$, and define B_t^{pred} and B_t inductively as follows:

$$B_t^{\text{pred}} := \begin{cases} \emptyset & \text{if } t \text{ is the root of } T, \\ B_{\text{pred}(t)} & \text{otherwise,} \end{cases}$$

$$B_t := \bigcup \sigma_t \cap (B_t^{\text{pred}} \cup \rho_t) \text{ for } t \in V(T).$$

⁴Note that while the monotone robber and cops game in [14] is only defined in the case that H is a graph, that definition can be translated literally from graphs to hypergraphs by replacing the word “graph” by “hypergraph” and “edge” by “hyperedge.” This does not really add any generality, since the robber and cops game on a hypergraph is essentially the same as the game on its underlying graph.

⁵Note that the winning strategy is monotone by definition.

We now show that (T, B, C) is a hypertree decomposition of (G, H) of width $\leq k$: Since $|C_t| = |\sigma_t|$ for all $t \in V(T)$, it is sufficient to prove that (T, B, C) satisfies the three conditions from the definition of a hypertree decomposition of (G, H) .

Condition 1. (T, B) is a tree-decomposition of G : We first show that every hyperedge $e \in E(H)$ is covered by some B_t . Choose $t = t_e$ for e according to Lemma 11 (ii). Then $e \subseteq B_{t_e} = \bigcup \sigma_{t_e} \cap (B_{t_e}^{\text{pred}} \cup \rho_{t_e})$: By the choice of t_e , $e \subseteq \bigcup \sigma_{t_e}$. Hence it remains to show that $e \subseteq B_{t_e}^{\text{pred}} \cup \rho_{t_e}$, or, equivalently, $e \setminus \rho_{t_e} \subseteq B_{t_e}^{\text{pred}}$. For showing this, choose for $v \in e \setminus \rho_{t_e}$ a node t_v according to Lemma 11 (i). Then $v \in B_{t_v}$ and t_v lies on the path from the root of T to t_e . For all $s \neq t_v$ on the path from t_v to t_e , v is a neighbor of $\rho_{t_e} \subseteq \rho_s$ and hence v is a neighbor of ρ_s . Therefore $v \in \bigcup \sigma_s$ for all $s \neq t_e$ on the path from t_v to t_e . Hence $v \in B_s$ for all $s \neq t_e$ on the path from t_v to t_e , in particular $v \in B_{t_e}^{\text{pred}}$.

Suppose for $v \in V(G)$ the set $Z = \{t \in V(T) \mid v \in B_t\}$ is not connected. Then Z has (at least) two connected components Z_1 and Z_2 . Let z_1 be the vertex from Z_1 with the shortest path from the root r to z_1 . Likewise, choose $z_2 \in Z_2$ with the shortest path from r to z_2 . Then $z_1 \neq z_2$ and both vertices satisfy Lemma 11 (i), a contradiction to uniqueness.

Condition 2. $B_t \subseteq \bigcup \sigma_t = \bigcup C_t$.

Condition 3. Claim: $B_{T_t} \subseteq B_t^{\text{pred}} \cup \rho_t$.

The claim implies condition 3: $\bigcup C_t \cap B_{T_t} = \bigcup \sigma_t \cap B_{T_t} \subseteq \bigcup \sigma_t \cap (B_t^{\text{pred}} \cup \rho_t) = B_t$.

Proof of the claim. Assuming that $v \in B_{T_t} \setminus (B_t^{\text{pred}} \cup \rho_t)$ exists, choose $s \in V(T)$ for v satisfying $v \in \bigcup \sigma_s \cap \rho_s \subseteq B_s$ as in Lemma 11 (i). Since $v \notin B_t^{\text{pred}}$, the connectedness condition of a tree-decomposition implies $s \in V(T_t)$. The monotonicity of the strategy implies $\rho_s \subseteq \rho_t$, hence $v \in \rho_t$. This is contrary to the choice of v . $\square$

To prove the “only if” part, we construct a winning strategy for k marshals from a hypertree decomposition of width $\leq k$ of (G, H) . For this purpose we introduce *weak strategies*. Let (T, σ, ρ) satisfy all the axioms of a strategy except for (f), and let r be the root of T . Then (T, σ, ρ) is called a *weak strategy* if every $t \in V(T) \setminus \{r\}$ satisfies $\rho_t \subseteq \rho_{\text{pred}(t)}$. *Weak winning strategies* are defined analogously to winning strategies, i.e. a weak strategy (T, σ, ρ) is a *weak winning strategy* for $\text{mon-mars}((G, H), k)$, if all infinite branches D of T satisfy $\bigcap_{t \in D} \rho_t = \emptyset$.

Clearly, every winning strategy is a weak winning strategy. Conversely we have the following lemma.

LEMMA 12. *Every weak winning strategy for $\text{mon-mars}((G, H), k)$ gives rise to a winning strategy.*

Proof. Intuitively, we modify a weak winning strategy by skipping all moves that do not change the robber’s escape space, thus obtaining a winning strategy. Let (T, σ, ρ) be a weak winning strategy for $\leq k$ marshals. Define a function $g : V(T) \rightarrow V(T)$ as follows: Set $g(t) = t$ if all children s of t satisfy $\rho_s \neq \rho_t$. Otherwise, t has exactly one child s and we set $g(t) = s$. Then for all $t \in V(T)$ there is an integer $n \geq 0$ such that $g^n(t) = g^{n+1}(t)$. (Otherwise, $\{g^n(t) \mid n \geq 0\}$ is an end piece of an infinite branch D of T and hence $\bigcap_{s \in D} \rho_s = \rho_t \neq \emptyset$, a contradiction to (T, σ, ρ) being a weak winning strategy.) Define $f : V(T) \rightarrow V(T)$ by setting $f(t) =$ the eventual value $g^n(t)$, for big n .

Now define (T', σ', ρ') by setting

- $V(T') := \{f(t) \mid t \in V(T)\}$,
- $E(T') = \{\{f(s), f(t)\} \mid \{s, t\} \in E(T) \text{ and } f(s) \neq f(t)\}$,
- $\sigma' = \sigma \upharpoonright V(T')$, and $\rho' = \rho \upharpoonright V(T')$.

(T', σ', ρ') is a winning strategy for $\text{mon-mars}((G, H), k)$:

T' is the result of contracting connected subsets of T , so T' is again a directed tree. Hence, T' satisfies the conditions (a), (b), and (c) of a marshal strategy.

The root r' of T' satisfies $\rho_{r'} = \rho_r = V(G)$. Let $s \in V(T')$, $s \neq r'$. We have to show that $\rho_s \in \mathcal{R}(\bigcup \sigma_t; \bigcup \sigma_t^{\text{pred}'}, \rho_t)$, where $t = \text{pred}'(s)$ is the predecessor of s in T' . For $t \in V(T)$ define $f'(t) := t'$, where $t' \in V(T)$ is the node nearest to the root of T such that $f(t) = f(t')$.

1. ρ_s is an escape space w.r.t. $\bigcup \sigma_t$, since $\rho_{f'(s)}$ is an escape space w.r.t. $\bigcup \sigma_t$ and $\rho_s = \rho_{f'(s)}$.

2. There is a path from ρ_t to ρ_s in $G \setminus (\bigcup \sigma_t^{\text{pred}'} \cap \bigcup \sigma_t)$, since $\rho_s \neq \emptyset$, $\rho_s \subseteq \rho_t$ and $\rho_s \subseteq V(G) \setminus \bigcup \sigma_t$.

Together, 1 and 2 show that (d) holds.

For (e) we have to show that $\mathcal{R}(\bigcup \sigma_t; \bigcup \sigma_t^{\text{pred}'}, \rho_t) = \mathcal{R}(\bigcup \sigma_t; \bigcup \sigma_t^{\text{pred}}, \rho_t)$.

For an escape space R w.r.t. X , define

$$\partial R := \{v \in X \mid v \text{ is a neighbor of some } u \in R\}.$$

Since (T, σ, ρ) is weakly monotone, $\partial R \subseteq \bigcup \sigma_p$ for all $p \in V(T)$ with $f(p) = t$. Therefore, $\partial R \subseteq \bigcup \sigma_t \cap \bigcup \sigma_t^{\text{pred}'} \cap \sigma_t^{\text{pred}}$. Thus during the flight from $\sigma_t^{\text{pred}'}$ to σ_t , as well as during the flight from σ_t^{pred} to σ_t , the robber can only move within $\rho_t = \rho_t'$. Hence $\mathcal{R}(\bigcup \sigma_t; \bigcup \sigma_t^{\text{pred}'}, \rho_t) = \mathcal{R}(\bigcup \sigma_t; \bigcup \sigma_t^{\text{pred}}, \rho_t)$.

(f) holds by construction. Moreover, (T', σ', ρ') is a winning strategy since (T, σ, ρ) is a weak winning strategy. $\square$

Thus, it is sufficient to construct a weak winning strategy for k marshals from a hypertree decomposition of width k . Let (T, B, C) be a width k hypertree decomposition of (G, H) . The edge labels yield the positions of the marshals: The marshals move to C_r , where r is the root of T . The robber chooses an escape space R . R satisfies $R \subseteq B_{T_t}$ for exactly one successor t of r . Then the marshals fly to C_t . During the flight, the robber cannot leave R .

Towards a formalization of this intuition, let (T, B) be a tree-decomposition of the hypergraph G , $t \in V(T)$. The *tree component* of t is defined as follows:

$$\beta_t := B_{T_t} \setminus B_t^{\text{pred}}.$$

LEMMA 13. *Let (T, B) be a tree-decomposition of a hypergraph G . Then*

- (i) $\beta_t \subseteq \beta_{\text{pred}(t)}$, if $t \in V(T)$ is not the root of T .
- (ii) Any infinite branch D of T satisfies $\bigcap_{t \in D} \beta_t = \emptyset$.

Proof. (i):

$$\begin{aligned} \beta_t \setminus \beta_{\text{pred}(t)} &= [B_{T_t} \setminus B_t^{\text{pred}}] \setminus [B_{T_{\text{pred}(t)}} \setminus B_{\text{pred}(t)}^{\text{pred}}] \\ &= (B_{T_t} \setminus [B_{T_{\text{pred}(t)}} \setminus B_{\text{pred}(t)}^{\text{pred}}]) \setminus B_t^{\text{pred}} \\ &= ([B_{T_t} \setminus B_{T_{\text{pred}(t)}}] \cup [B_{T_t} \cap B_{\text{pred}(t)}^{\text{pred}}]) \setminus B_t^{\text{pred}} \\ &= [B_{T_t} \cap B_{\text{pred}(t)}^{\text{pred}}] \setminus B_t^{\text{pred}} \\ &= \emptyset, \end{aligned}$$

where the last equation holds by connectedness.

(ii): Suppose $v \in \bigcap_{t \in D} \beta_t = \bigcap_{t \in D} [B_{T_t} \setminus B_t^{\text{pred}}] = \bigcap_{t \in D} B_{T_t} \setminus \bigcup_{t \in D} B_t^{\text{pred}}$. Then all $t \in D$ satisfy $v \notin B_t^{\text{pred}}$. Since D is infinite, this implies that all $t \in D$ satisfy

$v \notin B_t$. Since $v \in V(G)$, there is an $s \in V(T)$ such that $v \in B_s$. Let $t \in V(T)$ be the last vertex of D on the path from the root of T to s . Let t' be the successor of t in D . By assumption, we have $v \in B_{T_{t'}}$. Connectedness implies that $v \in B_t$ a contradiction. $\square$

LEMMA 14. *Let (T, B) be a tree-decomposition of G , and $t \in V(T)$. Let $R \in \mathcal{R}(B_t; B_t^{\text{pred}}, \beta_t)$. Then $R \subseteq \beta_s$ for a successor s of t .*

Proof. Let $v \in R$. Let W be a path in $G \setminus (B_t \cap B_t^{\text{pred}})$ that connects a vertex $u \in \beta_t$ to v . Every edge of W is covered by a node from T_t or from $T \setminus T_t$. But no edge of W can be covered by a node from $T \setminus T_t$, because otherwise a vertex $r \in W$ would be covered by T_t and $T \setminus T_t$, implying $r \in B_t \cap B_t^{\text{pred}}$, a contradiction. Therefore every vertex of W is covered by a node from T_t , and in particular we have $v \in B_{T_t}$. Thus $v \in B_{T_t} \setminus B_t$, showing that $v \in B_{T_s}$ for a successor s of t and hence $v \in \beta_s$. Equally, a neighbor v' of v satisfies $v' \in \beta_{s'}$ for a successor s' of t . Since $v, v' \notin B_t$ and since $\{v, v'\}$ must be covered by some node of T , connectedness implies that $s = s'$. $\square$

The following lemma proves the “only if” part of Theorem 9.

LEMMA 15. *Let (T, B, C) be a hypertree decomposition of (G, H) of width $\leq k$. Then there is a weak winning strategy (T', σ, ρ) of $\text{mon-mars}((G, H), k)$ and a function $f : V(T') \rightarrow V(T)$ such that*

- $f(r') = r$ for the roots r of T and r' of T' ,
- $f(\text{pred}(t')) = \text{pred}(f(t'))$ for all $t' \in V(T') \setminus \{r'\}$,
- $\sigma_{t'} = C_{f(t')}$ for all $t' \in V(T')$,
- $\rho_{t'} \subseteq \beta_{f(t')}$ for all $t' \in V(T')$.

Proof. We define (T', σ, ρ) and f by an informal induction, starting with the root r' of T' . (The function f keeps the induction going.)

We start by choosing a root r' for T' and setting $f(r') := r$, $\sigma_{r'} := C_r$ and $\rho_{r'} := V(G) = \beta_r$.

For each $t' \in V(T')$ we need to make sure that for every $R' \in \mathcal{R}(\bigcup \sigma_{t'}; \bigcup \sigma_{t'}^{\text{pred}}, \rho_{t'})$ there is exactly one successor s' of t' such that $\rho_{t'} = R'$.

Claim 1. There is a successor s of $f(t')$ such that $R' \subseteq \beta_s$.

Proof of Claim 1. Let $t' = \text{pred}(s')$ and $t = f(t')$. $B_t \subseteq \bigcup C_t = \bigcup \sigma_{t'}$ and therefore $R' \subseteq R$ for some connected component R of $G \setminus B_t$. Since $B_t^{\text{pred}} \subseteq \bigcup C_t^{\text{pred}} = \bigcup \sigma_{t'}^{\text{pred}}$ and $\rho_{t'} \subseteq \beta_t$, we have $R \in \mathcal{R}(B_t; B_t^{\text{pred}}, \beta_t)$. By Lemma 14 there is a successor s of t such that $R \subseteq \beta_s$. Therefore $R' \subseteq \beta_s$. Set $f(s') := s$, $\sigma_{s'} := C_s$ and $\rho_{s'} := R'$.

Claim 2. $\rho_{s'} \subseteq \rho_{t'}$.

Proof of Claim 2. Otherwise, there is a $v \in \rho_{s'} \setminus \rho_{t'}$. Let v_0 be the first vertex on a path from $\rho_{t'}$ to v in $G \setminus (\bigcup \sigma_{t'}^{\text{pred}} \cap \bigcup \sigma_{t'})$ such that $v_0 \notin \rho_{t'}$. $v_0 \notin \rho_{t'}$ implies $v_0 \in \bigcup \sigma_{t'}^{\text{pred}}$, and hence $v_0 \notin \bigcup \sigma_{t'}$. Therefore $v_0 \in \bigcup \mathcal{R}(\bigcup \sigma_{t'}; \bigcup \sigma_{t'}^{\text{pred}}, \rho_{t'}) \subseteq \bigcup \mathcal{R}(\bigcup B_t; \bigcup B_t^{\text{pred}}, \beta_t)$. Choose a successor $\hat{s}$ of t according to Lemma 14 such that $v_0 \in \beta_{\hat{s}}$. Since $v_0 \in \bigcup \sigma_{t'}^{\text{pred}} = \bigcup C_t^{\text{pred}}$ and $v_0 \in \beta_{\hat{s}} = B_{T_{\hat{s}}} \setminus B_t \subseteq B_{T_{\text{pred}(t)}}$, condition 3 of a hypertree decomposition implies that $v_0 \in B_t^{\text{pred}}$. But from this it follows that $v_0 \notin B_{T_t} \setminus B_t^{\text{pred}} = \beta_t$, hence $v_0 \in \beta_{\hat{s}} \setminus \beta_t$, a contradiction to Lemma 13 (i).

Since $\rho_{t'} \subseteq \beta_{f(t')}$ for all $t' \in V(T')$, Lemma 13 (ii) shows that all infinite branches D of T' satisfy $\bigcap_{t \in D} \rho_t = \emptyset$ and the proof is completed. $\square$

As an application, we sketch a proof of $k\text{-HW}^{\text{pair}} \in \text{P}$, which uses the game theoretic characterization.

COROLLARY 16. *For every fixed integer $k \geq 0$, the problem $k\text{-HW}^{\text{pair}}$ is solvable in polynomial time.*

Proof. Given a finite hypergraph pair (G, H) , the algorithm has to decide whether

$\text{hw}(G, H) \leq k$ holds. By game theoretic characterization (Theorem 9) this is equivalent to deciding whether $\text{mon-mw}(G, H) \leq k$ holds. There is a straightforward dynamic programming algorithm that decides whether k marshals have a winning strategy.

List all pairs (X, r) with $X \subseteq E(H)$, $|X| \leq k$ and $r \in V(G)$. Since k is fixed, this is possible in polynomial time. Then mark all pairs (X, r) with $r \in \bigcup X$ as *winning positions*. Given the list of all pairs, some of which are marked as winning positions, we mark a pair (X', r') as a winning position if there is a set $X \subseteq E(H)$ satisfying:

- The connected component R' of $G \setminus \bigcup X'$ containing r' is the component of $G \setminus ((\bigcup X') \cap (\bigcup X))$ that contains r' (monotonicity).
- For all vertices $r \in V(G)$ that are connected to r' in $G \setminus ((\bigcup X') \cap (\bigcup X))$, the pair (X, r) is marked as a winning position.

This can be done in polynomial time. Since the strategy must be monotone, after $|V(G)|$ iterations we can check whether all initial positions $(\emptyset, r)$ with $r \in V(G)$ are marked as winning positions. If this is the case, then k marshals have a monotone winning strategy the algorithm returns $\text{hw}(G, H) \leq k$. Otherwise it returns $\text{hw}(G, H) > k$. $\square$

5. Compactness of generalized hypertree-width.

5.1. A compactness theorem. Let (G, H) be a hypergraph pair. For a finite set $X_0 \subseteq V(G)$, we call

$$c_H(X_0) := \min \{ |Y| \mid Y \subseteq E(H), X_0 \subseteq \bigcup Y \}$$

the *cover number* of X_0 . A hypergraph H has *finite character*, if for all infinite $X \subseteq V(H)$ and for all integers $k \geq 0$ there is a finite subset $X_0 \subseteq X$ with $c_H(X_0) > k$. A hypergraph pair (G, H) has finite character, if H has finite character. We call a hypergraph *bounded* if there is a $\kappa \in \mathbb{N}$ such that for all $e \in E(H)$: $|e| \leq \kappa$. Obviously, if H is bounded, then (G, H) has finite character. An edge which joins two vertices of a cycle but is not itself an edge of the cycle is called a *chord*. A graph is *chordal* if every cycle of length ≥ 4 has a chord. G' is a *triangulation of G* if G' is chordal, then $V(G) = V(G')$ and $E(G) \subseteq E(G')$.

Our aim is to prove the following theorem.

THEOREM 17. *Let (G, H) be a hypergraph pair with finite character. Then*

$$\text{ghw}(G, H) \leq k$$

if and only if

$$\text{ghw}(G_0, H_0) \leq k \text{ for all finite induced subhypergraph pairs } (G_0, H_0) \text{ of } (G, H).$$

This implies the following corollary.

COROLLARY 18. *A hypergraph H with finite character satisfies*

$$\text{ghw}(H) \leq k$$

if and only if

$$\text{ghw}(H_0) \leq k \text{ for all finite induced subhypergraphs } H_0 \text{ of } H.$$

For the proof of Theorem 17 we use the following theorem.

THEOREM 19. *Let G be a graph containing no infinite clique. Then G is chordal if and only if G admits a tree-decomposition into complete pieces (i.e., for all $t \in V(T)$, B_t induces a clique in G).*

Theorem 19 can be derived from Halin's theory of simplicial decompositions, which will be done in section 5.2.⁶

LEMMA 20. *Let G be a finite graph. G is chordal if and only if G has a tree-decomposition (T, B) into complete pieces.*

A proof is given in [4, Proposition 12.3.11].

Proof of Theorem 17. For the “only if” part, restrict a generalized hypertree decomposition of (G, H) to (G_0, H_0) .

Proof of the “if” part:⁷ We may assume that G is a graph. Let $\text{ghw}(G_0, H_0) \leq k$ for all finite induced subhypergraph pairs (G_0, H_0) of (G, H) .

Claim 1. Every finite subgraph G_0 of G has a triangulation G'_0 such that every complete subgraph C of G'_0 satisfies $c_{H_0}(V(C)) \leq k$.

Proof of the claim. Let (T, B, C) be a generalized hypertree decomposition of (G_0, H_0) of width $\leq k$. We obtain G'_0 by setting $V(G'_0) := V(G_0)$ and

$$E(G'_0) := \{\{u, v\} \subseteq V(G_0) \mid \text{there is a } t \in V(T) \text{ such that } \{u, v\} \subseteq B_t\}.$$

By Lemma 20, G'_0 is chordal, and Remark 1 shows that every complete subgraph of G'_0 is covered by at most k hyperedges of H_0 .

Claim 2. G has a triangulation G' such that every complete subgraph C of G' satisfies $c_H(V(C)) \leq k$.

Proof of the claim. We use Zorn's lemma. For a hypergraph pair (G, H) let I be the set of all supergraphs G' of G with $V(G') = V(G)$ satisfying:

(*) Every finite subgraph G_0 of G' has a triangulation G'_0 s.t. every complete subgraph C of G'_0 satisfies $c_H(V(C)) \leq k$.

I is ordered inductively by inclusion: For a transfinite sequence $(G_\alpha)_{\alpha < \delta}$ in I with $G_\alpha \subseteq G_\beta$ for all $\alpha \leq \beta$, $\bigcup_{\alpha < \delta} G_\alpha$ satisfies (*), since a finite subgraph G_0 of $\bigcup_{\alpha < \delta} G_\alpha$ is already contained in some G_α .

Let $G' \in I$ be a maximal element. The following two claims finish the proof:

Claim (i). Let C be a clique in G' . Then $c_H(V(C)) \leq k$.

Claim (ii). G' is chordal.

(i) If C is finite, this is true since C is contained in some G_α . C cannot be infinite since (G, H) has finite character, and hence there would be a finite $C_0 \subseteq C$ with $c_H(V(C_0)) > k$, a contradiction.

(ii) Suppose $O = v_1, \dots, v_m, v_m = v_1$ is an induced cycle of length at least four in G' . By maximality of G' , we cannot add a chord $\{v_i, v_j\}$ to O without producing a finite subgraph $G_{ij} \subseteq G'$ that does not satisfy (*). But $G'[V(O) \cup \bigcup_{ij} V(G_{ij})]$ is a finite subgraph of G' and hence is contained in some G_α . A contradiction to (*).

If G' contains no infinite clique, we apply Theorem 19, obtaining a tree-decomposition (T, B) of G' into complete pieces. We extend (T, B) to a generalized hypertree decomposition (T, B, C) of (G', H) by choosing for each $t \in V(T)$ a cover $C_t \subseteq E(H)$ of B_t with $|C_t| \leq k$. Then (T, B, C) is a generalized hypertree decomposition of width $\leq k$ of (G, H) as well.

Let C be an infinite clique in G' . It suffices to show that in this case the generalized hypertree-width of the finite induced subhypergraph pairs of (H, G) is unbounded. By the finite character of (G, H) there is a finite subset $X_k \subseteq V(C)$ for each k , such that

⁶A slightly weaker version of Theorem 19 is proved in [12].

⁷Here we adapt Thomassen's elegant proof of compactness of tree-width of graphs, cf. [16].

$c_H(X_k) > k$. Remark 1 shows that each X_k induces a finite subhypergraph pair (H_k, G_k) of (H, G) with $\text{ghw}(G_k, H_k) > k$. $\square$

Now we give an example showing that generalized hypertree-width is not compact for hypergraphs without finite character.

LEMMA 21. $H = (\aleph_1, \mathcal{P}_{<\omega}(\aleph_1))$ satisfies the following.

1. $\text{ghw}(H_0) \leq 1$ for all finite induced subhypergraphs H_0 of H .
2. $\text{ghw}(H) = \infty$.

Proof. 1. is obvious. For 2., suppose (T, B, C) is a generalized hypertree decomposition of $(\underline{H}, H)$ of width $\leq k$. Then

- (i) Each $t \in V(T)$ has maximum one successor s s.t. $\beta_s \neq \emptyset$.
- (ii) Such a successor s satisfies $B_t \subseteq B_s$.

Proof of (i). Otherwise, there are (at least) two successors s and s' of t with $v \in \beta_s$ and $v' \in \beta_{s'}$. By connectedness, $v \neq v'$. But the edge $\{v, v'\}$ has to be covered in some node of T , a contradiction to connectedness.

Proof of (ii). Let s be a successor of t with $u \in \beta_s$. Suppose $v \in B_t \setminus B_s$. Again, the edge $\{u, v\}$ has to be covered in some node of T , a contradiction to connectedness.

Together, (i) and (ii) show that $V(H) = \aleph_1$ can be obtained as a countable union of an increasing sequence of finite subsets of $V(H)$, a contradiction. $\square$

Note that this example also shows that hypertree-width is not compact. In section 6.2 we give an example showing that even for bounded hypergraphs, hypertree-width is not compact.

5.2. Proof of Theorem 19.

Proof of the “if” part of Theorem 19. Let (T, B) be a tree-decomposition of G into complete pieces. Let O be a cycle of length ≥ 4 in G . For each edge e of O choose $t_e \in V(T)$ covering e . Let T' be the subtree of T spanned by the nodes $\{t_e \mid e \text{ an edge of } O\}$. Since T' is finite, we can apply Lemma 20 to the subgraph of G covered by T' . Therefore O has a chord. $\square$

For a proof of the “only if” part of Theorem 19, we develop Halin’s theory a little further.

Let G be a graph, δ an ordinal, and $(G_\alpha)_{\alpha < \delta}$ a family of subgraphs of G . $(G_\alpha)_{\alpha < \delta}$ is called a *simplicial decomposition* of G , if it satisfies

1. $G = \bigcup_{\alpha < \delta} G_\alpha$.
2. $C_\beta = G_\beta \cap \bigcup_{\alpha < \beta} G_\alpha$ is a clique for every $\beta < \delta$.
3. $G_\beta \supsetneq C_\beta$ and $\bigcup_{\alpha < \beta} G_\alpha \supsetneq C_\beta$.

LEMMA 22. Let $(G_\alpha)_{\alpha < \delta}$ be a simplicial decomposition of G such that all C_β , $\beta < \delta$ are finite. Then G has a tree-decomposition (T, B) with pieces exactly the G_α .

Proof. Define (T^0, B^0) by setting $T^0 = (\emptyset, \emptyset)$. Then (T^0, B^0) is a tree-decomposition of $\bigcup_{\alpha < 0} G_\alpha$. Let (T^α, B^α) be already defined, (T^α, B^α) a tree-decomposition of $\bigcup_{\beta < \alpha} G_\beta$. Since $C_\alpha = G_\alpha \cap \bigcup_{\beta < \alpha} G_\beta$ is finite, by Remark 1 there is a $t \in V(T^\alpha)$ such that $V(C_\alpha) \subseteq B_t^\alpha$. We now join $V(G_\alpha)$ to the node t :

- Set $V(T^{\alpha+1}) := V(T^\alpha) \cup \{s\}$ for a new node $s \notin V(T^\alpha)$,
- $E(T^{\alpha+1}) := E(T^\alpha) \cup \{\{s, t\}\}$, and
- $B_t^{\alpha+1} := B_t^\alpha$ for all $t \in V(T^\alpha)$, $B_s^{\alpha+1} := V(G_\alpha)$.

It is easy to see that $(T^{\alpha+1}, B^{\alpha+1})$ is a tree-decomposition of $\bigcup_{\beta < \alpha+1} G_\beta$.

Let λ be a limit ordinal. For each $\alpha < \lambda$, let (T^α, B^α) be already defined, (T^α, B^α) a tree-decomposition of $\bigcup_{\beta < \alpha} G_\beta$. Set $T^\lambda := \bigcup_{\alpha < \lambda} T^\alpha$ and $B^\lambda := \bigcup_{\alpha < \lambda} B^\alpha$. It is easy to see that (T^λ, B^λ) is a tree-decomposition of $\bigcup_{\beta < \lambda} G_\beta$. Hence, (T^δ, B^δ) is a tree-decomposition of G . $\square$

A *prime* graph G is a connected graph $G \neq \emptyset$ containing no clique C such that $G \setminus C$ is disconnected. We say G contains no *separating* clique.

LEMMA 23. *Let P be a prime chordal graph. Then P is complete.*

Proof. Suppose not. Then there are $p_1, p_2 \in V(P)$ with $\{p_1, p_2\} \notin E(P)$. Our aim is to find a clique in P separating p_1 and p_2 , a contradiction to P being prime.

Claim (i). There exists a minimal set $S \subseteq V(P)$ separating p_1 and p_2 .

Claim (ii). A minimal set S separating p_1 and p_2 induces a clique in P .

These claims complete the proof: Since P is prime P cannot contain such an S .

Proof of (i). Let $V(P) \setminus \{p_1, p_2\} = S_0 \supseteq S_1 \supseteq S_2 \supseteq \dots$ be a maximal decreasing sequence of subsets of $V(P)$ separating p_1 and p_2 . Then $S = \bigcap_i S_i$ separates p_1 and p_2 . If not, there is a path $p_1 = v_1, v_2, \dots, v_n = p_2$ in $P \setminus S$. Hence some S_{i_0} satisfies $S_{i_0} \cap \{v_i \mid i = 1, \dots, n\} = \emptyset$. But then S_{i_0} does not separate p_1 and p_2 , a contradiction to our assumption. Therefore S is a minimal set separating p_1 and p_2 .

Proof of (ii). If not, there are $s, t \in S$ such that $\{s, t\} \notin E(P)$. Let C_1 be the connected component of $P \setminus S$ containing p_1 , and let C_2 be the connected component of $P \setminus S$ containing p_2 . Since S is minimal, both s and t have a neighbor in C_1 and a neighbor in C_2 . Let P_1 be the shortest path from s to t in $C_1 \cup \{s, t\}$ and let P_2 be the shortest path from s to t in $C_2 \cup \{s, t\}$. Then $P_1 \cup P_2$ is a cycle of length ≥ 4 without a chord, a contradiction. $\square$

Our proof of Theorem 19 is based on a theorem of Halin (cf. [11, Satz 5]), which we do not prove here.

THEOREM 24 (Halin's theory). *Any graph G containing no infinite clique has a simplicial decomposition $(G_\alpha)_{\alpha < \delta}$ into prime induced subgraphs G_α of G .*

Proof of the "only if" part of Theorem 19. Let G be a chordal graph containing no infinite clique. By Fact 24, G has a simplicial decomposition $(G_\alpha)_{\alpha < \delta}$ into prime induced subgraphs. By Lemma 22, $(G_\alpha)_{\alpha < \delta}$ yields a tree-decomposition (T, B) with pieces exactly as G_α . Since all the G_α are prime and chordal, they are complete by Lemma 23. $\square$

6. Examples of noncompactness.

6.1. Noncompactness of hypertree-width of hypergraph pairs. In this section, we define a hypergraph pair (G, H) such that

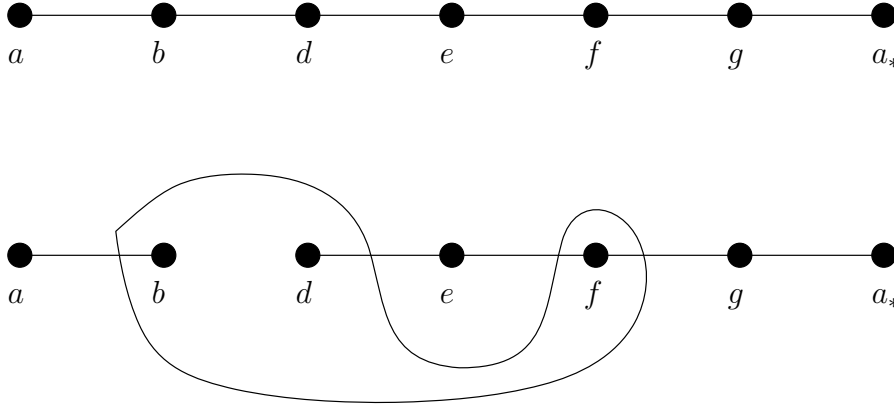
- (i) $\text{hw}(G, H) = 2$,
- (ii) $\text{hw}(G_0, H_0) \leq 1$ for all finite induced subhypergraph pairs (G_0, H_0) of (G, H) , and
- (iii) all hyperedges $e \in E(H)$ satisfy $|e| \leq 3$.

Let (G', H') be the hypergraph pair defined by $V(G') = \{a, b, d, e, f, g, a_*\}$, $E(G') = \{\{a, b\}, \{b, d\}, \{d, e\}, \{e, f\}, \{f, g\}, \{g, a_*\}\}$, and $E(H') = \{\{a, b\}, \{b, d, f\}, \{d, e\}, \{e, f\}, \{f, g\}, \{g, a_*\}\}$ (see Figure 6.1).

Intuitively, the idea is that one marshal can catch the robber monotonely on (G', H') by chasing him from right to left, but not by chasing him from left to right. Gluing together infinitely many copies of (G', H') as explained below allows the robber to make sure that he is always on the right side of the marshal and thus he can elude capture.

For each $i \in \mathbb{Z}$, let (G^i, H^i) be an isomorphic copy of (G', H') with vertex set $\{a^i, b^i, d^i, e^i, f^i, g^i, a_*^i\}$. By identifying the "rightmost" vertex of the i th copy with the first vertex of the $(i+1)$ th copy, we obtain the hypergraph pair of Figure 6.2,

$$(G, H) := \bigcup_{i \in \mathbb{Z}} (G^i, H^i) / (a_*^i = a^{i+1}).$$

FIG. 6.1. The hypergraph pair (G', H') .

The characterization of $\text{hw}(G, H)$ by the robber and marshals game (cf. section 4.2) allows us to argue the game theoretically: It is easy to see that two marshals have a winning strategy on (G, H) , and we will not prove this here.

Let (G_0, H_0) be a finite induced subhypergraph pair of (G, H) . Then one marshal has a winning strategy. He starts by occupying the rightmost hyperedge of (G_0, H_0) and then he chases the robber to the left.

Nevertheless, one marshal does not have a winning strategy on (G, H) . The robber can escape as follows: After the marshal's first move to a hyperedge $h \in E(H)$, the robber goes to the right side of h , choosing the unoccupied vertex e^i which is as far "left" as possible such that there is a vertex $v \in V(G)$ on the "left side" of e^i with $v \in h$. From then on, the robber stays on e^i . By monotonicity it is easy to see that each of the marshal's following moves is uniquely determined until at some point the marshal occupies the hyperedge $\{b^i, d^i, f^i\}$. Now he cannot make a move at all, and therefore the robber wins. Hence (G, H) satisfies (i), (ii), and (iii).

Let us remark that the hypergraph pair (G, H) has finite character and all finite subhypergraphs (G_0, H_0) of (G, H) satisfy $\text{ghw}(G_0, H_0) \leq 1$. With Theorem 17 it follows that $\text{ghw}(G, H) \leq 1$. It is left to the reader as an exercise to find a generalized hypertree decomposition of width 1 for (G, H) .

6.2. Noncompactness of hypertree-width. In this section we prove the following theorem.

THEOREM 25. *There is a hypergraph H satisfying*

- (i) $\text{hw}(H) = 4$,
- (ii) $\text{hw}(H_0) \leq 3$ for all finite induced subhypergraphs H_0 of H , and
- (iii) all hyperedges $e \in E(H)$ satisfy $|e| \leq 3$.

Although the presentation will be independent from the previous paragraph, the idea of the construction is to code the hypergraph pair (G, H) of the previous paragraph in a single hypergraph, in such a way that the winning strategies for few marshals on the "implementation" of (G, H) as a hypergraph are basically the same as on (G, H) . So we will have to implement the *marshal edges* (i.e., the edges from $E(H) \setminus E(G)$) and the *robber edges* (i.e., the edges from $E(G) \setminus E(H)$) using ordinary edges that can be used both by the marshals and by the robber. The implementation of the marshal edges will increase the number of marshals necessary to catch the robber by two. Intuitively, a marshal edge will be implemented by some highly con-

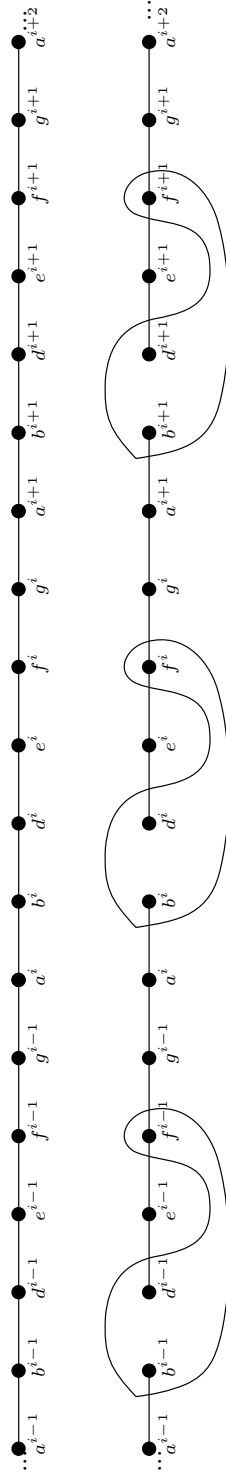


FIG. 6.2. The hypergraph pair (G, H) .

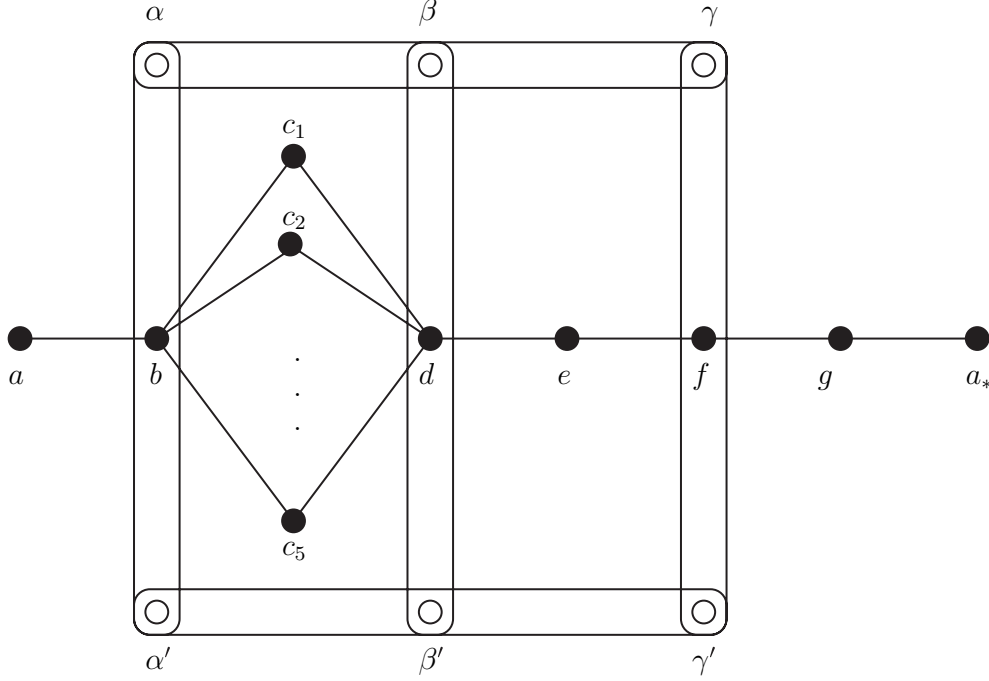


FIG. 6.3. The hypergraph M . In addition to the hyperedges shown in the figure, every vertex from P is connected to every vertex from Q by a 2-edge. Also, for every $v \in V(M)$, M contains the hyperedge $\{v\}$.

nected vertices, which the marshals have to occupy during the whole game (because otherwise the robber can reach almost every vertex he likes). In this way we make sure that the robber cannot use the marshal edge.

As in the previous section, we again construct the example by gluing together infinitely many copies of a certain “module.” Let M be the hypergraph from Figure 6.3, with $V(M) = P \cup Q \cup \{c_1, c_2, \dots, c_5\}$, where $P = \{\alpha, \beta, \gamma, \alpha', \beta', \gamma'\}$, and $Q = \{a, b, d, e, f, g, a_*\}$,

- $E(M) = \{\{a, b\}, \{d, e\}, \{e, f\}, \{f, g\}, \{g, a_*\}, \{\alpha, \beta, \gamma\}, \{\alpha', \beta', \gamma'\}\},$
- $\cup \{\{b, c_i\} \mid i = 1, \dots, 5\} \cup \{\{c_i, d\} \mid i = 1, \dots, 5\} \cup \{\{\alpha, b, \alpha'\}, \{\beta, d, \beta'\}, \{\gamma, f, \gamma'\}\},$
- $\cup \{\{p, q\} \mid p \in P, q \in Q\} \cup \{\{v\} \mid v \in V(M)\}.$

Intuitively, the vertices of P are the highly connected vertices, which the marshals will have to occupy during the whole game. We add the singletons $\{v\}$ for $v \in V(M)$ to the hyperedges so that it is no advantage for the marshals to use an edge of type $\{p, q\}$ for $p \in P, q \in Q$. It does not help the marshals chasing the robber on $Q \cup \{c_1, \dots, c_5\}$, since the marshal occupying $\{p, q\}$ can use the edge $\{q\}$ instead. Furthermore, occupying an edge $\{p, q\}$ does not allow to cover P with few marshals, so the marshals can avoid such edges.

The robber edge $\{b, d\} \in E(G')$ of the hypergraph pair of Figure 6.1 is implemented by 5 paths of length two between b and d (via the vertices $c_1, \dots, c_5$, respectively). This makes sure that 4 marshals cannot chase the robber from b to d or vice versa, without covering b and d at the same time—as in (G', H') (see Figure 6.1).

Let $(M^i)_{i \in \mathbb{Z}}$ be a family of isomorphic copies of M . We denote the vertices of M^i by $\{v^i \mid v \in V(M)\}$. By identifying the rightmost vertex of the i th copy with the

first vertex of the $(i + 1)$ th copy, we obtain the hypergraph

$$H := \bigcup_{i \in \mathbb{Z}} (M^i) / (a_*^i = a^{i+1}).$$

Proof of Theorem 25, (i). It is easy to see that four marshals can win on H . The following three lemmas show that $\text{mon-mw}(H) > 3$:

LEMMA 26. *On H as defined above, the robber can win against three marshals.*

Proof. We describe such a winning strategy, which is similar to the robber's winning strategy on the hypergraph pair (G, H) defined in the previous paragraph. Let $S := \bigcup_{i \in \mathbb{Z}} \{a^i, b^i, c_1^i, d^i, e^i, f^i, g^i, a_*^i\}$. As long as no vertex from S is occupied by a marshal, the robber stays in the connected component of H containing S . Sooner or later, the marshals will have to first occupy one or more vertices of S . Let v be the rightmost such. Then the robber reacts by moving to the leftmost vertex a^i that is to the right of v . Let H' be the restriction of H to M^i together with the path from v to a^i (in the sense of induced subhypergraph). The following lemma shows that the robber can elude capture. $\square$

LEMMA 27. *Let N be the hypergraph defined by $V(N) = \{v_i \mid i < n\} \cup V(M)$ and $E(N) = \{\{v_n, v_{n-1}\}, \{v_{n-1}, v_{n-2}\}, \dots, \{v_0, a\}\} \cup E(M)$. The robber can win on N against three marshals, if in the first move the marshals occupy v_n .*

Proof. The robber remains on a until a marshal occupies a . Then the robber plays according to his winning strategy from Lemma 28 below. $\square$

LEMMA 28. *The robber can win against three marshals on M , if the first game position is (X, a) with some $X \subseteq E(M)$, $|X| \leq 2$ and if in the second move some marshal occupies an edge containing the vertex a .*

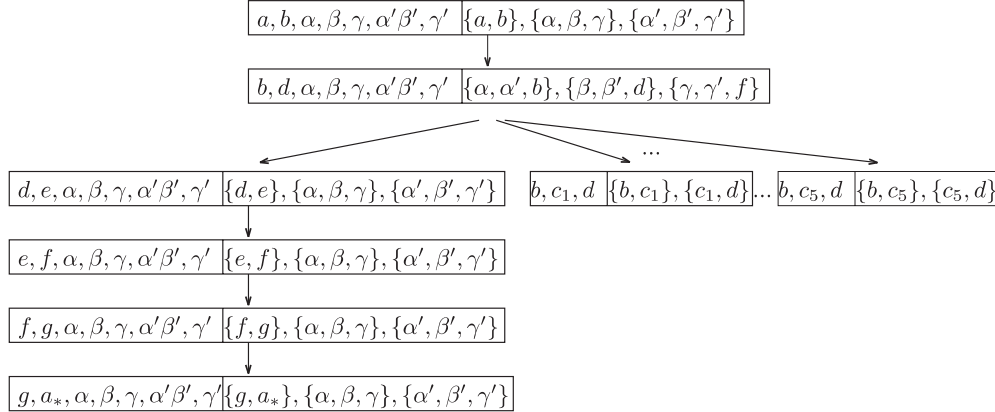
Proof. As long as a vertex from P is not occupied by the marshals, the robber stays on such a vertex. All this time the marshals have to occupy a . As soon as the marshals occupy all vertices from P (the marshals still have to occupy a), we have one of the two following situations:

The marshals occupy $\{\alpha, \beta, \gamma\}$, $\{\alpha', \beta', \gamma'\}$, and $\{a\}$, or they occupy $\{\alpha, \beta, \gamma\}$, $\{\alpha', \beta', \gamma'\}$, and $\{a, b\}$.

In the first case the robber stays on b until the marshals occupy b , by which we are in the second case. In the second case the robber moves to e . Now the marshals can either additionally occupy a vertex c_i (and then give up), or rearrange themselves to occupy $\{\alpha, b, \alpha'\}$, $\{\beta, d, \beta'\}$, and $\{\gamma, f, \gamma'\}$ (and then give up as well). $\square$

Proof of Theorem 25, (ii). Three marshals have a winning strategy on a finite induced subhypergraph H_0 of H . First, we may assume that H_0 is the union of finitely many “modules” M_i . The proof idea is that two marshals occupy hyperedges of the type $\{\alpha^i, \beta^i, \gamma^i\}$ and $\{\alpha'^i, \beta'^i, \gamma'^i\}$ almost all the time. The third marshal chases the robber from right to left, while the other two marshals do not change position except when the third marshal reaches an edge of type $\{d^j, e^j\}$. Then the next step for the marshals is to occupy $\{\alpha^j, b^j, \alpha'^j\}$, $\{\beta^j, d^j, \beta'^j\}$, and $\{\gamma^j, f^j, \gamma'^j\}$. After that, the robber is either caught on a vertex c_i^j , $i \in \{1, \dots, 5\}$, or two marshals again occupy the hyperedges $\{\alpha^i, \beta^i, \gamma^i\}$ and $\{\alpha'^i, \beta'^i, \gamma'^i\}$ and the third marshal continues chasing the robber towards the left end of the finite hypergraph, until the robber finally cannot move. Whenever the third marshal reaches a point of type a_i , the two other marshals move from $\{\alpha^i, \beta^i, \gamma^i\}$ and $\{\alpha'^i, \beta'^i, \gamma'^i\}$ to $\{\alpha^{i-1}, \beta^{i-1}, \gamma^{i-1}\}$ and $\{\alpha'^{i-1}, \beta'^{i-1}, \gamma'^{i-1}\}$.

Thus, $\text{mon-mw}(H_0) \leq 3$ for all finite induced subhypergraphs H_0 of H . Altogether, H satisfies (i), (ii), and (iii) as required. $\square$

FIG. 6.4. A width 3 generalized hypertree decomposition of M of Figure 6.3.

Let us remark that the hypergraph H of Theorem 25 has finite character and satisfies $\text{ghw}(H_0) \leq 3$ for all finite subhypergraphs H_0 of H . With Corollary 18 it follows that $\text{ghw}(H) \leq 3$. To obtain a generalized hypertree decomposition for H of width at most 3, the main ingredient is a width 3 generalized hypertree decomposition for M “from left to right.” Figure 6.4 shows such a generalized hypertree decomposition. Note that this is not a hypertree decomposition, because the second tree node does not satisfy condition 3.

Acknowledgments. The author thanks Prof. J. Flum for helpful hints, Prof. C. Thomassen for his elegant proof of compactness of tree-width and H. Scheuermann for many discussions.

REFERENCES

- [1] I. ADLER, *Marshals, monotone marshals, and hypertree-width*, 2004, J. Graph Theory, 47 (2004), pp. 275–296.
- [2] I. ADLER, *Width Functions for Hypertree Decompositions*, Dissertation, University of Freiburg, Freiburg, Germany, 2006; available online at http://www.freidok.uni-freiburg.de/volltexte/2468/pdf/thesis_adler.pdf.
- [3] I. ADLER, G. GOTTLÖB, AND M. GROHE, *Hypertree-width and related hypergraph invariants*, in Proceedings of the 3rd European Conference on Combinatorics, Graph Theory, and Applications, Volume AE of DMTCS Proceedings Series, S. Felsner, ed., 2005, pp. 5–10.
- [4] R. DIESTEL, *Graph Theory*, Springer, New York, 1997.
- [5] J. FLUM, M. FRICK, AND M. GROHE, *Query evaluation via tree-decompositions*, J. ACM, 49 (2002), pp. 716–752.
- [6] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability, A Guide to the Theory of NP-Completeness*, W. H. Freeman, San Francisco, 1979.
- [7] G. GOTTLÖB, N. LEONE, AND F. SCARCELLO, *Hypertree decompositions and tractable queries*, J. Comput. System Sci., 64 (2002), pp. 579–627.
- [8] G. GOTTLÖB, N. LEONE, AND F. SCARCELLO, *Robbers, marshals, and guards: Game theoretic and logical characterizations of hypertree width*, J. Comput. System Sci., 66 (2003), pp. 775–808.
- [9] G. GOTTLÖB, Z. MIKLÓS, AND TH. SCHWENTICK, *Generalized hypertree decompositions: NP-hardness and tractable variants*, in Proceedings of the Symposium on Principles of Database Systems (PODS’07), to appear.
- [10] M. GROHE AND D. MARX, *Constraint solving via fractional edge covers*, Proceedings of the of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2006), 2006, pp. 289–298.

- [11] R. HALIN, *Über simpliziale Zerfällungen beliebiger (endlicher oder unendlicher) graphen*, Math. Ann., 156 (1964), pp. 216–225.
- [12] I. KRÍŽ, AND R. THOMAS, *The menger-like property of the tree-width of infinite graphs*, J. Combin. Theory Ser. B, 52 (1991), pp. 86–91.
- [13] A. LUSTIG AND O. SHMUELI, *Acyclic hypergraph projections*, J. Algorithms, 30 (1999), pp. 400–422.
- [14] P. D. SEYMOUR AND R. THOMAS, *Graph searching and a min-max theorem for tree-width*, J. Combin. Theory Ser. B, 58 (1993), pp. 22–33.
- [15] R. THOMAS, *The tree-width compactness theorem for hypergraphs*, KAM-DIMATIA Series, preprint 92, 1988.
- [16] C. THOMASSEN, *Configurations in graphs of large minimum degree, connectivity, or chromatic number*, Combinatorial Mathematics, Proceedings of the Third International Conference, G. S. Bloom, R. L. Graham, J. Malkevitch, eds., New York Akad. Sci., New York, 1989, pp. 402–412.

RECONFIGURATIONS IN GRAPHS AND GRIDS*

GRUIA CĂLINESCU[†], ADRIAN DUMITRESCU[‡], AND JÁNOS PACH[§]

Abstract. Let G be a connected graph, and let V and V' be two n -element subsets of its vertex set $V(G)$. Imagine that we place a chip at each element of V and we want to move them into the positions of V' (V and V' may have common elements). A move is defined as shifting a chip from v_1 to v_2 ($v_1, v_2 \in V(G)$) on a path formed by edges of G so that no intermediate vertices are occupied. We give upper and lower bounds on the number of moves that are necessary and analyze the computational complexity of this problem under various assumptions: labeled versus unlabeled chips, arbitrary graphs versus the case when the graph is the rectangular (infinite) planar grid, etc. We prove hardness and inapproximability results for several variants of the problem. We also give a linear time algorithm which performs an optimal (minimum) number of moves for the unlabeled version in a tree, and a constant-ratio approximation algorithm for the unlabeled version in a graph. The graph algorithm uses the tree algorithm as a subroutine.

Key words. reconfiguration algorithms, approximation algorithms, local ratio method, proper function, NP-hardness

AMS subject classifications. 05C85, 68R05, 68R10, 68W25, 68W40

DOI. 10.1137/060652063

1. Introduction. Consider a set system (set) of n pairwise disjoint objects in the Euclidean space that need to be brought from a given *start* (initial) configuration S into a desired *target* (goal) configuration T . In many cases, the problem admits the following abstraction: we have an underlying finite or infinite graph, the start configuration is represented by a set of n chips at n distinct start vertices, and the target configuration is represented by another set of n distinct target vertices. A vertex can be both a start and target position. The cases when the chips are labeled or unlabeled give two different variants of the problem. In one move a chip can follow an arbitrary path in the graph and end up at another vertex, provided the path (including the end vertex) is free of other chips. The *motion planning* problem for such a system is that of computing a sequence of object motions (schedule) that achieves this task. If such a sequence of motions exists, we say that the problem is *feasible* and that it is *infeasible* otherwise. To avoid trivial questions, we always assume the graph is connected.

In certain applications, objects are indistinguishable, and therefore the chips are unlabeled; for instance, a modular robotic system consists of a number of identical modules (robots) that have identical capabilities [10, 11]. In another application, the chips are indivisible packets (copies) of the same data that need to be moved from

*Received by the editors February 13, 2006; accepted for publication (in revised form) July 12, 2007; published electronically February 15, 2008. A preliminary version of this paper appeared in *Proceedings of Latin American Theoretical Informatics Conference (LATIN '06)*, Lecture Notes in Comput. Sci. 3887, Springer, 2006, pp. 262–273.

<http://www.siam.org/journals/sidma/22-1/65206.html>

[†]Department of Computer Science, Illinois Institute of Technology, Chicago, IL 60616 (calinesc@iit.edu). This author's research was partially supported by NSF grant CCF-0515088.

[‡]Department of Computer Science, University of Wisconsin–Milwaukee, WI 53201-0784 (ad@cs.uwm.edu). This author's research was partially supported by NSF CAREER grant CCF-0444188.

[§]City College, CUNY and Courant Institute, NYU; 251 Mercer Street, New York, NY 10012-1185 (pach@cims.nyu.edu). This author's research was partially supported by grants from NSF, NSA, OTKA, and from the US-Israeli Binational Research Foundation.

one site to another of a wide-area communication network without ever exceeding the capacities of the communication buffers at each site [5, 16].

In this variant with unlabeled chips, the problem is easier and always feasible; therefore, one is interested in minimizing the number of moves. For the variant with labeled chips the problem may be infeasible: it is known, for instance, that the 15-puzzle on a 4×4 grid—introduced by Sam Loyd in 1878—has a solution if and only if the start permutation is an even permutation [15, 20] (see [3] for a recent approach).

Other reconfiguration rules (models) for systems of disks in the plane have been examined recently [1, 7, 8]; see also [9]. These models do not fall in the graph reconfiguration framework in this paper, because a disk may partially overlap several target positions. A model that fits in the graph reconfiguration framework has been analyzed in [10]: it deals with reconfiguration of modular systems acting in a grid-like environment, where moves must maintain connectivity of the whole system, and the motion rules are very local: a chip can only move to an adjacent position in one move. Denoting the configuration of the modules at time t by V_t , the system remains connected if for each $t = 0, 1, 2, \dots$ the graph $G_t = (V_t, E_t)$ is connected, where E_t is the set of edges connecting pairs of cells in V_t that are side-adjacent.

The general form of the reconfiguration problem we consider is to find a reconfiguration sequence with a minimum number of moves. Depending on whether we refer to the graph or grid version, or to the labeled or unlabeled version, we call the problem U-GRAPH-RP, L-GRAPH-RP, U-GRID-RP, or L-GRID-RP. In the grid version, the underlying graph is the infinite planar integer grid.

Consider, for example, the reconfiguration problem in the infinite grid with unlabeled (or labeled) chips (objects). The following simple algorithm does $2n$ moves for reconfiguration of n chips. In the first step (n moves), move in a suitable order all the chips away in the free grid space. In the second step (n moves), bring the chips “back” to their target positions. We will show that minimizing the number of moves is intractable in both (labeled and unlabeled) variants in the grid.

A move is called a *target move* if it moves a chip to a final target position. Otherwise it is called a *nontarget move*. Our lower bounds use the following argument: if no target chip coincides with a start chip (so each chip must move), then a schedule with x nontarget moves consists of at least $n + x$ moves.

It is worth noting that, in contrast with the variant with moves along free paths studied here, the variant of the reconfiguration problem (with unlabeled chips) in which we count as one move each edge traversed by a chip can be solved exactly (optimally). This will be explained at the end of section 2.

Previous related work. Most of the work done so far concerns labeled versions of the reconfiguration problem, and we give here only a very brief survey.

For the generalization of the 15-puzzle on an arbitrary graph (with $k = v - 1$ labeled chips in a connected graph on v vertices), Wilson [21] gave an efficiently checkable characterization of the solvable instances of the problem. Kornhauser, Miller, and Spirakis have extended his result to any $k \leq v - 1$ and provided bounds on the number of moves for solving any solvable instance [16].

Ratner and Warmuth have shown that finding a solution with a minimum number of moves for the $(N \times N)$ -extension of the 15-puzzle is intractable [19], so the reconfiguration problem in graphs with labeled chips is NP-hard.

Auletta et al. gave a linear time algorithm for the *pebble motion on a tree* [5]. This problem is the labeled variant of the same reconfiguration problem we study here; however, each move is along one edge only.

Papadimitriou et al. studied a problem of *motion planning on a graph* in which there is a mobile robot at one of the vertices s , which has to reach a designated vertex t using the smallest number of moves, in the presence of obstacles (pebbles) at some of the other vertices [17]. Robot and obstacle moves are done along edges. Obstacles have no destination assigned and may end up at any vertex of the graph. The problem has been shown to be NP-complete even for planar graphs, and a ratio $O(\sqrt{n})$ polynomial time approximation algorithm was given in [17].

Dumitrescu, Suzuki, and Yamashita have addressed several basic questions in the analysis of modular metamorphic systems [11]. In particular the next two questions have been shown to be decidable: (i) whether a given set of motion rules maintains connectivity, (ii) whether a goal configuration is reachable from a given initial configuration (at specified locations) using a given set of motion rules. Other seemingly similar questions have been shown to be undecidable.

Our results are the following:

- (1) The reconfiguration problem in graphs with unlabeled chips U-GRAPH-RP is NP-hard and even APX-hard.
- (2) The reconfiguration problem in graphs with labeled chips L-GRAPH-RP is APX-hard.
- (3) For the infinite planar rectangular grid, both the labeled and unlabeled variants L-GRID-RP and U-GRID-RP are NP-hard.
- (4) There exists a ratio 3 approximation algorithm for the unlabeled version in graphs U-GRAPH-RP. Thereby we also get a ratio 3 approximation algorithm for the unlabeled version U-GRID-RP in the (infinite) rectangular grid.
- (5) We show that n moves are always enough (and sometimes necessary) for the reconfiguration of n unlabeled chips in graphs. For the case of trees, we present a linear time algorithm which performs an optimal (minimum) number of moves.
- (6) We show that $7n/4$ moves are always enough, and $3n/2$ are sometimes necessary, for the reconfiguration of n labeled chips in the infinite planar rectangular grid (L-GRID-RP).

2. Chips in graphs. Let G be a connected graph, and let V and V' be two n -element subsets of its vertex set $V(G)$. Imagine that we place a chip at each element of V and we want to move them into the positions of V' (V and V' may have common elements). A move is defined as shifting a chip from v_1 to v_2 ($v_1, v_2 \in V(G)$) along a path in G so that no intermediate vertices are occupied.

THEOREM 1. *In G one can get from any n -element initial configuration V to any n -element final configuration V' using at most n moves, so that no chip moves twice. Moreover, for the case of a tree T with r vertices, there is an $O(r)$ -time algorithm which performs the optimal (minimum) number of moves.*

Proof. It is sufficient to prove the theorem for trees. We argue by induction on the number of chips. Take the smallest tree T containing V and V' , and consider an arbitrary leaf l of T . Assume first that the leaf l belongs to V : say $l = v$. If v also belongs to V' , then the result trivially follows by induction, so assume that this is not the case. Choose a path π in T , connecting v to an element v' of V' such that no internal point of π belongs to V' . Apply the induction hypothesis to $V \setminus \{v\}$ and $V' \setminus \{v'\}$ to obtain a sequence of at most $n - 1$ moves, and add a final (unobstructed) move from v to v' .

The remaining case when the leaf l belongs to V' is symmetric, say $l = v'$; choose a path π in T , connecting v' to an element v of V such that no internal point of π

belongs to V . First move v to v' and append the sequence of at most $n - 1$ moves obtained from the induction hypothesis applied to $V \setminus \{v\}$ and $V' \setminus \{v'\}$.

We further refine this algorithm so as to minimize the number of moves. We call a vertex that is both a start and a target position an *obstacle*. We have four types of vertices: free vertices, chip-only vertices, target-only vertices, and obstacles. Denote by c (resp., t) the number of chip-only (resp., target-only) vertices and by o the number of obstacles. We have $c + o = o + t = n$. We call a tree *balanced* if it contains an equal number of chip-only and target-only vertices. Clearly, the initial tree T is balanced. If there exists an obstacle whose removal from T breaks T into balanced subtrees, then we keep this obstacle fixed and proceed recursively (by induction) on the subtrees. If no obstacle removal breaks T into balanced subtrees, then all obstacles must move (each at least once), hence the number of moves necessary is at least $o + c = n$, and the algorithm in the first part of our proof can be used to obtain an optimal schedule.

We continue with the description of the linear time algorithm. We traverse the tree in postorder and compute for every vertex v four doubly linked lists, M_v , Q_v , A_v , and C_v . Let T_v denote the subtree rooted at v . We say that a sequence of moves *solves* T_v if, after executing the moves in this sequence, all the chips in T_v cover exactly every target of T_v (including obstacles). Note that by itself the definition of solving does not imply solving with a minimum number of moves.

We now discuss the four lists and give intuition on their use; a formal proof follows the description of the algorithm. A_v and C_v are lists of vertices inside T_v which require moves involving nodes outside T_v : for A_v those moves mean bringing a chip from outside T_v to a vertex of A_v , and for C_v those moves mean taking a chip from a vertex of C_v to a node outside T_v . Either A_v or C_v (possibly both) will be empty after node v is processed. A move is described by a pair of vertices: the place where we take hold of the chip and the place where we release it. M_v and Q_v are lists of moves inside T_v . Moves involving elements of A_v or C_v will use the edge from v to its parent. Moves not involving elements of $A_v \cup C_v$ will be done inside T_v and are all in M_v and Q_v . The goal is to obtain a sequence of moves solving T_v such that this sequence starts with the moves in M_v , followed by some moves on the edge from v to its parent (all these moves involve vertices from C_v or A_v), followed by the moves in Q_v . If v has no parent (i.e., v is the root of T), then the list of moves on the edge from v to its parent is empty. The final reconfiguration sequence (output by the algorithm) is the sequence of moves in the list M_v followed by the moves in the list Q_v at the root of the tree.

We now describe how A_v, C_v, M_v, Q_v are initialized and computed. Two examples of the execution of the algorithm are shown in Figure 1. If v is a leaf, then M_v and Q_v are empty, $A_v = \langle v \rangle$ if v is a target-only vertex and empty otherwise, and $C_v = \langle v \rangle$ if v is a chip-only vertex, and empty otherwise. We now show how the four lists M_v, Q_v, A_v , and C_v are obtained from the lists of the children of v . There are four cases, depending on whether v is free, target-only, chip-only, or an obstacle.

In all four cases, we first compute M_v, Q_v, A_v , and C_v by concatenating the corresponding lists of the children of v . We also make use of a procedure *Cancel*, which works as follows: iterate the following step until either A_v or C_v is empty; remove the first element from each list and add the corresponding pair at the end of M_v .

In the *first* case, v is a free vertex. After applying *Cancel*, we are done.

In the *second* case, v is a target-only vertex. We apply *Cancel*. Now, if C_v is empty, we put v at the end of A_v and we are done. If C_v is nonempty, we remove

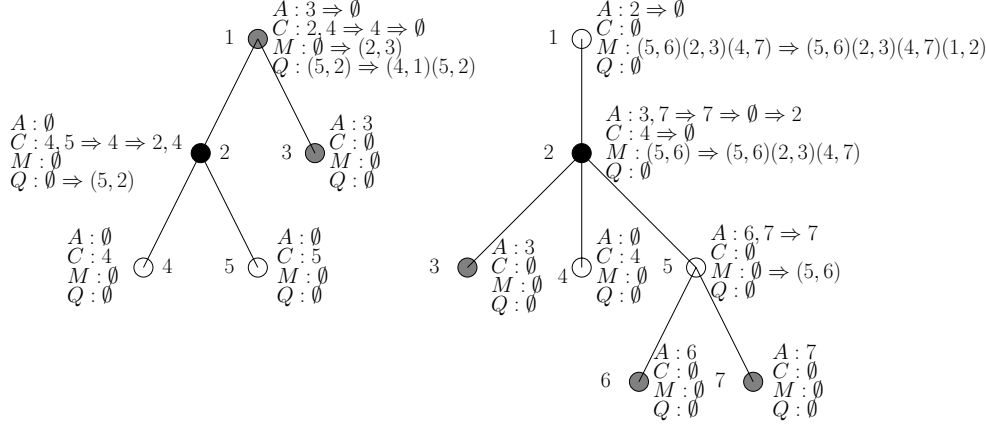


FIG. 1. Two examples of execution of the linear time algorithm for reconfiguration in a tree. Chips are drawn as empty circles, obstacles as black circles, and targets as gray circles.

from it the last element q . Insert the move (q, v) at the beginning of Q_v and we are done.

In the *third* case, v is a chip-only vertex. If A_v is empty, we put v at the beginning of C_v and we are done. If A_v is nonempty, we remove from it the first element q . We add the move (v, q) at the end of M_v , then we apply *Cancel* and we are done.

In the *fourth* and last case, v is an obstacle. (i) If A_v and C_v are both empty, we are done. (ii) If A_v is empty and C_v is nonempty, let q be the last chip from C_v . Remove q from C_v , insert the move (q, v) at the beginning of Q_v , add v at the beginning of C_v , and we are done. (iii) If A_v is nonempty, let p be the first target in A_v . Remove p from A_v and add the move (v, p) at the end of M_v . Apply *Cancel*. Now, if C_v is empty, add v at the end of A_v and we are done. If C_v is nonempty (note, A_v must be empty after *Cancel*), let q be the last chip from C_v . Remove q from C_v , insert the move (q, v) at the beginning of Q_v , and we are done.

We need to prove that this algorithm returns an optimum solution and then analyze its running time. First, we prove that it returns a correct solution, and then we justify separately its optimality. For the correctness, we prove by induction on the size of T_v that the following three conditions hold:

- If T_v contains an equal number of start and target vertices, then $A_v = C_v = \emptyset$. The sequence of moves obtained by concatenating M_v and Q_v solves T_v .
- If T_v contains q more start vertices than target vertices, then $|C_v| = q$ and $A_v = \emptyset$. Consider the sequence of moves obtained by concatenating (i) M_v , (ii) any q moves taking chips from vertices of C_v in the order they appear in C_v and placing them on vertices outside T_v , and (iii) Q_v . This sequence solves T_v .
- If T_v contains q more target vertices than start vertices, then $|A_v| = q$ and $C_v = \emptyset$. Consider the sequence of moves obtained by concatenating (i) M_v , (ii) any q moves taking chips from vertices outside T_v and placing them on vertices of A_v in the order they appear in A_v , and (iii) Q_v . This sequence solves T_v .

The base case is obvious, and the inductive step follows directly from the description of the algorithm, thereby proving that the algorithm returns a correct solution. Regarding the optimality of the number of moves, notice that an obstacle v whose

removal breaks T into balanced subtrees is not moved by the algorithm, and any other chip is moved only once; hence, the algorithm returns an optimum solution.

The time spent at vertex v is proportional to one plus the number of children of v plus the number of moves added to M_v and Q_v . Summing over the vertices, we obtain that the time is proportional to the number of vertices in the tree plus the total number of moves; hence, it is $O(n + r) = O(r)$, since $n \leq r$. $\square$

Remark. Theorem 1 implies that in the infinite rectangular grid, we can get from any starting position to any ending position of the same size n in at most n moves. It is perhaps interesting to compare this to the problem of sliding congruent unlabeled disks in the plane: here one can come up with “cage-like” constructions that require about $\frac{16n}{15}$ moves [7].

2.1. Hardness results.

THEOREM 2. *The unlabeled version in graphs U-GRAPH-RP is NP-complete. Moreover, assuming $P \neq NP$, there is an absolute constant $\epsilon_1 > 0$ such that no polynomial time algorithm has approximation guarantee at most $1 + \epsilon_1$. That is, U-GRAPH-RP is APX-hard.*

Proof. The decision version of U-GRAPH-RP is clearly in NP, so we only have to prove its NP-hardness. We reduce the *set cover* problem SC to U-GRAPH-RP. An instance of the set cover problem consists of a family $\mathcal{F}$ of subsets of a finite set U . The problem is to decide whether there is a set cover of size k for $\mathcal{F}$, i.e., a subset $\mathcal{F}' \subseteq \mathcal{F}$, with $|\mathcal{F}'| \leq k$, such that every element in U belongs to at least one member of $\mathcal{F}'$. SC is known to be NP-complete [12].

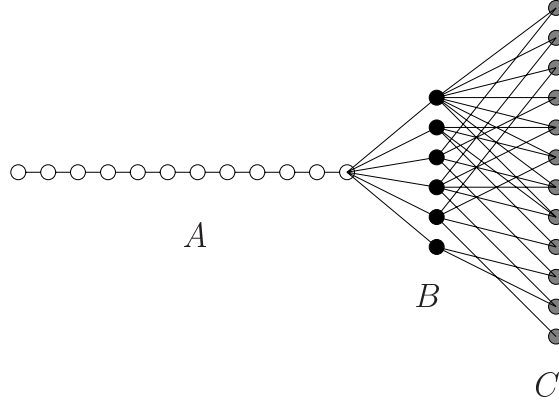


FIG. 2. The “broom” graph G corresponding to a set cover instance with $|U| = 12$ and $|\mathcal{F}| = 6$. The vertices occupied only by chips are white, those occupied by both chips and targets are black, and those occupied only by targets are shaded. An optimal reconfiguration takes 15 moves (an optimal set cover has size 3: the third, fourth, and fifth sets counting from the top).

Consider an instance of SC represented by a bipartite graph $(B \cup C, E)$, where $U = C$, $\mathcal{F} = B$, and edges describe the membership relation. Construct the undirected graph G shown in Figure 2 with $|A| = |C|$. The chips are $S = A \cup B$, and the targets are $T = B \cup C$. Clearly, G can be constructed in polynomial time. The reduction is complete once we establish the following claim.

Claim. There is a set cover consisting of at most q sets if and only if reconfiguration in G can be done using at most $|A| + q$ moves.

Proof of claim. Let q be the size of an optimal set cover. The direct implication is clear: move one chip from each element of B in the optimal cover to a target in

C ; then move the elements of A to targets in C and to the emptied targets in B using the required number of moves. For the converse implication, assume that there is a reconfiguration sequence with fewer than $|A| + q$ moves. Since all elements of A must move, fewer than q elements of B move away from their original positions. By the definition of set cover, it means that some target in C will not be filled, a contradiction.

To prove the approximation hardness we use the same reduction and the fact that 3-SC, the set cover problem in which the size of each set in $\mathcal{F}$ is bounded from above by 3, is APX-hard [2, 18]. Thus q , as in the previous paragraph, is between $|A|/3$ and $|A|$. Approximating $|A| + q$ within $1 + \epsilon$ will approximate q within $1 + 4\epsilon$. $\square$

A similar reduction can be made for the labeled version. The chips in A have targets in C , labeled as in Figure 3 (here $|A| = |C| = m$).

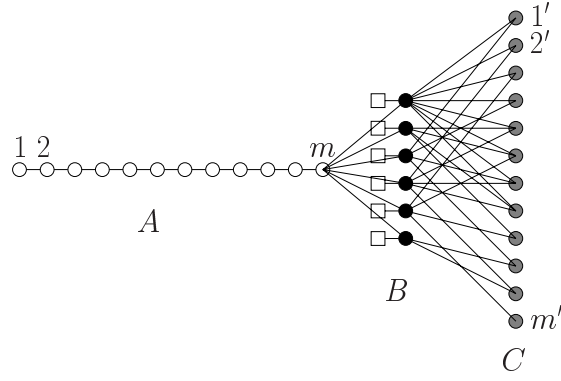


FIG. 3. The graph constructed in the reduction for labeled chips. Free vertices are drawn as squares. An optimal reconfiguration takes 18 moves (an optimal set cover has size 3: the third, fourth, and fifth sets counting from the top).

The obstacle chips in B coincide with their targets. Each vertex in B is adjacent to a “twin” free vertex. The reduction follows from the next claim. Its proof is similar to that for the previous claim, using the fact that a chip representing a selected set must move twice—once to make way for the chips whose targets are on the vertices representing elements, and once to come back.

Claim. There is a set cover consisting of at most q sets if and only if reconfiguration in G can be done using at most $|A| + 2q$ moves.

We thus obtain the following result.

THEOREM 3. *The labeled version in graphs L-GRAPH-RP is APX-hard.*

2.2. Approximation algorithms.

THEOREM 4. *There exists a 3-approximation algorithm for U-GRAPH-RP.*

Proof. The algorithm is obtained by applying the local ratio method of Bar-Yehuda [6] to a graph H whose construction we describe below.

The vertex set of the input graph G is partitioned into four sets:

- $C = V \setminus V'$, the chip-only vertices,
- $A = V' \setminus V$, the target-only vertices,
- $B = V \cap V'$, the obstacles,
- $F = V(G) \setminus (V \cup V')$, the free vertices.

Then $V(H) = A \cup B \cup C$. Observe that $|A| = |C|$. For every pair of vertices u and v of H , we put in $E(H)$ the edge uv if $uv \in E(G)$ or there is a path in G from u

to v having all the internal vertices from F .

A set of vertices (in $V(H)$) is called *balanced* if it contains an equal number of chip-only and target-only vertices. In H , we use the local ratio method to find a small set of edges Q such that every connected component D of $(V(H), Q)$ is balanced. We call this the U-STEINER problem. U-STEINER is a network design problem given by a 0-1 *proper function* [13, 14], for which both the primal-dual schema and the local ratio method were known to give a 2-approximation. The local-ratio 2-approximation algorithm is also implicit later in this section. We briefly recall here that for a ground set V , a 0-1 function $f : 2^V \rightarrow \{0, 1\}$ is proper (cf. [13, 14]) if it satisfies the following three conditions: (i) $f(V) = 0$, (ii) if A and B are disjoint, then $f(A \cup B) \leq \max(f(A), f(B))$, and (iii) f is symmetric, i.e., $f(S) = f(V - S)$ for all $S \subseteq V$.

In our case, the function, defined over all subsets of vertices, is one if the set is unbalanced and zero otherwise. It can be shown that this is a proper function; however, this fact is not needed in the ratio 3 approximation algorithm, whose proof is self-contained. From the next claim, one can easily obtain a 4-approximation for U-GRAPH-RP, as shown after the proof of claim.

One more piece of notation: given a solution R for U-GRAPH-RP in G , consider the edges of G traversed by the moving chips during the reconfiguration process. These edges, together with their endpoints (including the free vertices through which chips pass) form a number $k \geq 1$ of connected components. We then say that R has k connected components. Write $c = |C| = |A|$.

CLAIM 1. *Given a feasible solution R for U-GRAPH-RP in G with m moves and k connected components, there is a feasible solution Q for U-STEINER in H with at most $m + c - k$ edges. Conversely, given a feasible solution Q for U-STEINER in H with e edges, there is a feasible solution R for U-GRAPH-RP in G with at most $e - c + k$ moves, where k is the number of connected components of Q which intersect A (and C).*

Proof. For the first part, let S be the set of vertices of G involved in the moves of R , and let S_i , for $i = 1, 2, \dots, k$, be the connected components of R . Then the number of moves inside S_i is at least $|S_i \cap C| + |S_i \cap B|$. Let $S'_i = S_i \cap V(H)$, and note that S'_i is also connected in H . In each S'_i pick a spanning tree T_i ; the union of the trees T_i is the feasible solution Q for U-STEINER. In each T_i , the number of edges is $|V(T_i)| - 1 = |V(T_i) \cap C| + |V(T_i) \cap B| + |V(T_i) \cap A| - 1$. Summing over i gives the needed inequality.

For the second part, let T_i , for $i = 1, \dots, k$, be one such connected component (w.l.o.g. a tree) with e_i edges. Then $|A \cap V(T_i)| = |C \cap V(T_i)|$, and by using Theorem 1 one can move all the chips of $V(T_i)$, including those sitting on obstacles, to all the targets of $V(T_i)$ using $|V(T_i)| - |C \cap V(T_i)|$ moves: the chips from $(C \cup B) \cap V(T_i)$ move along the edges of T_i , passing, if necessary, through vertices of F (in G). Since $|V(T_i)| = |E(T_i)| + 1$, this second part of the claim follows by adding up over the components, since no move puts a chip on a vertex in F . (As a side remark, if Q is an optimal solution in H , then each T_i is a tree intersecting A and C .) $\square$

Here is a short account for the ratio 4 approximation algorithm. By the first part of Claim 1 applied to an optimal solution for U-GRAPH-RP in G with m_{OPT} moves and k_{OPT} components, the number of edges e_{OPT} in an optimal solution for U-STEINER in H satisfies

$$e_{OPT} \leq m_{OPT} + c - k_{OPT} \leq m_{OPT} + c \leq 2m_{OPT}.$$

Therefore, by the second part of Claim 1, the number of moves m in the solution for U-GRAPH-RP in G returned by the algorithm satisfies (since $k_1 \leq c$, where k_1 is

the number of components in the solution for U-STEINER in H which intersect A and C)

$$m \leq e - c + k_1 \leq e \leq 2e_{OPT} \leq 4m_{OPT}.$$

We now present the ratio 3 approximation algorithm. To this end, we have to enter the details of the local ratio method. The local ratio algorithm approximately solves U-STEINER instances with nonnegative weight δ on the edges. The algorithm below, given in [6], is recursive. Each recursive call is given a set S of already selected edges and a nonnegative edge weight function, and it returns a set of edges (in step 8). Given a set of edges $F \subseteq E(H)$, call a connected component X of $(V(H), F)$ *balanced* if $|V(X) \cap A| = |V(X) \cap C|$ and *unbalanced* otherwise. For the first call of the algorithm, $S = \emptyset$, and $\delta = 1$ on all edges. By feasible solution we mean feasible solution for the original U-STEINER instance.

1. The input parameters are the set of edges $S \subseteq E(H)$ and nonnegative weight δ on $E(H)$.
2. If all of the connected components of $(V(H), S)$ are balanced, then return $\emptyset$.
3. Define weight function δ_1 on edges of $E(H)$ as follows: edges going between two unbalanced components of $(V(H), S)$ get weight 1, edges going between one unbalanced component of $(V(H), S)$ and one balanced component of $(V(H), S)$ get weight $1/2$, and all other edges get weight 0.
4. Compute a positive real number α such that the weight function δ_2 on edges of H given by $\delta_2 = \delta - \alpha \cdot \delta_1$ is nonnegative, and for at least one edge e we have $\delta_2(e) = 0 < \delta(e)$.
5. Let $M = \{e \mid \delta_2(e) = 0\}$.
6. Recursively solve the instance with parameters $S \cup M$ and weight δ_2 on $E(H)$, producing a set of edges L such that $S \cup M \cup L$ is a feasible solution.
7. Obtain minimal $M' \subseteq M$ such that $S \cup M' \cup L$ is a feasible solution.
8. Return $L \cup M'$.

We are guaranteed that $M \neq \emptyset$ and thus the algorithm terminates. For the approximation ratio, we need the following claim.

CLAIM 2. *During the execution of a recursive call of the algorithm, let k be the number of unbalanced components of $(V(H), S)$. Then $\delta_1(P) \geq k/2$ for any feasible solution P . The set of edges Q returned by the recursive call of the algorithm satisfies $\delta_1(Q) \leq k - q$, where q is the number of connected components of $(V(H), Q \cup S)$.*

Proof. Consider a feasible solution P . If an edge of P goes between two unbalanced components of $(V(H), S)$, we assign $1/2$ to each such component, and if it goes from one unbalanced component of $(V(H), S)$ to one balanced component of $(V(H), S)$, we assign $1/2$ to the unbalanced component. Each unbalanced component of $(V(H), S)$ must have at least one edge of P going to some other component, and thus it is assigned at least $1/2$. Therefore $\delta_1(P) \geq k/2$.

Consider now the edges Q selected (returned) by one recursive call of the algorithm and let Q_i , for $i = 1, \dots, q$, be the connected components of $(V(H), Q \cup S)$. Fix one component Q_i . Inside Q_i , contract to a single vertex the vertices from the same component of $(V(H), S)$, obtaining $\bar{Q}_i$. The minimal property of Q as ensured by step 7 of the algorithm ensures that $\bar{Q}_i$ is acyclic (and thus it is a tree) and every leaf of $\bar{Q}_i$ is an unbalanced component of S . Note that all of the edges of Q_i with positive δ_1 -weight are in $\bar{Q}_i$. Root $\bar{Q}_i$ at an arbitrary vertex v given by an unbalanced component of S .

For every $u \in (V(\bar{Q}_i) \setminus \{v\})$ given by an unbalanced component of S , charge to u either one or two edges of total δ_1 -weight 1 as follows: consider the path from u to v

in $\bar{Q}_i$, and let u' be the next vertex on this path given by an unbalanced component of S . If the path has only one edge, then charge this edge to u . If this path has more than one edge, then charge to u the first and last edges of the path. It is easy to check that every edge of positive δ_1 -weight of $\bar{Q}_i$ is charged at least once. Thus $\delta_1(Q_i) = \delta_1(\bar{Q}_i) \leq s_i - 1$, where s_i is the number of unbalanced components of S included in Q_i ; here we used that v is not being charged.

Summing over i yields the second part of the claim. $\square$

We now continue with the proof of Theorem 4. First we note that, by using Claim 2 and induction, the local ratio algorithm ensures that its output LR satisfies, for any feasible solution P , the following:

$$(2.1) \quad \delta(LR) = \alpha\delta_1(LR) + \delta_2(LR) \leq \alpha 2\delta_1(P) + 2\delta_2(P) = 2\delta(P).$$

Let OPT be a solution with a minimum number of moves, $m(OPT)$ be the number of moves of OPT , and OP be the set of edges of the U-STEINER feasible solution obtained from OPT in Claim 1. Recall that LR is the set of edges selected by the local ratio approximation algorithm when applied to the U-STEINER instance, and let $k(LR)$ be the number of connected components of LR which intersect A , and $m(LR)$ be the number of moves of the solution obtained in Claim 1 from LR . The weight functions δ , δ_1 , and δ_2 refer to the first call of the local ratio algorithm, as does the real α , which we note is at least 1. We now have the following:

$$\begin{aligned} m(LR) &\leq |LR| + k(LR) - c \quad \text{by Claim 1} \\ &= \delta(LR) + k(LR) - c \\ &= \alpha\delta_1(LR) + \delta_2(LR) + k(LR) - c \\ &\leq \alpha(2\delta_1(OP) - k(LR)) + \delta_2(LR) + k(LR) - c \quad \text{by Claim 2} \\ &\leq 2\alpha\delta_1(OP) - k(LR) + \delta_2(LR) + k(LR) - c \quad \text{since } \alpha \geq 1 \\ &\leq 2\alpha\delta_1(OP) + 2\delta_2(OP) - c \quad \text{by (2.1)} \\ &\leq 2\delta(OP) - c \\ &= 2|OP| - c \\ &\leq 2(m(OPT) + c) - c \quad \text{by Claim 1} \\ &= 2m(OPT) + c \\ &\leq 3m(OPT) \quad \text{since } m(OPT) \geq c. \end{aligned}$$

This concludes the proof of Theorem 4. $\square$

Remark. In the graph version with unlabeled chips, if we count as one move every edge traversed by a chip, then minimizing the number of moves can be solved in polynomial time, as described below. Construct a complete weighted bipartite graph $B = (V \cup V', E)$ with bipartition V being the vertices containing chips and V' being the vertices containing targets (with obstacles in both sides of the bipartition). The weight of an edge in E is equal to the length of the shortest path in G connecting the endpoints of the edge. Now apply an algorithm for Minimum Weight Perfect Matching in B and move accordingly: if the path a chip c_1 would take to reach its destination has another chip c_2 on it, then have the two chips switch destinations and continue moving c_2 . One can check that the number of moves does not exceed the weight of the perfect matching. On the other hand, the optimum solution must move chips to targets and cannot do better than the total length of the shortest paths in a minimum matching.

3. Chips in grids. In this section we analyze the reconfiguration problem with labeled and unlabeled chips, respectively, in an infinite grid. However, a sufficiently large finite section of the grid containing all start and target positions clearly suffices for this purpose.

3.1. Hardness results.

THEOREM 5. *The unlabeled version in the grid U-GRID-RP is NP-complete.*

Proof. The decision version of U-GRID-RP is clearly in NP, so we only have to prove its NP-hardness. We reduce the *Rectilinear Steiner Tree* problem R-STEINER to U-GRID-RP. An instance of R-STEINER consists of n points in the plane, and the problem is to decide whether there is a rectilinear Steiner tree (that is, a tree containing only horizontal and vertical edges) of length at most q which connects all points. For convenience the points can be chosen with integer coordinates. R-STEINER is known to be strongly NP-complete [12], and thus we can assume all the coordinates are given in unary.

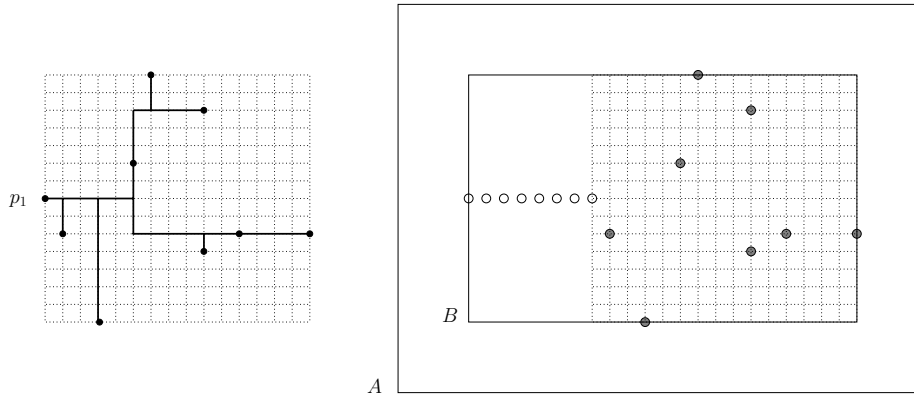


FIG. 4. Unlabeled chips in the grid obtained from a Rectilinear Steiner Tree instance with $n = 9$ (sketch). Chip-only vertices are drawn as empty circles, and target-only vertices are drawn as filled (black) circles. Obstacles (not shown) are placed at each other integer point in the axis-aligned bounding rectangle.

Consider an instance $P = \{p_1, \dots, p_n\}$ of R-STEINER with n points. Assume that the x -coordinate of a leftmost point $p_1 \in P$ is equal to zero. The instance of U-GRID-RP is illustrated in Figure 4. It has $n - 1$ chip-only vertices having the same y -coordinate with p_1 and x -coordinates $0, -1, -2, \dots, -n + 1$, and $n - 1$ target-only vertices located at the other points $\{p_2, \dots, p_n\}$ of the R-STEINER instance. Let B be a smallest axis-aligned bounding rectangle of all chip-only and target-only vertices and Δ_x, Δ_y be the dimensions of B . Consider a sufficiently large axis-aligned rectangle A enclosing the chip-only and target-only vertices (thus A encloses B) with obstacles placed at each other integer point. The boundary of A is at distance $\Delta_x + \Delta_y$ from the boundary of B . This construction is done in time polynomial in $\Delta_x \times \Delta_y$, which is polynomial in the size of the R-STEINER instance since the coordinates are given in unary. The reduction is complete once we establish the following claim.

Claim. There is a rectilinear Steiner tree of length at most q if and only if reconfiguration can be done using at most q moves.

Proof of claim. We start with the direct implication. Let T be a Steiner tree of length q connecting all target-only vertices and p_1 . Pick a leaf of T which is a target-only vertex and fill it with a chip (obstacle or chip-only) that is closest to it in

T by moving it along the corresponding path in T . If the chip comes from an obstacle vertex, then a new target-only vertex results in place of the leaf. The length of the resulting tree connecting all target-only vertices and p_1 is one less than the length of T . Continue by repeating this step until all targets have been filled (note that there is a leaf target-only vertex at each step); q moves are performed in total.

We continue with the converse implication. Denote by q the length of a minimum Steiner tree T connecting the points in P . Consider a valid reconfiguration sequence of moves $m_1, \dots, m_t$, w.l.o.g., having a minimum number of moves. Each move m_i is an orthogonal path in the grid covering a set of elements $V(m_i)$ in rectangle B (since A is large enough, each move going outside rectangle B would result in a nonoptimal reconfiguration sequence). In addition, we can assume that all chips pass through p_1 . Note that $V = V(m_1) \cup \dots \cup V(m_t)$ is the vertex set of a connected grid graph G that includes all chip-only elements, all target-only elements, and other obstacles.

The number of edges of G is at least $q + (n - 2)$, as in addition to the Steiner tree $n - 2$ edges appear to the left of p_1 ; thus $|V| \geq q + (n - 2) + 1$. Note that $V \setminus \{p_2, \dots, p_n\}$ consists of chips or obstacles. Each of these chips or obstacles must participate at least once in a move, and their number is at least $(q + 1) + (n - 2) - (n - 1) = q$, as required.

This also concludes the proof of Theorem 5. $\square$

Remark. From the results of [4], there exists a polynomial time approximation scheme for a Rectilinear Steiner Tree. Thus our reduction does not give an inapproximability result for U-GRID-RP, and we leave open the question of whether U-GRID-RP admits a polynomial time approximation scheme or not.

THEOREM 6. *The labeled version in the grid L-GRID-RP is NP-complete.*

Proof. The argument, similar to that in the proof of Theorem 5, still uses a reduction from R-STEINER. Consider an instance $P = \{p_1, \dots, p_n\}$ of R-STEINER with n grid points, and let p_n be a rightmost point in P .

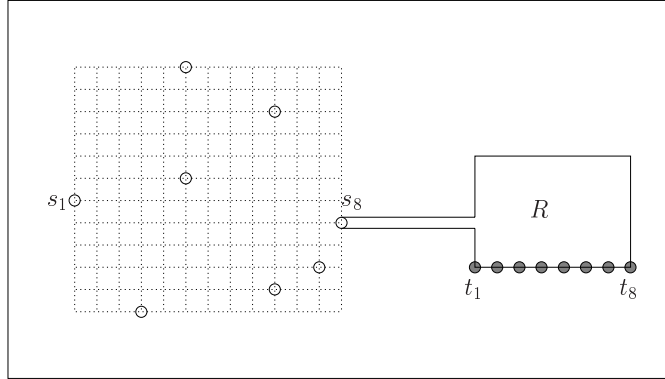


FIG. 5. Labeled chips in the grid obtained from a Rectilinear Steiner Tree instance with $n = 8$ (sketch). Chip-only vertices are drawn as empty circles, and target-only vertices are drawn as filled (black) circles. Obstacles (not shown) are placed at every other integer point in the axis-aligned bounding rectangle. The corridor and most of the target area R (except its bottom row) are free.

The instance of L-GRID-RP, illustrated in Figure 5, has n labeled chips $s_1, \dots, s_n$ located at the points in P (s_i at p_i). Consider a sufficiently large axis-aligned rectangle enclosing the n chips mentioned previously and filled with obstacles placed at every other integer point, except the following: make a thin (width is one), long corridor

connecting the rightmost chip s_n to a large rectangular “target area” R located to the right of all the n chips mentioned previously; both the corridor and all rows of R except its bottom row are free. The bottom row contains all n targets $t_1, \dots, t_n$; here t_i denotes the target of s_i . R has $\Delta_x + \Delta_y + 2$ rows and n columns, where Δ_x, Δ_y are as before.

Informally, the obstacles which form a minimum rectilinear Steiner tree of the n points must go out in the target area and then come back in to their original positions. More precisely, the reduction is complete once we establish the following claim (analogous to the claim in the proof of Theorem 5).

Claim. There is a rectilinear Steiner tree of length at most q if and only if reconfiguration can be done using at most $2q + 2 - n$ moves.

This concludes the proof of Theorem 6. $\square$

3.2. Labeled chips: Upper and lower bounds on the number of moves.

THEOREM 7. *For the reconfiguration of n labeled chips in the infinite planar rectangular grid (L-GRID-RP), $7n/4$ moves are always enough, while $3n/2$ moves are sometimes necessary.*

Proof. The lower bound is trivial (however, it does not appear to be trivial to improve on it): take a pair of chips labeled 1 and 2, say next to each other, and have the target positions switch them; that is, $t_1 = s_2$ and $t_2 = s_1$. Clearly three moves are needed to rearrange this group of two, and by repeating it (with different labels), one gets a pair of configurations which require $3n/2$ moves.

We now describe a reconfiguration algorithm which executes no more than $7n/4$ moves (as mentioned in the introduction, the problem can be solved trivially in $2n$ moves).

Let S and T be the start and target configurations. Consider the directed graph G with n edges (loops allowed) given by $S \rightarrow T$. Vertices are grid points of $S \cup T$ (the number of vertices is between n and $2n$). Each edge originates at a start chip and ends at some (free or occupied) target cell. Note that each in-degree and out-degree is at most one, so G can be partitioned into a collection of disjoint paths and cycles (and loops).

Consider the rows of S numbered from top to bottom: $1, 2, \dots, r$. Let D (resp., E) be the set of elements in the odd (resp., even) rows; we can assume w.l.o.g. that $|D| \leq |E|$, thus $|D| \leq n/2$. Let A be the set of elements of E whose target lies in E , and let B (resp., C) be the set of elements of E whose target lies in rows of D congruent to 1 (resp., to 3) modulo 4. Write $a = |A|$, $b = |B|$, $c = |C|$. We can assume without loss of generality that $c \leq b$.

1. Move (far) away the elements of D (row by row, and for each row, move elements one by one, say from left to right) to form a set of corridors.
2. Move away the elements of C (elements of the even rows are adjacent to corridors, therefore any subset of chips of an even row can be moved away).
3. Select and move away an element from each cycle of the directed graph G remaining among the elements of A (not from the loops).
4. Fill (say, from left to right) the odd rows congruent to 1 modulo 4 with the elements of B and elements far away as follows: note that each even row is adjacent to an odd row congruent to 3 modulo 4; take out an element of B from the even row through the empty corridor congruent to 3 modulo 4, and then move the element to its target on the odd row congruent to 1 modulo 4.
5. Fill the even rows using the adjacent empty corridors (congruent to 3 modulo 4), with elements from A and elements far away. The elements of A move

directly to their destination, and such a move is possible as long as some elements of A still need to move, since we moved away one element from each cycle of the directed graph G contained in A .

6. Fill (say, from left to right) the odd rows congruent to 3 modulo 4 (the corridors) with elements far away.

The number of nontarget moves is at most

$$n - (a + b + c) + \frac{a}{2} + c \leq n - \frac{a}{2} - b \leq \frac{3n}{4},$$

since $a + 2b \geq a + b + c \geq n/2$. Therefore the total number of moves is not more than $n + 3n/4 = 7n/4$. $\square$

Remark. The above lower bound clearly holds even in the stronger *lifting model*, when chips can be lifted and placed back in the plane (see [7, 8] for related aspects of disk reconfiguration problems).

Acknowledgment. The authors thank Sergey Bereg and Marius Zimand for several conversations on this topic.

REFERENCES

- [1] M. ABELLANAS, S. BEREG, F. HURTADO, A. G. OLAVERRI, D. RAPPAPORT, AND J. TEJEL, *Moving coins*, Comput. Geom., 34 (2006), pp. 35–48.
- [2] P. ALIMONTI AND V. KANN, *Hardness of approximating problems on cubic graphs*, in Proceedings of the 3rd Italian Conference on Algorithms and Complexity, Rome, Italy, Lecture Notes in Comput. Sci. 1203, Springer-Verlag, 1997, pp. 288–298.
- [3] A. F. ARCHER, *A modern treatment of the 15 puzzle*, Amer. Math. Monthly, 106 (1999), pp. 793–799.
- [4] S. ARORA, *Nearly linear time approximation schemes for Euclidean, TSP, and other geometric problems*, J. ACM, 45 (1998), pp. 1–30.
- [5] V. AULETTA, A. MONTI, M. PARENTE, AND P. PERSIANO, *A linear-time algorithm for the feasibility of pebble motion in trees*, Algorithmica, 23 (1999), pp. 223–245.
- [6] R. BAR-YEHUDA, *One for the price of two: A unified approach for approximating covering problems*, Algorithmica, 27 (2000), pp. 131–144.
- [7] S. BEREG, A. DUMITRESCU, AND J. PACH, *Sliding disks in the plane*, Internat. J. Comput. Geom. & Appl., to appear.
- [8] S. BEREG AND A. DUMITRESCU, *The lifting model for reconfiguration*, Discrete Comput. Geom., 35 (2006), pp. 653–669.
- [9] A. DUMITRESCU, *Motion planning and reconfiguration for systems of multiple objects*, in Mobile Robots: Perception & Navigation, S. Kolski, ed., Advanced Robotic Systems, 2007, pp. 523–542.
- [10] A. DUMITRESCU AND J. PACH, *Pushing squares around*, Graphs Combin., 22 (2006), pp. 37–50.
- [11] A. DUMITRESCU, I. SUZUKI, AND M. YAMASHITA, *Motion planning for metamorphic systems: Feasibility, decidability and distributed reconfiguration*, IEEE Trans. Robotics and Automation, 20 (2004), pp. 409–418.
- [12] M. GAREY AND D. JOHNSON, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, San Francisco, 1979.
- [13] M. X. GOEMANS AND D. P. WILLIAMSON, *A general approximation technique for constrained forest problems*, SIAM J. Comput., 24 (1995), pp. 296–317.
- [14] M. X. GOEMANS AND D. P. WILLIAMSON, *The primal-dual method for approximation algorithms and its application to network design problems*, in Approximation Algorithms for NP-Hard Problems, D. S. Hochbaum, ed., PWS Publishing, Boston, 1995, pp. 144–186.
- [15] W. W. JOHNSON, *Notes on the 15 puzzle. I.*, Amer. J. Math., 2 (1879), pp. 397–404.
- [16] D. KORNHAUSER, G. MILLER, AND P. SPIRAKIS, *Coordinating pebble motion on graphs, the diameter of permutation groups, and applications*, in Proceedings of the 25th Symposium on Foundations of Computer Science (FOCS '84), IEEE Press, Piscataway, NJ, 1984, pp. 241–250.

- [17] C. PAPADIMITRIOU, P. RAGHAVAN, M. SUDAN, AND H. TAMAKI, *Motion planning on a graph*, in Proceedings of the 35th Symposium on Foundations of Computer Science (FOCS '94), IEEE Press, Piscataway, NJ, 1994, pp. 511–520.
- [18] C. H. PAPADIMITRIOU AND M. YANNAKAKIS, *Optimization, approximation, and complexity classes*, J. Comput. System Sci., 43 (1991), pp. 425–440.
- [19] D. RATNER AND M. WARMUTH, *Finding a shortest solution for the $(N \times N)$ -extension of the 15-puzzle is intractable*, J. Symbolic Comput., 10 (1990), pp. 111–137.
- [20] W. E. STORY, *Notes on the 15 puzzle. II.*, Amer. J. Math., 2 (1879), pp. 397–404.
- [21] R. M. WILSON, *Graph puzzles, homotopy, and the alternating group*, J. Combin. Theory Ser. B, 16 (1974), pp. 86–96.

LIST-COLORING SQUARES OF SPARSE SUBCUBIC GRAPHS*

ZDENĚK DVOŘÁK[†], RISTE ŠKREKOVSKI[‡], AND MARTIN TANCER[†]

Abstract. The problem of coloring the square of a graph naturally arises in connection with the distance labelings, which have been studied intensively. We consider this problem for sparse subcubic graphs. We show that the choosability $\chi_\ell(G^2)$ of the square of a subcubic graph G of maximum average degree d is at most four if $d < 24/11$ and G does not contain a 5-cycle, at most five if $d < 7/3$, and at most six if $d < 5/2$. Wegner's conjecture claims that the chromatic number of the square of a subcubic planar graph is at most seven. Let G be a planar subcubic graph of girth g . Our result implies that $\chi_\ell(G^2)$ is at most four if $g \geq 24$, at most 5 if $g \geq 14$, and at most 6 if $g \geq 10$. For lower bounds, we find a planar subcubic graph G_1 of girth 9 such that $\chi(G_1^2) = 5$ and a planar subcubic graph G_2 of girth 5 such that $\chi(G_2^2) = 6$. As a consequence, we show that the problem of 4-coloring of the square of a subcubic planar graph of girth $g = 9$ is NP-complete. We conclude the paper by posing a few conjectures.

Key words. planar graphs, sparse graphs, square of a graph, list coloring, complexity, NP-complete

AMS subject classification. 05C15

DOI. 10.1137/050634049

1. Introduction. We study the coloring of squares of graphs and consider simple undirected graphs. Let $\chi(G)$ and $\chi_\ell(G)$ be the chromatic number and the choosability of a graph G , respectively. The *square* G^2 of a graph G is the graph with the same vertex set in which two vertices are joined by an edge if their distance in G is at most 2. It is easy to see that $\Delta + 1 \leq \chi(G^2) \leq \chi_\ell(G^2) \leq \Delta^2 + 1$, where Δ is the maximum degree of G . However, by Brooks' theorem, it is not hard to infer that there are only finitely many connected graphs for which the upper bound is attained. On the other hand, the chromatic number of the square of a planar graph is bounded by a function linear in the maximum degree (note that this bound does not follow directly from the 5-degeneracy of planar graphs [8]).

The notion of coloring of the square of a graph, as well as the other variants of a distance coloring, arise from the problem of assigning frequencies to transmitters. We are given a set of locations of transmitters, and we want to assign frequencies to them in such a way that they do not interfere. The transmitters interfere if they are close enough and the frequencies assigned to them are similar. The choosability version of the problem corresponds to the situation where we are not allowed to use all frequencies on all transmitters. The notion of closeness can often be approximated by forming the graph of the neighboring transmitters, and coloring of the square or a higher power of this graph enables us to take into account the interference between farther transmitters.

*Received by the editors June 20, 2005; accepted for publication (in revised form) August 8, 2007; published electronically February 15, 2008.

<http://www.siam.org/journals/sidma/22-1/63404.html>

[†]Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Malostranské náměstí 25, 118 00 Prague, Czech Republic (rakdver@kam.mff.cuni.cz, martin@atrey.karlin.mff.cuni.cz). These authors were partially supported by the Institute for Theoretical Computer Science (ITI) under project 1M0021620808 of the Ministry of Education of Czech Republic.

[‡]Department of Mathematics, University of Ljubljana, Jadranska 19, 1111 Ljubljana, Slovenia (skreko@fmf.uni-lj.si). This author was partially supported by the Ministry of Science and Technology of Slovenia under Research Program P1-0297.

Let us briefly survey the rich history of the coloring of the squares of planar graphs. Wegner [16] proved that the squares of cubic planar graphs are 8-colorable. He conjectured that his bound can be improved as follows.

CONJECTURE 1. *Let G be a planar graph with maximum degree Δ . The chromatic number of G^2 is at most 7 if $\Delta = 3$, at most $\Delta + 5$ if $4 \leq \Delta \leq 7$, and at most $\lfloor \frac{3\Delta}{2} \rfloor + 1$ otherwise.*

If this conjecture were true, then the bounds would be the best possible. The reader is referred to section 2.18 in [9] for more details. Although Conjecture 1 has been verified for several special classes of planar graphs, including the outerplanar graphs [11], it remains open for all values of $\Delta \geq 3$. The best known upper bounds are due to Molloy and Salavatipour [12, 13]: $\lceil 5\Delta/3 \rceil + 78$ for all Δ and $\lceil 5\Delta/3 \rceil + 25$ for $\Delta \geq 241$. Sharper results can be obtained for special classes of graphs. Dvořák et al. [5] proved that the chromatic number of the square of a planar graph G with a sufficiently large maximal degree is $\Delta + 1$ if the girth of G is at least 7 and is bounded by $\Delta + 2$ if the girth of G is 6. Independently, Borodin et al. [2] proved a similar result: The chromatic number of the square of a planar graph G with girth 7 and $\Delta \geq 30$ is $\Delta + 1$, and for every $\Delta \geq 3$, there exist planar graphs with girth 6 whose squares have chromatic number $\Delta + 2$.

A graph G is called *subcubic* if $\Delta(G) \leq 3$. We consider (not necessarily planar) subcubic graphs of small maximum average degree d and show the following bounds on the choosability of G^2 (section 3):

- If $d < 24/11$ and G does not contain a 5-cycle, then $\chi_\ell(G^2) \leq 4$.
- If $d < 7/3$, then $\chi_\ell(G^2) \leq 5$.
- If $d < 5/2$, then $\chi_\ell(G^2) \leq 6$.

For planar subcubic graphs of large girth g , this implies the following (section 4):

- If $g \geq 24$, then $\chi_\ell(G^2) \leq 4$.
- If $g \geq 14$, then $\chi_\ell(G^2) \leq 5$.
- If $g \geq 10$, then $\chi_\ell(G^2) \leq 6$.

Montassier and Raspaud [14] have shown similar results for coloring of the squares of planar subcubic graphs. They have shown that $\chi(G^2) \leq 5$ if $g \geq 14$ and $\chi(G^2) \leq 6$ if $g \geq 10$, and their proof can be easily generalized to work for choosability as well. They use some of the same reducible configurations that we do; however, their proofs of reducibility are simpler because they use the girth assumptions (not just the sparseness of the graph).

Regarding the lower bounds, we construct subcubic planar graphs G_1 and G_2 of girths $g(G_1) = 9$ and $g(G_2) = 5$ such that $\chi(G_1^2) = 5$ and $\chi(G_2^2) = 6$ (section 5).

The problem of finding the chromatic number of the square of a planar graph is NP-complete. It is NP-complete even to decide whether the square of a planar graph can be colored by seven colors [15]—note that this implies that determining the chromatic number of the square of a planar graph is NP-complete when restricted to graphs of maximum degree 6. It is also NP-complete to decide whether the square of a cubic (not necessarily planar) graph can be colored by four colors [7]. On the other hand, the problem can be solved in polynomial time for partial k -trees [17]. For further complexity and structural results we refer to the survey [3]. We are interested in the complexity of the problem for planar graphs of large girth. The graphs used in the reduction of Ramanathan [15] contain many triangles, and thus they do not provide any guidance in this direction.

We show that the problem of determining the chromatic number of the square of a graph is NP-complete even for subcubic planar graphs of girth 9 (section 6);

more precisely, we show that it is NP-complete to decide whether such a graph can be colored by four colors.

2. Notation. Let $\mathbb{N}$ denote the set of all nonnegative integers. A d -vertex is a vertex of degree d . The *maximum average degree* $d(G)$ of a graph G is the maximum of $2|E(H)|/|V(H)|$ over all induced subgraphs H of G .

A k -cycle is a cycle of length k . The *girth* g of G is the length of the shortest cycle in G , and it is infinity if G is a forest. A *thread* is an induced path in G whose vertices are all of degree 2 in G . A k -thread for $k \geq 1$ is a thread with k vertices. The *length* of a thread is the number of its vertices, i.e., a k -thread has length k .

If G is a connected plane graph, then let $F(G)$ be the set of all faces of G . Denote by $\ell(f)$ the length of a face f (we count multiple occurrences of an edge if f is not a simple cycle, which can happen if G is not 2-connected). An ℓ -face is a face of length ℓ .

We say that a graph G is k -minimal, if G^2 is not k -choosable, but the square of every proper subgraph of G is k -choosable. Obviously, a k -minimal graph is connected. A *configuration* is an induced subgraph of G , and we say that a configuration is k -reducible if it cannot appear in a k -minimal graph. Note that a k -reducible configuration is also k' -reducible for every $k' \geq k$.

We introduce the following notation to simplify the description of reducible configurations: Let $Y_{a,b,c}$ be a 3-vertex v together with an a -thread, a b -thread, and a c -thread incident to v such that the threads are vertex-disjoint and there are no edges between the end-vertices of the threads. Note, however, that the end-vertices may have a common neighbor. We allow the possibility that $a = 0$ or $b = 0$ or $c = 0$, and in this case v is adjacent to one or more vertices of degree 3.

Let $Y_{a,b} - j - Y_{c,d}$ be two 3-vertices v_1 and v_2 joined by a j -thread together with an a -thread and a b -thread incident to v_1 and a c -thread and a d -thread incident to v_2 . Again, the threads are vertex-disjoint and there are no edges between the end-vertices of the threads. We also allow that a, b, j, c , or d could be 0. Thus if $j = 0$, then v_1 and v_2 are adjacent.

Let G be a graph and let L be an assignment of lists to the vertices of G . We say that L is a p -list assignment for a given function $p : V(G) \rightarrow \mathbb{N}$ if $|L(v)| = p(v)$ for each vertex v of G . If p is a constant function with value k , we say that L is a k -list assignment.

3. Coloring sparse graphs. In this section we show that the squares of sparse subcubic graphs can be colored by few colors. We proceed by showing that if a subgraph induced by vertices close to any vertex v is too sparse, then it is k -reducible for given k . Thus we show that in the k -minimal graph, each such subgraph is dense, and thus that the k -minimal graph itself is dense, which is a contradiction.

We first identify the reducible configurations in subsection 3.1. Then, in subsection 3.2, we discuss the possible neighborhoods of vertices in a k -minimal graph. Finally, we derive the bounds on the density of k -minimal graphs for $k = 4, 5$, and 6.

3.1. Reducible configurations. In this section, we give k -reducible configurations for $k = 4, 5$, and 6. Let us first show that a k -minimal graph has minimum degree at least 2.

LEMMA 1. *A 1-vertex is a k -reducible configuration for each $k \geq 4$.*

Proof. Let G be a k -minimal graph and let L be a k -list assignment of G such that G^2 is not L -colorable. Assume for contradiction that G contains a 1-vertex v . By the k -minimality of G , the graph $(G - v)^2 = G^2 - v$ has an L -coloring c . The

vertex v has at most three neighbors in G^2 . We let $c(v)$ be a color that does not appear in the neighborhood of v in G^2 . Thus we extend c to an L -coloring of G^2 , a contradiction. $\square$

Let R be a configuration in a graph G . We define the function $s_R : V(R) \rightarrow \mathbb{N}$. The value $s_R(v)$ is the number of neighbors of v in G^2 that are not its neighbors in R^2 , i.e., $s_R(v) = \deg_{G^2}(v) - \deg_{R^2}(v)$. Note that some of such neighbors may belong to R .

Suppose that we want to show k -reducibility of a configuration R . We can use the following approach: We assume that R appears in a k -minimal graph G . We remove R from G and color the graph $G \setminus R$, which is always possible by k -minimality of G . We want to extend this coloring to the vertices of R . For each vertex v of R , we remove the colors used at the vertices in a distance of at most 2 from the list of v . At most $s_R(v)$ colors are removed in this way. Then we consider the coloring of R from these new lists. In the next lemma, we describe this idea more precisely.

LEMMA 2. *Let R be a configuration in a graph G and p a function from vertices of R such that $p(v) \leq k - s_R(v)$ for every vertex v of R . Suppose that*

- (a) *no two vertices of $G \setminus R$ have a common neighbor in R , and*
- (b) *no two vertices u and v of R such that $uv \notin E(R^2)$ have a common neighbor in $G \setminus R$.*

If R^2 can be colored from any p -list assignment L , then G is not k -minimal.

Proof. Let G be a k -minimal graph and let L be a k -list assignment such that G is not L -colorable. Assume for the contradiction that G contains the configuration R satisfying the assumptions of the lemma. Let $G' = G - R$. By the k -minimality of G , the graph G'^2 has an L -coloring c . We show that we can extend this coloring to G^2 , which is a contradiction.

We remove the colors used on the neighbors in G^2 from the lists of the vertices of R . The length of the new list of each vertex $v \in V(R)$ is at least $k - s_R(v) \geq p(v)$. We can color R^2 from these lists, thus obtaining the values of the coloring c on $V(R)$. By the assumptions of the lemma, the subgraph of G^2 induced by $V(R)$ is equal to R^2 and the subgraph induced by $V(G')$ is equal to G'^2 . Consequently, c is a proper L -coloring of G^2 , a contradiction. $\square$

Condition (a) of the above lemma is equivalent to saying that each vertex of R has at most one neighbor outside of R . This is satisfied in all of our applications of this lemma. Condition (b) is more complicated to satisfy. In the case $k = 4$, all the configurations for that lemma are used satisfy this property trivially. In the case of $k \geq 5$, we need to consider the cases when two separate vertices of R share a common neighbor outside of R .

Using $p(v) = k - s_R(v)$ in the applications of this lemma is not practical, since it requires us to know the exact neighborhood of R in G . Instead, we use upper bounds on s derived from the knowledge about the degrees of vertices in the neighborhood of R . In particular, we often use $p(v) = k - s'_R(v)$, where

$$s'_R(v) = 3(\deg_G(v) - \deg_R(v)) + \sum_{uv \in E(R)} (\deg_G(u) - \deg_R(u)).$$

A *block* of a graph G is its maximal 2-connected subgraph. A connected graph G is said to be a *Gallai tree* if each of the blocks of G is a complete graph or an odd cycle. In several proofs of the reducibility of configurations, we use the following theorem, which is proved independently by Borodin [1] and Erdős, Rubin, and Taylor [6].

THEOREM 3. *Let G be a connected graph with a list assignment L such that $|L(v)| \geq \deg(v)$ for each vertex v of G and G is not L -colorable. Then, $|L(v)| = \deg(v)$ for every v and G is a Gallai tree. Moreover, if G is 2-connected, then the lists $L(v)$ of all the vertices v of G are the same.*

Let us show a stronger version of this theorem for a particular graph.

LEMMA 4. *Let G be a path on four vertices $v_1v_2v_3v_4$ and let L be a list assignment to the vertices of G such that $|L(v_1)| = |L(v_4)| = 2$ and $|L(v_2)| = |L(v_3)| = 3$. Then G^2 has two L -colorings c_1 and c_2 such that $c_1(v_2) \neq c_2(v_2)$.*

Proof. Since G^2 is not a Gallai tree, at least one coloring c_1 of G^2 exists by Theorem 3.

Let $a_1 = c_1(v_2)$ and $L(v_2) = \{a_1, a_2, a_3\}$. Let us assume for the sake of contradiction that all colorings of G use the color a_1 on the vertex v_2 . This means that if we color v_2 by a_2 , then this coloring cannot be extended. Let $L'(v_1) = L(v_1) \setminus \{a_2\}$, $L'(v_3) = L(v_3) \setminus \{a_2\}$, and $L'(v_4) = L(v_4) \setminus \{a_2\}$. By Theorem 3, if v_1 , v_3 , and v_4 cannot be colored from lists L' , then $|L'(v_1)| = |L'(v_4)| = 1$ and $|L'(v_3)| = 2$. But then $a_2 \in L(v_1) \cap L(v_4)$. Now observe that the coloring c_2 defined by $c_2(v_1) = c_2(v_4) = a_2$, $c_2(v_2) = a_3$, and $c_2(v_3) \in L(v_3) \setminus \{a_2, a_3\}$ satisfies the requirements of the lemma. $\square$

Let us first consider cycles that contain at most one 3-vertex in k -minimal graphs.

LEMMA 5. *Let R be a configuration formed by a cycle $v_1v_2 \cdots v_s$ such that $v_2, v_3, \dots, v_s$ are 2-vertices. Then R is k -reducible for $k \geq 4$ unless $k = 4$ and $s = 5$.*

Proof. Since there is at most one edge joining R and $G \setminus R$, all the assumptions of Lemma 2 are trivially satisfied. It suffices to show 4-reducibility if $s \neq 5$ and 5-reducibility if $s = 5$, and we may assume that v_1 is a 3-vertex.

Suppose first that $s \neq 5$. We use the function p defined by $p(v_1) = 1$, $p(v_2) = p(v_s) = 3$, and $p(v_i) = 4$ for the remaining vertices. Let L be any p -list assignment. We need to show that R^2 can be colored from L . Observe that this is the case if $s = 3$ or $s = 4$.

In case $s > 5$, let $c(v_1)$ be the single color in $L(v_1)$. Let $p'(v_2) = p'(v_s) = 2$, $p'(v_3) = p'(v_{s-1}) = 3$, and $p'(v_i) = 4$ for the remaining vertices, and let L' be an arbitrary p' -list assignment such that $L'(v_i) \subseteq L(v_i) \setminus \{c(v_1)\}$ for $i = 2, 3, s-1$; and s and $L'(v_i) = L(v_i)$ for the remaining vertices. We need to find a coloring of vertices $v_2, \dots, v_s$ from the lists L' .

If $s > 6$, we choose $c(v_2) \in L'(v_2)$ and $c(v_s) \in L'(v_s)$ so that $c(v_2) \neq c(v_s)$. We remove the color $c(v_2)$ from the lists $L'(v_3)$ and $L'(v_4)$ and the color $c(v_s)$ from the lists $L'(v_{s-1})$ and $L'(v_{s-2})$ (note that all these vertices are mutually distinct, since $s > 6$). Finally, we color the subgraph of R^2 induced by the path $v_3 \cdots v_{s-1}$ from the new lists using Theorem 3, thus obtaining a proper coloring c .

Finally, consider the case $s = 6$. If $L'(v_2) \neq L'(v_6)$, we select $c(v_2)$ from $L'(v_2) \setminus L'(v_6)$, remove $c(v_2)$ from the lists of v_3 and v_4 , and color the vertices v_3, v_4, v_5 , and v_6 using Lemma 4. Therefore, suppose that $L'(v_2) = L'(v_6) = \{a_1, a_2\}$. If $L'(v_2) \subset L'(v_3)$ and $L'(v_2) \subset L'(v_5)$, then set $c(v_2) = c(v_5) = a_1$, $c(v_3) = c(v_6) = a_2$ and choose $c(v_4)$ from $L'(v_4) \setminus L'(v_2)$ arbitrarily. If this is not the case, we may assume that $a_1 \notin L'(v_3)$. Then we set $c(v_2) = a_1$, $c(v_6) = a_2$, remove the color a_1 from the list of v_4 and the color a_2 from the lists of v_4 and v_5 , and then color v_3, v_4 , and v_5 from their lists using Theorem 3.

In the case $s = 5$, we let $p(v_1) = 2$, $p(v_2) = p(v_5) = 4$ and $p(v_3) = p(v_4) = 5$. Note that we can color the vertices v_1, v_2, v_5, v_3 , and v_4 one by one in this order. $\square$

We are now ready to identify several 4-reducible configurations.

LEMMA 6. *The following configurations are 4-reducible:*

- (1) *a 6-thread;*
- (2) $Y_{1,4,5};$
- (3) $Y_{2,3,4};$
- (4) $Y_{5,5} - 0 - Y_{4,5};$ *and*
- (5) *a 7-cycle $v_1v_2 \cdots v_7$ such that v_1 incident to a 2-thread x_1x_2 , the vertex v_2 incident to a 2-thread y_1y_2 (where the 2-threads are vertex-disjoint), and each v_i is a 2-vertex for $3 \leq i \leq 7$.*

Proof. Let us prove the reducibility of each configuration separately. Let R be one of the configurations. We use Lemma 2 for the configuration R' obtained from R by removing all 2-vertices that have degree 1 in R , i.e., those that have a neighbor in $G - R$. We remove the vertices x_2 and y_2 in case (5) even if they are adjacent. Note that we then know that the neighbors of R' in $G - R'$ are mutually distinct 2-vertices, and thus the assumptions of Lemma 2 are trivially satisfied for R' . We can also use the function p defined for vertices of R' as

$$p(v) = 4 - 2(\deg_G(v) - \deg_{R'}(v)) - \sum_{uv \in E(R')} (\deg_G(u) - \deg_{R'}(u)).$$

It suffices to show that the configuration R' can be colored from any p -list assignment L . This will imply that R' with the additional assumptions on the degrees of surrounding vertices cannot appear in a 4-minimal graph, and hence that R is reducible.

- (1) Let R be a 6-thread. In this case, R' is a 4-thread $v_1v_2v_3v_4$ with $p(v_1) = p(v_4) = 2$ and $p(v_2) = p(v_3) = 3$. By Theorem 3, R'^2 can be colored from the lists L .
- (2) Let R be $Y_{1,4,5}$. In this case, the configuration R' is a path $u_3u_2u_1vw_1w_2w_3w_4$, with $p(w_4) = p(u_3) = p(v) = 2$, $p(w_1) = p(w_3) = p(u_1) = p(u_2) = 3$, and $p(w_2) = 4$.

First, we try to construct an L -coloring of R'^2 in the following way. We fix a coloring c' of the subgraph of R'^2 induced by vertices v , u_1 , u_2 , and u_3 . Notice that such a coloring always exists by Lemma 4. We choose $c'(w_1) \in L(w_1) \setminus \{c'(v), c'(u_1)\}$, and afterwards set $L'(w_2) = L(w_2) \setminus \{c'(v), c'(w_1)\}$, $L'(w_3) = L(w_3) \setminus \{c'(w_1)\}$, and $L'(w_4) = L(w_4)$. Each of the lists L' contains at least two colors. If we can color w_2 , w_3 , and w_4 from lists L' , then we obtain a proper coloring of R'^2 . If this is not the case, then by Theorem 3, $L'(w_2) = L'(w_3) = L'(w_4) = \{a_1, a_2\}$ for some two colors a_1 and a_2 . Let $a_3 = c'(w_1)$, $a_4 = c'(v)$, and $a_5 = c'(u_1)$. Then $L(w_3) = \{a_1, a_2, a_3\}$ and $L(w_2) = \{a_1, a_2, a_3, a_4\}$.

Note that if we can choose color for w_1 distinct from a_3 , i.e., if $|L(w_1) \setminus \{a_4, a_5\}| > 1$, then we are able to extend the coloring to the rest of the configuration. Therefore, let us assume that $L(w_1) = \{a_3, a_4, a_5\}$. The colors $a_4 = c'(v)$ and $a_5 = c'(u_1)$ such that the coloring c' cannot be extended to w_1 , w_2 , w_3 , and w_4 are determined uniquely. By Lemma 4, we can choose a coloring c of v , u_1 , u_2 , and u_3 such that $c(u_1) \neq a_5$. Hence, we can extend c to R'^2 .

- (3) Let R be $Y_{2,3,4}$. In this case, the configuration R' is $Y_{1,2,3}$. Let v be the 3-vertex of R' , and let u_1 , w_1w_2 , and $y_1y_2y_3$ be the threads incident to v , where u_1 , w_1 , and y_1 are adjacent to v . The lengths of the lists are $p(u_1) = p(w_2) = p(y_3) = 2$, $p(w_1) = p(y_2) = p(v) = 3$, and $p(y_1) = 4$.

Let us first fix a coloring c' of the subgraph of R'^2 induced by u_1, v, w_1 , and w_2 , which exists by Lemma 4. Let $L'(y_1) = L(y_1) \setminus \{c'(u_1), c'(w_1), c'(v)\}$, $L'(y_2) = L(y_2) \setminus \{c'(v)\}$, and $L'(y_3) = L(y_3)$. Note that $|L'(y_1)| \geq 1$, $|L'(y_2)| \geq 2$, and $|L'(y_3)| = 2$. If y_1, y_2 , and y_3 can be colored from lists L' , we obtain a proper coloring of R'^2 . Assume this is not the case. Then $|L'(y_2)| = 2$ and $L'(y_1) \subseteq L'(y_2) = L'(y_3)$. This implies that $L(y_2) = L(y_3) \cup \{c'(v)\}$. But by Lemma 4 we may choose a coloring c of a subgraph of R'^2 induced by u_1, v, w_1 , and w_2 such that $c(v) \neq c'(v)$, and such a coloring can be extended to R'^2 .

- (4) Let R be $Y_{5,5} - 0 - Y_{4,5}$. The configuration R' is $Y_{4,4} - 0 - Y_{3,4}$. Let v_1 and v_2 be the two adjacent 3-vertices of R' , and let $u_1 u_2 u_3 u_4, w_1 w_2 w_3 w_4, y_1 y_2 y_3 y_4$, and $z_1 z_2 z_3$ be the threads of R' , where u_1 and w_1 are adjacent to v_1 , and y_1 and z_1 are adjacent to v_2 . Notice that $p(u_4) = p(w_4) = p(y_4) = p(z_3) = 2$, $p(u_3) = p(w_3) = p(y_3) = p(z_2) = 3$, and $p(x) = 4$ for every other vertex x of R' .

Suppose that we have a coloring c' of the square of the path

$$P = w_4 w_3 w_2 w_1 v_1 v_2 z_1 z_2 z_3,$$

and suppose that this coloring cannot be extended to a coloring of R'^2 . The colorings of the threads $P_u = u_1 u_2 u_3 u_4$ and $P_y = y_1 y_2 y_3 y_4$ are independent, because the distance of u_1 and y_1 is 3. Assume that c' cannot be extended to P_u . Similarly as in the proof of the second claim of this lemma, we conclude that the lists of vertices of P_u uniquely determine the color $a_u = c'(v_1)$. Also, if c' cannot be extended to P_y , then the lists of vertices of P_y uniquely determine the color $a_y = c'(v_2)$. Let $L'(v_1) = L(v_1) \setminus \{a_u\}$, $L'(v_2) = L(v_2) \setminus \{a_y\}$, and $L'(x) = L(x)$ for the other vertices of P . If we can find an L' -coloring c of P^2 , then c can be extended to a proper coloring of R'^2 .

Let us choose a color $a_1 \in L'(z_2) \setminus L'(z_3)$. Using Theorem 3, fix an L' -coloring c of the square of the path $w_4 w_3 w_2 w_1 v_1 v_2$ such that $c(v_2) \neq a_1$. We extend c to the vertices z_1, z_2 , and z_3 in the following way. We assign lists $L''(z_1) = L'(z_1) \setminus \{c(v_1), c(v_2)\}$, $L''(z_2) = L'(z_2) \setminus \{c(v_2)\}$, and $L''(z_3) = L'(z_3)$ to these vertices. Note that each of these lists has size at least 2. If an L'' -coloring of the square of $z_1 z_2 z_3$ exists, it extends c to a proper L' -coloring of P^2 . If such a coloring does not exist, then by Theorem 3, $L''(z_1) = L''(z_2) = L''(z_3)$. Hence, $L'(z_2) = L'(z_3) \cup \{c(v_2)\}$. But this is impossible, because $c(v_2) \neq a_1$.

- (5) Let R be the last configuration given in the lemma. The configuration R' is induced by vertices $v_1, \dots, v_7, x_1$, and y_1 , and the lengths of the lists are $p(x_1) = p(y_1) = 2$, $p(v_1) = p(v_2) = 3$, and 4 for the remaining vertices. Let us construct the coloring c from the lists L .

We first color the subgraph induced by vertices x_1, v_1, v_2 , and y_1 , which is possible by Lemma 4. In the case when $L(v_4) \setminus L(v_5)$ contains precisely one element, say a , we may also assume that the color $c(v_2)$ is distinct from a . Let us choose colors $c(v_3)$ and $c(v_7)$ from their lists such that they extend this coloring. Then, we construct the assignment of lists to the rest of the vertices in the following way: $L'(v_6) = L(v_6) \setminus \{c(v_1), c(v_7)\}$, $L'(v_5) = L(v_5) \setminus \{c(v_3), c(v_7)\}$, and $L'(v_4) = L(v_4) \setminus \{c(v_2), c(v_3)\}$. If we can extend the coloring c to these vertices from lists L' , we properly color the whole configuration R'^2 . By Theorem 3, this is possible unless $L'(v_4) = L'(v_5) = L'(v_6) = S$ and the length of S is exactly 2. This implies that $L(v_4) = S \cup \{c(v_2), c(v_3)\}$ and

$L(v_5) = S \cup \{c(v_3), c(v_7)\}$. But this means that $c(v_2)$ is the single color of $L(v_4) \setminus L(v_5)$, which contradicts the choice of the coloring of vertices x_1 , v_1 , v_2 , and y_1 . $\square$

Now let us focus on 5-reducible configurations.

LEMMA 7. *The following two configurations are 5-reducible:*

- (1) a 3-thread, and
- (2) $Y_{1,2,2}$.

Proof. Let us prove the reducibility of each configuration R separately. We use Lemma 2. Let L be an arbitrary p -list assignment to the vertices of R , where $p(v) = 5 - s'_R(v)$ for each vertex v of R . In order to be able to apply the lemma, we need to consider the case when two of the vertices of R share a neighbor outside of R , and to show that R can be colored from any p -list assignment.

- (1) The configuration R is a 3-thread $v_1v_2v_3$ with the lengths of lists $p(v_1) = p(v_3) = 2$ and $p(v_2) = 3$. If v_1 and v_3 have a common neighbor outside of R , then R together with this vertex forms a 4-cycle with at least three 2-vertices. Such a configuration is reducible by Lemma 5. Also, if v_1 and v_3 do not have a common neighbor outside R , then we can apply Lemma 2, since R can be colored from any p -list assignment by Theorem 3.
- (2) Let v be the 3-vertex of the configuration and let the threads be u_1 , w_1w_2 , and y_1y_2 , where u_1 , w_1 , and y_1 are the neighbors of v . Let u_2 , w_3 , and y_3 be the neighbors (not necessarily distinct) of u_1 , w_2 , and y_2 outside of the configuration, respectively. Let L be any assignment of lists of length 5 to the vertices of a 5-minimal graph G that contains this configuration and let c be a coloring of the square of $G' = G \setminus \{w_1, v, y_1\}$ from these lists. Since y_2 is a 1-vertex in G' , it has at most three neighbors in G'^2 , and hence we may select one of at least two colors in $L(y_2)$ that are different from the colors of these neighbors for y_2 . Therefore, we may assume that if $c(y_3) \in \{c(w_2), c(w_3)\}$, then $c(y_2) \notin \{c(w_2), c(w_3)\}$.
Let $L'(w_1) = L(w_1) \setminus \{c(w_2), c(w_3), c(u_1)\}$, $L'(y_1) = L(y_1) \setminus \{c(y_2), c(y_3), c(u_1)\}$, and $L'(v) = L(v) \setminus \{c(u_1), c(u_2), c(y_2), c(w_2)\}$. If $|L'(w_1)| = 2$, then $L'(w_1) \neq L'(y_1)$, since otherwise $\{c(w_2), c(w_3)\} = \{c(y_2), c(y_3)\}$, which contradicts our choice of $c(y_2)$. It follows that we can assign mutually distinct colors from the lists L' to y_1 , v , and w_1 , and thus extend c to a proper coloring of G^2 . $\square$

Finally, we identify the 6-reducible configurations.

LEMMA 8. *The following configurations are 6-reducible:*

- (1) a 2-thread;
- (2) a 3-cycle containing two 3-vertices and one 2-vertex;
- (3) a 4-cycle containing two nonadjacent 3-vertices and two 2-vertices; and
- (4) $Y_{1,1} - 1 - Y_{0,1}$.

Proof. Let us prove the reducibility of each configuration R separately. In the first three cases, we use Lemma 2 directly. Let L be an arbitrary p -list assignment to the vertices of R , where $p(v) = 6 - s'_R(v)$ for each vertex v of R . Note that in each of these cases, R^2 is a complete graph; hence regardless of whether any vertices of R share a neighbor outside of R , the assumptions of Lemma 2 are satisfied.

- (1) The configuration R is a 2-thread v_1v_2 such that $p(v_1) = p(v_2) = 2$. The graph R^2 can obviously be colored from L .
- (2) Let v_1 , v_2 , and v_3 be the vertices of the configuration R , where v_3 is the 2-vertex. Then $p(v_1) = p(v_2) = 2$ and $p(v_3) = 4$. The graph R^2 can be colored by Theorem 3.

- (3) Let v_1, v_2, v_3 , and v_4 be the vertices of the configuration R , where v_2 and v_4 are the 2-vertices. Then $p(v_1) = p(v_3) = 3$ and $p(v_2) = p(v_4) = 4$. Again, the graph R^2 can be colored by Theorem 3.
- (4) For this configuration, we do not use Lemma 2. Let v_1 and v_2 be the 3-vertices of the configuration, with the common neighbor z . The 1-threads are u_1, w_1 , and y_1 , where u_1 and w_1 are the neighbors of v_1 and y_1 is a neighbor of v_2 . Let G be the 6-minimal graph in that this configuration appears, and let c be a coloring of the square of the graph $G' = G \setminus \{v_1, z, v_2, y_1\}$. Since u_1 is a 1-vertex in G' , it has at least three colors in its list that are not used on any of its neighbors in G'^2 ; hence we may choose its color so that $c(u_1) \neq c(w_1)$. For a vertex $x \in \{v_1, z, v_2, y_1\}$, let $L'(x)$ be the list be obtained from $L(x)$ by removing the colors of its neighbors in G^2 that are colored by c . The sizes of the lists are $|L'(v_1)| \geq 2$, $|L'(z)| \geq 3$, $|L'(v_2)| \geq 2$, and $|L'(y_1)| \geq 2$. Note that Lemma 4 implies that the square of the path $v_1zv_2y_1$ can be colored from the lists L' . This coloring extends c to a proper coloring of the whole G^2 . $\square$

3.2. Neighborhoods of vertices. Let us consider the possible neighborhoods of 3-vertices in k -minimal graphs. We prove that these neighborhoods cannot contain too many 2-vertices, which we later use to show that any k -minimal graph must be dense. All of the following lemmata are proved by a straightforward case analysis using the results of the previous subsection.

Given a graph G and a vertex v , let $G_2(v)$ be the subgraph of G induced by v and all 2-vertices reachable from v by paths consisting only from 2-vertices.

LEMMA 9. *Let G be a 4-minimal graph different from a 5-cycle and v a 3-vertex in G such that v is not adjacent to any 3-vertex. Then $G_2(v)$ is a subgraph of $Y_{1,3,5}$ or $Y_{1,4,4}$ or $Y_{2,2,5}$ or $Y_{3,3,3}$.*

Proof. Let x, y , and z be the 3-vertices (not necessarily distinct) that are joined by threads with v . By the 4-minimality, G cannot contain a 5-cycle, and any other cycle containing at most one 3-vertex is reducible by Lemma 5. Therefore, x, y , and z are distinct from v . Let l_x, l_y , and l_z be the lengths of the threads joining x, y , and z with v , and assume $1 \leq l_x \leq l_y \leq l_z$. By Lemma 6(1), $l_z \leq 5$, and by Lemma 6(3), $l_x \leq 3$. If $l_x = 1$, then $l_z < 5$ or $l_y < 4$ by Lemma 6(2), and $G_2(v)$ is a subgraph of $Y_{1,3,5}$ or $Y_{1,4,4}$. If $2 \leq l_x \leq 3$, then $l_x = l_y = 2$ (and $G_2(v)$ is a subgraph of $Y_{2,2,5}$) or $l_z \leq 3$ (and $G_2(v)$ is a subgraph of $Y_{3,3,3}$). $\square$

LEMMA 10. *Let G be a 4-minimal graph different from a 5-cycle and let v and w be two adjacent 3-vertices in G . Then, both $G_2(v)$ and $G_2(w)$ are subgraphs of $Y_{0,5,5}$, and at least one of the following conditions holds:*

- both $G_2(v)$ and $G_2(w)$ are subgraphs of $Y_{0,4,5}$; or
- $G_2(v)$ or $G_2(w)$ is a subgraph of $Y_{0,3,5}$; or
- $G_2(v)$ or $G_2(w)$ is a subgraph of $Y_{0,4,4}$.

Proof. Let x and y be the 3-vertices joined by threads with v and let s and t be the 3-vertices joined by threads with w . Let l_x, l_y, l_s , and l_t be the lengths of these threads, and assume $l_x \leq l_y \leq l_t$ and $l_s \leq l_t$. The vertices x and y are distinct from v and the vertices s and t are distinct from w by Lemma 5 and by the minimality of G . By Lemma 6(1), $l_t \leq 5$ and both $G_2(v)$ and $G_2(w)$ are subgraphs of $Y_{0,5,5}$.

Suppose that neither of the conditions of the lemma is satisfied by v and w . Then $l_y = l_s = l_t = 5$ and $4 \leq l_x \leq 5$. But if s and t are distinct from v , then this induces a copy of $Y_{5,5}-0-Y_{4,5}$ in G , which is reducible by Lemma 6(4). Thus we may assume $s = v$. Hence, v and w are a part of a 7-cycle $vwx_1 \cdots x_5$, where x_i are 2-vertices.

However, since $l_x \geq 4$, $l_y \geq 4$, $l_z \geq 4$, the vertices v and w are additionally incident to vertex-disjoint 2-threads, and the configuration is reducible by Lemma 6(5). $\square$

Let us now consider neighborhoods of vertices in 5-minimal graphs.

LEMMA 11. *Let G be a 5-minimal graph and v a 3-vertex in G . Then, $G_2(v)$ is a subgraph of $Y_{0,2,2}$ or $Y_{1,1,2}$.*

Proof. Let x , y , and z be the 3-vertices (not necessarily distinct) that are joined by threads with v . By Lemma 5, none of x , y , and z are equal to v . Let l_x , l_y , and l_z be the lengths of the threads joining x , y , and z with v , and assume $l_x \leq l_y \leq l_z$. By Lemma 7(1), $l_z \leq 2$, and by Lemma 7(2), $l_x = 0$ (then $G_2(v)$ is a subgraph of $Y_{0,2,2}$) or $l_y \leq 1$ (and $G_2(v)$ is a subgraph of $Y_{1,1,2}$). $\square$

For vertices in 6-minimal graphs, the situation may be a bit more complicated.

LEMMA 12. *Let G be a 6-minimal graph and v a 3-vertex in G . Then, either*

- $G_2(v)$ is a subgraph of $Y_{0,1,1}$, or
- $G_2(v)$ is $Y_{1,1,1}$, the 3-vertices x , y , and z incident with $G_2(v)$ are mutually distinct, and all the graphs $G_2(x)$, $G_2(y)$, and $G_2(z)$ are equal to $Y_{0,0,1}$.

Proof. Let x , y , and z be the 3-vertices (not necessarily distinct) that are joined by threads with v . By Lemma 5, x , y , and z are distinct from v . Let l_x , l_y , and l_z be the lengths of the threads joining x , y , and z with v , and again assume $l_x \leq l_y \leq l_z$. By Lemma 8(1), $l_z \leq 1$. If $l_x = 0$, then $G_2(v)$ is a subgraph of $Y_{0,1,1}$. Thus assume that $l_x = 1$ and $G_2(v)$ is $Y_{1,1,1}$. By Lemma 8(3), the vertices x , y , and z are mutually distinct. If $G_2(x)$ is not equal to $Y_{0,0,1}$, then neighborhoods of v and x form a supergraph of $Y_{1,1}-1-Y_{0,1}$, which is reducible by Lemma 8(4). Thus $G_2(x)$ (and similarly $G_2(y)$ and $G_2(z)$) is equal to $Y_{0,0,1}$. $\square$

3.3. Final step. In this section we combine the results of the previous subsections and prove the main results.

If G is a subcubic graph, let $n_2(G)$ be the number of 2-vertices of G and let $n_3(G)$ be the number of 3-vertices of G . Let $V_3(G)$ be the set of 3-vertices of G . By Lemmata 9–12, $G_2(v)$ is equal to $Y_{a,b,c}$ for some a , b , and c whenever G is a k -minimal graph for $k \in \{4, 5, 6\}$. We let $d_2(v) = a + b + c$. Note that

$$\sum_{v \in V_3} d_2(v) = 2n_2.$$

Let G be a 4-minimal graph and let v be a 3-vertex in G such that $G_2(v)$ is distinct from $Y_{0,5,5}$. Let $P(v)$ be the set of all 3-vertices u adjacent to v such that $G_2(u)$ is $Y_{0,5,5}$. Then, we define

$$d_3(v) = d_2(v) + \sum_{u \in P(v)} d_2(u) = d_2(v) + 10|P(v)|.$$

Let us show some estimates on d_2 and d_3 in the minimal graphs.

LEMMA 13. *Let G be a 4-minimal graph and let v be a 3-vertex of G such that $G_2(v)$ is distinct from $Y_{0,5,5}$. Then, $d_3(v) \leq 9(|P(v)| + 1)$.*

Proof. Let $P(v) = \{u_1, \dots, u_s\}$ with $s \leq 3$. Note that by Lemmata 9 and 10, $d_2(x) \leq 10$ for every 3-vertex x of G . Let us consider several cases regarding s :

$s = 0$: If v is not adjacent to any 3-vertex, then $d_2(v) \leq 9$ by Lemma 9. Otherwise, $G_2(v)$ is a subgraph of $Y_{0,5,5}$ by Lemma 10. The graph $G_2(v)$ is distinct from $Y_{0,5,5}$, and thus $d_2(v) \leq 9$ in this case as well.

$s = 1$: By Lemma 10, $G_2(v)$ is a subgraph of $Y_{0,3,5}$ or $Y_{0,4,4}$. Therefore, $d_3(v) = d_2(v) + 10 \leq 18$.

$s = 2$: Since $G_2(v)$ is a subgraph of $Y_{0,5,5}$ and v has at least two neighbors that are 3-vertices, $d_2(v) \leq 5$. Consequently, $d_3(v) = d_2(v) + 20 \leq 27$.

$s = 3$: In this case $d_2(v) = 0$ and $d_3(v) = 30 < 36$. $\square$

LEMMA 14. *Let G be a 6-minimal graph. If u and v are two 3-vertices of G joined by a path consisting of 2-vertices, then $d_2(u) + d_2(v) \leq 4$.*

Proof. We use Lemma 12. If $G_2(v)$ is $Y_{1,1,1}$, then $G_2(u)$ is $Y_{0,0,1}$ and $d_2(u) + d_2(v) = 4$. One can apply a similar argument if $G_2(u)$ is $Y_{1,1,1}$. If neither $G_2(u)$ nor $G_2(v)$ is $Y_{1,1,1}$, then both of them are subgraphs of $Y_{0,1,1}$ and $d_2(u) + d_2(v) \leq 4$. $\square$

We are now ready to prove the main results.

THEOREM 15. *Let G be a subcubic graph of maximum average degree $d < 24/11$. The graph G^2 is 4-choosable if and only if G does not contain a 5-cycle.*

Proof. If G contains a 5-cycle, then G^2 contains a clique of size 5 and therefore G^2 is not 4-colorable. For the other implication, assume for contradiction that G does not contain a 5-cycle and that G^2 is not 4-choosable. Hence G contains a 4-minimal subgraph G' whose average degree is at most $d < 24/11$. We show that the average degree of any 4-minimal graph is at least $24/11$, thus obtaining a contradiction.

By the 4-minimality, G' is connected and contains at least one 3-vertex. Let U be the set of all 3-vertices v of G' such that $G'_2(v)$ is not $Y_{0,5,5}$. Let us now calculate the following sum in two ways:

$$S = \sum_{v \in U} d_3(v).$$

Since $Y_{5,5} - 0 - Y_{5,5}$ is reducible,

$$S = \sum_{v \in V_3(G')} d_2(v) = 2n_2.$$

On the other hand, by Lemma 13,

$$S \leq \sum_{v \in U} 9(|P(v)| + 1) = 9n_3(G').$$

Therefore, $2n_2 = S \leq 9n_3$ and we infer that $n_2 \leq 4.5n_3$. This means that the average degree of G' is

$$\frac{2n_2 + 3n_3}{n_2 + n_3} = 2 + \frac{n_3}{n_2 + n_3} \geq 2 + \frac{n_3}{5.5n_3} = \frac{24}{11}. \quad \square$$

THEOREM 16. *Let G be a subcubic graph of maximum average degree $d < 7/3$. Then, G^2 is 5-choosable.*

Proof. Assume for contradiction that G^2 is not 5-choosable. Then G contains a 5-minimal subgraph G' whose average degree is at most $d < 7/3$. We show that the average degree of any 5-minimal graph is at least $7/3$, thus obtaining a contradiction.

By the 5-minimality, G' is connected and contains at least one 3-vertex. Let us now consider the following sum:

$$S = \sum_{v \in V_3(G')} d_2(v) = 2n_2(G').$$

By Lemma 11, $d_2(v) \leq 4$ for each 3-vertex of G' . Thus, $2n_2(G') = S \leq 4n_3(G')$ and we get that $n_2(G') \leq 2n_3(G')$. This means that the average degree of G' is

$$\frac{2n_2 + 3n_3}{n_2 + n_3} = 2 + \frac{n_3}{n_2 + n_3} \geq 2 + \frac{n_3}{3n_3} = \frac{7}{3}. \quad \square$$

THEOREM 17. *Let G be a subcubic graph of maximum average degree $d < 5/2$. Then G^2 is 6-choosable.*

Proof. Assume for contradiction that G^2 is not 6-choosable. Then G contains a 6-minimal subgraph G' whose average degree is at most $d < 5/2$. We show that the average degree of any 6-minimal graph is at least $5/2$, thus obtaining a contradiction.

By the 6-minimality, G' is connected and contains at least one 3-vertex. Let X be the set of all unordered pairs of 3-vertices $\{u, v\}$ such that u and v are connected by a thread. The length of the thread is at most 1 by Lemma 8(1). Note that two 3-vertices u and v are connected by at most one such path, since otherwise the configuration formed by u , v , and the two paths is reducible by Lemma 8(2) or (3). Let us now calculate the following sum in two ways:

$$S = \sum_{\{u,v\} \in X} [d_2(u) + d_2(v)].$$

Each vertex v appears in exactly three pairs in X (since a cycle with at most one 3-vertex is excluded by Lemma 5); thus the sum is equal to

$$S = 3 \sum_{v \in V_3(G')} d_2(v) = 6n_2(G').$$

On the other hand, consider an arbitrary pair $\{u, v\} \in X$. By Lemma 14, $d_2(u) + d_2(v) \leq 4$. Hence, $S \leq 4|X| = 6n_3(G')$. Thus, $6n_2(G') = S \leq 6n_3(G')$ and we get that $n_2(G') \leq n_3(G')$. This means that the average degree of G' is

$$\frac{2n_2 + 3n_3}{n_2 + n_3} = 2 + \frac{n_3}{n_2 + n_3} \geq 2 + \frac{n_3}{2n_3} = \frac{5}{2}. \quad \square$$

4. Coloring planar graphs of large girth. In this section we show bounds on the choosability of planar subcubic graphs of large girth. A well-known observation is that such graphs have a small maximum average degree.

LEMMA 18. *Let G be a planar graph of girth at least g . Then $d(G) < 2 + 4/(g-2)$.*

Proof. Since every induced subgraph of G is a planar graph of girth at least g , it suffices to show that $d = 2|E(G)|/|V(G)| < 2g/(g-2)$. We may also assume that G is connected. If G is a tree, then obviously $d(G) < 2$. So assume that G has a finite girth. Let $e = |E(G)|$, $v = |V(G)|$, and $f = |F(G)|$. By the Euler formula, $e + 2 = v + f$. The girth of G is at least g , and because G is not a tree every facial walk in G contains a cycle. Therefore, $2e \geq fg$. We obtain

$$\begin{aligned} e &< e + 2 \leq \frac{2e}{d} + \frac{2e}{g}, \\ 1 &< \frac{2}{d} + \frac{2}{g}, \\ d &< 2 + \frac{4}{g-2}. \quad \square \end{aligned}$$

Together with Theorems 15–17, this implies the following.

COROLLARY 19. *Let G be a planar subcubic graph of girth g . Then*

- if $g \geq 24$, then $\chi_\ell(G^2) \leq 4$.
- if $g \geq 14$, then $\chi_\ell(G^2) \leq 5$.
- if $g \geq 10$, then $\chi_\ell(G^2) \leq 6$.

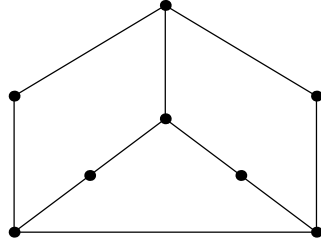
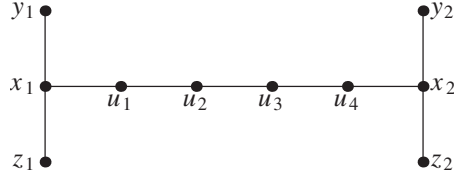


FIG. 1. A 5-minimal graph.

FIG. 2. The graph H_4 .

5. Lower bounds. In this section, we show lower bounds on the girth of the planar subcubic graphs for which the analogue of Corollary 19 holds. It is easy to find a 5-minimal graph of girth 5 as follows.

PROPOSITION 20. *The graph from Figure 1 has girth 5 and its square is not 5-colorable.*

We also show that there exists a 4-minimal graph of girth 9. First, we need to prove several auxiliary lemmata. In their proofs, the following notation is used (where H is the fixed graph that we are trying to color): $(u, v, w) \rightarrow c(x) = k$ stands for the statement “The vertices u , v , and w are the neighbors of vertex x in H^2 . Let $c : V(H) \rightarrow \{1, 2, 3, 4\}$ be a proper coloring of H^2 such that $c(u)$, $c(v)$, and $c(w)$ are three mutually different colors distinct from color k . Then necessarily $c(x) = k$.”

LEMMA 21. *Let H_4 be the graph in Figure 2. Then there is no proper coloring $c : V(H_4) \rightarrow \{1, 2, 3, 4\}$ of H_4^2 such that $c(x_1) = c(x_2)$ and $\{c(y_1), c(z_1)\} \neq \{c(y_2), c(z_2)\}$.*

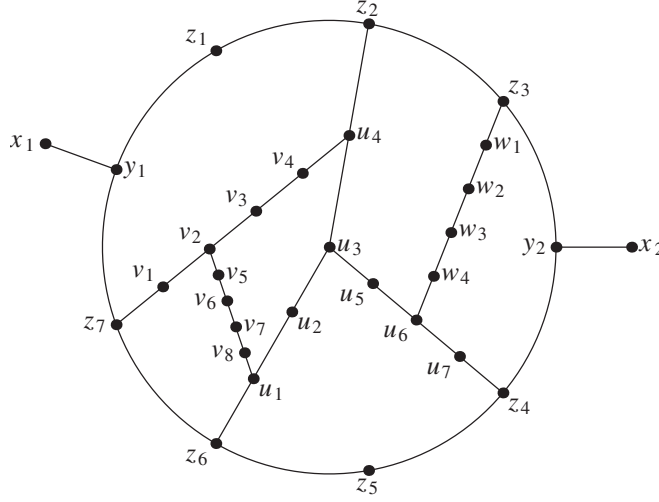
Proof. Without loss of generality, we may assume that $c(x_1) = c(x_2) = 1$, $c(y_1) = c(y_2) = 2$, $c(z_1) = 3$, and $c(z_2) = 4$. Then $(x_1, y_1, z_1) \rightarrow c(u_1) = 4$, and similarly $(x_2, y_2, z_2) \rightarrow c(u_4) = 3$, and then $(x_1, u_1, u_4) \rightarrow c(u_2) = 2$ and $(x_2, u_1, u_4) \rightarrow c(u_3) = 2$. But u_2 and u_3 are neighbors in H_4^2 , and thus the proper coloring does not exist. $\square$

LEMMA 22. *The graph $G_{\neq}(x_1, x_2)$ in Figure 3 has the following properties:*

1. $G_{\neq}$ is a planar subcubic graph.
2. The distance between the vertices x_1 and x_2 is 6.
3. The girth of $G_{\neq}$ is 9.
4. Every proper coloring $c : V(G_{\neq}) \rightarrow \{1, 2, 3, 4\}$ of $G_{\neq}^2$ satisfies $c(x_1) \neq c(x_2)$.

Proof. The first three properties are clear from the figure, and thus it is sufficient to prove the last property. For the contradiction suppose that $c : V(G_{\neq}) \rightarrow \{1, 2, 3, 4\}$ is a proper coloring of $G_{\neq}^2$ such that $c(x_1) = c(x_2)$. We may assume that $c(x_1) = c(x_2) = 1$.

First, we show that $c(y_1) \neq c(y_2)$. If $c(y_1) = c(y_2)$, then we may assume that $c(y_1) = c(y_2) = 2$ and $c(z_3) = 3$. Successively we observe that $(x_1, y_1, z_3) \rightarrow c(z_1) = 4$, $(x_1, y_1, z_1) \rightarrow c(z_7) = 3$, $(x_2, y_2, z_3) \rightarrow c(z_4) = 4$, $(y_1, z_4, z_7) \rightarrow c(z_6) = 1$, and

FIG. 3. The graph $G_{\neq}(x_1, x_2)$.

$(y_2, z_4, z_7) \rightarrow c(z_5) = 1$, but z_5 and z_6 are neighbors in $G_{\neq}^2$. That is a contradiction with assumption $c(y_1) = c(y_2)$.

Hence, we know that $c(y_1) \neq c(y_2)$. We may assume that $c(y_1) = 2$ and $c(y_2) = 3$. Then, $c(z_1) \in \{3, 4\}$ and $c(z_3) \in \{2, 4\}$. Since z_1 and z_3 are adjacent in $G_{\neq}^2$, they have distinct colors. Consider the following cases:

$c(z_1) = 3$ and $c(z_3) = 4$: We infer that $(y_1, z_1, z_3) \rightarrow c(z_2) = 1$, $(x_1, y_1, z_1) \rightarrow c(z_7) = 4$, $(z_1, z_2, z_3) \rightarrow c(u_4) = 2$, $(x_2, y_2, z_3) \rightarrow c(z_4) = 2$, $(z_4, y_2, z_7) \rightarrow c(z_5) = 1$, $(z_5, z_4, z_7) \rightarrow c(z_6) = 3$, $(z_5, z_6, z_7) \rightarrow c(u_1) = 2$, and $(y_2, z_4, z_5) \rightarrow c(u_7) = 4$. Vertex u_2 has neighbors in $G_{\neq}^2$ colored by 2 and 3, and thus $c(u_2) \in \{1, 4\}$. If $c(u_2) = 4$, then all the vertices u_3 , u_5 , and u_6 have neighbors colored by 2 and 4, and thus they are colored by 1 or 3. But it is impossible since they induce a triangle in $G_{\neq}^2$. Therefore, we know that $c(u_2) = 1$. Then $(u_2, u_4, u_7) \rightarrow c(u_5) = 3$, $(u_2, u_4, u_5) \rightarrow c(u_3) = 4$, $(z_2, u_3, u_4) \rightarrow c(v_4) = 3$, $(y_1, z_6, z_7) \rightarrow c(v_1) = 1$, $(z_7, v_1, v_4) \rightarrow c(v_2) = 2$, and $(v_1, v_2, v_4) \rightarrow c(v_3) = 4$. But by Lemma 21 this coloring cannot be extended to vertices v_5 , v_6 , v_7 , and v_8 .

$c(z_1) = 4$ and $c(z_3) = 2$: This case is straightforward: $(y_2, z_1, z_3) \rightarrow c(z_2) = 1$, $(x_1, y_1, z_1) \rightarrow c(z_7) = 3$, $(x_2, y_2, z_3) \rightarrow c(z_4) = 4$, $(y_1, z_4, z_7) \rightarrow c(z_6) = 1$, and $(z_4, z_6, z_7) \rightarrow c(z_5) = 2$. By Lemma 21, this coloring cannot be extended to vertices u_1 , u_2 , u_3 , and u_4 .

$c(z_1) = 3$ and $c(z_3) = 2$: First, observe that $(x_1, y_1, z_1) \rightarrow c(z_7) = 4$ and $(x_2, y_2, z_3) \rightarrow c(z_4) = 4$. Both of the vertices z_2 and u_4 have neighbors in $G_{\neq}^2$ colored by colors 2 and 3, and thus they are colored by 1 or 4. They are adjacent, and thus one of them has to be assigned the color 1 and the other the color 4. Hence, $c(u_3) \in \{2, 3\}$. The vertex z_6 has neighbors with colors 2 and 4, and thus $c(z_6) \in \{1, 3\}$. Now, we distinguish cases according to colors of u_3 and z_6 .

$c(z_6) = 1$ and $c(u_3) = 2$: In this case we infer $(z_6, z_7, u_3) \rightarrow c(u_1) = 3$, $(z_6, u_1, u_3) \rightarrow c(u_2) = 4$, $(z_6, z_7, u_1) \rightarrow c(z_5) = 2$, $(y_2, z_4, z_5) \rightarrow c(u_7) = 1$, $(z_4, u_3, u_7) \rightarrow c(u_6) = 3$, and $(u_2, u_3, u_7) \rightarrow c(u_5) = 3$, but this con-

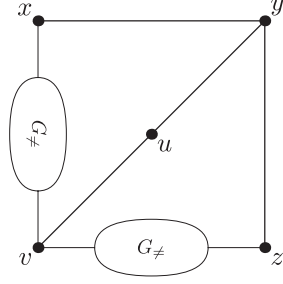


FIG. 4. A graph whose square is not 4-colorable.

tradicts the fact that coloring is proper since u_5 and u_6 are neighbors.

$c(z_6) = 1$ and $c(u_3) = 3$: This is an easy case; $(z_7, z_6, u_3) \rightarrow c(u_1) = 2$, $(z_6, u_1, u_3) \rightarrow c(u_2) = 4$, and $(z_7, z_6, u_1) \rightarrow c(z_5) = 3$. But z_5 and y_2 are neighbors in $G_{\neq}^2$ and both of them are colored by 3. This is a contradiction.

$c(z_6) = 3$ and $c(u_3) = 2$: Note that $(z_6, z_7, u_3) \rightarrow c(u_1) = 1$, $(z_6, u_1, u_3) \rightarrow c(u_2) = 4$, $(z_7, z_6, u_1) \rightarrow c(z_5) = 2$, $(y_2, z_4, z_5) \rightarrow c(u_7) = 1$, $(z_4, u_3, u_7) \rightarrow c(u_6) = 3$, and $(u_2, u_3, u_7) \rightarrow c(u_5) = 3$, but this is a contradiction, since u_5 and u_6 are adjacent in $G_{\neq}^2$.

$c(z_6) = 3$ and $c(u_3) = 3$: If $c(u_4) = 4$, then $(z_1, z_3, u_4) \rightarrow c(z_2) = 1$, but by Lemma 21 this coloring cannot be extended to vertices v_1 , v_2 , v_3 , and v_4 . Thus necessarily $c(u_4) = 1$. Similarly, if $c(u_6) = 2$, then by Lemma 21 the coloring cannot be extended to vertices w_1 , w_2 , w_3 , and w_4 (note that neither u_5 nor u_7 is colored by 3, and thus the assumptions of the lemma are satisfied). Hence, it follows that $c(u_6) = 1$. Next, $(y_2, z_4, u_6) \rightarrow c(u_7) = 2$, $(u_3, u_6, u_7) \rightarrow c(u_5) = 4$, $(u_3, u_4, u_5) \rightarrow c(u_2) = 2$, $(z_7, u_2, u_3) \rightarrow c(u_1) = 1$, and $(z_6, z_7, u_1) \rightarrow c(z_5) = 2$, but z_5 and u_7 are neighbors in $G_{\neq}^2$ colored by the same color, a contradiction. $\square$

THEOREM 23. *There exists a planar subcubic graph G with $g(G) = 9$ such that G^2 is not 4-colorable.*

Proof. The graph consists of vertices x , y , z , u , and v , edges xy , yz , yu , and uv , and copies of $G_{\neq}(z, v)$ and $G_{\neq}(x, y)$ (see Figure 4). By the first three properties of $G_{\neq}$ from Lemma 22, G is a planar graph of girth 9. For the contradiction suppose that G is 4-colorable. Because of the distances of the vertices x , y , z , u , and v , and the forth property of $G_{\neq}$ from Lemma 22, all the vertices x , y , z , u , and v must have mutually distinct colors. However, this contradicts the 4-colorability. $\square$

6. NP-completeness. In this section we discuss the complexity aspects of the problem. Let us first show an auxiliary structural lemma.

LEMMA 24. *There exists a subcubic planar graph $G_{\text{copy}}(x, y)$ of girth 9 with the properties described below. The graph $G_{\text{copy}}(x, y)$ contains two 1-vertices x and y on the outer face. Let v_x and v_y be the neighbors of x and y in $G_{\text{copy}}(x, y)$, respectively. The graph $G_{\text{copy}}(x, y)$ satisfies the following properties:*

- The distance of x and y in $G_{\text{copy}}(x, y)$ is greater than 9;
- $c(x) = c(y)$ in any proper coloring c of the square of $G_{\text{copy}}(x, y)$, and
- for any $a_1, a_2, a_3 \in \{1, 2, 3, 4\}$ such that $a_1 \neq a_2$ and $a_1 \neq a_3$, there exists a proper coloring c of the square of $G_{\text{copy}}(x, y)$ with $c(x) = c(y) = a_1$, $c(v_x) = a_2$, and $c(v_y) = a_3$.

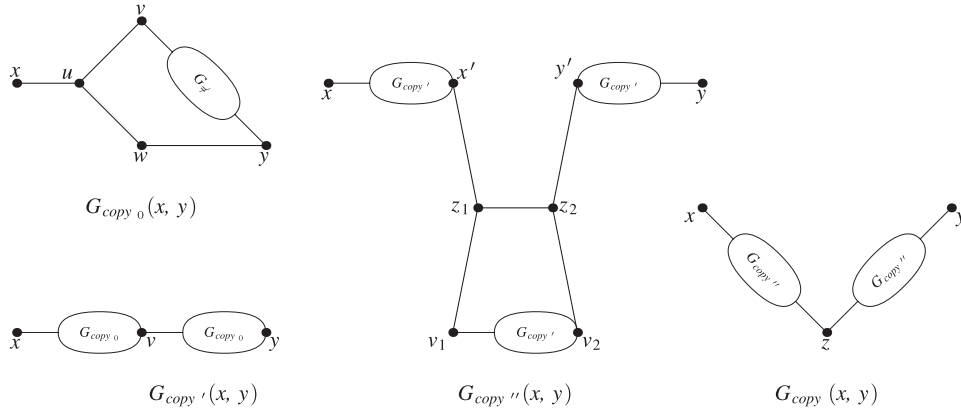


FIG. 5. Graphs used in Lemma 24.

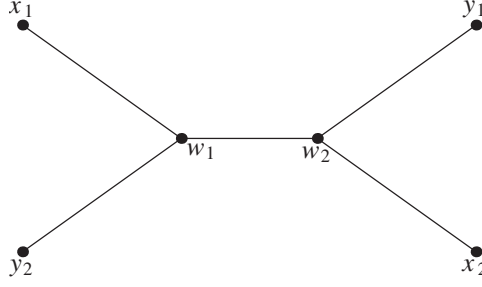
Proof. We use the graph $G_{\neq}$ from Lemma 22. The graph $G_{\neq}$ has the following coloring c_1 : $c_1^{-1}(\{1\}) = \{x_1, y_2, v_1, u_1, u_4, u_6, w_2\}$, $c_1^{-1}(\{2\}) = \{z_1, x_2, z_6, v_2, v_7, u_3, u_7, w_1\}$, $c_1^{-1}(\{3\}) = \{y_1, z_3, z_5, v_4, v_5, v_8, u_5, w_3\}$, and $c_1^{-1}(\{4\}) = \{z_2, z_4, z_7, v_3, v_6, u_2, w_4\}$.

The steps of the construction of $G_{copy}(x, y)$ are depicted in Figure 5. First, we construct an auxiliary subcubic planar graph $G_{copy_0}(x, y)$ of girth 9 such that the vertices x and y are on the outer face, x is a 1-vertex and y is a 2-vertex, and $c(x) = c(y)$ in any proper coloring c of the square of $G_{copy_0}(x, y)$. The graph $G_{copy_0}(x, y)$ consists of vertices x, y, u, v , and w , edges xu, uv, uw , and wy , and a copy of $G_{\neq}(v, y)$. The girth of $G_{copy_0}(x, y)$ is 9, since the distance of v and y in $G_{\neq}(v, y)$ is 6, and they are joined by path $vuwy$ in $G_{copy_0}(x, y)$. The distance of x and y is 3. In any coloring of the square of $G_{copy_0}(x, y)$ by four colors, the vertices x, y, u, v , and w must have mutually distinct colors, with the exception of vertices x and y that can have the same color. Since we color the graph by four colors, the colors of x and y must in fact be equal in any proper coloring of $G_{copy_0}^2(x, y)$. Also note that there exists a proper coloring c_2 of the square of $G_{copy_0}(x, y)$ by four colors—let us color the copy of $G_{\neq}(v, y)$ using the coloring c_1 , and set $c_2(x) = 2$, $c_2(u) = 4$, and $c_2(w) = 3$.

In the second step, we construct the second auxiliary subcubic planar graph $G_{copy'}(x, y)$ of girth 9 such that the vertices x and y are on the outer face, x is a 1-vertex and y is a 2-vertex, $c(x) = c(y)$ in any proper coloring c of the square of $G_{copy'}(x, y)$, and additionally the distance of x and y is at least 6. The graph $G_{copy'}(x, y)$ consists of vertices x, v and y and of the copies of the graphs $G_{copy_0}(x, v)$ and $G_{copy_0}(v, y)$. By properties of G_{copy_0} , the vertices x and y must get the same color in any proper coloring of $G_{copy'}^2(x, y)$. Also, such a coloring exists—we simply use the coloring c_2 on both copies of G_{copy_0} . Obviously, the girth of $G_{copy'}(x, y)$ is 9 and the distance between x and y is exactly 6.

Next we construct the third auxiliary subcubic planar graph $G_{copy''}(x, y)$ of girth 9 such that the 1-vertices x and y are on the outer face, and $c(x) = c(y)$ in any proper coloring c of $G_{copy''}^2(x, y)$. The graph $G_{copy''}(x, y)$ consists of vertices $x, x', y, y', z_1, z_2, v_1$ and v_2 , edges $x'z_1, y'z_2, z_1z_2, z_1v_1, z_2v_2$, and the following copies of graph $G_{copy'}$: $G_{copy'}(x, x')$, $G_{copy'}(y, y')$, and $G_{copy'}(v_1, v_2)$. The distance of x and y is 15, and the cycle consisting of vertices v_1, z_1, z_2, v_2 and of $G_{copy'}(v_1, v_2)$ has length 9, and hence the girth of $G_{copy''}(x, y)$ is 9.

In any coloring c of $G_{copy''}^2(x, y)$, by properties of $G_{copy'}$ it holds that $c(x) = c(x')$,

FIG. 6. $G_{\text{cross}}(x_1, y_1, x_2, y_2)$.

$c(y) = c(y')$, and $c(v_1) = c(v_2)$. Let us assume that exists such a coloring c with $c(x) \neq c(y)$. We may assume that $c(x) = 1$ and $c(y) = 2$, and also that $c(z_1) = 3$ and $c(z_2) = 4$. But then $c(v_1) = c(y') = c(y) = 2$ and $c(v_2) = c(x') = c(x) = 1$, which contradicts the properties of $G_{\text{copy}'}(v_1, v_2)$. On the other hand, there exists the following coloring c_3 of $G_{\text{copy}''}^2(x, y)$: color $G_{\text{copy}'}(x, x')$ by c_2 and $G_{\text{copy}'}(y, y')$ by c_2 with colors 3 and 4 swapped, so that $c_3(x) = c_3(x') = c_3(y) = c_3(y') = 2$, neighbors of x' have colors 1 and 3, and neighbors of y' have colors 1 and 4. Then set $c_3(z_1) = 4$, $c_3(z_2) = 3$, and finally color $G_{\text{copy}'}(v_1, v_2)$ by the coloring c_2 with colors 1 and 2 and colors 3 and 4 swapped, i.e., $c_3(v_1) = c_3(v_2) = 1$.

The graph $G_{\text{copy}''}(x, y)$ satisfies almost all properties of $G_{\text{copy}}(x, y)$, except possibly for the last property that any proper coloring of x, y , and their neighbors can be extended—we only constructed the coloring c_3 where $c_3(x) = c_3(y) = 2$, the color of the neighbor of x is 3, and the color of the neighbor of y is 4. We construct the graph $G_{\text{copy}}(x, y)$ by taking vertices x, z , and y and adding copies of $G_{\text{copy}''}(x, z)$ and $G_{\text{copy}''}(z, y)$. Obviously, the distance of x and y is greater than 9, and the graph has girth 9. Let v_x be the neighbor of x and v_y the neighbor of y in this graph. Let v_{zx} be the neighbor of z in $G_{\text{copy}''}(x, z)$ and v_{zy} be the neighbor of z in $G_{\text{copy}''}(z, y)$. Up to the permutation of colors, there are only two proper colorings of x, y, v_x , and v_y —either $c(x) = c(y) = 1, c(v_x) = 2, c(v_y) = 3$, or $c(x) = c(y) = 1, c(v_x) = 2 = c(v_y) = 2$. In both cases we can color $G_{\text{copy}''}(x, z)$ and $G_{\text{copy}''}(z, y)$ by the permutations of c_3 such that in addition to the prescribed colors on x, y, v_x , and v_y , we have $c(v_{zx}) = 3$ and $c(v_{zy}) = 4$. This shows that indeed, any precoloring of vertices x, y, v_x , and v_y can be extended to a proper coloring of $G_{\text{copy}}^2(x, y)$ and finishes the proof of the lemma. $\square$

The main result of this section follows.

THEOREM 25. *It is NP-complete to decide whether the square of a subcubic planar graph of girth 9 is 4-colorable.*

Proof. We use the copy gadget from Lemma 24. The special vertices x and y of $G_{\text{copy}}(x, y)$ have degree 1. Thus, we can form a binary tree, where the edges are replaced by the copies of $G_{\text{copy}}(x, y)$, and copy the color of the root of the tree to an arbitrary number of vertices of degree 1. Note that by the properties of $G_{\text{copy}}(x, y)$, this graph does not impose any additional constraints on colors of the neighbors of these vertices.

The other gadget we need is a crossover gadget $G_{\text{cross}}(x_1, y_1, x_2, y_2)$, which consists of vertices x_1, y_1, x_2, y_2, w_1 , and w_2 and edges $x_1w_1, w_1y_2, w_1w_2, y_1w_2$, and w_2x_2 ; see Figure 6. In any coloring c of the square of $G_{\text{cross}}(x_1, y_1, x_2, y_2)$ with $c(x_1) \neq c(y_1)$, it holds that $c(x_1) = c(x_2)$ and $c(y_1) = c(y_2)$, there exists such a coloring, and a cyclic order of the 1-vertices of the gadget on the outer face is x_1, y_1, x_2 ,

and y_2 . This gadget allows us to transfer information across the edges of the graph.

The proof of the NP-completeness proceeds by a reduction from the problem of 3-coloring of a planar graph, which is NP-complete due to Dailey [4]. Given an instance of this problem—a 2-connected planar graph G —we need to create a subcubic planar graph H of girth 9 of a size polynomial in $|V(G)|$ such that H^2 is 4-colorable if and only if G is 3-colorable. The idea of the construction is that we put a vertex into each face of G and force the color of all these vertices to be the same, say 4. We then join them to all the vertices of G , so the original vertices must have color 1, 2, or 3. Then we replace the edges by gadgets that ensure large enough girth. A 4-coloring of H^2 then straightforwardly corresponds to a 3-coloring of G .

The graph H is constructed in the following way (see Figure 7): We replace each d -vertex v of G by a tree of copy gadgets that copy the color of v to $2d$ vertices of degree 1. Let $e_1, f_1, e_2, f_2, \dots, e_d, f_d$ be a cyclic order of the edges and the faces incident to v , and let $v_{e_1}, v_{f_1}, \dots, v_{e_d}, v_{f_d}$ be a cyclic order of the copies of v .

We add a vertex u_f for each ℓ -face f of G and add a tree of copy gadgets that copy its color to 2ℓ vertices $u_{v_1}^f, u_{e_1}^f, \dots, u_{v_\ell}^f$, and $u_{e_\ell}^f$, where $v_1, e_1, v_2, e_2, \dots, v_\ell, e_\ell$ is a cyclic order of the vertices and the edges around the face f .

If a vertex v of G is incident to a face f , we add a new vertex x and a copy of the graph $G_{copy}(u_v^f, x)$ and join x and v_f by a 1-thread. This ensures that u_f and v have distinct colors.

If $e = vw$ is an edge of G incident to faces f_1 and f_2 , we add vertices x_1, x_2, x_3 , and x_4 and the following copies of gadgets: $G_{copy}(v_e, x_1)$, $G_{cross}(x_1, u_e^{f_2}, x_2, u_e^{f_1})$, $G_{copy}(x_2, x_3)$, and $G_{copy}(x_4, w_e)$, and join x_3 and x_4 by a 1-thread. Since we already know that u_{f_1} and v have distinct colors, this ensures that the vertices v and w have distinct colors, and that u_{f_1} and u_{f_2} have the same color.

Given a proper 3-coloring G by the colors $\{1, 2, 3\}$, we can color all the vertices u_f of H by the color 4 and preserve the coloring of the copies of the vertices of G inside H . We extend this coloring to the whole H^2 by the properties of the gadgets. Conversely, in a proper 4-coloring of H^2 , all the vertices u_f have the same color (say, 4) and the colors of the original vertices of G are distinct from it, i.e., belong to $\{1, 2, 3\}$. It follows that this coloring restricted to the vertices of G is a proper 3-coloring of G .

The construction of H can be performed in polynomial time, and H is a subcubic planar graph of girth 9. This finishes the proof of the NP-completeness of the problem. $\square$

7. Conclusion. The gap between the upper bounds and the examples showing the lower bounds obtained in sections 4 and 5 is large and it would be of independent interest to find sharper bounds. In particular, we believe that the upper bounds can be improved significantly by taking into account more complicated reducible configurations.

We were unable to find a 7-minimal planar subcubic graph of girth larger than 3. So we pose the following conjecture.

CONJECTURE 2. *The square of every triangle-free planar subcubic graph is 6-colorable.*

The graph in Figure 1 shows that the conjecture is tight.

In fact we propose the following more general conjecture. A coloring c of a plane graph is an ℓ -facial coloring if any two distinct vertices on a facial walk of length ℓ have distinct colors. Notice that 1-facial coloring is the usual coloring. Also notice that for subcubic graphs, 2-facial coloring corresponds to coloring of the square of the graph. This concept is introduced in [10], which poses the following conjecture.

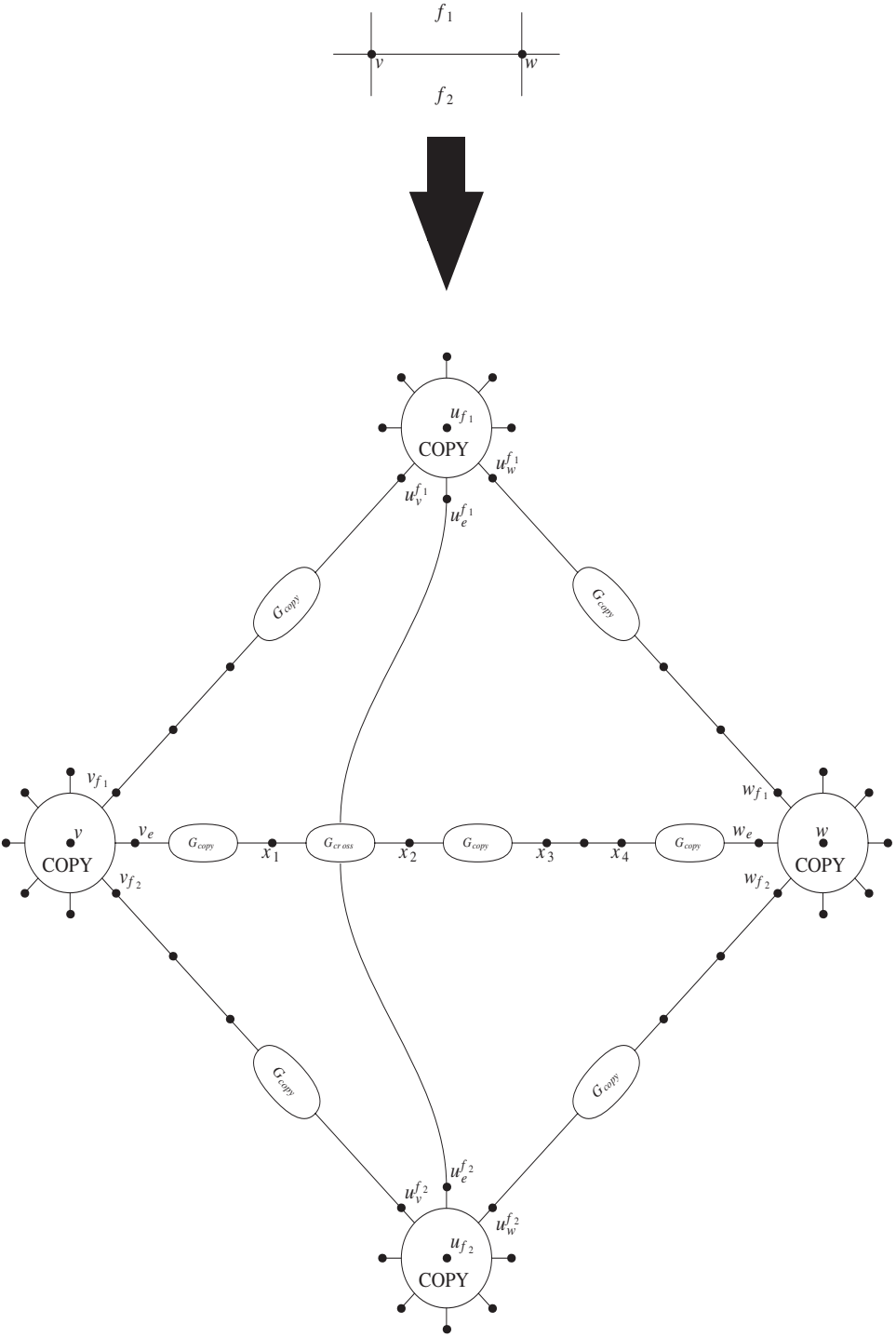
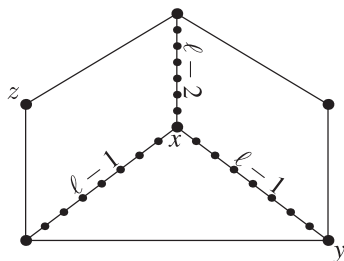


FIG. 7. The NP-completeness reduction.

FIG. 8. A graph that is not ℓ -facial colorable with $3\ell - 1$ colors.

$(3\ell + 1)$ -CONJECTURE. Let G be a plane graph and $\ell \geq 1$. Then, G is ℓ -facial colorable by $3\ell + 1$ colors.

For $\ell = 1$ the conjecture is equivalent to the four colors theorem. For $\ell = 2$ we obtain Wegner's conjecture for subcubic graphs. If the conjecture is true, it is tight for all values of ℓ , as shown by the graph obtained from K_4 in the following way: Let v be the vertex that does not lie on the outer face. Subdivide each of the edges incident with v by $\ell - 1$ new vertices of degree 2. In a similar fashion, we propose the following 3ℓ -conjecture.

3ℓ -CONJECTURE. Let G be a plane triangle-free graph and $\ell \geq 1$. Then G is ℓ -facial colorable by 3ℓ colors.

For $\ell = 1$ this statement is equivalent to Grötzsch's theorem. For $\ell = 2$, it implies Conjecture 2. Note that the bound in this conjecture is tight, as shown by graphs depicted in Figure 8. All the faces of the graph form cliques for the purposes of ℓ -facial coloring, and at most one of the vertices in thread between x and y may have the same color as vertex z .

If 3ℓ -Conjecture and Conjecture 2 are not false, then they are probably hard. The following nice consequence of these two conjectures might be easier for consideration.

CONJECTURE 3. The square of every bipartite planar subcubic graph is 6-colorable.

Acknowledgments. We would like to thank Jiří Fiala, Pascal Ochem, and Christian Sloper for fruitful discussions about the problem. We are also indebted to the anonymous referee, whose comments helped us to simplify some of our proofs and to fix several errors.

REFERENCES

- [1] O. V. BORODIN, *Criterion of chromaticity of a degree prescription*, in Abstracts of the Fourth All-Union Conference on Theoretical Cybernetics, Novosibirsk, 1997, pp. 127–128.
- [2] O. V. BORODIN, A. N. GLEBOV, A. O. IVANOVA, T. K. NEUSTROEVA, AND V. A. TASHKINOV, *A sufficient condition for a planar graph to be 2-distant $(\Delta + 1)$ -colorable*, Sib. Elektron. Math. Izv., 1 (2004), pp. 129–141 (in Russian).
- [3] T. CALAMONERI, *The $L(h, k)$ -Labelling Problem: a Survey*, Technical report 04/2004, Department of Computer Sciences, University of Rome “La Sapienza,” Rome, Italy, 2004.
- [4] D. P. DAILEY, *Uniqueness of colorability and colorability of planar 4-regular graphs are NP-complete*, Discrete Math., 30 (1980), pp. 289–293.
- [5] Z. DVOŘÁK, D. KRÁL', P. NEJEDLÝ, AND R. ŠKREKOVSKI, *Coloring squares of planar graphs with no short cycles*, submitted.

- [6] P. ERDŐS, A. L. RUBIN, AND H. TAYLOR, *Choosability in graphs*, in Proceedings of the West Coast Conference on Combinatorics, Graph Theory, and Computing (Humboldt State University, Arcata, CA, 1979), Congress. Numer. XXVI, Utilitas Math., Winnipeg, 1980, pp. 125–157.
- [7] P. HEGGERNES AND J. A. TELLE, *Partitioning graphs into generalized dominating sets*, Nordic J. Comput., 5 (1998), pp. 128–143.
- [8] J. VAN DEN HEUVEL AND S. MCGUINNESS, *Colouring of the square of a planar graph*, J. Graph Theory, 42 (2003), pp. 110–124.
- [9] T. R. JENSEN AND B. TOFT, *Graph Coloring Problems*, John Wiley and Sons, New York, 1995.
- [10] D. KRÁL', T. MADARAS, AND R. ŠKREKOVSKI, *Cyclic, diagonal, and facial colorings*, European J. Combin., 26 (2005), pp. 473–490.
- [11] K.-W. LIH AND W.-F. WANG, *Coloring the Square of an Outerplanar Graph*, Technical report, Academia Sinica, Taiwan, 2002.
- [12] M. MOLLOY AND M. R. SALAVATIPOUR, *A bound on the chromatic number of the square of a planar graph*, J. Combin. Theory Ser. B, 94 (2005), pp. 89–213.
- [13] M. MOLLOY AND M. R. SALAVATIPOUR, *Frequency channel assignment on planar networks*, in Algorithms—ESA 2002, R. H. Möhring and R. Raman, eds., Lecture Notes in Comput. Sci. 2461, Springer, Berlin, 2002, pp. 736–747.
- [14] M. MONTASSIER AND A. RASPAUD, *A Note on 2-Facial Coloring of Plane Graphs*, Technical report RR-1341-05, LaBRI, Bordeaux, France, 2005.
- [15] S. RAMANATHAN, *Scheduling Algorithms for Multi-Hop Radio Networks*, Ph.D. thesis, Department of Computer Science, University of Delaware, Newark, 1993.
- [16] G. WEGNER, *Graphs with Given Diameter and a Coloring Problem*, Technical report, University of Dortmund, Dortmund, Germany, 1977.
- [17] X. ZHOU, Y. KANARI, AND T. NISHIZEKI, *Generalized vertex-coloring of partial k -trees*, IEICE Trans. Fund. Electron. Commun. Comput. Sci., E83-A (2000), pp. 671–678.

PHASE CHANGES IN SUBTREE VARIETIES IN RANDOM RECURSIVE AND BINARY SEARCH TREES*

QUNQIANG FENG[†], HOSAM M. MAHMOUD[‡], AND ALOIS PANHOLZER[§]

Abstract. We study the variety of subtrees lying on the fringe of recursive trees and binary search trees by analyzing the distributional behavior of $X_{n,k}$, which counts the number of subtrees of size k in a random tree of size n , with $k = k(n)$ dependent on n . Using analytic methods we can characterize for both tree families the phase change behavior of $X_{n,k}$ as follows. In the subcritical case, when $k(n)/\sqrt{n} \rightarrow 0$, we show that $X_{n,k}$ is (after normalization) asymptotically normally distributed, whereas in the supercritical case, when $k(n)/\sqrt{n} \rightarrow \infty$, $X_{n,k}$ converges to 0. In the critical case, when $k(n) = \Theta(\sqrt{n})$, we show that if $k/\sqrt{n}$ approaches a limit, then $X_{n,k}$ converges in distribution to a Poisson random variable, whereas if $k/\sqrt{n}$ does not approach a finite nonzero limit, the size oscillates and does not converge in distribution to any random variable. In regard to recursive trees and binary search trees, this provides an understanding of the complete spectrum of phases of $X_{n,k}$ and the gradual change from the subcritical to the supercritical phase.

Key words. random trees, recurrence, moments, Riccati equation

AMS subject classifications. Primary, 05C05, 60C05; Secondary, 60F05

DOI. 10.1137/060653950

1. Introduction. The occurrence of patterns in random objects is an important area of modern research. The prime example is the interest one may have in the number of occurrences of a word in a given text, or the occurrence of words of a certain length in that text. Applications abound in linguistics, where one wishes to analyze grammatical frequencies, and in genetics, where one tries to identify genes in strands of DNA. The equivalent and equally important view in random trees is to find patterns (which are trees of a certain size or a certain shape) in a given tree generated randomly. This area has already attracted attention in the recent literature on random trees (see Flajolet, Gourdon, and Martínez (1997), Chyzak et al. (2007), and Feng, Mahmoud, and Su (2007)). In this paper we investigate this area for two of the most important random tree models. We look at the number of subtrees of a certain size on the fringe of random recursive trees and random binary search trees.

The random recursive tree is a naturally growing structure that underlies many stochastic phenomena, such as contagion, and algorithms such as the union-find algorithm. For numerous applications of recursive trees we refer the reader to the survey in Smythe and Mahmoud (1994). The binary search tree is another naturally growing structure that underlies many algorithms, such as combinatorial sorting and searching algorithms, and serves well as a data structure that supports fast retrieval of data.

*Received by the editors March 9, 2006; accepted for publication (in revised form) August 25, 2007; published electronically February 15, 2008.

<http://www.siam.org/journals/sidma/22-1/65395.html>

[†]Department of Statistics and Finance, University of Science and Technology of China, Hefei 230026, China (fengqq@mail.ustc.edu.cn). The work of this author is supported by the National Natural Science Foundation of China (grant 10371117) and the Special Foundation of USTC.

[‡]Department of Statistics, The George Washington University, Washington, DC 20052 (hosam@gwu.edu).

[§]Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien, Wiedner Hauptstr. 8-10/104, 1040 Wien, Austria (Alois.Panholzer@tuwien.ac.at). The work of this author is supported by the Austrian Science Foundation FWF (grant S9608-N13).

For numerous applications of binary search trees we refer the reader to Knuth (1998) and Mahmoud (2000).

2. Scope. We consider the variety of sizes that appear on the fringe of trees. For both classes of random recursive trees and binary search trees we analyze the random variable $X_{n,k}$, which counts the number of subtrees of size k in a random tree of size n , with both k fixed or $k = k(n)$ dependent on n . For k fixed it was shown in Feng, Mahmoud, and Su (2007) for recursive trees by using the contraction method and modeling via Pólya urns, and in Flajolet, Gourdon, and Martínez (1997) for binary search trees by using analytic techniques, that $X_{n,k}$ is (after normalization) asymptotically normally distributed. Furthermore, in Feng, Mahmoud, and Su (2007) the three “phases” of $X_{n,k}$ could be identified for recursive trees by computing explicit formulas for the first two moments of $X_{n,k}$.

To show the characterization of the distributional behavior of $X_{n,k}$ for $k = k(n)$ dependent on n obtained in the present paper we utilize analytic methods based on functional equations for generating functions. Such methods proved to be very effective in combinatorial analysis. It is well accepted folklore that when analytic methods are applicable, they provide fuller asymptotic expansion (hence better approximation) than purely probabilistic methods; compare, for example, the very detailed asymptotic expansion in Panholzer and Prodinger (1998) with the probabilistic approach in Lent and Mahmoud (1996).

The result we establish for recursive trees is the following.

THEOREM 2.1. *Let $X_{n,k}$ be the number of subtrees of size k on the fringe of a random recursive tree of size n .*

(a) *In the subcritical case, when $k/\sqrt{n} \rightarrow 0$,*

$$\frac{X_{n,k} - \frac{n}{k(k+1)}}{\sqrt{\frac{(2k^2-1)n}{k(k+1)^2(2k+1)}}} \xrightarrow{D} \mathcal{N}(0, 1).$$

(b) *We are in the critical case when $k = \Theta(\sqrt{n})$. In the critical case, when $k/\sqrt{n} \rightarrow c > 0$,*

$$X_{n,k} \xrightarrow{D} \text{Poi}\left(\frac{1}{c^2}\right),$$

and if $k/\sqrt{n}$ does not converge to a limit, no limiting distribution exists for $X_{n,k}$.

(c) *In the supercritical case, when $k/\sqrt{n} \rightarrow \infty$,*

$$X_{n,k} \xrightarrow{D} 0.$$

The analytic method employed for recursive trees extends naturally to binary search trees, which we include in this paper, and possibly to other tree classes, too, which we leave for future research. The result we establish for binary search trees is the following.

THEOREM 2.2. *Let $X_{n,k}$ be the number of subtrees of size k on the fringe of a random binary search tree of size n .*

(a) *In the subcritical case, when $k/\sqrt{n} \rightarrow 0$,*

$$\frac{X_{n,k} - \frac{2n}{(k+1)(k+2)}}{\sqrt{\frac{2k(4k^2+5k-3)n}{(k+1)(k+2)^2(2k+1)(2k+3)}}} \xrightarrow{D} \mathcal{N}(0, 1).$$

- (b) *We are in the critical case when $k = \Theta(\sqrt{n})$. In the critical case, when $k/\sqrt{n} \rightarrow c > 0$,*

$$X_{n,k} \xrightarrow{D} \text{Poi}\left(\frac{2}{c^2}\right),$$

and if $k/\sqrt{n}$ does not converge to a limit, no limiting distribution exists for $X_{n,k}$.

- (c) *In the supercritical case, when $k/\sqrt{n} \rightarrow \infty$,*

$$X_{n,k} \xrightarrow{D} 0.$$

We remark that $X_{n,k}$ can also be interpreted as the number of nodes in a random tree of size n , which have exactly k *descendants* (if we assume that a node itself is counted as its own descendant). From this point of view the “counterpart” of $X_{n,k}$ is the level polynomial $L_{n,k}$, i.e., the random variable that counts the number of nodes in a random tree of size n , which have exactly k *ascendants*. The distributional behavior $L_{n,k}$, i.e., the number of nodes at level k in a random tree of size n , has been analyzed recently for recursive trees and binary search trees by Fuchs, Hwang, and Neininger (2006).

The rest of the paper is organized as follows. In section 3 we first give the definition of random recursive trees and derive a partial differential equation for the number of subtrees of a given size (possibly dependent on n). This involves setting up a basic stochastic recurrence based on a decomposition property of random recursive trees. In subsection 3.1 we show how to compute the exact moments. The critical and supercritical cases fall out directly from the raw moments, as shown in subsections 3.2 and 3.3. However, the subcritical case needs centering (and scaling) to get a Gaussian limit. Therefore, to reach the limit distribution via moments, we need to have them centered. This is done in subsection 3.4, where several technical subsections are dedicated to deriving estimates for the centered moments (subsections 3.4.1–3.4.4).

After showing the mechanics of the method applied to recursive trees, we present a brief sketch for binary search trees, modeled after the exposition of recursive trees, starting with the definition of random binary search trees and its partial differential equation in section 4, then moving on to the exact moments (section 4.1). Then we show how to compute the limit distributions in the critical and supercritical cases (subsections 4.2 and 4.3). The subcritical case is dealt with in subsection 4.4, and the string of technical estimates is in subsections 4.4.1–4.4.5. We conclude in section 5 with a few remarks on the choice of an analytic methodology.

2.1. Notation. In the discourse we shall use the following notation. The symbol $\llbracket A \rrbracket$ denotes Iverson’s bracket for the predicate A , i.e., $\llbracket A \rrbracket = 1$ if A is true and 0 otherwise. For any number x , and nonnegative integer m , the notation $x^{\underline{m}}$ stands for the falling factorial $x(x-1)\dots(x-m+1)$; we interpret $x^{\underline{0}}$ as 1. The notation $\left\{ \begin{smallmatrix} r \\ s \end{smallmatrix} \right\}$ stands for the r th Stirling number (of the second kind) of order s . The operator $[z^n]$ extracts the n th coefficients when applied to a function; that is, $[z^n]f(z)$ is the coefficient of z^n in the power series expansion of $f(z)$.

For probabilistic convergence we use $\xrightarrow{D}$ to denote convergence in distribution. The standard random variables $\text{Poi}(\lambda)$ (the Poisson distributed with parameter λ) and $\mathcal{N}(\mu, \sigma^2)$ (the normally distributed with mean μ and variance σ^2) appear in the results as limiting random variables. Other standard nomenclature is employed, and we assume it is familiar to the reader and needs no special mention here.

3. The variety of subtree sizes in a recursive tree. The *random recursive tree* is an outgrowth from a single node labeled 1. Progressively, nodes are added in stages: At the n th stage a node in the existing tree is chosen at random as a parent for the n th entrant (labeled n). In this context *random* means that all nodes in the tree of size $n - 1$ are equally likely parents. According to this construction algorithm, the nodes along any root-to-leaf path carry increasing labels, and hence these trees are also members of the class of *increasing trees*.

The model of randomness in the growth of random recursive trees induces a uniform distribution on the trees: All $(n - 1)!$ recursive trees of size n are generated with equal probability. Many important properties of recursive trees have been analyzed from the convenient viewpoint of recurrence occurring naturally in the stochastic growth. Nonetheless, the uniform distribution of the trees also gave rise to analyses based on generating functions and analytic methods such as integral transforms (see Bergeron, Flajolet, and Salvy (1992)).

Let $X_{n,k}$ be the number of nodes in a random recursive tree of size n that are the roots of subtrees of size k . (A subtree rooted at a node is the entire structure of descendant nodes from the given one.) Equivalently, $X_{n,k}$ is the number of subtrees on the fringe of the recursive tree with size k . Our count of the size of a subtree includes its root. Thus, for example, a leaf roots a subtree of size 1. Applications of $X_{n,k}$ include finding out the number of participants in a chain letter scheme who will attain a certain profit (see Gastwirth and Bhattacharya (1984)), or the number of copies of a particular ancient text in a philological study (see Najock and Heyde (1982)).

We denote by $T_n = (n - 1)!$ the number of different recursive trees of size n . Let $M_k(z, v)$ be the moment generating function

$$M_k(z, v) = \sum_{n \geq 1} \sum_{m \geq 0} P\{X_{n,k} = m\} T_n \frac{z^n}{n!} v^m = \sum_{n \geq 1} \sum_{m \geq 0} P\{X_{n,k} = m\} \frac{z^n}{n} v^m.$$

According to the recursive nature of these trees, we can read right off the definition the probabilities that the subtrees have certain sizes. For all $k \geq 1$, the probabilities $P\{X_{n,k} = m\}$ satisfy (for $n > k$)

$$\begin{aligned} P\{X_{n,k} = m\} &= \sum_{r \geq 1} \frac{1}{r!} \sum_{\substack{n_1 + \dots + n_r = n-1 \\ n_i \geq 1, 1 \leq i \leq r}} \binom{n-1}{n_1, \dots, n_r} \frac{T_{n_1} \dots T_{n_r}}{T_n} \\ &\quad \times \sum_{\substack{m_1 + \dots + m_r = m \\ m_i \geq 0, 1 \leq i \leq r}} P\{X_{n_1,k} = m_1\} \dots P\{X_{n_r,k} = m_r\} \end{aligned}$$

with initial values $P\{X_{k,k} = 1\} = 1$ and $P\{X_{n,k} = 0\} = 1$ for $1 \leq n < k$.

Multiply this recurrence by $T_n \frac{z^{n-1}}{(n-1)!} v^m$ and sum up over $n > k$ and $m \geq 0$ to obtain the functional equation, for $k \geq 1$,

$$(3.1) \quad \frac{\partial}{\partial z} M_k(z, v) = e^{M_k(z, v)} + (v - 1)z^{k-1},$$

with initial condition $M_k(0, v) = 0$.

It can be easily checked that the function

$$(3.2) \quad M_k(z, v) = \frac{(v - 1)z^k}{k} + \log \left(\frac{1}{1 - \int_0^z e^{\frac{(v-1)t^k}{k}} dt} \right)$$

is a solution to the partial differential equation (3.1) that also satisfies the initial conditions. The solution (3.2) can be found by using the substitution $Q(z, v) := \exp(M_k(z, v))$ in (3.1), which leads to a solvable Riccati differential equation.

From the solution (3.2) we can obtain the r th factorial moments, which directly furnish a limiting distribution for the critical and supercritical cases.

3.1. Exact moments. To get the r th factorial moments we use the substitution $w := v - 1$ in $M_k(z, v)$, and extract coefficients:

$$\mathbf{E}(X_{n,k}^r) = \mathbf{E}(X_{n,k}(X_{n,k} - 1) \cdots (X_{n,k} - r + 1)) = nr! [z^n w^r] M_k(z, 1 + w).$$

In order to expand $M_k(z, 1 + w)$ around $w = 0$ we consider

$$\begin{aligned} \log \left(\frac{1}{1 - \int_0^z e^{\frac{wt^k}{k}} dt} \right) &= \log \left(\frac{1}{1 - \int_0^z \sum_{j \geq 0} \frac{w^j t^{kj}}{j! k^j} dt} \right) \\ &= \log \frac{1}{1 - z} + \log \left(\frac{1}{1 - \frac{z}{1-z} \sum_{j \geq 1} \frac{w^j z^{kj}}{j! k^j (kj+1)}} \right) \end{aligned}$$

and obtain from (3.2)

$$(3.3) \quad M_k(z, 1 + w) = \log \frac{1}{1 - z} + \frac{wz^k}{k} + \log \left(\frac{1}{1 - \frac{z}{1-z} \sum_{j \geq 1} \frac{w^j z^{kj}}{j! k^j (kj+1)}} \right).$$

Next we extract coefficients and get for $r \geq 1$,

$$\begin{aligned} [w^r] \log \left(\frac{1}{1 - \frac{z}{1-z} \sum_{j \geq 1} \frac{w^j z^{kj}}{j! k^j (kj+1)}} \right) &= [w^r] \sum_{\ell \geq 1} \frac{\left(\frac{z}{1-z}\right)^\ell}{\ell} \left(\sum_{j \geq 1} \frac{w^j z^{kj}}{j! k^j (kj+1)} \right)^\ell \\ &= \sum_{\ell \geq 1} \frac{\left(\frac{z}{1-z}\right)^\ell}{\ell} \sum_{\substack{j_1 + \dots + j_\ell = r \\ j_q \geq 1, 1 \leq q \leq \ell}} \frac{z^{kr}}{k^r \prod_{i=1}^\ell j_i! \prod_{i=1}^\ell (j_i k + 1)} \\ &= \sum_{\ell=1}^r \frac{\left(\frac{z}{1-z}\right)^\ell}{\ell} \times \frac{z^{kr}}{k^r} \sum_{\substack{j_1 + \dots + j_\ell = r \\ j_q \geq 1, 1 \leq q \leq \ell}} \frac{1}{\prod_{i=1}^\ell j_i! \prod_{i=1}^\ell (j_i k + 1)} \\ &= \frac{1}{r!} \sum_{\ell=1}^r \frac{\left(\frac{z}{1-z}\right)^\ell}{\ell} \times \frac{z^{kr}}{k^r} \sum_{\substack{j_1 + \dots + j_\ell = r \\ j_q \geq 1, 1 \leq q \leq \ell}} \binom{r}{j_1, \dots, j_\ell} \frac{1}{\prod_{i=1}^\ell (j_i k + 1)}. \end{aligned}$$

This immediately leads from (3.3) to the following formula for the r th coefficients ($r \geq 1$):

$$(3.4) \quad \begin{aligned} [w^r] M_k(z, v) &= \frac{z^k}{k} [r = 1] \\ &+ \frac{1}{r!} \sum_{\ell=1}^r \frac{\left(\frac{z}{1-z}\right)^\ell}{\ell} \times \frac{z^{kr}}{k^r} \sum_{\substack{j_1 + \dots + j_\ell = r \\ j_q \geq 1, 1 \leq q \leq \ell}} \binom{r}{j_1, \dots, j_\ell} \frac{1}{\prod_{i=1}^\ell (j_i k + 1)}. \end{aligned}$$

The formula (3.4) gives

$$\begin{aligned}\mathbf{E}(X_{n,k}) &= n[z^n w]M_k(z, 1+w) \\ &= n[z^n] \left(\frac{z^k}{k} + \frac{z^{k+1}}{k(k+1)(1-z)} \right) \\ &= n \left(\frac{1}{k} \llbracket n = k \rrbracket + \frac{1}{k(k+1)} \llbracket n \geq k+1 \rrbracket \right),\end{aligned}$$

and thus the following result ensues for the expectation of $X_{n,k}$, as given in Feng, Mahmoud, and Su (2007):

$$(3.5) \quad \mathbf{E}(X_{n,k}) = \begin{cases} \frac{n}{k(k+1)} & \text{for } n \geq k+1, \\ 1 & \text{for } n = k, \\ 0 & \text{for } 1 \leq n < k. \end{cases}$$

Owing to the explicit form of (3.4), we also get the following closed form solution for the r th factorial moments with $r \geq 2$:

$$\begin{aligned}(3.6) \quad \mathbf{E}(X_{n,k}^r) &= nr! [z^n w^r] M_k(z, v) \\ &= \frac{\llbracket n \geq kr+1 \rrbracket n}{k^r} \sum_{\ell=1}^r \frac{\binom{n-kr-1}{\ell-1}}{\ell} \\ &\quad \times \sum_{\substack{j_1+\dots+j_\ell=r \\ j_q \geq 1, 1 \leq q \leq \ell}} \binom{r}{j_1, \dots, j_\ell} \frac{1}{\prod_{i=1}^\ell (j_i k + 1)}.\end{aligned}$$

As an example, we compute the second factorial moment:

$$\mathbf{E}(X_{n,k}(X_{n,k} - 1)) = \frac{n(n-2k-1)}{k^2(k+1)^2} - \frac{n}{k^2(2k+1)} \quad \text{for } n \geq 2k+1,$$

which is 0 otherwise.

In view of the classical relation

$$\mathbf{E}(X_{n,k}^r) = \sum_{\ell=1}^r \left\{ \begin{matrix} r \\ \ell \end{matrix} \right\} \mathbf{E}(X_{n,k}^\ell),$$

one could also get (at least in principle) closed formulas for every ordinary r th moment.

3.2. The critical case. We are in the critical case when $k = \Theta(\sqrt{n})$. There are two flavors within this case: One flavor is when $k/\sqrt{n}$ converges to a limit, such as, for example, the case $k = 3\lfloor\sqrt{n} + \log n\rfloor$; the other is when k is asymptotic to $g(n)\sqrt{n}$ with $g(n)$ being a function of bounded variation, but oscillating and not converging to any limit, such as, for example, the case $k = \lfloor(2 + \sin n)\sqrt{n} + 6\rfloor$.

First, we consider the critical case having the flavor $\frac{n}{k^2} \rightarrow \lambda$ for some $\lambda > 0$. It is evident from (3.5) that we have $\mathbf{E}(X_{n,k}) \rightarrow \lambda$. Next we consider a fixed $r \geq 2$. We split the sum appearing in (3.6) as follows:

$$\begin{aligned} \mathbf{E}(X_{n,k}^r) &= \underbrace{\frac{\llbracket n \geq kr+1 \rrbracket n}{k^r} \times \frac{\binom{n-kr-1}{r-1}}{r} \sum_{\substack{j_1+\dots+j_r=r \\ j_q \geq 1, 1 \leq q \leq r}} \frac{\binom{r}{j_1, \dots, j_r}}{\prod_{i=1}^r (j_i k + 1)}}_{=:A} \\ &+ \underbrace{\frac{\llbracket n \geq kr+1 \rrbracket n}{k^r} \sum_{\ell=1}^{r-1} \frac{\binom{n-kr-1}{\ell-1}}{\ell} \sum_{\substack{j_1+\dots+j_\ell=r \\ j_q \geq 1, 1 \leq q \leq \ell}} \frac{\binom{r}{j_1, \dots, j_\ell}}{\prod_{i=1}^\ell (j_i k + 1)}}_{=:B}. \end{aligned}$$

For large n we can drop Iverson's brackets in both A and B . Using

$$\sum_{r \geq 0} \sum_{\substack{j_1+\dots+j_\ell=r \\ j_q \geq 1, 1 \leq q \leq \ell}} \binom{r}{j_1, \dots, j_\ell} \frac{z^r}{r!} = (e^z - 1)^\ell,$$

we obtain the inequality

$$\sum_{\substack{j_1+\dots+j_\ell=r \\ j_q \geq 1, 1 \leq q \leq \ell}} \binom{r}{j_1, \dots, j_\ell} = r! [z^r] (e^z - 1)^\ell \leq r! [z^r] e^{\ell z} = \ell^r,$$

and further by trivial estimates:

$$\begin{aligned} B &\leq \sum_{\ell=1}^{r-1} \frac{n^\ell}{\ell! k^r k^\ell} \ell^r \\ &\leq \frac{r^{r-1}}{k} \sum_{\ell=1}^{r-1} \frac{1}{k^{r-1-\ell}} \left(\frac{n}{k^2}\right)^\ell \frac{1}{(\ell-1)!} \\ &\leq \frac{r^{r-1}}{k} \times \frac{n}{k^2} \sum_{\ell \geq 0} \frac{\left(\frac{n}{k^2}\right)^\ell}{\ell!} \\ &= \frac{r^{r-1}}{k^3} n e^{\frac{n}{k^2}}. \end{aligned}$$

Since $\frac{n}{k^2} \rightarrow \lambda$, we further have

$$B = O\left(\frac{1}{k}\right) = O\left(\frac{1}{\sqrt{n}}\right).$$

The expression A simplifies considerably, since the only possible composition of r in r integers $j_q \geq 1$ is obtained for $j_1 = \dots = j_r = 1$; thus for large n we have

$$\begin{aligned} A &= \frac{n}{k^r} \times \frac{\binom{n-kr-1}{r-1}}{r} \times \frac{r!}{(k+1)^r} \\ &= \frac{n^r}{k^{2r}} \left(1 + O\left(\frac{k}{n}\right) + O\left(\frac{1}{k}\right)\right) \\ &= \left(\frac{n}{k^2}\right)^r \left(1 + O\left(\frac{1}{\sqrt{n}}\right)\right). \end{aligned}$$

Therefore, if $\frac{n}{k^2} \rightarrow \lambda$, we also get for every $r \geq 1$

$$\mathbf{E}(X_{n,k}^r) = \left(\frac{n}{k^2}\right)^r + O\left(\frac{1}{\sqrt{n}}\right) \rightarrow \lambda^r.$$

Since λ^r are the moments of $\text{Poi}(\lambda)$, and convergence of all moments to a random variable uniquely characterized by its moments implies weak convergence, we get

$$X_{n,k} \xrightarrow{D} \text{Poi}(\lambda).$$

This proves the convergence part of Theorem 2.1(b), where we used the substitution $c := \frac{1}{\sqrt{\lambda}}$, and thus $\frac{k}{\sqrt{n}} \rightarrow c$. When k is of the order $\sqrt{n}$, but fluctuations persist, as in the case involving sinusoids mentioned at the beginning of this subsection, all the moments oscillate, and no limit distribution can exist.

3.3. The supercritical case. We consider the supercritical case in passing, as it does not require additional effort. Assume that $k := k_n$ grows with n such that $\frac{n}{k^2} = o(1)$.

Crude estimates for (3.6) are sufficient for our purpose. We get for $r \geq 2$

$$\begin{aligned} \mathbf{E}(X_{n,k}^r) &\leq \frac{n}{k^r} \sum_{\ell=1}^r \frac{n^{\ell-1}}{\ell!} \sum_{\substack{j_1+\dots+j_\ell=r \\ j_q \geq 1, 1 \leq q \leq \ell}} \binom{r}{j_1, \dots, j_\ell} \frac{1}{k^\ell} = \frac{1}{k^r} \sum_{\ell=1}^r \frac{n^\ell}{\ell! k^\ell} \ell^r \\ &\leq r^{r-1} \sum_{\ell=1}^r \frac{1}{k^{r-\ell}} \left(\frac{n}{k^2}\right)^\ell \frac{1}{(\ell-1)!} \leq r^{r-1} \frac{n}{k^2} \sum_{\ell \geq 0} \frac{\left(\frac{n}{k^2}\right)^\ell}{\ell!} \\ &= r^{r-1} \frac{n}{k^2} e^{\frac{n}{k^2}}. \end{aligned}$$

Since $\frac{n}{k^2} \rightarrow 0$, we have that for all $r \geq 1$, $\mathbf{E}(X_{n,k}^r) \rightarrow 0$. Again we have convergence of all moments, and subsequently convergence of $X_{n,k}$ to a degenerate distribution with mass at 0. Part (c) of Theorem 2.1 is proved. Because the limit is constant, this convergence is in probability, too.

3.4. The subcritical case. For the critical and supercritical cases, it was sufficient to work with raw moments and establish their convergence to the moments of a known simple distribution. For the subcritical case, we need centering and scaling, too. We appeal to the method of recursive moments (see Chern, Hwang, and Tsai (2002)), although here we are not forced to “pump out” the behavior of the moments in an inductive way, but we are able to obtain exact (but rather involved) expressions for the moments directly by extracting coefficients of the moment generating function of the centered random variable. Consider the centered random variable $\tilde{X}_{n,k} := X_{n,k} - \mathbf{E}(X_{n,k})$ and introduce the generating function

$$\tilde{M}_k(z, s) := \sum_{n \geq 1} \mathbf{E}(e^{\tilde{X}_{n,k}s}) \frac{z^n}{n} = \sum_{n \geq 1} e^{-\mathbf{E}(X_{n,k})s} \mathbf{E}(e^{X_{n,k}s}) \frac{z^n}{n}.$$

Using the explicit formula for $\mathbf{E}(X_{n,k})$ as given by (3.5), we obtain by routine manipulations

$$\tilde{M}_k(z, s) = M_k\left(e^{-\frac{s}{k(k+1)}} z, e^s\right) + \left(1 - e^{\frac{ks}{k+1}}\right) \frac{z^k}{k} + \sum_{1 \leq n < k} \frac{z^n}{n} - \sum_{1 \leq n < k} \frac{\left(e^{-\frac{s}{k(k+1)}} z\right)^n}{n}.$$

Plugging in the formula for $M_k(z, v)$ given by (3.2), we further obtain after

simplifications

$$(3.7) \quad \begin{aligned} \tilde{M}_k(z, s) = \log & \left(\frac{1}{1 - \int_0^{e^{-\frac{s}{k(k+1)}} z} e^{\frac{(e^s-1)t^k}{k}} dt} \right) \\ & + \sum_{1 \leq n \leq k} \frac{z^n}{n} - \sum_{1 \leq n \leq k} \frac{\left(e^{-\frac{s}{k(k+1)}} z \right)^n}{n}. \end{aligned}$$

In the following chain of technical subsections, we analyze the centered moments, identify the main contribution, and put upper bounds on the negligible terms.

3.4.1. Expanding around $s = 0$. We first expand $\int_0^{e^{-\frac{s}{k(k+1)}} z} e^{\frac{(e^s-1)t^k}{k}} dt$ around $s = 0$ and $z = 1$:

$$\begin{aligned} & \int_0^{e^{-\frac{s}{k(k+1)}} z} e^{\frac{(e^s-1)t^k}{k}} dt \\ &= \int_0^{e^{-\frac{s}{k(k+1)}} z} \sum_{j \geq 0} \frac{(e^s-1)^j t^{kj}}{j! k^j} dt = \sum_{j \geq 0} \frac{(e^s-1)^j e^{-\frac{s(kj+1)}{k(k+1)}} z^{kj+1}}{j! (kj+1) k^j} \\ &= e^{-\frac{s}{k(k+1)}} z + \sum_{j \geq 1} \frac{(e^s-1)^j e^{-\frac{s(kj+1)}{k(k+1)}} z^{kj+1}}{j! (kj+1) k^j} \\ &= z + \sum_{j \geq 1} \left[\frac{(e^s-1)^j e^{-\frac{s(kj+1)}{k(k+1)}} z^{kj+1}}{j! (kj+1) k^j} + \frac{(-1)^j s^j z}{j! k^j (k+1)^j} \right] \\ &= z + \sum_{j \geq 1} \left[\frac{\sum_{m \geq j} \{ \begin{smallmatrix} m \\ j \end{smallmatrix} \} \frac{j! s^m}{m!} \sum_{m \geq 0} \frac{(-1)^m (kj+1)^m s^m}{k^m (k+1)^m m!} z^{kj+1} + \frac{(-1)^j s^j z}{j! k^j (k+1)^j} \right] \\ &= z + \sum_{\ell \geq 1} s^\ell \left(\sum_{j=1}^{\ell} \sum_{m=0}^{\ell-j} \frac{\{ \begin{smallmatrix} \ell-m \\ j \end{smallmatrix} \} (-1)^m (kj+1)^m z^{kj+1}}{m! (\ell-m)! (kj+1) k^{j+m} (k+1)^m} + \frac{(-1)^\ell z}{\ell! k^\ell (k+1)^\ell} \right) \\ &= z + \sum_{\ell \geq 1} s^\ell \left(\sum_{j=1}^{\ell} \sum_{m=0}^{\ell-j} \frac{\{ \begin{smallmatrix} \ell-m \\ j \end{smallmatrix} \} (-1)^m (kj+1)^m}{m! (\ell-m)! (kj+1) k^{j+m} (k+1)^m} \right. \\ & \quad \times \sum_{i=0}^{kj+1} (-1)^i \binom{kj+1}{i} (1-z)^i + \frac{(-1)^\ell}{\ell! k^\ell (k+1)^\ell} + \frac{(-1)^{\ell-1}}{\ell! k^\ell (k+1)^\ell} (1-z) \Big) \\ &= z + \sum_{\ell \geq 1} s^\ell \left(\sum_{i=0}^{\ell k+1} (-1)^i (1-z)^i \sum_{j=1}^{\ell} \sum_{m=0}^{\ell-j} \frac{\{ \begin{smallmatrix} \ell-m \\ j \end{smallmatrix} \} (-1)^m (kj+1)^m \binom{kj+1}{i}}{m! (\ell-m)! (kj+1) k^{j+m} (k+1)^m} \right. \\ & \quad \left. + \frac{(-1)^\ell}{\ell! k^\ell (k+1)^\ell} + \frac{(-1)^{\ell-1}}{\ell! k^\ell (k+1)^\ell} (1-z) \right). \end{aligned}$$

So, we obtain the following simple structure for the integral considered:

$$\int_0^{e^{-\frac{s}{k(k+1)}} z} e^{\frac{(e^s-1)t^k}{k}} dt = z + \sum_{\ell \geq 1} s^\ell \sum_{i=0}^{\ell k+1} c_{\ell,i}(k) (1-z)^i,$$

where the functions $c_{\ell,i}(k)$ are given by

$$(3.8) \quad c_{\ell,i}(k) = (-1)^i \sum_{j=1}^{\ell} \sum_{m=0}^{\ell-j} \frac{\{ \begin{smallmatrix} \ell-m \\ j \end{smallmatrix} \} (-1)^m (kj+1)^m \binom{kj+1}{i}}{m! (\ell-m)! (kj+1) k^{j+m} (k+1)^m} \\ + \frac{(-1)^{\ell} \llbracket i=0 \rrbracket}{\ell! k^{\ell} (k+1)^{\ell}} + \frac{(-1)^{\ell-1} \llbracket i=1 \rrbracket}{\ell! k^{\ell} (k+1)^{\ell}}.$$

Next we consider

$$\log \left(\frac{1}{1 - \int_0^e e^{-\frac{s}{k(k+1)}} z e^{\frac{(e^s-1)t^k}{k}} dt} \right) \\ = \log \frac{1}{1-z} + \log \left(\frac{1}{1 - \frac{1}{1-z} \sum_{\ell \geq 1} s^{\ell} \sum_{i=0}^{\ell k+1} c_{\ell,i}(k) (1-z)^i} \right),$$

and extract coefficients. This gives for $r \geq 1$

$$[s^r] \log \left(\frac{1}{1 - \int_0^e e^{-\frac{s}{k(k+1)}} z e^{\frac{(e^s-1)t^k}{k}} dt} \right) \\ = [s^r] \log \left(\frac{1}{1 - \frac{1}{1-z} \sum_{\ell \geq 1} s^{\ell} \sum_{i=0}^{\ell k+1} c_{\ell,i}(k) (1-z)^i} \right) \\ = \sum_{m=1}^r \frac{1}{m} [s^r] \left(\frac{1}{1-z} \sum_{\ell \geq 1} s^{\ell} \sum_{i=0}^{\ell k+1} (1-z)^i c_{\ell,i}(k) \right)^m \\ = \sum_{m=1}^r \frac{1}{m(1-z)^m} \sum_{\substack{r_1+\dots+r_m=r \\ r_q \geq 1, 1 \leq q \leq m}} \prod_{j=1}^m \sum_{i=0}^{r_j k+1} (1-z)^i c_{r_j,i}(k) \\ = \sum_{m=1}^r \frac{1}{m(1-z)^m} \sum_{t=0}^{rk+m} (1-z)^t \times \sum_{\substack{r_1+\dots+r_m=r \\ r_q \geq 1, 1 \leq q \leq m}} \sum_{\substack{t_1+\dots+t_m=t \\ 0 \leq t_q \leq r_q k+1, 1 \leq q \leq m}} \prod_{j=1}^m c_{r_j,t_j}(k),$$

and thus

$$[s^r] \log \left(\frac{1}{1 - \int_0^e e^{-\frac{s}{k(k+1)}} z e^{\frac{(e^s-1)t^k}{k}} dt} \right) = \sum_{p=-r}^{rk} f_{r,p}(k) (1-z)^p,$$

with

$$f_{r,p}(k) := \sum_{m=\max\{1,-p\}}^r \frac{1}{m} \sum_{\substack{r_1+\dots+r_m=r \\ r_q \geq 1, 1 \leq q \leq m}} \sum_{\substack{t_1+\dots+t_m=p+m \\ 0 \leq t_q \leq r_q k+1, 1 \leq q \leq m}} \prod_{j=1}^m c_{r_j,t_j}(k),$$

where the functions $c_{\ell,i}(k)$ are defined by (3.8).

Since

$$[s^r] \sum_{1 \leq n \leq k} \frac{(e^{-\frac{s}{k(k+1)}} z)^n}{n} = \sum_{1 \leq n \leq k} \frac{z^n}{n} [s^r] e^{-\frac{sn}{k(k+1)}} = \sum_{1 \leq n \leq k} z^n \frac{(-1)^r n^{r-1}}{r! k^r (k+1)^r},$$

we obtain for the coefficients at s^r of $\tilde{M}_k(z, s)$, as given by (3.7) for $r \geq 1$,

$$[s^r]\tilde{M}_k(z, s) = \sum_{p=-r}^{rk} f_{r,p}(k)(1-z)^p + \sum_{1 \leq n \leq k} z^n \frac{(-1)^{r-1} n^{r-1}}{r! k^r (k+1)^r},$$

which can be written as

$$[s^r]\tilde{M}_k(z, s) = \sum_{p=1}^r \frac{1}{(1-z)^p} \tilde{f}_{r,p}(k) + \sum_{n=0}^{rk} g_{r,n}(k) z^n$$

with

$$(3.9) \quad \tilde{f}_{r,p}(k) := f_{r,-p}(k) = \sum_{m=p}^r \frac{1}{m} \sum_{\substack{r_1+\dots+r_m=r \\ r_q \geq 1, 1 \leq q \leq m}} \sum_{\substack{t_1+\dots+t_m=m-p \\ 0 \leq t_q \leq r_q k+1, 1 \leq q \leq m}} \prod_{j=1}^m c_{r_j, t_j}(k)$$

for $1 \leq p \leq r$, and

$$g_{r,n}(k) := (-1)^n \sum_{p=n}^{rk} \binom{p}{n} f_{r,p}(k) + \frac{(-1)^{r-1} n^{r-1}}{r! k^r (k+1)^r} \llbracket 1 \leq n \leq k \rrbracket.$$

This representation of $[s^r]\tilde{M}_k(z, s)$ immediately leads to the following explicit formula for the r th moments of $\tilde{X}_{n,k}$, which is valid for $n \geq rk+1$:

$$(3.10) \quad \mathbf{E}(\tilde{X}_{n,k}^r) = r! n [z^n s^r] \tilde{M}_k(z, s) = r! n \sum_{p=1}^r \tilde{f}_{r,p}(k) \binom{n+p-1}{p-1},$$

where the functions $\tilde{f}_{r,p}(k)$ are given by (3.9).

We now examine the functions $c_{\ell,i}(k)$ as given by (3.8). Using the identity

$$\sum_{m=j}^{\ell} \binom{\ell}{m} \left\{ \begin{matrix} m \\ j \end{matrix} \right\} = \left\{ \begin{matrix} \ell+1 \\ j+1 \end{matrix} \right\},$$

we get by very crude estimates

$$\begin{aligned} & \left| (-1)^i \sum_{j=1}^{\ell} \sum_{m=0}^{\ell-j} \frac{\left\{ \begin{matrix} \ell-m \\ j \end{matrix} \right\} (-1)^m (kj+1)^m \binom{kj+1}{i}}{m! (\ell-m)! (kj+1) k^{j+m} (k+1)^m} \right| \\ & \leq \sum_{j=1}^{\ell} \frac{\binom{kj+1}{i}}{(kj+1) k^j} \sum_{m=0}^{\ell-j} \frac{\left\{ \begin{matrix} \ell-m \\ j \end{matrix} \right\} (kj+1)^m}{m! (\ell-m)! k^m (k+1)^m} \\ & \leq \sum_{j=1}^{\ell} \frac{k^i (j+1)^i}{\ell! i! k^{j+1}} \sum_{m=0}^{\ell-j} \frac{\left\{ \begin{matrix} \ell-m \\ j \end{matrix} \right\} (j+1)^m \binom{\ell}{m}}{(k+1)^m} \\ & \leq \sum_{j=1}^{\ell} \frac{k^i (j+1)^i (j+1)^{\ell-j}}{\ell! i! k^{j+1}} \sum_{m=0}^{\ell-j} \binom{\ell}{m} \left\{ \begin{matrix} \ell-m \\ j \end{matrix} \right\} \\ & = \sum_{j=1}^{\ell} \frac{k^i (j+1)^i (j+1)^{\ell-j}}{\ell! i! k^{j+1}} \left\{ \begin{matrix} \ell+1 \\ j+1 \end{matrix} \right\} \end{aligned}$$

$$\begin{aligned}
(3.11) \quad & \leq \frac{k^i(\ell+1)^i(\ell+1)^{\ell-1}}{\ell! i! k^2} \sum_{j=2}^{\ell+1} \left\{ \begin{matrix} \ell+1 \\ j \end{matrix} \right\} \\
& \leq \frac{(\ell+1)^\ell B_{\ell+1}}{(\ell+1)!} \times \frac{(\ell+1)^i}{i!} k^{i-2},
\end{aligned}$$

where B_ℓ denote the Bell numbers (number of partitions of a set with ℓ elements). Since

$$\left| \frac{(-1)^\ell}{k^\ell(k+1)^\ell \ell!} \right| \leq \frac{1}{k^2} = \frac{1}{k^2} \left\{ \begin{matrix} \ell+1 \\ 1 \end{matrix} \right\},$$

we can add this summand in the preceding computations at step (3.11) and thus get the same bound. Therefore, we arrive at the following bound, which holds uniformly for all ℓ , i , and k :

$$|c_{\ell,i}(k)| \leq q_{\ell,i} k^{i-2} \quad \text{with} \quad q_{\ell,i} := \frac{(\ell+1)^\ell B_{\ell+1}}{(\ell+1)!} \times \frac{(\ell+1)^i}{i!}.$$

The two functions that exert the most influence on the asymptotic distribution $c_{\ell,i}(k)$, and which can be computed easily from (3.8), are

$$\begin{aligned}
c_{1,0}(k) &= \frac{1}{k(k+1)} - \frac{1}{k(k+1)} = 0, \\
c_{2,0}(k) &= \frac{\nu(k)}{2} \quad \text{with} \quad \nu(k) := \frac{2k^2 - 1}{k(k+1)^2(2k+1)}.
\end{aligned}$$

3.4.2. Estimates for $\tilde{f}_{r,p}(k)$. We next treat the functions $\tilde{f}_{r,p}(k)$ as given by (3.9). Since $1 \leq p \leq r$ holds in (3.10), this implies that in the compositions appearing in the definition of $\tilde{f}_{r,p}(k)$ we always have $r_q, t_q \leq r$ for all $1 \leq q \leq m$. Thus, we obtain the following estimate, which holds for all $1 \leq p \leq r$ and all r_q, t_q :

$$|c_{r_j, t_j}(k)| \leq q_{r_j, t_j} k^{t_j-2} \leq c_r k^{t_j-2},$$

with the universal constant

$$(3.12) \quad c_r := \frac{(r+1)^{2r} B_{r+1}}{(r+1)!}.$$

This gives for $1 \leq p \leq r$ the estimate

$$\begin{aligned}
& \left| \sum_{\substack{r_1+\dots+r_m=r \\ r_q \geq 1, 1 \leq q \leq m}} \sum_{\substack{t_1+\dots+t_m=m-p \\ 0 \leq t_q \leq r_q k+1, 1 \leq q \leq m}} \prod_{j=1}^m c_{r_j, t_j}(k) \right| \\
& \leq \sum_{\substack{r_1+\dots+r_m=r \\ r_q \geq 1, 1 \leq q \leq m}} \sum_{\substack{t_1+\dots+t_m=m-p \\ 0 \leq t_q \leq r_q k+1, 1 \leq q \leq m}} \prod_{j=1}^m c_r k^{t_j-2} \\
& = c_r^m k^{-p-m} \sum_{\substack{r_1+\dots+r_m=r \\ r_q \geq 1, 1 \leq q \leq m}} \sum_{\substack{t_1+\dots+t_m=m-p \\ 0 \leq t_q \leq r_q k+1, 1 \leq q \leq m}} 1 \\
& \leq c_r^m k^{-p-m} \sum_{\substack{r_1+\dots+r_m=r \\ r_q \geq 1, 1 \leq q \leq m}} \sum_{\substack{t_1+\dots+t_m=m-p \\ t_q \geq 0, 1 \leq q \leq m}} 1
\end{aligned}$$

$$\begin{aligned}
&= c_r^m k^{-p-m} [z^r] \left(\frac{z}{1-z} \right)^m [z^{m-p}] \left(\frac{1}{1-z} \right)^m \\
&= c_r^m k^{-p-m} \binom{r-1}{p-1} \binom{2m-p-1}{m-1} \\
&\leq c_r^m \binom{2m-2}{m-1} \binom{r-1}{m-1} k^{-p-m},
\end{aligned}$$

and further

$$\begin{aligned}
|\tilde{f}_{r,p}(k)| &= \left| \sum_{m=p}^r \frac{1}{m} \sum_{\substack{r_1+\dots+r_m=r \\ r_q \geq 1, 1 \leq q \leq m}} \sum_{\substack{t_1+\dots+t_m=m-p \\ 0 \leq t_q \leq r_q k+1, 1 \leq q \leq m}} \prod_{j=1}^m c_{r_j, t_j}(k) \right| \\
&\leq \sum_{m=p}^r \frac{1}{m} c_r^m \binom{2m-2}{m-1} \binom{r-1}{m-1} k^{-p-m} \\
&\leq \binom{2r-2}{r-1} (r-1)! c_r^r \sum_{m=p}^r \frac{1}{k^{p+m}} = \binom{2r-2}{r-1} (r-1)! c_r^r \frac{1}{k^{2p}} \sum_{q=0}^{r-p} \frac{1}{k^q} \\
&\leq \binom{2r-2}{r-1} (r-1)! c_r^r \frac{1}{k^{2p}} r.
\end{aligned}$$

Thus, we obtain the following estimate, which holds for all $1 \leq p \leq r$ and $r, k \geq 1$:

$$|\tilde{f}_{r,p}(k)| \leq \kappa_r \frac{1}{k^{2p}}$$

with

$$\kappa_r := \binom{2r-2}{r-1} r! c_r^r.$$

The constants c_r appearing here are given by (3.12).

3.4.3. Cancellation of $\tilde{f}_{r,p}(\mathbf{k})$. We shall demonstrate that the functions $\tilde{f}_{r,p}(k)$ as defined by (3.9) satisfy $\tilde{f}_{r,p}(k) = 0$ for all $p \geq \lfloor \frac{r}{2} \rfloor + 1$. We will do this by simply showing that for every composition $m, r_1, \dots, r_m, t_1, \dots, t_m$ of r and $m-p$, respectively, there exists a factor $c_{1,0}(k)$, which is zero, and thus the product $\prod_{j=1}^m c_{r_j, t_j}(k)$ vanishes.

Let us take $p \geq \lfloor \frac{r}{2} \rfloor + 1$. This implies then by (3.9) that also $m \geq \lfloor \frac{r}{2} \rfloor + 1$. We consider now an arbitrary but fixed composition $m, r_1, \dots, r_m, t_1, \dots, t_m$ with $r_q \geq 1$ and $t_q \geq 0$ such that $r_1 + \dots + r_m = r$ and $t_1 + \dots + t_m = m-p$. We define the sets $T = \{q : t_q = 0\}$ and $R = \{q : r_q = 1\}$. Then the relations

$$|T| \geq m - (m-p) = p \quad \text{and} \quad |R| \geq 2m - r$$

hold. Otherwise, if $|T| \leq p-1$ this would imply $\sum_{q=1}^m t_q \geq m-p+1 > m-p$, and if $|R| \leq 2m-r-1$ this would imply $\sum_{q=1}^r r_q \geq (2m-r-1) + 2(m-(2m-r-1)) = r+1 > r$, which is a contradiction to the choice of r_q and t_q .

Next we consider the set $R \cap T = \{q : t_q = 0 \wedge r_q = 1\}$. Since $|R \cup T| \leq m$, we get by a simple application of the inclusion-exclusion principle for $p \geq \lfloor \frac{r}{2} \rfloor + 1$ the estimate

$$|R \cap T| = |R| + |T| - |R \cup T| \geq 2m - r + p - m = m + p - r \geq 2 \left\lfloor \frac{r}{2} \right\rfloor - r + 2 \geq 1.$$

The latter equation guarantees that in every composition there exists a number q such that $r_q = 1$ and $t_q = 0$, and consequently there is always a factor $c_{1,0}(k) = 0$ in every product $\prod_{j=1}^m c_{r_j, t_j}(k)$. Hence, we have

$$\tilde{f}_{r,p}(k) = 0 \quad \text{for } p \geq \left\lfloor \frac{r}{2} \right\rfloor + 1.$$

We also consider the special case $r = 2d$ and $p = d$; i.e., we examine

$$\tilde{f}_{2d,d}(k) = \sum_{m=d}^{2d} \frac{1}{m} \sum_{\substack{r_1+\dots+r_m=2d \\ r_q \geq 1, 1 \leq q \leq m}} \sum_{\substack{t_1+\dots+t_m=m-d \\ 0 \leq t_q \leq r_q k+1, 1 \leq q \leq m}} \prod_{j=1}^m c_{r_j, t_j}(k).$$

If $m \geq d+1$, we can show by arguments as above that there always exists a factor $c_{1,0}(k) = 0$ in every product $\prod_{j=1}^m c_{r_j, t_j}(k)$. Thus all summands with $m \geq d+1$ are zero. This gives

$$\tilde{f}_{2d,d}(k) = \frac{1}{d} \sum_{\substack{r_1+\dots+r_d=2d \\ r_q \geq 1, 1 \leq q \leq d}} \sum_{\substack{t_1+\dots+t_d=0 \\ 0 \leq t_q \leq r_q k+1, 1 \leq q \leq d}} \prod_{j=1}^d c_{r_j, t_j}(k) = \frac{1}{d} \sum_{\substack{r_1+\dots+r_d=2d \\ r_q \geq 1, 1 \leq q \leq d}} \prod_{j=1}^d c_{r_j, 0}(k).$$

Now, since $c_{1,0}(k) = 0$, we see that the only nonzero contribution is obtained if $r_1 = \dots = r_d = 2$, which implies

$$\tilde{f}_{2d,d}(k) = \frac{1}{d} c_{2,0}^d(k) = \frac{\nu^d(k)}{d2^d},$$

with $\nu(k) = \frac{2k^2-1}{k(k+1)^2(2k+1)}$.

3.4.4. Asymptotics of the centered moments. Summarizing the results of the previous subsections, we have for $r \geq 1$ that for all $n \geq rk+1$ the r th centered moments are given by

$$\mathbf{E}(\tilde{X}_{n,k}^r) = r! n \sum_{p=1}^{\lfloor \frac{r}{2} \rfloor} \tilde{f}_{r,p}(k) \binom{n+p-1}{p-1}$$

with

$$|\tilde{f}_{r,p}(k)| \leq \frac{\kappa_r}{k^{2p}} \quad \text{and} \quad \tilde{f}_{2d,d}(k) = \frac{\nu^d(k)}{d2^d}.$$

We are now going to evaluate the centered moments asymptotically for the subcritical case, i.e., $\frac{k^2}{n} = o(1)$; thus, we may always assume that we choose n large enough such that $\frac{(d-1)^2}{n} \leq 1$ and $\frac{k^2}{n} \leq \frac{1}{2}$ are satisfied.

We consider the case $r = 2d$ with $d \geq 1$, which gives

$$\begin{aligned} \mathbf{E}(\tilde{X}_{n,k}^{2d}) &= (2d)! \frac{\nu^d(k)}{d2^d} \times \frac{(n+d-1)^{d-1}n}{(d-1)!} + \sum_{p=1}^{d-1} (2d)! \tilde{f}_{2d,p}(k) \binom{n+p-1}{p-1} n \\ &= \frac{(2d)!}{d! 2^d} \nu^d(k) n^d (1 + R(n, k)), \end{aligned}$$

with

$$R(n, k) := \prod_{j=1}^{d-1} \left(1 + \frac{j}{n}\right) - 1 + \sum_{p=1}^{d-1} \frac{2^d d! \tilde{f}_{2d,p}(k)}{\nu^d(k) n^{d-1}} \binom{n+p-1}{p-1}.$$

We obtain the simple estimate

$$\begin{aligned} \left| \prod_{j=1}^{d-1} \left(1 + \frac{j}{n}\right) - 1 \right| &= \prod_{j=1}^{d-1} \left(1 + \frac{j}{n}\right) - 1 \leq \left(1 + \frac{d-1}{n}\right)^{d-1} - 1 \\ &\leq e^{\frac{(d-1)^2}{n}} - 1 \leq \frac{(d-1)^2}{n} e^{\frac{(d-1)^2}{n}} \\ &\leq \frac{(d-1)^2 e}{n}. \end{aligned}$$

Using the trivial bounds

$$\frac{1}{12k^2} \leq \nu(k) \leq \frac{1}{k^2},$$

which hold for all $k \geq 1$, we obtain the following estimates:

$$\begin{aligned} \left| \sum_{p=1}^{d-1} \frac{2^d d! \tilde{f}_{2d,p}(k) \binom{n+p-1}{p-1}}{\nu^d(k) n^{d-1}} \right| &\leq 2^d d! \kappa_{2d} \sum_{p=1}^{d-1} \frac{n^{p-1} \prod_{j=1}^{p-1} \left(1 + \frac{j}{n}\right)}{\nu^d(k) n^{d-1} k^{2p}} \\ &\leq 2^d d! \kappa_{2d} e^{\frac{(d-1)^2}{n}} \sum_{p=1}^{d-1} \frac{1}{n^{d-p} \nu^d(k) k^{2p}} \\ &\leq 2^d d! \kappa_{2d} e \sum_{p=1}^{d-1} \frac{12^d k^{2d}}{n^{d-p} k^{2p}} \\ &= (24)^d d! \kappa_{2d} e \frac{k^2}{n} \sum_{q=0}^{d-2} \left(\frac{k^2}{n}\right)^q \\ &\leq (24)^d d! \kappa_{2d} e \frac{k^2}{n} \times \frac{1}{1 - \frac{k^2}{n}} \\ &\leq 2e(24)^d d! \kappa_{2d} \frac{k^2}{n}. \end{aligned}$$

Combining these estimates yields, in the subcritical case,

$$|R(n, k)| \leq \frac{(d-1)^2 e}{n} + 2e(24)^d d! \kappa_{2d} \frac{k^2}{n} = o(1),$$

and further

$$\mathbf{E}(\tilde{X}_{n,k}^{2d}) \sim \frac{(2d)!}{2^d d!} (\nu(k)n)^d \quad \text{for } d \geq 1.$$

We have to consider also the case $r = 2d + 1$ with $d \geq 0$, which gives

$$\begin{aligned}
\left| \mathbf{E}(\tilde{X}_{n,k}^{2d+1}) \right| &= \left| (2d+1)! n \sum_{p=1}^d \tilde{f}_{2d+1,p}(k) \binom{n+p-1}{p-1} \right| \\
&\leq (2d+1)! \sum_{p=1}^d \frac{\kappa_{2d+1}}{k^{2p}} \frac{n^p}{(p-1)!} \prod_{j=1}^{p-1} \left(1 + \frac{j}{n}\right) \\
&\leq (2d+1)! \kappa_{2d+1} e^{\frac{(d-1)^2}{n}} \sum_{p=1}^d \left(\frac{n}{k^2}\right)^p \\
&\leq (2d+1)! \kappa_{2d+1} e \left(\frac{n}{k^2}\right)^d \sum_{q=0}^{d-1} \left(\frac{k^2}{n}\right)^q \\
&\leq (2d+1)! \kappa_{2d+1} e \left(\frac{n}{k^2}\right)^d \frac{1}{1 - \frac{k^2}{n}} \\
&\leq 2e[(2d+1)!] \kappa_{2d+1} \left(\frac{n}{k^2}\right)^d \\
&\leq 2e(12)^d [(2d+1)!] \kappa_{2d+1} (\nu(k)n)^d,
\end{aligned}$$

and thus

$$\mathbf{E}(\tilde{X}_{n,k}^{2d+1}) = O((\nu(k)n)^d) \quad \text{for } d \geq 0.$$

Therefore, considering

$$\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}} = \frac{X_{n,k} - \mathbf{E}(X_{n,k})}{\sqrt{\nu(k)n}},$$

with $\nu(k) = \frac{2k^2-1}{k(k+1)^2(2k+1)}$, we have for $k = o(\sqrt{n})$

$$\mathbf{E} \left(\left(\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}} \right)^{2d} \right) \rightarrow \frac{(2d)!}{d! 2^d} \quad \text{for } d \geq 1,$$

and

$$\mathbf{E} \left(\left(\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}} \right)^{2d+1} \right) = O \left(\frac{1}{\sqrt{\nu(k)n}} \right) = O \left(\frac{k}{\sqrt{n}} \right) \rightarrow 0 \quad \text{for } d \geq 0.$$

For the subcritical case, the moments of $\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}}$ converge to the moments of a standard normal distribution, which proves the convergence

$$\frac{X_{n,k} - \frac{n}{k(k+1)}}{\sqrt{\frac{(2k^2-1)n}{k(k+1)^2(2k+1)}}} \xrightarrow{D} \mathcal{N}(0, 1).$$

Part (a) of Theorem 2.1 has been established.

4. The variety of subtree sizes in binary search trees. A binary search tree is constructed from the permutation $(\pi_1, \dots, \pi_n)$ of the set $\{1, 2, \dots, n\}$ by the following algorithm. The first element of the permutation is inserted in an empty tree, and a root node is allocated for it. A subsequent element π_j (with $j \geq 2$) is directed to the left subtree if $\pi_j < \pi_1$; otherwise it is directed to the right subtree. In whichever subtree π_j goes, it is subjected to the same insertion algorithm recursively, until it is inserted in an empty subtree, in which case a node is allocated for it and linked appropriately as a left (resp., right) child if its rank is less than (resp., at least as much as) the value of the last node on the path.

Several models of randomness are in common use on binary trees. The uniform model in which all trees are equally likely has been proposed for applications in formal languages, compilers, computer algebra, etc. (see Kemp (1984)). However, for the searching and sorting algorithms alluded to, the *random permutation model* is considered to be more appropriate. In this model of randomness we assume that the tree is built from permutations of $\{1, \dots, n\}$, where a uniform probability model is imposed on the *permutations* instead of the trees. When all $n!$ permutations are equally likely or *random*, binary search trees are not equally likely. Several permutations give rise to the same tree, favoring shorter and well balanced trees rather than scrawny and tall shapes, which is a desirable property in searching and sorting algorithms (see Mahmoud (1992)). The term *random binary search tree* will refer to a binary search tree built from a random permutation. The random permutation model is not restrictive, as it covers a rather wide variety of instances, such as when the input is a sample drawn from *any* continuous probability distribution, and the construction algorithm is concerned only with the ranks of the keys, not their actual values.

Let $X_{n,k}$ be the random variable that counts the number of subtrees of size k on the fringe of a random binary search tree of size n , and let $M_k(z, v)$ be its generating function

$$M_k(z, v) = \sum_{n \geq 1} \sum_{m \geq 0} P\{X_{n,k} = m\} z^n v^m.$$

The binary dichotomy of the ranks in the permutation (relative to the first element) preserves the probabilistic structure in the conditionally independent subtrees. For all $n > k \geq 1$ the probabilities $P\{X_{n,k} = m\}$ satisfy the following recurrence:

$$P\{X_{n,k} = m\} = \frac{1}{n} \sum_{\substack{n_1 + n_2 = n-1 \\ m_1 + m_2 = m \\ m_1, m_2 \geq 0, n_1, n_2 \geq 0}} P\{X_{n_1,k} = m_1\} P\{X_{n_2,k} = m_2\},$$

with initial values $P\{X_{k,k} = 1\} = 1$ and $P\{X_{n,k} = 0\} = 1$, for $1 \leq n < k$. After multiplication with $nz^{n-1}v^m$ and summing up over $n > k$ and $m \geq 0$, we obtain a functional equation for the generating function and find

$$\frac{\partial}{\partial z} M_k(z, v) = (1 + M_k(z, v))^2 + (v - 1)kz^{k-1}$$

for $k \geq 1$, with initial condition $M_k(0, v) = 0$.

Substituting $Q_k(z, v) := M_k(z, v) + 1$, we obtain the Riccati differential equation

$$\frac{\partial}{\partial z} Q_k(z, v) = Q_k^2(z, v) + (v - 1)kz^{k-1}, \quad Q_k(0, v) = 1.$$

The solution to this differential equation already appears in Flajolet, Gordon, and Martínez (1997). Using this solution, we obtain a representation for our generating function:

$$(4.1) \quad M_k(z, v) = -1 + \frac{1 + \sum_{j \geq 1} \delta_j(k)(v-1)^j z^{(k+1)j} - \sum_{j \geq 1} \gamma_j(k)(v-1)^j z^{(k+1)j-1}}{1 - z - \sum_{j \geq 1} \beta_j(k)(v-1)^j z^{(k+1)j+1} + \sum_{j \geq 1} \alpha_j(k)(v-1)^j z^{(k+1)j}},$$

with the functions

$$(4.2) \quad \alpha_j(k) = \xi_{j,-1,0}(k), \quad \beta_j(k) = \xi_{j,1,0}(k),$$

$$(4.3) \quad \gamma_j(k) = \xi_{j,-1,-1}(k), \quad \delta_j(k) = \xi_{j,1,1}(k),$$

where

$$\xi_{j,m,s}(k) = \frac{(-1)^j k^j ((k+1)j + \llbracket s=1 \rrbracket)^s}{\prod_{i=1}^j [(ik+i)(ik+i+m)]}.$$

4.1. The factorial moments for binary search trees. From the solution (4.1) we can obtain the r th factorial moments, which lead directly to the limiting distribution for the critical and supercritical cases.

To get the r th factorial moments we use the substitution $w := v - 1$ in $M_k(z, v)$ and expand this function in powers of w . Straightforward computations then lead to the following expression:

$$M_k(z, 1+w) = -1 + \frac{1}{1-z} \left(1 + \sum_{j \geq 1} w^j (\delta_j(k) z^{(k+1)j} - \gamma_j(k) z^{(k+1)j-1}) \right) \\ \times \sum_{\ell \geq 0} \left(\frac{z}{1-z} \sum_{j \geq 1} w^j (\beta_j(k) z^{(k+1)j} - \alpha_j(k) z^{(k+1)j-1}) \right)^\ell.$$

Extract coefficients of w^r in $M_k(z, 1+w)$ eventually give for $r \geq 1$

$$[w^r]M_k(z, 1+w) \\ = \frac{z^{(k+1)r}}{(1-z)^{r+1}} (\beta_1(k)z - \alpha_1(k))^r \\ + z^{(k+1)r} \sum_{\ell=1}^{r-1} \frac{1}{(1-z)^{\ell+1}} \sum_{\substack{j_1+\dots+j_\ell=r \\ j_i \geq 1}} \prod_{i=1}^{\ell} (\beta_{j_i}(k)z - \alpha_{j_i}(k)) \\ + z^{(k+1)r-1} \sum_{j=1}^{r-1} (\delta_j(k)z - \gamma_j(k)) \sum_{\ell=1}^{r-j} \frac{1}{(1-z)^{\ell+1}} \\ \times \sum_{\substack{j_1+\dots+j_\ell=r-j \\ j_i \geq 1}} \prod_{i=1}^{\ell} (\beta_{j_i}(k)z - \alpha_{j_i}(k)) + \frac{z^{(k+1)r-1}}{1-z} (\delta_r(k)z - \gamma_r(k)).$$

Using

$$(\beta_1(k)z - \alpha_1(k))^r = (\beta_1(k) - \alpha_1(k))^r + (\beta_1(k) - \alpha_1(k))^r \\ \times \sum_{m=1}^r \binom{r}{m} (-1)^m \left(\frac{\beta_1(k)}{\beta_1(k) - \alpha_1(k)} \right)^m (1-z)^m$$

and analogous expansions, and $[z^n] \frac{1}{(1-z)^{\alpha+1}} = \binom{n+\alpha}{n}$, we can expand the above expression also around $z = 1$ and get the following formula for the r th factorial moment of $X_{n,k}$, which is valid for all $n \geq (k+1)r$ and $r \geq 1$:

$$\begin{aligned}
\mathbf{E}(X_{n,k}^r) &= r! [z^n w^r] M_k(z, 1+w) \\
&= r! \binom{n-kr}{r} \left(\frac{2}{(k+1)(k+2)} \right)^r \\
&\quad + r! \left(\frac{1}{(k+1)(k+2)} \right)^r \sum_{p=1}^r \binom{n-(k+1)r+p-1}{p-1} \binom{r}{p-1} \left(\frac{k}{2} \right)^{r+1-p} \\
&\quad + r! \sum_{p=1}^r \binom{n-(k+1)r+p-1}{p-1} \sum_{\ell=\max(1,p-1)}^{r-1} (-1)^{\ell+1-p} \\
&\quad \times \sum_{\substack{j_1+\dots+j_\ell=r \\ j_i \geq 1}} \prod_{i=1}^{\ell} (\beta_{j_i}(k) - \alpha_{j_i}(k)) \\
&\quad \times \sum_{1 \leq i_1 < \dots < i_{\ell+1-p} \leq \ell} \prod_{q=1}^{\ell+1-p} \left(\frac{\beta_{j_{i_q}}(k)}{\beta_{j_{i_q}}(k) - \alpha_{j_{i_q}}(k)} \right) \\
&\quad + r! \sum_{p=1}^r \binom{n-(k+1)r+p}{p-1} \sum_{j=1}^{\min(r+1-p, r-1)} (\delta_j(k) - \gamma_j(k)) \\
&\quad \times \sum_{\ell=\max(1,p-1)}^{r-j} (-1)^{\ell+1-p} \sum_{\substack{j_1+\dots+j_\ell=r-j \\ j_i \geq 1}} \prod_{i=1}^{\ell} (\beta_{j_i}(k) - \alpha_{j_i}(k)) \\
&\quad \times \sum_{1 \leq i_1 < \dots < i_{\ell+1-p} \leq \ell} \prod_{q=1}^{\ell+1-p} \left(\frac{\beta_{j_{i_q}}(k)}{\beta_{j_{i_q}}(k) - \alpha_{j_{i_q}}(k)} \right) \\
&\quad - r! \sum_{p=1}^{r-1} \binom{n-(k+1)r+p}{p-1} \sum_{j=1}^{\min(r-p, r-1)} \delta_j(k) \\
&\quad \times \sum_{\ell=\max(1,p)}^{r-j} (-1)^{\ell-p} \sum_{\substack{j_1+\dots+j_\ell=r-j \\ j_i \geq 1}} \prod_{i=1}^{\ell} (\beta_{j_i}(k) - \alpha_{j_i}(k)) \\
&\quad \times \sum_{1 \leq i_1 < \dots < i_{\ell-p} \leq \ell} \prod_{q=1}^{\ell-p} \left(\frac{\beta_{j_{i_q}}(k)}{\beta_{j_{i_q}}(k) - \alpha_{j_{i_q}}(k)} \right) + r! (\delta_r(k) - \gamma_r(k)).
\end{aligned} \tag{4.4}$$

This yields, in particular, results for the expectation and the variance:

$$\begin{aligned}
(4.5) \quad \mathbf{E}(X_{n,k}) &= \frac{2(n+1)}{(k+1)(k+2)} \quad \text{for } n \geq k+1, \\
\mathbf{V}(X_{n,k}) &= \frac{2k(4k^2+5k-3)(n+1)}{(k+1)(k+2)^2(2k+1)(2k+3)} \quad \text{for } n \geq 2(k+1).
\end{aligned}$$

4.2. The critical case. We consider the critical case and assume that $k := k_n$ grows with n such that $\frac{n}{k^2} \rightarrow \lambda$ for some $\lambda > 0$.

Finding the limiting distribution requires some estimates on the functions appearing in (4.4). Using the definition of $\alpha_j(k)$, $\beta_j(k)$, $\gamma_j(k)$, and $\delta_j(k)$ given by (4.2)–(4.3), the following estimates, which hold for all $k \geq 2$ and $j \geq 1$, are not hard to show:

$$(4.6) \quad |\beta_j(k) - \alpha_j(k)| \leq \frac{2}{k^{j+1}}, \quad \left| \frac{\beta_j(k)}{\beta_j(k) - \alpha_j(k)} \right| \leq 4^j k,$$

$$(4.7) \quad |\delta_j(k) - \gamma_j(k)| \leq \frac{3}{k^j}, \quad |\delta_j(k)| \leq \frac{1}{k^{j-1}}.$$

We can now deduce the asymptotic behavior of the r th factorial moments of $X_{n,k}$ by inspecting the summands of (4.4). Considering a fixed $r \geq 1$, we trivially obtain for the first summand

$$\begin{aligned} r! \binom{n-kr}{r} \left(\frac{2}{(k+1)(k+2)} \right)^r &= \left(\frac{2n}{k^2} \right)^r \left(1 + O\left(\frac{1}{k}\right) + \left(\frac{k}{n}\right) \right) \\ &= (2\lambda)^r \left(1 + O\left(\frac{1}{\sqrt{n}}\right) \right). \end{aligned}$$

Let us denote by B the remaining summands of (4.4). They can then be estimated for fixed $r \geq 1$ by using (4.6) and (4.7), which finally gives

$$|B| = O\left(\frac{1}{k}\right) = O\left(\frac{1}{\sqrt{n}}\right).$$

Summarizing, we obtain for all $r \geq 1$ and sequences (n, k) such that $\frac{n}{k^2} \rightarrow \lambda$ the following asymptotic expansion:

$$\mathbf{E}(X_{n,k}^r) = (2\lambda)^r \left(1 + O\left(\frac{1}{\sqrt{n}}\right) \right).$$

This shows convergence in distribution of $X_{n,k}$ to a Poisson random variable with parameter 2λ . Thus we proved the convergence asserted in Theorem 2.2(b), where we used the substitution $c := \frac{1}{\sqrt{\lambda}}$, and thus $\frac{k}{\sqrt{n}} \rightarrow c$. When we have a critical case with $k/\sqrt{n}$ not converging to a limit, all factorial (and ordinary) moments oscillate, and no limit distribution exists for $X_{n,k}$.

4.3. The supercritical case. Again, by using the estimates (4.6) and (4.7), one can show $\mathbf{E}(X_{n,k}) \rightarrow 0$ for $\frac{n}{k^2} = o(1)$ under the natural restriction $n > k$. Thus, we have convergence of $X_{n,k}$ to a degenerate distribution with mass at 0.

4.4. The subcritical case. We consider the centered random variables $\tilde{X}_{n,k} := X_{n,k} - \mathbf{E}(X_{n,k})$ and introduce the generating function

$$\tilde{M}_k(z, s) := \sum_{n \geq 1} \mathbf{E}(e^{\tilde{X}_{n,k}s}) z^n = \sum_{n \geq 1} e^{-\mathbf{E}(X_{n,k})s} \mathbf{E}(e^{X_{n,k}s}) z^n.$$

Using the explicit formula for $\mathbf{E}(X_{n,k})$ as given by (4.5), we obtain

$$\begin{aligned} \tilde{M}_k(z, s) &= e^{-\frac{2s}{(k+1)(k+2)}} M_k(e^{-\frac{2s}{(k+1)(k+2)}} z, e^s) + (1 - e^{\frac{ks}{k+2}}) z^k \\ &\quad + \sum_{1 \leq n < k} z^n - \sum_{1 \leq n < k} e^{-\frac{2(n+1)s}{(k+1)(k+2)}} z^n, \end{aligned}$$

which, by using (4.1) and (4.2)–(4.3), further leads to

$$\begin{aligned}\tilde{M}_k(z, s) &= \frac{e^{-\frac{2s}{(k+1)(k+2)}} U}{V} - e^{-\frac{2s}{(k+1)(k+2)}} + (1 - e^{\frac{ks}{k+2}}) z^k \\ &\quad + \sum_{1 \leq n < k} z^n - \sum_{1 \leq n < k} e^{-\frac{2(n+1)s}{(k+1)(k+2)}} z^n,\end{aligned}$$

with the expressions

$$\begin{aligned}U &:= 1 + \sum_{j \geq 1} \delta_j(k) (e^s - 1)^j e^{-\frac{2js}{k+2}} z^{(k+1)j} \\ &\quad - \sum_{j \geq 1} \gamma_j(k) (e^s - 1)^j e^{-\frac{2((k+1)j-1)s}{(k+1)(k+2)}} z^{(k+1)j-1}, \\ V &:= 1 - e^{-\frac{2s}{(k+1)(k+2)}} z - \sum_{j \geq 1} \beta_j(k) (e^s - 1)^j e^{-\frac{2((k+1)j+1)s}{(k+1)(k+2)}} z^{(k+1)j+1} \\ &\quad + \sum_{j \geq 1} \alpha_j(k) (e^s - 1)^j e^{-\frac{2js}{k+2}} z^{(k+1)j}.\end{aligned}$$

4.4.1. Expanding around $s = 0$. We expand $\tilde{M}_k(z, s)$ around $s = 0$ and $z = 1$. It is straightforward, though rather laborious, to come up with the following representation:

$$M_k(e^{-\frac{2s}{(k+1)(k+2)}} z, e^s) = -1 + \frac{\sum_{\ell \geq 0} s^\ell \sum_{i=0}^{\ell(k+1)} d_{\ell,i}(k) (1-z)^i}{1 - z - \sum_{\ell \geq 1} s^\ell \sum_{i=0}^{\ell(k+1)+1} c_{\ell,i}(k) (1-z)^i},$$

where the functions $d_{\ell,i}(k)$ and $c_{\ell,i}(k)$ are given as follows:

$$\begin{aligned}(4.8) \quad d_{\ell,i}(k) &= (-1)^i \sum_{j=1}^{\ell} \sum_{m=0}^{\ell-j} \frac{j! \{ \begin{smallmatrix} \ell-m \\ j \end{smallmatrix} \} (-1)^m}{m! (\ell-m)! (k+2)^m} \\ &\quad \times \left(\binom{(k+1)j-1}{i} \left((2j)^m \delta_j(k) - \frac{(2((k+1)j-1))^m \gamma_j(k)}{(k+1)^m} \right) \right. \\ &\quad \left. + \binom{(k+1)j-1}{i-1} (2j)^m \delta_j(k) \mathbb{I}[i \geq 1] \right) \quad \text{for } \ell \geq 1,\end{aligned}$$

$$d_{0,0}(k) = 1,$$

$$\begin{aligned}(4.9) \quad c_{\ell,i}(k) &= (-1)^i \sum_{j=1}^{\ell} \sum_{m=0}^{\ell-j} \frac{\{ \begin{smallmatrix} \ell-m \\ j \end{smallmatrix} \} j! (-1)^m}{m! (\ell-m)! (k+2)^m} \\ &\quad \times \left(\binom{(k+1)j}{i} \left(\frac{(2((k+1)j+1))^m \beta_j(k)}{(k+1)^m} - (2j)^m \alpha_j(k) \right) \right. \\ &\quad \left. + \binom{(k+1)j}{i-1} \frac{(2((k+1)j+1))^m \beta_j(k)}{(k+1)^m} \mathbb{I}[i \geq 1] \right) \\ &\quad + \frac{(-1)^\ell 2^\ell}{\ell! (k+1)^\ell (k+2)^\ell} \mathbb{I}[i=0] - \frac{(-1)^\ell 2^\ell}{\ell! (k+1)^\ell (k+2)^\ell} \mathbb{I}[i=1].\end{aligned}$$

Thus, $\tilde{M}_k(z, s)$ can be written as

$$(4.10) \quad \tilde{M}_k(z, s) = \frac{e^{-\frac{2s}{(k+1)(k+2)}} \left(\sum_{\ell \geq 0} s^\ell \sum_{i=0}^{\ell(k+1)} d_{\ell,i}(k) (1-z)^i \right)}{1 - z - \sum_{\ell \geq 1} s^\ell \sum_{i=0}^{\ell(k+1)+1} c_{\ell,i}(k) (1-z)^i} - e^{-\frac{2s}{(k+1)(k+2)}} \\ + (1 - e^{\frac{ks}{k+2}}) z^k + \sum_{1 \leq n < k} z^n - \sum_{1 \leq n < k} e^{-\frac{2(n+1)s}{(k+1)(k+2)}} z^n.$$

The only interesting part in (4.10) is the first summand, since the remaining summands do not give a contribution for $n > k$ (the remainder is a polynomial in z of degree k). The first summand of (4.10) can be expanded around $s = 0$ and $z = 1$, and for $r \geq 1$ finally leads to the following representation:

$$[s^r] \frac{e^{-\frac{2s}{(k+1)(k+2)}} \left(\sum_{\ell \geq 0} s^\ell \sum_{i=0}^{\ell(k+1)} d_{\ell,i}(k) (1-z)^i \right)}{1 - z - \sum_{\ell \geq 1} s^\ell \sum_{i=0}^{\ell(k+1)+1} c_{\ell,i}(k) (1-z)^i} \\ = \sum_{p=0}^r \frac{1}{(1-z)^{p+1}} f_{r,p}(k) + \sum_{p=0}^{r-1} \frac{1}{(1-z)^{p+1}} g_{r,p}(k) + \sum_{p=0}^{r(k+1)-1} h_{r,p}(k) z^p$$

with

$$(4.11) \quad f_{r,p}(k) = \sum_{m=p}^r \sum_{\substack{r_1 + \dots + r_m = r \\ r_q \geq 1}} \sum_{\substack{t_1 + \dots + t_m = m-p \\ 0 \leq t_q \leq r_q(k+1)+1}} \prod_{j=1}^m c_{r_j, t_j}(k), \\ (4.12) \quad g_{r,p}(k) = \sum_{c=1}^r \sum_{a=p}^{r-c} \left(\sum_{m=a}^{r-c} \sum_{\substack{r_1 + \dots + r_m = r-c \\ r_q \geq 1}} \sum_{\substack{t_1 + \dots + t_m = m-a \\ 0 \leq t_q \leq r_q(k+1)+1}} \prod_{j=1}^m c_{r_j, t_j}(k) \right) \\ \times \left(\sum_{b=\lceil \frac{a-p}{k+1} \rceil}^c \frac{(-1)^{c-b} 2^{c-b} d_{b, a-p}(k)}{(k+1)^{c-b} (k+2)^{c-b} (c-b)!} \right)$$

and certain functions $h_{r,p}(k)$, which are of course irrelevant for our purpose.

This also leads to the required representation

$$[s^r] \tilde{M}_k(z, s) = \sum_{p=0}^r \frac{1}{(1-z)^{p+1}} f_{r,p}(k) + \sum_{p=0}^{r-1} \frac{1}{(1-z)^{p+1}} g_{r,p}(k) + \sum_{p=0}^{r(k+1)-1} \tilde{h}_{r,p}(k) z^p,$$

where the functions $f_{r,p}(k)$, and $g_{r,p}(k)$ are defined by (4.11) and (4.12). The functions $\tilde{h}_{r,p}(k)$ are irrelevant and thus not given explicitly.

The representation in the last display gives the following explicit formula for the r th moments of $\tilde{X}_{n,k}$, which is valid for $n \geq r(k+1)$:

$$\mathbf{E}(\tilde{X}_{n,k}^r) = r! [z^n s^r] \tilde{M}_k(z, s) = r! \sum_{p=0}^r \binom{n+p}{p} f_{r,p}(k) + r! \sum_{p=0}^r \binom{n+p}{p} g_{r,p}(k),$$

where the functions $f_{r,p}(k)$ and $g_{r,p}(k)$ are given by (4.11) and (4.12).

4.4.2. Estimates for $c_{\ell,i}(\mathbf{k})$ and $d_{\ell,i}(\mathbf{k})$. One obtains easily the following estimates of the functions $c_{\ell,i}(\mathbf{k})$ and $d_{\ell,i}(\mathbf{k})$ given by (4.8) and (4.9):

$$\begin{aligned} |d_{\ell,i}(\mathbf{k})| &\leq 2\ell^2(2\ell)^{2\ell+1}B_\ell(2\ell)^i k^{i-1} && \text{for } \ell \geq 1, i \geq 0, k \geq 1, \\ |c_{\ell,i}(\mathbf{k})| &\leq B_\ell 2^{2\ell+3}\ell^{\ell+3}(2\ell)^i k^{i-2} && \text{for } \ell \geq 1, i \geq 0, k \geq 1. \end{aligned}$$

Moreover, we have the following values:

$$c_{1,0}(\mathbf{k}) = 0 \quad \text{and} \quad c_{2,0}(\mathbf{k}) = \frac{\nu(\mathbf{k})}{2},$$

with

$$\nu(\mathbf{k}) := \frac{2k(4k^2 + 5k - 3)}{(k+1)(k+2)^2(2k+1)(2k+3)}.$$

4.4.3. Estimates for $f_{r,p}(\mathbf{k})$ and $g_{r,p}(\mathbf{k})$. It is possible to give suitable estimates on the growth of $f_{r,p}(\mathbf{k})$ and $g_{r,p}(\mathbf{k})$: There exist constants κ_r and η_r (depending only on r) such that

$$|f_{r,p}(\mathbf{k})| \leq \kappa_r \frac{1}{k^{2p}} \quad \text{and} \quad |g_{r,p}(\mathbf{k})| \leq \eta_r \frac{1}{k^{2p+2}}$$

for all $0 \leq p \leq r$ (resp., $0 \leq p \leq r-1$) and $k \geq 1$.

For instance one can choose the constants

$$\kappa_r = \frac{(2r-1)!(r+1)}{r!} c_r^r, \quad \eta_r = 2^{r+2} \binom{2r-1}{r-1} r! r c_r^r d_r,$$

with

$$c_r := B_r 2^{2r+3} r^{r+3} (2r)^r \quad \text{and} \quad d_r := 2B_r r^r (2r)^{3r+1}.$$

4.4.4. Cancellation of $\tilde{f}_{r,p}(\mathbf{k})$. Using the same “combinatorial argumentation” as for recursive trees, we can show that

$$\begin{aligned} f_{r,p}(\mathbf{k}) &= 0 && \text{for } p \geq \left\lfloor \frac{r}{2} \right\rfloor + 1, \\ g_{r,p}(\mathbf{k}) &= 0 && \text{for } p \geq \left\lceil \frac{r}{2} \right\rceil. \end{aligned}$$

Furthermore, one can show

$$f_{2d,d}(\mathbf{k}) = c_{2,0}^d(\mathbf{k}) = \frac{\nu^d(\mathbf{k})}{2^d}.$$

4.4.5. Asymptotics of the centered moments. In view of the development in subsection 4.4.3 we obtain the following explicit formulas for the r th moments of $X_{n,k}$:

$$\begin{aligned} \mathbf{E}(\tilde{X}_{n,k}^{2d}) &= (2d)! \left(\binom{n+d}{d} f_{2d,d}(\mathbf{k}) \right. \\ &\quad \left. + \sum_{p=0}^{d-1} \binom{n+p}{p} (f_{2d,p}(\mathbf{k}) + g_{2d,p}(\mathbf{k})) \right), \quad d \geq 1, \\ \mathbf{E}(\tilde{X}_{n,k}^{2d+1}) &= (2d+1)! \sum_{p=0}^d \binom{n+p}{p} (f_{2d+1,p}(\mathbf{k}) + g_{2d+1,p}(\mathbf{k})), \quad d \geq 0. \end{aligned}$$

Using the previous estimates, we can easily show that, for $\frac{k^2}{n} \rightarrow 0$,

$$\begin{aligned}\mathbf{E}(\tilde{X}_{n,k}^{2d}) &\rightarrow \frac{(2d)!}{d! 2^d} n^d \nu^d(k), \\ \mathbf{E}(\tilde{X}_{n,k}^{2d+1}) &= O\left(\left(\frac{n}{k^2}\right)^d\right) = O(n^d \nu^d(k)).\end{aligned}$$

Therefore, considering

$$\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}} = \frac{X_{n,k} - \mathbf{E}(X_{n,k})}{\sqrt{\nu(k)n}},$$

with $\nu(k) = \frac{2k(4k^2+5k-3)}{(k+1)(k+2)^2(2k+1)(2k+3)}$, we have for $k = o(\sqrt{n})$

$$\begin{aligned}\mathbf{E}\left(\left(\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}}\right)^{2d}\right) &\rightarrow \frac{(2d)!}{d! 2^d} \quad \text{for } d \geq 1, \\ \mathbf{E}\left(\left(\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}}\right)^{2d+1}\right) &= O\left(\frac{k}{\sqrt{n}}\right) \rightarrow 0 \quad \text{for } d \geq 0,\end{aligned}$$

whence for the subcritical case we have

$$\frac{X_{n,k} - \frac{2n}{(k+1)(k+2)}}{\sqrt{\frac{2k(4k^2+5k-3)n}{(k+1)(k+2)^2(2k+1)(2k+3)}}} \xrightarrow{D} \mathcal{N}(0,1),$$

completing the proof for the subcritical case (part (a) of Theorem 2.2).

5. Concluding remarks. Several methods have recently come into the repertoire of the analysis of algorithms, and there is a question of which method to choose for an analysis such as the one we conducted in this investigation. For example, the contraction method (introduced in Rösler (1991)), Pólya urns (Johnson and Kotz (1977)), the Chen–Stein method (Barbour, Holst, and Janson (1992)), methods for additive functions in subtrees (Devroye (1991) and Devroye (2002)) are all attractive for this kind of analysis. These methods have been used successfully in the analysis of algorithms (see, for example Neininger (2002) or Rösler and Rüschendorf (2001) for applications of the contraction method). However, experience shows that several of these methods will work with ease in some of the phases, but not systematically across all the phases. For instance, in our multiphase problem, Pólya urns work well in the very low range (small fixed k), as was attempted in Feng, Mahmoud, and Su (2007). The difficulty in this method is that one needs a detailed description of the fringe in terms of urns that are different from each other for each k and that become more complex as k increases. As for the contraction method, again it can be used in the case of small fixed k , but it does not lend itself easily to $k = k_n \rightarrow \infty$ in the subcritical case when k increases at a rate slower than $\sqrt{n}$. The difficulty stems from complicated toll functions. It is not clear how to apply the contraction method, or the other methods above, in the critical case. It is our experience that the analytic methods, based on generating functions, are systematic enough to cover all the ranges.

REFERENCES

- A. BARBOUR, L. LARS HOLST, AND S. JANSON (1992), *Poisson Approximation*, Oxford University Press, Oxford, UK.
- F. BERGERON, P. FLAJOLET, AND B. SALVY (1992), *Varieties of increasing trees*, in Proceedings of the 17th Colloquium on Trees in Algebra and Programming (CAAP '92), Lecture Notes in Comput. Sci. 581, Springer-Verlag, Berlin, pp. 24–48.
- H. CHERN, H. HWANG, AND T. TSAI (2002), *An asymptotic theory for Cauchy-Euler differential equations with applications to the analysis of algorithms*, J. Algorithms, 44, pp. 177–225.
- F. CHYZAK, M. DRMOTA, T. KLAUSNER, AND G. KOK (2007), *The distribution of patterns in random trees*, Combin. Probab. Comput., to appear. (Published online September 28, 2007.)
- L. DEVROYE (1991), *Limit laws for local counters in random binary search trees*, Random Structures Algorithms, 3, pp. 303–315.
- L. DEVROYE (2002), *Limit laws for sums of functions of subtrees of random binary search trees*, SIAM J. Comput., 32, pp. 152–171.
- Q. FENG, H. MAHMOUD, AND C. SU (2007), *On the Variety of Subtrees in a Random Recursive Tree*, Technical report, The George Washington University, Washington, DC.
- P. FLAJOLET, X. GOURDON, AND C. MARTÍNEZ (1997), *Patterns in random binary search trees*, Random Structures Algorithms, 11, pp. 223–244.
- M. FUCHS, H.-K. HWANG, AND R. NEININGER (2006), *Profiles of random trees: Limit theorems for random recursive trees and binary search trees*, Algorithmica, 46, pp. 367–407.
- J. GASTWIRTH AND P. BHATTACHARYA (1984), *Two probability models of pyramids or chain letter schemes demonstrating that their promotional claims are unreliable*, Oper. Res., 32, pp. 527–536.
- N. JOHNSON AND S. KOTZ (1977), *Urn Models and Their Applications*, Wiley, New York.
- R. KEMP (1984), *Fundamentals of the Average Case Analysis of Particular Algorithms*, Wiley, New York.
- D. KNUTH (1998), *The Art of Computer Programming, Vol. 3: Sorting and Searching*, 2nd ed., Addison-Wesley, Reading, MA.
- J. LENT AND H. MAHMOUD (1996), *Average-case analysis of multiple Quickselect: An algorithm for finding order statistics*, Statist. Probab. Lett., 28, pp. 299–310.
- H. MAHMOUD (1992), *Evolution of Random Search Trees*, Wiley, New York.
- H. MAHMOUD (2000), *Sorting: A Distribution Theory*, Wiley, New York.
- D. NAJOCK AND C. HEYDE (1982), *On the number of terminal vertices in certain random trees with an application to stemma construction in philology*, J. Appl. Probab., 19, pp. 675–680.
- R. NEININGER (2002), *The Wiener index of random trees*, Combin. Probab. Comput., 11, pp. 587–597.
- U. RÖSLER (1991), *A limit theorem for “Quicksort,”* RAIRO Inform. Théor. Appl., 25, pp. 85–100.
- U. RÖSLER AND L. RÜSCHENDORF (2001), *The contraction method for recursive algorithms*, Algorithmica, 29, pp. 3–33.
- A. PANHOLZER AND H. PRODINGER (1998), *A generating functions approach for the analysis of grand averages for multiple QUICKSELECT*, Random Structures Algorithms, 13, pp. 189–209.
- R. SMYTHE AND H. MAHMOUD (1994), *A survey of recursive trees*, Teor. ĭmovĭr. Mat. Stat., 51, pp. 1–27 (in Ukrainian); Theory Probab. Math. Statist., 51 (1996), pp. 1–27 (in English).

l -CHAIN PROFILE VECTORS*

DÁNIEL GERBNER[†] AND BALÁZS PATKÓ[‡]

Abstract. The l -chain profile vector of a set system $\mathcal{F}$ on an underlying set X of size n is defined to be the vector of length $\binom{n+1}{l}$ in which the α th component ($\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)$ $0 \leq \alpha_1 < \alpha_2 < \dots < \alpha_l \leq n$) is the number of l -chains in $\mathcal{F}$ with the smallest set having size α_1 , the second smallest α_2 , and so on. We modify the method of Erdős, Frankl, and Katona to determine the l -chain profile polytope of some sets of families including k -Sperner, complement-free, and complement-free k -Sperner families.

Key words. profile vectors, k -Sperner, intersecting, complement-free families

AMS subject classification. 05D05

DOI. 10.1137/060666366

1. Introduction. Basic problems in extremal set theory consider a set $\mathbb{A}$ of families of subsets of an n -element set all having some fixed property (Sperner, intersecting, complement-free, etc.) and ask for the family with the largest size. That is, we sum over all sets in the family, adding one for each set, and we ask for the family with the largest sum (largest *weight*). A natural generalization of this problem is in the summation we allow *weights* different from 1, but depending only on the size of the subset in question. For example, in the famous LYM-inequality a k -element set has weight $\frac{1}{\binom{n}{k}}$.

The tool that helps in dealing with this generalized problem is the so-called *profile vector* of a family $\mathcal{F}$. It is a vector $f(\mathcal{F}) \in \mathbb{R}^{n+1}$ whose components are defined by

$$f(\mathcal{F})_i = |\{F \in \mathcal{F} : |F| = i\}|.$$

Given a set of families $\mathbb{A}$, let $\mu(\mathbb{A})$ denote the set of the profile vectors of families in $\mathbb{A}$ (i.e., $\{f(\mathcal{F}) : \mathcal{F} \in \mathbb{A}\}$). Now determining the family in $\mathbb{A}$ with the largest weight (for a weight function w as above, where we denote the weight of an i -element set by w_i) is equivalent to finding the family of which the profile maximizes

$$w \cdot f(\mathcal{F}) = \sum_{i=0}^n w_i f(\mathcal{F})_i$$

among profiles in $\mu(\mathbb{A})$.

A basic fact of linear programming says that this maximum is attained at some *extreme point* of the *convex hull* of $\mu(\mathbb{A})$, which we call the *profile polytope* of $\mathbb{A}$; we denote it by $\langle \mu(\mathbb{A}) \rangle$. And if all w_i s are nonnegative, then the maximum is taken at an *essential extreme point* (i.e., an extreme point that is maximal with respect to the componentwise ordering). Therefore, to establish “all linear inequalities” concerning a set $\mathbb{A}$ of families, one has to determine its profile polytope and its extreme points. The

*Received by the editors July 30, 2006; accepted for publication (in revised form) November 7, 2007; published electronically February 15, 2008.

<http://www.siam.org/journals/sidma/22-1/66636.html>

[†]Department of Information Systems, Eötvös University, Pázmány Péter sétány 1/B, Budapest, 1117 Hungary (gerbner@cs.elte.hu).

[‡]Department of Mathematics and Its Applications, Central European University, Nádor u. 9, Budapest, 1051 Hungary (patkos@renyi.hu, tphab01@ceu.hu).

first result in this area (implicitly, without using the notion of the profile polytope) is due to Katona [7], and the systematic investigation of profile polytopes was started by Erdős, Frankl, and Katona in [4], [5], [6]. A survey on this topic can be found in the book of Engel [3].

However, there are problems dealing with other kinds of weight functions, and there are problems dealing not with sets of some families, but subfamilies of families. A family of sets is called k -Sperner (and Sperner or *antichain* if $k = 1$) if it does not contain a chain of length $k + 1$ (a sequence of sets of length $k + 1$ in which every set contains the previous one). By the dual version of Dilworth's theorem we know that being k -Sperner is equivalent to being the union of at most k antichains. A natural question is the following: Letting $l \leq k$ be two integers, how many l -chains can be contained in a family without a $k + 1$ -chain (i.e., in a k -Sperner family)? Note that this problem is analogous to the celebrated theorem of Turán [12], generalized by Sauer [11] (see also [2]), stating that for all integers $2 \leq r \leq s$ a graph on n vertices without a clique of size $s + 1$ can contain at most as many cliques of size r as does the s -partite Turán graph on n vertices (the complete s -partite graph with equipartite partition).

To deal with the above problem we introduce the notion of an l -chain profile vector of a family $\mathcal{F}$ on an n -element underlying set. This has $\binom{n+1}{l}$ components, and the α th component f_α , where $\alpha = (\alpha_1, \dots, \alpha_l)$ with $0 \leq \alpha_1 < \alpha_2 < \dots < \alpha_l \leq n$, denotes the number of l -chains in $\mathcal{F}$ in which the smallest set has size α_1 , the second smallest has size α_2 , and so on. Note that for $l = 1$ this is just the original notion of the profile vector.

The rest of the paper is organized as follows. In section 2 we provide some more definitions and make some easy remarks on the connection between the original notion of a profile vector and the newly introduced notion of an l -chain profile. In section 3 we prove our main tool in determining l -chain profile polytopes, and in section 4 we use this tool (the reduction method) to get some results on some sets of families.

2. Definitions and remarks. In this section we give some further definitions and describe some basic connections between the extreme points in the l -chain case and the extreme points in the original (1-chain) case.

Notation. For $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)$, $0 \leq \alpha_1 < \alpha_2 < \dots < \alpha_l \leq n$, we define the following multinomial coefficient:

$$\binom{n}{\alpha} = \prod_{i=1}^l \binom{n - \alpha_{i-1}}{\alpha_i - \alpha_{i-1}} = \frac{n!}{\alpha_1!(\alpha_2 - \alpha_1)! \dots (\alpha_l - \alpha_{l-1})!(n - \alpha_l)!},$$

where $\alpha_0 = 0$ and $0! = 1$ as usual. Note that $\binom{n}{\alpha}$ is the number of l -chains that can be formed from subsets of an n -element set in such a way that the smallest set has size α_1 , the second smallest has size α_2 , and so on.

Definition. Given an underlying set X and a family $\mathcal{F}$ of its subsets, the upset of $\mathcal{F}$ is $\mathcal{U}(\mathcal{F}) = \{G \subseteq X : \exists F \in \mathcal{F} \text{ such that } F \subseteq G\}$, and the downset of $\mathcal{F}$ is $\mathcal{D}(\mathcal{F}) = \{G \subseteq X : \exists F \in \mathcal{F} \text{ such that } F \supseteq G\}$.

Definition. A set $\mathbb{A}$ of families is upward (downward) closed if $\mathcal{F} \in \mathbb{A}$ implies $\mathcal{U}(\mathcal{F}) \in \mathbb{A}$ ($\mathcal{D}(\mathcal{F}) \in \mathbb{A}$).

Examples. Clearly the set of t -intersecting (t -cointersecting) families is upward (downward) closed. (A family $\mathcal{F}$ is said to be t -intersecting if, for any two $F_1, F_2 \in \mathcal{F}$, $|F_1 \cap F_2| \geq t$, and a family $\mathcal{G}$ is said to be t -cointersecting if $\overline{\mathcal{G}} = \{\overline{G} : G \in \mathcal{G}\}$ is t -intersecting or equivalently if, for any two $G_1, G_2 \in \mathcal{G}$, $|\overline{G_1} \cup \overline{G_2}| \geq t$.)

Definition. Let $\mu_l(\mathbb{A})$ denote the set of all l -chain profile vectors of families in $\mathbb{A}$, $\langle \mu_l(\mathbb{A}) \rangle$ its convex hull, $\mathcal{E}_l(\mathbb{A})$ the extreme points of $\langle \mu_l(\mathbb{A}) \rangle$, and $E_l(\mathbb{A})$ the families from $\mathbb{A}$ with an l -chain profile in $\mathcal{E}_l(\mathbb{A})$. Let, furthermore, $\mathcal{E}_l^*(\mathbb{A})$ denote the essential extreme points and $E_l^*(\mathbb{A})$ the corresponding families.

THEOREM 1. *For any upward or downward closed set of families $\mathbb{A} \subseteq 2^{2^X}$ and for any $l \geq 1$*

$$\mathcal{E}_l^*(\mathbb{A}) \subseteq \mu_l(E_1^*(\mathbb{A})).$$

Note that equality does not always hold, as the set of intersecting families, the family $\mathcal{F} = \{F \subseteq X : |F| > |X|/2\}$, and any $l > |X|/2$ show.

Proof. The proof is the same for downward and upward closed sets of families, so we assume that $\mathbb{A}$ is upward closed.

Let $E_1^*(\mathbb{A}) = \{\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_m\}$, and let f^i be the profile of $\mathcal{F}_i$, $f^{i,l}$ be the l -chain profile of $\mathcal{F}_i$, and $f_\alpha^{i,l}$ be its α th component.

We have to prove that the l -chain profile f^l of any family $\mathcal{F}$ in $\mathbb{A}$ can be dominated by a convex combination of the $f^{i,l}$'s. Denote the profile of $\mathcal{F}$ by f . Clearly we have

$$f_\alpha^l \leq f_{\alpha_1} \binom{n - \alpha_1}{\alpha^*},$$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)$, $\alpha^* = (\alpha_2 - \alpha_1, \alpha_3 - \alpha_1, \dots, \alpha_l - \alpha_1)$. Inequality holds with equality for the $f_\alpha^{i,l}$'s and the $f_\alpha^{i,l}$'s. The fact that the f^i 's are the essential extreme points of $\langle \mu_1(\mathbb{A}) \rangle$ means that for some convex combination $c_i, i = 1, \dots, m$,

$$f \leq \sum_{i=1}^m c_i f^i.$$

But then

$$f_\alpha^l \leq f_{\alpha_1} \binom{n - \alpha_1}{\alpha^*} \leq \binom{n - \alpha_1}{\alpha^*} \sum_{i=1}^m c_i f_{\alpha_1}^i = \sum_{i=1}^m c_i f_\alpha^{i,l},$$

which completes the proof. $\square$

Since the convex hull of the profile polytope of the set of intersecting families was determined by Erdős, Frankl, and Katona in [5], Theorem 1 provides the essential extreme points of the convex hull of the l -chain profile polytopes.

Definition. For any family $\mathcal{F}$ on a base set X let $\text{conv}(\mathcal{F}) = \{G \subseteq X : \exists F, F' \in \mathcal{F} (F \subseteq G \subseteq F')\}$ denote its *convex closure*. $\mathcal{F}$ is said to be *convex* if $\mathcal{F} = \text{conv}(\mathcal{F})$.

Definition. A set of families $\mathbb{A}$ is said to be *convex closed* if $\mathcal{F} \in \mathbb{A}$ implies $\text{conv}(\mathcal{F}) \in \mathbb{A}$.

Example. The basic example for a convex closed set is the set of intersecting and counterintersecting families.

THEOREM 2. *For any convex closed set of families $\mathbb{A} \subseteq 2^{2^X}$ and for any $l \geq 2$*

$$\mathcal{E}_l^*(\mathbb{A}) \subseteq \mu_l(E_2^*(\mathbb{A})).$$

Proof. The proof is analogous to that of Theorem 1. The inequality needed is

$$f_\alpha^l \leq f_{\alpha_1, \alpha_l}^2 \binom{\alpha_l - \alpha_1}{\alpha^*},$$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)$, $\alpha^* = (\alpha_2 - \alpha_1, \alpha_3 - \alpha_1, \dots, \alpha_{l-1} - \alpha_1)$, and for families with an essential extreme profile, the above inequality holds with equality. $\square$

Unfortunately the extreme points of neither the 1-chain nor the 2-chain profile polytope are known for the set of intersecting and counterintersecting families.

3. The reduction method. In this section we describe our main tool in determining the l -chain profile polytope of families of sets with some given property. We call this tool the reduction method. In fact, this is not a new method. Most of the proofs of results already obtained proceeded similarly. What we observed was that the method works for the l -chain case as well, and—what seems more important—in some cases it is enough to reduce the original problem to a maximal chain or to a pair of complementary chains (for the precise definitions, see below) instead of to the more complicated structures that previous proofs (see, for example, [4], [5]) used. We have to mention that Sali in [10] uses arguments involving reduction to a maximal chain, but we are not aware of any proofs with reduction to a pair of complementary chains.

Definition. For any l let $T_{\mathbf{C}}^l$ denote the following operator acting on the $\binom{n+1}{l}$ -dimensional $\mathbf{R}$ -space whose coordinates are indexed by l -tuples of integers $(\alpha_1, \alpha_2, \dots, \alpha_l)$, where $0 \leq \alpha_1 < \alpha_2 < \dots < \alpha_l \leq n$:

$$T_{\mathbf{C}}^l : e \mapsto T_{\mathbf{C}}^l(e), \quad \text{where} \quad T_{\mathbf{C}}^l(e)_{\alpha} = \binom{n}{\alpha} e_{\alpha}.$$

Definition. For a family $\mathcal{F}$ on a base set X and a maximal chain $\mathbf{C}$ in X let $\mathcal{F}(\mathbf{C}) = \{F \in \mathcal{F} \cap \mathbf{C}\}$, and for a set of families $\mathbb{A}$ let $\mathbb{A}(\mathbf{C}) = \{\mathcal{F}(\mathbf{C}) : \mathcal{F} \in \mathbb{A}\}$.

THEOREM 3. *For any set of families $\mathbb{A} \subseteq 2^{2^X}$ if the extreme points $e_1, e_2, \dots, e_m$ of $\langle \mu_l(\mathbb{A}(\mathbf{C})) \rangle$ do not depend on the choice of $\mathbf{C}$, then*

$$\langle \mu_l(\mathbb{A}) \rangle \subseteq \langle \{T_{\mathbf{C}}^l(e_1), \dots, T_{\mathbf{C}}^l(e_m)\} \rangle.$$

Proof. The modification of the argument in [5] works. Let $\mathcal{F}$ be an element of $\mathbb{A}$ with an l -profile $f = (\dots, f_{\alpha}, \dots)$. For $\mathbf{F} = \{F_1 \subset F_2 \subset \dots \subset F_l\}$ with $|F_i| = \alpha_i, i = 1, \dots, l$, let $\underline{w}(\mathbf{F})$ be the vector of length $\binom{n+1}{l}$ with $1/n!$ in the α th component and 0 everywhere else (where n is the size of the base set). Consider the sum $\sum \underline{w}(\mathbf{F})$ for all pairs $(\mathbf{C}, \mathbf{F})$, where $\mathbf{C}$ is a maximal chain on X and $\mathbf{F} \subset \mathcal{F} \cap \mathbf{C}$ is an l -chain. For a fixed $\mathbf{C}$ we have

$$\sum_{\mathbf{F} \in \mathcal{F}(\mathbf{C})} \underline{w}(\mathbf{F}) = \frac{1}{n!} (\text{profile of } \mathcal{F}(\mathbf{C})).$$

Here the profile of $\mathcal{F}(\mathbf{C})$ is a convex linear combination $\sum_{i=1}^m \lambda_i(\mathbf{C}) e_i$ of the e_i 's. Therefore,

$$(1) \quad \sum_{\mathbf{C}, \mathbf{F}} \underline{w}(\mathbf{F}) = \sum_{\mathbf{C}} \sum_{\mathbf{F}} \underline{w}(\mathbf{F}) = \sum_{\mathbf{C}} \frac{1}{n!} \sum_{i=1}^m \lambda_i(\mathbf{C}) e_i = \sum_{i=1}^m \frac{1}{n!} \left(\sum_{\mathbf{C}} \lambda_i(\mathbf{C}) \right) e_i$$

holds where $\sum_{\mathbf{C}} \frac{1}{n!} \sum_{i=1}^m \lambda_i(\mathbf{C}) = 1$. Thus $\sum \underline{w}(\mathbf{F})$ is a convex linear combination of the e_i 's.

Summing in the other way around, we have

$$(2) \quad \begin{aligned} \sum_{\mathbf{C}, \mathbf{F}} \underline{w}(\mathbf{F}) &= \sum_{\mathbf{F}} \sum_{\mathbf{C}} \underline{w}(\mathbf{F}) \\ &= \sum_{\mathbf{F}} \left(0, 0, \dots, \frac{|F_1|! (|F_2| - |F_1|)! \dots (|F_l| - |F_{l-1}|)! (n - |F_l|)!}{n!}, \dots, 0 \right) \\ &= \left(\dots, \frac{f_{\alpha}}{\binom{n}{\alpha}}, \dots \right), \end{aligned}$$

since for a fixed $\mathbf{F} = \{F_1 \subset F_2 \subset \dots \subset F_l\}$ there are exactly $|F_1|!(|F_2|-|F_1|)!\dots(|F_l|-|F_{l-1}|)!(n-|F_l|)!$ chains containing $\mathbf{F}$. So (1) and (2) give that this last vector is a convex linear combination of the e_i 's, which implies that f is the linear combination of $T_{\mathbf{C}}^l(e_1), \dots, T_{\mathbf{C}}^l(e_m)$. $\square$

The structure of maximal chains is too simple, so using only them is not enough to determine the l -chain profile polytope of more difficult sets of families. But the proof of Theorem 3 works if we replace the chain by a pair of complementary maximal chains (i.e., for $i = 1, 2$ $\mathbf{C}^i = \{C_0^i, C_1^i, \dots, C_n^i\}$ with $C_j^i = X \setminus C_{n-j}^{3-i} = \overline{C}_{n-j}^{3-i}$ for all $j = 0, 1, \dots, n$). In the proof one has to write (instead of $\frac{1}{n!}$) $\frac{2}{(n!)}$ in the definition of $w(\mathbf{F})$ and modify the definition of the T -operator to

$$(T_{\mathbf{C}_1, \mathbf{C}_2}^l(e))_{\alpha} = \frac{1}{d_{\alpha}} \binom{n}{\alpha},$$

where d_{α} is the number of α -type l -chains in the pair of complementing chains.

4. Applications. In this section we determine the profile polytope of some sets of families using the reduction method. In the first subsection the problem will be reduced to the case of the maximal chain, while in the second subsection we will consider reduction to a pair of complementary chains. Using the results obtained by the latter we will give examples of when the extreme families of the l -profile polytope can really depend on l .

4.1. Reduction to the chain.

THEOREM 4. *For all $l \geq 1$ the extreme points of the convex hull of the l -chain profile vectors of convex families are the following: the all zero vector,*

$$\mathbf{0} = (0, \dots, 0),$$

and for all $0 \leq i \leq j \leq n$ the vectors $v_{i,j}$,

$$(3) \quad (v_{i,j})_{\alpha} = \begin{cases} \binom{n}{\alpha} & \text{if } i \leq \alpha_1 < \alpha_l \leq j, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. The vector $v_{i,j}$ is the l -profile of the family $\mathcal{F}_{i,j} = \{F \subseteq [n] : i \leq |F| \leq j\}$, which is convex.

On a chain any convex family must consist of some consecutive subsets of the chain. The theorem follows now from Theorem 3. $\square$

Note that the set of convex families is not hereditary; therefore, the extreme points (for the original profile vectors) need not be the ones obtained from the essential extreme points (in this case there is only one, and that is the profile of $2^{[n]}$) by changing some of the nonzero components to zero, and, as Theorem 4 shows, they are not those vectors, indeed.

THEOREM 5. *For any $l \leq k$ the extreme points of the l -chain profile polytope of k -Sperner families are the following: the all zero vector,*

$$\mathbf{0} = (0, \dots, 0, \dots, 0),$$

and for all $l \leq z \leq k$ and $\beta = \{\beta_1, \dots, \beta_z\}$ with $0 \leq \beta_1 < \dots < \beta_z \leq n$, the vectors v_{β} ,

$$(4) \quad (v_{\beta})_{\alpha} = \begin{cases} \binom{n}{\alpha} & \text{if } \alpha \subseteq \beta, \\ 0 & \text{otherwise.} \end{cases}$$

The case $l = 1$ is a result of Erdős, Frankl, and Katona [5].

Proof. It is trivial to see that these vectors are l -chain profiles of the corresponding levels, and they are convex linearly independent.

A k -Sperner family on a maximal chain consists of at most k sets; therefore, its l -chain profile vector has 1's in those components $\alpha = (\alpha_1, \dots, \alpha_l)$ for which there is an element in the family with size α_i for all $i = 1, \dots, l$. All these vectors are convex independent. Therefore, they form the convex hull of the profile polytope on the chain, and Theorem 3 now implies Theorem 5. $\square$

Applying Theorem 5 for the constant 1 weight function, one gets the following corollary.

COROLLARY. *For any $l \leq k$ if a family $\mathcal{F}$ on an n -element base set X does not contain a chain of length $k + 1$, then the number of l -chains in $\mathcal{F}$ is at most*

$$\max_{\beta \subset [0, n]; |\beta| = k} \sum_{\alpha \subseteq \beta; |\alpha| = l} \binom{n}{\alpha}.$$

Remarks. In the case $l = k$, even the very simple argument of [7] works. First we need an LYM-type inequality. To get this we double-count the pairs $(\mathbf{C}, \mathbf{F})$, where $\mathbf{C}$ is a maximal chain and $\mathbf{F}$ is an l -chain contained in $\mathbf{C}$. If we decompose the k -Sperner family into k antichains, then all sets of an $\mathbf{F}$ come from different antichains, and any $\mathbf{C}$ can contain at most k sets from our family, so by a standard calculation we obtain

$$(5) \quad \sum_{\alpha} \frac{f_{\alpha}}{\binom{n}{\alpha}} \leq \binom{k}{l}.$$

If $l = k$, then the right-hand side is 1, and we can finish the proof as follows:

$$(6) \quad \sum_{\alpha} w_{\alpha} f_{\alpha} = \sum_{\alpha} \frac{f_{\alpha} w_{\alpha}}{\binom{n}{\alpha}} \binom{n}{\alpha} \leq \max_{\alpha} \left\{ w_{\alpha} \binom{n}{\alpha} \right\} \sum_{\alpha} \frac{f_{\alpha}}{\binom{n}{\alpha}} \leq \max_{\alpha} \left\{ w_{\alpha} \binom{n}{\alpha} \right\},$$

where w_{α} is any nonnegative weight function and the last inequality in (6) uses (5).

If $l = k$, then (5) gives the maximum number of k -chains that a k -Sperner family can contain. This is $\binom{n}{\alpha}$, where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$ and the numbers $\alpha_1, \alpha_2 - \alpha_1, \dots, \alpha_k - \alpha_{k-1}$ differ by at most one. If $k + 1$ divides n , then we get the uniqueness of the extremal system (take all $F \subseteq X$ with $|F| = \alpha_i$ for some $i = 1, \dots, k$) automatically. If $k + 1$ does not divide n , then we can lift (5) up to an AZ-type identity (for more details see [9], and for the original AZ-identity see the paper of Ahlswede and Zhang [1]), which will ensure the uniqueness.

With the notation of section 3, Theorem 5 implies (if $\mathbb{S}_k$ denotes the set of k -Sperner families) $E_1(\mathbb{S}_k) = E_l(\mathbb{S}_k)$. But the bordering faces of the convex hulls $\langle \mu_1(\mathbb{S}_k) \rangle$ and $\langle \mu_l(\mathbb{S}_k) \rangle$ are “not the same.” If $l = 1$, the convex hull is determined by the faces given by the inequalities $0 \leq f_i \leq \binom{n}{i}$ and the LYM-inequality $\sum_i f_i / \binom{n}{i} \leq k$ (see [5]), while if $l > 1$, the hyperplanes given by $0 \leq f_{\alpha} \leq \binom{n}{\alpha}$ and the LYM-type inequality of (5) are bordering faces, but there are some additional bordering faces, which can be seen by the following observation. Choosing $\binom{k}{l}$ α 's in such a way that their union has size strictly larger than k and putting $f_{\alpha} = \binom{k}{\alpha}$ for these α 's and 0 for the others, we obtain an essential extreme point of the polytope determined by the above inequalities which is not an l -chain profile of any k -Sperner families.

In [8] the second author defined the distance of $\mathcal{F}$ -free families. For two $\mathcal{F}$ -free families $\mathcal{H}_1$ and $\mathcal{H}_2$ this is the number of copies of $\mathcal{F}$ in $\mathcal{H}_1 \cup \mathcal{H}_2$ (the edge (vertex) set of $\mathcal{H}_1 \cup \mathcal{H}_2$ is the union of the edge (vertex) sets of $\mathcal{H}_1$ and $\mathcal{H}_2$). For a collection

$\mathbb{C}$ of forbidden families, the $\mathbb{C}$ -free distance of two $\mathbb{C}$ -free families is the sum of $\mathcal{F}$ -free distances taken over all $\mathcal{F} \in \mathbb{C}$. If in the collection $\mathbb{C}$ we enumerate all kinds of $k+1$ -chains, then the $\mathbb{C}$ -free families are the k -Sperner systems. The union of two such systems is a $2k$ -Sperner family, and of course all $2k$ -Sperner families can be decomposed into two k -Sperner systems. So the maximum distance that two k -Sperner families can have is the maximum number of $k+1$ -chains that a $2k$ -Sperner family can contain, which by the corollary of Theorem 5 is

$$\max_{\beta \subset [0, n]; |\beta|=2k} \sum_{\alpha \subseteq \beta; |\alpha|=k+1} \binom{n}{\alpha}.$$

4.2. Reduction to a pair of complement chains.

THEOREM 6. *Let $n = 2m + 1$ and $k \leq m + 1$. Then the extreme points of the 1-chain profile polytope (i.e., the ordinary profile polytope) of the set of complement-free k -Sperner families are the following vectors (indexed with a z -element ($z \leq k$) subset α of $\{0, 1, 2, \dots, n\}$, where $\alpha_i \in \alpha$ implies $n - \alpha_i \notin \alpha$):*

$$v_\alpha = \left(0, \dots, 0, \binom{n}{\alpha_1}, 0, \dots, 0, \binom{n}{\alpha_2}, 0, \dots, 0, \dots, 0, \binom{n}{\alpha_z}, 0, \dots, 0 \right).$$

Proof. By the version of Theorem 3 just mentioned in the first paragraph of this section, it is enough to prove the following lemma.

LEMMA 1. *If $n = 2m + 1$ and $k \leq m + 1$, then the extreme points of the profile polytope of complement-free k -Sperner families on a pair of maximal complement chains are the vectors with at most k nonzero components, where all the nonzero components are 2 (except for the first or the last component; if one of them is nonzero, it equals 1), and if the i th component is nonzero, then the $(n - i)$ th component is zero.*

Proof of Lemma 1. If the nonzero components of such a vector are $\alpha_1, \alpha_2, \dots, \alpha_z$ (satisfying the condition of the lemma), then the sets in the two chains with cardinality α_i for some $i = 1, \dots, z$ form a complement-free k -Sperner family with the vector as profile.

Now let $\mathcal{F}$ be a complement-free k -Sperner family on a pair of complement chains $\mathbf{C}_1, \mathbf{C}_2$ with profile vector f . Let α be the set of indices of the nonzero components of f . Partition α into three subsets. Let CL (complete levels) denote the indices α_i with $f_{\alpha_i} = 2$, and, furthermore, CL contains 0 (n) provided $f_0 = 1$ ($f_n = 1$). Let, furthermore, CP (complementing pairs) denote the indices $\alpha_i \in \alpha$ with $n - \alpha_i \in \alpha$, and let $R = \alpha \setminus (CL \cup CP)$. Note that $CP \cap CL = \emptyset$, for otherwise $\mathcal{F}$ would not be complement-free. Now form two subsets α^1, α^2 of α in the following way. Put all indices in CL into both α^1 and α^2 . For all pairs of indices $i, n - i$ in CP (note that these are really pairs, for n is odd) put one of the indices into α^1 and the other into α^2 . Finally, choose α^1 or α^2 for all indices of R in such a way that $|\alpha^1| \leq k$ and $|\alpha^2| \leq k$ hold. (This is possible, for $\mathcal{F}$ is k -Sperner; therefore, $|\alpha| \leq 2k$.) Now let f^i , $i = 1, 2$, be the following vectors:

$$(7) \quad f_j^i = \begin{cases} 2 & (1) \text{ if } j \neq 0, n \text{ } (j = 0, n), \text{ } j \in \alpha^i, \\ 0 & \text{otherwise.} \end{cases}$$

By the facts that both f^i 's are of the form of the statement of the lemma and $f = \frac{1}{2}f^1 + \frac{1}{2}f^2$, furthermore f^1 and f^2 are both profile vectors of k -Sperner families on $\mathbf{C}_1, \mathbf{C}_2$, and the proof is completed. $\square$

The case of complement-free families is analogous (and even simpler); therefore, we just sketch the proof.

THEOREM 7. *The extreme points of the convex hull of the 1-chain profile vectors of complement-free families are the vectors corresponding to the families consisting of*

- (i) *a set I of levels with the property that the i th and the $(n - i)$ th levels cannot be both in I if n is odd, and*
- (ii) *a set I of levels with the property that the i th and the $(n - i)$ th levels cannot be both in I and possibly half of the sets with size $n/2$, one from each pair of complementary sets, if n is even.*

Proof. It is easy to see that it is enough to solve the problem reduced to a pair of maximal complement chains. There the statement holds, since there a complement-free family can contain at most two sets out of the four with size i or $n - i$, and the vectors $(1, 1), (0, 1), (1, 0)$ are convex combinations of the vectors $(2, 0), (0, 2), (0, 0)$. $\square$

Theorems 1 and 2 state that for a certain class of sets of families all candidates for the families with essential extreme l -chain profiles are among the families with essential extreme 1-chain (2-chain) profiles. Theorem 5 states that for k -Sperner families the above statement is true for all extreme profiles (not only for essential extreme profiles). It seems natural to conjecture (with the notation of section 2) that, for all sets of families $\mathbb{A}$ and $l > 1$, $E_l(\mathbb{A}) \subseteq E_1(\mathbb{A})$ and/or $E_l^*(\mathbb{A}) \subseteq E_1^*(\mathbb{A})$. But this is false. Here we present two counterexamples.

The first example is based on Theorem 6. Note that the families corresponding to the extreme points cannot contain sets of size i and $n - i$ at the same time. Hence all 2-chain profiles of those families have 0 in their components indexed with the sets $\{i, n - i\}$, and therefore all their convex combinations have 0 in those components. But a pair of subsets in inclusion with size i and $n - i$ is of course a complement-free k -Sperner family (if $k \geq 2$), and its profile is not in the convex hull of the above-mentioned vectors.

The second example is absolutely analogous to the first one. According to Theorem 7 in the extremal families of the set of complement-free systems there are no pairs of sets in inclusion with size i and $n - i$ (so the corresponding component is 0 in any convex combinations), but there are complement-free families with such pairs.

Acknowledgment. We would like to thank the two anonymous referees for helpful comments on the paper.

REFERENCES

- [1] R. AHLWEDE AND Z. ZHANG, *An identity in combinatorial extremal theory*, Adv. Math., 80 (1990), pp. 137–151.
- [2] B. BOLLOBÁS, *On complete subgraphs of different order*, Math. Proc. Cambridge Philos. Soc., 80 (1976), pp. 19–24.
- [3] K. ENGEL, *Sperner Theory*, Encyclopedia Math. Appl. 65, Cambridge University Press, Cambridge, UK, 1997.
- [4] P. L. ERDŐS, P. FRANKL, AND G. O. H. KATONA, *Intersecting Sperner families and their convex hulls*, Combinatorica, 4 (1984), pp. 21–34.
- [5] P. L. ERDŐS, P. FRANKL, AND G. O. H. KATONA, *Extremal hypergraph problems and convex hulls*, Combinatorica, 5 (1985), pp. 11–26.
- [6] P. FRANKL AND G. O. H. KATONA, *Polytopes determined by hypergraph classes*, European J. Combin., 6 (1985), pp. 233–243.
- [7] G. O. H. KATONA, *Two applications (for search theory and truth functions) of Sperner type theorems*, Period. Math. Hungar., 3 (1973), pp. 19–26.
- [8] B. PATKÓS, *The distance of $\mathcal{F}$ -free families*, Studia Sci. Math. Hungar., to appear.
- [9] B. PATKÓS, *AZ-type Identities*, <http://www.ceu.hu/math/Preprints/Patkos3.pdf>.

- [10] A. SALI, *A note on convex hulls of more-part Sperner families*, J. Combin. Theory Ser. A, 49 (1988), pp. 188–190.
- [11] N. SAUER, *A Generalization of a Theorem of Turán*, Research paper 61, Department of Mathematics, University of Calgary, Calgary, Alberta, Canada, 1968.
- [12] P. TURÁN, *An extremal problem in graph theory*, Mat. Fiz. Lapok, 48 (1941), pp. 436–452 (in Hungarian).

PLANARITY, COLORABILITY, AND MINOR GAMES*

DAN HEFETZ[†], MICHAEL KRIVELEVICH[‡], MILOŠ STOJAKOVIĆ[§], AND TIBOR SZABÓ[¶]

Abstract. Let m and b be positive integers, and let $\mathcal{F}$ be a hypergraph. In an (m, b) Maker-Breaker game $\mathcal{F}$ two players, called Maker and Breaker, take turns selecting previously unclaimed vertices of $\mathcal{F}$. Maker selects m vertices per move, and Breaker selects b vertices per move. The game ends when every vertex has been claimed by one of the players. Maker wins if he claims all of the vertices of some hyperedge of $\mathcal{F}$; otherwise Breaker wins. An (m, b) Avoider-Enforcer game $\mathcal{F}$ is played in a similar way. The only difference is in the determination of the winner: Avoider loses if he claims all of the vertices of some hyperedge of $\mathcal{F}$; otherwise Enforcer loses. In this paper we consider the Maker-Breaker and Avoider-Enforcer versions of the planarity game, the k -colorability game, and the K_t -minor game.

Key words. combinatorial games, planar graph, graph coloring, graph minors

AMS subject classifications. 91A46, 05C10, 05C15, 05C83

DOI. 10.1137/060654414

1. Introduction. Let m and b be two positive integers. We are given a set X and a hypergraph $\mathcal{F} \subseteq 2^X$. During the (m, b) positional game $\mathcal{F}$, two players take turns claiming previously unclaimed elements of X . In every round, the first player claims m elements, and then the second player claims b elements. The set X is called the “board”; m and b are the biases of the first and second players, respectively. For the purposes of this paper $\mathcal{F}$ is assumed to be monotone-increasing. We investigate positional games with two different kinds of rules for determining the winner.

In a *Maker-Breaker-type* positional game, the two players are called Maker and Breaker, and $\mathcal{F}$ is referred to as the family of “winning sets.” Maker wins the game if the set M he has claimed by the end of the game (i.e., when every element of the board has been claimed by one of the players) is a winning set, that is, $M \in \mathcal{F}$. Otherwise Breaker wins. Observe that, since $\mathcal{F}$ is assumed to be monotone-increasing, the game could essentially be stopped as soon as Maker occupied a *minimal* (with respect to inclusion) winning set $F \in \mathcal{F}$; the position on $X \setminus F$ can no longer influence the outcome of the game. Hence, Breaker wins if and only if he claims at least one element in every minimal winning set. Since a monotone-increasing family and the family of its minimal elements uniquely determine each other, often, when there is no risk of confusion, we use $\mathcal{F}$ for the family of minimal winning sets as well. A classical example of the Maker-Breaker setting is the popular board game HEX.

*Received by the editors March 16, 2006; accepted for publication (in revised form) August 31, 2007; published electronically February 20, 2008.

<http://www.siam.org/journals/sidma/22-1/65441.html>

[†]School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel (dannyh@post.tau.ac.il). This paper is a part of the author’s Ph.D. thesis, written under the supervision of Prof. Michael Krivelevich.

[‡]Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel (krivelev@post.tau.ac.il). The research of this author was supported in part by a U.S.–Israeli BSF grant and a grant from the Israeli Science Foundation.

[§]Department of Mathematics and Informatics, University of Novi Sad, Novi Sad, Serbia and Montenegro (smilos@inf.ethz.ch). The research of this author was partly supported by the Ministry of Science and Environmental Protection, Republic of Serbia, and Provincial secretariat for science, Province of Vojvodina.

[¶]Institute of Theoretical Computer Science, ETH Zurich, CH-8092 Zurich, Switzerland (szabo@inf.ethz.ch).

In an *Avoider-Enforcer-type* positional game, the players are called Avoider and Enforcer, and $\mathcal{F}$ is called the family of “losing sets.” Avoider wins the game if the set A he has claimed by the end of the game (i.e., when every element of the board has been claimed by one of the players) is not a losing set, that is, $A \notin \mathcal{F}$. Otherwise Enforcer wins. Again, observe that the game could be stopped as soon as Avoider fully occupies a minimal losing set, since then Enforcer surely wins no matter how the game proceeds. Having this in mind, we will sometimes use $\mathcal{F}$ to denote the family of minimal losing sets.

The identity of the player making the first move does not affect the asymptotics of our results. However, for convenience, we will assume that Maker (Avoider) is the first player in all Maker-Breaker (Avoider-Enforcer) games we discuss.

In this paper, we investigate Maker-Breaker and Avoider-Enforcer positional games played on the board $E(K_n)$ —the edges of the complete graph on n vertices.

Maker-breaker games. The study of positional games on the edges of a (complete) graph was initiated by Lehman [20], who, in particular, proved that in the $(1, 1)$ game Maker can easily build a spanning tree (by “easily” we mean that he can do so within $n - 1$ moves). Chvátal and Erdős [10] suggested to “even out the odds” by giving Breaker more power, that is, by increasing his bias. They determined that the $(1, b)$ game $\mathcal{T}_n$ of “connectivity,” where the family $\mathcal{T}_n$ of minimal winning sets consists of the edge sets of all spanning trees of K_n , is won by Maker even when the bias b of Breaker is as large as $cn/\log n$ for some small constant $c > 0$, while for some large enough constant $C > 0$, Breaker wins if his bias is at least $Cn/\log n$. They also showed that the $(1, 1)$ game $\mathcal{H}_n$ of “Hamiltonicity,” where Maker’s goal is to build a Hamiltonian cycle (the family $\mathcal{H}_n$ of minimal winning sets consists of the edge sets of all Hamiltonian cycles of K_n), is won by Maker for sufficiently large n . They conjectured that in fact Maker can win the $(1, b)$ Hamiltonicity game for some b that tends to infinity with n . This was proved by Bollobás and Papaioannou [9], who showed that Maker wins Hamiltonicity against a bias of $\Omega(\log n / \log \log n)$. Finally, Beck [1] gave a winning strategy for Maker against a bias of $\Omega(n / \log n)$.

Following [14], it will be convenient to introduce the following notation. For a family $\mathcal{F}$ of winning sets, let $b_{\mathcal{F}}$ be the nonnegative integer for which Breaker has a winning strategy in the $(1, b)$ game $\mathcal{F}$ if and only if $b \geq b_{\mathcal{F}}$. Note that $b_{\mathcal{F}}$ is well-defined for any (monotone-increasing) family $\mathcal{F}$ (unless $\mathcal{F}$ contains a hyperedge of size at most one). By the above, $b_{\mathcal{T}_n} = \Theta(n / \log n)$ and $b_{\mathcal{H}_n} = \Theta(n / \log n)$.

There is an intriguing relation between the threshold biases of the connectivity and Hamiltonicity games and the threshold probability at which the random graph $G(n, p)$ first possesses these properties. The number of edges of the random graph $G(n, p)$ around this threshold $p = \log n / n$ has the same order of magnitude as the number of edges Maker has in his graph after playing against the threshold bias $b_{\mathcal{T}_n}$ or $b_{\mathcal{H}_n}$. This phenomenon (known as the Erdős paradigm) was pointed out by Beck [1], where this observation is attributed to Erdős (indeed the first such result appeared already in [10]). Since then, several other games have been found to have a threshold bias which is closely linked to a meaningful random graph threshold related to that particular game (see, e.g., [2, 4, 5, 6, 25]). Note that the asymptotic values of the threshold biases $b_{\mathcal{T}_n}$ and $b_{\mathcal{H}_n}$ are not known in general. It can still turn out that these threshold biases are asymptotically equal to the inverse of the corresponding sharp thresholds for the appropriate properties of random graphs.

In this paper we investigate this connection further and find three more instances where such an intuition proves to be correct. Let $\mathcal{N}_{\mathcal{P}_n}$, $\mathcal{N}\mathcal{C}_n^k$, $\mathcal{M}_n^t$ consist of the

edge sets of all nonplanar graphs on n vertices, the edge sets of all non- k -colorable graphs on n vertices, and the edge sets of all graphs on n vertices containing a K_t -minor (a graph G contains an H -minor if H can be obtained from G by the deletion and contraction of some of its edges), respectively. Obviously, all three families are monotone-increasing.

In the game $\mathcal{NP}_n$, which we call the “planarity game” (although perhaps “non-planarity” would have been a more appropriate name), Maker’s goal is to claim all edges of a nonplanar graph. In Theorem 2.1 we show that the corresponding threshold bias $b_{\mathcal{NP}_n}$ is asymptotically $n/2$. Note that this result is a consequence of the result for the minor game we obtain in Theorem 4.1, as the presence of a K_5 -minor guarantees nonplanarity. Nevertheless, we include an alternative, direct proof of Theorem 2.1 which approaches planarity from a different angle and, in our opinion, is more instructive and serves as a good example of how Maker can win by adopting the role of Avoider in a different game.

Coming back to the relation between the probability thresholds of random graphs and the thresholds for the game bias, we note that the threshold probability $p_{\mathcal{NP}_n}$ of nonplanarity in the random graph is $1/n$, that is, $p_{\mathcal{NP}_n} = \Theta(1/b_{\mathcal{NP}_n})$. However, the number of edges of $G(n, p_{\mathcal{NP}_n})$ is concentrated around *one-half* of the number of edges in Maker’s graph in the $(1, b_{\mathcal{NP}_n}, \mathcal{NP}_n)$ game (the nonplanarity threshold is sharp, so such a comparison makes sense).

In the “ k -colorability game,” the family of winning sets is $\mathcal{NC}_n^k$; that is, Maker wins the game if he claims a non- k -colorable graph. In Theorem 3.1 we show that the threshold bias is linear for a fixed k by establishing that there are absolute constants c_1, c_2 such that $c_1 \frac{n}{k \log k} \leq b_{\mathcal{NC}_n^k} \leq c_2 \frac{n}{k \log k}$ for any k . For the special case $k = 2$, that is, the bipartite game, a more accurate result was proved by Bednarska and Pikhurko [8]. They showed that $(1 - 1/\sqrt{2} - o(1))n \leq b_{\mathcal{NC}_n^2} \leq \lceil n/2 \rceil - 1$. Again, one can compare these results with the corresponding random graph threshold $p_{\mathcal{NC}_n^k} = \frac{2k \log k}{n}$ for non- k -colorability and find the reciprocal relation.

Finally, we turn to a more general setting. In the K_t -minor game the family of winning sets is $\mathcal{M}_n^t$; hence, Maker wins the game if he is able to claim the edges of some K_t -minor of K_n . This game is in some sense a generalization of both the planarity and the colorability games we have discussed. Indeed, Wagner’s theorem gives a full description of planarity via the language of forbidden minors. Furthermore, if a graph is not r -colorable, then it contains a K_s -minor for $s = r/(c\sqrt{\log r})$, where c is an absolute constant (see [17, 26]), and the famous and long-standing Hadwiger conjecture asserts that in fact it contains a K_r -minor. The opposite implication is trivially false, as, for every positive integer n , the complete bipartite graph $K_{n,n}$ admits a K_n -minor.

In Theorem 4.1 we prove that the corresponding threshold bias $b_{\mathcal{M}_n^t}$ is asymptotically $n/2$ for every $3 \leq t \leq c\sqrt{n/\log n}$, for an appropriate constant c . Note that for the case $t = 3$ an accurate threshold bias follows from a result of Bednarska and Pikhurko [7], as a graph is K_3 -minor free if and only if it is a forest (see the exact statement in Theorem 2.2 in section 2).

Avoider-enforcer games. At first sight, the definition of Avoider-Enforcer games seems less natural than that of Maker-Breaker games, and accordingly the theory is much less developed. We, however, argue that, besides being interesting in their own right (see [14] for a comprehensive discussion), they are essential in studying Maker-Breaker games. Avoider-enforcer games arise naturally whenever one would like to play (what looks like) Maker-Breaker games on a *monotone-decreasing family*. For

example, the set $\mathcal{P}_n$ of planar graphs is a monotone-decreasing family. If, for example, Maker's goal is to keep his graph planar to the end of the game, then he can be thought of as Avoider, playing an Avoider-Enforcer game on the (monotone-increasing) family of losing sets $2^{E(K_n)} \setminus \mathcal{P}_n = \mathcal{NP}_n$.

Moreover, for certain Maker-Breaker games the best known maker strategies involve building a pseudorandom graph with certain parameters (see [13, 15]). It is proved that the particular pseudorandom properties of Maker's graph imply the graph-theoretic properties in question, entailing his win. Here pseudorandom properties include bounds on the number of edges between pairs of disjoint vertex sets from *below* and from *above*. Hence, in such a game the family $\mathcal{F}$ is the intersection of a monotone-increasing family and a monotone-decreasing family.

Avoider-enforcer games were studied in [3, 14, 21, 22, 23]. Similarly to Maker-Breaker games, one would like to define for each monotone-increasing family $\mathcal{F}$ the Avoider-Enforcer threshold bias $f_{\mathcal{F}}$. A reasonable choice for $f_{\mathcal{F}}$ would be the non-negative integer for which Avoider wins the $(1, b)$ game $\mathcal{F}$ if and only if $b \geq f_{\mathcal{F}}$. While the similar threshold $b_{\mathcal{F}}$ does exist for Maker-Breaker games on (essentially) any hypergraph, for Avoider-Enforcer games it generally does not (see [14]).

Following [14], it will be convenient to introduce the following notation. For a hypergraph $\mathcal{F}$ we define the *lower threshold bias* $f_{\mathcal{F}}^-$ to be the largest integer such that Enforcer can win the $(1, b)$ game $\mathcal{F}$ for every $b \leq f_{\mathcal{F}}^-$ and the *upper threshold bias* $f_{\mathcal{F}}^+$ to be the smallest nonnegative integer such that Avoider can win the $(1, b)$ game $\mathcal{F}$ for every $b > f_{\mathcal{F}}^+$. Except for certain degenerate cases, $f_{\mathcal{F}}^-$ and $f_{\mathcal{F}}^+$ always exist and satisfy $f_{\mathcal{F}}^- \leq f_{\mathcal{F}}^+$. Whenever $f_{\mathcal{F}}^- = f_{\mathcal{F}}^+$, the threshold bias $f_{\mathcal{F}}$ of the Avoider-Enforcer game $\mathcal{F}$ does exist and satisfies $f_{\mathcal{F}} = f_{\mathcal{F}}^+$.

In [14] the existence of the threshold $f_{\mathcal{T}_n}$ was established for the connectivity game, and it was proved that $f_{\mathcal{T}_n}$ is roughly $n/2$. For the perfect matching game and Hamiltonicity game it was shown that $f_{\mathcal{M}_n}^- = \Omega(n/\log n)$ and $f_{\mathcal{H}_n}^- = \Omega(n \log \log \log \log n / \log n \log \log \log n)$, respectively. Note that it is not known if the threshold bias for any of these natural games exists.

In this paper, we give bounds on the lower and upper threshold biases for the planarity, k -colorability, and K_t -minor games. Note that, unlike the games of connectivity and Hamiltonicity, these Avoider-Enforcer games seem to be more natural than their Maker-Breaker analogues.

In Theorem 2.3 we prove that Avoider can keep his graph planar against any bias which is larger than $2n^{5/4}$, whereas Enforcer wins when playing with a bias no larger than $\frac{n}{2} - o(n)$. As in the case of the Maker-Breaker planarity game, the second part of this result is a direct consequence of our result for the minor game presented in Theorem 4.7. Nevertheless, we include an alternative, direct proof which relies on other properties of planar graphs. We believe it is more illustrative and gives more insight into the course of the game.

In Theorem 3.4 we show that, by playing against a bias of at least $2kn^{1+1/(2k-3)}$, Avoider can keep his graph k -colorable, whereas Enforcer wins if his bias is at most $\frac{n}{ck \log k}$.

In the Avoider-Enforcer version of the K_t -minor game, Avoider's task is to build a K_t -minor free graph. As in the planarity and colorability games, Avoider's goal in this game is very natural. Indeed, many graph-theoretic properties can be expressed in a "forbidden minor" fashion.

We show that, by playing with a linear bias, Enforcer can make Avoider build a graph that admits a K_t -minor for t which is as large as $c\sqrt{n/\log n}$. Moreover, in

Theorem 4.7 we prove that, even when playing with a bias which is arbitrarily close (from below) to $n/2$, Enforcer can make sure that by the end of the game Avoider will claim a K_t -minor for $t = n^{a_0}$, where $a_0 > 0$ is an appropriate constant.

For the sake of simplicity and clarity of presentation, we do not make a particular effort to optimize the constants obtained in theorems we prove. We also omit floor and ceiling signs whenever these are not crucial. All of our results are asymptotic in nature, and whenever necessary we assume that n is sufficiently large. Throughout the paper, $\log$ stands for the natural logarithm. Our graph-theoretic notation is standard and follows that of [11].

The rest of the paper is organized as follows: In section 2 we discuss the Maker-Breaker and Avoider-Enforcer planarity games, in section 3 we discuss the colorability games, and in section 4 we discuss the minor games. Finally, in section 5 we present some open problems.

2. Planarity games.

2.1. The Maker-Breaker planarity game. The following theorem states that the threshold bias at which Maker's win turns into a Breaker's win in the planarity game is "around" $n/2$.

THEOREM 2.1.

$$b_{\mathcal{NP}_n} = \frac{n}{2} - o(n).$$

Proof. Let $b \geq n/2$. The existence of a winning strategy for Breaker in the planarity game is an easy consequence of the following result of Bednarska and Pikhurko.

THEOREM 2.2 (see [7, Corollary 10]). *Suppose that CycleMaker and CycleBreaker select, respectively, 1 and b edges of K_n and CycleMaker wins if he builds a cycle. Then CycleMaker has a winning strategy (no matter who starts) if and only if $b < \lceil n/2 \rceil$.*

The assertion of Theorem 2.2 implies that with the bias $b \geq n/2$, Breaker can prevent Maker from building a cycle. It follows that at the end of the game Maker's graph will be a forest which is obviously planar.

Next, let $0 < \varepsilon < 1/3$ (the restriction $\varepsilon < 1/3$ is technical), and let $b \leq (1/2 - \varepsilon)n$, where $n = n(\varepsilon)$ is sufficiently large. We will provide Maker with a strategy for building a nonplanar graph. Let $\alpha = \frac{2\varepsilon}{1-2\varepsilon}$, and let $\alpha_n = \alpha_n(\varepsilon)$ be the real number satisfying the equation

$$(1 + \alpha_n)n = \frac{\binom{n}{2}}{(\frac{1}{2} - \varepsilon)n + 1}.$$

Then $\lim_{n \rightarrow \infty} \alpha_n = \alpha$. Let m_n denote the number of edges that Maker will claim by the end of the game on K_n . We have $m_n - (1 + \frac{\alpha}{2})n = \Omega(n)$.

Let $k = k(\varepsilon)$ be the smallest positive integer such that

$$\left(1 + \frac{\alpha}{2}\right) > \frac{k}{k-2}.$$

Maker's goal is to avoid cycles of length smaller than k , which we will call "short cycles," during the first $(1 + \frac{\alpha}{2})n$ moves. If he succeeds, Maker's graph will at that point of the game have

$$\left(1 + \frac{\alpha}{2}\right)n > \frac{k}{k-2}n$$

edges and girth at least k . But, it is well known that a planar graph with girth at least k cannot have more than $\frac{k}{k-2}(n-2)$ edges. Hence, Maker's graph will already be nonplanar, and he will win no matter how the game continues.

It remains to show that Maker can indeed avoid claiming a short cycle during the first $(1 + \frac{\alpha}{2})n$ moves. His strategy is the following. For as long as possible he claims edges (u, v) that satisfy the following two properties:

- (a) (u, v) does not close a short cycle;
- (b) the degrees of both u and v in Maker's graph are less than $n^{1/(k+1)}$.

It suffices to prove that when this is no longer possible, that is, every remaining unclaimed edge violates either (a) or (b), Maker has already claimed at least $(1 + \frac{\alpha}{2})n$ edges.

Every edge that violates property (b) must have at least one end point of degree $n^{1/(k+1)}$ in Maker's graph. Since Maker's graph at any point of the game contains at most $(1 + \alpha)n$ edges, there are at most $2(1 + \alpha)n^{1-1/(k+1)}$ vertices of degree at least $n^{1/(k+1)}$. Therefore, the number of edges that violate property (b) is at most

$$n \cdot 2(1 + \alpha)n^{1-1/(k+1)} = o(n^2).$$

For any fixed $s < k$ and every vertex v , the number of paths of length s that have v as one end point is at most Δ^s , where Δ is the maximum degree in Maker's graph. If we assume that property (b) has not been violated, then $\Delta \leq n^{1/(k+1)}$. Therefore, there are at most

$$n \cdot \sum_{s=3}^{k-1} n^{s/(k+1)} = o(n^2)$$

edges that close a short cycle.

Thus, the total number of edges that violate (a) or (b) if claimed by Maker is $o(n^2)$. On the other hand, after $(1 + \frac{\alpha}{2})n$ moves have been played, the number of unclaimed edges is $\Theta(n^2)$. Hence, in the first $(1 + \frac{\alpha}{2})n$ moves Maker can claim edges that satisfy properties (a) and (b), which means that he does not claim a short cycle. This completes the proof of the theorem. $\square$

2.2. The Avoider-Enforcer planarity game. In the following theorem we give an upper bound and a lower bound for the threshold bias at which Enforcer's win turns into Avoider's win in the planarity game.

THEOREM 2.3.

$$\frac{n}{2} - o(n) \leq f_{\mathcal{NP}_n}^- \leq f_{\mathcal{NP}_n}^+ \leq 2n^{5/4}.$$

Proof. Assume first that $b \geq 2n^{5/4}$. We will provide Avoider with a strategy for building a planar graph. The game is divided into four stages. Avoider's strategy is the following.

In the first stage, Avoider builds a matching by repeatedly claiming an edge that connects two vertices, neither of which is incident with any other edge previously claimed by him. The first stage ends when no such unclaimed edge remains and so Avoider cannot further extend his matching. We denote the set of vertices that are covered by Avoider's matching by M .

Next, in the second stage Avoider claims edges with one end point in M and the other in $V \setminus M$ such that throughout the second stage every vertex of $V \setminus M$ has degree at most one in Avoider's graph. The second stage ends when no such unclaimed edge remains.

In the third stage Avoider builds another matching on M . More precisely, he repeatedly claims edges that connect two vertices of M , neither of which is incident with any other edge previously claimed by him in the *third* stage. The third stage ends when no such unclaimed edge remains and Avoider cannot further extend this second matching.

In the fourth and final stage, Avoider claims edges arbitrarily to the end of the game. If we prove that in this stage Avoider will claim at most one edge, then the upper bound of the theorem will follow. Indeed, the graph that is spanned by Avoider's edges from the first and third stages is a union of two matchings, that is, a union of disjoint paths and cycles. Furthermore, if we add Avoider's edges from the second stage to this graph, then we simply add "hanging" edges (edges with one end point having degree one). Clearly, if we now add any single edge that may have been claimed by Avoider in the fourth stage to that graph, it remains planar.

Let e be the number of edges that Avoider claims in the entire game. By the end of the first stage, Enforcer must have claimed all of the edges with both end points in $V \setminus M$. Since Avoider's matching on M consists of at most e edges, we have $|V \setminus M| \geq n - 2e$, and therefore Enforcer has already claimed at least $\binom{n-2e}{2}$ edges. It follows that there are at most

$$\binom{n}{2} - \binom{n-2e}{2} \leq 2en$$

unclaimed edges left in the graph, and Avoider will claim at most $\frac{2en}{b}$ of them.

In the second stage, Avoider claims edges between M and $V \setminus M$. When this is no longer possible, every unclaimed edge between M and $V \setminus M$ is incident with a vertex of $V \setminus M$ which has degree one in Avoider's graph. It follows that, at this point, the number of unclaimed edges between M and $V \setminus M$ is at most

$$2e \cdot \frac{2en}{b} = \frac{4e^2n}{b}.$$

In the third stage, Avoider builds his second matching on M . When this is no longer possible, the number of unclaimed edges with both end points in M is at most

$$\binom{2e}{2} - \binom{2e - 4en/b}{2} \leq \frac{8e^2n}{b}.$$

To see this, it is enough to observe that the number of vertices that are incident with the second matching is at most $4en/b$ and that all edges with end points in M that are not adjacent to the second matching must be claimed by Enforcer after the third stage.

Putting everything together, the total number of unclaimed edges after the third stage is at most

$$\frac{4e^2n}{b} + \frac{8e^2n}{b} = \frac{12e^2n}{b}.$$

Since $e < n^2/2b$, we have that the number of edges to be played in the fourth stage is at most $\frac{12e^2n}{b} \leq \frac{3n^5}{b^3} \leq b$, which means that in the fourth stage Avoider will claim at most one edge.

Next, fix an $\varepsilon > 0$, and assume that $b \leq \frac{n}{2}(1 - \varepsilon)$. We will provide Enforcer with a strategy, which guarantees that Avoider will occupy the edges of a nonplanar graph.

If $b \leq n/7$, then the number of edges Avoider claims in the entire game is at least $\lfloor \binom{n}{2} \cdot (b+1)^{-1} \rfloor > 3n$, and thus Avoider surely loses regardless of Enforcer's strategy. Hence, from now on we can assume that $b > n/7$. Let $k = k(\varepsilon)$ be the smallest positive integer such that $\frac{1}{1-\varepsilon/2} > \frac{k}{k-2}$. Enforcer's strategy will be to prevent Avoider from claiming a cycle of length smaller than k , which we will call a "short cycle." If he succeeds, then at the end of the game Avoider's graph will have at least

$$\left\lfloor \frac{\binom{n}{2}}{b+1} \right\rfloor \geq \frac{n}{1-\varepsilon/2} > \frac{k}{k-2}n$$

edges, for sufficiently large n , and girth at least k . As we have mentioned before, a graph with such properties cannot be planar, and thus Enforcer wins.

It remains to show that Enforcer can indeed prevent Avoider from claiming a short cycle. In order to do that we will use the following theorem of Bednarska and Luczak.

THEOREM 2.4 (see [5, Theorem 1]). *For every graph G which contains at least three nonisolated vertices there exist positive constants c and n_0 such that, playing the $(1, q)$ game on K_n , G -Breaker can prevent G -Maker from building a copy of G provided that $n > n_0$ and $q > cn^{1/m_2(G)}$, where*

$$m_2(G) = \max_{\substack{H \subseteq G \\ v(H) \geq 3}} \frac{e(H) - 1}{v(H) - 2}.$$

For a cycle C_i of length i , we have $m_2(C_i) = \frac{i-1}{i-2}$. Therefore, there exist constants c_i , $i = 3, \dots, k-1$, such that, for sufficiently large n , Enforcer can prevent Avoider from claiming a copy of C_i , if the number of edges he is allowed to claim per move is at least $c_i n^{\frac{i-2}{i-1}}$. Since for sufficiently large n

$$\sum_{i=3}^{k-1} c_i n^{\frac{i-2}{i-1}} \leq \frac{n}{7} \leq b,$$

Enforcer can simultaneously prevent Avoider from claiming any short cycle C_i , $3 \leq i < k$, by simply playing all $k-3$ games in parallel. That is, after Avoider claims an edge, Enforcer responds by claiming $c_3 n^{\frac{1}{2}}$ edges according to the strategy in the "triangle avoidance game," then he claims $c_4 n^{\frac{2}{3}}$ edges according to the strategy in the "4-cycle avoidance game," and so on. His different strategies, for the different cycle games, might call for claiming the same edge more than once, in which case he just claims an arbitrary unclaimed edge instead. It is easy to see that this cannot harm him. This concludes the proof of the theorem. $\square$

3. k -colorability games.

3.1. The Maker-Breaker k -colorability game. The following theorem states that the threshold bias at which Maker's win turns into Breaker's win in the k -colorability game, where k is fixed and n is sufficiently large, is of order n . This is true for every $k \geq 2$. However, for convenience, and since the case $k = 2$ was treated in [8], we will assume that $k \geq 3$.

THEOREM 3.1. *For every $k \geq 3$ and sufficiently large n , there exist constants s_k and s'_k such that*

$$s'_k n \leq b_{NC_n^k} \leq s_k n,$$

where $s_k \sim \frac{2}{k \log k}$ and $s'_k \sim \frac{\log 2}{2k \log k}$ as $k \rightarrow \infty$.

Proof. Assume first that $b \leq \frac{n}{ck \log k}$, where $c > 2/\log 2$. We will provide Maker with a strategy for building a non- k -colorable graph. Maker's goal will be to prevent Breaker from building a clique of size $\lceil n/k \rceil$, and this is enough to ensure his win. Indeed, Maker's graph is surely not k -colorable if it does not admit an independent set of size $\lceil n/k \rceil$.

Let $\mathcal{F}$ be the hypergraph whose vertices are the edges of K_n and whose hyperedges are the $\lceil n/k \rceil$ -cliques of K_n . We name the players of the $(b, 1, \mathcal{F})$ game CliqueMaker and CliqueBreaker. As was mentioned before, Maker wins the k -colorability game if he claims a vertex in every hyperedge of $\mathcal{F}$, that is, if he is able to win the $\mathcal{F}$ game as CliqueBreaker. We will use Beck's criterion, which is applicable to any Maker-Breaker-type game.

THEOREM 3.2 (see [1, Theorem 1]). *If*

$$\sum_{D \in \mathcal{H}} (1+q)^{-|D|/p} < \frac{1}{q+1},$$

then Breaker wins the (p, q) game $\mathcal{H}$.

We have

$$\begin{aligned} \sum_{D \in \mathcal{F}} 2^{-|D|/b} &\leq \binom{n}{\lceil n/k \rceil} 2^{-\binom{\lceil n/k \rceil}{2}/b} \leq (ek)^{\lceil n/k \rceil} 2^{-\binom{\lceil n/k \rceil}{2}/b} \\ &\leq 2^{\frac{n \log_2 e}{k} + \frac{n \log_2 k}{k} + \log_2(ek) - \frac{cn^2 k \log k}{2k^2 n} + \frac{nc k \log k}{2kn}} = o(1). \end{aligned}$$

Hence, Maker can win the k -colorability game.

Assume now that $b \geq s_k n$, where s_k is a constant depending on k that will be determined later. We will provide Breaker with a strategy, to force Maker into building a k -colorable graph. We will make use of the following theorem of Kim.

THEOREM 3.3 (see [16, Corollary 1.2]). *If G is a graph with maximum degree Δ and girth at least 5, then*

$$\chi(G) \leq (1 + \nu(\Delta)) \frac{\Delta}{\log \Delta},$$

where $\nu(\Delta) \rightarrow 0$ as $\Delta \rightarrow \infty$.

Let Δ_0 be the maximal value of Δ for which

$$(1 + \nu(\Delta)) \frac{\Delta}{\log \Delta} \leq k$$

(if no such Δ_0 exists or if $\Delta_0 < 2$, then we take $s_k = 1/2$, and so Breaker wins by Theorem 2.2).

Since $\nu(\Delta) \rightarrow 0$ as $\Delta \rightarrow \infty$, we have that $\Delta_0 \sim k \log k$ as $k \rightarrow \infty$. Breaker's goal will be to force Maker to build a graph with maximum degree at most Δ_0 and girth at least 5. By Theorem 3.3 Maker's graph will then be k -colorable. In each move, Breaker will use $c_3 n^{1/2}$ of his edges to prevent Maker from building a triangle (recall that $m_2(C_3) = 2$) and $c_4 n^{2/3}$ of his edges to prevent Maker from building a cycle of length 4 ($m_2(C_4) = 3/2$), where c_3 and c_4 are the constants whose existence is guaranteed by Theorem 2.4. Breaker will use all of his remaining $b' := b - c_3 n^{1/2} - c_4 n^{2/3} = (1 - o(1))b$ edges to make sure that the maximum degree in Maker's graph does not surpass Δ_0 . Hence, if Maker claims the edge (u, v) , then Breaker will claim $\frac{1}{2}b'$ edges incident with u and $\frac{1}{2}b'$ edges incident with v (if there

are only $r < \frac{1}{2}b'$ unclaimed edges incident with u or v , then Breaker will claim all of them and additional $\frac{1}{2}b' - r$ arbitrary unclaimed edges). It follows that the maximum degree in Maker's graph will be at most

$$1 + \frac{n-1}{b'/2} \leq 1 + \frac{2n}{(1-o(1))b} \leq 1 + \frac{2}{s_k} + o(1),$$

where the $o(1)$ term tends to zero as n tends to infinity. Therefore, if $s_k = \lceil \frac{2}{\Delta_0 - 1.5} \rceil$, then Maker's graph will have maximum degree at most Δ_0 . Hence, Breaker can force Maker to build a graph with maximum degree at most Δ_0 and girth at least 5, and thus he can win. Note that s_k , defined this way, satisfies $s_k \sim \frac{2}{k \log k}$ as $k \rightarrow \infty$. This concludes the proof of the theorem. $\square$

3.2. The Avoider-Enforcer k -colorability game. In the following theorem we give an upper bound and a lower bound for the threshold bias at which Enforcer's win turns into Avoider's win in the k -colorability game.

THEOREM 3.4. *For every $k \geq 3$ there exists a constant s'_k such that*

$$s'_k n \leq f_{\mathcal{NC}_n}^- \leq f_{\mathcal{NC}_n}^+ \leq 2kn^{1+\frac{1}{2k-3}}.$$

Moreover, $s'_k \sim \frac{\log 2}{2k \log k}$ as $k \rightarrow \infty$.

Proof. Assume first that $b \leq \frac{n}{ck \log k}$. We will provide Enforcer with a strategy which ensures that by the end of the game Avoider's graph will not be k -colorable. Enforcer's goal will be to avoid building a clique of size $\lceil n/k \rceil$. If he achieves this goal, Avoider's graph will not contain an independent set of size $\lceil n/k \rceil$ and so will not be k -colorable; thus Enforcer will win. Let $\mathcal{F}$ be the hypergraph whose vertices are the edges of K_n and whose hyperedges are the $\lceil n/k \rceil$ -cliques of K_n . We name the players of the $(b, 1, \mathcal{F})$ game CliqueAvoider and CliqueEnforcer. As mentioned above, Enforcer will win the k -colorability game if he will not claim all vertices in any hyperedge of $\mathcal{F}$, that is, if he is able to win as CliqueAvoider.

We will use the following criterion, applicable to any Avoider-Enforcer-type game.

THEOREM 3.5 (see [14, Theorem 1.1]). *If*

$$\sum_{D \in \mathcal{H}} \left(1 + \frac{1}{b}\right)^{-|D|} < \left(1 + \frac{1}{b}\right)^{-b},$$

then Avoider wins the $(b, 1)$ game $\mathcal{H}$.

We have

$$\begin{aligned} \sum_{D \in \mathcal{F}} \left(1 + \frac{1}{b}\right)^{-|D|} &\leq \binom{n}{\lceil n/k \rceil} \left(1 + \frac{1}{b}\right)^{-\binom{\lceil n/k \rceil}{2}} \leq (ek)^{\lceil n/k \rceil} 2^{-\binom{\lceil n/k \rceil}{2}/b} \\ &\leq 2^{\frac{n \log 2 e}{k} + \frac{n \log 2 k}{k} + \log_2(ek) - \frac{cn^2 k \log k}{2k^2 n} + \frac{nc k \log k}{2kn}} = o(1). \end{aligned}$$

By applying Theorem 3.5 we conclude that there exists a winning strategy for CliqueAvoider, and thus Enforcer wins the k -colorability game.

Next, let $b > 2kn^{1+\frac{1}{2k-3}}$. We will provide Avoider with a strategy for building a $(k-1)$ -degenerate graph (a graph G is called r -degenerate if there is an ordering of the vertices $v_1, \dots, v_n$ such that every vertex has at most r neighbors with a higher index). Clearly, that would entail Avoider's win in the k -colorability game as every $(k-1)$ -degenerate graph is k -colorable.

Avoider will play several auxiliary minigames one after the other, never starting a new minigame before finishing the previous one, until all edges are claimed and the k -colorability game is over. Before we describe his strategy in detail, let us define two basic types of *minigames*.

Minigame type I. For a set of vertices A , the (A) -minigame is played on those edges with both end points in A which are still unclaimed at the beginning of this minigame. Note that some edges within A may have already been claimed during previous minigames. We say that the vertices of A are *designated* to the (A) -minigame. When we say that *Avoider is playing the (A) -minigame*, we mean that Avoider is repeatedly claiming independent edges with both end points in A for as long as possible; that is, he extends a matching on A until it is no longer possible. When Avoider cannot further extend his matching, the (A) -minigame is over. At this point we denote the set of vertices of A incident with an edge, claimed by Avoider in *this* minigame by A_1 , and let $A_2 = A \setminus A_1$. Note that by the end of the (A) -minigame all edges with both end points in A_2 have already been claimed by one of the players.

Minigame type II. Let A and B be two disjoint sets of vertices. The $(A : B)$ -minigame is played on those edges with one end point in A and the other in B which are still unclaimed at the beginning of this minigame. Again, we assume that the (big) game is in progress, meaning that some of the edges between A and B may have already been claimed in previous minigames. We say that the vertices of B are *designated* to the $(A : B)$ -minigame. When we say that *Avoider is playing the $(A : B)$ -minigame*, we mean that Avoider is repeatedly claiming edges between A and B such that no vertex in B is incident with more than one of Avoider's edges claimed in *this* minigame. When this is no longer possible, the $(A : B)$ -minigame is over. At this point, let B_1 denote the set of vertices of B that are incident with an edge claimed by Avoider in *this* minigame, and let $B_2 = B \setminus B_1$. Note that all edges with one end point in A and the other in B_2 have already been claimed by one of the players. The vertices in B_2 are called *finished*.

Now we can describe the way Avoider plays the game. We introduce a *minigame pool* $\mathcal{P}$, which is a dynamic collection of minigames that will be updated during the game—it will contain minigames waiting to be played by Avoider. At each moment $\mathcal{P}$ will contain at most one minigame of type I and at most $k - 1$ minigames of type II.

Avoider will maintain a partial ordering of the vertices, which he will refine whenever a minigame is over. In this partial ordering, the vertices designated to the same minigame will be incomparable to each other, the vertices designated to the lone minigame of type I in the pool will be above all of the other vertices, and for any minigame $(A : B)$ of type II every vertex of A will be above every vertex of B .

Given a partial ordering, let the *up-degree* of v be the number of Avoider's edges (v, u) where either u is above v or they are incomparable.

To each minigame in the pool, we assign an integer parameter that will help us keep track of the degeneracy of Avoider's graph throughout the game. Thus, instead of the (A) -minigame (or the $(A : B)$ -minigame), we will consider the $(A)_l$ -minigame (or the $(A : B)_l$ -minigame) for an appropriate integer l . During play, Avoider maintains the following property: If a vertex is designated to a minigame with parameter l , then its up-degree is at most l .

In the beginning of the game $\mathcal{P}$ contains only one minigame—the $(V(K_n))_0$ -minigame. The partially ordered set on $V(K_n)$ contains no relations.

We say that the size of an (A) -minigame is $\frac{1}{2}|A|^2$ and the size of an $(A : B)$ -minigame is $|A| \cdot |B|$. Note that the size of a minigame is an upper bound on the number of edges it contains.

When the game is played, Avoider repeatedly chooses a minigame of the largest size in the pool $\mathcal{P}$, removes it from the pool, plays it to its end, and then updates $\mathcal{P}$ and the partial ordering as follows. If the minigame played was an $(A)_l$ -minigame, then he places two new minigames into $\mathcal{P}$: the $(A_1)_{l+1}$ -minigame and the $(A_1 : A_2)_l$ -minigame. Furthermore, the designation of the vertices of A is lifted and replaced by that of the vertices of A_1 to the $(A_1)_{l+1}$ -minigame and that of the vertices of A_2 to the $(A_1 : A_2)_l$ -minigame. The partial order is refined by placing every vertex of A_1 above every vertex of A_2 .

On the other hand, if the minigame played was an $(A : B)_l$ -minigame, then Avoider places only the $(A : B_1)_{l+1}$ -minigame back into $\mathcal{P}$. Furthermore, the designation of the vertices of B is lifted—the vertices of B_2 are finished, and the vertices of B_1 are designated to the $(A : B_1)_{l+1}$ -minigame. The partial order is not affected.

This shows that indeed in every stage of the game $\mathcal{P}$ will contain at most (in fact, exactly) one minigame of type I.

Note that at any point of the game every unclaimed edge is in exactly one of the minigames in $\mathcal{P}$. Moreover, every vertex of K_n will be either finished or designated to exactly one minigame in the pool.

After having played an $(A : B)$ -minigame of type II, the up-degree of the finished vertices, that is, the vertices of B_2 , is fixed and in particular will not be increased in later stages of the game. This is because there are simply no more unclaimed edges which go to higher or incomparable vertices left. Indeed, the edges with both end points in B were all claimed during that minigame of type I after which the vertices of B were designated to a type II minigame. The edges (u, v) , where $u \in B_2$ and v is above u , were all claimed during the $(A : B)$ -minigame. Furthermore, the up-degree of every vertex of B_2 was not changed during the $(A : B)$ -minigame, so if the parameter of the $(A : B)$ -minigame was l , then the up-degree of the vertices in B_2 is at most l at the end of the game.

It is clear that, as long as Avoider follows this strategy and the parameter of every minigame in $\mathcal{P}$ is at most l , Avoider's graph is l -degenerate. Therefore, it suffices to prove that, after the first minigame with parameter $k-2$ is taken out of the pool to be played, Avoider plays at most one more move in the whole game. Note that whenever a minigame of type II is played the size of the pool $\mathcal{P}$ is not changed, and whenever a minigame of type I is played both the size of the pool $\mathcal{P}$ and the parameter of the new minigame of type I are increased by one. It follows that proving the above will show that indeed, throughout the game, there will be at most $k-1$ minigames of type II in $\mathcal{P}$.

We will prove by induction on l that any minigame in the pool which has parameter $0 \leq l \leq k-2$ is of size at most $n^2 \left(\frac{2k^2 n^2}{b^2} \right)^l$. First, for the base step, note that the size of any minigame with parameter $l=0$ is less than n^2 . Now let us assume that l is an integer with $0 < l \leq k-2$ and the induction hypotheses holds for all games with parameter less than l . For a minigame M in the pool with parameter l we consider three cases.

Case 1. M is an $(A_1)_l$ -minigame that was inserted into the pool after the $(A)_{l-1}$ -minigame has ended. Just before Avoider started playing the $(A)_{l-1}$ -minigame, there was no minigame in the pool of larger size. Since the total number of games in the pool was at most k , the total number of unplayed edges at that point was at most k times the size of the $(A)_{l-1}$ -minigame. By the induction hypotheses, this is at most $kn^2 \left(\frac{2k^2 n^2}{b^2} \right)^{l-1}$. The number of edges Avoider will play during the $(A)_{l-1}$ -minigame is certainly bounded from above by the total number of edges that Avoider will claim un-

til the end of the whole k -colorability game, which is at most $\frac{kn^2}{b} \left(\frac{2k^2n^2}{b^2}\right)^{l-1}$. Avoider's strategy for the $(A)_{l-1}$ -minigame guarantees that the set A_1 will be of size at most twice this much, and hence the $(A_1)_l$ -minigame will be of size at most

$$\frac{1}{2}|A_1|^2 \leq \frac{1}{2} \left(\frac{2kn^2}{b} \left(\frac{2k^2n^2}{b^2} \right)^{l-1} \right)^2 \leq n^2 \cdot \left(\frac{2k^2 \cdot n^2}{b^2} \right)^l.$$

Case 2. M is an $(A_1 : A_2)_l$ -minigame that was inserted into the pool after the $(A_1 \cup A_2)_l$ -minigame has ended. The size of the $(A_1 : A_2)_l$ -minigame is obviously bounded from above by the size of the $(A_1 \cup A_2)_l$ -minigame, which we already upper-bounded in Case 1.

Case 3. M is an $(A : B)_l$ -minigame that was inserted into the pool after the $(A : B)_{l-1}$ -minigame has ended. As in Case 1, we can bound the number of edges Avoider will play during the $(A : B)_{l-1}$ -minigame from above by the total number of edges that Avoider will claim until the end of the whole k -colorability game. Thus, by knowing that the $(A : B)_{l-1}$ -minigame was of maximal size in $\mathcal{P}$ before it was played, we get that Avoider will make at most

$$\frac{k|A||B|}{b} \leq \frac{kn^2}{b} \left(\frac{2k^2n^2}{b^2} \right)^{l-1}$$

moves until the end of the game. Therefore, the size of B_1 is also at most that much. Since the size of A is at most n^2/b (the total number of vertices that can have positive degree in Avoider's graph), the total size of the $(A : B)_l$ -minigame is at most $n^2 \left(\frac{2k^2n^2}{b^2}\right)^l$. This concludes the induction step.

At the point when a minigame with parameter $k - 2$ becomes the largest size in the pool, then the total number of edges to be played in the remainder of the game is at most $kn^2 \left(\frac{2k^2n^2}{b^2}\right)^{k-2}$, which is less than b , meaning that Avoider will play at most one move before the game ends. However, at this point Avoider's graph is $(k - 2)$ -degenerate, so we are done. $\square$

Remark. The graph built by Avoider in the proof of Theorem 2.3 is clearly 3-colorable. It follows that if $k \geq 3$ and $b \geq 2n^{5/4}$, then Avoider can win the $(1, b)$ k -colorability game. For $k = 3$ this yields a better result than the one given in Theorem 3.4. Moreover, it is easy to see that if $b > n^{3/2}$, then Avoider can build a graph which consists of a matching and one additional edge; clearly such a graph is 2-colorable. Hence, by using Enforcer's strategy from Theorem 3.4 we get

$$cn \leq f_{\mathcal{NC}_n^2}^- \leq f_{\mathcal{NC}_n^2}^+ \leq n^{3/2}$$

for an appropriate constant $c > 0$.

4. Minor games.

4.1. The Maker-Breaker minor game. In the Maker-Breaker version of the game, Maker's goal is to build a graph that contains a K_t -minor. The following theorem states that the threshold bias at which Maker's win turns into Breaker's win in the K_t -minor game for every $3 \leq t \leq c\sqrt{n/\log n}$, for an appropriate constant c , is asymptotically $n/2$.

THEOREM 4.1. *For every fixed $\varepsilon > 0$, there exists a constant $C = C(\varepsilon)$ such that if n is sufficiently large and $b \leq (1 - \varepsilon)n/2$, then Maker has a winning strategy for the $(1, b)$ game $\mathcal{M}_n^t$ for every $t < C\sqrt{n/\log n}$.*

As a corollary we have the asymptotics of $b_{\mathcal{M}_n^t}$ for arbitrary fixed t .

COROLLARY 4.2. *Let $t \geq 3$ be a positive integer. Then*

$$b_{\mathcal{M}_n^t} = \frac{n}{2} - o_t(n).$$

The lower bound in the corollary follows from Theorem 4.1, while if $b \geq n/2$, then Breaker, as in the proof of Theorem 2.1, can force Maker to build a forest, which does not contain a K_t -minor for any $t \geq 3$.

In the proof of Theorem 4.1 we will use the following result of Kostochka and of Thomason.

THEOREM 4.3 (see [17, 26]). *There exists a constant c' such that every graph of average degree at least $c'r\sqrt{\log r}$ admits a K_r -minor.*

Proof of Theorem 4.1. Assume that $b \leq (1 - \varepsilon)n/2$. We will provide Maker with a strategy for building a graph that admits a large minor. Maker's strategy is divided into two stages. In the first stage that lasts exactly $m-1$ (where m is to be determined later) rounds, Maker builds a tree $T = (V, E)$ that satisfies the following properties:

1. $|V| = m \geq \varepsilon n$;
2. the degree of every $u \in V$ is at most 3;
3. there remain at least $\varepsilon^2 n^2/3$ unclaimed edges with both end points in V .

Maker's strategy for building such a tree is very simple: He starts by claiming an arbitrary edge, and then, for as long as possible, he claims a previously unclaimed edge (u, v) such that, in his current graph, u has degree 1 or 2 and v is isolated; such an edge, when it exists, is chosen arbitrarily. Clearly this results in a tree with maximum degree at most 3. Now assume that by using this strategy Maker was able to build a tree on m vertices but could not extend it to a tree on $m+1$ vertices (while maintaining the maximum degree criterion). This means that every edge (u, v) such that, in Maker's graph, u has degree 1 or 2 and v is isolated must have been claimed by Breaker. Since Maker's graph is a tree, at least half its vertices have degree at most 2, and so Breaker must have claimed at least $\frac{m}{2}(n-m)$ edges. But at this point, Breaker has at most $m(1 - \varepsilon)n/2$ edges entailing $m \geq \varepsilon n$. Furthermore, the number of edges with both end points in V that Breaker could have claimed is at most $m(1 - \varepsilon)n/2 - \frac{m}{2}(n-m) = \frac{m^2}{2} - \frac{m\varepsilon n}{2}$. It follows that there must be at least $\frac{m\varepsilon n}{2} - \frac{3m}{2} \geq \varepsilon^2 n^2/3$ unclaimed edges with both end points in V . This ends the first stage.

Before claiming edges in the second stage, Maker would like to partition T into roughly $\sqrt{n}$ connected components of roughly the same size. The following result of Krivelevich and Nachmias asserts that he can.

LEMMA 4.4 (see [18, Proposition 4.5]). *Let $G = (V, E)$ be a connected graph on r vertices with maximum degree at most k . Then for every positive integer l there exist pairwise disjoint sets $V_1, \dots, V_s \subseteq V$, with the following properties:*

1. $lk \leq |V_i| \leq lk^2$ for every $1 \leq i \leq s$.
2. $\sum_{i=1}^s |V_i| \geq r - lk$.
3. $G[V_i]$ is connected for every $1 \leq i \leq s$.

By using Lemma 4.4 with $r = m$, $l = \varepsilon^{-1}\sqrt{m}$, and $k = 3$ we conclude that at least $m - 3\varepsilon^{-1}\sqrt{m}$ of the vertices of T can be partitioned into parts $V_1, \dots, V_s$ such that $3\varepsilon^{-1}\sqrt{m} \leq |V_i| \leq 9\varepsilon^{-1}\sqrt{m}$ and $T[V_i]$ is connected for every $1 \leq i \leq s$. Note that $\varepsilon\sqrt{m}/10 \leq s \leq \varepsilon\sqrt{m}/3$.

A pair (V_i, V_j) will be called *good* if there are at least $b+1$ unclaimed edges (u, v) where $u \in V_i$ and $v \in V_j$. We claim that at least an $\varepsilon^2/20$ fraction of the total number

of pairs is good. Indeed, assume for the sake of contradiction that there are less than $\varepsilon^2 \binom{s}{2}/20$ good pairs; then there are at most

$$3m\varepsilon^{-1}\sqrt{m} + \sum_{i=1}^s \binom{|V_i|}{2} + b \binom{s}{2} + \varepsilon^2 \binom{s}{2} (9\varepsilon^{-1}\sqrt{m})^2/20$$

unclaimed edges in V (the first term stands for the edges incident with vertices outside $\bigcup_{i=1}^s V_i$, the second term stands for edges inside the V_i 's, the third term stands for unclaimed edges that might remain between any pair, even if it is not good, and the fourth term stands for edges between good pairs). For sufficiently large n this is strictly less than $\varepsilon^2 n^2/3$; this contradicts Maker's strategy for the first stage.

For every good pair (V_i, V_j) , let $A_{i,j}$ be any set of $b+1$ unclaimed edges with one end point in V_i and the other in V_j . Trivially, by simply not claiming more than one edge from any $A_{i,j}$ (and not claiming edges outside the $A_{i,j}$'s for as long as possible), Maker can claim an edge of $A_{i,j}$ for at least half of the good pairs (V_i, V_j) . In the second stage, Maker will use this strategy; denote Maker's graph at the end of the second stage by H . Consider the graph $\tilde{H}$ on the vertex set $V(\tilde{H}) = \{V_1, \dots, V_s\}$, where (V_i, V_j) is an edge if and only if Maker has claimed an edge (x, y) such that $x \in V_i$ and $y \in V_j$. The average degree in $\tilde{H}$ is at least

$$\frac{\varepsilon^2 \binom{s}{2}/20}{s} \geq \varepsilon^3 \sqrt{m}/400 \geq \varepsilon^4 \sqrt{n}/400,$$

and so by Theorem 4.3 it admits a K_t -minor for $t = C\sqrt{n/\log n}$, for an appropriate constant C . Since $T[V_i]$ is connected for every $1 \leq i \leq s$, H admits the same minor, and the proof of Theorem 4.1 is complete. $\square$

4.2. The Avoider-Enforcer minor game. In this game, Enforcer would like to make Avoider build a graph that admits a K_t -minor. As in the Maker-Breaker case, we allow t to be a function of n .

In the first theorem, we show that with a linear bias Enforcer can win the game for t which is as large as $c\sqrt{n/\log n}$.

THEOREM 4.5. *If $b \leq n/19$, then Enforcer has a winning strategy for the $(1, b)$ game $\mathcal{M}_n^t$ for every $t < c\sqrt{n/\log n}$, where c is some absolute constant.*

In the proof of Theorem 4.5, we will use a result of Plotkin, Rao, and Smith. Before we can state their result, we need the following definition.

DEFINITION. *Let $G = (V, E)$ be a graph on n vertices. A set $S \subset V$ is called a separator if every connected component of $G[V \setminus S]$ contains at most $2n/3$ vertices.*

THEOREM 4.6 (see [24, Corollary 2.4]). *Let G be a graph on n vertices, and let h be a function of n . If G does not have a separator of size at most $O(h\sqrt{n \log n})$, then G admits a K_h -minor.*

Proof of Theorem 4.5. Assume that $b \leq n/19$; we will present a winning strategy for Enforcer. We will prove that Enforcer can make Avoider build a graph in which there is an edge between any two disjoint vertex sets of size at least $s = (1/3 - \varepsilon)n$ each, where $\varepsilon > 0$ is some small constant. Let $\mathcal{H}_n$ denote the hypergraph whose vertices are the edges of K_n and whose hyperedges are all of the subgraphs of K_n , isomorphic to $K_{s,s}$. Enforcer's goal is to avoid claiming any hyperedge of $\mathcal{H}_n$. By using the criterion given by Theorem 3.5, we obtain

$$\sum_{D \in \mathcal{H}_n} \left(1 + \frac{1}{b}\right)^{-|D|} \leq \binom{n}{n/3} \binom{n}{n/3} 2^{-n^2/(9+\varepsilon')b} = o(1).$$

Thus, Enforcer can make sure that Avoider's graph will not contain a separator of size at most εn , and so, by Theorem 4.6, Avoider's graph will admit a K_h -minor for $h = \Theta(\sqrt{n/\log n})$. This concludes the proof of the theorem. $\square$

In the following theorem we show that, even when the bias b is almost as large as $n/2$, Enforcer can still make Avoider claim all edges of a K_t -minor, where t is some constant power of n . On the other hand, we also prove an upper bound for the upper threshold bias of this game.

THEOREM 4.7. *For every $\varepsilon > 0$ there exists a constant $a_0 = a_0(\varepsilon) > 0$ such that for every $4 \leq t \leq n^{a_0}$ we have*

$$\left(\frac{1}{2} - \varepsilon\right)n \leq f_{\mathcal{M}_n^t}^- \leq f_{\mathcal{M}_n^t}^+ \leq 2n^{5/4}.$$

Before proving this theorem, we will state and prove a graph-theoretic lemma, which may also be of independent interest.

LEMMA 4.8. *Let $G = (V, E)$ be a graph with average degree $2 + \alpha$, for some $\alpha > 0$, and girth $g^* \geq (1 + \frac{2}{\alpha})(4 \log_2 t + 2 \log_2 \log_2 t + c)$, where c is an appropriate constant. Then G admits a K_t -minor.*

Proof. In the proof of the lemma we will use the following result of Kühn and Osthus (a similar result was also obtained by Diestel and Rempel [12]).

THEOREM 4.9 (see [19, Corollary 5]). *Let $t \geq 3$ be an integer. There exists a constant c such that every graph of minimum degree at least 3 and girth at least $4 \log_2 t + 2 \log_2 \log_2 t + c$ contains a K_t -minor.*

We repeatedly apply two deletion operations on G , which do not decrease the average degree and (trivially) do not decrease the girth. The first operation is the deletion of a vertex of degree at most one. Such an operation obviously does not decrease the average degree. In the second type of operation, given a path $u_1, u_2, \dots, u_k$, with $k \geq 2 + 2/\alpha$, such that each of the internal vertices $u_2, u_3, \dots, u_{k-1}$ has degree two in G , we remove $u_2, \dots, u_{k-1}$. Again the average degree of the new graph is at least $2 + \alpha$. To verify this, let us assume that from a graph with e edges and v vertices satisfying $\frac{2e}{v} \geq 2 + \alpha$ we remove the internal vertices of a path with $k \geq 2 + 2/\alpha$ vertices. Then we obtain a graph with average degree at least

$$\frac{2(e - k + 1)}{v - (k - 2)} \geq 2 + \alpha,$$

as claimed. Let G_2 be the graph we obtain from G by repeated applications of these two operations. Since the average degree was not decreased in any step of the process, G_2 is not empty. It also follows that $\delta(G_2) \geq 2$, the girth of G_2 , is at least g^* and every path of G_2 , with internal vertices of degree two, is of length at most $1 + 2/\alpha$. Let G_3 denote the graph obtained from G_2 by contracting every path $u_1, u_2, \dots, u_k$, such that u_i has degree two in G_2 for every $1 < i < k$, into a single edge. Again, this operation does not decrease the average degree, and therefore G_3 is not empty. Clearly $\delta(G_3) \geq 3$. Moreover, since every such path in G_2 is of length at most $1 + 2/\alpha$ it follows that the girth of G_3 is at least $g = \frac{g^*}{1 + 2/\alpha} \geq 4 \log_2 t + 2 \log_2 \log_2 t + c$. By applying Theorem 4.9 we conclude that G_3 admits a K_t -minor. Since G_3 was obtained from G by the deletion and contraction of edges (and the removal of isolated vertices), G admits the same minor, and the proof is complete. $\square$

Proof of Theorem 4.7. Let $\varepsilon > 0$, and let $n = n(\varepsilon)$ be sufficiently large. Assume that $b \leq (1/2 - \varepsilon)n$, and $4 \leq t \leq n^{a_0}$, where $a_0 = a_0(\varepsilon)$ is a “small” positive constant

whose value will be determined later. If $b \leq n/19$, then by Theorem 4.5 Enforcer wins. Hence, from now on we can assume that $n/19 < b \leq (1/2 - \varepsilon)n$.

Let $\alpha = \alpha(n, \varepsilon) > 0$ be the real number that satisfies

$$(1 + \alpha)n = \frac{\binom{n}{2}}{(1/2 - \varepsilon)n + 1},$$

and let

$$\ell = 5a_0 \left(1 + \frac{2}{\alpha}\right) \log_2 n.$$

Enforcer's goal is to make sure that Avoider claims the edges of at most $o(n)$ cycles of length at most ℓ , throughout the game. In order to prove that Enforcer is able to do so, we will use a generalization of the Erdős–Selfridge theorem that applies to Maker–Breaker games in which the goal of Maker is to claim d (instead of just one) winning sets.

THEOREM 4.10 (see [1, 5]). *If for positive integers d and b we have*

$$\sum_{A \in \mathcal{F}} (1 + b)^{-|A|} < d,$$

then Breaker (as the first player) has a winning strategy in the $(1, b)$ Maker–Breaker game $\{\cup_{B \in \mathcal{F}} B : F \in \binom{\mathcal{F}}{d}\}$ (a minimal hyperedge of $\{\cup_{B \in \mathcal{F}} B : F \in \binom{\mathcal{F}}{d}\}$ is the union of d , not necessarily disjoint, hyperedges of $\mathcal{F}$).

Assume that, on his first move, Avoider claims some edge e . Enforcer will claim the role of Breaker, as the first player, in the game on $E(K_n) \setminus \{e\}$, in which the winning sets are the cycle of length at most ℓ . We have

$$\begin{aligned} \sum_{A \in \mathcal{F}} (1 + b)^{-|A|} &\leq \sum_{A \in \mathcal{F}} \left(\frac{n}{19}\right)^{-|A|} \\ &\leq \sum_{3 \leq i \leq \ell} \frac{\binom{n}{i}}{2^i} \left(\frac{n}{19}\right)^{-i} \\ &\leq \sum_{3 \leq i \leq \ell} 19^i \\ &\leq \ell \cdot n^{5a_0(1 + \frac{2}{\alpha}) \log_2 19} = o(n), \end{aligned}$$

where the last equality follows from an appropriate choice of a_0 .

By applying Theorem 4.10, we conclude that, in the entire game, Avoider will claim the edges of at most $o(n)$ cycles of length at most ℓ in $E(K_n) \setminus \{e\}$.

Let G_A denote Avoider's graph after he has played exactly $(1 + \alpha)n$ moves. The average degree in G_A is $2 + 2\alpha$. We will prove that Enforcer has won the game already at this point. Remove the edge e , as well as one edge from every cycle of length at most ℓ in G_A ; denote the resulting graph by G'_A . Clearly the girth of G'_A is at least ℓ . Moreover, the average degree in G'_A is at least $2 + \alpha$ since we removed just $o(n)$ edges from G_A .

Since $\ell \geq (1 + \frac{2}{\alpha})(4 \log_2 t + 2 \log_2 \log_2 t + c)$, where c is the constant given by Theorem 4.9, we conclude that G'_A admits a K_t -minor by Lemma 4.8. Clearly G_A admits the same minor.

If $b \geq 2n^{5/4}$, then by playing as in the proof of Theorem 2.3 Avoider builds a graph that does not admit a K_4 -minor.

Remark. As in the remark following the proof of Theorem 3.4, if $b > n^{3/2}$, then Avoider can create a K_3 -minor-free graph (i.e., a forest). Moreover, our strategy for Enforcer is valid also for $t = 3$.

5. Concluding remarks and open problems.

Maker-breaker thresholds. In section 3 it was proved that $b_{\mathcal{NC}_n^k} = \Theta(n)$ for every fixed k . We believe that in fact the following stronger statement holds.

CONJECTURE 5.1. *There is a constant c such that for every $k \geq 3$ and sufficiently large n we have*

$$b_{\mathcal{NC}_n^k} = \frac{(c + o(1))}{k \log k} n.$$

For both the planarity and the K_t -minor Maker-Breaker games, the second order terms are unknown; this is worth studying, in particular the dependence of the threshold $b_{\mathcal{M}_n^t}$ on t .

Asymptotic monotonicity of Avoider-Enforcer games. We say that an Avoider-Enforcer game $\mathcal{F}$ is *monotone* if the existence of Avoider's winning strategy for the $(1, q, \mathcal{F})$ game implies his win in the $(1, q + 1, \mathcal{F})$ game or, equivalently, if $f_{\mathcal{F}}^+ = f_{\mathcal{F}}^-$. Following [14], the function $f(n)$ is called an *asymptotic threshold bias* of the game $\mathcal{B}_n$ if both $f_{\mathcal{B}_n}^- = \Theta(f(n))$ and $f_{\mathcal{B}_n}^+ = \Theta(f(n))$. If an asymptotic threshold bias exists, that is, if $f_{\mathcal{B}_n}^- = \Theta(f_{\mathcal{B}_n}^+)$, then the game $\mathcal{B}_n$ is called *asymptotically monotone*. In [14] it was conjectured that the perfect matching and Hamiltonicity games are monotone. Here we conjecture the following.

CONJECTURE 5.2. *The Avoider-Enforcer nonplanarity, non- k -colorability, and K_t -minor games are asymptotically monotone.*

Avoider's strategies. For all of the Avoider-Enforcer games studied in this paper, there is a significant gap between the upper and lower bounds on the threshold biases. It would be interesting to close, or at least to reduce, these gaps. We do not believe that any of the upper bounds for $f_{\mathcal{F}}^+$ proved in this paper are tight. As a first step, it would be desirable to obtain an upper bound for the upper threshold bias $f_{\mathcal{M}_n^t}^+$ of the K_t -minor game which does depend on t .

The reason for the difficulty of finding upper bounds is due partly to the lack of a usable criterion for Avoider's win, similar to Theorem 3.5, in a $(1, q)$ game where $q > 1$. This is reminiscent of the difficulties that arise in the Hamiltonicity game $\mathcal{H}_n$ and in the perfect matching game $\mathcal{M}_n$ (see [14]). Let us recall that it is not even known whether Avoider can win $(1, n/100, \mathcal{M}_n)$ or $(1, n/100, \mathcal{H}_n)$. It seems quite reasonable that he can, knowing that Avoider wins $(1, n, \mathcal{M}_n)$ and $(1, n/2, \mathcal{H}_n)$ regardless of his strategy.

Acknowledgment. The authors thank the anonymous referee for helpful comments.

REFERENCES

- [1] J. BECK, *Remarks on positional games*, Acta Math. Acad. Sci. Hungar., 40 (1982), pp. 65–71.
- [2] J. BECK, *Deterministic graph games and a probabilistic intuition*, Combin. Probab. Comput., 3 (1994), pp. 13–26.
- [3] J. BECK, *Ramsey games*, Discrete Math., 249 (2002), pp. 3–30.
- [4] J. BECK, *Positional games and the second moment method*, Combinatorica, 22 (2002), pp. 169–216.

- [5] M. BEDNARSKA AND T. LUCZAK, *Biased positional games for which random strategies are nearly optimal*, *Combinatorica*, 20 (2000), pp. 477–488.
- [6] M. BEDNARSKA AND T. LUCZAK, *Biased positional games and the phase transition*, *Random Structures Algorithms*, 18 (2001), pp. 141–152.
- [7] M. BEDNARSKA AND O. PIKHURKO, *Biased positional games on matroids*, *European J. Combin.*, 26 (2005), pp. 271–285.
- [8] M. BEDNARSKA AND O. PIKHURKO, *Odd and even cycles in Maker-Breaker games*, *European J. Combin.*, to appear.
- [9] B. BOLLOBÁS AND A. PAPAIOANNOU, *A biased Hamiltonian game*, *Congr. Numer.*, 35 (1982), pp. 105–115.
- [10] V. CHVÁTAL AND P. ERDŐS, *Biased positional games*, *Ann. Discrete Math.*, 2 (1978), pp. 221–228.
- [11] R. DIESTEL, *Graph Theory*, 2nd ed., Springer, New York, 1999.
- [12] R. DIESTEL AND C. REMPEL, *Dense minors in graphs of large girth*, *Combinatorica*, 25 (2005), pp. 111–116.
- [13] A. FRIEZE, M. KRIVELEVICH, O. PIKHURKO, AND T. SZABÓ, *The game of JumbleG*, *Combin. Probab. Comput.*, 14 (2005), pp. 783–793.
- [14] D. HEFETZ, M. KRIVELEVICH, AND T. SZABÓ, *Avoider-Enforcer games*, *J. Combin. Theory Ser. A*, 114 (2007), pp. 840–853.
- [15] D. HEFETZ, M. KRIVELEVICH, AND T. SZABÓ, *Bart-Moe games, JumbleG and Discrepancy*, *European J. Combin.*, 28 (2007), pp. 1131–1143.
- [16] J. H. KIM, *On Brooks’ theorem for sparse graphs*, *Combin. Probab. Comput.*, 4 (1995), pp. 97–132.
- [17] A. V. KOSTOCHKA, *A lower bound for the Hadwiger number of graphs by the average degree*, *Combinatorica*, 4 (1984), pp. 307–316.
- [18] M. KRIVELEVICH AND A. NACHMIAS, *Colouring complete bipartite graphs from random lists*, *Random Structures Algorithms*, 29 (2006), pp. 436–449.
- [19] D. KÜHN AND D. OSTHUS, *Minors in graphs of large girth*, *Random Structures Algorithms*, 22 (2003), pp. 213–225.
- [20] A. LEHMAN, *A solution of the Shannon switching game*, *J. Soc. Indust. Appl. Math.*, 12 (1964), pp. 687–725.
- [21] X. LU, *A Matching game*, *Discrete Math.*, 94 (1991), pp. 199–207.
- [22] X. LU, *Hamiltonian games*, *J. Combin. Theory Ser. B*, 55 (1992), pp. 18–32.
- [23] X. LU, *A Hamiltonian game on $K_{n,n}$* , *Discrete Math.*, 142 (1995), pp. 185–191.
- [24] S. A. PLOTKIN, S. RAO, AND W. D. SMITH, *Shallow excluded minors and improved graph decompositions*, in *Proceedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA’94)*, 1994, pp. 462–470.
- [25] M. STOJAKOVIĆ AND T. SZABÓ, *Positional games on random graphs*, *Random Structures Algorithms*, 26 (2005), pp. 204–223.
- [26] A. THOMASON, *An extremal function for contractions of graphs*, *Math. Proc. Cambridge Philos. Soc.*, 95 (1984), pp. 261–265.

$L(2, 1)$ -LABELING OF HAMILTONIAN GRAPHS WITH MAXIMUM DEGREE 3*

JEONG-HYUN KANG[†]

Abstract. An integer coloring f of the vertices of a graph G is an $L(2, 1)$ -labeling if $|f(u) - f(v)| \geq 2$ for each edge uv and $|f(u) - f(v)| \geq 1$ for each pair $u, v \in V(G)$ at distance 2 apart. The $L(2, 1)$ -labeling span of G , denoted by $\lambda(G)$, is the smallest number t such that G has an $L(2, 1)$ -labeling using labels $\{0, 1, \dots, t\}$. Griggs and Yeh [*SIAM J. Discrete Math.*, 5 (1992), pp. 586–595] conjectured that always $\lambda(G) \leq (\Delta(G))^2$, where $\Delta(G)$ is the maximum degree of G and $\Delta(G) \geq 2$. In this paper, we will prove that the conjecture holds if G is a Hamiltonian graph with maximum degree 3.

Key words. $L(2, 1)$ -labeling, channel assignment, graph labeling, generalized coloring

AMS subject classifications. 05C78, 05C15

DOI. 10.1137/050632609

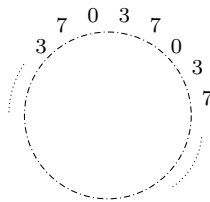
1. Introduction. In ordinary graph coloring, adjacent vertices must be given different colors, and the actual values of the colors used are irrelevant. However, in many applications, it is also important to separate labels on vertices at farther distances, where the labels used have some numerical meaning. A natural problem of this type is the channel assignment problem, where channels (nonnegative integers) are assigned to each radio transmitter (vertex) so that interfering (adjacent) transmitters get channels that are far apart. Roberts proposed a variation of the channel assignment problem, which Griggs and Yeh [5] introduced in 1992 and called the $L(2, 1)$ -labeling problem. Keeping the radio transmitter analogy in mind, vertices in a graph need to be labeled such that “close” vertices (at distance 2) get different labels, while “very close” vertices (at distance 1) get labels that are farther apart. More precisely, for a given graph G , a mapping $f : V(G) \rightarrow \mathbb{N} \cup \{0\}$ is called an $L(2, 1)$ -labeling if $|f(u) - f(v)| \geq 2$ for each edge uv of G and $|f(u) - f(v)| \geq 1$ for each pair $u, v \in V(G)$ at distance 2 apart. The $L(2, 1)$ -labeling number of G , denoted by $\lambda(G)$, is the smallest number t such that G has an $L(2, 1)$ -labeling that does not use any label greater than t .

As with the chromatic number of a graph G , it is natural to seek a bound on $\lambda(G)$ in terms of the maximum degree $\Delta(G)$. The obvious lower bound for $\lambda(G)$ is $\Delta(G) + 1$, which holds with equality for stars. A greedy labeling (as shown in [5]) gives $\lambda(G) \leq \Delta^2(G) + 2\Delta(G)$. This upper bound was improved to $\Delta^2(G) + \Delta(G)$ by Chang and Kuo [1], $\Delta^2(G) + \Delta(G) - 1$ by Král’ and Škrekovski [8], and $\Delta^2(G) + \Delta(G) - 2$ by Gonçalves [4]. Griggs and Yeh [5] conjectured that $\lambda(G) \leq \Delta^2(G)$ for every graph G . This has been a motivating problem for research in this field, and some results are known. Note that it is enough to consider connected regular graphs. If G is a 2-regular graph (a cycle), then $\lambda(G) = 4$, as proved in [5]. Tight bounds have been obtained for special classes of graphs like paths, cycles, wheels, complete k -partite graphs, and graphs with diameter 2 [5], trees [1, 5], the incidence graphs of projective planes [2], etc. Bounds have also been obtained for various other graph families like

*Received by the editors May 27, 2005; accepted for publication (in revised form) September 14, 2007; published electronically February 20, 2008.

<http://www.siam.org/journals/sidma/22-1/63260.html>

[†]Department of Mathematics, University of West Georgia, Carrollton, GA 30118 (jkang@westga.edu).

FIG. 1. $L(2, 1)$ -labeling of C .

chordal graphs and unit interval graphs [9], hypercubes [5, 6, 10], planar graphs [6], and certain classes of Kneser graphs [7]. The list version of the problem has been also investigated [8]. All of the known bounds for these families satisfy the conjecture of Griggs and Yeh.

In this paper, we prove the Griggs–Yeh conjecture for Hamiltonian graphs with maximum degree 3.

THEOREM 1. *If G is a Hamiltonian graph with maximum degree 3, then G has a $L(2, 1)$ -labeling of $\lambda(G) \leq 9$.*

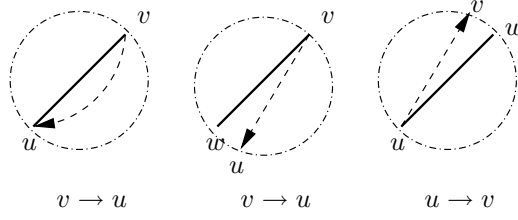
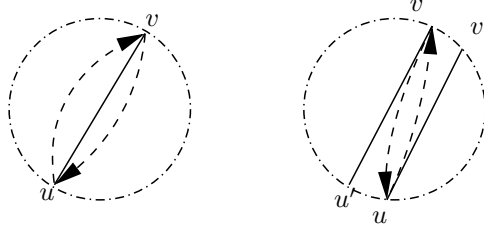
The graphs considered here consist of a spanning cycle and a perfect matching. Georges and Mauro [3] worked on a similar graph, namely a generalized Petersen graph that is two identical cycles plus a matching between them. The best known upper bound for general graphs of maximum degree 3 is 10, by Goncalves [4]; that is just one above the conjecture. The conjectured bound 9 is achieved by the Petersen graph, and no other examples achieving it are known for the maximum degree being 3.

Sections 2, 3, 4, and 5.1 deal exclusively with the case of vertex sets of size 0 (mod 3). In sections 2, 3, and 4, we consider 3-regular Hamiltonian graphs and lay down the foundations for the proof of this case. We first label the vertices based on the Hamiltonian cycle and then explore the properties of an auxiliary graph of badly labeled pairs of vertices. In section 5.1, using the results from sections 2, 3, and 4, we describe an explicit labeling for the case $n(G) \equiv 0 \pmod{3}$. In section 5.2, we adapt the labeling to the other two residue classes.

2. Prelabeling and a graph of bad pairs.

2.1. Definitions and notation. Let G be a 3-regular Hamiltonian graph on n vertices with $n \equiv 0 \pmod{3}$, and let C be a specific Hamiltonian cycle in G . Deleting the edges of C from G leaves a perfect matching, say M . For $u, v \in V$, we will often call u and v G -neighbors, C -neighbors, or the M -neighbors/ M -mates to each other if $uv \in E(G)$, $uv \in E(C)$, or $uv \in M$, respectively. An edge is called *monochromatic* if both its endpoints have the same labels.

Let ρ be a reference coloring that assigns to the vertices of C colors 0, 3, 7, 0, 3, 7, $\dots$, 0, 3, 7 in order along C . Note that ρ is an $L(2, 1)$ -labeling of C . (See Figure 1.) Adding the edges of M to C introduces “bad” pairs of vertices in G . Let H be the graph of these bad pairs; that is, $V(H) = V(G)$, and $uv \in E(H)$ if $\rho(u)$ and $\rho(v)$ violate the conditions of $L(2, 1)$ -labeling for G . Since the colors $\{0, 3, 7\}$ pairwise differ by at least two, $uv \in E(H)$ if and only if u and v receive the same color, and (i) $uv \in M$, (ii) there exists $w \in V$ such that $vw \in M$ and $wu \in E(C)$, or (iii) there exists $w \in V$ such that $vw \in E(C)$ and $wu \in M$. Observe that, for each $v \in V(G)$, there is exactly one vertex $u \in V(G)$ of either (i) or (ii). Hence no vertex of $V(G)$ is an isolated vertex in H , and $\Delta(H) \leq 3$.

FIG. 2. Orienting an edge uv of H .FIG. 3. Double edges $v \rightarrow u, u \rightarrow v$ in D .

In the figures of this paper, we denote the edges of C, M , and H by irregular dashed lines $\cdots\cdots$, thick solid lines — , and regular dashed lines $- - -$, respectively.

Let $X = \rho^{-1}(0)$, $Y = \rho^{-1}(3)$, and $Z = \rho^{-1}(7)$. Let $X_1, X_2, \dots, X_r$ be the vertex sets of the components of $H[X]$, and similarly define $Y_1, Y_2, \dots, Y_s$ and $Z_1, Z_2, \dots, Z_t$ for the vertex sets of the components of $H[Y]$ and $H[Z]$ in colors 3 and 7, respectively. Note that $|X| = |Y| = |Z|$.

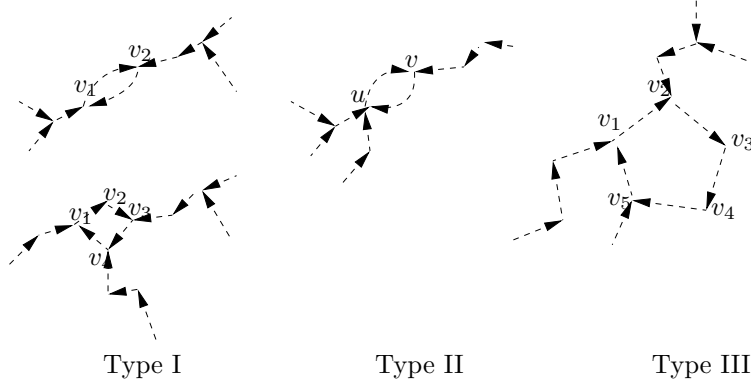
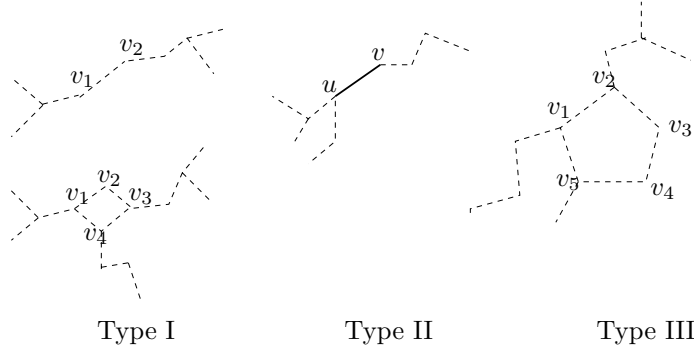
To study the structure of H , we will orient its edges (possibly both ways). For each ordered pair (u, v) with $vu \in E(H)$, orient $v \rightarrow u$ if u occurs as either the M -neighbor of v or one of the C -neighbors of the M -mate of v . (See Figure 2). It is helpful to view the result as a digraph D , with opposed edges vu and uv when we orient $uv \in E(H)$ as both $v \rightarrow u$ and $u \rightarrow v$. A double edge can arise in D in two ways as follows (see Figure 3):

- (i) a monochromatic M -edge uv in G ,
- (ii) a cycle (u, u', v, v') of length 4 with edges alternating between C and M , in which the M -edges are not monochromatic. In particular, suppose that $uu', vv' \in E(C)$, $u'v, v'u \in M$ such that $u'v$ and $v'u$ are not monochromatic. By the pigeonhole principle and the definition of the coloring, at least one of the two pairs $\{u, v\}$ and $\{u', v'\}$ must receive the same color, since we have only three color classes. If $\{u, v\}$ is the monochromatic pair, then we obtain $uv, vu \in E(D)$.

2.2. Structure of the graph H of bad pairs. We show that every component of H is either a tree or a unicyclic graph. For this, we show first that every component of D is a unicyclic digraph, which requires a proposition about the vertex degrees in D .

PROPOSITION 1.

- (a) $d_D^+(v) = 1$ for all $v \in V(D)$.
- (b) $d_D^-(v) \leq 3$ for all $v \in V(D)$, with equality only if v is incident to a double edge in D arising from a monochromatic edge in M .

FIG. 4. Components of the orientation D .FIG. 5. Types of components of H : Type I, Type II, Type III.

Proposition 1 yields the unicyclic structure for components of D .

LEMMA 1 (see Figure 4). *Each component of D is unicyclic (i.e., has exactly one cycle), and each vertex outside the cycle has exactly one out-neighbor toward the cycle and at most two in-neighbors farther from the cycle.*

Proof. Proposition 1(a) implies that every component of D is unicyclic. Proposition 1(b) implies that three edges enter only when one of them is on the cycle. $\square$

By collapsing the double edges in D , we can describe our “bad graph” H precisely. We show the possible types of components in Figure 5, where the drawing of edges follows the conventions adopted in section 2.1.

LEMMA 2 (see Figure 5). *Each component of H is one of the following three types:*

Type I: *Either a tree obtained by collapsing the double edge created by a C , M -alternating 4-cycle in G , or a unicyclic graph with an even cycle.*

Type II: *A tree containing one M -edge of G obtained by collapsing the double edge arising from a monochromatic M -edge.*

Type III: *A unicyclic graph whose cycle has odd length.*

It is illustrated in Figure 6 how a cycle in D (hence in H) can occur. The cycle $(x_1, x_2, x_3, x_4, x_5)$ in $D[X]$ is introduced by a cycle $(x_1, y_2, x_2, y_3, x_3, z_4, x_4, z_5, x_5, y_1)$ in G , where each x_s belongs to X and y_s or z_s is one of the C -neighbors of x_s in $Y \cup Z$.

From the structure of H , a desired coloring for H that yields a $L(2, 1)$ -labeling for G can be specified as follows.

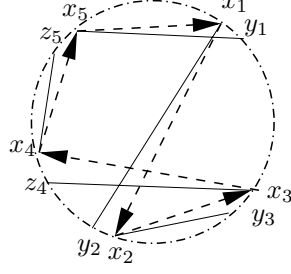


FIG. 6. A cycle $(x_1, x_2, x_3, x_4, x_5)$ in D and a cycle $(x_1, y_1, y_2, y_3, x_3, z_4, x_4, z_5, x_5, y_1)$ in G .

THEOREM 2. *Given a 3-regular Hamiltonian graph G , $|V(G)| \equiv 0 \pmod{3}$, with precoloring ρ and the resulting “bad graph” H (as defined in section 2.1), an $L(2, 1)$ -labeling for G can be obtained by giving ordinary proper colorings of the components of $H[X]$, $H[Y]$, and $H[Z]$ with colors $\{0, 1, 2\}$, $\{3, 4, 5\}$, and $\{7, 8, 9\}$, respectively, satisfying the following conditions:*

- (a) *the M -edges of the components of $H[X]$, $H[Y]$, and $H[Z]$ of Type II receive $\{0, 2\}$, $\{3, 5\}$, and $\{7, 9\}$, respectively, on their endpoints, and*
- (b) *the vertices of X colored with 2 are not adjacent to those of Y colored with 3.*

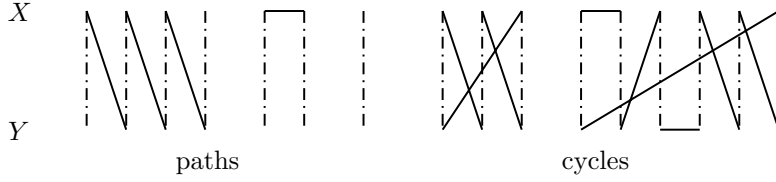
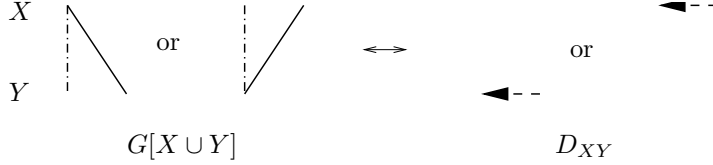
Proof. The only edges of H that are also edges of G are the M -edges in components of Type II. Hence we may use two consecutive colors in modifying ρ on the vertices of a component of Type I or II, except that the colors on the endpoints of an M -edge must differ by at least 2. On components of Type III, we may use any proper 3-coloring. The edges of G between X and Y should satisfy part (b) to ensure that it is an $L(2, 1)$ -labeling. Note that the edges of G between Y and Z automatically satisfy the condition of $L(2, 1)$ -labeling due to the buffer color 6. $\square$

Vertices that have to be colored with 2 or 3 belong to a monochromatic M -edge or to an odd cycle of a component in H (while the rest of the vertices can receive colors from $\{0, 1\}$ or $\{4, 5\}$ that are far apart). To take care of the condition (b) of Theorem 2, it is critical to know how monochromatic M -edges or odd cycles in H are connected to each other. In this sense, we define *trouble part* to be the monochromatic M -edge or the odd cycle of the component in H of Type II or III.

3. Structure of subgraphs. Theorem 2(b) requires us to study the interaction between color classes under ρ in G . In this section, we study how certain edges in D imply the existence of some substructure in G , and vice versa.

3.1. Definitions and notation. When F is a graph and A and B are disjoint subsets of $V(F)$, we write $F[A]$ for the subgraph of F induced by A and $F[A, B]$ for the set of edges of F with one endpoint in A and the other endpoint in B . We will abuse these notations even when F is not a graph (such as matching M) when the context is clear. Continuing with the partition X, Y, Z of $V(G)$ as introduced in section 2, the structure of $G[X \cup Y]$ consists of a perfect matching $C[X, Y]$ and a (not necessarily perfect) matching $M[X \cup Y]$. Hence each component of $G[X \cup Y]$ is an M -alternating cycle or an M -augmenting path (see Figure 7). The same description holds for $G[Y \cup Z]$ and $G[X \cup Z]$ using $C[Y, Z] \cup M[Y \cup Z]$ and $C[X, Z] \cup M[X \cup Z]$, respectively. Note that a vertex of degree 1 in $G[X \cup Y]$ has its M -neighbor in Z .

Let D_{XY} be the subgraph D defined as follows. The vertex set is $X \cup Y$. The edge set is the edges in $D[X \cup Y] = D[X] \cup D[Y]$ whose endpoints have a common neighbor (in G) in Y or X . Such a common neighbor must be an M -neighbor of the

FIG. 7. Examples of components of $G[X \cup Y]$.FIG. 8. Wedges in $G[X \cup Y]$ and directed edges in D_{XY} forced by each other.

tail and a C -neighbor of the head. Precisely,

$$E(D_{XY}) = \{uv \in E(D[X]) \mid uy \in M \text{ and } vy \in C \text{ for some } y \in Y\} \\ \cup \{uv \in E(D[Y]) \mid ux \in M \text{ and } vx \in C \text{ for some } x \in X\}.$$

Similarly, let $V(H_{XY}) = X \cup Y$, and $uv \in E(H_{XY})$ if $uv \in H[X] \cup H[Y]$ have a common neighbor (in G) in Y or X . Note that H_{XY} is the underlying graph of D_{XY} , as before, by collapsing the double edges in D_{XY} .

Remark. H_{XY} differs from $H[X] \cup H[Y]$, and so does D_{XY} from $D[X] \cup D[Y]$. Although $E(H_{XY}) \subset E(H[X] \cup H[Y])$ and $E(D_{XY}) \subset E(D[X] \cup D[Y])$, H_{XY} lacks all the edges of $M[X] \cup M[Y] \subset H[X] \cup H[Y]$ and, more importantly, the edges of $H[X] \cup H[Y]$ whose endpoints have a common neighbor (in G) in Z . From Figure 6, $H[\{x_1, \dots, x_5\}]$ (resp., $D[\{x_1, \dots, x_5\}]$) is an odd cycle $(x_1, x_2, x_3, x_4, x_5)$ in $H[X \cup Y]$ (resp., $D[X \cup Y]$), while $H_{XY}[\{x_1, \dots, x_5\}]$ (resp., $D_{XY}[\{x_1, \dots, x_5\}]$) is a path (x_5, x_1, x_2, x_3) of length 3.

3.2. Relation between subgraphs of H and G . Now, we will show a closer relation between D_{XY} and $G[X \cup Y]$: a given configuration in one forces a certain configuration in the other, and vice versa. The relation holds between D_{YZ} and $G[Y \cup Z]$ analogously. Let us start with a restatement combining the definition of D and that of D_{XY} . A *wedge* from $x_1 \in X$ to $x_2 \in X$ in $G[X \cup Y]$ is a path of length 2 from x_1 to x_2 in which the first edge (incident to x_1) is in $M[X, Y]$ and the second edge (incident to x_2) is in $C[X, Y]$. By this definition, we get the following proposition.

PROPOSITION 2 (see Figure 8). $D_{XY}[X]$ has an edge x_1x_2 (resp., y_1y_2) if and only if $G[X \cup Y]$ has a wedge from $x_1 \in X$ to $x_2 \in X$ (resp., $y_1 \in Y$ to $y_2 \in Y$).

Applying Proposition 2 twice, we can obtain the next proposition that describes configurations occurring in $D_{XY}[Y]$ forced by edges in $D_{XY}[X]$ (or vice versa). This is easily extended to the corresponding results for paths and cycles in $D_{XY}[X]$, Corollary 1, and this will be used later to help produce the desired coloring.

PROPOSITION 3 (see Figure 9). Let $x_1x_2 \in E(D_{XY}[X])$ be a directed edge. If $y_i \in Y$ is the C -neighbor of x_i for $i = 1, 2$, then

- (a) y_2 is the M -neighbor of x_1 , and hence $y_1x_1y_2x_2$ is a path in $G[X \cup Y]$;
- (b) $y_2y_1 \in E(D_{XY}[Y])$.

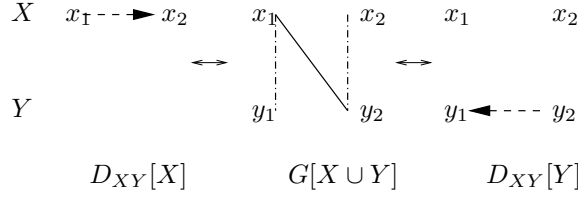


FIG. 9. The path (y_1, x_1, y_2, x_2) in $G[X \cup Y]$ and the directed edge $y_2 y_1$ in $D_{XY}[Y]$ forced by $x_1 x_2$ in $D_{XY}[X]$.

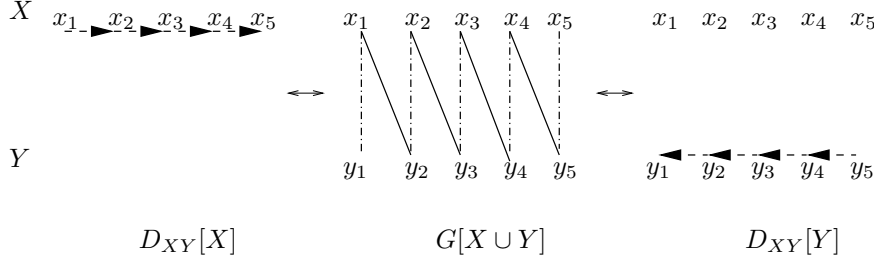


FIG. 10. The path $(y_1, x_1, y_2, x_2, \dots, y_k, x_k)$ in $G[X \cup Y]$ and the directed path $(y_k, y_{k-1}, \dots, y_1)$ in $D_{XY}[Y]$ forced by $(x_1, x_2, \dots, x_k)$.

Proof. We apply both directions of Proposition 2 to work back and forth between D_{XY} and $G[X \cup Y]$.

(a) By Proposition 2, the existence of an edge $x_1 x_2$ in $D_{XY}[X]$ requires a vertex in Y that is both the M -neighbor of x_1 and the C -neighbor of x_2 , and it must be y_2 .

(b) With y_2 being the M -neighbor of x_1 by part (a), we apply Proposition 2 to the path $y_2 x_1 y_1$ in $G[X \cup Y]$ to obtain $y_2 y_1 \in E(D_{XY}[Y])$. $\square$

Proposition 3 immediately extends to transfer paths and cycles between $D_{XY}[X]$ and $D_{XY}[Y]$.

COROLLARY 1 (see Figure 10). *If $x_1 x_2 \dots x_k$ is a directed path in $D_{XY}[X]$, then*

- (a) *there is a path $y_1 x_1 y_2 x_2 \dots y_k x_k$, in $G[X \cup Y]$, where y_i is the C -neighbor of x_i for $1 \leq i \leq k$, and $y_{i+1} \in Y$ is both the M -neighbor of x_i and the C -neighbor of x_{i+1} for $1 \leq i \leq k-1$;*
- (b) *the vertices $y_k, y_{k-1}, \dots, y_1$ defined in part (a) form a path in that order in $D_{XY}[Y]$;*
- (c) *(see Figure 11) if also $x_k x_1 \in D_{XY}[X]$, then also $x_k y_1 \in M$ and $y_1 y_k \in D_{XY}[Y]$.*

We henceforth call the two corresponding cycles $(x_1, x_2, \dots, x_k)$ and $(y_k, y_{k-1}, \dots, y_1)$ in Corollary 1(c) *twin cycles*. A pair of twin cycles of odd length will play an important role later.

Remark. We comment on the significance of choosing $D_{XY}[X]$ and $D_{XY}[Y]$ as subgraphs of D_{XY} instead of choosing $D[X]$ and $D[Y]$ as subgraphs of $D[X \cup Y]$ (see Figure 12). Unlike the results $D_{XY}[X]$ and $D_{XY}[Y]$ in Corollary 1, a cycle of length k in $D[X]$ need not force a cycle of length $2k$ in $G[X \cup Y]$ or a cycle of length k in $D[Y]$, because a cycle in $D[X]$ may use edges that result from wedges that pass through Z . For example, in Figure 6, $(x_1, x_2, x_3, x_4, x_5)$ is a directed cycle in $D[X]$. But it corresponds to only a path $(y_5, x_5, y_1, x_1, y_2, x_2, y_3, x_3)$ of length 7 in $G[X \cup Y]$ plus one isolated edge $x_4 y_4$, where $y_4, y_5 \in Y$ are the C -neighbors of x_4, x_5 , respectively. This forces a directed path (y_3, y_2, y_1, y_5) and an isolated vertex y_4 in $D[Y]$.

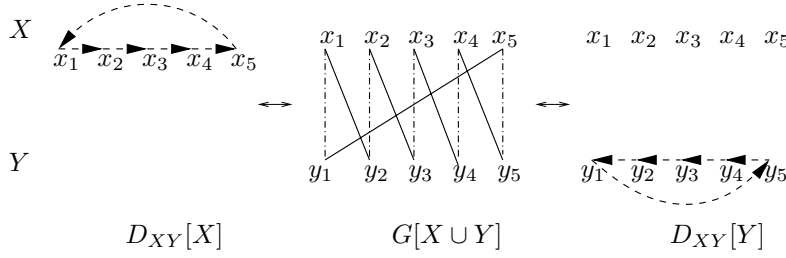


FIG. 11. The cycle $(y_1, x_1, y_2, x_2, \dots, y_k, x_k)$ in $G[X \cup Y]$, and the directed cycle $(y_k, y_{k-1}, \dots, y_1, y_k)$ in $D_{XY}[Y]$ forced by $(x_1, x_2, \dots, x_k, x_1)$.

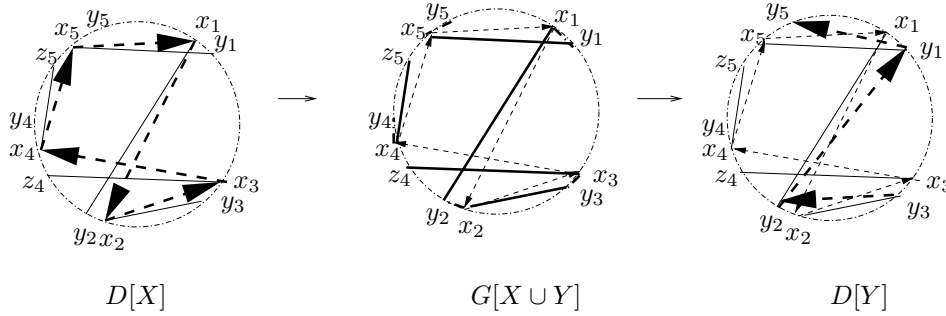


FIG. 12. The significance of choosing $D_{XY}[X]$ instead of $D[X]$.

4. Trouble parts in color classes. To give a coloring as described in Theorem 2, we need to study the interaction between M -edges and odd cycles of $H[X \cup Y]$. Let us introduce some vertex subsets and subgraphs of G and H on $X \cup Y$ that intersect not all but most of such M -edges and odd cycles of $H[X]$ and $H[Y]$.

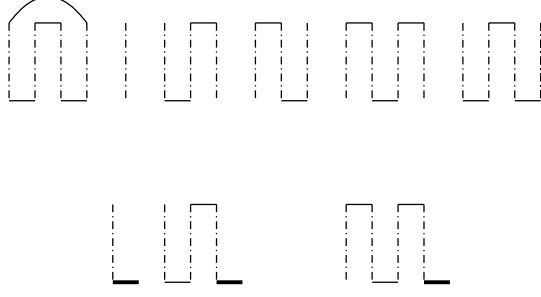
4.1. Definitions and notation. Let X' (resp., Y') be the subset of X (resp., Y) consisting of all the elements $x \in X$ (resp., $y \in Y$) whose M -neighbor is *not* in Y (resp., X). Let G_0 be the subgraph of $G[X \cup Y]$ induced by $X' \cup Y_{X'} \cup Y'$, where $Y_{X'}$ is the set of neighbors of the vertices of X' in Y . Let M_0 and C_0 be the set of M -edges and C -edges, respectively, of G_0 . Note that all the edges of $M[X] \cup M[Y]$ appear in G_0 . Note also that each component of G_0 is a C, M -alternating cycle, a C, M -alternating path, or an isolated vertex v that is in Y' such that the C -neighbor of v in X is not in X' . (See Figure 13.) Note that $M_0 = M[X] \cup M[Y]$ and $M_0 \cap M[X, Y] = \emptyset$.

Let $\mathcal{O}$ be the family of odd cycles in $H[X] \cup H[Y]$ that have a common vertex with $X' \cup Y_{X'} \cup Y'$. It is *not* necessary that all the vertices of such a cycle lie in $X' \cup Y_{X'} \cup Y'$.

4.2. Interaction of trouble parts between color classes. By Lemma 2, two trouble parts of components of H , a cycle in $\mathcal{O}$ and an M_0 -edge, cannot share a vertex. Therefore, a component of G_0 has a common vertex with a cycle in $\mathcal{O}$ only if the component is an M_0 -alternating path and the common vertex is an endpoint of the path incident to a C_0 -edge.

LEMMA 3.

- (a) Let $yy' \in E(D[Y])$ with $y \notin Y', y' \in Y'$. If y' is not incident to an $M[Y]$ -edge, then y' is an isolated vertex in G_0 .
- (b) For a cycle $O \in \mathcal{O}$ in $H[Y]$, there is always $y \in V(O)$ such that $y \in Y'$.

FIG. 13. *Types of components of G_0 .*

- (c) If a cycle O in $H[Y]$ has exactly one vertex from $Y_{X'} \cup Y'$, then the unique vertex is an isolated vertex in G_0 .

Proof. (a) Since $y \notin Y'$, y has its M -neighbor in X , say x . By the uniqueness of the out-neighbor (of y) in H , the C -neighbor of x in Y must be y' . Since y' is not incident to an $M[Y]$ -edge and its C -neighbor x in X is not in X' , the vertex y' must be isolated in G_0 .

(b) Suppose that $y \in V(O) \in \mathcal{O}$ in $H[Y]$ and $y \in Y_{X'} \setminus Y'$. Let $xy \in C$ so that $x \in X'$. Consider the in-neighbor y' of y in O . Then we claim that $y' \in Y'$. Otherwise, the uniqueness of the out-neighbor (of y') in H implies that the M -neighbor of y' in X must be x , which contradicts $x \in X'$.

(c) Let y' be the unique vertex. Then (b) implies that $y' \in Y'$. If y is the in-neighbor of y' in O , the conclusion is immediate from (a). $\square$

LEMMA 4.

- (a) Suppose that a cycle $O \in \mathcal{O}$ in $H[X]$ has exactly one vertex from X' . Then the unique vertex and its C -neighbor in Y form an isolated edge in G_0 .
(b) Let $O \in \mathcal{O}$ be a cycle and xy be an isolated edge in (a), where $x \in V(O)$ and $y \in Y$. If y is a vertex of a cycle $O' \in \mathcal{O}$ in $H[Y]$, the cycle O' has an isolated vertex in G_0 .

Proof. (a) Let $x \in V(O)$ be the unique vertex, and y be its C -neighbor in Y . If x' is the in-neighbor of x in O , then $x' \notin X'$. The uniqueness of the out-neighbor (of x') in H implies that $x'y \in M$ and hence $y \notin Y'$. The M_0 -augmenting path in G_0 beginning with the edge xy has to stop at y .

(b) We have seen that $y \notin Y'$ in the proof of (a). On the other hand, since $O' \in \mathcal{O}$ is a cycle in $H[Y]$, there is a vertex in O' from Y' by Lemma 3(b). Since the cycle O' has vertices in Y' and outside of Y' , we can find an edge of Lemma 3(a), and O' has an isolated vertex in G_0 . $\square$

The following proposition is critical for choosing vertices of colors 2 and 3 from trouble parts in $H[X \cup Y]$.

PROPOSITION 4. $V(G_0)$ has a bipartition A, B with the property that A has at least one vertex in

- (a) every edge of M_0 , and
(b) every cycle of $\mathcal{O}$.

Proof. We have observed that the components of G_0 are C, M -alternating cycles, C, M -alternating paths, and isolated vertices (arising from a cycle of $\mathcal{O}$ in $H[Y]$). Hence G_0 is bipartite. Furthermore, we have observed that $V(G_0)$ intersects cycles in $\mathcal{O}$ only at the ends of its path components. If a path component meets a cycle in $\mathcal{O}$

only at one end, then choose the bipartition on that component to put that end in A . The bipartition can be chosen arbitrarily on C , M -alternating cycles in G_0 , since they do not intersect cycles in $\mathcal{O}$. Isolated vertices can freely be placed in A . It remains to choose the bipartition of the M -augmenting paths in G_0 that have both end cycles in $\mathcal{O}$.

Define an auxiliary multigraph Q whose vertices are the cycles in $\mathcal{O}$. Join O and O' by an edge in Q when G_0 has a path with endpoints in O and O' (loop and multiple edges may arise).

We show that Q has an orientation such that every vertex has positive outdegree except possibly for one vertex of degree 1 in each component of Q as follows. If a component of Q has a cycle, then orient some cycle consistently. For each vertex not on the cycle, there is a neighbor whose distance from the cycle is smaller; orient that edge toward the cycle. If a component of Q has no cycle, then choose a leaf as a root. Orient the component as a directed tree of paths toward this root.

Now we complete the desired bipartition of G_0 as follows: for each oriented edge of Q , the end of the corresponding component of G_0 corresponding to the tail goes into A , and the end corresponding to the head goes into B . Except for the root, all vertices have positive outdegree.

Let O be a cycle corresponding to the root of a tree in Q . If $|V(O) \cap X'| \geq 2$ or $|V(O) \cap (Y_{X'} \cup Y')| \geq 2$, there is one vertex in $V(O)$ from which an M_0 -augmenting path starts and does not end at another cycle in $\mathcal{O}$. This vertex is in A . If O has only one vertex from $X' \cup Y_{X'} \cup Y'$, the cycle O must be in $H[X]$ by Lemma 3(c). Let x be the unique vertex of $V(O) \cap X'$. Lemma 4(b) gives us that $O' \in \mathcal{O}$ and $y, y' \in V(O')$ such that xy is the C_0 -edge connecting O and O' , and y' is an isolated vertex in G_0 . We can swap the A, B on x, y so that $x \in A$, $y \in B$, and then put $y' \in A$. $\square$

5. $L(2, 1)$ -labeling for Hamiltonian graphs with maximum degree 3. In this section, we will use the results from the previous sections to give a proof of the main theorem, Theorem 1 (see section 1).

In section 5.1, we give the proof when $n(G) \equiv 0 \pmod{3}$, and in section 5.2, we handle the other residue classes.

5.1. Vertex sets of size 0 (mod 3). First, we apply Theorem 2 to obtain an $L(2, 1)$ -labeling for 3-regular Hamiltonian graphs. A discussion for Hamiltonian graphs with maximum degree 3 follows after that.

Case 1. G is a 3-regular Hamiltonian graph. We will color componentwise depending on the types given in Lemma 2.

Step 1: Coloring of $H[X]$. If a component $H[X_i]$ is of Type I, a bipartite graph, then color X_i with $\{0, 1\}$.

Let us suppose that $H[X_i]$ is of Type II or III. Then, recall that $H[X_i]$ needs three colors, but deleting one vertex from the trouble part makes $H[X_i]$ 2-colorable.

If the trouble part is covered by A of Proposition 4, which is all $M[X]$ and odd cycles of $\mathcal{O}$, we choose exactly one vertex from A , color it with 2, and color the rest of the vertices of X_i with $\{0, 1\}$. Otherwise, the trouble part has to be an odd cycle all of whose vertices have their M -neighbors in $H[Y]$. By Corollary 1(c), this odd cycle induces its twin cycle in $H_{XY}[Y]$. We pick one arbitrary vertex of the cycle for color 2, then color the rest with $\{0, 1\}$.

Step 2: Coloring of $H[Y]$. If a component $H[Y_i]$ is of Type I, color Y_i with $\{4, 5\}$.

Suppose that $H[Y_i]$ is of Type II or III. We want to choose one vertex from the trouble part for color 3, and color the rest with $\{4, 5\}$.

If the trouble part is covered by A , we choose exactly one vertex from A for color 3. Two vertices colored with 2 and 3 are not adjacent because both of them are in A , which is an independent set in G . A trouble part of the Type II and cycles of $\mathcal{O}$ will always be covered by A by Proposition 4. An odd cycle O that is not contained in $\mathcal{O}$ yields its twin cycle O' in $H[X]$ as in Corollary 1. We chose one vertex from $V(O')$ for color 2 in Step 1, and it has two neighbors in O . Since $|V(O)| \geq 3$, there is a vertex in $V(O)$ that is not adjacent to the vertex of color 2. That is the right vertex for color 3.

Step 3: Coloring of $H[Z]$. For a component $H[Z_i]$ of Type I, color Z_i with $\{8, 9\}$.

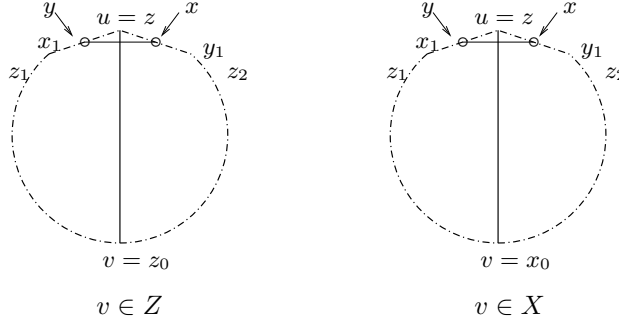
For a component $H[Z_i]$ of Type II or III, we pick an arbitrary vertex from the trouble part for color 7, and color the rest with $\{8, 9\}$.

Case 2. G is a Hamiltonian graph with maximum degree 3. Since G has a Hamiltonian cycle C , each vertex has degree either 2 or 3, and those of degree 2 miss their M -mates. Intuitively, there is less conflict than in the regular case since we have potentially fewer M -edges. In detail, if there are at least 2 vertices of degree 2, we modify the graph as follows. If two vertices of degree 2 are not adjacent in C , we throw in an edge between them. If they are adjacent in C , we duplicate the graph G and connect each of those two vertices to its own copy. Now, we may assume that there is exactly one vertex of degree 2. The fact that there is exactly one degree-2 vertex in G makes the graph “almost” regular; hence it is easier to relate the regular case. Let v be the unique vertex of degree 2 in G . We construct a bad graph H and an orientation D as in section 2.1. All the structure of D is still the same except that (i) v has no out-neighbor, (ii) v can occur as an isolated vertex in D when v does not have any in-neighbor, and (iii) one component of D containing v does not have a cycle. However, the types of H in Lemma 2 are still the same, and hence Theorem 2 is valid for a Hamiltonian graph with maximum degree 3. Observe that v is not a vertex of the trouble part. Hence it does not affect the structure of twin cycles in section 3 nor Proposition 4 in section 4 that are the keys for choosing vertices for colors 2 and 3 in trouble parts—the only nontrivial labeling in Theorem 2.

5.2. Vertex sets of size 1 (mod 3) or 2 (mod 3). In this section, we will construct $L(2, 1)$ -labelings for Hamiltonian graphs with maximum degree 3 of order 1 (mod 3) or 2 (mod 3). We add a few vertices and matching edges to a given graph G to obtain an “almost” supergraph $\widehat{G}$ whose vertex-set is of size 0 (mod 3). Then we give an $L(2, 1)$ -labeling to the supergraph according to the labeling in section 5.1 and rearrange the labeling for G .

The rearrangement will occur locally at the vertices at distance 1 or 2 from the added vertices. In particular, we will have cases depending on the color classes (in ρ) of the relevant M -neighbors. Hence having vertices of degree 2 near the added vertices reduces the number of cases that we have to consider. In the following proof, we will assume that all the vertices of G have M -mates, instead of always writing “if it exists” next to the color class of a possible M -neighbor being considered.

Let us start by reminding readers of the “freedom of the coloring of $H[Z]$ ” (FCZ) from Step 3 in section 5.1: For a component of $H[Z]$ of Type II or III, we choose a vertex to have the color 7 from the trouble part arbitrarily. Moreover, an $L(2, 1)$ -labeling of $H[Z]$ using $\{7, 8, 9\}$ is independent of the labeling on $X \cup Y$, since our $L(2, 1)$ -labeling of $G[X \cup Y]$ does not use color 6.

FIG. 14. $\widehat{G}$ on $V(G) \cup \{x, y\}$.

Proof of Theorem 1 for $1 \pmod{3}$. Let uv be a chord in a Hamiltonian cycle C of G . We define a related graph $\widehat{G}$ on a vertex-set of size $0 \pmod{3}$: Subdivide two edges of C that are incident to u , and make the two new vertices adjacent in $\widehat{G}$. (See Figure 14.) Now $\widehat{G}$ has a $L(2, 1)$ -labeling as in section 5.1, say $\widehat{f}$. We will put $\widehat{}$ on top of the names of sets or graphs precisely defined for graph of order divisible by 3 to denote the corresponding sets or graphs for $\widehat{G}$ such as $\widehat{M}, \widehat{X}, \widehat{H}$, etc. We will call the resulting subsets $\widehat{X}, \widehat{Y}, \widehat{Z}$ the *pre-color classes* of the vertices. By symmetry, we may assume that u belongs to $\widehat{Z}$ and that v is contained in $\widehat{Z} \cup \widehat{X}$. We name u, v and their nearby vertices so that x and y are vertices not in G . We also rename u as z and rename v as z_0 if $v \in \widehat{Z}$ and x_0 if $v \in \widehat{X}$ (see Figure 14). Name the vertices along the modified Hamiltonian cycle so that they are $z_1, x_1, y, z, x, y_1, z_2$ in order around z . We have named them so that $z_1, z, z_2 \in \widehat{Z}$, $x_1, x \in \widehat{X}$, and $y, y_1 \in \widehat{Y}$.

For $\widehat{f}$ to be an $L(2, 1)$ -labeling not only for $\widehat{G}$ but also restricted to an $L(2, 1)$ -labeling for G , we need to check the following conditions:

- (a) $|\widehat{f}(z) - \widehat{f}(x_1)| \geq 2$, $|\widehat{f}(z) - \widehat{f}(y_1)| \geq 2$,
- (b) $\widehat{f}(x_1) \neq \widehat{f}(y_1)$,
- (c) $\widehat{f}(z) \neq \widehat{f}(z_1)$, $\widehat{f}(z) \neq \widehat{f}(z_2)$, and
 $\widehat{f}(z) \neq \widehat{f}(z_3)$, $\widehat{f}(z) \neq \widehat{f}(z_4)$ if $x_1 z_3, y_1 z_4 \in M$,
- (d) $\widehat{f}(x_1) \neq \widehat{f}(x_0)$ when $v = x_0 \in X$.

Some of the conditions are guaranteed, and the others need a relabeling on some vertices. During the relabeling, new requirements can arise on $\widehat{f}$ at nearby vertices of the newly labeled vertices. We give an explicit relabeling f , according to the conditions (a)–(d) above, that satisfies all the requirements as follows.

- Conditions (a) and (b) are guaranteed by the precolor classes.
- Condition (c): We relabel z as $f(z) = 6$. This relabeling requires

- (1) $\widehat{f}(y_1) \neq 5$,
- (2) $\widehat{f}(z_0) \neq 7$ when $v = z_0 \in Z$.

For (2), FCZ allows us to assume that $\widehat{f}(z_0) = 9$ and $\widehat{f}(z) = 7$. For (1), suppose that $\widehat{f}(y_1) = 5$. Let $\widehat{Y}_j$ induce the component of $\widehat{H}[\widehat{Y}]$ containing y_1 , and let y' be the vertex of color 3 if $\widehat{H}[\widehat{Y}_j]$ is of Type II or Type III. If $\widehat{H}[\widehat{Y}_j]$ is a bipartite graph as Type I, or if y_1 belongs to a bipartite subgraph $\widehat{H}[\widehat{Y}_j - y']$ in case that $\widehat{H}[\widehat{Y}_j]$ is of Type III, we exchange the bipartition of $\widehat{H}[\widehat{Y}_j - y']$ to change the color on y_1 to 4.

Suppose that $\widehat{H}[\widehat{Y}_j]$ is of Type II. Let $y'y''$ be the M -edge in the component. In the subgraph $\widehat{H}[\widehat{Y}_j - y']$, we denote the component containing y'' by $T(y'')$. If y_1 belongs to components $\widehat{H}[\widehat{Y}_j - y'] \setminus T(y'')$, then we exchange the bipartition of $\widehat{H}[\widehat{Y}_j - y] \setminus T(y'')$ as before. If $y_1 \in V(T(y''))$, we cannot simply exchange the bipartition of $T(y'')$ because y'' must receive the color 5. Consider the component $\mathcal{T}$ of the auxiliary graph Q , as in the proof of Proposition 4, corresponding to the component in G_0 containing the edge $y'y''$. If $y'y''$ corresponds to an edge of a cycle in $\mathcal{T}$, we give a consistent orientation on the cycle so that $y'' \in A$. If $y'y''$ corresponds to an edge of a tree in $\mathcal{T}$, we flip the orientation along the longest path containing $y'y''$ so that $y'' \in A$. This would not uncover any odd cycle by A ; the possible odd cycle uncovered by A is the one that corresponds to a leaf of $\mathcal{T}$. Observe that this cycle is guaranteed to be covered by A , not by an orientation, as shown at the end of the proof of Proposition 4; we used an extra vertex from G_0 or an isolated vertex of the cycle, or its adjacent cycle in Q .

- Condition (d): Observe that (x_0, x, x_1) is a path in a component $\widehat{H}[\widehat{X}_i]$. If deleting x separates x_0 and x_1 in $\widehat{H}[\widehat{X}_i]$, we can choose colors freely on the vertices of $\widehat{H}[\widehat{X}_i - x]$ so that $f(x_1) \neq f(x_0)$. If deleting x does not separate x_0 and x_1 in $\widehat{H}[\widehat{X}_i - x]$, the path (x_0, x, x_1) is part of the cycle of $\widehat{H}[\widehat{X}_i]$. If it is an odd cycle, the path from x_0 to x_1 in $\widehat{H}[\widehat{X}_i - x]$ has odd length and guarantees that $\widehat{f}(x_1) \neq \widehat{f}(x_0)$. If it is an even cycle, all the vertices of $\widehat{X}_i$ receive colors from $\{0, 1\}$ under $\widehat{f}$. Hence we can relabel

$$f(x_1) = 2,$$

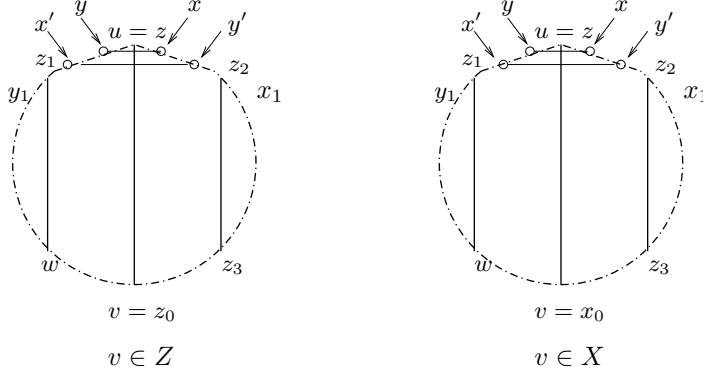
as long as

$$\widehat{f}(y_2) \neq 3 \quad \text{if } x_1 y_2 \in M \text{ for some } y_2 \in Y.$$

If $\widehat{f}(y_2) = 3$, then y_2 is a vertex of an odd cycle in $\widehat{G}$. Observe that (y_2, y, y_1) is a path in the cycle by the uniqueness of the out-neighbor in $\widehat{D}$. In G , this odd cycle is destroyed, and we can color the component left in G with only colors 4 and 5. $\square$

Proof of Theorem 1 for 2 (mod 3). We will relate $\widehat{G}$ and $\widehat{f}$ again. We consider two cases depending on whether G has a chord st that induces a cycle of length 1 (mod 3) with a path in C ending at s and t . Observe that the other path of C with endpoints s and t induces a cycle of length 0 (mod 3). Hence, if no chord induces a cycle of length 1 (mod 3), every chord induces cycles of length only 2 (mod 3) with both paths in C ending at the vertices of the chord.

Case 1. G has a chord that induces a cycle of length 1 (mod 3) with a path in C . Let st be the chord, and let u be the C -neighbor of s at the side of the cycle of 0 (mod 3). For $\widehat{G}$ and $\widehat{f}$, we name the vertices in a similar way as before, except that we subdivide the two new edges incident to u twice so that four new vertices and two new M -edges are introduced in a way that a pair of vertices closer to u forms one edge, and the other pair the other edge: x', y, x, y' are new vertices, $xy, x'y' \in E(\widehat{G})$. Note that s is z_1 or z_2 , so that one of z_1 and z_2 has its M -neighbor t in Z , say z_3 . We set $z_2 z_3 \in M$ and $z_1 w \in M$ (while we do not know the precolor class of w); see Figure 15. For $\widehat{f}$ to be an $L(2, 1)$ -labeling for G , we need to check on some vertices as in the case of 1 (mod 3). We give detailed relabeling f for the cases $v = z_0 \in Z$ and $v = x_0 \in X$ satisfying the requirements.

FIG. 15. $\widehat{G}$ on $V(G) \cup \{x, y, x', y'\}$.

Case $v = z_0 \in Z$. We need to check the following conditions:

- (a) $\widehat{f}(z) \neq \widehat{f}(x_1)$, $\widehat{f}(z) \neq \widehat{f}(y_1)$,
- (b) $|\widehat{f}(z) - \widehat{f}(z_1)| \geq 2$, $|\widehat{f}(z) - \widehat{f}(z_2)| \geq 2$,
- (c) $\widehat{f}(z) \neq \widehat{f}(z_3)$, $\widehat{f}(z) \neq \widehat{f}(w)$ if $w \in Z$,
- (d) all the colors of $\widehat{f}(z_0)$, $\widehat{f}(z_1)$, $\widehat{f}(z_2)$ are distinct.

A relabeling f goes as follows:

- Condition (a) is guaranteed by the precolor classes.
- Conditions (b) and (c): Assign

$$f(z) \in \{0, 1, 2, 3, 4, 5\} \setminus \{\widehat{f}(x_1), \widehat{f}(x_2), \widehat{f}(y_1), \widehat{f}(y_2), \widehat{f}(w)\}$$

to ensure that the color of u is distinct from the colors of the vertices at distance 2 in G . Also, the color $f(u)$ is two apart from the colors of its G -neighbors, z_0 , z_1 , and z_2 .

- Condition (d): Our aim is having all distinct colors for $f(z_0)$, $f(z_1)$, and $f(z_2)$ in $\{7, 8, 9\}$. Note that the edges zz_0 and z_2z_3 are in distinct components because both of them are trouble parts. Let $\widehat{Z}_i$ and $\widehat{Z}_j$ be the components containing zz_0 and z_2z_3 , respectively. If $z_1 \in \widehat{Z}_i$, a proper coloring of $\widehat{H}[\widehat{Z}_i - z]$ with $\{8, 9\}$ gives $f(z_0) \neq f(z_1)$ in $\{8, 9\}$. We can set $\widehat{f}(z_2) = 7$ by FCZ. Suppose that $z_1 \notin \widehat{Z}_i$. We may assume $\widehat{f}(z_1) \neq \widehat{f}(z_2)$ by FCZ, regardless of whether z_1 is in $\widehat{Z}_j$ or not. If $\{\widehat{f}(z_1), \widehat{f}(z_2)\} = \{8, 9\}$, we will assume $\widehat{f}(z_0) = 7$ on $\widehat{Z}_i$ by FCZ. If $\{\widehat{f}(z_1), \widehat{f}(z_2)\} \ni 7$, relabel on a bipartite graph $\widehat{H}[\widehat{Z}_i - z]$, if necessary, to obtain f in colors $\{8, 9\}$ such that $f(z_0) \in \{8, 9\} \setminus \{\widehat{f}(z_1), \widehat{f}(z_2)\}$.

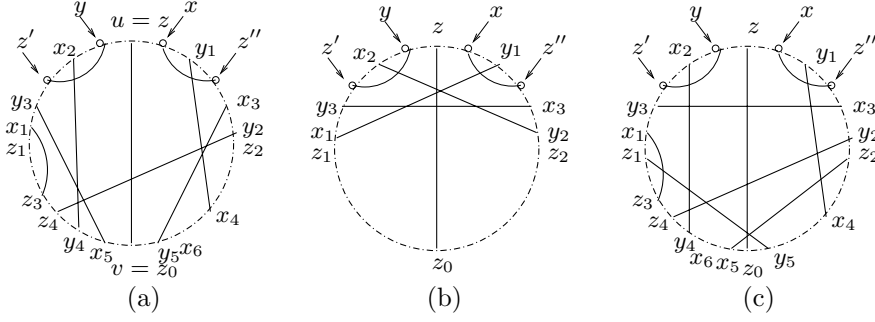
Case $v = x_0 \in X$. We need to check the following conditions:

- (a) $\widehat{f}(z) \neq \widehat{f}(x_1)$, $\widehat{f}(z) \neq \widehat{f}(y_1)$,
- (b) $|\widehat{f}(z) - \widehat{f}(z_1)| \geq 2$, $|\widehat{f}(z) - \widehat{f}(z_2)| \geq 2$,
- (c) $\widehat{f}(z) \neq \widehat{f}(z_3)$, $\widehat{f}(z) \neq \widehat{f}(w)$ if $w \in Z$,
- (d) $\widehat{f}(z_1) \neq \widehat{f}(z_2)$.

A relabeling f is as follows:

- Condition (a) is guaranteed by the precolor classes.
- Conditions (b) and (c): If $w \in \widehat{X} \cup \widehat{Z}$, assign

$$f(u = z) \in \{3, 4, 5\} \setminus \{\widehat{f}(y_1), \widehat{f}(y_2)\}.$$

FIG. 16. $\widehat{G}$ on $V(G) \cup \{x, y, z', z''\}$.

Suppose that $w \in \widehat{Y}$. First, we observe that $x_0x \in E(\widehat{D})$ and xx' is a double edge in $\widehat{D}[\widehat{X}]$. Hence, the component of $\widehat{H}$ containing x_0 is of Type I, and we may assume that

$$\widehat{f}(x_0) = 0.$$

We want to relabel

$$f(z) = 2,$$

which requires

$$\widehat{f}(x_1) \neq 2.$$

If $\widehat{f}(x_1) = 2$, then x_1 is either in a twin cycle or a vertex of the trouble part of Type II or III that is chosen for $A \subset X'$ as in Proposition 4. If it is the former, we can freely choose another vertex of the cycle to receive color 2. Suppose the case is the latter. If more than one vertex of the trouble part is in X' , we can choose another vertex to be in A by flipping the orientation of the auxiliary graph Q in the proof of Proposition 4, if needed. (A similar iteration was done to Y in the case of $1 \pmod{3}$ for the condition (c).) Note that this takes care of Type II. Suppose that $O \in \mathcal{O}$ is the odd cycle containing x_1 , and x_1 is the only vertex of O in X' . We will find a vertex of O that has an M -neighbor in Y , yet is not adjacent to a vertex colored 3, so that it can receive color 2. Let (x_3, x_2, x_1) be a path in O and $\widehat{D}[\widehat{X}]$, and let $y'_2, y'_1 \in Y$ be the C -neighbors of x_2, x_1 , respectively. The edge $x_1y'_1$ is an isolated edge in G_0 by Lemma 4(a), and furthermore, $x_3y'_2, x_2y'_1 \in M$. Observe that $\widehat{f}(y'_1) \neq 3$ because $\widehat{f}(x_1) = 2$, and that $\widehat{f}(y'_2) \neq 3$ because $y_2 \notin Y_{X'} \cup Y'$. The vertex x_2 , whose neighbors in Y are y'_1 and y'_2 , is the desired one.

Case 2. Every chord of G induces cycles of length $2 \pmod{3}$ with a path in C . We relate, again, $\widehat{G}$ and $\widehat{f}$, and name the vertices in a similar way as before except that we subdivide four consecutive edges of C near u , and make the first two new vertices adjacent and the last two vertices adjacent in $\widehat{G}$: The vertices along $\widehat{C}$ are $x_1, y_3, z', x_2, y, z, x, y_1, z'', x_3, y_2$ in order around z , where $u = z$ as before, y, z', x, z'' are new vertices, $z'y, xz'' \in E(\widehat{G})$. (See Figure 16.)

The assumption that every chord of G induces cycles of length $2 \pmod{3}$ determines the precolor class of the M -neighbor of each vertex. The vertex v always belongs to Z so that $v = z_0$, and no other vertex has its M -neighbor in its own

precolor class. Hence there is no component of Type II in $\widehat{H}$, except for the one containing zz_0 . Having no M -edge in $\widehat{H}[\widehat{X}] \cup \widehat{H}[\widehat{Y}]$ gives us a convenient property that is a similar but weaker version of FCZ, namely “weak freedom of the coloring of $H[X]$ or $H[Y]$ ” (WFCX or WFCY, respectively): We can freely exchange colors 0 and 1, or 4 and 5, respectively.

We consider cases depending on whether there is an M -edge in $G[x_1, y_3, x_2, z, y_1, z, x_3, y_2]$. Observe that the possible M -edges are $x_i y_i$ for $i = 1, 2, 3$.

Case of no M -edge in $G[x_1, y_3, x_2, z, y_1, x_3, y_2]$. We know the precolor classes of the M -neighbors of $x_1, x_2, x_3, y_1, y_2, y_3$. Set $x_1 z_3, x_2 y_4, x_3 y_5, y_1 x_4, y_2 z_4, y_3 x_5 \in M$. (See Figure 16(a).) We need to check the following conditions:

- (a) $|\widehat{f}(x_2) - \widehat{f}(y_3)| \geq 2$,
- (b) $\widehat{f}(x_2) \neq \widehat{f}(x_1), \widehat{f}(x_2) \neq \widehat{f}(x_5)$,
- (c) $\widehat{f}(y_3) \neq \widehat{f}(y_4)$,
- (d) $|\widehat{f}(y_1) - \widehat{f}(x_3)| \geq 2$,
- (e) $\widehat{f}(y_1) \neq \widehat{f}(y_2), \widehat{f}(y_1) \neq \widehat{f}(y_5)$,
- (f) $\widehat{f}(x_3) \neq \widehat{f}(x_4)$.

A relabeling f goes as follows:

- Conditions (a)–(c): Recolor

$$f(y_3) = 6, \quad f(x_2) = 3.$$

This requires

$$\widehat{f}(y_4) = 5.$$

If $\widehat{f}(y_4) = 4$, then we can exchange colors 4 and 5 by WFCY. For the case $\widehat{f}(y_4) = 3$, observe that (y_4, y, y_3) is a path in $\widehat{D}[\widehat{Y}]$. The labeling $\widehat{f}(y_4) = 3$ implies that y_4 is a vertex of an odd cycle in $\widehat{D}[\widehat{Y}]$, and so are y and y_3 by the uniqueness of out-neighbors. Since $y \notin V(G)$, this odd cycle is destroyed in H , and we can color the component left in G with $\{4, 5\}$ so that $f(y_4) = 5$.

- Conditions (d)–(f): We will recolor y_1 and x_3 depending on $\widehat{f}(y_2)$ and $\widehat{f}(y_5)$.

To begin with, note that $y_5 y_2 \in E(\widehat{D})$, and hence $\widehat{f}(y_2) \neq \widehat{f}(y_5)$.

If $\{\widehat{f}(y_2), \widehat{f}(y_5)\} = \{4, 5\}$, recolor

$$f(x_3) \in \{0, 1, 2\} \setminus \{\widehat{f}(x_4), \widehat{f}(x_6)\}, \quad f(y_1) = 6,$$

where x_6 is the C -neighbor of y_5 . This requires

$$\widehat{f}(z) \neq 7.$$

By FCZ, we may assume that $\widehat{f}(z) = 9, \widehat{f}(z_0) = 7$.

If $\{\widehat{f}(y_2), \widehat{f}(y_5)\} \ni 3$, we may assume that $\{\widehat{f}(y_2), \widehat{f}(y_5)\} = \{3, 4\}$ by WFCY.

Recolor

$$f(x_3) = 6, \quad f(y_1) = 2.$$

This requires

$$\widehat{f}(x_4) = 0.$$

If $\widehat{f}(x_4) = 1$, then we can exchange colors 0 and 1 by WFCX. If $\widehat{f}(x_4) = 2$, then x_4 is a vertex of an odd cycle, and the path (x_4, x, x_3) in $\widehat{D}[\widehat{X}]$ is a part of the cycle. With an argument along the line of the case $\widehat{f}(y_4) = 3$ above, we may assume $f(x_4) \in \{0, 1\}$ so that $f(x_4) = 0$.

It is worth mentioning that the previous assumption that none of the possible M -edges $x_i y_i$, $1 \leq i \leq 3$, occurs is critical; two vertices of color 6 under f can be at distance 1 or 2. If there are seven consecutive vertices along C without induced M -edges, it falls into the previous case. Now, assume that we cannot find seven such vertices.

Case that $G[x_1, y_3, x_2, z, y_1, x_3, y_2]$ has an M -edge. If $x_i y_i \in M$ for all $i = 1, 2, 3$, (see Figure 16(b)), then we assign

$$(3) \quad \text{to } x_1, y_3, x_2, y_1, x_3, y_2 \quad \text{colors } 2, 4, 0, 5, 1, 3 \quad \text{in order under } f.$$

Note that $z_0 = z_1 = z_2$ can occur, in which case we can set $f(z) = 7$ and $f(z_0) = 6$.

If exactly one of $x_i y_i$ is missing, say $x_j y_j$, then recolor

$$f(x_j) = f(y_j) = 6,$$

while keeping the other colors in (3) the same. However, in case $x_3 y_3 \notin M$, we recolor $f(y_1) = 4$ to have $|f(y_1) - f(x_3)| \geq 2$.

If there is exactly one M -edge, we may assume $x_3 y_3 \in M$: Suppose that we have seven consecutive vertices along C , say $u_1, u_2, \dots, u_7$, that induce one M -edge. We want that $u_2 u_6 \in M$. If either $u_1 u_5 \in M$ or $u_3 u_7 \in M$, say $u_3 u_7 \in M$, consider shifted seven vertices $u_2, u_3, \dots, u_8$, where u_8 is the vertex next to u_7 in C . If u_8 does not introduce an M -edge with any of $u_2, u_3, \dots, u_7$, then $u_2, u_3, \dots, u_8$ are the desired ones. If u_8 does, then it is u_4 , and $G[u_2, u_3, \dots, u_8]$ has exactly two M -edges, $u_3 u_7$ and $u_4 u_8$. We were in this situation before.

As before, the precolor classes of M -neighbors are determined. Set $x_1 z_3, x_2 y_4, y_1 x_4, y_2 z_4, z_1 y_5, z_2 x_5 \in M$, where z_1, z_2 are the C -neighbors of x_1, y_2 , respectively. (See Figure 16(c).) We need to check the following:

- (a) $|\widehat{f}(x_2) - \widehat{f}(y_3)| \geq 2$,
- (b) $\widehat{f}(x_2) \neq \widehat{f}(x_1), \widehat{f}(x_2) \neq \widehat{f}(x_3)$,
- (c) $|\widehat{f}(y_1) - \widehat{f}(x_3)| \geq 2$,
- (d) $\widehat{f}(y_1) \neq \widehat{f}(y_2), \widehat{f}(y_1) \neq \widehat{f}(y_3)$.

Observe that (x_4, x, x_3, x_1) is a path in $\widehat{D}[\widehat{X}]$, and that the in-degrees of x, x_3 and x_1 in $\widehat{D}$ are 1. Hence, if one of those vertices is in a cycle, then the path is also in the cycle, by the uniqueness of the in-neighbor and also by the uniqueness of the out-neighbor in general. Since $y \notin Y$, we can assume that

$$(4) \quad \widehat{f}(x_1), \widehat{f}(x_3), \widehat{f}(x_4) \in \{0, 1\}.$$

Applying the same argument to (y_4, y, y_3, y_2) in $\widehat{D}[\widehat{Y}]$, we assume that

$$(5) \quad \widehat{f}(y_2), \widehat{f}(y_3), \widehat{f}(y_4) \in \{4, 5\}.$$

For the conditions (c) and (d), recolor

$$f(y_1) = 6.$$

Moreover, this recoloring and the assumptions (4) and (5) enable us to use color 2 freely on x_2 or x_3 , as long as x_2 receives a different color from the C -neighbor of y_4 , say x_6 . We recolor x_2 or x_3 depending on $\widehat{f}(x_6)$:

$$f(x_2) = 2 \quad \text{if } \widehat{f}(x_6) \in \{0, 1\}$$

or

$$f(x_3) = 2 \quad \text{and} \quad f(x_2) \in \{0, 1\} \setminus \widehat{f}(x_1) \quad \text{if } \widehat{f}(x_6) = 2. \quad \square$$

Remarks. Arguments of sections 2, 3, 4, and 5.1 can easily be extended to the case that a 3-regular graph G has a 2-factor, each cycle of which is $0 \pmod{3}$: We can color each cycle with only 0, 3, 7, as in the Hamiltonian case. Then, we have the same bad graph H and the same subgraph of G between color classes. This does not work when the 2-factor has cycles of length $1 \pmod{3}$ or $2 \pmod{3}$ because such cycles cannot be labeled with only 0, 3, 7. Our $L(2, 1)$ -labeling in section 5.1 depends only on H and its interaction with G .

The author, with Douglas West, is currently working on extending the result in this paper to 3-regular graphs. As a by-product of this work, an explicit structure of $\mathcal{O}$ in $H[Y]$ in section 4 has been found.

Acknowledgments. This work was done as a part of the author's Ph.D thesis at the University of Illinois at Urbana-Champaign. She thanks Professor D. West for fruitful conversations and his help with improving the manuscript. She also thanks the referees for helpful comments and updated progress toward the problem.

REFERENCES

- [1] G. J. CHANG AND D. KUO, *The $L(2, 1)$ -labelling problem on graphs*, SIAM J. Discrete Math., 9 (1996), pp. 309–316.
- [2] Z. FÜREDI AND J.-H. KANG, *$L(2, 1)$ -Labeling and Packing of Bipartite Graphs*, preprint, University of Illinois at Urbana-Champaign, 2004.
- [3] J. P. GEORGES AND D. W. MAURO, *On generalized Petersen graphs labeled with a condition at distance two*, Discrete Math., 259 (2002), pp. 311–318.
- [4] D. GONCALVES, *On the $L(p, 1)$ -labelling of graphs*, in Discrete Math. Theoret. Comput. Sci., Proceedings of the European Conference on Combinatorics, Graph Theory, and Applications, 2005, pp. 81–86.
- [5] J. R. GRIGGS AND R. K. YEH, *Labelling graphs with a condition at distance 2*, SIAM J. Discrete Math., 5 (1992), pp. 586–595.
- [6] K. JONAS, *Graph colorings analogues with a condition at distance two: $L(2, 1)$ -labeling and list λ -labelings*, Ph.D. thesis, Department of Mathematics, University of South Carolina, Columbia, SC, 1993.
- [7] J.-H. KANG, *Labeling for Kneser Graph*, preprint, University of Illinois at Urbana-Champaign, 2004.
- [8] D. KRÁL' AND R. ŠKREKOVSKI, *A theorem about the channel assignment problem*, SIAM J. Discrete Math., 16 (2003), pp. 426–437.
- [9] D. SAKAI, *Labeling chordal graphs: Distance two condition*, SIAM J. Discrete Math., 7 (1994), pp. 133–140.
- [10] M. A. WHITTLESEY, J. P. GEORGES, AND D. W. MAURO, *On the λ -number of Q_n and related graphs*, SIAM J. Discrete Math., 8 (1995), pp. 499–506.

3-FACIAL COLORING OF PLANE GRAPHS*

FRÉDÉRIC HAVET[†], JEAN-SÉBASTIEN SERENI[‡], AND RISTE ŠKREKOVSKI[§]

Abstract. A plane graph is ℓ -facially k -colorable if its vertices can be colored with k colors such that any two distinct vertices on a facial segment of length at most ℓ are colored differently. We prove that every plane graph is 3-facially 11-colorable. As a consequence, we derive that every 2-connected plane graph with maximum face-size at most 7 is cyclically 11-colorable. These two bounds are just one higher than those that are proposed by the $(3\ell + 1)$ -conjecture and the cyclic conjecture.

Key words. facial coloring, cyclic coloring, plane graph

AMS subject classification. 05C15

DOI. 10.1137/060664124

1. Introduction. The concept of facial colorings, introduced by Král', Madaras, and Škrekovski [11, 12], extends the well-known concept of cyclic colorings. A *facial segment* of a plane graph G is a sequence of vertices in the order obtained when traversing a part of the boundary of a face. The *length* of a facial segment is the number of its edges. Two vertices u and v of G are ℓ -*facially adjacent* if there exists a facial segment of length at most ℓ between them. An ℓ -*facial coloring* of G is a function which assigns a color to each vertex of G such that any two distinct ℓ -facially adjacent vertices are assigned with distinct colors. Notice that a vertex of G that is ℓ -facially adjacent to itself does not prevent G from being colored. A graph admitting an ℓ -facial coloring with k colors is called ℓ -*facially k -colorable*.

The following conjecture is called the $(3\ell + 1)$ -conjecture [11].

CONJECTURE 1 (see Král', Madaras, and Škrekovski). *Every plane graph is ℓ -facially colorable with $3\ell + 1$ colors.*

Observe that the bound offered by Conjecture 1 is tight: as shown by Figure 1, for every $\ell \geq 1$, there exists a plane graph that is not ℓ -facially 3ℓ -colorable.

Conjecture 1 can be considered as a counterpart for ℓ -facial coloring of the following famous conjecture by Ore and Plummer [13] concerning the cyclic coloring. A plane graph G is *cyclically k -colorable* if it admits a vertex coloring with k colors such that any two vertices incident to the same face are assigned distinct colors.

CONJECTURE 2 (see Ore and Plummer). *Every plane graph is cyclically $\lfloor \frac{3\Delta^*}{2} \rfloor$ -colorable, where Δ^* is the size of the largest face of G .*

Note that Conjecture 1 implies Conjecture 2 for odd values of Δ^* . The best known result towards Conjecture 2 has been obtained by Sanders and Zhao [16], who

*Received by the editors July 2, 2006; accepted for publication (in revised form) November 2, 2007; published electronically February 20, 2008. This work was partially supported by the Égide ECO-NET project 16305SB.

<http://www.siam.org/journals/sidma/22-1/66412.html>

[†]Project MASCOTTE, I3S (CNRS–UNSA)—INRIA, 2004 Route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex, France (havet@sophia.inria.fr). This author's work was partially supported by the European project ist fet Aeolus.

[‡]Institute for Theoretical Computer Science (ITI) and Department of Applied Mathematics (KAM), Faculty of Mathematics and Physics, Charles University, Malostranské náměstí 25, 118 00 Prague, Czech Republic (sereni@kam.mff.cuni.cz). This author's work was supported by the European project IST FET AEOLUS.

[§]Department of Mathematics, University of Ljubljana, Jadranska 19, 1111 Ljubljana, Slovenia (bluesky2high@yahoo.com). This author's research was supported in part by the Ministry of Science and Technology of Slovenia, Research Program P1-0297.

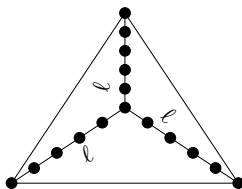


FIG. 1. The plane graph $G_\ell = (V, E)$: each thread represents a path of length ℓ . The graph G_ℓ is not ℓ -facially 3ℓ -colorable: any two vertices are ℓ -facially adjacent, and therefore any ℓ -facial coloring must use $|V| = 3\ell + 1$ colors.

proved the bound $\lceil \frac{5\Delta^*}{3} \rceil$.

Define $f_c(x)$ to be the minimum number of colors needed to cyclically color every plane graph of maximum face size x . The value of $f_c(x)$ is known for $x \in \{3, 4\}$: $f_c(3) = 4$ (the problem of finding $f_c(3)$ being equivalent to the four color theorem proved by Appel and Haken [1]) and $f_c(4) = 6$ (see [3, 5]). It is also known that $f_c(5) \in \{7, 8\}$ and $f_c(6) \leq 10$ [6], and that $f_c(7) \leq 12$ [4].

Conjecture 1 is trivially true for $\ell = 0$, and is equivalent to the four color theorem for $\ell = 1$. It is open for all other values of ℓ . As noted by Král', Madaras, and Škrekovski [11], if Conjecture 1 were true for $\ell = 2$, it would have several interesting corollaries. Besides giving the exact value of $f_c(5)$ (which would then be 7), it would allow the upper bound on the number of colors needed to 1-diagonally color every plane quadrangulation to decrease from 16 to 14 (by applying a method from [11]). (For more details on this problem, consult [9, 14, 15, 11].) It would also imply Wegner's conjecture on 2-distance colorings (i.e., colorings of squares of graphs) restricted to plane cubic graphs since colorings of the square of a plane cubic graph are precisely its 2-facial colorings (refer to the book by Jensen and Toft [10, Problem 2.18] for more details on Wegner's conjecture).

Let $f_f(\ell)$ be the minimum number of colors needed to ℓ -facially color every plane graph. Note that $f_c(2\ell + 1) \leq f_f(\ell)$. So far, no value of ℓ is known for which this inequality is strict. The following problem is offered by [11].

PROBLEM 1. *Is it true that, for every integer $\ell \geq 1$, $f_c(2\ell + 1) = f_f(\ell)$?*

Another conjecture that should maybe be mentioned is the so-called 3ℓ -conjecture proposed by Dvořák, Škrekovski, and Tancer [7], stating that every plane triangle-free graph is ℓ -facially 3ℓ -colorable. As for the $(3\ell + 1)$ -conjecture, if this conjecture were true, then its bound would be tight and it would have several interesting corollaries (see [7] for more details).

Král', Madaras, and Škrekovski [11] proved that every plane graph has an ℓ -facial coloring using at most $\lfloor \frac{18}{5}\ell \rfloor + 2 \lfloor \frac{18}{5}\ell \rfloor + 2$ colors (and this bound is decreased by 1 for $\ell \in \{2, 4\}$). So, in particular, every plane graph has a 3-facial 12-coloring. In this paper, we improve this last result by proving the following theorem.

THEOREM 1. *Every plane graph is 3-facially 11-colorable.*

To prove this result, we suppose that it is false. In section 2, we exhibit some properties of a minimal graph (regarding the number of vertices) that contradict Theorem 1. Relying on these properties, we use the discharging method in section 3 to obtain a contradiction.

2. Properties of (3,11)-minimal graphs. Let us start this section by introducing some definitions. A vertex of degree d (at least d , at most d) is said to be a d -vertex (a $(\geq d)$ -vertex, a $(\leq d)$ -vertex, respectively). The notion of a d -face (a

($\leq d$)-face, a ($\geq d$)-face, respectively) is defined analogously regarding the size of a face. An ℓ -path is a path of length ℓ .

Two faces are *adjacent*, or *neighboring*, if they share a common edge. A 5-face is *bad* if it is incident to at least four 3-vertices. It is said to be *very bad* if it is incident to five 3-vertices.

If u and v are 3-facially adjacent, then u is a 3-facial neighbor of v . The set of all 3-facial neighbors of v is $\mathcal{N}_3(v)$. The 3-facial degree of v is $\deg_3(v) = |\mathcal{N}_3(v)|$. A vertex is *dangerous* if it has degree three and is incident to a face of size three or four. A 3-vertex is *safe* if it is not dangerous, i.e., is not incident to a (≤ 4)-face.

Let $G = (V, E)$ be a plane graph, and $\mathcal{U} \subseteq V$. Let $G_3[\mathcal{U}]$ be the graph with vertex set $\mathcal{U}$ such that xy is an edge in $G_3[\mathcal{U}]$ if and only if x and y are 3-facially adjacent vertices in G . If c is a partial coloring of G and u an uncolored vertex of G , we let $L_c(u)$ (or just $L(u)$) be the set $\{x \in \{1, 2, \dots, 11\} : \text{for all } v \in \mathcal{N}_3(u), c(v) \neq x\}$. The graph $G_3[\mathcal{U}]$ is *L-colorable* if there exists a proper vertex coloring of the vertices of $G_3[\mathcal{U}]$ such that for every $u \in \mathcal{U}$ we have $c(u) \in L(u)$.

The next two results are used by Král', Madaras, and Škrekovski [11].

LEMMA 1. *Let v be a vertex whose incident faces in a 2-connected plane graph G are $f_1, f_2, \dots, f_d$. Then*

$$\deg_3(v) \leq \left(\sum_{i=1}^d \min(|f_i|, 7) \right) - 2d,$$

where $|f_i|$ is the size of the face f_i .

Suppose that Theorem 1 is false: a $(3, 11)$ -minimal graph G is a plane graph that is not 3-facially 11-colorable, with $|V(G)| + |E(G)|$ as small as possible.

LEMMA 2. *Let G be a $(3, 11)$ -minimal graph. Then the following hold:*

1. G is 2-connected.
2. G has no separating cycle of length at most 7.
3. G contains no adjacent f_1 -face and f_2 -face with $f_1 + f_2 \leq 9$.
4. G has no vertex whose 3-facial degree is less than 11. In particular, the minimum degree of G is at least three.
5. G contains no edge uv separating two (≥ 4)-faces with $\deg_3(u) \leq 11$ and $\deg_3(v) \leq 12$. In particular, if two adjacent dangerous vertices do not lie on a same (≤ 4)-face, then none of them is incident to a 3-face.

In the remainder of this section, we give additional local structural properties of $(3, 11)$ -minimal graphs.

LEMMA 3. *Let G be a $(3, 11)$ -minimal graph. Suppose that v and w are two adjacent 3-vertices of G , both incident to a same 5-face and a same 6-face. Then the size of the third face incident to w is at least 7.*

Proof. By contradiction, suppose that the size of the last face incident to w is at most 6. Then, according to Lemma 1, we infer that $\deg_3(v) \leq 12$ and $\deg_3(w) \leq 11$, but this contradicts Lemma 2(v). $\square$

A *reducible configuration* is a (plane) graph that cannot be an induced subgraph of a $(3, 11)$ -minimal graph. The usual method to prove that a configuration is reducible is the following: first, we suppose that a $(3, 11)$ -minimal graph G contains a prescribed induced subgraph H . Then we contract some subgraphs $H_1, H_2, \dots, H_k$ of H . In most of the cases, $k \leq 2$. This yields a proper minor G' of G , which by the minimality of G admits a 3-facial 11-coloring c' . The goal is to derive from c' a 3-facial 11-coloring c of G , which would give a contradiction. To do so, each noncontracted vertex v of G keeps its color $c'(v)$. Let h_i be the vertex of G' created by the contraction of the

vertices of H_i : some vertices of H_i are assigned the color $c'(h_i)$ (in doing so, we must take care that these vertices are not 3-facially adjacent in G). Last, we show that the remaining uncolored vertices can also be colored.

In other words, we show that the graph $G_3[\mathcal{U}]$ is L -colorable, where for each $u \in \mathcal{U}$, $L(u)$ is the list of the colors that are assigned to no vertex in $\mathcal{N}_3(u) \setminus \mathcal{U}$ (defined in section 1) and $\mathcal{U}$ is the set of uncolored vertices. In most of the cases, the vertices of $\mathcal{U}$ will be greedily colored.

In all figures of the paper, the following conventions are used: a triangle represents a 3-vertex, a square represents a 4-vertex, and a circle may be any kind of vertex whose degree is at least the maximum between three and the one it has in the figure. The edges of each subgraph H_i are drawn in bold, and the circled vertices are the vertices of $\mathcal{U} = \{u_1, u_2, \dots\}$. A dashed edge between two vertices indicates a path of length at least one between those two vertices. An (in)equality written in a bounded region indicates a face whose size achieves the (in)equality. Last, vertices that are assigned the color $c'(h_i)$ are v, w , and t if a unique subgraph is contracted or x_1, x_2 for $i = 1$ and y_1, y_2 for $i = 2$ if two subgraphs are contracted.

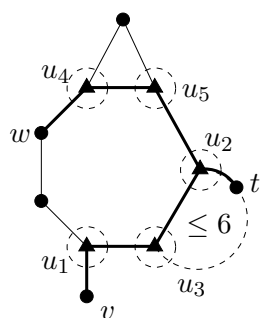
Note that the graph G' may contain loops and parallel edges. One way to consider them is as follows. If there is a face of size at most two, then we can just remove the loop, or one of the parallel edges. Otherwise, there is a separating cycle C of length at most two, and we can first color the subgraph of G' induced by C and the vertices inside C , and then the subgraph of G' induced by C and the vertices outside C .

LEMMA 4. *Configurations in Figures 2, 3, and 4 are reducible.*

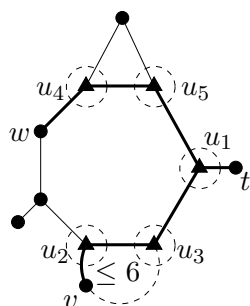
Proof. Let H be an induced subgraph of G . We suppose that H is isomorphic to one of the configurations stated and derive a way to construct a 3-facial 11-coloring of G , a contradiction.

L1. Suppose that H is isomorphic to the configuration (L1) of Figure 2. The edge u_2u_3 cannot be incident to a 3-face, since otherwise the edge u_2u_5 would contradict Lemma 2.5. More precisely, it would be incident to two (≥ 7)-faces by Lemma 2.3, and the 3-facial degree of u_2 and u_5 would be at most 11. Let H_1 be the subgraph induced by the bold edges. Contract the vertices of H_1 , thereby creating a new vertex h_1 . By minimality of G , let c' be a 3-facial 11-coloring of the obtained graph. Assign to each vertex x not in H_1 the color $c'(x)$, and to each of v, w, t the color $c'(h_1)$. Observe that, since the edge u_2u_3 does not lie on a 3-face, no two vertices among v, w, t are 3-facially adjacent in G ; otherwise there would be a (≤ 7)-separating cycle in G , thereby contradicting Lemma 2.2. According to Lemma 1, $\deg_3(u_1) \leq 15$, $\deg_3(u_i) \leq 14$ if $i \in \{2, 3\}$, and $\deg_3(u_i) \leq 11$ if $i \in \{4, 5\}$. Note that any two vertices of $\mathcal{U} = \{u_1, u_2, \dots, u_5\}$ are 3-facially adjacent; that is, $G_3[\mathcal{U}] \simeq K_5$. Hence, the number of colored 3-facial neighbors of u_1 is at most 11; i.e., $|\mathcal{N}_3(u_1) \setminus \{u_2, u_3, u_4, u_5\}| \leq 11$. Moreover, at least two of them are assigned the same color, namely v and w . Therefore, $|L(u_1)| \geq 1$. For $i \in \{2, 3\}$, the vertex u_i has at most 10 colored 3-facial neighbors. Furthermore, at least two 3-facial neighbors of u_2 are identically colored, namely w and t . Thus, $|L(u_2)| \geq 2$. Now, observe that at least three 3-facial neighbors of u_3 are colored the same, namely v, w , and t . Hence, $|L(u_3)| \geq 3$. For $i \in \{4, 5\}$, the vertex u_i has at most 7 colored 3-facial neighbors. Thus, $|L(u_4)| \geq 4$, and because at least two 3-facial neighbors of u_5 are identically colored (w and t), $|L(u_5)| \geq 5$. So, the graph $G_3[\mathcal{U}]$ is greedily L -colorable, according to the ordering u_1, u_2, u_3, u_4, u_5 . This allows us to extend c to a 3-facial 11-coloring of G .

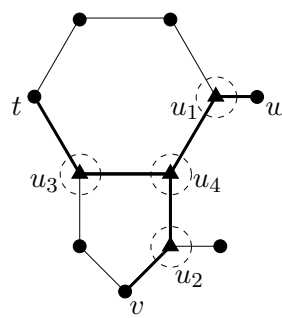
L2. Suppose that H is isomorphic to the configuration (L2) of Figure 2. Assume first that the edge u_2u_3 is not incident to a 3-face. Let c' be a 3-facial 11-coloring



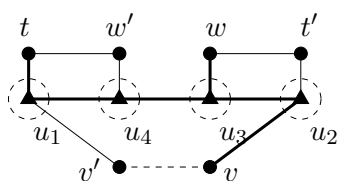
(L1)



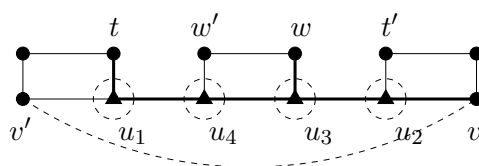
(L2)



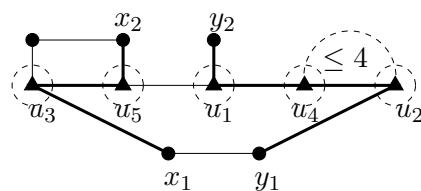
(L3)



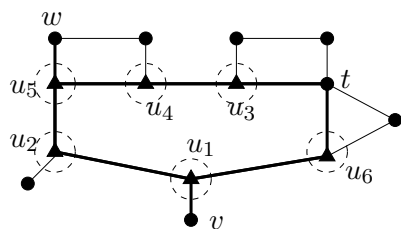
(L4)



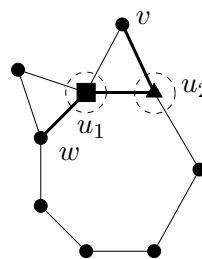
(L5)



(L6)



(L7)



(L8)

FIG. 2. Reducible configurations (L1)–(L8).

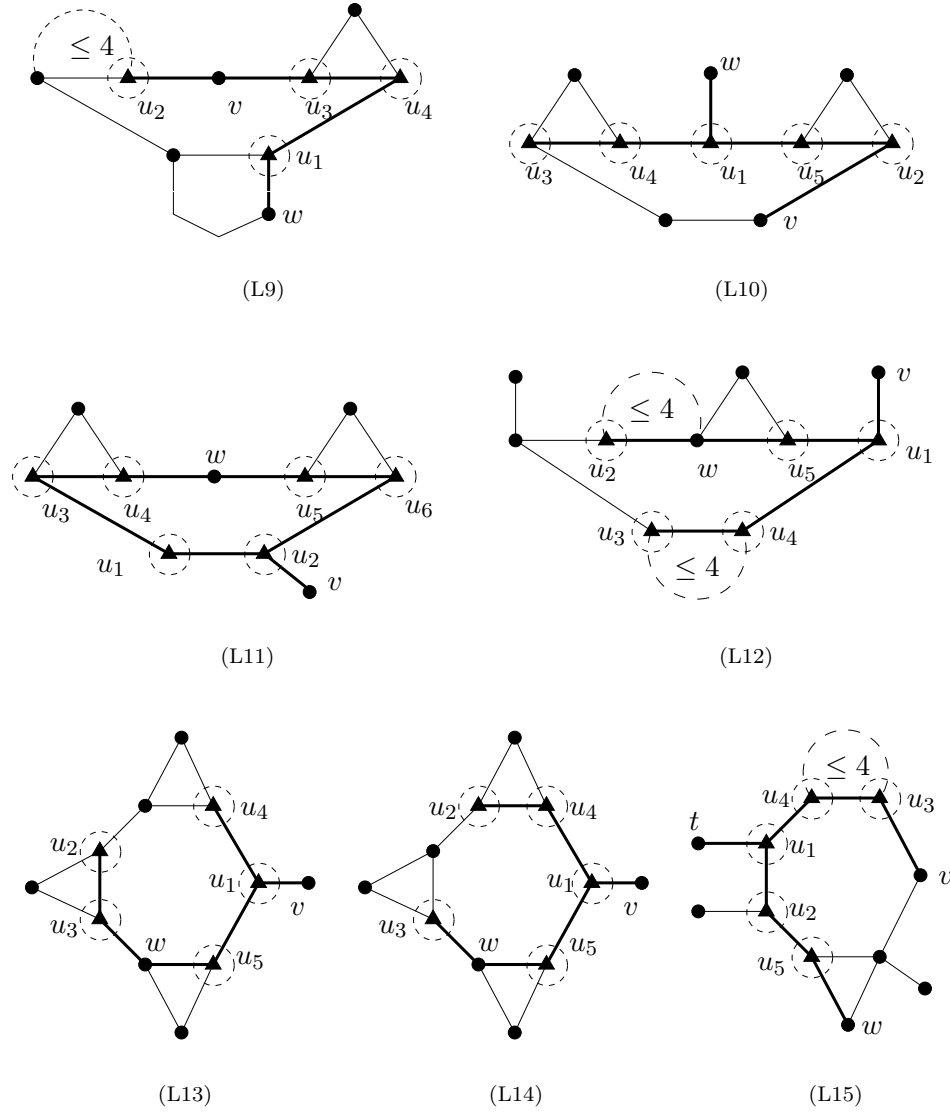


FIG. 3. Reducible configurations (L9)–(L15).

of the minor of G obtained by contracting the bold edges into a single vertex h_1 . Let $c(x) = c'(x)$ for every vertex $x \neq h_1$. Define $c(v) = c(w) = c(t) = c'(h_1)$. The obtained coloring is still 3-facial since no two vertices among v, w, t are 3-facially adjacent in G by Lemma 2.2, and because of our assumption. Note that $G_3[\mathcal{U}] \simeq K_5$. In particular, each vertex u_i has four uncolored 3-facial neighbors. By Lemma 1, $\deg_3(u_1) \leq 15$, $\deg_3(u_i) \leq 14$ if $i \in \{2, 3\}$, and $\deg_3(u_i) \leq 11$ if $i \in \{4, 5\}$. Moreover, each of u_1 and u_2 has at least two 3-facial neighbors colored the same; for u_1 , these vertices are w, t , and for u_2 they are w, v . So, there exists at least one color which is assigned to no vertex of $\mathcal{N}_3(u_1)$ and at least two colors assigned to no vertex of $\mathcal{N}_3(u_2)$. Also, u_3 has at least three 3-facial neighbors colored the same, namely w ,

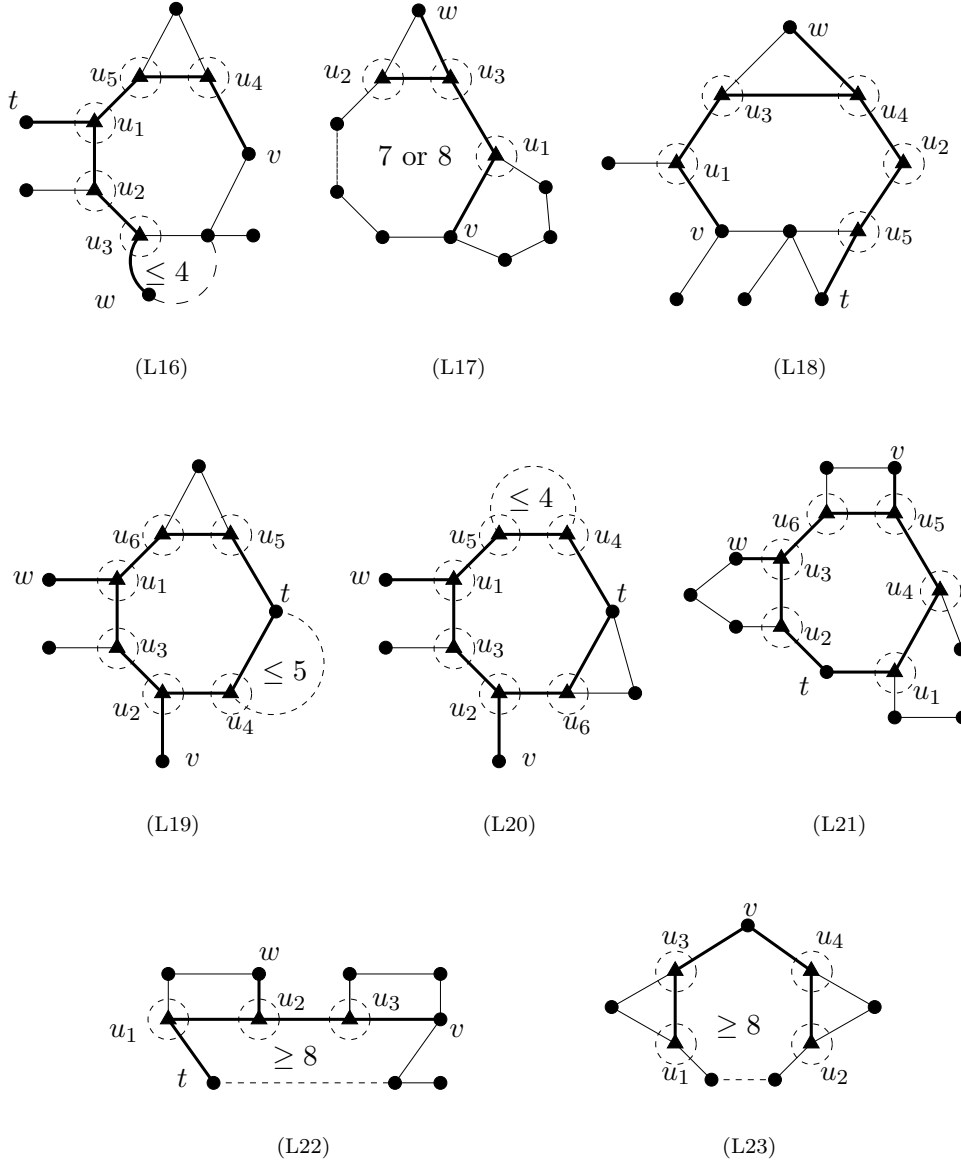


FIG. 4. Reducible configurations (L16)–(L23).

v , and t ; hence at least three colors are assigned to no vertex of $\mathcal{N}_3(u_3)$. Therefore, $|L(u_1)| \geq 1$, $|L(u_2)| \geq 2$, and $|L(u_3)| \geq 3$. Furthermore, $|L(u_4)| \geq 4$ and $|L(u_5)| \geq 5$ because w and t are both 3-facial neighbors of u_5 . So $G_3[\mathcal{U}]$ is L -colorable, and hence G is 3-facially 11-colorable.

If the edge u_2u_3 is incident to a 3-face, then the same proof works, except that at the beginning the edge u_2v is not contracted. Thus, only the vertices w and t have the same color, but the partial coloring extends as previously to G since u_2 and u_3 both have now 3-facial degree 11.

L3. Suppose that H is isomorphic to the configuration (L3) of Figure 2. Contract the bold edges into a new vertex h_1 , and let c' be a 3-facial 11-coloring of the obtained graph. This coloring can be extended to a 3-facial 11-coloring c of G as follows: first, let $c(v) = c(w) = c(t) = c'(h_1)$. Note that no two of these vertices can be 3-facially adjacent in G without contradicting Lemma 2.2. By Lemma 1, $\deg_3(u_1) \leq 14$, $\deg_3(u_2) \leq 13$ and for $i \in \{3, 4\}$, and $\deg_3(u_i) \leq 12$. Observe that $G_3[\mathcal{U}] \simeq K_4$. Moreover, each of u_1, u_2, u_3 has a set of two 3-facial neighbors colored by $c'(h_1)$. These sets are $\{w, t\}$, $\{w, v\}$, and $\{v, t\}$ for u_1, u_2 , and u_3 , respectively. Thus, $|L(u_1)| \geq 1$, $|L(u_2)| \geq 2$, and $|L(u_3)| \geq 3$. Also $|L(u_4)| \geq 4$ because u_4 has at least three identically colored 3-facial neighbors, namely v, w , and t . Hence, $G_3[\mathcal{U}]$ is L -colorable, and so G is 3-facially 11-colorable.

L4. First, observe that if $v \in \mathcal{N}_3(t)$, then $v' \notin \mathcal{N}_3(t')$, since G is a plane graph. So, by symmetry, we may assume that v and t are not 3-facially adjacent in G . Now, contract the bold edges into a new vertex h_1 . Again, let c' be a 3-facial 11-coloring of the obtained graph, and define c to be equal to c' on all vertices of $V(G) \setminus \{v, w, t, u_1, u_2, u_3, u_4\}$. Let $c(v) = c(w) = c(t) = c'(h_1)$. Note that the partial coloring c is still 3-facial due to the above assumption. The graph $G_3[\mathcal{U}]$ is isomorphic to K_4 , and according to Lemma 1, $\deg_3(u_i) \leq 12$ for all $i \in \{1, 2, 3, 4\}$. Moreover, for $i \in \{2, 3\}$, the vertex u_i has at least two 3-facial neighbors that are colored the same, namely, v and w . Last, the vertex u_4 has at least three such 3-facial neighbors, namely v, w, t . Therefore, $|L(u_1)| \geq 2$, $|L(u_i)| \geq 3$ for $i \in \{2, 3\}$, and $|L(u_4)| \geq 4$. So, $G_3[\mathcal{U}]$ is L -colorable, and hence G is 3-facially 11-colorable.

L5. The same remark as in the previous configuration allows us to assume that $t \notin \mathcal{N}_3(v)$. Again, the graph obtained by contracting the bold edges into a new vertex h_1 admits a 3-facial 11-coloring c' . As before, define a 3-facial 11-coloring c of the graph induced by $V(G) \setminus \mathcal{U}$. Then, for every $i \in \{1, 2, 3, 4\}$, $\deg_3(u_i) \leq 12$ and $G_3[\mathcal{U}] \simeq K_4$. Thus, $|L(u_1)| \geq 2$ and $|L(u_2)| \geq 2$. Note that u_3 has at least two identically colored 3-facial neighbors, namely v and w , so $|L(u_3)| \geq 3$. Last, the vertex u_4 has at least three such neighbors, hence $|L(u_4)| \geq 4$. Therefore, the graph $G_3[\mathcal{U}]$ is L -colorable, and so the graph G admits a 3-facial 11-coloring.

L6. Let H_1 be the path $x_1u_3u_5x_2$, H_2 the path $y_1u_2u_4u_1y_2$, and c' a 3-facial coloring of the graph obtained from G by contracting each path H_i into a vertex h_i . Notice that $c'(h_1) \neq c'(h_2)$. For every $v \notin V(H_1) \cup V(H_2)$, let $c(v) = c'(v)$. Observe that x_1 and x_2 cannot be 3-facially adjacent in G , otherwise G would have a separating (≤ 7)-cycle, contradicting Lemma 2.2. Note that the same holds for y_1 and y_2 ; therefore defining $c(x_1) = c(x_2) = c'(h_1)$ and $c(y_1) = c(y_2) = c'(h_2)$ yields a partial 3-facial 11-coloring of G , since $c'(h_1) \neq c'(h_2)$. It remains to color the vertices of $\mathcal{U} = \{u_1, u_2, \dots, u_5\}$. Note that $G_3[\mathcal{U}] \simeq K_5$. According to Lemma 2.2, $\deg_3(u_1) \leq 15$ and $\deg_3(u_i) \leq 12$ if $i \geq 2$. The number of colored 3-facial neighbors of u_1 , i.e., its number of 3-facial neighbors in $V(G) \setminus \{u_2, u_3, u_4, u_5\}$, is at most 11 because each u_i with $i \geq 2$ is a 3-facial neighbor of u_1 . Furthermore, u_1 has two 3-facial neighbors colored with the same color, namely x_1 and x_2 . Hence, $|L(u_1)| \geq 1$. The vertex u_2 has four uncolored 3-facial neighbors, so $|L(u_2)| \geq 3$. For $i \in \{3, 4\}$, the vertex u_i has at least two 3-facial neighbors colored the same, namely x_1, x_2 for u_3 , and y_1, y_2 for u_4 , so $|L(u_i)| \geq 4$. Finally, observe that u_5 has two pairs of identically colored 3-facial neighbors; the first pair being x_1, x_2 and the second y_1, y_2 . Thus, $|L(u_5)| \geq 5$, and hence the graph $G_3[\mathcal{U}]$ is L -colorable, which yields a contradiction.

L7. We contract the bold edges into a new vertex h_1 , take a 3-facial 11-coloring of the graph obtained, and define a 3-facial 11-coloring c of $V(G) \setminus \mathcal{U}$ as usual. By

Lemma 1, $\deg_3(u_i) \leq 15$ if $i \in \{1, 2\}$, $\deg_3(u_i) \leq 12$ if $i \in \{3, 4, 5\}$, and $\deg_3(u_6) \leq 11$. Moreover, $G_3[\mathcal{U}] \simeq K_6$. As v, w , and t are colored the same, and $\{v, w\} \subset \mathcal{N}_3(u_2)$, $\{w, t\} \subset \mathcal{N}_3(u_i)$ for $i \in \{4, 5\}$, and $\{v, t\} \subset \mathcal{N}_3(u_6)$, we obtain $|L(u_i)| \geq i$ for every $i \in \{1, 2, 3, 4, 5, 6\}$. Thus, the graph $G_3[\mathcal{U}]$ is L -colorable, and hence G admits a 3-facial 11-coloring.

L8. We contract the bold edges into a new vertex, take a 3-facial 11-coloring of the graph obtained, and define a 3-facial 11-coloring of $V(G) \setminus \mathcal{U}$ as usual. Then, $G_3[\mathcal{U}] \simeq K_2$. Moreover, $\deg_3(u_1) \leq 12$ and $\deg_3(u_2) \leq 11$. Furthermore, $\{v, w\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 2\}$. Thus, we infer $|L(u_i)| \geq i$ for $i \in \{1, 2\}$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L9. We contract the bold edges into a new vertex, take a 3-facial 11-coloring of the graph obtained, and define a 3-facial 11-coloring of $V(G) \setminus \mathcal{U}$ as usual. Then, $G_3[\mathcal{U}] \simeq K_4$. Moreover, $\deg_3(u_1) \leq 13$, $\deg_3(u_2) \leq 12$, and $\deg_3(u_i) \leq 11$ for $i \in \{3, 4\}$. Furthermore, $\{v, w\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 4\}$. Thus, we infer $|L(u_i)| \geq 2$ for $i \in \{1, 2\}$, and $|L(u_i)| \geq i$ for $i \in \{3, 4\}$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L10. We contract the bold edges into a new vertex h_1 , take a 3-facial 11-coloring of the graph obtained, and define a 3-facial 11-coloring c of $V(G) \setminus \mathcal{U}$ as usual. By Lemma 1, $\deg_3(u_1) \leq 15$ and $\deg_3(u_i) \leq 11$ if $i \in \{2, 3, 4, 5\}$. Moreover, $G_3[\mathcal{U}] \simeq K_5$. As v and w are colored the same, and $\{v, w\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 4, 5\}$, we obtain $|L(u_1)| \geq 1$, $|L(u_i)| \geq 4$ if $i \in \{2, 3\}$, and $|L(u_i)| \geq 5$ if $i \in \{4, 5\}$. Thus, the graph $G_3[\mathcal{U}]$ is L -colorable, and hence G admits a 3-facial 11-coloring.

L11. Let c' be a 3-facial 11-coloring of the graph G' obtained by contracting the bold edges into a new vertex h_1 . Define $c(x) = c'(x)$ for every vertex $x \in V(G) \cap V(G')$, and let $c(v) = c(w) = c'(h_1)$. By Lemma 1, $\deg_3(u_i) \leq 15$ for $i \in \{1, 2\}$ and $\deg_3(u_i) \leq 11$ for $i \in \{3, 4, 5\}$. Moreover, $G_3[\mathcal{U}] \simeq K_6$. Hence, $|L(u_1)| \geq 1$ and $|L(u_i)| \geq i$ for $i \in \{3, 4, 5\}$. As v and w are colored the same, and $\{v, w\} \subset \mathcal{N}_3(u_i)$ for $i \in \{2, 6\}$, we infer that $|L(u_2)| \geq 2$ and $|L(u_6)| \geq 6$. Thus, the graph G is 3-facially 11-colorable.

L12. Let us define the partial 3-facial 11-coloring c as always, regarding the bold edges and the vertices v and w . From Lemma 1 we obtain that $\deg_3(u_1) \leq 15$, $\deg_3(u_i) \leq 12$ for $i \in \{2, 3, 4\}$, and $\deg_3(u_5) \leq 11$. Moreover, since $G_3[\mathcal{U}] \simeq K_5$ and $\{v, w\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 4, 5\}$, we obtain $|L(u_1)| \geq 1$, $|L(u_i)| \geq 3$ for $i \in \{2, 3\}$, $|L(u_4)| \geq 4$, and $|L(u_5)| \geq 5$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L13. Define the partial 3-facial 11-coloring c as usual, regarding the bold edges and the vertices v and w . By Lemma 1, $\deg_3(u_1) \leq 15$ and $\deg_3(u_i) \leq 11$ for $i \in \{2, 3, 4, 5\}$. Moreover, since $G_3[\mathcal{U}] \simeq K_5$ and $\{v, w\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 5\}$, we obtain $|L(u_1)| \geq 1$, $|L(u_i)| \geq 4$ for $i \in \{2, 3, 4\}$, and $|L(u_5)| \geq 5$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L14. Let us define the partial 3-facial 11-coloring c as always, regarding the bold edges and the vertices v and w . Again, $G_3[\mathcal{U}] \simeq K_5$. From Lemma 1 we obtain that $\deg_3(u_1) \leq 15$ and $\deg_3(u_i) \leq 11$ if $i \in \{2, 3, 4, 5\}$. Moreover, since $\{v, w\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 5\}$, we obtain $|L(u_1)| \geq 1$, $|L(u_i)| \geq 4$ for $i \in \{2, 3, 4\}$, and $|L(u_5)| \geq 5$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L15. Define the partial 3-facial 11-coloring c as always, regarding the bold edges and the vertices v, w , and t . Then, $G_3[\mathcal{U}] \simeq K_5$ and $\deg_3(u_i) \leq 15$ for $i \in \{1, 2\}$, $\deg_3(u_i) \leq 12$ for $i \in \{3, 4\}$, and $\deg_3(u_5) \leq 11$. Moreover, notice that $\{v, t\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 4\}$, $\{v, w, t\} \subset \mathcal{N}_3(u_2)$, and $\{v, w\} \subset \mathcal{N}_3(u_5)$. Thus, we obtain $|L(u_1)| \geq 1$, $|L(u_2)| \geq 2$, $|L(u_3)| \geq 3$, $|L(u_4)| \geq 4$, and $|L(u_5)| \geq 5$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L16. Define the partial 3-facial 11-coloring c as always, regarding the bold edges and the vertices v, w , and t . Then, $G_3[\mathcal{U}] \simeq K_5$ and $\deg_3(u_i) \leq 15$ for $i \in \{1, 2\}$, $\deg_3(u_3) \leq 12$, and $\deg_3(u_i) \leq 11$ for $i \in \{4, 5\}$. Moreover, notice that $\{v, t\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 5\}$, $\{v, w, t\} \subset \mathcal{N}_3(u_2)$, and $\{v, w\} \subset \mathcal{N}_3(u_3)$. Thus, we obtain $|L(u_1)| \geq 1$, $|L(u_2)| \geq 2$, $|L(u_i)| \geq 4$ for $i \in \{3, 4\}$, and $|L(u_5)| \geq 5$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L17. Let us define the partial 3-facial 11-coloring c as always, regarding the bold edges and the vertices v and w . Then, $G_3[\mathcal{U}] \simeq K_3$, $\deg_3(u_1) \leq 13$, and $\deg_3(u_i) \leq 11$ for $i \in \{2, 3\}$. Moreover, $\{v, w\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 2, 3\}$. Thus, we obtain $|L(u_1)| \geq 1$ and $|L(u_i)| \geq 3$ for $i \in \{2, 3\}$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L18. Again, $G_3[\mathcal{U}] \simeq K_5$ and $\deg_3(u_i) \leq 15$ for $i \in \{1, 2\}$, while $\deg_3(u_i) \leq 11$ for $i \in \{3, 4, 5\}$. Furthermore, $\{v, w\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 3, 4\}$, $\{v, t\} \subset \mathcal{N}_3(u_5)$, and $\{v, w, t\} \subset \mathcal{N}_3(u_2)$. Thus, we deduce $|L(u_1)| \geq 1$, $|L(u_2)| \geq 2$, and $|L(u_i)| \geq 5$ for $i \in \{3, 4, 5\}$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L19. Here, $G_3[\mathcal{U}] \simeq K_6$. Also, $\deg_3(u_i) \leq 15$ for $i \in \{1, 2, 3\}$, $\deg_3(u_4) \leq 13$, and $\deg_3(u_i) \leq 11$ for $i \in \{5, 6\}$. Furthermore, $\{w, t\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 6\}$, $\{v, w, t\} \subset \mathcal{N}_3(u_3)$, and $\{v, t\} \subset \mathcal{N}_3(u_i)$ for $i \in \{2, 4\}$. Thus, we infer $|L(u_i)| \geq 2$ for $i \in \{1, 2\}$, $|L(u_3)| \geq 3$, $|L(u_4)| \geq 4$, $|L(u_5)| \geq 5$, and $|L(u_6)| \geq 6$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L20. Again $G_3[\mathcal{U}] \simeq K_6$. Also, $\deg_3(u_i) \leq 15$ for $i \in \{1, 2, 3\}$, $\deg_3(u_i) \leq 12$ for $i \in \{4, 5\}$, and $\deg_3(u_6) \leq 11$. Furthermore, $\{w, t\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 5\}$, $\{v, w, t\} \subset \mathcal{N}_3(u_3)$, and $\{v, t\} \subset \mathcal{N}_3(u_i)$ for $i \in \{2, 6\}$. Thus, we infer $|L(u_i)| \geq 2$ for $i \in \{1, 2\}$ and $|L(u_i)| \geq i$ for $i \in \{3, 4, 5, 6\}$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L21. In this case, $G_3[\mathcal{U}] \simeq K_6$. Also, $\deg_3(u_i) \leq 13$ for $i \in \{1, 2, 3, 4\}$, and $\deg_3(u_i) \leq 12$ for $i \in \{5, 6\}$. Furthermore, $\{v, t\} \subset \mathcal{N}_3(u_i)$ for $i \in \{4, 5\}$, $\{v, w, t\} \subset \mathcal{N}_3(u_6)$, and $\{w, t\} \subset \mathcal{N}_3(u_i)$ for $i \in \{2, 3\}$. Thus, we infer $|L(u_1)| \geq 3$, $|L(u_i)| \geq 4$ for $i \in \{2, 3, 4\}$, $|L(u_5)| \geq 5$, and $|L(u_6)| \geq 6$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L22. In this case, $G_3[\mathcal{U}] \simeq K_3$. Also, $\deg_3(u_i) \leq 12$ for $i \in \{1, 2, 3\}$. Moreover, $\{v, w, t\} \subset \mathcal{N}_3(u_i)$ for $i \in \{1, 2, 3\}$. Thus, we infer $|L(u_i)| \geq 3$ for $i \in \{1, 2, 3\}$. Therefore, $G_3[\mathcal{U}]$ is L -colorable.

L23. Define the partial coloring c as always, regarding the bold edges and the vertex v . Note that $G_3[\mathcal{U}]$ is isomorphic to the complete graph on four vertices minus one edge K_4^- , since $u_1 \notin \mathcal{N}_3(u_2)$ (because the face has size at least 8). By Lemma 1, $\deg_3(u_i) \leq 11$ for every $i \in \{1, 2, 3, 4\}$. Thus, $|L(u_i)| \geq 2$ for $i \in \{1, 2\}$, and $|L(u_i)| \geq 3$ for $i \in \{3, 4\}$. Hence, the graph $G_3[\mathcal{U}]$ is L -colorable. This assertion can be directly checked, or seen as a consequence of a theorem independently proved by Borodin [2] and Erdős, Rubin, and Taylor [8] (see also [17]), stating that a connected graph is degree-choosable unless it is a *Gallai tree*, that is, each of its blocks is either complete or an odd cycle. $\square$

COROLLARY 1. *Every (3, 11)-minimal graph G has the following properties:*

1. *Let f_1, f_2 be two 5-faces of G with a common edge xy . Then, x and y are not both 3-vertices.*
2. *Let f be a 7-face whose every incident vertex is a 3-vertex. If f is adjacent to a 3-face, then every other face adjacent to f is a (≥ 7) -face.*
3. *Two dangerous vertices incident to a same 6-face are not adjacent.*
4. *There cannot be four consecutive dangerous vertices incident to a same (≥ 6) -face.*
5. *A very bad face is adjacent to at least three (≥ 7) -faces.*
6. *A bad face is adjacent to at least two (≥ 7) -faces.*

Proof. 1. By Lemma 2.5, $\deg_3(x) + \deg_3(y) \geq 23$. By Lemma 1, the 3-facial degree of a 3-vertex incident to two 5-faces is at most 11. Hence at least one of x and y is a (≥ 4) -vertex.

2. First note that, according to Lemma 2.3, the faces adjacent to both f and the 3-face have sizes at least 7. Hence, f is adjacent to at most four (≤ 6) -faces. Now, the assertion directly follows from the reducibility of the configurations (L1) and (L2) of Figure 2.

3. Suppose on the contrary that x and y are two such vertices. By Lemma 2.3, a 6-face is not adjacent to a 3-face; hence both x and y are incident to a 4-face. Then, $\deg_3(x) \leq 11$ and $\deg_3(y) \leq 11$, which contradicts Lemma 2.5.

4. Suppose that the assertion is false. Then, according to the third item of this corollary, the graph G must contain the configuration (L4) or (L5) of Figure 2, which are both reducible.

5. Let f be a very bad face. By the first item of this corollary and Lemma 3, two adjacent (≤ 6) -faces cannot be both adjacent to f . Hence, f is adjacent to at most two such faces.

6. Let f be a bad face, and $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\}$ its incident vertices in clockwise order. Without loss of generality, assume that, for every $i \in \{1, 2, 3, 4\}$, α_i is a 3-vertex. For $i \in \{1, 2, 3, 4\}$, let f_i be the face adjacent to f and incident to both α_i and α_{i+1} . According to the first item of this corollary and Lemma 3, at most two faces among f_1, f_2, f_3, f_4 can be (≤ 6) -faces. This concludes the proof. $\square$

3. Proof of Theorem 1. Suppose that Theorem 1 is false, and let G be a $(3, 11)$ -minimal graph. We obtain a contradiction by using the discharging method. Here is an overview of the proof: each vertex and face is assigned an initial charge. The total sum of the initial charges is known to be negative by Euler's formula. Then, some redistribution rules are defined, and each vertex and face gives or receives some charge according to these rules. The total sum of the charges is not changed during this step, but at the end we show, by case analysis, that the charge of each vertex and each face is nonnegative, a contradiction.

Initial charge. First, we assign a charge to each vertex and face. For every $v \in V(G)$, we define the initial charge

$$\text{ch}(v) = d(v) - 4,$$

where $d(v)$ is the degree of the vertex v in G . Similarly, for every $f \in F(G)$, where $F(G)$ is the set of faces of G , we define the initial charge

$$\text{ch}(f) = r(f) - 4,$$

with $r(f)$ the size of the face f . By Euler's formula the total sum is

$$\sum_{v \in V(G)} \text{ch}(v) + \sum_{f \in F(G)} \text{ch}(f) = -8.$$

Rules. We use the following discharging rules to redistribute the initial charge.

RULE R1. A (≥ 5) -face sends $1/3$ to each of its incident safe vertices and $1/2$ to each of its incident dangerous vertices.

RULE R2. A (≥ 7) -face sends $1/3$ to each adjacent 3-face.

RULE R3. A (≥ 7) -face sends $1/6$ to each adjacent bad face.

RULE R4. A 6-face sends $1/12$ to each adjacent very bad face.

RULE R5. A (≥ 5) -vertex v gives $2/3$ to an incident face f if and only if there exist two 3-faces both incident to v and both adjacent to f . (Note that the size of such a face f is at least 7.)

We now prove that the final charge $\text{ch}^*(x)$ of every $x \in V(G) \cup F(G)$ is non-negative. Therefore, we obtain

$$-8 = \sum_{v \in V(G)} \text{ch}(v) + \sum_{f \in F(G)} \text{ch}(f) = \sum_{v \in V(G)} \text{ch}^*(v) + \sum_{f \in F(G)} \text{ch}^*(f) \geq 0,$$

a contradiction.

Final charge of vertices. First, as noticed in Lemma 2.4, G has minimum degree at least three. Let v be an arbitrary vertex of G . We prove that its final charge $\text{ch}^*(v)$ is nonnegative. To this end, we consider a few cases regarding its degree. So, suppose first that v is a 3-vertex. If v is a safe vertex, then by Rule R1 its final charge is $\text{ch}^*(v) = -1 + 3 \cdot \frac{1}{3} = 0$. Similarly, if v is dangerous, then $\text{ch}^*(v) = -1 + 2 \cdot \frac{1}{2} = 0$. If v is a 4-vertex, then it neither receives nor sends any charge. Thus, $\text{ch}^*(v) = \text{ch}(v) = 0$.

Finally, suppose that v is of degree $d \geq 5$. Notice that v may send charge only by Rule R5. This may occur at most $d/2$ times if d is even, and at most $\lfloor d/2 \rfloor - 1$ times if d is odd (since two 3-faces are not adjacent). Thus, $\text{ch}^*(v) \geq d - 4 - \lfloor \frac{d}{2} \rfloor \cdot \frac{2}{3}$, which is nonnegative if $d \geq 6$. For $d = 5$, $\text{ch}^*(v) \geq 5 - 4 - \frac{2}{3} > 0$.

Final charge of faces. Let f be an arbitrary face of G . We define **fce** and **bad** to be the number of 3-faces and the number of bad faces, respectively, adjacent to f . We define **sfe** and **dgs** to be the number of safe vertices and the number of dangerous vertices, respectively, incident to f . We prove that the final charge $\text{ch}^*(f)$ of f is nonnegative. To this end, we consider a few cases regarding the size of f .

f is a 3-face. It is adjacent only to (≥ 7) -faces by Lemma 2.3. Thus, by Rule R2, f receives $1/3$ from each of its three adjacent faces, so we obtain $\text{ch}^*(f) = 0$.

f is a 4-face. It neither receives nor sends any charge. Thus, $\text{ch}^*(f) = \text{ch}(f) = 0$.

f is a 5-face. Then, f is adjacent only to (≥ 5) -faces due to Lemma 2.3. So a 5-face may send charge only to its incident 3-vertices, which are all safe. Consider the following cases regarding the number **sfe** of such vertices:

sfe ≤ 3 : Then, $\text{ch}^*(v) \geq 1 - 3 \cdot \frac{1}{3} = 0$.

sfe $= 4$: In this case, f is a bad face. According to Corollary 1.6, at least two of the faces that are adjacent to f have size at least 7. Thus, according to Rule R3, f receives $1/6$ from at least two of its adjacent faces. Hence, we conclude that $\text{ch}^*(v) \geq 1 - 4 \cdot \frac{1}{3} + 2 \cdot \frac{1}{6} = 0$.

sfe $= 5$: Then f is a very bad face, and so, according to Corollary 1.5, at least three faces adjacent to f have size at least 7. Moreover, all faces adjacent to f have size at least 6, by Lemma 2.3 and Corollary 1.1. By Rules R3 and R4, it follows that the neighboring faces of f send at least $4 \cdot 1/6$ to f , which implies that $\text{ch}^*(v) \geq 1 - 5 \cdot \frac{1}{3} + 4 \cdot \frac{1}{6} = 0$.

f is a 6-face. By Lemma 2.3, **fce** $= 0$. Let **vbd** be the number of very bad faces adjacent to f . The final charge of f is $2 - \text{dgs} \cdot \frac{1}{2} - \text{sfe} \cdot \frac{1}{3} - \text{vbd} \cdot \frac{1}{12}$ due to Rules R1 and R4.

According to Corollary 1.3, two dangerous vertices on f cannot be adjacent, so there are at most three dangerous vertices on f . Observe also that **vbd** $\leq \text{sfe}/2$ by Corollary 1.1 and because a very bad face adjacent to f is incident to two safe vertices of f . Let us consider the final charge of f regarding its number of dangerous vertices.

dgs $= 3$: Since a safe vertex is not incident to a (≤ 4) -face, there is at most one safe vertex incident to f , i.e., **sfe** ≤ 1 . Thus, **vbd** $= 0$, and hence $\text{ch}^*(f) \geq$

$$2 - 3 \cdot \frac{1}{2} - \frac{1}{3} > 0.$$

$\text{dgs} = 2$: Then, $\text{sfe} \leq 3$. Let us distinguish two cases according to the value of sfe .

$\text{sfe} = 3$: Notice that $\text{vbd} = 0$; otherwise it would contradict the reducibility of (L3). Hence, $\text{ch}^*(f) \geq 2 - 2 \cdot \frac{1}{2} - 3 \cdot \frac{1}{3} = 0$.

$\text{sfe} \leq 2$: In this case, there is at most one very bad face adjacent to f , so $\text{ch}^*(f) \geq 2 - 2 \cdot \frac{1}{2} - 2 \cdot \frac{1}{3} - \frac{1}{12} > 0$.

$\text{dgs} = 1$: Then, $\text{sfe} \leq 4$ and $\text{vbd} \leq 1$ because (L3) is reducible. So, $\text{ch}^*(f) \geq 2 - \frac{1}{2} - \frac{4}{3} - \frac{1}{12} > 0$.

$\text{dgs} = 0$: If $\text{sfe} \geq 5$ then, because (L3) is reducible, $\text{vbd} = 0$; therefore $\text{ch}^*(f) \geq 2 - \frac{6}{3} = 0$. And, if $\text{sfe} \leq 4$, then $\text{vbd} \leq 2$, so $\text{ch}^*(f) \geq 2 - 4 \cdot \frac{1}{3} - 2 \cdot \frac{1}{12} > 0$.

f is a 7-face. The final charge of f is at least $3 - \text{dgs} \cdot \frac{1}{2} - (\text{fce} + \text{sfe}) \cdot \frac{1}{3} - \text{bad} \cdot \frac{1}{6}$.

According to Corollary 1.4, four dangerous vertices cannot be consecutive on f ; hence there cannot be more than five dangerous vertices on f . Let $\alpha_1, \alpha_2, \dots, \alpha_7$ be the vertices of f in clockwise order. Let $\mathcal{D}$ be the set of dangerous vertices of f , so $\text{dgs} = |\mathcal{D}|$. We look at the final charge of f , regarding its number dgs of dangerous vertices.

$\text{dgs} = 5$: Up to symmetry, $\mathcal{D} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_5, \alpha_6\}$. Suppose first that α_5 and α_6 are not incident to a same (≤ 4)-face. Then, there can be neither a safe vertex incident to f nor a bad face adjacent to f , because a safe vertex is not incident to a (≤ 4)-face, and also a bad face is not adjacent to a (≤ 4)-face. Moreover, by Corollary 1.3, there is no 3-face adjacent to f . Therefore, $\text{ch}^*(f) \geq 3 - \frac{5}{2} > 0$. Now, if α_5 and α_6 are incident to a same (≤ 4)-face, then α_4 and α_7 must be (≥ 4)-vertices, by the reducibility of (L6) and because none of them is a dangerous vertex. Hence, there is no safe vertex and no bad face adjacent to f , so its charge is $\text{ch}^*(f) \geq 3 - \frac{5}{2} - \frac{1}{3} > 0$.

$\text{dgs} = 4$: We consider several subcases, according to the relative position of the dangerous vertices on f . Recall that, by Corollary 1.4, there are at most three consecutive dangerous vertices. Without loss of generality, we need only to consider the following three possibilities:

$\mathcal{D} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_5\}$: The charge of f is $\text{ch}^*(f) = 1 - (\text{fce} + \text{sfe}) \cdot \frac{1}{3} - \text{bad} \cdot \frac{1}{6}$. Moreover, $\text{sfe} \leq 2$, $\text{bad} \leq 1$, and $\text{fce} + \text{sfe} \leq 3$ by Corollary 1.3 and because a safe vertex is not incident to a (≤ 4)-face. So, $\text{ch}^*(f)$ is negative if and only if $\text{sfe} = 2$, $\text{bad} = 1$, and $\text{fce} = 1$. But in this case, the obtained configuration is (L7), which is reducible.

$\mathcal{D} = \{\alpha_1, \alpha_2, \alpha_4, \alpha_5\}$: As a bad face is neither adjacent to a (≤ 4)-face nor incident to a dangerous vertex, we obtain that $\text{bad} \leq 1$. Observe also that, as α_3 is not dangerous, it has degree at least four by the reducibility of (L6) and (L10). Thus, $\text{sfe} \leq 2$. Suppose first that $\text{bad} = 1$, then sfe is one or two. According to the reducibility of (L9), we infer $\text{sfe} + \text{fce} \leq 2$. Hence, $\text{ch}^*(f) \geq 3 - 4 \cdot \frac{1}{2} - 2 \cdot \frac{1}{3} - \frac{1}{6} > 0$. Suppose now that $\text{bad} = 0$. We have $\text{fce} \leq 3$ and $\text{sfe} \leq 2$. If $\text{fce} = 3$, then $\text{sfe} = 0$, and if $\text{fce} = 2$, then $\text{sfe} \leq 1$ according to the reducibility of (L11). So, $\text{fce} + \text{sfe} \leq 3$. Therefore, $\text{ch}^*(f) \geq 3 - 4 \cdot \frac{1}{2} - (\text{fce} + \text{sfe}) \cdot \frac{1}{3} \geq 0$.

$\mathcal{D} = \{\alpha_1, \alpha_2, \alpha_4, \alpha_6\}$: In this case, there is no bad face adjacent to f . Furthermore, by Corollary 1.3, $\text{fce} \leq 3$ and $\text{sfe} \leq 2$, as the dangerous vertices α_4 and α_6 prevent at least one nondangerous vertex from being safe. Observe that $\text{fce} + \text{sfe} \neq 5$ since otherwise it would contradict the reducibility of (L12). According to the reducibility of (L12), if $\text{fce} + \text{sfe} = 4$, then $\text{fce} = 3$ and no two 3-faces have a common vertex. Hence, the obtained configuration is isomorphic to

(L13) or (L14), which are both reducible. So, $\mathbf{fce} + \mathbf{sfe} \leq 3$, and thus $\mathbf{ch}^*(f) \geq 3 - 2 - (\mathbf{fce} + \mathbf{sfe}) \cdot \frac{1}{3} \geq 0$.

$\mathbf{dgs} = 3$: Again, we consider several subcases according to the relative position of the dangerous vertices on f :

$\mathcal{D} = \{\alpha_1, \alpha_2, \alpha_3\}$: Then $\mathbf{fce} + \mathbf{sfe} \leq 3$ by Corollary 1.3, and $\mathbf{bad} \leq 2$. Thus, $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 3 \cdot \frac{1}{3} - 2 \cdot \frac{1}{6} > 0$.

$\mathcal{D} = \{\alpha_1, \alpha_2, \alpha_4\}$: Then, $\mathbf{fce} \leq 4$. We now examine the situation according to each possible value of $\mathbf{fce}$.

$\mathbf{fce} = 4$: Necessarily, $\mathbf{sfe} \leq 1$ and $\mathbf{bad} = 0$. Now, if $\mathbf{sfe} = 0$, then $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} > 0$. And, if $\mathbf{sfe} = 1$, then the safe vertex must be α_3 . Moreover, α_5 must be a (≥ 5) -vertex because (L8) is reducible. Hence, f is incident to α_5 between two 3-faces, so by Rule R5 the vertex α_5 gives $\frac{2}{3}$ to f . Thus, $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 5 \cdot \frac{1}{3} + \frac{2}{3} > 0$.

$\mathbf{fce} = 3$: Suppose first that one of the dangerous vertices is incident to a 4-face. Necessarily, $\mathbf{sfe} \leq 1$ and $\mathbf{bad} \leq 1$. Thus, $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} - \frac{1}{6} = 0$.

Suppose now that no dangerous vertex is incident to a 4-face. In particular, $\mathbf{sfe} \leq 2$. If $\mathbf{sfe} = 2$, then the obtained configuration contradicts the reducibility of (L18). Hence, $\mathbf{sfe} \leq 1$ and $\mathbf{bad} \leq 1$. Therefore, $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} - \frac{1}{6} = 0$.

$\mathbf{fce} = 2$: We prove that $\mathbf{sfe} \leq 2$. This is true if α_1 and α_2 are not incident to a same 3-face. So, we may assume that the edge $\alpha_1\alpha_2$ lies on a 3-face. But then we obtain the inequality due to the reducibility of (L18) and (L19). Using Corollary 1.1, the reducibility of (L17), and $\mathbf{sfe} \leq 2$, we infer that $\mathbf{bad} \leq 1$. Hence, $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} - \frac{1}{6} = 0$.

$\mathbf{fce} = 1$: Then $\mathbf{sfe} \leq 3$ and $\mathbf{bad} \leq 2$. If $\mathbf{sfe} = 3$ and $\mathbf{bad} = 2$, the obtained configuration contradicts the reducibility of (L19) or of (L20). So, $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} - \frac{1}{6} = 0$.

$\mathbf{fce} = 0$: Again, $\mathbf{sfe} \leq 3$ and $\mathbf{bad} \leq 2$, so $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 3 \cdot \frac{1}{3} - 2 \cdot \frac{1}{6} > 0$.

$\mathcal{D} = \{\alpha_1, \alpha_2, \alpha_5\}$: As in the previous case, $\mathbf{fce} \leq 4$, and we look at all the possible cases according to the value of $\mathbf{fce}$. Since a bad face is not incident to a dangerous vertex, notice that only edges $\alpha_3\alpha_4$ and $\alpha_6\alpha_7$ can be incident to a bad face. In particular, $\mathbf{bad} \leq 2$.

$\mathbf{fce} = 4$: In this case, $\mathbf{sfe} = 0$ and $\mathbf{bad} = 0$. Therefore, $\mathbf{ch}^*(f) = 3 - 3 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} > 0$.

$\mathbf{fce} = 3$: If one of the dangerous vertices is incident to a 4-face, then $\mathbf{sfe} = 0$; hence $\mathbf{bad} = 0$. Thus, $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 3 \cdot \frac{1}{3} \geq 0$. So now we infer that $\mathbf{sfe}$ cannot be 2; otherwise it would contradict the reducibility of (L15). Therefore, $\mathbf{sfe}$ is at most one, and so $\mathbf{bad} \leq 1$ by Corollary 1.1. Thus, $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} - \frac{1}{6} = 0$.

$\mathbf{fce} = 2$: According to the reducibility of (L15) and (L16), $\mathbf{sfe} \leq 2$. As $\mathbf{ch}^*(f) = 3 - 3 \cdot \frac{1}{2} - (\mathbf{fce} + \mathbf{sfe}) \cdot \frac{1}{3} - \mathbf{bad} \cdot \frac{1}{6}$, we deduce $\mathbf{ch}^*(f) < 0$ if and only if $\mathbf{sfe} = 2$ and $\mathbf{bad} = 2$. In this case, the obtained configuration is (L17), which is reducible.

$\mathbf{fce} = 1$: Because (L15) and (L16) are reducible, $\mathbf{sfe} \leq 2$. So, $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 3 \cdot \frac{1}{3} - 2 \cdot \frac{1}{6} > 0$.

$\mathbf{fce} = 0$: Then $\mathbf{sfe} \leq 3$, and so $\mathbf{ch}^*(f) \geq 3 - 3 \cdot \frac{3}{2} - 3 \cdot \frac{1}{3} - 2 \cdot \frac{1}{6} > 0$.

$\mathcal{D} = \{\alpha_1, \alpha_3, \alpha_5\}$: In this case, $\mathbf{sfe} \leq 2$ since a safe vertex is not incident to a (≤ 4)-face, and $\mathbf{bad} \leq 1$, since a bad face cannot be incident to a dangerous vertex. Moreover, $\mathbf{fce} \leq 4$. Let us examine the possible cases regarding the value of $\mathbf{fce}$.

$\mathbf{fce} = 4$: Observe that $\mathbf{sfe} \leq 1$ and $\mathbf{bad} = 0$. Note also that one of $\alpha_2, \alpha_4, \alpha_6, \alpha_7$ is adjacent to a dangerous vertex, and incident to f between two triangles. Hence, by the reducibility of (L8), it has degree at least five, and by Rule R5, it sends $\frac{2}{3}$ to f . Thus, $\text{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 5 \cdot \frac{1}{3} + \frac{2}{3} > 0$.

$\mathbf{fce} = 3$: If $\mathbf{sfe} \leq 1$, then $\text{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} - \frac{1}{6} = 0$. And, if $\mathbf{sfe} = 2$, then, up to symmetry, the two safe vertices are either α_6 and α_7 or α_2 and α_6 . In the former case, one of α_2, α_4 is incident to f at the intersection of two 3-faces. Furthermore, it must be a (≥ 5)-vertex due to the reducibility of (L8). In the latter case, the same holds for α_4 due to the reducibility of (L8). Hence, in both cases the face f receives $2/3$ from one of its incident vertices by Rule R5.

Recall that $\mathbf{bad} \leq 1$, and therefore $\text{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 5 \cdot \frac{1}{3} - \frac{1}{6} + \frac{2}{3} > 0$.

$\mathbf{fce} \leq 2$: As $\mathbf{sfe} \leq 2$ and $\mathbf{bad} \leq 1$, we infer that $\text{ch}^*(f) \geq 3 - 3 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} - \frac{1}{6} = 0$.

$\mathbf{dgs} = 2$: Again, we consider several subcases, regarding the position of the dangerous vertices on f .

$\mathcal{D} = \{\alpha_1, \alpha_2\}$: Observe that $\mathbf{bad} \leq 3$, and according to Corollary 1.3, $\mathbf{fce} + \mathbf{sfe} \leq 6$. We consider three cases, according to the value of $\mathbf{fce} + \mathbf{sfe}$:

$\mathbf{fce} + \mathbf{sfe} = 6$: All the vertices incident to f have degree three, and f is adjacent to a 3-face. Thus, by Corollary 1.2, f is not adjacent to any (≤ 6)-face. In particular, no bad face is adjacent to f ; i.e., $\mathbf{bad} = 0$. Hence, $\text{ch}^*(f) \geq 3 - 1 - 6 \cdot \frac{1}{3} = 0$.

$\mathbf{fce} + \mathbf{sfe} = 5$: If $\mathbf{bad} \leq 2$, then $\text{ch}^*(f) \geq 3 - 1 - 5 \cdot \frac{1}{3} - 2 \cdot \frac{1}{6} = 0$. Otherwise, $\mathbf{bad} = 3$. Note that the edge $\alpha_1\alpha_2$ must be incident to a (≤ 4)-face. If this face is of size four, then we obtain configuration (L21). Suppose now that this face is of size three. Since there are no three consecutive bad faces around f , we can assume that each of the edges $\alpha_3\alpha_4$ and $\alpha_6\alpha_7$ lies on a bad face. By the reducibility of (L17), we conclude that α_3 and α_7 have degree at least four. But then, $\mathbf{fce} + \mathbf{sfe} < 5$.

$\mathbf{fce} + \mathbf{sfe} \leq 4$: In this case, $\text{ch}^*(f) \geq 3 - 1 - 4 \cdot \frac{1}{3} - 3 \cdot \frac{1}{6} > 0$.

$\mathcal{D} = \{\alpha_1, \alpha_3\}$ or $\mathcal{D} = \{\alpha_1, \alpha_4\}$: Again $\mathbf{fce} + \mathbf{sfe} \leq 6$, and we consider two cases regarding the value of $\mathbf{fce} + \mathbf{sfe}$. Since a bad face is not incident to a dangerous vertex, we infer that $\mathbf{bad} \leq 3$.

$\mathbf{fce} + \mathbf{sfe} = 6$: Suppose first that $\mathcal{D} = \{\alpha_1, \alpha_3\}$. Let $P_1 = \alpha_1\alpha_2\alpha_3$ and $P_2 = \alpha_3\alpha_4\alpha_5\alpha_6\alpha_7\alpha_1$. In order to assure $\mathbf{fce} + \mathbf{sfe} = 6$, observe that all edges of P_1 are incident to 3-faces, and all inner vertices of P_2 are safe, or vice versa. Thus, α_2 or α_4 is a (≥ 5)-vertex by the reducibility of (L8). Hence, it gives $\frac{2}{3}$ to f by Rule R5. Therefore, $\text{ch}^*(f) \geq 3 - 2 \cdot \frac{1}{2} - 6 \cdot \frac{1}{3} - 3 \cdot \frac{1}{6} + \frac{2}{3} > 0$.

Suppose now that $\mathcal{D} = \{\alpha_1, \alpha_4\}$. Similarly as above, one can show that α_2 or α_5 is a (≥ 5)-vertex that donates $\frac{2}{3}$ to f . Hence, $\text{ch}^*(f) \geq 3 - 2 \cdot \frac{1}{2} - 6 \cdot \frac{1}{3} - \frac{3}{6} + \frac{2}{3} > 0$.

$\mathbf{fce} + \mathbf{sfe} \leq 5$: Notice that $\mathbf{bad} \leq 2$. Therefore, $\text{ch}^*(f) \geq 3 - 2 \cdot \frac{1}{2} - 5 \cdot \frac{1}{3} - 2 \cdot \frac{1}{6} = 0$.

dgs = 1: Then **fce** + **sfe** ≤ 6 and, by Corollary 1.1, we infer that **bad** ≤ 3 . So, $\text{ch}^*(f) \geq 3 - \frac{1}{2} - 6 \cdot \frac{1}{3} - 3 \cdot \frac{1}{6} = 0$.

dgs = 0: By Corollary 1.1, **fce** + **sfe** ≤ 7 and **bad** ≤ 4 . So, $\text{ch}^*(f) \geq 3 - 7 \cdot \frac{1}{3} - 4 \cdot \frac{1}{6} = 0$.

f is an 8-face. By Lemma 2.5 and because (L22) is reducible, there cannot be three consecutive dangerous vertices on *f*. Hence, **dgs** ≤ 5 . Let α_i , $i \in \{1, 2, \dots, 8\}$, be the vertices incident to *f* in clockwise order, and let $\mathcal{D}$ be the set of dangerous vertices incident to *f*.

dgs = 5: Up to symmetry, $\mathcal{D} = \{\alpha_1, \alpha_2, \alpha_4, \alpha_5, \alpha_7\}$. Since a bad face is not incident to a dangerous vertex, necessarily **bad** = 0. For $i \in \{1, 4\}$, let f_i be the face adjacent to *f* and incident to both α_i and α_{i+1} . Since (L23) is reducible, at most one of f_1 and f_4 is a 3-face. Furthermore, at most two of $\alpha_3, \alpha_6, \alpha_8$ can be safe vertices, since at least one of α_6, α_8 is a (≥ 4) -vertex. Therefore, **fce** ≤ 2 , **sfe** ≤ 2 , and so $\text{ch}^*(f) \geq 4 - 5 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} > 0$.

dgs = 4: Up to symmetry the set of dangerous vertices comprises $\{\alpha_1, \alpha_2, \alpha_4, \alpha_5\}$, $\{\alpha_1, \alpha_2, \alpha_5, \alpha_6\}$, $\{\alpha_1, \alpha_2, \alpha_4, \alpha_6\}$, $\{\alpha_1, \alpha_2, \alpha_4, \alpha_7\}$, or $\{\alpha_1, \alpha_3, \alpha_5, \alpha_7\}$. In any case, **bad** ≤ 2 and **fce** + **sfe** ≤ 6 . Moreover, **fce** + **sfe** + **bad** ≤ 6 . Indeed, if **bad** = 0, then the inequality holds by the prior remark. This solves the fifth case for $\mathcal{D}$ since a bad face is not incident to a dangerous vertex. If **bad** = 2, then we are in the first or second case for $\mathcal{D}$. We infer that **fce** + **sfe** ≤ 4 by the reducibility of (L17) and (L23), and by Corollary 1.1. Finally, if **bad** = 1, then we deduce that **fce** + **sfe** ≤ 5 by the reducibility of (L17). Hence, $\text{ch}^*(f) \geq 4 - \frac{4}{2} - \frac{6}{3} = 0$.

dgs = 3: Then, **fce** + **sfe** ≤ 6 and **bad** ≤ 3 . So, $\text{ch}^*(f) \geq 4 - \frac{3}{2} - \frac{6}{3} - \frac{3}{6} = 0$.

dgs = 2: Then, **fce** + **sfe** ≤ 7 , and by Corollary 1.1, **bad** ≤ 4 . Thus, $\text{ch}^*(f) \geq 4 - \frac{2}{2} - \frac{7}{3} - \frac{4}{6} = 0$.

dgs = 1: Again, **fce** + **sfe** ≤ 7 and **bad** ≤ 4 , so $\text{ch}^*(f) \geq 4 - \frac{1}{2} - \frac{7}{3} - \frac{4}{6} > 0$.

dgs = 0: By Corollary 1.1, **bad** ≤ 5 . So, $\text{ch}^*(f) \geq 4 - \frac{8}{3} - \frac{5}{6} > 0$.

f is a (≥ 9) -face. Let *f* be a *k*-face with $k \geq 9$. We use the following averaging scheme for the rules followed by *f*: the face *f* sends the charge to incident vertices or adjacent faces *through* its incident edges. More precisely, for Rules R2 and R3, if f' is a bad face or a triangle incident to *f*, then we say that *f* sends the corresponding charge *through* the edge that is incident to both *f* and f' . As for Rule R1, let *u* be a vertex incident to *f*, and let *v* and *w* be its two neighbors on *f*. If *u* is a safe vertex, then *f* sends $\frac{1}{6}$ to *u* through each of the two edges *uv* and *uw*. If *u* is dangerous, let *uv* be the edge incident with a (≤ 4) -face: if *v* is dangerous, then *f* sends $\frac{1}{9}$ to *u* through *uv* and $\frac{7}{18}$ to *u* through *uw*. Otherwise (i.e., when *v* is a (≥ 4) -vertex), *f* sends $\frac{2}{9}$ to *u* through *uv* and $\frac{5}{18}$ to *u* through *uw*.

By Lemma 2.5 and the reducibility of (L22), there cannot be three consecutive dangerous vertices on *f*. So, we deduce that all the vertices incident to *f* receive the same charge as if *f* applied the original Rule R1. We prove now that *f* sends at most $\frac{5}{9}$ to each of its edges, and hence $\text{ch}^*(f) \geq k(1 - \frac{5}{9}) - 4 \geq 0$ since $k \geq 9$.

Let *uv* be an edge incident to *f*. We consider three cases.

uv is incident to a bad face f' : Then, *u* and *v* are not dangerous. So *f* sends through *uv* $\frac{1}{6}$ to *u* plus $\frac{1}{6}$ to *v* and $\frac{1}{6}$ to f' . Thus, the charge sent by *f* through the edge *uv* is at most $3 \cdot \frac{1}{6} = \frac{1}{2} < \frac{5}{9}$.

uv is incident to a triangle f' : In this case, *f* sends $\frac{1}{3}$ to f' . If neither of *u* or *v* is dangerous, then *f* sends nothing more through *uv*. If exactly one of *u* and *v* is dangerous, say *u*, then *f* sends $\frac{2}{9}$ to *u* through *uv*. Thus, the charge sent by *f* through *uv* is $\frac{1}{3} + \frac{2}{9} = \frac{5}{9}$. Finally, assume that both *u* and *v* are dangerous.

Then, f sends $\frac{1}{9}$ to each of u and v through uv . Hence, f sends $\frac{1}{3} + 2 \cdot \frac{1}{9} = \frac{5}{9}$ through uv .

uv is incident to neither a bad face nor a triangle: Again, if neither of u or v is dangerous, then f sends at most $2 \cdot \frac{1}{6} = \frac{1}{3}$ through uv . Suppose that both u and v are dangerous. If uv is incident to a 4-face, then f sends $2 \cdot \frac{1}{9} = \frac{2}{9}$ through uv . Otherwise, let t be the neighbor of u on f different from v , and let w be the neighbor of v on f different from u . By Lemma 2.5, each of tu and vw is incident to a 4-face, and t and w are not dangerous since (L22) is reducible. Therefore, f sends $\frac{5}{18}$ to each of u and v through uv , and thus f sends $\frac{5}{9}$ through uv . Finally, if exactly one of u and v is dangerous, say u , then f sends at most $\frac{7}{18}$ to u through uv , and at most $\frac{1}{6}$ to v through uv . In total, f sends at most $\frac{7}{18} + \frac{1}{6} = \frac{5}{9}$ through uv .

The proof of Theorem 1 is now complete. $\square$

Acknowledgments. The authors are grateful to Mr. Aron Koldenbach, whose remarks helped to improve this paper. They are also grateful to both the referees for their careful reading; they are especially indebted to one of them for pointing out the averaging argument for (≥ 9) -faces, which allowed a shortening of the very last part of the proof.

REFERENCES

- [1] K. APPEL AND W. HAKEN, *Every Planar Map is Four Colorable*, Contemp. Math. 98, American Mathematical Society, Providence, RI, 1989.
- [2] O. V. BORODIN, *A criterion of chromaticity of a degree prescription*, in Abstracts of the fourth All-Union Conference on Theoretical Cybernetics (Novosibirsk), 1977, pp. 127–128 (in Russian).
- [3] O. V. BORODIN, *Solution of the Ringel problem on vertex-face coloring of planar graphs and coloring of 1-planar graphs*, Metody Diskret. Analiz., 41 (1984), pp. 12–26.
- [4] O. V. BORODIN, *Cyclic coloring of plane graphs*, Discrete Math., 100 (1992), pp. 281–289.
- [5] O. V. BORODIN, *A new proof of the 6 color theorem*, J. Graph Theory, 19 (1995), pp. 507–521.
- [6] O. V. BORODIN, D. P. SANDERS, AND Y. ZHAO, *On cyclic colorings and their generalizations*, Discrete Math., 203 (1999), pp. 23–40.
- [7] Z. DVOŘÁK, R. ŠKREKOVSKI, AND M. TANCER, *List-coloring squares of sparse subcubic graphs*, SIAM J. Discrete Math., 22 (2008), pp. 139–159.
- [8] P. ERDŐS, A. L. RUBIN, AND H. TAYLOR, *Choosability in graphs*, in Proceedings of the West Coast Conference on Combinatorics, Graph Theory and Computing (Humboldt State University, Arcata, CA, 1979), Congress. Numer., XXVI, Winnipeg, MB, 1980, Utilitas Math., pp. 125–157.
- [9] M. HORŇÁK AND S. JENDROŤ, *On some properties of 4-regular plane graphs*, J. Graph Theory, 20 (1995), pp. 163–175.
- [10] T. R. JENSEN AND B. TOFT, *Graph Coloring Problems*, Wiley–Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, New York, 1995.
- [11] D. KRÁĚ, T. MADARAS, AND R. ŠKREKOVSKI, *Cyclic, diagonal and facial colorings*, European J. Combin., 26 (2005), pp. 473–490.
- [12] D. KRÁĚ, T. MADARAS, AND R. ŠKREKOVSKI, *Cyclic, diagonal and facial colorings—A missing case*, European J. Combin., 28 (2007), pp. 1637–1639.
- [13] Ø. ORE AND M. D. PLUMMER, *Cyclic coloration of plane graphs*, in Recent Progress in Combinatorics (Proceedings of the Third Waterloo Conference on Combinatorics, 1968), Academic Press, New York, 1969, pp. 287–293.
- [14] D. P. SANDERS AND Y. ZHAO, *On d-diagonal colorings*, J. Graph Theory, 22 (1996), pp. 155–166.
- [15] D. P. SANDERS AND Y. ZHAO, *On d-diagonal colorings of embedded graphs of low maximum face size*, Graphs Combin., 14 (1998), pp. 81–94.
- [16] D. P. SANDERS AND Y. ZHAO, *A new bound on the cyclic chromatic number*, J. Combin. Theory Ser. B, 83 (2001), pp. 102–111.
- [17] C. THOMASSEN, *Color-critical graphs on a fixed surface*, J. Combin. Theory Ser. B, 70 (1997), pp. 67–100.

NORMALIZED MATCHING PROPERTY OF RESTRICTED SUBSPACE LATTICES*

JUN WANG[†] AND HUAJUN ZHANG[‡]

Abstract. Let V be the n -dimensional vector space over the finite field with q elements and $L(V)$ the lattice of subspaces of V ordered by inclusion. Let $V_1, V_2, \dots, V_r$ be r selected subspaces of V such that $\{0\} = V_0 \subseteq V_1 \subseteq \dots \subseteq V_r = V$ and $\dim(V_i) = \sum_{j=1}^i k_j$. Let $\mathcal{I}_i$ be a subinterval of $[0, k_i]$, $i = 1, \dots, r$, let $I = \{\mathcal{I}_1, \dots, \mathcal{I}_r\}$, and let $C[n, r, I] = \{K \in L(V) : \dim(K \cap V_i) - \dim(K \cap V_{i-1}) \in \mathcal{I}_i, i = 1, 2, \dots, r\}$. Then $C[n, r, I]$ is a graded poset. In this paper, using group actions and flow morphisms we show that $C[n, r, I]$ is log concave and has the NM property, which yields that $C[n, r, I]$ has the strong Sperner property and the LYM property.

Key words. poset, subspace lattice, Sperner property, normalized matching property

AMS subject classifications. 05D05, 06A07

DOI. 10.1137/050635110

1. Introduction. A finite poset P is *graded* if every maximal chain of P has the same length, which is called the rank of P denoted by $r(P)$. Thus, for each $x \in P$, every maximal chain with x as the top element has the same length, which is called the rank of the element x and denoted by $r(x)$. Here the length of a chain with k elements is $k - 1$. Let P_i denote the i th level of P , which consists of all $x \in P$ with $r(x) = i$ ($0 \leq i \leq r(P)$). An *antichain* is a subset $A \subseteq P$ of which no two elements are comparable in P . A *k-family* is a subset of P that contains no chains of length k . Clearly, each P_i is an antichain, and each union of k levels of P is a k -family. We say P is *unimodal* if the sequence of the Whitney numbers of P is unimodal, i.e., $W_0 \leq \dots \leq W_{\ell-1} \leq W_{\ell} \geq W_{\ell+1} \geq \dots \geq W_n$ for some ℓ , where $n = r(P)$ and $W_i = |P_i|$. Similarly, P is *log concave* if the sequence of Whitney numbers of P is log concave: $W_m^2 \geq W_{m-1}W_{m+1}$ for all m . It is clear that the log concavity of P implies that P is unimodal. For $A \subseteq P_i$ where $0 \leq i < r(P)$, define $\nabla(A) = \{b \in P_{i+1} : b \geq a \text{ for some } a \in A\}$.

As stated in [3], some extremal problems can be considered in a weighted poset (P, w) , which is a poset P together with a function (called a weight function) w from P into the set of nonnegative real numbers. The weight $w(A)$ of a subset A of P is defined by $w(A) = \sum_{a \in A} w(a)$. Every poset P can be considered to be a weighted poset (P, w) , where $w \equiv 1$; that is, $w(x) = 1$ for all $x \in P$.

We say a graded poset (P, w) is *k-Sperner*, or has the *k-Sperner property*, if a union of k ranks of P is a maximal-weight k -family in P . 1-Sperner is referred to simply as Sperner. We say (P, w) is *strongly Sperner*, or has the *strong Sperner property* if it is k -Sperner for all k . We say (P, w) has the normalized matching (NM) property if

$$(1) \quad \frac{w(\nabla(A))}{w(A)} \geq \frac{w(P_{i+1})}{w(P_i)} \quad \text{for all } A \subseteq P_i, \quad i = 0, 1, 2, \dots, r(P) - 1.$$

*Received by the editors July 4, 2005; accepted for publication (in revised form) July 2, 2007; published electronically February 22, 2008. This work was supported by the National Natural Science Foundation of China grant 10471016.

<http://www.siam.org/journals/sidma/22-1/63511.html>

[†]Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, P. R. of China (junwang@dlut.edu.cn).

[‡]Correspondence author. Current address: Department of Mathematics, Zhejiang Normal University, Jinhua 321004, P. R. of China (zhjhj1979@yahoo.com.cn).

We say (P, w) has the *Lubell–Yamamoto–Meshalkin (LYM) property* if for every antichain A in P the following inequality holds:

$$\sum_{i=0}^{r(P)} \frac{w(A \cap P_i)}{w(P_i)} \leq 1.$$

These properties are called the properties of Sperner type. It is well known that the NM property and the LYM property are equivalent, and that they imply the strong Sperner property [8].

Let $C(n, k)$ be the collection of all subsets of an n -set Y which intersect a fixed k -subset of X . In fact, $C(n, k)$ constitutes a natural generalization for the subset lattice. Lih [14] first observed this and showed that $C(n, k)$ has the Sperner property. Griggs [9] further showed that $C(n, k)$ has several strong properties.

Let V be the n -dimensional vector space over the finite field with q elements and $L(V)$ the lattice of subspaces of V ordered by inclusion. Let $V_1, V_2, \dots, V_r$ be r selected subspaces of V such that $\{0\} = V_0 \subseteq V_1 \subseteq \dots \subseteq V_r = V$ and $\dim(V_i) = \sum_{j=1}^i k_j$. Let $\mathcal{I}_i$ be a subinterval of $[0, k_i]$, $i = 1, \dots, r$, let $I = \{\mathcal{I}_1, \dots, \mathcal{I}_r\}$, and let $C[n, r, I] = \{K \in L(V) : \dim(K \cap V_i) - \dim(K \cap V_{i-1}) \in \mathcal{I}_i, i = 1, 2, \dots, r\}$.

In this paper we prove that $C[n, r, I]$ is log concave and has the NM property, which can be regarded as a vector-space analog of some results in [9, 14]. We shall make preparations for proving the theorem and present some preliminary results in the next section. The main results will be proved in sections 3 and 4.

2. Quotient posets and flow morphism. From algebra we know that some properties of a group, ring or module etc. can be partly reflected by their homomorphic image. Is this approach available for properties of Sperner type of a poset? In other words, if there is an order preserving map (homomorphism) from (P, w) onto (Q, v) and (Q, v) has a property of Sperner type, can one correctly infer this to be a property of (P, w) ? A first step in this direction is a theorem of Kleitman, Edelberg, and Lubell in [13] (see also [5]) as follows.

THEOREM 2.1 (see [13]). *Let P be a poset and Γ a permutation group acting on P which preserves the order relation on P . Then P contains a maximal sized antichain which is invariant under the action of Γ .*

This theorem provides a powerful method to study the Sperner property of a poset. For each $x \in P$, let $\Gamma(x) = \{\gamma(x) : \gamma \in \Gamma\}$. Then we have the quotient poset $(P/\Gamma, w_\Gamma)$, where P/Γ consists of the Γ -orbits ordered as follows: $\Gamma(x) \leq \Gamma(y)$ in P/Γ if $x' \leq y'$ for some $x' \in \Gamma(x)$ and $y' \in \Gamma(y)$, and the weight function w_Γ is given by $w_\Gamma(\Gamma(x)) = |\Gamma(x)|$. Then, $\psi : x \mapsto \Gamma(x)$ for $x \in P$ is an order and rank preserving map from P onto P/Γ . Then, the above theorem says that P has the Sperner property if and only if $(P/\Gamma, w_\Gamma)$ does. By this result the first author of this paper proved a conjecture on the Sperner property of the subgroup lattice of an abelian p -group [15]. The group theoretic approach was also used for intersecting Sperner families [4].

As shown in [13], Theorem 2.1 may be generalized in two directions. First, we can consider a k -family instead of an antichain, as done in [7]. Second, if we consider the weighted poset (P, v) where $v(x)$ is defined to be $1/|P_i|$ for $x \in P_i$, then the same argument will yield the LYM version or, equivalently, the NM version of the theorem, of which a direct and short proof was given in [16]. We state these results in the following theorem.

THEOREM 2.2. *P has the Sperner (strong Sperner, LYM or NM) property if and only if $(P/\Gamma, w_\Gamma)$ does.*

A remarkable step in this direction was given by Harper [10] (see also [11]), who introduced the concept of flow morphism between two bipartite graphs. For the purpose of this paper, we adopt the terminologies in [3, p. 157].

DEFINITION 2.3. *Let (P, w) and (Q, v) be two weighted posets. A flow morphism is a mapping $\phi : P \rightarrow Q$ such that*

- (i) ϕ is surjective,
- (ii) ϕ is order preserving and rank preserving, that is, that p_1 covers p_2 in P , which implies that $\phi(p_1)$ covers $\phi(p_2)$ in Q ,
- (iii) $v(q) = w(\phi^{-1}(q))$ for all $q \in Q$,
- (iv) the weighted poset induced by $\phi^{-1}(q_1) \cup \phi^{-1}(q_2)$ has rank 1 and the NM property for all $q_1, q_2 \in Q$ such that q_1 covers q_2 in Q .

By the fundamental lemma and note 2 in [10] we obtain the following theorem.

THEOREM 2.4. *Suppose that there is a flow morphism ϕ from (P, w) onto (Q, v) . Then (P, w) has the NM property if and only if (Q, v) does.*

Note that the fundamental lemma in [10] is contained in an earlier paper by Graham and Harper [6]. Also, as pointed out by a referee, the proof given in [13] works for Theorem 2.4 without changes if one replaces orbits under a group action by preimages of a flow morphism.

Note also that for P and Γ as in Theorem 2.1, the map ψ from P onto P/Γ satisfies conditions (i)–(iii) by definition, and condition (iv) by Theorem 2.2. Therefore, ψ is a flow morphism from P onto P/Γ .

To achieve our aims, we first find a group acting on $C[n, r, I]$ which preserves its rank and partial order and obtain a quotient poset of $C[n, r, I]$, written as $\overline{C}[n, r, I]$. Next, we construct a flow morphism from $\overline{C}[n, r, I]$ to another “smaller” poset. Then Theorems 2.2 and 2.4 are seen to be applicable.

In the remainder of this section we shall introduce the group, describe the quotient poset $\overline{C}[n, r, I]$, and determine its weight function.

Let $GL(V)$ denote the general linear group over V consisting of all invertible linear transformations of V . Set $X = (V_1, \dots, V_r)$ and let $GL(X|V)$ denote the subgroup of $GL(V)$ that fixes every V_i for $1 \leq i \leq r$. It is easily seen that $GL(X|V)$ induces an order preserving permutation group acting on $C[n, r, I]$.

For any nonnegative integers $i_1, i_2, \dots, i_r$, let $\langle i_1, i_2, \dots, i_r \rangle$ denote the set of all subspaces $A \in C[n, r, I]$ such that $\dim(A) = i_1 + i_2 + \dots + i_r$ and $\dim(A \cap V_j) = i_1 + i_2 + \dots + i_j$. Then $\langle i_1, i_2, \dots, i_r \rangle$ is an orbit of $GL(X|V)$ and

$$C[n, r, I] = \bigcup_{i_j \in \mathcal{I}_j} \langle i_1, i_2, \dots, i_r \rangle.$$

Then we obtain the quotient poset as follows:

$$\overline{C}[n, r, I] = C[n, r, I]/GL(X|V) = \{\langle i_1, i_2, \dots, i_r \rangle : i_j \in \mathcal{I}_j\},$$

the ordering of which is defined as

$$\langle i_1, i_2, \dots, i_r \rangle \leq \langle i'_1, i'_2, \dots, i'_r \rangle \text{ if and only if } i_j \leq i'_j \text{ for } j = 1, \dots, r.$$

In order to obtain a suitable weight function for $\overline{C}[n, r, I]$ we introduce a few notations.

Define $\{1\} = 1$, $\{i\} = 1 + q + \dots + q^{i-1} = \frac{q^i - 1}{q - 1}$, and $\{i\}! = \{1\}\{2\}\dots\{i\}$. Then $\begin{bmatrix} n \\ m \end{bmatrix} := \frac{\{n\}!}{\{n-m\}!\{m\}!}$, the q -binomial coefficient, which enumerates the number of the m -subspaces of V . For convenience, we define $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1$ and $\begin{bmatrix} i \\ j \end{bmatrix} = 0$ for $i > n$ or $i < 0$.

PROPOSITION 2.5. *Let $\langle i_1, i_2, \dots, i_r \rangle$ be an orbit of $C[n, r, I]$ under $G(X|V)$ and set $\bar{i}_j = k_j - i_j$, $j = 1, \dots, r$. Then the cardinality of $\langle i_1, i_2, \dots, i_r \rangle$ is given by*

$$\# \langle i_1, i_2, \dots, i_r \rangle = \begin{bmatrix} k_1 \\ i_1 \end{bmatrix} \dots \begin{bmatrix} k_r \\ i_r \end{bmatrix} q^{i_2 \bar{i}_1 + i_3 (\bar{i}_1 + \bar{i}_2) + \dots + i_r (\bar{i}_1 + \dots + \bar{i}_{r-1})}.$$

Proof. Set $\ell = i_1 + \dots + i_r$. If $r = 1$, then the result is trivial. Suppose $r > 1$ and let Z be an arbitrarily selected subspace of V_{r-1} with $\dim(Z) = i_1 + \dots + i_{r-1}$. As a little generalization of Proposition 2.2 in [2], it is straightforward that there are $\begin{bmatrix} k_r \\ i_r \end{bmatrix} q^{i_r (\bar{i}_1 + \dots + \bar{i}_{r-1})}$ ℓ -dimensional subspaces Y of $V_r = V$ such that $Y \cap V_{r-1} = Z$. If $r = 2$, then $\dim(Z) = i_1$ and we have $\begin{bmatrix} k_1 \\ i_1 \end{bmatrix}$ choices for Z , which implies the results. Suppose $r > 2$ and the result holds for $r - 1$. This means that in V_{r-1} , there are $\begin{bmatrix} k_1 \\ i_1 \end{bmatrix} \dots \begin{bmatrix} k_{r-1} \\ i_{r-1} \end{bmatrix} q^{i_2 \bar{i}_1 + i_3 (\bar{i}_1 + \bar{i}_2) + \dots + i_{r-1} (\bar{i}_1 + \dots + \bar{i}_{r-2})}$ subspaces Z with $\dim(Z) = i_1 + \dots + i_{r-1}$ such that $\dim(Z \cap V_j) - \dim(Z \cap V_{j-1}) = i_j$ for $j = 1, \dots, r - 1$. $\square$

For $1 \leq t \leq r$, set $n_t = k_1 + \dots + k_t$, $I_t = (\mathcal{I}_1, \dots, \mathcal{I}_t)$ and $X_t = (V_1, \dots, V_t)$. Define $C[n_t, t, I_t] = \{K \in L(V_t) : \dim(K \cap V_i) - \dim(K \cap V_{i-1}) \in \mathcal{I}_i, i = 1, 2, \dots, t\}$ and $\bar{C}[n_t, t, I_t]$ the quotient poset of $C[n_t, t, I_t]$ under the action of the group $GL(X_t|V_t)$. Also, we define a weight function on $\bar{C}[n_t, t, I_t]$ as

$$w_t(\langle i_1, \dots, i_t \rangle) = \begin{bmatrix} k_1 \\ i_1 \end{bmatrix} \dots \begin{bmatrix} k_t \\ i_t \end{bmatrix} q^{i_2 \bar{i}_1 + i_3 (\bar{i}_1 + \bar{i}_2) + \dots + i_t (\bar{i}_1 + \dots + \bar{i}_{t-1})}.$$

For $0 \leq \ell \leq n_t$, put $\bar{C}[n_t, t, I_t]_\ell = \{\langle i_1, i_2, \dots, i_t \rangle \in \bar{C}[n_t, t, I_t] : i_1 + i_2 + \dots + i_t = \ell\}$ and let $W_{(\ell, t)}$ denote the weight of $\bar{C}[n_t, t, I_t]_\ell$, that is,

$$W_{(\ell, t)} = \sum_{\mu \in \bar{C}[n_t, t, I_t]_\ell} w_t(\mu).$$

3. Log concavity of $C[n, r, I]$. We first present two lemmas.

LEMMA 3.1 (see [1]). *For $m \leq l$,*

$$\begin{bmatrix} n \\ m \end{bmatrix} \begin{bmatrix} n \\ l \end{bmatrix} - q^{l-m+1} \begin{bmatrix} n \\ m-1 \end{bmatrix} \begin{bmatrix} n \\ l+1 \end{bmatrix}$$

has nonnegative coefficients as a polynomial in q .

LEMMA 3.2. *Let $\{x_i\}_{i=0}^\infty$ be a sequence of positive numbers and let I be a subinterval of $[0, k]$ for an integer $k > 0$. If $x_i^2 \geq qx_{i-1}x_{i+1}$ for $i \geq 1$, then $y_\ell^2 \geq qy_{\ell-1}y_{\ell+1}$, where y_ℓ ($\ell \geq 1$) is defined to be*

$$y_\ell = \sum_{i \in I} x_{\ell-i} \begin{bmatrix} k \\ i \end{bmatrix} q^{i(n-\ell+i)}.$$

Proof. Put $a_i = x_i q^{i^2/2}$ for $i \geq 0$ and $b_i = \begin{bmatrix} k \\ i \end{bmatrix} q^{in+i^2/2}$ if $i \in I$ and $b_i = 0$ if $i \geq 0$ but $i \notin I$. It is easy to check that both $\{a_i\}_{i=0}^\infty$ and $\{b_i\}_{i=0}^\infty$ are log concave. Set $z_\ell = \sum_{i=0}^\ell a_{\ell-i} b_i$ for $\ell \geq 0$. Then $\{z_\ell\}_{\ell=0}^\infty$ is log concave by the *convolution theorem* of log concave sequences (see [10] or [12]). Moreover, we have that

$$\begin{aligned} y_\ell &= \sum_{i \in I} x_{\ell-i} \begin{bmatrix} k \\ i \end{bmatrix} q^{i(n-\ell+i)} \\ &= \sum_{i=0}^\ell a_{\ell-i} b_i q^{-\ell^2/2} \\ &= z_\ell q^{-\ell^2/2}. \end{aligned}$$

The log concavity of $\{z_\ell\}_{\ell=0}^\infty$ implies that $y_\ell^2 \geq qy_{\ell-1}y_{\ell+1}$ for $\ell \geq 1$. $\square$

Recall that $W_{(\ell, t)}$ denotes the Whitney numbers of $C[n_t, t, I_t]$ for $1 \leq t \leq r$.

THEOREM 3.3. *We have that*

$$W_{(\ell, r)}^2 \geq qW_{(\ell-1, r)}W_{(\ell+1, r)}.$$

Hence $C[n, r, I]$ is log concave and unimodal.

Proof. We apply induction on r . When $r = 1$, $W_{(\ell, 1)} = \begin{bmatrix} k_1 \\ \ell \end{bmatrix}$. Then the statement holds true by Lemma 3.1. Suppose $r \geq 2$ and the statement holds for $r - 1$. By Proposition 2.5 we have that

$$\begin{aligned} W_{(\ell, r)} &= \sum_{i_1 + \dots + i_r = \ell} \begin{bmatrix} k_1 \\ i_1 \end{bmatrix} \dots \begin{bmatrix} k_r \\ i_r \end{bmatrix} q^{i_2 \bar{i}_1 + i_3(\bar{i}_1 + \bar{i}_2) + \dots + i_r(\bar{i}_1 + \dots + \bar{i}_{r-1})} \\ &= \sum_{j \in \mathcal{I}_r} W_{(\ell-j, r-1)} \begin{bmatrix} k_r \\ j \end{bmatrix} q^{j(n_{r-1} - \ell + j)}. \end{aligned}$$

Then the assertion follows from the induction and Lemma 3.2. $\square$

4. Normalized matching property of $C[n, r, I]$. We first prove a useful lemma.

LEMMA 4.1. *Let $\{a_i\}_{i=0}^\infty$ and $\{b_i\}_{i=0}^\infty$ be two sequences and a and b be two real numbers. Suppose k_1, k_2, r , and s are indices such that $k_1 \leq r \leq s \leq k_2$. Write*

$$S_k[r, s] = \sum_{r \leq i \leq s} a_i b_{k-i} q^{-(i+a)(k-i+b)}.$$

If $\{a_i\}_{i=0}^\infty$ and $\{b_i\}_{i=0}^\infty$ satisfy that $a_i^2 \geq qa_{i-1}a_{i+1}$ and $b_i^2 \geq qb_{i-1}b_{i+1}$ for $i \geq 1$, then

$$\frac{S_{k+1}[r, s+1]}{S_k[r, s]} \geq \frac{S_{k+1}[k_1, k_2+1]}{S_k[k_1, k_2]}.$$

Proof. By the condition that $a_i^2 \geq qa_{i-1}a_{i+1}$ and $b_i^2 \geq qb_{i-1}b_{i+1}$, we can deduce that

$$(2) \quad \frac{a_j b_{k+1-j} q^{-(j+a)(k+1-j+b)}}{a_j b_{k-j} q^{-(j+a)(k-j+b)}} \leq \frac{a_{j+1} b_{k-j} q^{-(j+1+a)(k-j+b)}}{a_{j+1} b_{k-j-1} q^{-(j+1+a)(k-j-1+b)}}$$

and

$$(3) \quad \frac{a_{j+1} b_{k-j} q^{-(j+1+a)(k-j+b)}}{a_j b_{k-j} q^{-(j+a)(k-j+b)}} \geq \frac{a_{j+2} b_{k-j-1} q^{-(j+2+a)(k-j-1+b)}}{a_{j+1} b_{k-j-1} q^{-(j+1+a)(k-j-1+b)}}.$$

If $k_1 < r$, then by (2) we have

$$\frac{S_{k+1}[r, s+1]}{S_k[r, s]} \leq \frac{S_{k+1}[r-1, s+1]}{S_k[r-1, s]} \leq \dots \leq \frac{S_{k+1}[k_1, s+1]}{S_k[k_1, s]}.$$

If $s < k_2$, then by (3) we have

$$\frac{S_{k+1}[r, s+1]}{S_k[r, s]} \geq \frac{S_{k+1}[r, s+2]}{S_k[r, s+1]} \geq \dots \geq \frac{S_{k+1}[r, k_2+1]}{S_k[r, k_2]}.$$

Then, for each r and s such $k_1 \leq r \leq s \leq k_2$, we have

$$\frac{S_{k+1}[r, s+1]}{S_k[r, s]} \geq \frac{S_{k+1}[k_1, s+1]}{S_k[k_1, s]} \geq \frac{S_{k+1}[k_1, k_2+1]}{S_k[k_1, k_2]}.$$

Thus the statement follows and we complete the proof. $\square$

THEOREM 4.2. $C[n, r, I]$ has the NM property.

Proof. We apply induction on r . When $r = 1$, $C[n, r, I]_\ell = L(V)_\ell$ for $\ell \in I$. Then the statement holds since $L(V)$ has the NM property. Suppose $r \geq 2$ and $C[n_t, t, I_t]$ has the NM property for all $t < r$. We now proceed to prove that $C[n, r, I]$ has the property. By Theorem 2.2 it suffices to prove that $(\overline{C}[n, r, I], w)$ has the NM property.

Let $\Sigma_r = \{i_1 + i_2 + \cdots + i_{r-1} : i_j \in \mathcal{I}_j\}$ and $Q_r = \{(k, j) : k \in \Sigma_r \text{ and } j \in \mathcal{I}_r\}$. Define an order " $\leq$ " on Q_r as follows:

$$(k, j) \leq (k', j') \text{ in } Q_r \text{ if and only if } k \leq k' \text{ and } j \leq j'.$$

Then $(Q_r, \leq)$ is a graded poset of the same rank of that of $\overline{C}[n_r, r, I_r]$.

An order-preserving and rank-preserving map from $\overline{C}[n_r, r, I_r]$ onto Q_r is defined as follows:

$$\phi_r(\langle i_1, i_2, \dots, i_r \rangle) = (i_1 + i_2 + \cdots + i_{r-1}, i_r),$$

and a weight function v_r on Q_r is defined by

$$\begin{aligned} v_r(k, j) &= \sum_{\mu \in \phi^{-1}(k, j)} w_r(\mu) = w_{r-1}(\overline{C}[n_{r-1}, r-1, I_{r-1}]_k) \begin{bmatrix} k_r \\ j \end{bmatrix} q^{j(\bar{i}_1 + \cdots + \bar{i}_{r-1})} \\ (4) \quad &= W_{(k, r-1)} \begin{bmatrix} k_r \\ j \end{bmatrix} q^{j(n_{r-1}-k)}, \end{aligned}$$

where

$$\phi_r^{-1}(k, j) = \{\langle i_1, \dots, i_{r-1}, j \rangle \in \overline{C}[n_r, r, I_r] : i_1 + \cdots + i_{r-1} = k\}.$$

By Theorem 2.4 it suffices to show that ϕ_r is a flow morphism from $\overline{C}[n, r, I]$ onto (Q_r, v_r) and (Q_r, v_r) has the NM property. For brevity, we write $\phi_r = \phi$, $w_r = w$, and $v_r = v$.

First, we show that ϕ is a flow morphism. By definition we see that ϕ satisfies the conditions (i), (ii), and (iii) in Definition 2.3. We prove that it satisfies the fourth one, that is, for any $(k, j), (k+1, j), (k, j+1) \in (Q_r, v)$ (if they, in fact, exist), $\phi^{-1}(k, j) \cup \phi^{-1}(k+1, j)$ and $\phi^{-1}(k, j) \cup \phi^{-1}(k, j+1)$ have the NM property.

- For each $\mu = \langle i_1, \dots, i_{r-1}, j \rangle \in \phi^{-1}(k, j) \cup \phi^{-1}(k+1, j)$, define $f(\mu) = \langle i_1, \dots, i_{r-1} \rangle$, written as $\bar{\mu}$. Then, it is easy to see that f is a poset-isomorphism from $\phi^{-1}(k, j) \cup \phi^{-1}(k+1, j)$ to $\overline{C}[n_{r-1}, r-1, I_{r-1}]_k \cup \overline{C}[n_{r-1}, r-1, I_{r-1}]_{k+1}$. Also, for each $\mu \in \phi^{-1}(k, j) \cup \phi^{-1}(k+1, j)$, by Proposition 2.5 we have that

$$w(\mu) = w_{r-1}(\bar{\mu}) \begin{bmatrix} k_r \\ j \end{bmatrix} q^{j(n_{r-1}-k')},$$

where $k' = k$ if $\mu \in \phi^{-1}(k, j)$ and $k' = k+1$ if $\mu \in \phi^{-1}(k+1, j)$. Then, for each $S \subseteq \phi^{-1}(k, j)$, it is easy to verify that

$$\frac{w(\nabla(S))}{w(S)} = q^{-j} \frac{w_{r-1}(\overline{\nabla(S)})}{w_{r-1}(\overline{S})},$$

from which, along with induction, we obtain that $\phi^{-1}(k, j) \cup \phi^{-1}(k+1, j)$ has the NM property.

- If $(k, j), (k, j+1) \in Q_r$, it is easy to see that for $\mu = \langle i_1, \dots, i_{r-1}, j \rangle \in \phi^{-1}(k, j)$, there exists only $\nu = \langle i_1, \dots, i_{r-1}, j+1 \rangle \in \phi^{-1}(k, j+1)$ which covers μ , and from Proposition 2.5 it follows that

$$\frac{w(\nu)}{w(\mu)} = \frac{\begin{bmatrix} k_r \\ j+1 \end{bmatrix}}{\begin{bmatrix} k_r \\ j \end{bmatrix}} q^{(n_{r-1}-k)},$$

a constant. This implies that $\phi^{-1}(k, j) \cup \phi^{-1}(k, j+1)$ has the NM property. Hence ϕ is a flow morphism.

Next, we show that (Q_r, v) has the NM property.

Let $Q_{r,\ell} = \{(i, j) \in Q_r : i+j = \ell\}$, written as $\{(i_0, j_0), (i_0-1, j_0+1), \dots, (i_1, j_1)\}$. Let $T = \{(a, \ell-a), (a-1, \ell-a+1), \dots, (b, \ell-b)\} \subseteq Q_{r,\ell}$, where $i_1 \leq b \leq a \leq i_0$. If $(a+1, \ell-a) \notin Q_{r,\ell+1-a}$, then $W_{(r-1, a+1)} = 0$. For $i \geq 0$, set

$$\begin{bmatrix} k_r \\ i \end{bmatrix}' = \begin{cases} \begin{bmatrix} k_r \\ i \end{bmatrix} & \text{if } i \in \mathcal{I}_r, \\ 0 & \text{otherwise.} \end{cases}$$

Then, we have that

$$v(T) = \sum_{b \leq i \leq a} W_{(i, r-1)} \begin{bmatrix} k_r \end{bmatrix} \ell - i' q^{-(i-n_{r-1})(\ell-i)}$$

and

$$v(\nabla(T)) = \sum_{b \leq i \leq a+1} W_{(i, r-1)} \begin{bmatrix} k_r \end{bmatrix} \ell + 1 - i' q^{-(i-n_{r-1})(\ell+1-i)}.$$

Moreover, it is easy to find that $\begin{bmatrix} k_r \\ i \end{bmatrix}'^2 \geq q \begin{bmatrix} k_r \\ i-1 \end{bmatrix}' \begin{bmatrix} k_r \\ i+1 \end{bmatrix}'$ for all $i \geq 1$. Then, by Theorem 3.3 and Lemmas 3.1 and 4.1 we obtain that

$$\begin{aligned} \frac{v(\nabla(T))}{v(T)} &= \frac{\sum_{b \leq i \leq a+1} W_{(i, r-1)} \begin{bmatrix} k_r \end{bmatrix} \ell + 1 - i' q^{-(i-n_{r-1})(\ell+1-i)}}{\sum_{b \leq i \leq a} W_{(i, r-1)} \begin{bmatrix} k_r \end{bmatrix} \ell - i' q^{-(i-n_{r-1})(\ell-i)}} \\ &\geq \frac{\sum_{i_1 \leq i \leq i_0+1} W_{(i, r-1)} \begin{bmatrix} k_r \end{bmatrix} \ell + 1 - i' q^{-(i-n_{r-1})(\ell+1-i)}}{\sum_{i_1 \leq i \leq i_0} W_{(i, r-1)} \begin{bmatrix} k_r \end{bmatrix} \ell - i' q^{-(i-n_{r-1})(\ell-i)}} \\ (5) \quad &= \frac{v(Q_{r, \ell+1})}{v(Q_{r, \ell})}. \end{aligned}$$

Given a subset M of $Q_{r,\ell}$, let $M = T_1 \cup T_2 \cup \dots \cup T_r$ ($r \geq 1$), where $T_k = \{(a_k, \ell - a_k), (a_k - 1, \ell - a_k + 1), \dots, (b_k, \ell - b_k)\}$ ($k = 1, 2, \dots, r$), $i_1 \leq b_1 \leq a_1 < b_2 \leq a_2 < \dots < b_r \leq a_r \leq i_0$. Then $T_i \cap T_j = \emptyset$ and $\nabla(T_i) \cap \nabla(T_j) = \emptyset$ for $i \neq j$. We thus obtain that $v(M) = v(T_1) + \dots + v(T_r)$ and $v(\nabla(M)) = v(\nabla(T_1)) + \dots + v(\nabla(T_r))$. Then the inequality $\frac{v(\nabla(M))}{v(M)} \geq \frac{v(Q_{r, \ell+1})}{v(Q_{r, \ell})}$ follows from (5). Thus (Q_r, v) has the NM property, which completes the proof. $\square$

Acknowledgments. The authors wish to express their deep indebtedness to David Torney and two anonymous referees for many valuable suggestions and comments.

REFERENCES

- [1] L. M. BUTLER, *The q -log-concavity of q -binomial coefficients*, J. Combin. Theory Ser. A, 54 (1990), pp. 54–63.
- [2] W. Y. C. CHEN AND G.-C. ROTA, *q -analogs of the inclusion-exclusion principle and permutations with restricted position*, Discrete Math., 104 (1992), pp. 7–22.
- [3] K. ENGEL, *Sperner Theory*, Cambridge University Press, Cambridge, UK, 1997.
- [4] P. L. ERDÖS, U. FAIGLE, AND W. KERN, *A group-theoretic setting for intersecting Sperner families*, Combin. Probab. Comput., 1 (1992), pp. 323–334.
- [5] R. FREESE, *An application of Dilworth's lattice of maximal antichains*, Discrete Math., 7 (1974), pp. 107–109.
- [6] R. L. GRAHAM AND L. H. HARPER, *Some results on matching in bipartite graphs*, SIAM J. Appl. Math., 17 (1969), pp. 1017–1022.
- [7] C. GREENE AND D. J. KLEITMAN, *Proof techniques in the finite sets*, in Studies in Combinatorics, G.-C. Rota, ed., MAA Stud. Math. 17, Math. Assoc. America, Washington, DC, 1978, pp. 22–79.
- [8] J. R. GRIGGS, *On chains and Sperner k -families in ranked posets*, J. Combin. Theory Ser. A, 28 (1980), pp. 156–168.
- [9] J. R. GRIGGS, *Collections of subsets with the Sperner property*, Trans. Amer. Math. Soc., 269 (1982), pp. 575–591.
- [10] L. H. HARPER, *The morphology of partially ordered sets*, J. Combin. Theory Ser. A, 17 (1974), pp. 44–58.
- [11] L. H. HARPER, *The global theory of flows in networks*, Adv. Appl. Math., 1 (1980), pp. 158–181.
- [12] W. N. HSIEH AND D. J. KLEITMAN, *Normalized matching in direct products of partial orders*, Studies in Appl. Math., 52 (1973), pp. 285–289.
- [13] D. KLEITMAN, M. EDELBURG, AND D. LUBELL, *Maximal sized antichains in partial orders*, Discrete Math., 1 (1971), pp. 47–53.
- [14] K. W. LIH, *Sperner families over a subset*, J. Combin. Theory Ser. A, 29 (1980), pp. 182–185.
- [15] J. WANG, *Proof of a conjecture on the Sperner property of the subgroup lattice of an abelian p -group*, Ann. Comb., 2 (1998), pp. 85–101.
- [16] J. WANG, *Normalized matching property of the subgroup lattice of an abelian p -group*, Discrete Math., 257 (2002), pp. 559–574.

MATROID PATHWIDTH AND CODE TRELLIS COMPLEXITY*

NAVIN KASHYAP†

Abstract. We relate the notion of matroid pathwidth to the minimum trellis state-complexity (which we term trellis-width) of a linear code and to the pathwidth of a graph. By reducing from the problem of computing the pathwidth of a graph, we show that the problem of determining the pathwidth of a representable matroid is NP-hard. Consequently, the problem of computing the trellis-width of a linear code is also NP-hard. For a finite field $\mathbb{F}$, we also consider the class of $\mathbb{F}$ -representable matroids of pathwidth at most w , and correspondingly, the family of linear codes over $\mathbb{F}$ with trellis-width at most w . These are easily seen to be minor-closed. Since these matroids (and codes) have branchwidth at most w , a result of Geelen and Whittle shows that such matroids (and the corresponding codes) are characterized by finitely many excluded minors. We provide the complete list of excluded minors for $w = 1$ and give a partial list for $w = 2$.

Key words. matroids, pathwidth, linear codes, trellis complexity, NP-hard

AMS subject classifications. 05B35, 94B05

DOI. 10.1137/070691152

1. Introduction. The notion of pathwidth of a matroid has received some recent attention in the matroid theory literature [7], [9]. This notion has long been studied in coding theory literature, where it is used as a measure of trellis complexity of a linear code [13], [5], [16]. However, there appears to be no standard coding-theoretic nomenclature for this notion. It has been called the state complexity of a code in [10], but the use of this term there conflicts slightly with its use in [16]. So to avoid ambiguity, we will give it a new name here—*trellis-width*—which acknowledges its roots in trellis complexity.

The relationship between matroid pathwidth and code trellis-width can be made precise as follows. To an arbitrary linear code $\mathcal{C}$ over a finite field $\mathbb{F}$, we associate a matroid, $M(\mathcal{C})$, which is simply the vector matroid, over $\mathbb{F}$, of any generator matrix of the code. Recall that in coding theory, a matrix G is called a generator matrix of a code $\mathcal{C}$ if $\mathcal{C}$ is the row space of G . Consequently, the matroid $M(\mathcal{C})$ does not depend on the actual choice of the generator matrix and, therefore, is a characteristic of the code $\mathcal{C}$. The code $\mathcal{C}$ may in fact be viewed as a representation over $\mathbb{F}$ of the matroid $M(\mathcal{C})$. The trellis-width of $\mathcal{C}$ is simply the pathwidth of $M(\mathcal{C})$; we will give the precise definition of matroid pathwidth in section 2.2.

It has repeatedly been conjectured in coding theory literature that computing the trellis-width of a linear code over a fixed finite field $\mathbb{F}$ is NP-hard [10], [11], [16, section 5]. This would imply that the corresponding decision problem (over a fixed finite field $\mathbb{F}$)—given a generator matrix for a code $\mathcal{C}$ over $\mathbb{F}$, and a positive integer w , deciding whether or not the trellis-width of $\mathcal{C}$ is at most w —is NP-complete. This decision problem has been given various names—“Maximum Partition Rank Permutation” [10], “Maximum Width” [11], and “Trellis State-Complexity” [16].

*Received by the editors May 9, 2007; accepted for publication (in revised form) July 25, 2007; published electronically February 27, 2008. This work was supported in part by a research grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

<http://www.siam.org/journals/sidma/22-1/69115.html>

†Department of Mathematics and Statistics, Queen’s University, Kingston, ON, K7L 3N6, Canada (nkashyap@mast.queensu.ca).

An equivalent statement of the trellis-width conjecture is the following:
 given a matrix A over $\mathbb{F}$, the problem of computing the pathwidth of
 the vector matroid $M[A]$ is NP-hard.

In this paper, we prove the above statement for any fixed field $\mathbb{F}$, not necessarily finite. Our proof is by reduction from the problem of computing the pathwidth of a graph, which is known to be NP-hard [1], [2]. Thus, in particular, computing the trellis-width of a linear code over $\mathbb{F}$ is NP-hard, which settles the aforementioned coding-theoretic conjecture.

The situation is rather different if we weaken the trellis-width decision problem above by *not* considering the integer w to be a part of the input to the problem. In other words, for a fixed finite field $\mathbb{F}$ and a *fixed* integer $w > 0$, consider the following problem:

given a length- n linear code $\mathcal{C}$ over $\mathbb{F}$, decide whether or not $\mathcal{C}$ has
 trellis-width at most w .

The equivalent decision problem for matroid pathwidth would be to decide (for a fixed finite field $\mathbb{F}$ and integer $w > 0$) whether or not a given $\mathbb{F}$ -representable matroid has pathwidth at most w . Based on results from the structure theory of matroids [6], we strongly believe that these problems are solvable in polynomial time.

In the process of studying matroids of bounded pathwidth, we observe that for any finite field $\mathbb{F}_q = GF(q)$ and integer $w > 0$, the class, $\mathcal{P}_{w,q}$, of $\mathbb{F}_q$ -representable matroids having pathwidth at most w , is minor-closed and has finitely many excluded minors. As a relatively easy exercise, we show that the list of excluded minors for $\mathcal{P}_{1,q}$ consists of¹ $U_{2,4}$, $M(K_4)$, $M(K_{2,3})$, and $M^*(K_{2,3})$. Unfortunately, the problem of finding excluded-minor characterizations of $\mathcal{P}_{w,q}$ for $w > 1$ becomes difficult very quickly. We give a list of excluded minors for $\mathcal{P}_{2,q}$, which is probably not complete.

The rest of the paper is organized as follows. In section 2, we lay down the definitions and notation used in the paper. In section 3, we prove that, for any fixed field $\mathbb{F}$, the problem of computing the pathwidth of an $\mathbb{F}$ -representable matroid is NP-hard, and therefore, so is the problem of computing the trellis-width of a linear code over $\mathbb{F}$. Finally, in section 4, we consider the class of matroids $\mathcal{P}_{w,q}$. We give the complete lists of excluded minors for $\mathcal{P}_{1,q}$ and the corresponding family of linear codes over $\mathbb{F}_q$ having trellis-width at most one. We also give a partial list of excluded minors for $\mathcal{P}_{2,q}$.

2. Preliminaries. We assume familiarity with the basic definitions and notation of matroid theory, as expounded by Oxley [14]. The main results and proofs in this paper will be given in the language of matroid theory, rather than that of coding theory, as it is easier to do so. However, as our results may be of some interest to coding theorists, we make an effort in this section to provide the vocabulary necessary to translate the language of matroid theory into that of coding theory. Definitions of coding-theoretic terms not explicitly defined here can be found in any text on coding theory (e.g., [12]).

2.1. Codes and their associated matroids. Let $\mathcal{C}$ be a linear code of length n over the finite field $\mathbb{F}_q = GF(q)$. The dimension of $\mathcal{C}$ is denoted by $\dim(\mathcal{C})$, and the coordinates of $\mathcal{C}$ are indexed by the integers from the set $[n] = \{1, 2, \dots, n\}$ as usual. We will also associate with the coordinates of $\mathcal{C}$ a set, $E(\mathcal{C})$, of *coordinate labels*, so

¹In this paper, we take the connectivity function of a matroid M with ground set E and rank function r to be $\lambda_M(X) = r(X) + r(E - X) - r(E)$ for $X \subseteq E$. Therefore, what we consider to be matroids of pathwidth one would be matroids of pathwidth two in [7], [9].

that there is a bijection $\alpha_C : [n] \rightarrow E(C)$. The *label sequence* of C is defined to be the n -tuple $(\alpha_1, \alpha_2, \dots, \alpha_n)$, where $\alpha_i = \alpha_C(i)$. For notational convenience, we will simply let α_C denote the label sequence of C . Unless specified otherwise (as in the case of code minors and duals below), we will, by default, set $E(C)$ to be $[n]$, and α_C to be the n -tuple $(1, 2, 3, \dots, n)$. In such a case, the label of each coordinate is the same as its index.

Given a code C over $\mathbb{F}_q$, specified by a generator matrix G , we define its *associated matroid* $M(C)$ to be the vector matroid, $M[G]$, of G . We identify the ground set of $M(C)$ with $E(C)$. Note that if G and G' are distinct generator matrices of the code C , then $M[G] = M[G']$, and hence, $M(C)$ is independent of the choice of generator matrix. Thus, any generator matrix of C is an $\mathbb{F}_q$ -representation of $M(C)$.

Conversely, if M is an $\mathbb{F}_q$ -representable matroid and G is an $\mathbb{F}_q$ -representation of M , then $M = M(C)$ for the code C generated by G . Thus, each $\mathbb{F}_q$ -representable matroid is associated with some code C over $\mathbb{F}_q$.

For any code C , the dual code, $C^\perp$, is specified to have the same label sequence as C , i.e., $\alpha_{C^\perp} = \alpha_C$. It is a particularly nice fact [14, Theorem 2.2.8] that the matroids associated with C and $C^\perp$ are dual to each other, i.e., $M(C^\perp) = (M(C))^* \stackrel{\text{def}}{=} M^*(C)$.

Given a $J \subseteq E(C)$, we will denote by $C \setminus J$ (resp., C/J) the code obtained from C by puncturing (resp., shortening at) those coordinates having labels in J . Thus, $C/J = (C^\perp \setminus J)^\perp$. A *minor* of C is a code of the form $C/X \setminus Y$ for disjoint subsets $X, Y \subseteq E(C)$. A minor of C that is not C itself is called a *proper minor* of C . The coordinates of a minor of C retain their labels from $E(C)$. More precisely, we set $E(C/X \setminus Y) = E(C) - (X \cup Y)$, and take the label sequence of $C/X \setminus Y$ to be the $(n - |X \cup Y|)$ -tuple obtained from $\alpha_C = (\alpha_1, \alpha_2, \dots, \alpha_n)$ by simply removing those entries that are in $X \cup Y$. The operations of puncturing and shortening correspond to the matroid-theoretic operations of deletion and contraction, respectively: for $J \subseteq E(C)$,

$$M(C \setminus J) = M(C) \setminus J \quad \text{and} \quad M(C/J) = M(C)/J.$$

We will find it convenient to use $C|_J$ to denote the restriction of C to the coordinates with labels in J , i.e., $C|_J = C \setminus J^c$, where J^c denotes the set difference $E(C) - J$. This allows us to express the rank function, $r : E(C) \rightarrow \mathbb{Z}$, of the matroid $M(C)$ as follows: for $J \subseteq E(C)$, $r(J) = \dim(C|_J)$.

Two length- n linear codes C and C' over $\mathbb{F}_q$ are defined to be *equivalent* if there is an $n \times n$ permutation matrix Π and an invertible $n \times n$ diagonal matrix Δ such that C' is the image of C under the vector space isomorphism $\phi : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^n$ defined by $\phi(\mathbf{x}) = (\Pi\Delta)\mathbf{x}$. Informally, C' is equivalent to C if C' can be obtained by first multiplying the coordinates of C by some nonzero elements of $\mathbb{F}_q$, and then applying a coordinate permutation. In such a case, we write $C \equiv C'$. The equivalence class of codes equivalent to C will be denoted by $[C]$. It is clear that if codes C and C' are equivalent, then their associated matroids are isomorphic.

We remark that code equivalence has been defined above according to the coding-theoretic convention. Note that, under this definition, if C' is obtained by applying an automorphism of the field $\mathbb{F}_q$ to C , then C and C' would, in general, be considered to be inequivalent.

A family, $\mathfrak{C}$, of codes over $\mathbb{F}_q$ is said to be *minor-closed* if, for each $C \in \mathfrak{C}$, any code equivalent to a minor of C is also in $\mathfrak{C}$. A code, $\mathcal{D}$, over $\mathbb{F}_q$ is said to be an *excluded minor* for a minor-closed family $\mathfrak{C}$ if $\mathcal{D} \notin \mathfrak{C}$, but every proper minor of $\mathcal{D}$ is

in $\mathfrak{C}$. It is easily verified that if $\mathfrak{C}$ is a minor-closed family, then a code $\mathcal{C}$ is in $\mathfrak{C}$ iff no minor of $\mathcal{C}$ is an excluded minor for $\mathfrak{C}$.

Given a collection, $\mathcal{M}$, of $\mathbb{F}_q$ -representable matroids, define the code family

$$(1) \quad \mathfrak{C}(\mathcal{M}) = \{\mathcal{C} : \mathcal{C} \text{ is a linear code over } \mathbb{F}_q \text{ such that } M(\mathcal{C}) \in \mathcal{M}\}.$$

Evidently, if $\mathcal{M}$ is a minor-closed class of $\mathbb{F}_q$ -representable matroids, then $\mathfrak{C}(\mathcal{M})$ is also minor-closed. In this case, if $\mathcal{F}$ is the set of all excluded minors for $\mathcal{M}$, then $\mathfrak{C}(\mathcal{F})$ is the set of all excluded minors for $\mathfrak{C}(\mathcal{M})$.

2.2. Pathwidth, trellis-width, and branchwidth. The definitions in this section rely on the notion of the connectivity function of a matroid. Let M be a matroid with ground set $E(M)$ and rank function r_M . Its *connectivity function*, λ_M , is defined by $\lambda_M(X) = r_M(X) + r_M(E(M) - X) - r_M(E(M))$ for $X \subseteq E(M)$. Note that $\lambda_M(X) = \lambda_M(E(M) - X)$ and $\lambda_M(E(M)) = \lambda_M(\emptyset) = 0$. It should be pointed out that in the matroid theory literature, the prevalent definition of the connectivity function adds a “+1” to the expression we have given. We have chosen not to follow suit in order that we can give a minimum-fuss definition of trellis-width below.

The connectivity function is nonnegative, i.e., $\lambda_M(X) \geq 0$ for all $X \subseteq E(M)$, and submodular, i.e., $\lambda_M(X \cup Y) + \lambda_M(X \cap Y) \leq \lambda_M(X) + \lambda_M(Y)$ for all $X, Y \subseteq E(M)$. It is monotone under the action of taking minors—if N is a minor of M , then for all $X \subseteq E(N)$, $\lambda_N(X) \leq \lambda_M(X)$. Finally, the connectivity function of a matroid is identical to that of its dual, i.e., $\lambda_M(X) = \lambda_{M^*}(X)$ for all $X \subseteq E(M)$.

Given an ordering $(e_1, e_2, \dots, e_n)$ of the elements of M , define the *width* of the ordering to be $w_M(e_1, e_2, \dots, e_n) = \max_{i \in [n]} \lambda_M(e_1, e_2, \dots, e_i)$. (For simplicity of notation, we use $\lambda_M(e_1, e_2, \dots, e_i)$ instead of $\lambda_M(\{e_1, e_2, \dots, e_i\})$.) The *pathwidth* of M is defined as $\text{pw}(M) = \min w_M(e_1, e_2, \dots, e_n)$, the minimum being taken over all orderings $(e_1, e_2, \dots, e_n)$ of $E(M)$. An ordering $(e_1, e_2, \dots, e_n)$ of $E(M)$ such that $w_M(e_1, e_2, \dots, e_n) = \text{pw}(M)$ is called an *optimal* ordering.

Since $\lambda_M \equiv \lambda_{M^*}$, it is clear that $\text{pw}(M) = \text{pw}(M^*)$. Another useful and easily verifiable property of pathwidth is that, for matroids M_1 and M_2 , the pathwidth of their direct sum, $\text{pw}(M_1 \oplus M_2)$, equals $\max\{\text{pw}(M_1), \text{pw}(M_2)\}$. The property of pathwidth most important for our purposes is stated in the following lemma.

LEMMA 2.1. *If N is a minor of M , then $\text{pw}(N) \leq \text{pw}(M)$.*

Proof. Let $(e_1, \dots, e_n)$ be an optimal ordering of $E(M)$. It is enough to show the result in the case when $N = M \setminus e_i$ or $N = M/e_i$ for some $i \in [n]$. In such a case, consider the ordering $(e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n)$ of $E(N)$. For $j \in \{1, \dots, i-1\}$, we have $\lambda_N(e_1, \dots, e_j) \leq \lambda_M(e_1, \dots, e_j)$. For $j \in \{i+1, \dots, n\}$, we have

$$\begin{aligned} \lambda_N(e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_j) &= \lambda_N(e_{j+1}, \dots, e_n) \\ &\leq \lambda_M(e_{j+1}, \dots, e_n) = \lambda_M(e_1, \dots, e_j). \end{aligned}$$

It follows that $w_N(e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n) \leq w_M(e_1, \dots, e_n) = \text{pw}(M)$, and hence, $\text{pw}(N) \leq \text{pw}(M)$. $\square$

The *trellis-width* of a linear code $\mathcal{C}$ over $\mathbb{F}_q$ is defined to be $\text{tw}(\mathcal{C}) = \text{pw}(M(\mathcal{C}))$. For a discussion of the motivation and practical implications of this definition, we refer the reader to [16, section 5].

The pathwidth of a matroid is an upper bound on its branchwidth, a more well-known measure of matroid complexity. The branchwidth of a matroid is defined via cubic trees. A *cubic tree* is a tree in which the degree of any vertex is either one or three. The vertices of degree one are called *leaves*. A *branch-decomposition* of a

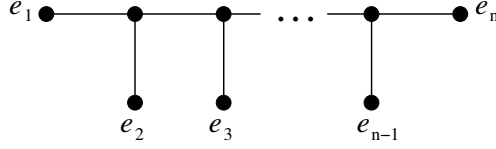


FIG. 1. A branch-decomposition of M having width equal to $w_M(e_1, e_2, \dots, e_n)$.

matroid M is a cubic tree, T , with $|E(M)|$ leaves, labelled in a one-to-one fashion by the elements of M . Each edge e of such a branch-decomposition T connects two subtrees of T , so $T \setminus e$ has two components. We say that edge e *displays* a subset $X \subseteq E(M)$ if X is the set of labels of leaves of one of the components of $T \setminus e$. The *width* of an edge e of T is defined to be $\lambda_M(X)$, where X is one of the label sets displayed by e . The *width* of T is the maximum among the widths of its edges.

The *branchwidth* of M is the minimum among the widths of all its branch-decompositions. Note that if T is the branch-decomposition of M shown in Figure 1, then the width of T is precisely $w_M(e_1, e_2, \dots, e_n)$. Indeed, the width of any edge of T is either $\lambda_M(e_i)$ or $\lambda_M(e_1, \dots, e_i)$ for some $i \in [n]$. Now, for any $i \in [n]$,

$$\begin{aligned} \lambda_M(e_1, \dots, e_{i-1}) + \lambda_M(e_1, \dots, e_i) &= \lambda_M(e_i, e_{i+1}, \dots, e_n) + \lambda_M(e_1, \dots, e_i) \\ &\geq \lambda_M(E(M)) + \lambda_M(e_i) = \lambda_M(e_i), \end{aligned}$$

the inequality above arising from the submodularity of λ_M . Since $\lambda_M(e_i) \in \{0, 1\}$, either $\lambda_M(e_1, \dots, e_{i-1})$ or $\lambda_M(e_1, \dots, e_i)$ is at least as large as $\lambda_M(e_i)$. Therefore, the width of T is given by $\max_{i \in [n]} \lambda_M(e_1, \dots, e_i) = w_M(e_1, \dots, e_n)$. It follows that the branchwidth of M is bounded from above by $\text{pw}(M)$.

3. NP-hardness of matroid pathwidth and code trellis-width. In this section, we prove that for any fixed field $\mathbb{F}$, the problem of computing the pathwidth of an $\mathbb{F}$ -representable matroid M , given a representation of M over $\mathbb{F}$, is NP-hard. We accomplish this by reduction from the known NP-hard problem of computing the pathwidth of a graph [1], [2].

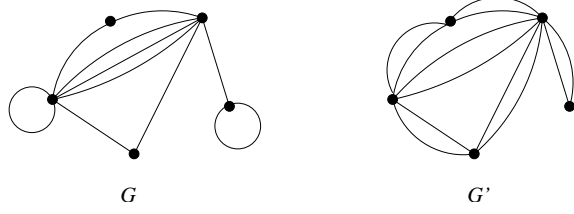
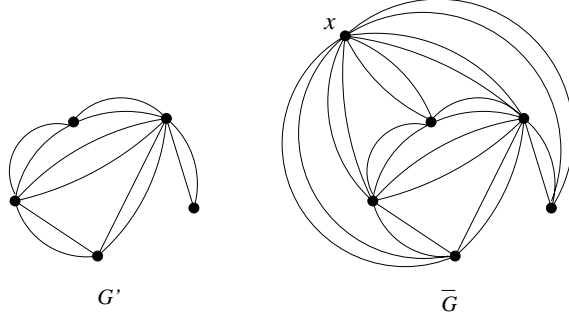
The notion of graph pathwidth was introduced by Robertson and Seymour in [15]. Let $\mathcal{G}$ be a graph with vertex set V . An ordered collection $\mathcal{V} = (V_1, \dots, V_t)$, $t \geq 1$, of subsets of V is called a *path-decomposition* of $\mathcal{G}$ if

- (i) $\bigcup_{i=1}^t V_i = V$;
- (ii) for each pair of adjacent vertices $u, v \in V$, we have $\{u, v\} \subseteq V_i$ for some $i \in [t]$; and
- (iii) for $1 \leq i < j < k \leq t$, $V_i \cap V_k \subseteq V_j$.

The *width* of such a path-decomposition $\mathcal{V}$ is defined to be $w_{\mathcal{G}}(\mathcal{V}) = \max_{i \in [t]} |V_i| - 1$. The *pathwidth* of $\mathcal{G}$, denoted by $\text{pw}(\mathcal{G})$, is the minimum among the widths of all its path-decompositions. A path-decomposition $\mathcal{V}$ such that $w_{\mathcal{G}}(\mathcal{V}) = \text{pw}(\mathcal{G})$ is called an *optimal* path-decomposition of $\mathcal{G}$.

Let $\mathbb{F}$ be an arbitrary field. Given a graph $\mathcal{G}$ with vertex set V , our aim is to produce, in time polynomial in $|V|$, a matrix A over $\mathbb{F}$ such that $\text{pw}(\mathcal{G})$ can be directly computed from $\text{pw}(M[A])$. The NP-hardness of computing graph pathwidth then implies the NP-hardness of computing the pathwidth of an $\mathbb{F}$ -representable matroid.

The obvious idea of taking A to be a representation of the cycle matroid of $\mathcal{G}$ does not work. As observed by Robertson and Seymour [15], trees can have arbitrarily large pathwidth; however, the cycle matroid of any tree is $U_{n,n}$ for some n , and $\text{pw}(U_{n,n}) = 0$. What actually turns out to work is to take A to be a representation of


 FIG. 2. Construction of $\mathcal{G}'$ from $\mathcal{G}$.

 FIG. 3. Construction of $\bar{\mathcal{G}}$ from $\mathcal{G}'$.

the cycle matroid of a certain graph constructible from $\mathcal{G}$ in polynomial time, as we describe next.

Let $\mathcal{G}'$ be a graph defined on the same vertex set, V , as $\mathcal{G}$, having the following properties (see Figure 2):

- (P1) $\mathcal{G}'$ is loopless;
- (P2) a pair of distinct vertices is adjacent in $\mathcal{G}'$ iff it is adjacent in $\mathcal{G}$; and
- (P3) in $\mathcal{G}'$, there are exactly two edges between each pair of adjacent vertices.

It is evident from the definition that $(V_1, \dots, V_t)$ is a path-decomposition of $\mathcal{G}$ iff it is a path-decomposition of $\mathcal{G}'$. Therefore, $\text{pw}(\mathcal{G}') = \text{pw}(\mathcal{G})$.

Define $\bar{\mathcal{G}}$ to be the graph obtained by adding an extra vertex, henceforth denoted by x to $\mathcal{G}'$, along with a pair of parallel edges from x to each $v \in V$ (see Figure 3). Clearly, $\bar{\mathcal{G}}$ is constructible directly from $\mathcal{G}$ in $O(|V|^2)$ time. But more importantly, the pathwidth of the cycle matroid, $M(\bar{\mathcal{G}})$, of $\bar{\mathcal{G}}$ relates very simply to the pathwidth of $\mathcal{G}$, as made precise by the following proposition.

PROPOSITION 3.1. $\text{pw}(M(\bar{\mathcal{G}})) = \text{pw}(\mathcal{G}) + 1$.

Before proving the result, we present some of its implications. For any field $\mathbb{F}$, $M(\bar{\mathcal{G}})$ is $\mathbb{F}$ -representable. Indeed, if $D(\bar{\mathcal{G}})$ is any directed graph obtained by arbitrarily assigning orientations to the edges of $\bar{\mathcal{G}}$, then the vertex-arc incidence matrix of $D(\bar{\mathcal{G}})$ is an $\mathbb{F}$ -representation of $M(\bar{\mathcal{G}})$ [14, Proposition 5.1.2]. It is easily verified that such an $\mathbb{F}$ -representation of $M(\bar{\mathcal{G}})$ can be constructed directly from $\mathcal{G}$ in $O(|V|^3)$ time. Now, suppose there were a polynomial-time algorithm for computing the pathwidth of an arbitrary $\mathbb{F}$ -representable matroid, given an $\mathbb{F}$ -representation for it. Then, given any graph $\mathcal{G}$, we can construct an $\mathbb{F}$ -representation, A , of $M(\bar{\mathcal{G}})$, and then compute the pathwidth of $M[A] = M(\bar{\mathcal{G}})$, all in polynomial time. Therefore, by Proposition 3.1, we have a polynomial-time algorithm to compute the pathwidth of $\mathcal{G}$. However, the graph pathwidth problem is NP-hard. So, if there exists a polynomial-time algorithm for it, then we must have $P = NP$. This implies the following result.

THEOREM 3.2. *Let $\mathbb{F}$ be a fixed field. The problem of computing the pathwidth of $M[A]$, for an arbitrary matrix A over $\mathbb{F}$, is NP-hard.*

As a corollary, we have that computing the trellis-width of a code is NP-hard.

COROLLARY 3.3. *Let $\mathbb{F}$ be a fixed finite field. The problem of computing the trellis-width of an arbitrary linear code over $\mathbb{F}$, specified by any of its generator matrices, is NP-hard.*

The remainder of this section is devoted to the proof of Proposition 3.1. Since $\text{pw}(\mathcal{G}') = \text{pw}(\mathcal{G})$, for the purpose of our proof, we may assume that $\mathcal{G}' = \mathcal{G}$. Thus, from now until the end of this section, we take $\mathcal{G}$ to be a loopless graph satisfying property (P3) above. Note that $\bar{\mathcal{G}}$ also satisfies (P3). For each pair of adjacent vertices u, v in $\mathcal{G}$ or $\bar{\mathcal{G}}$, we denote by l_{uv} and r_{uv} the two edges between u and v . Let V and E denote the sets of vertices and edges of $\mathcal{G}$, and let $\bar{V}$ and $\bar{E}$ denote the corresponding sets of $\bar{\mathcal{G}}$. We thus have $\bar{V} = V \dot{\cup} \{x\}$ and $\bar{E} = E \dot{\cup} (\bigcup_{v \in V} \{l_{xv}, r_{xv}\})$.

Set $M = M(\bar{\mathcal{G}})$, so that $E(M) = \bar{E}$. Note that since $\bar{\mathcal{G}}$ is connected (each $v \in V$ is adjacent to x), we have $\text{rank}(M) = |\bar{V}| - 1 = |V|$.

We will first prove that $\text{pw}(M) \leq \text{pw}(\mathcal{G}) + 1$. Let $\mathcal{V} = (V_1, \dots, V_t)$ be a path-decomposition of $\mathcal{G}$. We need the following fact about $\mathcal{V}$: for each $j \in [t]$,

$$(2) \quad \bigcup_{i \leq j} V_i \cap \bigcup_{k \geq j} V_k = V_j.$$

The above equality follows from the fact that a path-decomposition, by definition, has the property that for $1 \leq i < j < k \leq t$, $V_i \cap V_k \subseteq V_j$.

For $j \in [t]$, let F_j be the set of edges of $\mathcal{G}$ that have both their endpoints in V_j . By condition (ii) in the definition of path-decomposition, $\bigcup_{j=1}^t F_j = E$. Now, let $\bar{F}_j = F_j \cup (\bigcup_{v \in V_j} \{l_{xv}, r_{xv}\})$, so that $\bigcup_{j=1}^t \bar{F}_j = \bar{E}$.

DEFINITION 3.1. *An ordering $(e_1, \dots, e_n)$ of the elements of a matroid M is said to induce an ordered partition $(E_1, \dots, E_t)$ of $E(M)$ if for each $j \in [t]$, $\{e_{n_{j-1}+1}, e_{n_{j-1}+2}, \dots, e_{n_j}\} = E_j$, where $n_j = |\bigcup_{i \leq j} E_i|$ (and $n_0 = 0$).*

Let $\pi = (e_1, \dots, e_n)$ be any ordering of $\bar{E}$ that induces the ordered partition $(E_1, E_2, \dots, E_t)$, where for each $j \in [t]$, $E_j = \bar{F}_j - \bigcup_{i < j} \bar{F}_i$. We claim that the width of π is at most one more than the width of the path-decomposition $\mathcal{V}$.

LEMMA 3.4. $w_M(\pi) \leq w_{\mathcal{G}}(\mathcal{V}) + 1$.

Proof. Observe first that

$$(3) \quad \begin{aligned} w_M(\pi) &= \max_{j \in [t]} \max_{1 \leq k \leq n_j - n_{j-1}} \lambda_M \left(\bigcup_{i < j} E_i \cup \{e_{n_{j-1}+1}, \dots, e_{n_{j-1}+k}\} \right) \\ &\leq \max_{j \in [t]} \max_{E' \subseteq E_j} \lambda_M \left(\bigcup_{i < j} E_i \cup E' \right). \end{aligned}$$

Let $X = \bigcup_{i < j} E_i \cup E'$ for some $j \in [t]$ and $E' \subseteq E_j$, and consider $\lambda_M(X) = r_M(X) + r_M(\bar{E} - X) - r_M(\bar{E})$. Since $\bar{\mathcal{G}}$ is a connected graph, $r_M(\bar{E}) = |\bar{V}| - 1 = |V|$.

If v is a vertex of $\bar{\mathcal{G}}$ incident with an edge in X , then $v \in \bigcup_{i \leq j} V_i \dot{\cup} \{x\}$. So, the subgraph of $\bar{\mathcal{G}}$ induced by X has its vertices contained in $\bigcup_{i \leq j} V_i \dot{\cup} \{x\}$. Therefore, $r_M(X) \leq |\bigcup_{i \leq j} V_i \dot{\cup} \{x\}| - 1 = |\bigcup_{i \leq j} V_i|$.

Next, consider $\bar{E} - X = (\bigcup_{k>j} E_k) \cup (E_j - E')$. Reasoning as before, the subgraph of $\mathcal{G}$ induced by $\bar{E} - X$ has its vertices contained in $\bigcup_{k \geq j} V_k \cup \{x\}$. Hence, $r_M(\bar{E} - X) \leq |\bigcup_{k \geq j} V_k|$.

Therefore, we have

$$\begin{aligned} \lambda_M(X) &\leq \left| \bigcup_{i \leq j} V_i \right| + \left| \bigcup_{k \geq j} V_k \right| - |V| \\ &= \left| \bigcup_{i \leq j} V_i \cap \bigcup_{k \geq j} V_k \right| = |V_j|, \end{aligned}$$

the last equality arising from (2). Hence, carrying on from (3),

$$w_M(\pi) \leq \max_{j \in [t]} |V_j| = w_{\mathcal{G}}(\mathcal{V}) + 1,$$

as desired. $\square$

The fact that $\text{pw}(M) \leq \text{pw}(\mathcal{G}) + 1$ easily follows from the above lemma. Indeed, we may choose $\mathcal{V}$ to be an optimal path-decomposition of $\mathcal{G}$. Then, by Lemma 3.4, there exists an ordering $(e_1, \dots, e_n)$ of $E(M)$ such that $w_M(e_1, \dots, e_n) \leq \text{pw}(\mathcal{G}) + 1$. Hence, $\text{pw}(M) \leq w_M(e_1, \dots, e_n) \leq \text{pw}(\mathcal{G}) + 1$.

We prove the reverse inequality in two steps, first showing that $\text{pw}(\bar{\mathcal{G}}) = \text{pw}(\mathcal{G}) + 1$, and then showing that $\text{pw}(M) \geq \text{pw}(\bar{\mathcal{G}})$.

LEMMA 3.5. $\text{pw}(\bar{\mathcal{G}}) = \text{pw}(\mathcal{G}) + 1$.

Proof. Clearly, if $\mathcal{V} = (V_1, \dots, V_t)$ is a path-decomposition of $\mathcal{G}$, then $\bar{\mathcal{V}} = (V_1 \cup \{x\}, \dots, V_t \cup \{x\})$ is a path-decomposition of $\bar{\mathcal{G}}$. Hence, choosing $\mathcal{V}$ to be an optimal path-decomposition of $\mathcal{G}$, we have that $\text{pw}(\bar{\mathcal{G}}) \leq w_{\bar{\mathcal{G}}}(\bar{\mathcal{V}}) = w_{\mathcal{G}}(\mathcal{V}) + 1 = \text{pw}(\mathcal{G}) + 1$.

For the inequality in the other direction, we will show that there exists an optimal path-decomposition, $\tilde{\mathcal{V}} = (\tilde{V}_1, \dots, \tilde{V}_s)$, of $\bar{\mathcal{G}}$ such that $x \in \tilde{V}_i$ for all $i \in [s]$. We then have $\mathcal{V} = (\tilde{V}_1 - \{x\}, \dots, \tilde{V}_s - \{x\})$ being a path-decomposition of $\mathcal{G}$, and hence, $\text{pw}(\mathcal{G}) \leq w_{\mathcal{G}}(\mathcal{V}) = w_{\bar{\mathcal{G}}}(\tilde{\mathcal{V}}) - 1 = \text{pw}(\bar{\mathcal{G}}) - 1$.

Let $\bar{\mathcal{V}} = (\bar{V}_1, \dots, \bar{V}_t)$ be an optimal path-decomposition of $\bar{\mathcal{G}}$, and let $i_0 = \min\{i : x \in \bar{V}_i\}$ and $i_1 = \max\{i : x \in \bar{V}_i\}$. Since $\bar{V}_i \cap \bar{V}_k \subseteq \bar{V}_j$ for $i < j < k$, we must have $x \in \bar{V}_i$ for each $i \in [i_0, i_1]$.

We claim that $(\bar{V}_{i_0}, \bar{V}_{i_0+1}, \dots, \bar{V}_{i_1})$ is a path-decomposition of $\bar{\mathcal{G}}$. We only have to show that $\bigcup_{i=i_0}^{i_1} \bar{V}_i = \bar{V}$, and that for each pair of adjacent vertices $u, v \in \bar{V}$, $\{u, v\} \subseteq \bar{V}_i$ for some $i \in [i_0, i_1]$. To see why the first assertion is true, consider any $v \in \bar{V}$, $v \neq x$. Since x is adjacent to v and $\bar{\mathcal{V}}$ is a path-decomposition of $\bar{\mathcal{G}}$, $\{x, v\} \subseteq \bar{V}_i$ for some $i \in [t]$. However, $x \in \bar{V}_i$ iff $i \in [i_0, i_1]$, and so, $\{x, v\} \subseteq \bar{V}_i$ for some $i \in [i_0, i_1]$. In particular, $v \in \bar{V}_i$ for some $i \in [i_0, i_1]$.

For the second assertion, suppose that u, v is a pair of vertices adjacent in $\bar{\mathcal{G}}$. Obviously, $\{u, v\} \subseteq \bar{V}_j$ for some $j \in [t]$. Suppose that $j \notin [i_0, i_1]$. We consider the case when $j > i_1$; the case when $j < i_0$ is similar. As $\bigcup_{i=i_0}^{i_1} \bar{V}_i = \bar{V}$, there exist $i_2, i_3 \in [i_0, i_1]$ such that $u \in \bar{V}_{i_2}$ and $v \in \bar{V}_{i_3}$. Without loss of generality, $i_2 \leq i_3$. If $i_2 = i_3$, then there exists $i \in [i_0, i_1]$ such that $\{u, v\} \subseteq \bar{V}_i$. If $i_2 < i_3$, then we have $u \in \bar{V}_{i_2} \cap \bar{V}_j$ and $i_2 < i_3 < j$. Hence, $u \in \bar{V}_{i_3}$ as well, and so once again, we have an $i \in [i_0, i_1]$ such that $\{u, v\} \subseteq \bar{V}_i$.

Thus, $(\bar{V}_{i_0}, \bar{V}_{i_0+1}, \dots, \bar{V}_{i_1})$ is a path-decomposition of $\bar{\mathcal{G}}$, with the property that $x \in \bar{V}_i$ for all $i \in [i_0, i_1]$. It must be an optimal path-decomposition, since it is a subsequence of the optimal path-decomposition $\bar{\mathcal{V}}$. $\square$

To complete the proof of Proposition 3.1, it remains to show that $\text{pw}(M) \geq \text{pw}(\overline{\mathcal{G}})$. We introduce some notation at this point. Recall that the two edges between a pair of adjacent vertices u and v in $\overline{\mathcal{G}}$ (or $\mathcal{G}$) are denoted by l_{uv} and r_{uv} . We define

$$\begin{aligned} L_{\mathcal{G}} &= \{l_{uv} : u, v \text{ are adjacent vertices in } \mathcal{G}\}, \\ R_{\mathcal{G}} &= \{r_{uv} : u, v \text{ are adjacent vertices in } \mathcal{G}\}, \end{aligned}$$

$L_x = \bigcup_{v \in V} \{l_{xv}\}$, and $R_x = \bigcup_{v \in V} \{r_{xv}\}$, where x is the distinguished vertex in $\overline{V} - V$. Thus, $L_{\mathcal{G}} \cup R_{\mathcal{G}} = E$ and $E \cup L_x \cup R_x = \overline{E}$. Note that, by construction of $\overline{\mathcal{G}}$, $\text{cl}_M(L_x) = \text{cl}_M(R_x) = \overline{E}$, where cl_M denotes the closure operator of M .

We will need the fact that there exists an optimal ordering $(e_1, \dots, e_n)$ of $\overline{E}$ that induces a certain ordered partition of $\overline{E}$ of the form

$$(L_1, A_1, B_1, R_1, L_2, A_2, B_2, R_2, \dots, L_t, A_t, B_t, R_t),$$

where for each $j \in [t]$, $L_j \subseteq L_x$, $A_j \subseteq L_{\mathcal{G}}$, $B_j \subseteq R_{\mathcal{G}}$, and $R_j \subseteq R_x$. This will follow from a reordering argument given below. But first, we make some simple observations about orderings of $\overline{E}$. Given an ordering of $\overline{E}$, we may assume, without loss of generality, that for each pair of adjacent vertices $u, v \in V$, l_{uv} appears before r_{uv} in the ordering; we denote this by $l_{uv} < r_{uv}$. We call such an ordering of $\overline{E}$ a *normal* ordering.

LEMMA 3.6. *Let $(e_1, \dots, e_n)$ be a normal ordering of $\overline{E}$. Then, for $1 \leq j \leq n-1$, we have*

- (a) $\lambda_M(e_1, \dots, e_{j+1}) = \lambda_M(e_1, \dots, e_j) + 1$ iff $e_{j+1} \notin \text{cl}_M(e_1, \dots, e_j)$ and
- (b) $\lambda_M(e_1, \dots, e_{j+1}) = \lambda_M(e_1, \dots, e_j) - 1$ iff $e_{j+1} \notin \text{cl}_M(e_{j+2}, \dots, e_n)$.

Proof. We prove only (a), as the proof of (b) is similar. It is easy to deduce from the definition of the connectivity function that $\lambda_M(e_1, \dots, e_{j+1}) = \lambda_M(e_1, \dots, e_j) + 1$ iff $e_{j+1} \notin \text{cl}_M(e_1, \dots, e_j)$ and $e_{j+1} \in \text{cl}_M(e_{j+2}, \dots, e_n)$.

Now, if $e_{j+1} \notin \text{cl}_M(e_1, \dots, e_j)$, then $e_{j+1} = l_{uv}$ for some u, v . (If not, i.e., if $e_{j+1} = r_{uv}$, then since $l_{uv} < r_{uv}$, we must have $l_{uv} \in \{e_1, \dots, e_j\}$, and so, $e_{j+1} = r_{uv} \in \text{cl}_M(l_{uv}) \subseteq \text{cl}_M(e_1, \dots, e_j)$, a contradiction.) Therefore, $\{e_{j+2}, \dots, e_n\}$ contains r_{uv} , and hence, $e_{j+1} = l_{uv} \in \text{cl}_M(e_{j+2}, \dots, e_n)$. We have thus shown that if $e_{j+1} \notin \text{cl}_M(e_1, \dots, e_j)$, then $e_{j+1} \in \text{cl}_M(e_{j+2}, \dots, e_n)$. Part (a) of the lemma now follows. $\square$

We now describe a procedure that takes as input a normal ordering of $\overline{E}$ and produces as output a reordering of $\overline{E}$ with certain desirable properties.

REORDERING ALGORITHM.

Input: a normal ordering $(e_1, \dots, e_n)$ of $\overline{E}$.

Initialization: $j = 0$.

Step 0: If $j = 0$, set $X_j = \emptyset$;
else, set $X_j = \text{cl}_M(e_1, \dots, e_j) - \{e_1, \dots, e_j\}$.

Step 1: If $X_j = \emptyset$,
find the least $k > j$ such that
for some $m > j$, $e_m \in L_x \cap \text{cl}_M(e_1, \dots, e_k)$;
set $(e'_1, \dots, e'_n) = (e_1, \dots, e_j, e_m, e_{j+1}, \dots, e_{m-1}, e_{m+1}, \dots, e_n)$.
If $X_j \neq \emptyset$,
if $L_x \cap X_j \neq \emptyset$, find an $m > j$ such that $e_m \in L_x \cap X_j$;
else, if $L_{\mathcal{G}} \cap X_j \neq \emptyset$, find an $m > j$ such that $e_m \in L_{\mathcal{G}} \cap X_j$;

else, if $R_G \cap X_j \neq \emptyset$, find an $m > j$ such that $e_m \in R_G \cap X_j$;
 else, if $R_x \cap X_j \neq \emptyset$, find an $m > j$ such that $e_m \in R_x \cap X_j$;
 set $(e'_1, \dots, e'_n) = (e_1, \dots, e_j, e_m, e_{j+1}, \dots, e_{m-1}, e_{m+1}, \dots, e_n)$.

Step 2: Replace j by $j + 1$.

If $j < n$, replace $(e_1, \dots, e_n)$ by $(e'_1, \dots, e'_n)$
 and return to Step 0;
 else, output $(e'_1, \dots, e'_n)$.

Denote by $(e_1^*, \dots, e_n^*)$ the final output generated by the above algorithm. Set $X_0^* = \emptyset$, and for $j \in [n]$, $X_j^* = \text{cl}_M(e_1^*, \dots, e_j^*) - \{e_1^*, \dots, e_j^*\}$. Stepping through the algorithm, one may easily check that $(e_1^*, \dots, e_n^*)$ has the following property: for $0 \leq j \leq n - 1$, if $X_j^* = \emptyset$, then $e_{j+1}^* \in L_x$, and if $X_j^* \neq \emptyset$, then

$$e_{j+1}^* \in \begin{cases} L_x \cap X_j^* & \text{if } L_x \cap X_j^* \neq \emptyset, \\ L_G \cap X_j^* & \text{if } L_x \cap X_j^* = \emptyset, \text{ but } L_G \cap X_j^* \neq \emptyset, \\ R_G \cap X_j^* & \text{if } L_x \cap X_j^* = L_G \cap X_j^* = \emptyset, \text{ but } R_G \cap X_j^* \neq \emptyset, \\ R_x \cap X_j^* & \text{if } L_x \cap X_j^* = L_G \cap X_j^* = R_G \cap X_j^* = \emptyset, \text{ but } R_x \cap X_j^* \neq \emptyset. \end{cases}$$

The following claim can be readily deduced from this property, and we leave the details to the reader.

CLAIM 3.7. (a) *The ordering $(e_1^*, \dots, e_n^*)$ induces an ordered partition of $\overline{E}$ of the form*

$$(L_1, A_1, B_1, R_1, L_2, A_2, B_2, R_2, \dots, L_t, A_t, B_t, R_t),$$

where for each $j \in [t]$, $L_j \subseteq L_x$, $A_j \subseteq L_G$, $B_j \subseteq R_G$, and $R_j \subseteq R_x$. Moreover, for each $u, v \in \overline{V}$, $l_{uv} \in L_j \cup A_j$ iff $r_{uv} \in B_j \cup R_j$.

(b) *For the ordered partition in (a), we have for each $j \in [t]$,*

$$A_j \cup B_j \subseteq \text{cl}_M \left(\bigcup_{i \leq j} L_i \right) - \text{cl}_M \left(\bigcup_{i < j} L_i \right).$$

The crucial property of $(e_1^*, \dots, e_n^*)$ is the following.

LEMMA 3.8. *If $(e_1^*, \dots, e_n^*)$ is the output of REORDERING ALGORITHM in response to the input $(e_1, \dots, e_n)$, then $w_M(e_1^*, \dots, e_n^*) \leq w_M(e_1, \dots, e_n)$.*

Proof. Steps 0–1 of the algorithm go through n iterations, indexed by $j \in \{0, 1, \dots, n - 1\}$. In the j th iteration, Step 1 is given a normal ordering $(e_1, \dots, e_n)$, in response to which it produces an ordering $(e'_1, \dots, e'_n)$, which is also normal. To prove the lemma, it is enough to show that $w_M(e'_1, \dots, e'_n) \leq w_M(e_1, \dots, e_n)$.

So, suppose that the algorithm is in its j th iteration ($0 \leq j \leq n - 1$). We first dispose of the case when $X_j \neq \emptyset$. Then,

$$(e'_1, \dots, e'_n) = (e_1, \dots, e_j, e_m, e_{j+1}, \dots, e_{m-1}, e_{m+1}, \dots, e_n)$$

for some $m > j$ such that $e_m \in X_j$. Observe that if $1 \leq s \leq j$ or if $m \leq s \leq n$, then $(e'_1, \dots, e'_s)$ is just a reordering of $(e_1, \dots, e_s)$, and hence, $\lambda_M(e'_1, \dots, e'_s) = \lambda_M(e_1, \dots, e_s)$.

So, consider $j < s < m$. In this case, $(e'_1, \dots, e'_s) = (e_1, \dots, e_j, e_m, e_{j+1}, \dots, e_{s-1})$. Since $e_m \in \text{cl}_M(e_1, \dots, e_j)$, we have $r_M(e'_1, \dots, e'_s) = r_M(e_1, \dots, e_j, e_{j+1}, \dots, e_{s-1})$.

On the other hand,

$$(4) \quad \begin{aligned} r_M(e'_{s+1}, \dots, e'_n) &= r_M(e_s, \dots, e_{m-1}, e_{m+1}, \dots, e_n) \\ &\leq r_M(e_s, \dots, e_{m-1}, e_m, e_{m+1}, \dots, e_n). \end{aligned}$$

Hence, $\lambda(e'_1, \dots, e'_s) \leq \lambda(e_1, \dots, e_{s-1})$. Therefore, for any $s \in [n]$, we have shown that there exists a $t \in [n]$ such that $\lambda(e'_1, \dots, e'_s) \leq \lambda(e_1, \dots, e_t)$. It follows that $w_M(e'_1, \dots, e'_n) \leq w_M(e_1, \dots, e_n)$.

We must now deal with the case when $X_j = \emptyset$, i.e., $\text{cl}_M(e_1, \dots, e_j) = \{e_1, \dots, e_j\}$. Note that if $L_x \subseteq \{e_1, \dots, e_j\}$, then since $\text{cl}_M(L_x) = \overline{E}$, we have $\text{cl}_M(e_1, \dots, e_j) = \overline{E}$. Therefore, $\{e_1, \dots, e_j\} = \overline{E}$, which means that $j = n$, a contradiction. Therefore, there must exist some $m > j$ such that $e_m \in L_x$.

Let k^* be the least integer $k > j$ such that there exists $e_m \in L_x \cap \text{cl}_M(e_1, \dots, e_k)$ for some $m > j$. By choice of k^* , we have $m \geq k^*$. For this m , we again have

$$(e'_1, \dots, e'_n) = (e_1, \dots, e_j, e_m, e_{j+1}, \dots, e_{m-1}, e_{m+1}, \dots, e_n).$$

As before, if $1 \leq s \leq j$ or if $m \leq s \leq n$, then $\lambda_M(e'_1, \dots, e'_s) = \lambda_M(e_1, \dots, e_s)$. For $k^* < s < m$, we have

$$r_M(e'_1, \dots, e'_s) = r_M(e_1, \dots, e_j, e_m, e_{j+1}, \dots, e_{s-1}) = r_M(e_1, \dots, e_j, e_{j+1}, \dots, e_{s-1}),$$

as $e_m \in \text{cl}_M(e_1, \dots, e_{k^*}) \subseteq \text{cl}_M(e_1, \dots, e_{s-1})$. And as in (4), $r_M(e'_{s+1}, \dots, e'_n) \leq r_M(e_s, \dots, e_n)$. Hence, $\lambda(e'_1, \dots, e'_s) \leq \lambda(e_1, \dots, e_{s-1})$.

We are left with $j+1 \leq s \leq k^*$. Note that by choice of k^* , $e_m \notin \text{cl}_M(e_1, \dots, e_{s-1})$. Therefore,

$$\begin{aligned} r_M(e'_1, \dots, e'_s) &= r_M(e_1, \dots, e_j, e_m, e_{j+1}, \dots, e_{s-1}) \\ &= 1 + r_M(e_1, \dots, e_j, e_{j+1}, \dots, e_{s-1}). \end{aligned}$$

Since (4) again applies, we have that

$$(5) \quad \lambda_M(e'_1, \dots, e'_s) \leq 1 + \lambda_M(e_1, \dots, e_{s-1}).$$

Observe that, since $e_{j+1} \notin \text{cl}_M(e_1, \dots, e_j)$, by Lemma 3.6(a),

$$(6) \quad \lambda_M(e_1, \dots, e_{j+1}) = \lambda_M(e_1, \dots, e_j) + 1.$$

Furthermore, by choice of k^* , $e_m \notin \text{cl}_M(e_1, \dots, e_{k^*-1})$, but $e_m \in \text{cl}_M(e_1, \dots, e_{k^*})$, which together imply that $e_{k^*} \notin \text{cl}_M(e_1, \dots, e_{k^*-1})$. Hence, again by Lemma 3.6(a),

$$(7) \quad \lambda_M(e_1, \dots, e_{k^*}) = \lambda_M(e_1, \dots, e_{k^*-1}) + 1.$$

Therefore, from (5)–(7), we find that for $s = j+1$ or $s = k^*$ we have $\lambda_M(e'_1, \dots, e'_s) \leq \lambda_M(e_1, \dots, e_s)$.

We claim that for $j+1 < s < k^*$ we have $\lambda_M(e_1, \dots, e_{s-1}) \leq \lambda_M(e_1, \dots, e_s)$, so that by induction, $\lambda_M(e_1, \dots, e_{s-1}) \leq \lambda_M(e_1, \dots, e_{k^*-1})$. This would then imply, via (5) and (7), that $\lambda_M(e'_1, \dots, e'_s) \leq \lambda_M(e_1, \dots, e_{k^*})$. Thus, for any $s \in [n]$, we have a $t \in [n]$ such that $\lambda(e'_1, \dots, e'_s) \leq \lambda(e_1, \dots, e_t)$. Therefore, $w_M(e'_1, \dots, e'_n) \leq w_M(e_1, \dots, e_n)$, which would complete the proof of the lemma.

To prove our claim, it is enough to show that when $j+1 < s < k^*$, we have $e_s \in \text{cl}_M(e_{s+1}, \dots, e_n)$. Indeed, it then follows from Lemma 3.6(b) that $\lambda_M(e_1, \dots, e_s) \geq$

$\lambda_M(e_1, \dots, e_{s-1})$. So, suppose that $e_s \notin \text{cl}_M(e_{s+1}, \dots, e_n)$ for some $j+1 < s < k^*$. Then, $e_s = r_{uv}$ for some $u, v \in \bar{V}$. (Otherwise, if $e_s = l_{uv}$, then since $r_{uv} > l_{uv}$, we would have $e_s \in \text{cl}_M(e_{s+1}, \dots, e_n)$.) Note that $l_{uv} \notin \{e_1, \dots, e_j\}$; otherwise, the fact that $\{e_1, \dots, e_j\}$ is a flat of M would imply that $r_{uv} \in \{e_1, \dots, e_j\}$. So, $l_{uv} \in \{e_{j+1}, \dots, e_{s-1}\}$.

Suppose that $e_s = r_{uv}$ for some $v \in \bar{V}$. Then, $l_{xv} \in \{e_{j+1}, \dots, e_{s-1}\}$, which contradicts the choice of k^* . Therefore, $e_s \notin R_x$, meaning that $e_s = r_{uv}$ for some $u, v \in V$.

Now, if $l_{xu}, l_{xv} \in \{e_{s+1}, \dots, e_n\}$, then $e_s \in \text{cl}_M(e_{s+1}, \dots, e_n)$, as (l_{xu}, l_{xv}, r_{uv}) is a triangle in $\bar{\mathcal{G}}$. So, without loss of generality, $l_{xu} \notin \{e_{s+1}, \dots, e_n\}$. By choice of k^* , $l_{xu} \notin \{e_{j+1}, \dots, e_s\}$. Therefore, $l_{xu} \in \{e_1, \dots, e_j\}$. But now, $l_{xv} \in \text{cl}_M(e_1, \dots, e_{s-1})$, as (l_{xu}, l_{uv}, l_{xv}) is a triangle in $\bar{\mathcal{G}}$. However, $l_{xv} \notin \{e_1, \dots, e_j\}$; otherwise, we would have $l_{xu}, l_{xv} \in \{e_1, \dots, e_j\}$, which, since $\{e_1, \dots, e_j\}$ is a flat and (l_{xu}, l_{uv}, l_{xv}) is a triangle, would imply that $l_{uv} \in \{e_1, \dots, e_j\}$. Thus, $l_{xv} = e_{m^*}$ for some $m^* > j$. As already noted, $l_{xv} \in \text{cl}_M(e_1, \dots, e_{s-1})$, and so once again, our choice of k^* is contradicted.

Therefore, our assumption that $e_s \notin \text{cl}_M(e_{s+1}, \dots, e_n)$ always leads to a contradiction, from which we conclude that the assumption is false. This completes the proof of the lemma. $\square$

We can now furnish the last remaining piece of the proof of Proposition 3.1.

LEMMA 3.9. $pw(M) \geq pw(\bar{\mathcal{G}})$.

Proof. Let $(e_1, \dots, e_n)$ be an optimal ordering of $\bar{E}$. Without loss of generality, $(e_1, \dots, e_n)$ may be assumed to be normal. Let $(e_1^*, \dots, e_n^*)$ be the output of REORDERING ALGORITHM in the response to the input $(e_1, \dots, e_n)$. Then, $(e_1^*, \dots, e_n^*)$ has the properties listed in Claim 3.7, and, by Lemma 3.8, is also an optimal ordering of $\bar{E}$.

Now, $(e_1^*, \dots, e_n^*)$ induces an ordered partition $(L_1, A_1, B_1, R_1, \dots, L_t, A_t, B_t, R_t)$ of $\bar{E}$, as in Claim 3.7(a). For $j \in [t]$, define $Y_j = \bigcup_{i < j} (L_i \cup A_i \cup B_i \cup R_i) \cup (L_j \cup A_j)$, and $Y'_j = \bar{E} - Y_j = \bigcup_{i > j} (L_i \cup A_i \cup B_i \cup R_i) \cup (B_j \cup R_j)$. Letting $\bar{\mathcal{G}}[Y_j]$ and $\bar{\mathcal{G}}[Y'_j]$ denote the subgraphs of $\bar{\mathcal{G}}$ induced by Y_j and Y'_j , respectively, set $V_j = V(\bar{\mathcal{G}}[Y_j]) \cap V(\bar{\mathcal{G}}[Y'_j])$. In other words, V_j is the set of vertices common to both $\bar{\mathcal{G}}[Y_j]$ and $\bar{\mathcal{G}}[Y'_j]$. It is easily checked that $\mathcal{V} = (V_1, \dots, V_t)$ is a path-decomposition of $\bar{\mathcal{G}}$. Note that

$$|V_j| = |V(\bar{\mathcal{G}}[Y_j])| + |V(\bar{\mathcal{G}}[Y'_j])| - |\bar{V}|.$$

We next observe that $\bar{\mathcal{G}}[Y_j]$ and $\bar{\mathcal{G}}[Y'_j]$ are connected graphs. From Claim 3.7(b), we have that $Y_j \subseteq \text{cl}_M(\bigcup_{i \leq j} L_i)$. Therefore, for any edge l_{uv} (or r_{uv}) in $Y_j - \bigcup_{i \leq j} L_i$, both l_{xu} and l_{xv} must be in some L_i , $i \leq j$. Thus, in $\bar{\mathcal{G}}[Y_j]$, each vertex $v \neq x$ is adjacent to x , which shows that $\bar{\mathcal{G}}[Y_j]$ is connected.

Consider any vertex $v \neq x$ in $\bar{\mathcal{G}}[Y'_j]$ such that $r_{xv} \notin Y'_j$. Then, $r_{uv} \in Y'_j$ for some $u \neq x$. So, $r_{uv} \in B_k$ for some $k \geq j$. By Claim 3.7(b), $r_{uv} \in \text{cl}_M(\bigcup_{i \leq k} L_i) - \text{cl}_M(\bigcup_{i < k} L_i)$. This implies that either $l_{xu} \in L_k$ or $l_{xv} \in L_k$. Hence, either $r_{xu} \in R_k$ or $r_{xv} \in R_k$. However, r_{xv} cannot be in R_k , since $r_{xv} \notin Y'_j$, and so, $r_{xu} \in R_k$. Thus, (r_{xu}, r_{uv}) forms a path in $\bar{\mathcal{G}}[Y'_j]$ from x to v . It follows that $\bar{\mathcal{G}}[Y'_j]$ is connected.

Therefore,

$$\begin{aligned} \lambda_M(Y_j) &= r_M(Y_j) + r_M(Y'_j) - r_M(\bar{E}) \\ &= (|V(\bar{\mathcal{G}}[Y_j])| - 1) + (|V(\bar{\mathcal{G}}[Y'_j])| - 1) - (\bar{V} - 1) = |V_j| - 1. \end{aligned}$$

Hence,

$$\text{pw}(M) = w_M(e_1^*, \dots, e_n^*) \geq \max_{j \in [t]} \lambda_M(Y_j) = \max_{j \in [t]} |V_j| - 1 = w_{\overline{\mathcal{G}}}(\mathcal{V}) \geq \text{pw}(\overline{\mathcal{G}}),$$

which proves the lemma. $\square$

The proof of Proposition 3.1 is now complete.

4. Matroids of bounded pathwidth. Theorem 3.2 shows that the following decision problem is NP-complete.

Problem: MATROID PATHWIDTH.

Let $\mathbb{F}$ be a fixed field.

Instance: An $m \times n$ matrix A over $\mathbb{F}$ and an integer $w > 0$.

Question: Is there an ordering $(e_1, \dots, e_n)$ of the elements of $M = M[A]$ such that $w_M(e_1, \dots, e_n) \leq w$?

Similarly, Corollary 3.3 shows that the corresponding decision problem for code trellis-width (over a fixed finite field $\mathbb{F}$) is NP-complete.

In this section, we consider the situation when the parameter w above is a fixed constant and therefore not considered to be part of the problem instance. In contrast to the NP-completeness of MATROID PATHWIDTH, we believe that the following decision problem and its coding-theoretic counterpart are solvable in polynomial time.

Problem: WEAK MATROID PATHWIDTH.

Let $\mathbb{F}_q = GF(q)$ be a fixed finite field and w a fixed positive integer.

Instance: An $m \times n$ matrix A over $\mathbb{F}_q$.

Question: Is there an ordering $(e_1, \dots, e_n)$ of the elements of $M = M[A]$ such that $w_M(e_1, \dots, e_n) \leq w$?

Our optimism above stems from the fact that the property of having pathwidth bounded by w is preserved by the minors of a matroid. To be precise, let $\mathcal{P}_{w,q}$ be the class of matroids representable over the finite field $\mathbb{F}_q = GF(q)$, which have pathwidth of at most w . By Lemma 2.1, $\mathcal{P}_{w,q}$ is minor-closed. Since pathwidth is an upper bound on the branchwidth of a matroid, all matroids in $\mathcal{P}_{w,q}$ have branchwidth at most w . Now, Geelen and Whittle have shown that if $\mathcal{M}$ is any minor-closed class of $\mathbb{F}_q$ -representable matroids having bounded branchwidth, then $\mathcal{M}$ has finitely many excluded minors [8, Theorem 1.4]. As a result, we have the following theorem.

THEOREM 4.1. *For any integer $w > 0$ and finite field $\mathbb{F}_q$, $\mathcal{P}_{w,q}$ has finitely many excluded minors. Consequently, the code family*

$$\mathfrak{C}(\mathcal{P}_{w,q}) = \{\mathcal{C} : \mathcal{C} \text{ is a linear code over } \mathbb{F}_q \text{ such that } \text{tw}(\mathcal{C}) \leq w\}$$

also has finitely many excluded minors.

Theorem 4.1 shows that deciding whether or not a given $\mathbb{F}_q$ -representable matroid M belongs to $\mathcal{P}_{w,q}$ can be accomplished by testing whether or not M contains as a minor one of the finitely many excluded minors of $\mathcal{P}_{w,q}$. The Minor-Recognition Conjecture of Geelen, Gerards, and Whittle [6, Conjecture 1.3] states that, for any fixed $\mathbb{F}_q$ -representable matroid N , testing a given $\mathbb{F}_q$ -representable matroid for the presence of an N -minor can be done in polynomial time. So, if this conjecture is true—and there is evidence to support its validity [6]—then membership of an $\mathbb{F}_q$ -representable matroid in the class $\mathcal{P}_{w,q}$ can be decided in polynomial time. Hence, assuming the validity of the Minor-Recognition Conjecture, WEAK MATROID PATHWIDTH is solvable in polynomial time.

While the finiteness of the list of excluded minors for $\mathcal{P}_{w,q}$ implies, modulo the Minor-Recognition Conjecture, the existence of a polynomial-time algorithm for

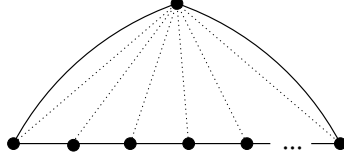


FIG. 4. An “umbrella” graph. A dotted line between a pair of vertices represents zero or more parallel edges between them.

WEAK MATROID PATHWIDTH, an actual implementation of such an algorithm would require the explicit determination of the excluded minors. As a relatively easy exercise, we prove the following theorem.

THEOREM 4.2. *A matroid is in $\mathcal{P}_{1,q}$ iff it contains no minor isomorphic to any of the matroids $U_{2,4}$, $M(K_4)$, $M(K_{2,3})$, and $M^*(K_{2,3})$.*

We first verify the easy “only if” part of the above theorem.

PROPOSITION 4.3. *$U_{2,4}$, $M(K_4)$, $M(K_{2,3})$, and $M^*(K_{2,3})$ are not in $\mathcal{P}_{1,q}$.*

Proof. If (e_1, e_2, e_3, e_4) is any ordering of the elements of $M = U_{2,4}$, then $\lambda_M(e_1, e_2) = r_M(e_1, e_2) + r_M(e_3, e_4) - \text{rank}(M) = 2 + 2 - 2 = 2$. It follows that $\text{pw}(U_{2,4}) = 2$.

Now consider $M = M(K_4)$. For any ordering $(e_1, \dots, e_6)$ of $E(K_4)$, we have $r_M(e_1, e_2, e_3) \geq 2$, with equality iff $\{e_1, e_2, e_3\}$ is a triangle, in which case $\{e_4, e_5, e_6\}$ is a triad. It follows that $r_M(e_1, e_2, e_3) + r_M(e_4, e_5, e_6) \geq 5$. Hence, $w_M(e_1, \dots, e_6) \geq \lambda_M(e_1, e_2, e_3) \geq 2$.

The proof for $M = M(K_{2,3})$ is very similar. For any $J \subseteq E(K_{2,3})$ with $|J| = 3$, $r_M(J) = 3$, since $K_{2,3}$ has no circuits of size less than 4. Therefore, for any ordering $(e_1, \dots, e_6)$ of $E(K_{2,3})$, $w_M(e_1, \dots, e_6) \geq \lambda_M(e_1, e_2, e_3) = 3 + 3 - 4 = 2$. Thus, $\text{pw}(M(K_{2,3})) \geq 2$, and by duality, $\text{pw}(M^*(K_{2,3})) \geq 2$ as well. $\square$

We now prove the “if” part of Theorem 4.2. For the duration of the proof, we take M to be a matroid that contains no minor isomorphic to the matroids listed in the statement of the theorem. Since $M(K_4)$ is a minor of each of the matroids F_7 , F_7^* , $M(K_5)$, $M^*(K_5)$, $M(K_{3,3})$, and $M^*(K_{3,3})$, M contains none of these as minors. Therefore, $M = M(\mathcal{G})$ for some planar graph $\mathcal{G}$ (cf. [14, Theorem 13.3.1 and Proposition 5.2.6]). Evidently, we may take $\mathcal{G}$ to be connected as a graph.

Since $\mathcal{P}_{1,q}$ is closed under direct sums, we may assume that M is 2-connected. Therefore, either $\mathcal{G}$ is a graph consisting of a single vertex with a self-loop incident with it, or $\mathcal{G}$ is a loopless graph. In the former case, $M \cong U_{0,1}$, which is in $\mathcal{P}_{1,q}$. So, we may assume that $\mathcal{G}$ is loopless. If $\mathcal{G}$ has exactly two vertices, then $M \cong U_{1,n}$ for some n , which is also in $\mathcal{P}_{1,q}$. Hence, we may assume that $|V(\mathcal{G})| \geq 3$, in which case $\mathcal{G}$ is 2-connected as a graph [14, Corollary 8.2.2]. Moreover, if $\mathcal{G}^*$ is any geometric dual of $\mathcal{G}$, then, since $M^* = M(\mathcal{G}^*)$ is 2-connected, by the same argument as above, we may assume that $\mathcal{G}^*$ is also 2-connected as a graph.

At this point, we need the following definition. We call a graph an *umbrella* if it is of the form shown in Figure 4. Formally, an umbrella is a graph H that consists of a circuit on $m + 1$ vertices $u_0, u_1, \dots, u_m$, and in addition, for each $i \in [m]$, zero or more parallel edges between u_0 and u_i . Note that $H - u_0$ is a simple path, where $H - u_0$ denotes the graph obtained from H by deleting the vertex u_0 and all edges incident with it.

Returning to our proof, we have $M = M(\mathcal{G})$ for a loopless, 2-connected, planar graph $\mathcal{G}$, such that any geometric dual of $\mathcal{G}$ is also 2-connected.

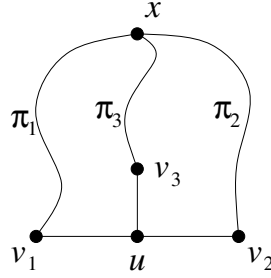


FIG. 5. If $\mathcal{G}^* - x$ has a vertex of degree at least 3, then $\mathcal{G}^*$ has a $K_{2,3}$ minor.

LEMMA 4.4. $\mathcal{G}$ has a geometric dual $\mathcal{G}^*$ that is isomorphic to an umbrella.

We prove the lemma using the concept of an outerplanar graph. A planar graph is said to be *outerplanar* if it has a planar embedding in which every vertex lies on the exterior (unbounded) face. We will refer to such a planar embedding of the graph as an *outerplanar embedding*. Outerplanar graphs were characterized by Chartrand and Harary [3] as graphs that do not contain K_4 or $K_{2,3}$ as a minor.

Proof of Lemma 4.4. Since $M(\mathcal{G})$ contains no $M(K_4)$ - or $M(K_{2,3})$ -minor, $\mathcal{G}$ cannot contain K_4 or $K_{2,3}$ as a minor. Therefore, by the Chartrand–Harary result mentioned above, $\mathcal{G}$ is outerplanar. Let $\mathcal{G}^*$ be the geometric dual of an outerplanar embedding of $\mathcal{G}$.

Let x be the vertex of $\mathcal{G}^*$ corresponding to the exterior face of the outerplanar embedding of $\mathcal{G}$. By a result of Fleischner, Geller, and Harary [4, Theorem 1], $\mathcal{G}^* - x$ is a forest. In fact, since $\mathcal{G}^*$ is 2-connected, $\mathcal{G}^* - x$ is a tree.

We claim that no vertex of $\mathcal{G}^* - x$ has degree greater than two, and hence, $\mathcal{G}^* - x$ is a simple path. Indeed, suppose that $\mathcal{G}^* - x$ has a vertex u adjacent to three other vertices v_1, v_2, v_3 . Since $\mathcal{G}^*$ is 2-connected, there are paths π_1, π_2 , and π_3 in $\mathcal{G}^*$ from v_1, v_2 , and v_3 , respectively, to x that do not pass through u . Also, since $\mathcal{G}^* - x$ is a tree, these paths must be internally disjoint in $\mathcal{G}^*$. The graph $\mathcal{G}^*$ thus has a subgraph as depicted in Figure 5. But this subgraph is obviously contractible to $K_{2,3}$, and hence $\mathcal{G}^*$ has $K_{2,3}$ as a minor. However, this is impossible as $M^* = M(\mathcal{G}^*)$ does not have $M(K_{2,3})$ as a minor.

Thus, $\mathcal{G}^* - x$ is a simple path. The two degree-one vertices (endpoints) of this path must be adjacent to x in $\mathcal{G}^*$; otherwise, $\mathcal{G}^*$ is not 2-connected. It follows that $\mathcal{G}^*$ is isomorphic to an umbrella. $\square$

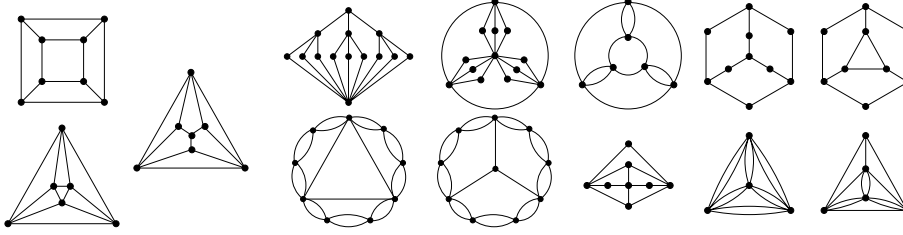
To complete the proof of Theorem 4.2, we show that $M(\mathcal{G}^*) \in \mathcal{P}_{1,q}$, so that by duality, $M = M^*(\mathcal{G}^*) \in \mathcal{P}_{1,q}$. This is done by the following lemma.

LEMMA 4.5. If H is an umbrella, then $M(H) \in \mathcal{P}_{1,q}$.

Proof. Let H be an umbrella on $m + 1$ vertices $u_0, u_1, \dots, u_m$, where u_0 is the vertex such that $H - \{u_0\}$ is a simple path. For $i \in [m]$, let E_i denote the set of edges between u_0 and u_i . Also, for $j \in [m - 1]$, let e_j denote the edge between u_j and u_{j+1} . Consider any ordering of the edges of H that induces the ordered partition

$$(E_1, e_1, E_2, e_2, \dots, E_{m-1}, e_{m-1}, E_m).$$

Let $J = (\bigcup_{i=1}^{j-1} (E_i \cup \{e_i\})) \cup X$, with $X \subseteq E_j$ (X may be empty). Note that the subgraph $H[J]$ of H induced by the edges in J is incident only with vertices in $\{u_0, u_1, \dots, u_j\}$. Therefore, setting $M = M(H)$, $r_M(J) = |V(H[J])| - 1 \leq j$. Similarly, the subgraph of H induced by the edges in $E(H) - J$ is incident only with vertices in $\{u_j, u_{j+1}, \dots, u_m, u_0\}$, and so, $r_M(E(H) - J) \leq m - j + 1$.


 FIG. 6. Some of the planar graphs whose cycle matroids are excluded minors for $\mathcal{P}_{2,q}$.

Thus, $\lambda_M(J) \leq j + (m - j + 1) - m = 1$, and it follows that $\text{pw}(M) \leq 1$. Being graphic, M is $\mathbb{F}_q$ -representable, and hence, $M \in \mathcal{P}_{1,q}$. $\square$

This completes the proof of Theorem 4.2.

As a corollary to the theorem, we give a coding-theoretic characterization of the code family $\mathfrak{C}(\mathcal{P}_{1,q})$. In coding theory, an $\mathbb{F}_q$ -representation of a uniform matroid is called a *maximum-distance separable (MDS) code*. For any field $\mathbb{F}_q$, the matrices G_4 , $G_{2,3}$, and $G_{2,3}^*$ below are $\mathbb{F}_q$ -representations of $M(K_4)$, $M(K_{2,3})$, and $M^*(K_{2,3})$, respectively,

$$G_4 = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 & 1 & -1 \end{bmatrix},$$

$$G_{2,3} = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}, \quad G_{2,3}^* = \begin{bmatrix} 1 & -1 & 0 & -1 & 1 & 0 \\ 1 & 0 & -1 & -1 & 0 & 1 \end{bmatrix}.$$

The matroids $M(K_4)$, $M(K_{2,3})$, and $M^*(K_{2,3})$, being binary, are uniquely representable over $\mathbb{F}_q$ in the matroid-theoretic sense [14, section 6.3 and Theorem 10.1.3]. We let $\mathcal{C}(K_4)$, $\mathcal{C}(K_{2,3})$, and $\mathcal{C}(K_{2,3})^\perp$ denote the codes over $\mathbb{F}_q$ generated by the matrices G_4 , $G_{2,3}$, and $G_{2,3}^*$, respectively.

COROLLARY 4.6. *Let $\mathbb{F}_q$ be an arbitrary finite field. A linear code $\mathcal{C}$ over $\mathbb{F}_q$ has trellis-width at most one iff it contains no minor equivalent to any of the following:*

- (i) a $[4, 2]$ MDS code;
- (ii) a code obtainable by applying an automorphism of $\mathbb{F}_q$ to one of the codes $\mathcal{C}(K_4)$, $\mathcal{C}(K_{2,3})$, and $\mathcal{C}(K_{2,3})^\perp$.

The problem of finding the complete set of excluded minors for $\mathcal{P}_{w,q}$ quickly becomes difficult for $w > 1$. The main obstacle is that we may only assume the basic property of 2-connectedness for such excluded minors. The class $\mathcal{P}_{w,q}$ is not even closed under 2-sums, so excluded minors for the class need not be 3-connected. An illustration of this is given by the following result, which provides a partial list of excluded minors for $\mathcal{P}_{2,q}$.

PROPOSITION 4.7. *For any finite field $\mathbb{F}_q$, the matroids F_7 , F_7^* , $M(K_5)$, $M^*(K_5)$, $M(K_{3,3})$, $M^*(K_{3,3})$, and $M(\mathcal{G})$, where $\mathcal{G}$ is any of the planar graphs in Figure 6, are excluded minors for $\mathcal{P}_{2,q}$. If $q \geq 4$, then $U_{3,6}$ is also an excluded minor for $\mathcal{P}_{2,q}$.*

We omit the proof, as it is only a matter of verifying that for each matroid M listed in the proposition, $M \notin \mathcal{P}_{2,q}$, but $M \setminus e, M/e \in \mathcal{P}_{2,q}$ for any $e \in E(M)$. We point out that the cycle matroids of all but the three leftmost graphs in Figure 6 are not 3-connected.

Acknowledgments. The author would like to thank Jim Geelen for contributing some of his ideas to this paper, and Alexander Vardy for pointers to the prior literature on trellis complexity.

REFERENCES

- [1] S. ARNBORG, D. G. CORNEIL, AND A. PROSKUROWSKI, *Complexity of finding embeddings in a k -tree*, SIAM J. Algebraic Discrete Methods, 8 (1987), pp. 277–284.
- [2] H. L. BODLAENDER, *A tourist guide through treewidth*, Acta Cybernet., 11 (1993), pp. 1–21.
- [3] G. CHARTRAND AND F. HARARY, *Planar permutation graphs*, Ann. Inst. H. Poincaré Sect. B, 3 (1967), pp. 433–438.
- [4] H. J. FLEISCHNER, D. P. GELLER, AND F. HARARY, *Outerplanar graphs and weak duals*, J. Indian Math. Soc., 38 (1975), pp. 215–219.
- [5] G. D. FORNEY, JR., *Dimension/length profiles and trellis complexity of linear block codes*, IEEE Trans. Inform. Theory, 40 (1994), pp. 1741–1752.
- [6] J. GEELEN, B. GERARDS, AND G. WHITTLE, *Towards a matroid-minor structure theory*, in Combinatorics, Complexity and Chance (A tribute to Dominic Welsh), G. Grimmett and C. McDiarmid, eds., Oxford University Press, 2007, pp. 72–82.
- [7] J. GEELEN, B. GERARDS, AND G. WHITTLE, *On Rota’s conjecture and excluded minors containing large projective geometries*, J. Combin. Theory Ser. B, 96 (2006), pp. 405–425.
- [8] J. GEELEN AND G. WHITTLE, *Branch-width and Rota’s conjecture*, J. Combin. Theory Ser. B, 86 (2002), pp. 315–330.
- [9] R. HALL, J. OXLEY, AND C. SEMPLE, *The structure of 3-connected matroids of path width three*, European J. Combin., 28 (2007), pp. 964–989.
- [10] G. B. HORN AND F. R. KSCHISCHANG, *On the intractability of permuting a block code to minimize trellis complexity*, IEEE Trans. Inform. Theory, 42 (1996), pp. 2042–2048.
- [11] K. JAIN, I. MĂNDOIU, AND V. V. VAZIRANI, *The “art of trellis decoding” is computationally hard—for large fields*, IEEE Trans. Inform. Theory, 44 (1998), pp. 1211–1214.
- [12] F. J. MACWILLIAMS AND N. J. A. SLOANE, *The Theory of Error-Correcting Codes*, North-Holland, Amsterdam, 1977.
- [13] D. J. MUDER, *Minimal trellises for block codes*, IEEE Trans. Inform. Theory, 34 (1988), pp. 1049–1053.
- [14] J. G. OXLEY, *Matroid Theory*, Oxford University Press, Oxford, UK, 1992.
- [15] N. ROBERTSON AND P. D. SEYMOUR, *Graph minors. I. Excluding a forest*, J. Combin. Theory Ser. B, 35 (1983), pp. 39–61.
- [16] A. VARDY, *Trellis structure of codes*, in Handbook of Coding Theory, R. Brualdi, C. Huffman, and V. Pless, eds., North-Holland, Amsterdam, The Netherlands, 1998, pp. 1989–2118.

ON THE DUALS OF BINARY HYPER-KLOOSTERMAN CODES*

MARKO MOISIO†

Abstract. Binary hyper-Kloosterman codes $C(r, m)$ of length $(2^r - 1)^{m-1}$ are a quasi-cyclic generalization of the dual of the Melas code of length $2^r - 1$. In this paper the duals $C^\perp(r, m)$, i.e., a generalization of the Melas code $C^\perp(r, 2)$ itself, are studied. In particular, the minimum distance of $C^\perp(r, m)$ for all $r, m \geq 2$, the weight distribution of $C(2, m)$ and $C^\perp(2, m)$ for all $m \geq 2$, and the weight distribution of $C(r, 3)$ and $C^\perp(r, 3)$ for all $r \geq 2$ are obtained.

Key words. exponential sum, Fermat curve, hyper-Kloosterman code, Kloosterman sum, Melas code, Pless power moments, weight distribution

AMS subject classifications. 11T23, 11T71

DOI. 10.1137/060668900

1. Introduction. Let $r, m \geq 2$ be integers and let $q = 2^r$. Let $\mathbb{F} := \mathbb{F}_q$ denote the finite field of q elements and let $\mathbb{F}^* := \mathbb{F} \setminus \{0\}$. For $\mathbf{a} := (a_1, \dots, a_m) \in \mathbb{F}^m$ we define a rational function in $m - 1$ variables:

$$f_{\mathbf{a}}(\mathbf{X}) := a_1 X_1 + \dots + a_{m-1} X_{m-1} + \frac{a_m}{X_1 \cdots X_{m-1}}.$$

Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be a fixed ordering of the elements of $(\mathbb{F}^*)^{m-1}$.

In [4] the following linear codes $C(r, m)$ were introduced and were called a *hyper-Kloosterman code*:

$$C(r, m) = \left\{ c(\mathbf{a}) := (\text{tr}(f_{\mathbf{a}}(\mathbf{x}_1)), \dots, \text{tr}(f_{\mathbf{a}}(\mathbf{x}_n))) \mid \mathbf{a} \in \mathbb{F}^m \right\};$$

here tr is the trace function from $\mathbb{F}$ onto $\mathbb{F}_2$. These codes are a quasi-cyclic generalization of the Kloosterman code, i.e., the dual of the Melas code, of length $2^r - 1$. For the proof of the quasi cyclicity we refer to [3, Theorem 4.2].

In this paper we are interested in the duals $C^\perp(r, m)$, which are a generalization of the Melas code $C^\perp(r, 2)$ ($r > 2$). We shall show that the minimum distance of $C^\perp(r, m)$ is three if $m > 2$, give the weight distribution of $C(2, m)$ and $C^\perp(2, m)$ for all $m \geq 2$, and give the weight distribution of $C(r, 3)$ and $C^\perp(r, 3)$ for all $r \geq 2$. We remark that the weight distributions of $C(r, 2)$ and $C^\perp(r, 2)$ ($r > 2$) were obtained in [5] and in [14], respectively.

The rest of this paper is organized as follows. In section 2 we first consider some simple basic properties of hyper-Kloosterman codes. Next, the weight distribution of $C(r, m)$ is given in terms of certain monomial exponential sums (Theorem 2), and then a recursion formula for the weight distribution of $C^\perp(r, m)$ involving the moments M_j of those exponential sums is obtained by using the Pless power moment identity (Theorem 4).

In section 3 we first connect the moments M_j to a Fermat curve $\mathcal{X}$ and then obtain the number of weight three codewords in $C^\perp(r, m)$ in terms of the number of

*Received by the editors September 2, 2006; accepted for publication (in revised form) July 29, 2007; published electronically February 27, 2008.

<http://www.siam.org/journals/sidma/22-1/66890.html>

†Department of Mathematics and Statistics, Faculty of Technology, University of Vaasa, P.O. Box 700, FIN-65101 Vaasa, Finland (mamo@uwasa.fi).

rational points on $\mathcal{X}$ (Theorem 5). Finally, we determine the minimum distance of $C^\perp(r, m)$ either by using our explicit knowledge of the number of rational points on $\mathcal{X}$ or by estimating that number by either the Hasse–Weil bound or a bound which we shall derive by using Deligne’s bound on hyper-Kloosterman sums (Theorem 6). In a few cases we are forced to calculate the number of rational points numerically since neither of those bounds is then strong enough.

In sections 4 and 5 we determine the weight distribution of the codes $C(r, m)$ and $C^\perp(r, m)$ in the special cases $r = 2$, $m > 2$ ($r > 2$, $m = 3$, respectively). In the latter case a relation between one- and two-dimensional Kloosterman sums from [1] is used, and then the weight distribution of $C(r, 3)$ is obtained by using results on the distribution of values of Kloosterman sums obtained in [5] (Theorem 11). Finally, the weight distribution of $C^\perp(r, 3)$ is obtained in terms of even moments of Kloosterman sums calculated in [10] by using results from [14]. In particular, explicit formulae for the number of codewords of weights from 3 to 5 is given (Theorem 13).

2. On the weight distribution of $C(r, m)$ and $C^\perp(r, m)$. Let χ be the canonical additive character of $\mathbb{F}$. Let

$$k_{m-1}(\mathbf{a}) = \sum_{x_1, \dots, x_{m-1} \in \mathbb{F}^*} \chi(a_1 x_1 + \dots + a_{m-1} x_{m-1} + a_m (x_1 \cdots x_{m-1})^{-1})$$

be an $(m-1)$ -dimensional Kloosterman sum. If $\mathbf{a} = (1, 1, \dots, 1, a)$ with $a \neq 0$, we use the notation

$$k_{m-1}(a) := k_{m-1}(\mathbf{a}).$$

Let v be the number of zero-components of $\mathbf{a}$. Assume $v > 0$. If $a_m = 0$, then by the orthogonality of characters we get

$$k_{m-1}(\mathbf{a}) = \sum_{x_1, \dots, x_{m-1} \in \mathbb{F}^*} \chi(a_1 x_1 + \dots + a_{m-1} x_{m-1}) = (-1)^{m-v} (q-1)^{v-1}.$$

If $a_m \neq 0$ and, e.g., $a_1 = 0$, then by the substitution $y = x_1^{-1}$ we obtain

$$\begin{aligned} k_{m-1}(\mathbf{a}) &= \sum_{x_2, \dots, x_{m-1} \in \mathbb{F}^*} \chi(a_2 x_2 + \dots + a_{m-1} x_{m-1}) \sum_{y \in \mathbb{F}^*} \chi\left(\frac{a_m}{x_2 \cdots x_{m-1}} y\right) \\ &= - \sum_{x_2, \dots, x_{m-1} \in \mathbb{F}^*} \chi(a_2 x_2 + \dots + a_{m-1} x_{m-1}) \\ &= -(-1)^{m-2-(v-1)} (q-1)^{v-1}. \end{aligned}$$

Hence we have the following lemma.

LEMMA 1. *If exactly $v > 0$ of the components of $\mathbf{a} \in \mathbb{F}^m$ are zeros, then*

$$k_{m-1}(\mathbf{a}) = (-1)^{m-v} (q-1)^{v-1}.$$

If $v = 0$, then we have the following well-known bound by Deligne:

$$|k_{m-1}(\mathbf{a})| \leq m q^{\frac{m-1}{2}}.$$

LEMMA 2. *The dimension k of $C(r, m)$ over $\mathbb{F}_2$ is rm if $rm > 4$. If $r = m = 2$, then $k = 2$.*

Proof. Consider the group homomorphism

$$(1) \quad \Psi : (\mathbb{F}^m, +) \longrightarrow C(r, m), \mathbf{a} \mapsto (\mathrm{tr}(f_{\mathbf{a}}(\mathbf{x}_1)), \dots, \mathrm{tr}(f_{\mathbf{a}}(\mathbf{x}_n))).$$

If $\mathbf{a}$ belongs to $\mathrm{Ker}(\Psi)$, then $k_{m-1}(\mathbf{a}) = (q-1)^{m-1}$. If $rm > 4$, this can happen if and only if $\mathbf{a} = \mathbf{0}$, by Deligne's bound and by Lemma 1, and therefore ψ is an isomorphism.

If $r = m = 2$ and $a, b \in \mathbb{F}_4^*$, then $k_{m-1}((a, b)) = k_{m-1}(ab)$. If $ab = 1$, then $k_{m-1}(ab) = 3$; otherwise $k_{m-1}(ab) = -1$. Hence, in this case $|\mathrm{Ker}(\Psi)| = 4$, and consequently $|C(r, m)| = 16/4 = 4$. $\square$

Remark 1. A different proof for this result is given in [3, Theorem 3.1].

The Hamming weight $w(c(\mathbf{a}))$ of codeword $c(\mathbf{a})$ is given by

$$(2) \quad w(c(\mathbf{a})) = \sum_{\mathbf{x} \in (\mathbb{F}^*)^{m-1}} \frac{1}{2} \left(1 - (-1)^{\mathrm{tr}(f_{\mathbf{a}}(\mathbf{x}))} \right) = \frac{1}{2} ((q-1)^{m-1} - k_{m-1}(\mathbf{a})).$$

Next we express $w(c(\mathbf{a}))$ by means of a monomial exponential sum over $\mathbb{F}_{q^m}$. Let e denote the canonical additive character of $\mathbb{F}_{q^m}$. Let $t = (q^m - 1)/(q - 1)$ and let $N(\alpha) := \alpha^t$ denote the norm of α from $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$. Let γ be a primitive element of $\mathbb{F}_{q^m}$ and let

$$s(\alpha) = \sum_{i=0}^{t-1} e(\alpha \gamma^{(q-1)i}).$$

We have the following result from [9, Theorem 3].

THEOREM 1. *Let $\alpha \in \mathbb{F}_{q^m}^*$ and let $a = N(\alpha)$. Then*

$$\sum_{x \in \mathbb{F}_{q^m}^*} e(\alpha x^{q-1}) = (-1)^{m-1} (q-1) k_{m-1}(a),$$

or, equivalently,

$$k_{m-1}(a) = (-1)^{m-1} s(\alpha).$$

LEMMA 3. *Let $\mathbf{a} \in (\mathbb{F}^*)^m$ and let $b = a_1 \cdots a_m$. Let $g := N(\gamma)$ be a primitive element of $\mathbb{F}$, $i = \mathrm{ind}_g(b)$, and $\beta = \gamma^i$. Then*

$$\begin{aligned} w(c(\mathbf{a})) &= \frac{1}{2} ((q-1)^{m-1} - k_{m-1}(b)) \\ &= \frac{1}{2} ((q-1)^{m-1} + (-1)^m s(\beta)). \end{aligned}$$

Proof. The first equality follows easily by (2), and the second one then follows by Theorem 1. $\square$

Let S denote the range of $s(\gamma^i)$ as i varies over the set $I := \{0, \dots, q-2\}$, and, for $j \in S$, let N_j denote the number of elements i in I such that $s(\gamma^i) = j$, i.e.,

$$S = \{s(\gamma^i) \mid i \in I\}$$

and

$$N_j = |\{i \in I \mid s(\gamma^i) = j\}|.$$

THEOREM 2. Assume $rm > 4$. For $\mathbf{a} \in \mathbb{F}^m$ let v be the number of zero-components of $\mathbf{a}$. If $v > 0$, there are

$$\binom{m}{v} (q-1)^{m-v} \text{ codewords } c(\mathbf{a}) \text{ of weight } ((q-1)^{m-1} - (-1)^{m-v}(q-1)^{v-1})/2,$$

and otherwise, for each $j \in S$, there are

$$N_j (q-1)^{m-1} \text{ codewords } c(\mathbf{a}) \text{ of weight } ((q-1)^{m-1} + (-1)^m j)/2$$

in $C(r, m)$. Moreover, these are the only weights in $C(r, m)$.

Proof. First, for each $\mathbf{a} \in \mathbb{F}^m$, there exists exactly one codeword $c(\mathbf{a}) \in C(r, m)$, by isomorphism (1). If $v > 0$, the claim follows now by Lemmas 1 and 3.

Assume $v = 0$. For each $b \in \mathbb{F}^*$ there are exactly $(q-1)^{m-1}$ vectors $\mathbf{a} \in (\mathbb{F}^*)^m$ such that the product of the components of $\mathbf{a}$ equals b . The second claim follows now by Lemma 3 since there is exactly one i in I such that $N(\gamma^i) = b$. The last claim is now obvious. $\square$

COROLLARY 1. The weights in $C(r, m)$ are divisible by $2^{\ell-1}$, where $\ell = \min\{r, m\}$.

Proof. Let $\alpha \in \mathbb{F}_{q^m}$. By [12, Theorem 2], the exponential sum

$$\sum_{x \in \mathbb{F}_{q^m}} e(\alpha x^{q-1})$$

is divisible by $2^{\lceil rm/s \rceil}$, where s is the binary weight of $q-1$. Now $s = r$, and therefore $\sum_{x \in \mathbb{F}_{q^m}} e(\alpha x^{q-1}) = 2^m z$ for some $z \in \mathbb{Z}$.

Let $\mathbf{a} \in (\mathbb{F}^*)^m$, and let $\beta \in \mathbb{F}_{q^m}^*$ such that $N(\beta) = a_1 \cdots a_m$. Now

$$\begin{aligned} (q-1)w(c(\mathbf{a})) &= \frac{1}{2}((q-1)^m + (-1)^m (q-1)s(\beta)) \\ &= \frac{1}{2} \left((q-1)^m + (-1)^m \sum_{x \in \mathbb{F}_{q^m}^*} e(\alpha x^{q-1}) \right) \\ &= \frac{1}{2} \left((q-1)^m + (-1)^m \sum_{x \in \mathbb{F}_{q^m}} e(\alpha x^{q-1}) - (-1)^m \right) \\ &= \frac{1}{2} (q^m - mq^{m-1} + \cdots + (-1)^{m-1} mq + (-1)^m 2^m z), \end{aligned}$$

and, as $q = 2^r$, the claim follows in this case. If some of the components of $\mathbf{a}$ are zero, then it is easily seen that 2^{r-1} is a factor of $w(c(\mathbf{a}))$. $\square$

Remark 2. A different proof for this result is given in [4, Corollary 4.3].

To obtain the weight distribution of $C^\perp(r, m)$ we use the Pless power moment identity proved in [13] (see also, e.g., [6, p. 131]).

THEOREM 3 (power moment identity). Let B be a binary linear $[n, k]$ code, and let B_i (resp., $B_i^\perp$) denote the number of codewords of weight i in B (resp., in $B^\perp$). Then, for $h = 0, 1, \dots$, we have

$$\sum_{i=0}^n i^h B_i = \sum_{i=0}^n (-1)^i B_i^\perp \sum_{t=0}^h t! S(h, t) 2^{k-t} \binom{n-i}{n-t},$$

where

$$S(h, t) := \frac{1}{t!} \sum_{j=0}^t (-1)^{t-j} \binom{t}{j} j^h \quad (\text{a Stirling number of the second kind}),$$

and the binomial coefficient $\binom{u}{v}$ is defined to be zero whenever $v > u$ or $v < 0$.

For a nonnegative integer j we denote by M_j the j th moment of the period $s(\gamma^l)$, or

$$M_j := \sum_{l=0}^{q-2} s(\gamma^l)^j.$$

THEOREM 4. Assume $rm > 4$, and let $w = 1 - q$. The number $C_h^\perp$ of codewords of weight h in $C^\perp(r, m)$ is given by

$$\begin{aligned} q^m h! C_h^\perp &= f(C_0^\perp, \dots, C_{h-1}^\perp) + g(M_0, \dots, M_h) \\ &\quad + (-1)^{(m+1)(h+1)} w^{-h} \left(w^m (1 - w^m)^h - \sum_{j=0}^h \binom{h}{j} (-1)^j (w^{j+1} - w^h)^m \right), \end{aligned}$$

where

$$\begin{aligned} f(C_0^\perp, \dots, C_{h-1}^\perp) &= q^m \sum_{i=0}^{h-1} (-1)^{h+i+1} C_i^\perp \sum_{t=i}^h t! S(h, t) 2^{h-t} \binom{n-i}{n-t}, \\ g(M_0, \dots, M_h) &= \sum_{j=0}^h \binom{h}{j} (-1)^{mj+h} (q-1)^{(m-1)(h-j+1)} M_j. \end{aligned}$$

Moreover, if $m = 3$, the formula simplifies to

$$\begin{aligned} q^3 h! C_h^\perp &= f(C_0^\perp, \dots, C_{h-1}^\perp) + g(M_0, \dots, M_h) \\ &\quad + 3(q-1)^2 (-q)^h ((q-2)^h + (q-1)^{h-1}). \end{aligned}$$

Proof. We choose $B = C(r, m)$ in the power moment identity. Then, by Theorem 2,

$$\begin{aligned} \sum_{i=0}^n i^h C_i &= \sum_{v=1}^m \binom{m}{v} (q-1)^{m-v} 2^{-h} ((q-1)^{m-1} - (-1)^{m-v} (q-1)^{v-1})^h \\ &\quad + \sum_{l=0}^{q-2} 2^{-h} (q-1)^{m-1} ((q-1)^{m-1} + (-1)^m s(\gamma^l))^h =: S_1 + S_2, \end{aligned}$$

where

$$2^h S_1 = \sum_{v=1}^m \binom{m}{v} (q-1)^{m-v} ((q-1)^{m-1} - (-1)^{m-v} (q-1)^{v-1})^h$$

and

$$2^h S_2 = (q-1)^{m-1} \sum_{l=0}^{q-2} ((q-1)^{m-1} + (-1)^m s(\gamma^l))^h.$$

First, we manipulate S_2 somewhat:

$$\begin{aligned} 2^h S_2 &= (q-1)^{m-1} \sum_{l=0}^{q-2} \sum_{j=0}^h \binom{h}{j} (-1)^{mj} s(\gamma^l)^j (q-1)^{(m-1)(h-j)} \\ &= \sum_{j=0}^h \binom{h}{j} (-1)^{mj} (q-1)^{(m-1)(h-j+1)} \sum_{l=0}^{q-2} s(\gamma^l)^j \\ &= \sum_{j=0}^h \binom{h}{j} (-1)^{mj} (q-1)^{(m-1)(h-j+1)} M_j. \end{aligned}$$

Second, we consider S_1 . If $m = 3$, then

$$2^h S_1 = 3(q-1)^2 q^h ((q-2)^h + (q-1)^{h-1}).$$

Next we write S_1 in the form from which we can derive explicit formulae for the number of low-weight codewords in the duals $C^\perp(r, m)$ for an arbitrary integer $m \geq 2$:

$$\begin{aligned} 2^h S_1 &= (q-1)^{m-1} \sum_{v=1}^m \binom{m}{v} (q-1)^{(v-1)(h-1)} ((q-1)^{m-v} + (-1)^{m-v-1})^h \\ &= (q-1)^{m-1} \sum_{v=1}^m \binom{m}{v} (q-1)^{(v-1)(h-1)} (-1)^{(m-v-1)h} (1 - (1-q)^{m-v})^h \\ &= (-1)^{(m-1)h} (-w)^{m-1} \sum_{j=0}^h \binom{h}{j} (-1)^j \sum_{v=1}^m \binom{m}{v} (-1)^{vh} (-w)^{(v-1)(h-1)} w^{(m-v)j} \\ &= (-1)^{(m-1)h} (-w)^{-h} \sum_{j=0}^h \binom{h}{j} (-1)^j \sum_{v=1}^m \binom{m}{v} w^{vh} (-w)^{m-v} w^{(m-v)j} \\ &= (-1)^{(m-1)h+m} (-w)^{-h} \sum_{j=0}^h \binom{h}{j} (-1)^j \sum_{v=1}^m \binom{m}{v} (-1)^v w^{vh} w^{(m-v)(j+1)} \\ &= (-1)^{m(h+1)} w^{-h} \sum_{j=0}^h \binom{h}{j} (-1)^j \left((w^{j+1} - w^h)^m - w^{m(j+1)} \right). \end{aligned}$$

Since

$$\begin{aligned} \sum_{j=0}^h \binom{h}{j} (-1)^j w^{m(j+1)} &= (1-q)^m \sum_{j=0}^h \binom{h}{j} (-1)^j (1-q)^{mj} \\ &= w^m (1 - w^m)^h, \end{aligned}$$

we have

$$\begin{aligned} 2^h S_1 &= (-1)^{m(h+1)} w^{-h} \sum_{j=0}^h \binom{h}{j} (-1)^j (w^{j+1} - w^h)^m \\ &\quad - (-1)^{m(h+1)} w^{m-h} (1 - w^m)^h. \end{aligned}$$

As the left-hand side of the power moment identity equals $S_1 + S_2$, and the right-hand side equals

$$\begin{aligned} &\sum_{i=0}^n (-1)^i C_i^\perp \sum_{t=0}^h t! S(h, t) 2^{rm-t} \binom{n-i}{n-t} \\ &= \sum_{i=0}^h (-1)^i C_i^\perp \sum_{t=i}^h t! S(h, t) 2^{rm-t} \binom{n-i}{n-t} \\ &= \frac{q^m}{2^h} \sum_{i=0}^h (-1)^i C_i^\perp \sum_{t=i}^h t! S(h, t) 2^{h-t} \binom{n-i}{n-t}, \end{aligned}$$

the claims now easily follows. $\square$

3. The minimum distance of $C^\perp(r, m)$. To determine the minimum distance of $C^\perp(r, m)$ we need some auxiliary results. We recall that $t = (q^m - 1)/(q - 1)$ and $\mathbb{F}_{q^m}^* = \langle \gamma \rangle$.

LEMMA 4. *The first four moments M_j in Theorem 4 are given by*

$$\begin{aligned} M_0 &= q - 1, \quad M_1 = -1, \quad M_2 = q^m - t, \\ M_3 &= \frac{|\mathcal{X}(\mathbb{F}_{q^m})| - 3(q - 1)}{(q - 1)^2} q^m - t^2, \end{aligned}$$

where $|\mathcal{X}(\mathbb{F}_{q^m})|$ is the number of rational points on the projective curve $\mathcal{X}$ over $\mathbb{F}_{q^m}$ defined by the equation

$$\mathcal{X} : x^{q-1} + y^{q-1} + z^{q-1} = 0.$$

Proof. Obviously $M_0 = q - 1$, and

$$M_1 = \sum_{l=0}^{q-2} s(\gamma^l) = \frac{q-1}{q^m-1} \sum_{l=0}^{q^m-2} s(\gamma^l) = \frac{1}{t} \sum_{i=0}^{t-1} \sum_{l=0}^{q^m-2} e(\gamma^{(q-1)i} \gamma^l) = -\frac{t}{t},$$

where the last equality follows by the orthogonality of characters. To prove the formula for M_2 we count the number N of solutions of the equation $x + y = 0$ in the group H of $(q-1)$ th powers in $\mathbb{F}_{q^m}^*$. On the one hand, $N = t$, and on the other hand, by the orthogonality of characters,

$$\begin{aligned} q^m t &= \sum_{x, y \in H} \sum_{u \in \mathbb{F}_{q^m}^*} e(u(x + y)) = t^2 + \sum_{u \in \mathbb{F}_{q^m}^*} \left(\sum_{x \in H} e(ux) \right)^2 = t^2 + t \sum_{l=0}^{q-2} s(\gamma^l)^2 \\ &= t^2 + t M_2, \end{aligned}$$

from which the formula for M_2 follows.

Let N denote the number of solutions of equation

$$(3) \quad x^{q-1} + y^{q-1} + z^{q-1} = 0$$

in $\mathbb{F}_{q^m}^3$. It is easy to see [9, section 3] that $N = N'_m + N_m$, where

$$N'_m = 3(q-1)(q^m-1) + 1$$

and

$$\begin{aligned} q^m N_m &= \sum_{u \in \mathbb{F}_{q^m}^*} \left(\sum_{x \in \mathbb{F}_{q^m}^*} e(ux^{q-1}) \right)^3 + (q^m - 1)^3 \\ &= (q-1)^3 \sum_{u \in \mathbb{F}_{q^m}^*} s(u)^3 + (q^m - 1)^3 \\ &= (q-1)^3 tM_3 + (q^m - 1)^3, \end{aligned}$$

and consequently,

$$q^m N = 3(q-1)(q^m-1)q^m + q^m + (q-1)^3 tM_3 + (q^m - 1)^3.$$

Since $|\mathcal{X}(\mathbb{F}_{q^m})| = (N-1)/(q^m-1)$, we obtain

$$(q-1)^3 \frac{q^m-1}{q-1} M_3 = q^m(q^m-1)|\mathcal{X}(\mathbb{F}_{q^m})| - 3(q-1)(q^m-1)q^m - (q^m-1)^3,$$

which simplifies to

$$(q-1)^2 M_3 = q^m |\mathcal{X}(\mathbb{F}_{q^m})| - 3(q-1)q^m - (q^m-1)^2.$$

Since $(q^m-1)^2 = (q-1)^2 t^2$, we see that the claim is true also for M_3 . $\square$

THEOREM 5. *The minimum distance of $C^\perp(r, m)$ is at least three. Moreover, if $rm > 4$, the number $C_3^\perp$ of weight three codewords in $C^\perp(r, m)$ is given by*

$$C_3^\perp = \frac{(q-1)^{m-3}((q-2)^m + (-1)^m(q^m + 3q - |\mathcal{X}(\mathbb{F}_{q^m})| - 5))}{6}.$$

Proof. Let $n = (q-1)^{m-1}$, and let $\mathbf{c} \in C^\perp(r, m)$. If $w(\mathbf{c}) = 2$, then

$$\text{tr}(f_{\mathbf{a}}(\mathbf{x}_i) + f_{\mathbf{a}}(\mathbf{x}_j)) = \text{tr}(f_{\mathbf{a}}(\mathbf{x}_i)) + \text{tr}(f_{\mathbf{a}}(\mathbf{x}_j)) = 0$$

for some $1 \leq i < j \leq n$, say $i = 1, j = 2$, and for all $\mathbf{a} \in \mathbb{F}_q^m$. Let $1 \leq l \leq m-1$ be the index of the coordinate place where $\mathbf{x}_1$ and $\mathbf{x}_2$ differ, say $l = 1$. By choosing $\mathbf{a} = (a, 0, \dots, 0)$, we have $\text{tr}(a(x+y)) = 0$ for all $a \in \mathbb{F}_q$ and for some $x, y \in \mathbb{F}_q^*$ with $x \neq y$. (Here x and y are the first components of $\mathbf{x}_1$ and $\mathbf{x}_2$.) This contradicts the surjectivity of tr . A similar argument also proves that $w(\mathbf{c}) \neq 1$.

Next we use Theorem 4 and Lemma 4 to prove the claimed formula for $C_3^\perp$. First,

$$\begin{aligned} f(C_0^\perp, C_1^\perp, C_2^\perp) &= f(1, 0, 0) = q^m \sum_{t=0}^3 t! S(3, t) 2^{3-t} \binom{n}{n-t} \\ &= q^m (4n + 6n(n-1) + (n-2)(n-1)n) \\ &= (q-1)^{2m-2} ((q-1)^{m-1} + 3) q^m, \end{aligned}$$

and second,

$$\begin{aligned}
g(M_0, M_1, M_2, M_3) &= \sum_{j=0}^3 \binom{3}{j} (-1)^{mj+h} (q-1)^{(m-1)(h-j+1)} M_j \\
&= -(q-1)^{4m-3} + 3(-1)^m (q-1)^{3(m-1)} - 3(q-1)^{2(m-1)} \left(q^m - \frac{q^m - 1}{q-1} \right) \\
&\quad - (-1)^m (q-1)^{m-1} \left(\frac{|\mathcal{X}(\mathbb{F}_{q^m})| - 3(q-1)}{(q-1)^2} q^m - \frac{(q^m - 1)^2}{(q-1)^2} \right),
\end{aligned}$$

or, equivalently,

$$\begin{aligned}
(q-1)^{3-m} g(M_0, M_1, M_2, M_3) &= (-1)^m - 3(q-1)^m + 3(-1)^m (q-1)^{2m} \\
&\quad - (q-1)^{3m} + ((-1)^m (q^m + 3q - |\mathcal{X}(\mathbb{F}_{q^m})| - 5) - 3(q-2)(q-1)^m) q^m.
\end{aligned}$$

Finally, since

$$\begin{aligned}
&(q-1)^{3-m} w^{-3} \left(w^m (1-w^m)^3 - \sum_{j=0}^3 \binom{3}{j} (w^{j+1} - w^3)^m \right) \\
&= -(-1)^m + 3(q-1)^m - 3(-1)^m (q-1)^{2m} + (q-1)^{3m} \\
&\quad - ((q-1)^m + 3)(q-1)^m q^m + (q-2)^m q^m,
\end{aligned}$$

we obtain

$$6C_3^\perp = (q-1)^{m-3} ((q-2)^m + (-1)^m (q^m + 3q - |\mathcal{X}(\mathbb{F}_{q^m})| - 5))$$

by Theorem 4. $\square$

Example 1. If $r = m = 2$, then $n = 3$, and therefore the minimum distance of $C^\perp(2, 2)$ is three. Hence, $C^\perp(2, 2)$ is a repetition code.

Example 2. Consider the Melas code $C^\perp(r, 2)$. By [9, Theorem 1],

$$|\mathcal{X}(\mathbb{F}_{q^2})| = (1 - (-1)^r)(q-1)^2 + 3(q-1),$$

and consequently,

$$C_3^\perp = (1 + (-1)^r)(q-1)/6,$$

which is in accordance with [14, Table 6.1].

Example 3. Consider code $C^\perp(r, 3)$. By [9, Theorem 2],

$$|\mathcal{X}(\mathbb{F}_{q^3})| = (2q + 1 - (-1)^r)(q-1)^2 + 3(q-1),$$

and therefore,

$$C_3^\perp = (2q - 5 - (-1)^r)(q-1)^2/6.$$

Remark 3. By generalizing the argument used in the proof of Theorem 5 to prove the nonexistence of weight two codewords, it is easy to see that a check matrix for

$C^\perp(r, m)$ is $(\mathbf{y}_1^T \mathbf{y}_2^T \dots \mathbf{y}_n^T)$, where $\mathbf{y}_i = (\mathbf{x}_i \ z_i)$ and z_i is the product of the inverses of the components of $\mathbf{x}_i$.

We shall see soon that the minimum distance of $C^\perp(r, m)$ is always three if $m > 2$. It will turn out that Theorem 5, together with the Hasse–Weil bound, proves most of the cases. On the other hand, in case $m = 4$ it is too weak, and we shall use the following upper bound.

LEMMA 5.

$$|\mathcal{X}(\mathbb{F}_{q^m})| < q^m + 3q + (q-1)^3 m^3 q^{\frac{m-3}{2}} - 4.$$

Proof. As we pointed out in the proof of Lemma 4, the number of solutions N of (3) satisfies

$$q^m N = (q-1)^3 \left(\sum_{u \in \mathbb{F}_{q^m}^*} e(ux^{q-1}) \right) + (q^m - 1)^3 + 3(q-1)(q^m - 1)q^m + q^m,$$

and then it is easily seen (see [9, section 3]) that

$$q^m N = (-1)^{m-1} t (q-1)^3 \sum_{u \in \mathbb{F}_q^*} k_{m-1}(u)^3 + (q^m - 1)^3 + 3(q-1)(q^m - 1)q^m + q^m.$$

Since $|\mathcal{X}(\mathbb{F}_{q^m})| = (N-1)/(q^m - 1)$, we obtain

$$(4) \quad q^m |\mathcal{X}(\mathbb{F}_{q^m})| = (-1)^{m-1} (q-1)^2 \sum_{u \in \mathbb{F}_q^*} k_{m-1}(u)^3 + (q^m - 1)^2 + 3(q-1)q^m.$$

Now Deligne's bound gives the inequality

$$(5) \quad \left| \sum_{u \in \mathbb{F}_q^*} k_{m-1}(u)^3 \right| \leq (q-1) m^3 q^{\frac{3(m-1)}{2}},$$

and therefore,

$$|\mathcal{X}(\mathbb{F}_{q^m})| \leq (q-1)^3 m^3 q^{\frac{m-3}{2}} + q^m - 2 + 3(q-1) + q^{-m}. \quad \square$$

THEOREM 6. *The minimum distance of $C^\perp(r, m)$ is three unless r is odd and $m = 2$, in which case it is at least five.*

Proof. Assume $m = 2$. If $r = 2$, the minimum distance $d = 3$ by Example 1, and if $r > 2$, then it is well known that $d = 3$ or $d \geq 5$ according to whether r is even or odd (see, e.g., [14]).

CLAIM. *If $m > 2$, then $d = 3$.*

If $m = 3$, the claim is true by Example 3. Assume $m > 3$. To prove the claim it is enough, by Theorem 5, to show that

$$\epsilon := (q-2)^m + (-1)^m (q^m + 3q - |\mathcal{X}(\mathbb{F}_{q^m})| - 5)$$

is positive. By separating the cases according to the parity of m , and by using the Hasse–Weil bounds

$$q^m + 1 - (q-2)(q-3)q^{\frac{m}{2}} \leq |\mathcal{X}(\mathbb{F}_{q^m})| \leq q^m + 1 + (q-2)(q-3)q^{\frac{m}{2}},$$

TABLE 3.1

r	ϵ
3	$2^3 \cdot 3 \cdot 7^2$
4	$2 \cdot 3^3 \cdot 5^2 \cdot 37$
5	$2^4 \cdot 3^2 \cdot 5 \cdot 31^2$
6	$2 \cdot 3^5 \cdot 7^2 \cdot 641$
7	$2^3 \cdot 3^2 \cdot 7 \cdot 31 \cdot 127^2$
8	$2 \cdot 3^3 \cdot 5^2 \cdot 7 \cdot 17^2 \cdot 1531$
9	$2^8 \cdot 3 \cdot 5 \cdot 7^2 \cdot 67 \cdot 73^2$
10	$2 \cdot 3^3 \cdot 11^5 \cdot 31^2 \cdot 131$
11	$2^3 \cdot 3^3 \cdot 11 \cdot 23^2 \cdot 89^2 \cdot 1759$
12	$2 \cdot 3^5 \cdot 5^2 \cdot 7^2 \cdot 13^2 \cdot 113 \cdot 24709$

we obtain

$$\epsilon > (q-2)^m - (q-2)(q-3)q^{\frac{m}{2}} - 3q - 6,$$

which is obviously positive if $m \geq 5$ and $r \geq 3$ (i.e., $q \geq 8$).

Assume $m = 4$ and $\epsilon = 0$. Then, by Lemma 5, we must have

$$\begin{aligned} (q-2)^4 &< 64\sqrt{q}(q-1)^3 + 1 \\ \Leftrightarrow (q-1)^4 &< 64\sqrt{q}(q-1)^3 + 4(q-1)^3 - 6(q-1)^2 + 4(q-1) \\ \Leftrightarrow q-1 &< 64\sqrt{q} + 4 - \frac{6}{q-1} + \frac{4}{(q-1)^2} < 64\sqrt{q} + 5. \end{aligned}$$

The inequality $q-6 < 64\sqrt{q}$ implies that we must have $q \leq 2^{12}$, i.e., $r \leq 12$. Hence, if $m = 4$ and $r > 12$, the minimum distance is three. In the cases $m = 4$, $3 \leq r \leq 12$, we have verified this by calculating $|\mathcal{X}(\mathbb{F}_{q^4})|$ numerically (see Table 3.1).

In the remaining cases $r = 2$, $m \geq 4$, the claim follows by Theorem 9 below, by which $C_3^\perp = 3^{m-3}(2^{m-1} \pm 1)$. $\square$

We computed $|\mathcal{X}(\mathbb{F}_{q^4})|$ by using (4). In the calculation of the three-dimensional Kloosterman sums $k_3(a)$ over $\mathbb{F}_q$, $q = 2^r$ with $3 \leq r \leq 12$, we took advantage of the following result by Carlitz from [1] which related one- and two-dimensional Kloosterman sums.

THEOREM 7. *For any $a \in \mathbb{F}_q^*$, we have*

$$k_2(a) = k(a)^2 - q,$$

where $k(a) := k_1(a)$.

By Theorem 7 we have

$$\begin{aligned} k_3(a) &= \sum_{x,y,z \in \mathbb{F}_q^*} \chi(x+y+z+a(xyz)^{-1}) = \sum_{x \in \mathbb{F}_q^*} \chi(x)k_2(ax^{-1}) \\ &= \sum_{x \in \mathbb{F}_q^*} \chi(x)k(ax^{-1})^2 - q \sum_{x \in \mathbb{F}_q^*} \chi(x) \\ &= \sum_{x \in \mathbb{F}_q^*} \chi(x)k(ax^{-1})^2 + q, \end{aligned}$$

and now it is easy to see that

$$\begin{aligned} k_3(a) &= 2 \sum_{\substack{x \in \mathbb{F}_q^* \\ \text{tr}(x)=0}} k(ax^{-1})^2 - \sum_{x \in \mathbb{F}_q^*} k(ax^{-1})^2 + q \\ &= 2 \sum_{\substack{x \in \mathbb{F}_q^* \\ \text{tr}(x)=0}} k(ax^{-1})^2 - (q^2 - q - 1) + q. \end{aligned}$$

By tabulating the traces of elements of $\mathbb{F}_q^*$, the indices of those elements of $\mathbb{F}_q^*$ having the trace equal to zero, and then the range of $k(u)$ as u varies over $\mathbb{F}_q^*$, before using the formula above, the data of Table 3.1 can quickly be verified.

Remark 4. The traces were calculated by making use of [3, Theorem 5.1].

4. The weight distribution of $C(2, m)$ and $C^\perp(2, m)$. In this section we assume that $m > 2$. Let γ be a primitive element of $\mathbb{F}_{2^{2m}}$. To determine the weight distribution of $C(2, m)$ and $C^\perp(2, m)$ we need the following result, which has been proved already in [2] (see, e.g., [8] for a different proof).

LEMMA 6. *Let $\alpha \in \mathbb{F}_{2^{2m}}^*$. Then*

$$\sum_{x \in \mathbb{F}_{2^{2m}}^*} e(\alpha x^3) = \begin{cases} (-1)^m 2^m - 1 & \text{if } 3 \nmid \text{ind}_\gamma \alpha, \\ (-1)^{m+1} 2^{m+1} - 1 & \text{if } 3 \mid \text{ind}_\gamma \alpha. \end{cases}$$

Lemma 6, together with Theorem 2, gives the weight distribution of $C(2, m)$.

THEOREM 8. *The weight distribution of $C(2, m)$ is given in the following table, where v runs over the integers $1, \dots, m$:*

Weight	Frequency
$\frac{3^{m-1} - (-1)^{m-v} 3^{v-1}}{2}$	$\binom{m}{v} 3^{m-v}$
$\frac{1}{2} \left(3^{m-1} + \frac{2^m - (-1)^m}{3} \right)$	$2 \cdot 3^{m-1}$
$\frac{1}{2} \left(3^{m-1} - \frac{2^{m+1} + (-1)^m}{3} \right)$	3^{m-1}

THEOREM 9. *For every nonnegative integer h the number $C_h^\perp$ of codewords of weight h in the dual $C^\perp(2, m)$ of $C(2, m)$ is given by the recursion of Theorem 4 with*

$$M_j = 2 \left(\frac{(-2)^m - 1}{3} \right)^j + \left(\frac{(-2)^{m+1} - 1}{3} \right)^j \quad \forall j = 0, 1, \dots$$

In particular,

$$C_0^\perp = 1, \quad C_1^\perp = C_2^\perp = 0, \quad C_3^\perp = 3^{m-3} (2^{m-1} \pm 1),$$

$$C_4^\perp = 3^{m-5} \left(\frac{7^m - 3^{m+3} + 66}{8} + 3 \cdot 2^{2m-2} \pm 2^m \right),$$

$$C_5^\perp = 3^{m-6} \left((5^{m-1} \pm 6) 2^{2m-3} - 3^{m+2} 2^{m-2} + 2^{3m-2} + 7 \cdot 2^{m+1} \pm \frac{55 - 3^{m+1}}{2} \right),$$

where $\pm = (-1)^m$.

Proof. By Lemma 6 the moments M_j in Theorem 4 are of the claimed form; the claimed formulae for the low-weight codewords can be verified, e.g., by using Mathematica. $\square$

Remark 5. In a similar manner as was done above, the weight distribution of the codes $C(r, m)$ and $C^\perp(r, m)$ with $r = 3$ and $r = 4$ can be calculated as well.

5. The weight distribution of $C(r, 3)$ and $C^\perp(r, 3)$. In this section we assume that $r > 2$. Let γ be a primitive element of $\mathbb{F}_{q^3}$, and let $g = N(\gamma)$ be a primitive element of $\mathbb{F} = \mathbb{F}_q$.

Now, by Theorems 1 and 7, we have the following.

LEMMA 7. *For each integer i satisfying $0 \leq i \leq q - 2$, we have*

$$s(\gamma^i) = k(g^i)^2 - q.$$

Hence, the question about the distribution of the values of $s(\gamma^i)$ is equivalent to the question about the distribution of the values of (one-dimensional) Kloosterman sums over $\mathbb{F}^*$. This question has been answered by Lachaud and Wolfmann in [5, Theorem 3.4 and Proposition 9.1].

THEOREM 10. *The set of values S of $k(a)$ as a runs over $\mathbb{F}_q^*$ is*

$$S = \{j \in \mathbb{Z} \mid |j| < 2\sqrt{q} \text{ and } j \equiv -1 \pmod{4}\}.$$

Moreover, each value $j \in S$ is attained exactly $H(j^2 - 4q)$ times, where $H(d)$ is the Kronecker class number of d .

As a corollary we obtain, by using Theorem 2, the weight distribution of $C(r, 3)$.

THEOREM 11. *The weight distribution of $C(r, 3)$ is given in the following table, where j runs over the set $\{|j| < 2^{(r+2)/2} \text{ and } j \equiv -1 \pmod{4}\}$:*

Weight	Frequency
0	1
$2^r(2^{r-1} - 1)$	$3(2^r - 1)^2$
$2^{r-1}(2^r - 1)$	$3(2^r - 1)$
$(2^r(2^r - 1) - j^2 + 1)/2$	$H(j^2 - 2^{r+2})(2^r - 1)^2$

To give the weight distribution of $C^\perp(r, 3)$ we denote by K_h the h th moment of the Kloosterman sum $k(a)$ over the field $\mathbb{F}$, i.e.,

$$K_h = \sum_{a \in \mathbb{F}^*} k(a)^h,$$

and use the following result from [10] which was proved by using results from [14].

THEOREM 12. *Let $q = 2^r$. Then*

$$K_0 = q - 1, \quad K_1 = 1, \quad K_2 = q^2 - q - 1, \quad K_3 = \pm q^2 + 2q + 1,$$

$$K_4 = 2q^3 - 2q^2 - 3q - 1,$$

$$K_5 = (t_7 \pm 4)q^3 + 5q^2 + 4q + 1,$$

$$K_6 = 5q^4 - (5 + (-1)^r)q^3 - 9q^2 - 5q - 1,$$

$$K_7 = (t_9 + 6t_7 \pm 14 + 1)q^4 + 14q^3 + 14q^2 + 6q + 1,$$

$$K_8 = 14q^5 - (15 \pm 7)q^4 - 28q^3 - 20q^2 - 7q - 1,$$

$$K_9 = (t_{11} + 8t_9 + 27t_7 + 8 \pm 48)q^5 + 42q^4 + 48q^3 + 27q^2 + 8q + 1,$$

$$K_{10} = 42q^6 - (51 \pm 35)q^5 - 90q^4 - 75q^3 - 35q^2 - 9q - 1 + 2048\tau(q/4) - \tau(q),$$

where $\pm$ denotes $(-1)^r$, $t_7 = \alpha_7^r + \bar{\alpha}_7^r$ with $\alpha_7 = (1 + \sqrt{-15})/4$, $t_9 = \alpha_9^r + \bar{\alpha}_9^r$ with $\alpha_9 = (-5 + \sqrt{-39})/8$, $t_{11} = \beta_{11}^r + \bar{\beta}_{11}^r + \eta_{11}^r + \bar{\eta}_{11}^r$ with $\beta_{11} = (-3 + \sqrt{505} + \sqrt{-510 - 6\sqrt{505}})/32$, $\eta_{11} = (-3 - \sqrt{505} + \sqrt{-510 + 6\sqrt{505}})/32$, and τ is the Ramanujan tau-function.

Remark 6. It is not hard to see that

$$\tau(q) - 2048\tau(q/4) = \mu_2^r + \bar{\mu}_2^r = D_r(-24, 2048),$$

where $\mu_2 = -12 + 4\sqrt{-119}$ and $D_r(x, 2048)$ is the Dickson polynomial of the first kind of degree r with parameter 2048 (see [11, section 2]).

THEOREM 13. *For every nonnegative integer h the number $C_h^\perp$ of codewords of weight h in the dual $C^\perp(r, 3)$ of $C(r, 3)$ is given by*

$$\begin{aligned} q^3 h! C_h^\perp &= f(C_0^\perp, \dots, C_{h-1}^\perp) + g(M_0, \dots, M_h) \\ &+ 3(q-1)^2 (-q)^h ((q-2)^h + (q-1)^{h-1}), \end{aligned}$$

where

$$\begin{aligned} f(C_0^\perp, \dots, C_{h-1}^\perp) &= q^3 \sum_{i=0}^{h-1} (-1)^{h+i+1} C_i^\perp \sum_{t=i}^h t! S(h, t) 2^{h-t} \binom{n-i}{n-t}, \\ g(M_0, \dots, M_h) &= \sum_{j=0}^h \binom{h}{j} (-1)^{j+h} (q-1)^{2(h-j+1)} \sum_{i=0}^j \binom{j}{i} (-q)^{j-i} K_{2i}. \end{aligned}$$

In particular,

$$\begin{aligned} C_0^\perp &= 1, \quad C_1^\perp = C_2^\perp = 0, \quad C_3^\perp = (q-1)^2 (2q-5 \mp 1)/3!, \\ C_4^\perp &= (q-1)^2 (q^3 - 6q^2 + (17 \mp 3)q - 24)/4!, \\ C_5^\perp &= (q-1)^2 (q^5 - 8q^4 + 14q^3 + 24q^2 - 4(7 \pm 5)q - 109 \mp 10 \\ &\quad + (2048\tau(q/4) - \tau(q))/q^3)/5!. \end{aligned}$$

Proof. The moments M_j in Theorem 4 are, by Lemma 7, of the form

$$\begin{aligned} M_j &= \sum_{l=0}^{q-2} (k(g^l)^2 - q)^j = \sum_{l=0}^{q-2} \sum_{i=0}^j \binom{j}{i} k(g^l)^{2i} (-q)^{j-i} \\ &= \sum_{i=0}^j \binom{j}{i} (-q)^{j-i} \sum_{l=0}^{q-2} k(g^l)^{2i} \\ &= \sum_{i=0}^j \binom{j}{i} (-q)^{j-i} K_{2i}, \end{aligned}$$

and the first claim follows now by Theorem 4. The validity of the formulae for the number of low-weight codewords can be verified by using Mathematica. $\square$

Remark 7. By Theorem 10, moments K_h can be calculated effectively for each nonnegative integer h by

$$K_h = \sum_{\substack{|j| < 2\sqrt{q} \\ j \equiv -1 \pmod{4}}} H(j^2 - 4q)j^h,$$

provided that r is not too large (an “ $H(d)$ -calculator” can be found in [7]).

Acknowledgments. The author would like to thank the anonymous reviewers for their detailed comments.

REFERENCES

- [1] L. CARLITZ, *A note on exponential sums*, Pacific J. Math., 30 (1969), pp. 35–37.
- [2] L. CARLITZ, *Explicit evaluation of certain exponential sums*, Math. Scand., 44 (1979), pp. 5–16.
- [3] K. CHINEN, *On some properties of the hyper-Kloosterman codes*, Tokyo J. Math., 26 (2003), pp. 55–65.
- [4] K. CHINEN AND T. HIRAMATSU, *Hyper-Kloosterman sums and their applications to the coding theory*, Appl. Algebra Engrg. Comm. Comput., 12 (2001), pp. 381–390.
- [5] G. LACHAUD AND J. WOLFMANN, *The weights of orthogonals of the extended quadratic binary Goppa codes*, IEEE Trans. Inform. Theory, 36 (1990), pp. 686–692.
- [6] F. J. MACWILLIAMS AND N. J. A. SLOANE, *The Theory of Error Correcting Codes*, North-Holland, Amsterdam, 1977.
- [7] K. MATTHEWS, *Some BMath/PHP Number Theory Programs*, <http://www.numbertheory.org/php> (2006).
- [8] M. MOISIO, *A note on evaluations of some exponential sums*, Acta Arith., 93 (2000), pp. 117–119.
- [9] M. MOISIO, *On the number of rational points on some families of Fermat curves over finite fields*, Finite Fields Appl., 13 (2007), pp. 546–562.
- [10] M. MOISIO, *The moments of a Kloosterman sum and the weight distribution of a Zetterberg type binary cyclic code*, IEEE Trans. Inform. Theory, 53 (2007), pp. 843–847.
- [11] M. MOISIO, *On the moments of Kloosterman sums and fibre products of Kloosterman curves*, Finite Fields Appl., to appear. (Published online June 16, 2007.)
- [12] C. J. MORENO AND O. MORENO, *The MacWilliams-Sloane conjecture on the tightness of the Carlitz-Uchiyama bound and the weights of duals of BCH codes*, IEEE Trans. Inform. Theory, 40 (1994), pp. 1894–1907.
- [13] V. PLESS, *Power moment identities on weight distributions in error correcting codes*, Inform. and Control, 6 (1963), pp. 147–152.
- [14] R. SCHOOF AND M. VAN DER VLUGT, *Hecke operators and the weight distributions of certain codes*, J. Combin. Theory Ser. A, 57 (1991), pp. 163–186.

ORE CONDITION AND NOWHERE-ZERO 3-FLOWS*

GENGHUA FAN[†] AND CHUIXIANG ZHOU[†]

Abstract. Let G be a simple graph on n vertices, $n \geq 3$. It is well known that if G satisfies the Ore condition that $d(x) + d(y) \geq n$ for every pair of nonadjacent vertices x and y , then G has a Hamiltonian circuit, which implies that G has a nowhere-zero 4-flow. But it is not necessary for G to have a nowhere-zero 3-flow. In this paper, we prove that with six exceptions, all graphs satisfying the Ore condition have a nowhere-zero 3-flow. More precisely, if G is a graph on n vertices, $n \geq 3$, in which $d(x) + d(y) \geq n$ for every pair of nonadjacent vertices x and y , then G has no nowhere-zero 3-flow if and only if G is one of six completely described graphs.

Key words. Ore condition, degree sum, nowhere-zero 3-flow

AMS subject classifications. 05C15, 05C45

DOI. 10.1137/060677744

1. Introduction. The vertex set and edge set of a graph G are denoted by $V(G)$ and $E(G)$, respectively. For $xy \in E(G)$, we call y a *neighbor* of x , and the set of neighbors of x in G is denoted by $N_G(x)$, or simply $N(x)$. Let H be a subgraph of G and $v \in V(G)$, and let $d_H(v) = |N(v) \cap V(H)|$ be the number of the neighbors of v in H . When $H = G$, $d_G(v)$ is called the *degree* of v and is abbreviated as $d(v)$. Denote by $\delta(G)$ and $\Delta(G)$ the minimum and maximum degree of G , respectively. For subgraphs A and B , $e(A, B)$ denotes the number of edges with one end in A and the other end in B .

An edge is *contracted* if it is deleted and its two ends are identified into a single vertex. Let H be a connected subgraph of G . G/H denotes the graph obtained from G by contracting all the edges of H and deleting all the resulting loops. For $S \subseteq V(G)$, $G - S$ denotes the graph obtained from G by deleting all the vertices of S together with all the edges with at least one end in S . When $S = \{v\}$, we simplify this notation to $G - v$. The complete graph on n vertices is denoted by K_n . Denote by K_n^- the graph obtained from K_n by deleting an edge.

A k -circuit is a circuit of k vertices. A *wheel* W_k is the graph obtained from a k -circuit by adding a new vertex, called the *center* of the wheel, which is joined to every vertex of the k -circuit. W_k is an *odd (even) wheel* if k is odd (even). For simplicity, A 3-circuit (triangle) on vertices $\{x, y, z\}$ is denoted by xyz .

Let G be a graph with an orientation. For each vertex $v \in V(G)$, $E^+(v)$ is the set of nonloop edges with tail v , and $E^-(v)$ is the set of nonloop edges with head v . Let $\mathbb{Z}_k$ denote an abelian group of k elements with identity 0. Let f be a function from $E(G)$ to $\mathbb{Z}_k$. Set

$$f^*(v) = \sum_{e \in E^+(v)} f(e) - \sum_{e \in E^-(v)} f(e).$$

*Received by the editors December 15, 2006; accepted for publication (in revised form) August 17, 2007; published electronically February 27, 2008.

<http://www.siam.org/journals/sidma/22-1/67774.html>

[†]Center for Discrete Mathematics, Fuzhou University, Fuzhou, Fujian 350002, China (fan@fzu.edu.cn, cxzhou@fzu.edu.cn). The research of the first author was supported by the National Science Foundation of China.

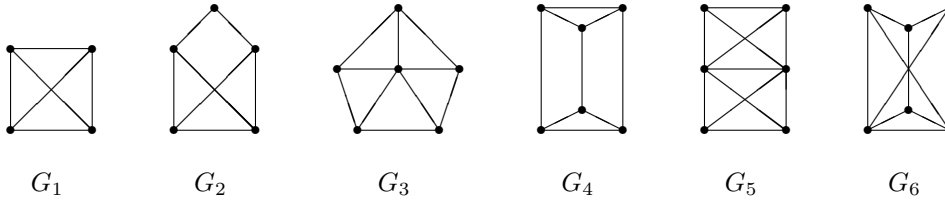


FIG. 1.

f is a $\mathbb{Z}_k$ -flow in G if $f^*(v) = 0$ for each vertex $v \in V(G)$. For an edge $e \in E(G)$, we call $f(e)$ the *flow value* of e . The *support* of f is defined by $S(f) = \{e \in E(G) : f(e) \neq 0\}$. f is *nowhere-zero* if $S(f) = E(G)$. It is well known that a graph G has a nowhere-zero $\mathbb{Z}_k$ -flow if and only if there is an integer-valued function f on $E(G)$ such that $0 < |f(e)| < k$ for each $e \in E(G)$, and $f^*(v) = 0$ for each $v \in V(G)$, which is called a *nowhere-zero k -flow* in G . Therefore, we also call a $\mathbb{Z}_k$ -flow a *k -flow*. Tutte [7] conjectured that every 2-edge-connected graph has a nowhere-zero 5-flow. Seymour [6] proved that every 2-edge-connected graph has a nowhere-zero 6-flow. In this paper, we focus on the nowhere-zero 3-flow. Since loops play no role with respect to existence of nowhere-zero flows, we consider only loopless graphs. The well-known 3-flow conjecture of Tutte (see unsolved problem 48 of [1]) is as follows.

CONJECTURE 1.1. *Every 4-edge-connected graph has a nowhere-zero 3-flow.*

A subgraph H of G is *3-flow contractible* if G/H having a nowhere-zero 3-flow implies that G has a nowhere-zero 3-flow. It is well known that a 2-circuit (two parallel edges) is 3-flow contractible. By results in [2, Proposition 2.4 and Observation 1.3], the even wheel W_{2k} is 3-flow contractible for $k \geq 2$. Similarly, by [3, Corollary 3.5] and [2, Observation 1.3], K_n^- and K_n are 3-flow contractible for $n \geq 5$. In summary, we have the following.

PROPOSITION 1.2 (see [2, 3]). (i) W_{2k} is 3-flow contractible for $k \geq 2$; (ii) K_n^- and K_n are 3-flow contractible for $n \geq 5$.

PROPOSITION 1.3. *None of the six graphs in Figure 1 has a nowhere-zero 3-flow.*

Proof. G_1 is the complete graph K_4 , which does not have nowhere-zero 3-flows, and neither does G_2 since it is obtained from K_4 by subdividing an edge. It is known that the odd wheel has no nowhere-zero 3-flow, and thus G_3 has no nowhere-zero 3-flow. Since G_4 is a cubic graph with triangles, it is not bipartite, and thus has no nowhere-zero 3-flows. For the last two graphs G_5 and G_6 , suppose, to the contrary, that they have nowhere-zero 3-flows. By choosing orientations, we see that they have a $\mathbb{Z}_3$ -flow, in which each edge has a flow value of 1. Then, at each vertex of degree 3, all three edges are either out of or into the vertex, which means that the edge between the two vertices of degree 5 (in G_5) or degree 4 (in G_6) must have a flow value of zero, which is a contradiction. This shows that all six graphs in Figure 1 have no nowhere-zero 3-flows. $\square$

Let G be a simple graph on n vertices. It is well known that if G satisfies the Ore condition [5] that $d(x) + d(y) \geq n$ for every pair of nonadjacent vertices x and y , then G has a Hamiltonian circuit (an n -circuit in G), which implies that G has a nowhere-zero 4-flow. But it is not necessary for G to have a nowhere-zero 3-flow. It is easy to check that each of the six graphs in Figure 1 satisfies the Ore condition but has no nowhere-zero 3-flow. The main result of this paper is that these six graphs are the only ones that satisfy the Ore condition and have no nowhere-zero 3-flow.

MAIN THEOREM. *Let G be a simple graph on n vertices, $n \geq 3$. If $d(x) + d(y) \geq n$ for each $xy \notin E(G)$, then G has no nowhere-zero 3-flow if and only if G is one of the six graphs in Figure 1.*

The bound on the degree sum is best possible. For each $n \geq 7$, there exists a simple graph G on n vertices without a nowhere-zero 3-flow, in which $d(x) + d(y) \geq n - 1$ for each $xy \notin E(G)$. Let G be the graph obtained by joining vertex-disjoint K_3 and K_{n-3} with three independent edges, where $n \geq 7$. Then, $d(x) + d(y) \geq n - 1$ for every $xy \notin E(G)$. However, the contraction of the K_{n-3} results in a K_4 , which has no nowhere-zero 3-flow. Thus, G cannot have a nowhere-zero 3-flow.

2. Lemmas.

LEMMA 2.1. *If G is a cubic simple graph on six vertices, then either G is $K_{3,3}$ or G is G_4 in Figure 1.*

Proof. Clearly, if G contains no odd circuit, then G is $K_{3,3}$. Suppose that G has an odd circuit. Let C be the shortest odd circuit in G . If $|V(C)| = 5$, then C has no chord, which means that the only vertex not in C has degree 5, which is a contradiction. Thus, $|V(C)| = 3$. Since G is cubic and simple, vertices of C have no common neighbor in $V(G) \setminus V(C)$, and hence, G is G_4 in Figure 1. $\square$

LEMMA 2.2. *Let G be a simple graph on n vertices, where $3 \leq n \leq 6$. If $d(x) + d(y) \geq n$ for each $xy \notin E(G)$, then G has no nowhere-zero 3-flow if and only if G is one of the six graphs in Figure 1.*

Proof. If G is one of the six graphs in Figure 1, then, by Proposition 1.3, G has no nowhere-zero 3-flow. Conversely, suppose that G has no nowhere-zero 3-flow. We shall prove that it must be one of the six graphs. By the condition that $d(x) + d(y) \geq n$ for each $xy \notin E(G)$, $\delta(G) \geq 2$, and since G has no nowhere-zero 3-flow, we have that $n \geq 4$. If $n = 4$, by the given condition, G must be K_4 , the graph G_1 in Figure 1. Suppose therefore that $n \geq 5$.

(i) $n = 5$. If $\delta(G) = 2$, let $w \in V(G)$ with $d(w) = 2$ and $N(w) = \{u_1, u_2\}$. Let G' be the graph obtained from G by deleting w and adding u_1u_2 . If G' has a nowhere-zero 3-flow, then so does G , which is impossible. Therefore, G' must be K_4 , which implies that G is G_2 in Figure 1. If $\delta(G) \geq 3$, since n is odd, there is a vertex $u \in V(G)$ such that $d(u) = 4$. If $d(v) = 3$ for each $v \in V(G) \setminus \{u\}$, then G is the even wheel W_4 centered at u ; if there is $v \in V(G) \setminus \{u\}$ such that $d(v) = 4$, then G is K_5 or K_5^- . In either case, by Proposition 1.2, G has a nowhere-zero 3-flow, which is a contradiction.

(ii) $n = 6$. If $\delta(G) = 2$, let w be a vertex of degree 2 in G . Then, by the condition that $d(x) + d(y) \geq n$ for each $xy \notin E(G)$, all three vertices, not the neighbors of w , have degree at least 4, which means that $G - w$ is either K_5 or K_5^- . It follows from Proposition 1.2(ii) that G has a nowhere-zero 3-flow, which is a contradiction. Thus we may assume that $\delta(G) \geq 3$.

If $\Delta(G) = 3$, then G is cubic. By Lemma 2.1, G is G_4 in Figure 1.

If $\Delta(G) = 4$, since the number of the vertices of odd degrees in a graph must be even, G has at least two vertices of degree 4. If all the vertices of G have degree 4, then G has a nowhere-zero 2-flow, which is a contradiction. Suppose thus that G has at most four vertices of degree 4. Let x and y be two vertices of degree 4.

If $xy \notin E(G)$, then $N(x) = N(y)$. If there is a vertex $v \in N(x)$ with $d(v) = 4$, then there is an even wheel W_4 centered at v , and by Proposition 1.2, G has a nowhere-zero 3-flow; if $d(v) = 3$ for each $v \in N(x)$, then the subgraph induced by $N(x)$ consists of two independent edges. Then it is easy to check that G has a nowhere-zero 3-flow. In either case, we have a contradiction. Therefore, $xy \in E(G)$. By the arbitrariness of x

and y , we have that the subgraph induced by vertices of degree 4 in G is a complete subgraph.

If there are precisely two vertices of degree 4, say x and y , let G' be the graph obtained from G by removing the edge xy . Then G' is a cubic simple graph on six vertices. By Lemma 2.1, G' is either $K_{3,3}$ or G_4 in Figure 1. In the former case, G is the graph G_6 in Figure 1; in the later case, G is the graph obtained from G_4 by adding an edge between the two triangles, and it is easy to check that G has a nowhere-zero 3-flow, which is impossible. Suppose therefore that there are four vertices, say u, v, x , and y , of degree 4 in G . Recall that the subgraph induced by $\{x, y, u, v\}$ is K_4 . Let G' be the graph obtained from G by removing the edges xy and uv . Then G' is a cubic simple graph on six vertices. As before, by Lemma 2.1, G' is either $K_{3,3}$ or G_4 in Figure 1. It is not difficult to check that in either case, G has a nowhere-zero 3-flow, which is a contradiction.

If $\Delta(G) = 5$, let $u \in V(G)$ with $d(u) = 5$ and let N denote the subgraph induced by $N(u)$. We note that an even circuit in N together with u gives an even wheel centered at u , which implies, by Proposition 1.2(i), that G has a nowhere-zero 3-flow. Thus, we may assume that there is no even circuit in N . Then, each block of N is either K_1 or K_2 , or an odd circuit (see exercise 3.2.3 in [1]). But, $\delta(G) \geq 3$ and $|N(x)| = 5 = |V(G)| - 1$, and so each block of N is an odd circuit and N has at most two blocks. If N has exactly one block, then N is a circuit of length 5, and hence G is G_3 in Figure 1; if N has two blocks, then N consists of two triangles with exactly one vertex in common, and hence G is G_5 in Figure 1. This completes the proof of Lemma 2.2. $\square$

LEMMA 2.3. *Let G be a simple graph on n vertices, where $n \geq 7$. If $d(x) + d(y) \geq n$ for each $xy \notin E(G)$, then either G has a nowhere-zero 3-flow or G contains K_4^- .*

Proof. Suppose first that $\delta(G) \leq \frac{n-1}{2}$. Let $u \in V(G)$ with $d(u) = \delta(G)$. Let $R = G - (N(u) \cup \{u\})$. Then $|V(R)| \geq \frac{n-1}{2} \geq 3$. For any $v \in V(R)$, by the given condition, $d(v) + d(u) \geq n$, and so $d(v) \geq \frac{n+1}{2}$, which implies that there is an edge $xy \in E(R)$. Note that $d(x) + d(y) \geq \frac{n+1}{2} + \frac{n+1}{2} = n + 1$, but $u \notin N(x) \cup N(y)$, and so $|N(x) \cap N(y)| \geq 2$. Any two vertices in $N(x) \cap N(y)$ together with x and y induce a K_4^- . Therefore, we assume that $\delta(G) \geq \frac{n}{2}$. Let $e = |E(G)|$. Then $e \geq \frac{n^2}{4}$, and by the Turan theorem, either G contains a triangle or G is the complete bipartite graph $K_{m,m}$, where $m = \frac{n}{2}$.

(i) G contains a triangle $T = v_1v_2v_3$. For any $u \in V(G) \setminus V(T)$, if u has two neighbors in T , then we have a K_4^- . Suppose therefore that each vertex in $V(G) \setminus V(T)$ has at most one neighbor in T . It follows that

$$d(v_1) + d(v_2) + d(v_3) \leq (n - 3) + 6 = n + 3.$$

But $\delta(G) \geq \frac{n}{2}$, and thus we have that

$$d(v_1) + d(v_2) + d(v_3) \geq \frac{3n}{2}.$$

Combining the two inequalities yields that $n \leq 6$, which is a contradiction.

(ii) G is the complete bipartite graph $K_{m,m}$, where $m = \frac{n}{2}$. It is known that $K_{m,m}$ has a nowhere-zero 3-flow (see [4]). For completeness, we present a proof here. If m is even, then $K_{m,m}$ is Eulerian and has a nowhere-zero 2-flow; if m is odd, $K_{m,m}$ can be decomposed into a set of 2-factors (each has a nowhere-zero 2-flow) and a bipartite cubic graph, which is known to have a nowhere-zero 3-flow, and combining these flows yields a nowhere-zero 3-flow in $K_{m,m}$. In either case, G has a nowhere-zero 3-flow. This completes the proof of Lemma 2.3. $\square$

3. Proof of main theorem. If G is one of the six graphs in Figure 1, then, by Proposition 1.3, G has no nowhere-zero 3-flow. Conversely, suppose that G is not any of the six graphs in Figure 1. We shall prove that G has a nowhere-zero 3-flow. By the given degree-sum condition, we have that $\delta(G) \geq 2$.

We use induction on $n = |V(G)|$. If $n \leq 6$, the theorem holds by Lemma 2.2. Suppose thus that $n \geq 7$ and the theorem holds for any graph G' with $|V(G')| < n$. If G has a vertex of degree 2, let $w \in V(G)$ with $d(w) = 2$ and $H = G - w$. By the given degree-sum condition, each vertex in $V(H) \setminus N(w)$ has degree at least $n - 2$, which implies that H is either K_{n-1} or K_{n-1}^- . It follows from Proposition 1.2(ii) that G has a nowhere-zero 3-flow. In what follows, we assume that $\delta(G) \geq 3$.

CLAIM 1. *If there are distinct vertices $x, y \in V(G)$ such that $d(x) = d(y) = 3$, then G has a nowhere-zero 3-flow.*

By the given degree-sum condition, and since $n \geq 7$, we have that $xy \in E(G)$. Let $B = N(x) \cup N(y)$ and $R = G - B$. By the given condition again, for each vertex $v \in V(R)$, $d(v) \geq n - 3$. If $|V(R)| \geq 3$, then R together with any two vertices of $B \setminus \{x, y\}$ induces a K_m or K_m^- , where $m \geq 5$. By Proposition 1.2(ii), G has a nowhere-zero 3-flow. Therefore, we assume that $|V(R)| \leq 2$.

(i) $|V(R)| = 1$, then $n = 7$, and by the given degree-sum condition, there is a W_4 centered at the unique vertex of R . By Proposition 1.2(i), G has a nowhere-zero 3-flow.

(ii) $|V(R)| = 2$, then $|V(B)| = 6$ or 5 . If $|V(B)| = 6$, then as in case (i), we have a W_4 and are done. If $|V(B)| = 5$, then G is a graph on seven vertices, which either contains a K_5 or K_5^- , or each vertex, except for x and y , has degree 4. In any case, G has a nowhere-zero 3-flow. This proves the claim.

By Lemma 2.3, we may assume that G contains a K_4^- . Consider the K_4^- as the union of two triangles xyz and xyw , with edge xy in common. By Claim 1, either $d(z) \geq 4$ or $d(w) \geq 4$. Without loss of generality, we assume that $d(z) \geq 4$. Let G' be the graph obtained from G by deleting zx , zy and adding xy . In G' , we have a 2-circuit on $\{x, y\}$. Let G^1 be the graph obtained from G' by contracting the 2-circuit on $\{x, y\}$ into a single vertex u^* . Then, G^1 has a 2-circuit on $\{u^*, w\}$. Let G^2 be the graph obtained from G^1 by contracting the 2-circuit on $\{u^*, w\}$; for convenience, the resulting new vertex is still denoted by u^* . If there is a 2-circuit containing u^* , we continue to contract the 2-circuit into u^* and denote the resulting graph by G^3 . Keep going this way until no 2-circuit exists. Let $G^1, G^2, \dots, G^t$ be a sequence of graphs obtained in the above-described way of contracting 2-circuits. So, for each i , G^i is obtained from G^{i-1} by contracting a 2-circuit into u^* , $1 \leq i \leq t$, where $G^0 = G'$. Note that G^t is a simple graph, in which all vertices, except for u^* and z , have the same degree as in G . Since G^t is obtained from G' by consecutively contracting 2-circuits, if G^t has a nowhere-zero 3-flow, then so does G' . Moreover, there is a subgraph H such that $G^t = G'/H$. Let $|V(G^t)| = n^*$. Since $t \geq 2$, we have that $n^* \leq n - 2$. If $n^* \leq 3$, then G^t has a nowhere-zero 3-flow, which implies that G' , and so G , has a nowhere-zero 3-flow. Thus, assume that $n^* \geq 4$.

CLAIM 2. *For any distinct $u, v \in V(G^t)$ and $uv \notin E(G^t)$, $d_{G^t}(u) + d_{G^t}(v) \geq n^*$.*

If $u, v \in V(G^t) \setminus \{z, u^*\}$, then

$$d_{G^t}(u) + d_{G^t}(v) = d_G(u) + d_G(v) \geq n > n^*.$$

If $u \neq u^*$ and $v = z$, then, using that $d_{G^t}(z) = d_G(z) - 2$, we have

$$d_{G^t}(u) + d_{G^t}(v) = d_G(u) + d_G(v) - 2 \geq n - 2 \geq n^*.$$

If $u = u^*$ and $v \neq z$, then, since $d_{G^t}(u^*) \geq d_G(w) - (|V(H)| - 1)$, we have that

$$d_{G^t}(u) + d_{G^t}(v) \geq d_G(w) - (|V(H)| - 1) + d_G(v) \geq n - (|V(H)| - 1) = n^*.$$

What remains is the case that $u = u^*$ and $v = z$. Let $R = G - V(H)$. For any $a \in V(H) \setminus \{x, y\}$, we have that

$$d_{G^t}(u^*) \geq d_G(a) - (|V(H)| - 1) + e(H - a, R - z).$$

Since $d_{G^t}(z) = d_G(z) - 2$, it follows that

$$d_{G^t}(z) + d_{G^t}(u^*) \geq d_G(z) - 2 + d_G(a) - (|V(H)| - 1) + e(H - a, R - z).$$

Using that $d_G(z) + d_G(a) \geq n$ and $n - (|V(H)| - 1) = n^*$, we obtain that

$$d_{G^t}(z) + d_{G^t}(u^*) \geq n^* - 2 + e(H - a, R - z).$$

If $e(H - a, R - z) \geq 2$, then $d_{G^t}(z) + d_{G^t}(u^*) \geq n^*$ as claimed. Suppose therefore that for any $a \in V(H) \setminus \{x, y\}$,

$$(3.1) \quad e(H - a, R - z) \leq 1.$$

If $|V(H)| = 3$, that is, $V(H) = \{x, y, w\}$, then by (3.1), $e(xy, R - z) \leq 1$. We may assume that $e(x, R - z) = 0$. Then, $d_G(x) = 3$, and by the given degree-sum condition, for each $r \in V(R - z)$, we have $d_G(r) \geq n - 3$, which implies that $e(r, \{z, y, w\}) \geq 2$, and so

$$e(\{z, y, w\}, R - z) \geq 2|V(R - z)| = 2(n - 4).$$

But $d_{G^t}(z) + d_{G^t}(u^*) = e(\{z, y, w\}, R - z)$, and so $d_{G^t}(z) + d_{G^t}(u^*) \geq 2(n - 4) > n - 2 \geq n^*$ as claimed.

If $|V(H)| \geq 4$, for any $a \in V(H) \setminus \{x, y\}$, we have that

$$(3.2) \quad e(a, R) \geq 1;$$

otherwise, for each $r \in V(R - z)$, by the given degree-sum condition, we have $e(r, H) \geq 2$, contradicting the fact that G^t is simple. By (3.1) and (3.2), we see that $|V(H)| = 4$, say $V(H) = \{x, y, w, b\}$. Furthermore, since $u^*z \notin E(G^t)$, we have that $e(z, V(H) \setminus \{x, y\}) = 0$. It follows that

$$(3.3) \quad e(xy, R - z) = 0 \quad \text{and} \quad e(w, R - z) = e(b, R - z) = 1.$$

Since $n^* \geq 4$, we have that $|V(R - z)| \geq 2$. By Claim 1, there is a vertex in $R - z$ with degree at least 4 in G . It follows from (3.3) that $|V(R)| \geq 4$, and there is an $r \in V(R - z)$ such that $e(r, H) = 0$. Then $|N(r) \cap N(w)| \leq 1$, which implies that $d_G(r) + d_G(w) \leq n - 1$. This contradiction completes the proof of Claim 2.

By Claim 2 and the induction hypothesis, either G^t has a nowhere-zero 3-flow or G^t is one of the six graphs in Figure 1. In the former case, since G^t is obtained from G' by consecutively contracting 2-circuits, we see that G' , and thus G , has a nowhere-zero 3-flow, and we are done. In the latter case, G^t contains at least four vertices of degree 3, at least two of which have degree 3 in the original graph G , and by Claim 1, G has a nowhere-zero 3-flow. This completes the proof of the main theorem. $\square$

REFERENCES

- [1] J. A. BONDY AND U. S. R. MURTY, *Graph Theory with Applications*, North-Holland, New York, 1976.
- [2] M. DEVOS, R. XU, AND G. YU, *Nowhere-zero $\mathbb{Z}_3$ -flows through $\mathbb{Z}_3$ -connectivity*, Discrete Math., 306 (2006), pp. 26–30.
- [3] H.-J. LAI, *Group connectivity of 3-edge-connected chordal graphs*, Graphs Combin., 16 (2000), pp. 165–176.
- [4] H.-J. LAI AND X. YAO, *Group connectivity of graphs with diameter at most 2*, European J. Combin., 27 (2006), pp. 436–447.
- [5] O. ORE, *Note on Hamilton circuits*, Amer. Math. Monthly, 67 (1960), p. 55.
- [6] P. D. SEYMOUR, *Nowhere-zero 6-flows*, J. Combin. Theory Ser. B, 30 (1981), pp. 130–155.
- [7] W. T. TUTTE, *A contribution to the theory of chromatic polynomials*, Canad. J. Math., 6 (1954), pp. 80–91.

LARGE MONOCHROMATIC COMPONENTS IN TWO-COLORED GRIDS*

JIŘÍ MATOUŠEK[†] AND ALEŠ PŘÍVĚTIVÝ[‡]

Abstract. Let D_n^d denote the d -dimensional *grid with diagonals*, that is, the graph with vertex set $\{1, 2, \dots, n\}^d$ and with edges connecting every two vertices that differ by at most 1 in every coordinate. We prove that for an arbitrary coloring of the vertices of D_n^d by two colors there exists a monochromatic connected subgraph with at least $n^{d-1} - d^2 n^{d-2}$ vertices; and thus the “horizontal layer” coloring (by the parity of the first coordinate) is almost optimal. We also consider a d -dimensional *triangulated grid*; this is the graph (1-skeleton) of a triangulation of the solid cube $[1, n]^d$ that refines the subdivision of $[1, n]^d$ into the grid of unit cubes. Here every two-coloring has a monochromatic connected subgraph with $\Omega(n^{d-1}/\sqrt{d})$ vertices. These results are proved by combining combinatorial and topological arguments with suitable isoperimetric inequalities, and they can be viewed as d -dimensional generalizations of the planar HEX lemma.

Key words. graph coloring, HEX lemma, algebraic topology in combinatorics, isoperimetric inequality

AMS subject classifications. 05C15, 05C35, 57M15

DOI. 10.1137/070684112

1. Introduction. *The HEX lemma.* Let T_n^2 be the graph of an $n \times n$ triangulated square grid in the plane as in Figure 1. Here n counts the number of rows and columns of vertices; so the picture has $n = 6$. The well-known planar HEX lemma asserts that if the vertices of T_n^2 are colored red and blue in an arbitrary way, then there exists a path in T_n^2 connecting the top and bottom sides and having all vertices red, or a path connecting the left and right sides and having all vertices blue. In particular, under any two-coloring of the vertices of T_n^2 there is a monochromatic connected subgraph with at least n vertices.

Triangulated grids. Our results can be regarded as d -dimensional generalizations of the HEX lemma. First, we consider a d -dimensional *triangulated grid* T_n^d , which is defined geometrically as follows. We begin with the d -dimensional solid cube $[1, n]^d$ and subdivide it into the grid of unit cubes; each unit cube is of the form $[i_1, i_1 + 1] \times [i_2, i_2 + 1] \times \dots \times [i_d, i_d + 1]$, $1 \leq i_1, i_2, \dots, i_d \leq n - 1$. Then we triangulate each of the unit cubes in such a way that the simplices of all of these triangulations taken together form a triangulation of $[1, n]^d$ (here a triangulation is a simplicial complex in the sense of algebraic topology, i.e., a face-to-face triangulation with no additional vertices; see section 2). Then T_n^d is the graph (or 1-skeleton) of such a triangulation; that is, the vertex set is $[n]^d$ (where we use the shorthand $[n]$ for the set $\{1, 2, \dots, n\}$), and the edges of T_n^d are the edges of the triangulation. Thus, for given n and d , T_n^d is not defined uniquely, but rather it stands for an arbitrary graph from a (finite) family.

*Received by the editors March 1, 2007; accepted for publication (in revised form) October 22, 2007; published electronically February 27, 2008. Part of this work has been published in a preliminary form in the Proceedings of Eurocomb 2007.

<http://www.siam.org/journals/sidma/22-1/68411.html>

[†]Department of Applied Mathematics and Institute of Theoretical Computer Science (ITI), Charles University, Malostranské nám. 25, 118 00 Praha 1, Czech Republic (matousek@kam.mff.cuni.cz).

[‡]Department of Applied Mathematics, Charles University, Malostranské nám. 25, 118 00 Praha 1, Czech Republic (privetivy@kam.mff.cuni.cz). This author’s research is supported by a grant from the Czech Science Foundation GD201/05/H014.

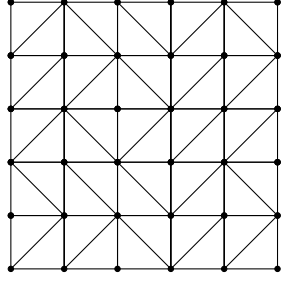


FIG. 1. A triangulated square grid.

We consider an arbitrary coloring of $[n]^d$ by two colors and we ask, What can be said about the number of vertices of the largest monochromatic connected subgraph of T_n^d ? The obvious “horizontal layer” coloring, by the parity of the first coordinate, shows an upper bound of n^{d-1} for this quantity in the worst case, and one might suspect this coloring to be optimal. However, it turns out that for $d \geq 3$, there are better colorings, at least for *some* of the possible triangulated grids. Namely, let us define the i th diagonal layer $L_n^d(i)$ by $L_n^d(i) = \{x \in [n]^d : \sum_{j=1}^d x_j = i\}$, and the *diagonal-layer coloring* by coloring all $L_n^d(i)$ with even i red and odd i blue. As we will verify in section 7, there are triangulated grids T_n^d with no edges connecting $L_n^d(i)$ and $L_n^d(i')$ with $|i - i'| \geq 2$, and, for these, the largest monochromatic connected component is the largest diagonal layer. For $d = 3$, for example, $\max_i |L_n^d(i)|$ is approximately $\frac{3}{4}n^2$.

We prove that the diagonal-layer coloring is not far from optimal in the worst case. To state the result precisely, let us define, for $\alpha \in [0, 1]$ and given n and d ,

$$i_\alpha = \min \left\{ i : \sum_{j \leq i} |L_n^d(j)| > \alpha n^d \right\}.$$

It can be checked that $i_{1/2} = \lceil \frac{(n+1)d}{2} \rceil$ and that $L_n^d(i_{1/2})$ is either the single largest diagonal layer or one of the two largest diagonal layers; see Lemma 5.3. We prove the following theorem.

THEOREM 1.1. *For an arbitrary two-coloring of the vertices of a d -dimensional triangulated grid T_n^d , $n \geq 3$, there exists a monochromatic connected subgraph with at least $|L_n^d(i_{2/3})| - |L_n^{d-1}(i_{1/2})|$ vertices. The last quantity is of order $\frac{n^{d-1}}{\sqrt{d}}$.*

We conjecture that the right answer is $|L_n^d(i_{1/2})|$ instead of $|L_n^d(i_{2/3})| - |L_n^{d-1}(i_{1/2})|$; that is, the diagonal-layer coloring is optimal in the worst case. This would follow from a natural and very plausible-looking conjecture stated in section 8. The numerical values of $|L_n^d(i_{1/2})|$ and $|L_n^d(i_{2/3})| - |L_n^{d-1}(i_{1/2})|$ are not too far from each other, as is illustrated by Table 1 (showing approximate values for small fixed d and $n \rightarrow \infty$).

TABLE 1

	$ L_n^d(i_{1/2}) $	$ L_n^d(i_{2/3}) - L_n^{d-1}(i_{1/2}) $
$d = 3$	$0.750n^2$	$0.698n^2$
$d = 4$	$0.667n^3$	$0.609n^3$
$d = 5$	$0.599n^4$	$0.549n^4$
$d = 10$	$0.430n^9$	$0.393n^9$
$d = 20$	$0.307n^{19}$	$0.280n^{19}$

Let us remark that the proof of Theorem 1.1 works for graphs G somewhat more general than T_n^d . However, since we do not see any good use of such greater generality at the moment, we stick to the concrete geometric formulation as before. If needed, the properties of the graph actually required can be reconstructed from the proof.

Grid with all diagonals. Next, we consider the d -dimensional grid with *all* diagonals of the unit cubes added. That is, we define the graph D_n^d , the d -dimensional *grid with diagonals*, as the graph with vertex set $[n]^d$ and edge set $\{\{u, v\} : u, v \in [n]^d, \|u - v\|_\infty \leq 1\}$, where $\|u - v\|_\infty = \max_i |u_i - v_i|$. In this case we show that the horizontal-layer coloring is almost optimal (the remaining lower order term is probably an artifact of the proof method).

THEOREM 1.2. *For an arbitrary two-coloring of the vertices of the grid with diagonals D_n^d , there exists a monochromatic connected subgraph with at least $n^{d-1} - d^2n^{d-2}$ vertices.*

Related work. There is a different and well-known d -dimensional generalization of the planar HEX lemma appearing, e.g., in Gale [G79] (also see Linial and Saks [LS93] for a different proof and an application in computer science). It asserts that if the vertices of T_n^d are colored by d colors, then there is a monochromatic path connecting two opposite facets of the cube $[1, n]^d$. In particular, there is a monochromatic connected component with at least n vertices.

The problem considered in this paper fits in the following general context. For an arbitrary graph G and an integer k , let us define $\text{mcc}_k(G)$ as the smallest m such that there exists a coloring of the vertices of G by k colors with no monochromatic connected subgraph having more than m vertices. (The letters mcc abbreviate “monochromatic connected component.” The usual chromatic number $\chi(G)$ equals $\min\{k : \text{mcc}_k(G) = 1\}$.)

The quantity $\text{mcc}_k(G)$ was investigated by Kleinberg, Motwani, Raghavan, and Venkatasubramanian [KMR+97] in the context of data storage management. They call a coloring witnessing $\text{mcc}_k(G) \leq c$ a (k, c) -*fragmented coloring*. They proved, among other things, that mcc_3 can be arbitrarily large for planar graphs (we note that $\text{mcc}_4 = 1$ for planar graphs by the four-color theorem), and that for a suitable absolute constant $\varepsilon > 0$ and for every Δ there is an infinite family $\{G_n\}$ of graphs of maximum degree Δ with $\text{mcc}_{\varepsilon\sqrt{\Delta}}(G_n) > \varepsilon|V(G_n)|$.

Apparently independent of the work of Kleinberg et al., $\text{mcc}_k(G)$ has been studied extensively by graph-theorists for graphs of bounded degree, mainly for $k = 2$. It is easy to see that any graph G of maximum degree 3 has $\text{mcc}_2(G) \leq 2$. Alon, Ding, Oporowski, and Vertigan [ADO+03] proved that every graph G of maximum degree 4 satisfies $\text{mcc}_2(G) \leq 57$. Haxell, Szabó, and Tardos [HST03] improved this to $\text{mcc}_2(G) \leq 6$ and proved that $\text{mcc}_2(G) \leq 20000$ for every graph G of maximum

degree 5. On the other hand, Alon et al. [ADO+03] constructed 6-regular graphs G with $\text{mcc}_2(G)$ arbitrarily large. For graphs G of maximum degree 3 it was also shown in [BS05] that they admit two-coloring where one color induces an independent set, while the other color induces components of size at most 189. Earlier work on this subject, [DOS+96] and [JW96], mainly focused on more specific questions concerning line graphs of 3-regular graphs. These investigations culminated in [Tho99] showing that the edges of every 3-regular graph can be two-colored so that each monochromatic component is a path of length at most 5. Recently Linial, Matoušek, Sheffet, and Tardos [LMST07] investigated the order of magnitude of mcc_k for several graph classes; for example, they showed that $\text{mcc}_2(G) = O(n^{2/3})$ for every planar graph on n vertices, and this bound is tight for some planar graphs.

Outline of the proofs of Theorems 1.1 and 1.2. To prove Theorem 1.1, we consider a red-blue coloring of $[n]^d$. Assuming that no monochromatic connected component in T_n^d is very large, we prove that T_n^d has a connected monochromatic *separator*, where a separator in a graph G is a subset $S \subseteq V(G)$ whose removal disconnects G into components of size at most $\frac{1}{2}|V(G)|$. The only property of T_n^d that we use for this part is that it is the graph of a simply connected simplicial complex. Theorem 1.1 then follows from a known result about the vertex expansion of the ordinary grid graph (not triangulated). Theorem 1.2 follows along similar lines: We verify that D_n^d is also the graph of a simply connected simplicial complex (which is not obvious here, unlike in the case of T_n^d), and then we apply a bound on the vertex expansion of D_n^d . We derive this bound from an edge-isoperimetric inequality for the ordinary grid.

2. Topological preliminaries. Here we present some material from elementary algebraic topology. We briefly recall even very standard things in order to make the paper more accessible; see, e.g., [Hat01] or [Mun84] for more details and background.

Simplicial complexes. A (geometric) *simplex* σ is the convex hull of a finite affinely independent set A in some $\mathbf{R}^d$. The points of A are the *vertices* of σ . The *dimension* of σ is $\dim \sigma := |A| - 1$. Thus a *k-simplex* (*k-dimensional simplex*) has $k + 1$ vertices. The convex hull of an arbitrary subset of vertices of a simplex σ is a *face* of σ (every face is itself a simplex).

A nonempty family K of simplices is a (geometric) *simplicial complex* if the following two conditions hold:

1. Each face of any simplex $\sigma \in K$ is also a simplex of K .
2. The intersection $\sigma_1 \cap \sigma_2$ of any two simplices $\sigma_1, \sigma_2 \in K$ is a face of both σ_1 and σ_2 .

In this paper we consider only simplicial complexes with finitely many simplices. A *subcomplex* of a simplicial complex K is a subset L of the simplices of K that constitutes a simplicial complex.

The union of all simplices in a simplicial complex K is the *polyhedron* of K and is denoted by $\|K\|$. In this situation K is called a *triangulation* of the topological space $\|K\|$.

The *vertex set* of K , denoted by $V(K)$, is the union of the vertex sets of all simplices of K . The *graph* (or *1-skeleton*) of K , denoted by $G(K)$, has the vertex set $V(K)$; two vertices are connected by an edge if they are contained in a common 1-simplex (edge) of K .

Simplicial maps and simplicial approximation. Let K and L be simplicial complexes. A *simplicial map* of K into L is a continuous map $f: \|K\| \rightarrow \|L\|$ such that the image of every simplex $\sigma \in K$ is a simplex of L (in particular, every vertex of K is mapped to a vertex of L); moreover, f restricted to each simplex $\sigma \in K$ is an affine

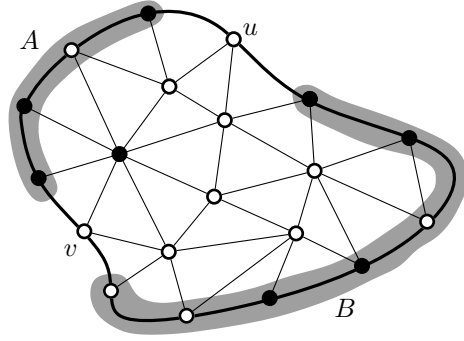


FIG. 2. A variant of the HEX lemma.

map.

Let K and $\tilde{K}$ be simplicial complexes. We call $\tilde{K}$ a *refinement* of K if $\|\tilde{K}\| = \|K\|$ and every simplex of $\tilde{K}$ is contained in some simplex of K (and thus for every nonempty simplex σ of K there is a subcomplex of $\tilde{K}$ that is a triangulation of σ).

We will need the following proposition, which is a variation of the standard *simplicial approximation theorem*.

PROPOSITION 2.1. *Let K and L be simplicial complexes, and let $f: \|K\| \rightarrow \|L\|$ be an arbitrary continuous map. Then there is a refinement $\tilde{K}$ of K and a simplicial map $\tilde{f}$ of $\tilde{K}$ into L such that*

- (*) *If $\sigma \in K$ is a simplex such that $f(\sigma)$ is completely contained in some simplex $\tau \in L$, then $\tilde{f}$ maps all vertices of K lying in σ to vertices of τ . In particular, if $f(v) \in V(L)$ for some $v \in V(K)$, then $\tilde{f}(v) = f(v)$.*

The usual simplicial approximation theorem also yields $\tilde{f}$ homotopic to f (but we do not need this part). On the other hand, condition (*) does not appear in the formulations of the simplicial approximation theorem in the literature, but it immediately follows from the standard proof; see, for example, [Hat01] or [Mun84].

A variant of the HEX lemma. Let T be a 2-dimensional simplicial complex with $\|T\|$ homeomorphic to B^2 , the unit disk in the plane, and let C_T be the cycle in the graph $G(T)$ corresponding to the boundary of B^2 . (So $G(T)$ can be drawn in the plane with the outer face bounded by C_T and with all inner faces being triangles.) We need the following variant of the HEX lemma.

LEMMA 2.2. *Let T and C_T be as above, and let the vertices of T be colored red and blue. Let the vertices of C_T be partitioned into four sets, forming consecutive segments along C_T (see Figure 2), as follows: a set A of red and blue vertices, a single red vertex u , a set B of red and blue vertices, and a single red vertex v . If there is no blue path (path consisting of blue vertices) from A to B , then u and v are connected by a red path. See Figure 2.*

Proof. The case where C_T has four vertices (and thus A and B are singletons) and A and B are blue is Claim 6.1.4 in [MN98]. The case of A, B arbitrarily reduces to the previous case by adding two new vertices a and b in the outer face (see Figure 3), coloring them blue, connecting a to all vertices of $A \cup \{u, v\}$, and connecting b to all vertices of $B \cup \{u, v\}$. $\square$

3. Monochromatic connected separators. We recall that a topological space X is *simply connected* if each continuous map $f: S^1 \rightarrow X$ can be extended to a continuous map $\tilde{f}: B^2 \rightarrow X$ (S^1 denotes the unit circle, the boundary of B^2). A sim-

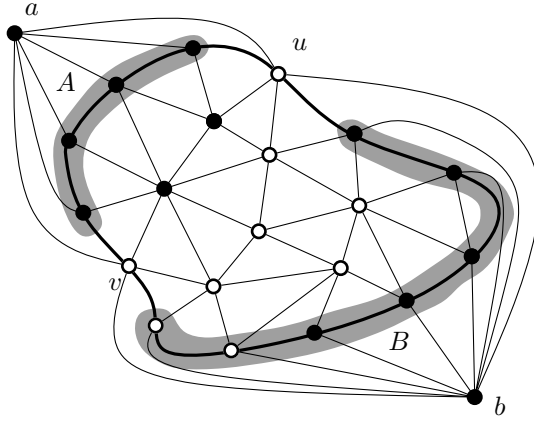


FIG. 3. A reduction to the 4-cycle case.

plicial complex K is simply connected if the topological space $\|K\|$ is simply connected.

Here is the main result of this section.

PROPOSITION 3.1. *Let $G = G(K)$ be the graph of a (finite) simply connected simplicial complex K . For an arbitrary coloring of $V(G)$ by two colors there exists a monochromatic connected separator, that is, a subset $S \subseteq V(G)$ such that all vertices in S have the same color, the subgraph induced by S is connected, and each component of $G \setminus S$ has at most $\frac{1}{2}|V(G)|$ vertices.*

We need the following lemma.

LEMMA 3.2. *Let G be a graph as in Proposition 3.1 with $V(G)$ colored red and blue. Let C be a cycle in G , let $u, v \in V(C)$ be two distinct red vertices, and let A and B be the two connected components (arcs) of $C \setminus \{u, v\}$. If there is no blue path from a vertex of A to a vertex of B in G , then u and v are connected by a red path.*

Proof. The goal is to apply Lemma 2.2. Let us take a triangulated disk T with $\|T\|$ homeomorphic to the unit disk and with the boundary cycle C_T such that C_T has the same length as C . Let $f_0: V(C_T) \rightarrow V(C)$ be an isomorphism of the two cycles, and let $f: \|C_T\| \rightarrow \|C\|$ be the (unique) extension of f_0 to a simplicial map. (So here we also regard C_T as a 1-dimensional subcomplex of T and C as a 1-dimensional subcomplex of K .) Since $\|T\|$ is homeomorphic to B^2 and since K is simply connected, f can be extended to a continuous map $\tilde{f}: \|T\| \rightarrow \|K\|$.

Applying our version of the simplicial approximation theorem (Proposition 2.1) to the mapping $\tilde{f}$, we obtain a refinement $\tilde{T}$ of T and a simplicial map $\tilde{f}$ of $\tilde{T}$ into K . Let $\tilde{C}_T$ be the subcomplex of $\tilde{T}$ that refines C_T (so $\|\tilde{C}_T\| = \|C_T\|$ is the boundary of the disk $\|T\|$).

According to condition (*) in Proposition 2.1, we have $\tilde{f}(v) = f(v)$ for every vertex v of C_T , and for every edge (1-simplex) e of C_T with vertices u and v , all vertices of $\tilde{C}_T$ lying in e are mapped by $\tilde{f}$ to either $f(u)$ or $f(v)$.

Each vertex of $\tilde{T}$ is mapped to a vertex $\tilde{f}(v) \in V(G)$ since $\tilde{f}$ is a simplicial map, so we can define the color of each $v \in V(\tilde{T})$ as the color of $\tilde{f}(v)$. We note that if $x_1, x_2, \dots, x_k$ is a monochromatic path in $\tilde{T}$, then $\tilde{f}(x_1), \tilde{f}(x_2), \dots, \tilde{f}(x_k)$ is a monochromatic walk in G . A monochromatic walk can be shortcut to a monochromatic path.

Now we check that $\tilde{T}$, $\tilde{C}_T$, and the two red vertices $u_T = f^{-1}(u)$ and $v_T = f^{-1}(v)$ satisfy the conditions of Lemma 2.2. Since the two components A and B of

$C \setminus \{u, v\}$ are not connected by a blue path in G , it follows that the two components of $\tilde{C}_T \setminus \{u_T, v_T\}$ are not connected by a blue path in $\tilde{T}$ either. Therefore, by Lemma 2.2, u_T and v_T are connected by a red path, and u and v are connected by a red path too, as claimed. $\square$

Proof of Proposition 3.1. If the vertices of G are colored with only one color, we can take $V(G)$ for S and we are done. Otherwise, we choose S as the vertex set of a monochromatic connected subgraph of G that is

- inclusion-maximal (no vertex of the same color can be added), and
- such that the largest component G_1 of $G \setminus S$ has the smallest possible number of vertices.

Without loss of generality, let S be colored blue.

We claim that G_1 has at most $\frac{1}{2}|V(G)|$ vertices. For contradiction, we thus assume $|V(G_1)| > \frac{1}{2}|V(G)|$. We define S' as the union of all inclusion-maximal red connected subgraphs of G_1 that have at least one vertex connected to S (so S' is the “red foam” on the blue surface of S facing G_1). We will show that S' is “better” than S ; namely, that S' induces an inclusion-maximal connected red subgraph of G and that the largest component of $G \setminus S'$ is smaller than G_1 . The set S' is red by definition.

Next, we show that S' induces a connected subgraph (then it will also be clear that it is inclusion-maximal, by the choice of S'). It suffices us to show that any two vertices u and v of S' are connected by a red path in G . Moreover, it suffices us to show this only for u and v , that are both connected to S by an edge. Let $u' \in S$ be a blue neighbor of u , and let $v' \in S$ be a blue neighbor of v . Since S is a blue connected subgraph, there is a blue path $\mathcal{P}_1$ connecting u' and v' . Since G_1 is connected, u and v are connected by a path $\mathcal{P}_2$ in G_1 , which is vertex-disjoint from $\mathcal{P}_1$. The path $\mathcal{P}_2$ may consist of vertices of both colors, but there is no blue path from a vertex of $\mathcal{P}_2$ to $\mathcal{P}_1$ (by the inclusion-maximality of S). The paths $\mathcal{P}_1$ and $\mathcal{P}_2$ together form a cycle C in G . The graph G and the cycle C satisfy all conditions of Lemma 3.2, and so we have a red path connecting u and v . Hence S' indeed induces a connected subgraph.

Each component of $G \setminus S'$ is contained either in $G_1 \setminus S'$, or in $G \setminus G_1$. We have $|V(G_1) \setminus S'| < |V(G_1)|$ since $S' \neq \emptyset$, and $|V(G) \setminus V(G_1)| < \frac{1}{2}|V(G)|$. Thus the largest component of $G \setminus S'$ is strictly smaller than G_1 , and this contradicts the choice of S . $\square$

4. D_n^d as the graph of a simply connected complex. In order to apply Proposition 3.1 to some graph G , we need to know that G is the graph of a simply connected simplicial complex. This is obviously true for T_n^d , a graph of a triangulated grid, but for D_n^d , the grid with all diagonals, this has to be established. (One could say that this is a somewhat artificial-looking way of proving a purely combinatorial statement about D_n^d , but it is a convenient one.)

LEMMA 4.1. *For all $n, d \geq 2$ there exists a simply connected (2-dimensional) simplicial complex K such that D_n^d is (isomorphic to) the graph of K .*

Let us note that the simplicial complexes K constructed in the following proofs cannot be embedded in $\mathbf{R}^d$ (and they are probably hard to visualize).

First proof. (This proof was simplified using a suggestion of a referee.) First, let us fix a triangulation T of the grid, i.e., a geometric simplicial complex refining the subdivision of $[1, n]^d$ into unit cubes as in the definition of the graphs T_n^d . Let K_0 be the 2-skeleton of the triangulation (consisting of all simplices of this T_n^d of dimension at most 2). This K_0 is simply connected since T is simply connected (we use the fact that the fundamental group of a simplicial complex is determined by the 2-skeleton).

Let $E = E(D_n^d) \setminus E(K_0)$ be the set of all edges of D_n^d not present in K_0 , and let $e_1, e_2, \dots, e_t$ be an enumeration of E such that shorter edges precede longer ones; that is, $i < j$ implies that $\|e_i\|_1 \leq \|e_j\|_1$, where we write $\|e\|_1 = \sum_{k=1}^d |u_k - v_k|$ for $e = \{u, v\}$.

We are going to define a sequence $K_0, K_1, \dots, K_t$ of simplicial complexes inductively, and $K = K_t$ will be as in Lemma 4.1. The simplicial complex K_i is obtained from K_{i-1} by adding the edge e_i plus a triangle with e_i as one of the edges, while the remaining two edges of the triangle are already present in K_{i-1} . Obviously, K_i is homotopy equivalent to K_{i-1} and hence by induction we get that each K_i is simply connected. We also have $E(K_t) = E(D_n^d)$. It remains to check that the triangles can indeed be added as described.

Let us write $e_i = \{u, v\}$, and let $a = \min\{k \in [d] : u_k \neq v_k\}$. Let $x \in [n]^d$ be the vertex

$$x = (u_1, \dots, u_a, v_{a+1}, \dots, v_d).$$

Since all edges $e \in E(D_n^d)$ with $\|e\|_1 = 1$ are already present in K_0 , we have $\|e_i\|_1 \geq 2$ and $u \neq x \neq v$.

The edges $\{u, x\}$ and $\{x, v\}$ belong to $E(D_n^d)$ and have length strictly smaller than $\|e_i\|_1$, and hence they are present in K_{i-1} . The triangle added to K_{i-1} to form K_i is $\{u, x, v\}$. $\square$

Second proof. (This proof was suggested by a referee.) For each $u \in [n]^d$ let C_u be the axis-parallel unit cube centered at u , and let K be the nerve of the set system $\{C_u : u \in [n]^d\}$ (that is, a set $I \subseteq [n]^d$ forms a simplex if $\bigcap_{u \in I} C_u \neq \emptyset$). The 1-skeleton of K is exactly D_n^d . By the nerve theorem, K is homotopy equivalent to $\bigcup_{u \in [n]^d} C_u = [\frac{1}{2}, n + \frac{1}{2}]^d$, which is simply connected. $\square$

5. Isoperimetric inequalities. For a graph G and a set $A \subseteq V(G)$, the *vertex boundary* of A in G is the set

$$\text{vert-bd}_G(A) = \{v \in V(G) : v \notin A, \{u, v\} \in E(G) \text{ for some } u \in A\}.$$

An *isoperimetric inequality* for G (or more precisely, a vertex-isoperimetric inequality) bounds from below the quantity $\min\{|\text{vert-bd}_G(A)| : A \subseteq V(G), |A| = s\}$ as a function of s .

The ordinary grid. Let G_n^d be the graph of the ordinary grid (not triangulated). Explicitly, $V(G_n^d) = [n]^d$ and $u, v \in V(G_n^d)$ are connected by an edge if $\|u - v\|_1 = \sum_{i=1}^d |u_i - v_i| = 1$.

A vertex-isoperimetric inequality for G_n^d was established by Bollobás and Leader [BL91a]. To state the inequality, we define the *simplicial order* on $[n]^d$ by setting $x <_{\text{simpl}} y$ if either $\sum_{i=1}^d x_i < \sum_{i=1}^d y_i$, or if $\sum_{i=1}^d x_i = \sum_{i=1}^d y_i$ and x precedes y lexicographically; that is, for some $i \in [d]$ we have $x_j = y_j$ for $j = 1, 2, \dots, i-1$ and $x_i < y_i$.

THEOREM 5.1 ([BL91a]). *For every $A \subseteq [n]^d$ we have $|\text{vert-bd}_{G_n^d}(A)| \geq |\text{bd}_{G_n^d}(I_{\text{simpl}}(|A|))|$, where $I_{\text{simpl}}(s)$ denotes the set of the first s elements of $[n]^d$ in the ordering $<_{\text{simpl}}$.*

To bound from below the size of the vertex boundary of arbitrary set $A \subseteq [n]^d$ we use the following lemma.

LEMMA 5.2. *Let $A \subseteq [n]^d$ with $|A| = \beta n^d$. Then we have $|\text{vert-bd}_{G_n^d}(A)| \geq |L_n^d(i_\beta)| - |L_n^{d-1}(i_{1/2})|$.*

Proof. By Theorem 5.1 it suffices to prove the statement only for A of the form $I_{\text{simpl}}(|A|)$. Then A is a subset of the union of the first i_β diagonal layers and the first $i_\beta - 1$ layers are fully contained in A .

Let us write D for the intersection of A with $L_n^d(i_\beta)$. If $D = \emptyset$, then $\text{vert-bd}_{G_n^d}(A) = L_n^d(i_\beta)$ and we are done. If $D \neq \emptyset$, then $\text{vert-bd}_{G_n^d}(A)$ is a disjoint union of two sets B_1 and B_2 , where $B_1 \subseteq L_n^d(i_\beta)$ and $B_2 \subseteq L_n^d(i_\beta + 1)$. It is easy to check that $B_1 = L_n^d(i_\beta) \setminus D$, and if we show that $|B_2| \geq |D| - |L_n^{d-1}(i_{1/2})|$, then we are done.

Let D' be D shifted by 1 in the first component; that is, $D' = D + (1, 0, \dots, 0)$. Similarly, let L' be $L_n^d(i_\beta) + (1, 0, \dots, 0)$. Since $D' \subseteq L'$, we have $D' \setminus [n]^d \subseteq L' \setminus [n]^d$, and from the fact that $D' \cap [n]^d \subseteq B_2$ we arrive at $|B_2| \geq |D| - |L' \setminus [n]^d|$.

The rest follows from the fact that $L' \setminus [n]^d$ is an exact copy of the set $L_n^{d-1}(i_\beta - n)$, which always has a smaller size than the middle layer $L_n^{d-1}(i_{1/2})$, as will be shown in Lemma 5.3. $\square$

Let us remark that some error term in the previous estimate of $\text{vert-bd}_{G_n^d}(A)$ is needed. For example, $\text{vert-bd}_{G_n^d}(I_{\text{simpl}}(21)) = 5$ for the case $n = 3$, $d = 3$, but $|L_3^3(i_{21/27})| = 6$.

LEMMA 5.3. *For any $n, d \in \mathbf{N}$ the following hold:*

- (a) $|L_n^d(\frac{(n+1)d}{2} - i)| = |L_n^d(\frac{(n+1)d}{2} + i)|$ for every $i = 0, 1, 2, \dots$, and therefore, $i_{1/2} = \lceil \frac{(n+1)d}{2} \rceil$.
- (b) The function $i \mapsto |L_n^d(i)|$ is nondecreasing on integers $i \leq \lceil \frac{(n+1)d}{2} \rceil = i_{1/2}$ and nonincreasing on integers $i \geq \lfloor \frac{(n+1)d}{2} \rfloor$.

Proof. (a) The function $f((x_1, \dots, x_d)) = (n+1 - x_1, \dots, n+1 - x_d)$ maps $L_n^d(\frac{(n+1)d}{2} - i)$ to $L_n^d(\frac{(n+1)d}{2} + i)$ and it is easy to check from definition that $i_{1/2} = \lceil \frac{(n+1)d}{2} \rceil$.

(b) The proof is by induction on d . For $d = 1, 2$ and any $n \geq 1$ the statement holds trivially. Fix some n and $d > 2$. We can partition the set $L_n^d(i)$ into n sets $L_1, \dots, L_n$ according to the last component of the vertices; that is

$$L_t = \{x \in L_n^d(i) : x_d = t\} \quad \text{for all } t \in [n].$$

It is easy to check that each L_t is isomorphic to $L_n^{d-1}(i - t)$ and thus

$$\begin{aligned} |L_n^d(i+1)| - |L_n^d(i)| &= \sum_{t=1}^n |L_n^{d-1}(i+1-t)| - \sum_{t=1}^n |L_n^{d-1}(i-t)| \\ &= |L_n^{d-1}(i)| - |L_n^{d-1}(i-n)|. \end{aligned}$$

We distinguish four cases:

1. If $i \leq \lceil \frac{(n+1)(d-1)}{2} \rceil$, then by the inductive hypothesis we have $|L_n^{d-1}(i-n)| \leq |L_n^{d-1}(i)|$.
2. If $\lceil \frac{(n+1)(d-1)}{2} \rceil < i < \lceil \frac{(n+1)d}{2} \rceil$, then by (a) we have

$$|L_n^{d-1}(i)| = |L_n^{d-1}((n+1)(d-1) - i)|.$$

Since

$$(i-n) \leq (n+1)(d-1) - i \leq \lceil \frac{(n+1)(d-1)}{2} \rceil,$$

we have by the inductive hypothesis that

$$|L_n^{d-1}(i-n)| \leq |L_n^{d-1}((n+1)(d-1) - i)| = |L_n^{d-1}(i)|.$$

3. If $\lfloor \frac{(n+1)d}{2} \rfloor \leq i < \lfloor \frac{(n+1)(d+1)}{2} + n \rfloor$, then by (a) we have

$$|L_n^{d-1}(i-n)| = |L_n^{d-1}((n+1)(d-1) - (i-n))|.$$

Since

$$i \geq (n+1)(d-1) - (i-n) \geq \lfloor \frac{(n+1)(d-1)}{2} \rfloor,$$

we have by the inductive hypothesis that

$$|L_n^{d-1}(i)| \leq |L_n^{d-1}((n+1)(d-1) - (i-n))| = |L_n^{d-1}(i-n)|.$$

4. If $i \geq \lfloor \frac{(n+1)(d+1)}{2} + n \rfloor$, then by the inductive hypothesis we have $|L_n^{d-1}(i-n)| \geq |L_n^{d-1}(i)|$. $\square$

PROPOSITION 5.4. For every $\alpha \in (0, \frac{1}{2})$ and every $A \subseteq [n]^d$ with $\alpha n^d \leq |A| \leq (1-\alpha)n^d$ we have $|\text{vert-bd}_{G_n^d}(A)| \geq |L_n^d(i_{1-\alpha})| - |L_n^{d-1}(i_{1/2})|$.

Proof. Let $|A| = \beta n^d$. By Lemma 5.2 we have $|\text{vert-bd}_{G_n^d}(A)| \geq |L_n^d(i_\beta)| - |L_n^{d-1}(i_{1/2})|$. If $\alpha \leq \beta \leq \frac{1}{2}$, then also $i_\alpha \leq i_\beta \leq i_{1/2}$, and by Lemma 5.3 we have $|L_n^d(i_\beta)| \geq |L_n^d(i_\alpha)|$. Similarly, for $\frac{1}{2} < \beta \leq (1-\alpha)$ we have $|L_n^d(i_\beta)| \geq |L_n^d(i_{1-\alpha})|$. Since $i_{1/2} \leq (n+1)d - i_\alpha \leq i_{(1-\alpha)}$, by Lemma 5.3 we have $|L_n^d(i_\alpha)| = |L_n^d((n+1)d - i_\alpha)| \geq |L_n^d(i_{(1-\alpha)})|$, and thus $|\text{vert-bd}_{G_n^d}(A)| \geq |L_n^d(i_{(1-\alpha)})| - |L_n^{d-1}(i_{1/2})|$. $\square$

The grid with diagonals. Here we derive the following vertex-isoperimetric inequality.

PROPOSITION 5.5. For any set $A \subseteq [n]^d$ with $\frac{1}{4}n^d \leq |A| \leq \frac{3}{4}n^d$ we have $|\text{vert-bd}_{D_n^d}(A)| \geq n^{d-1} - d^2 n^{d-2}$.

We are going to derive this result from an edge-isoperimetric inequality for the ordinary grid G_n^d . For $A \subseteq [n]^d$, let

$$\text{edge-bd}_{G_n^d}(A) = \{\{u, v\} \in E(G_n^d) : u \in A, v \notin A\}$$

be the edge boundary of A in G_n^d .

THEOREM 5.6 (see Bollobás and Leader [BL91b]). Let $A \subseteq [n]^d$. Then

$$|\text{edge-bd}_{G_n^d}(A)| \geq \begin{cases} 4|A|/n & \text{for } |A| < n^d/4, \\ n^{d-1} & \text{for } n^d/4 \leq |A| \leq 3n^d/4, \\ 4(n^d - |A|)/n & \text{for } |A| > 3n^d/4. \end{cases}$$

First we show, following [BL91a] almost verbatim, that it suffices to prove Proposition 5.5 for sets A that are *down-sets*. We begin with the necessary definitions from [BL91a].

A set $A \subseteq [n]^d$ is a *down-set* if $x \in A$ implies $y \in A$ for all $y \in [n]^d$ satisfying $y_i \leq x_i$ for all $i \in [d]$. For $A \subseteq [n]^d$ and $1 \leq i \leq d$, let the *i-section* of A at $x \in [n]^{d-1}$ be defined as

$$A^i(x) = \{t \in [n] : (x_1, \dots, x_{i-1}, t, x_i, \dots, x_{d-1}) \in A\}.$$

For $A \subseteq [n]^d$ and $1 \leq i \leq d$, let the *i-compression* of A be defined by giving its *i-sections*:

$$(C_i(A))^i(x) = \{1, \dots, |A^i(x)|\} \quad \text{for all } x \in [n]^{d-1}.$$

In other words, C_i “compresses” each i -section of A downwards. We note that $|C_i(A)| = |A|$, and it is also easy to check that if a set A is i -compressed, then so is $C_j(A)$. So the set $A' = C_n(C_{n-1}(\dots C_1(A)\dots))$ is i -compressed for every $i \in [d]$, and thus it is a down-set.

LEMMA 5.7. *Let $A \subseteq [n]^d$. Then there is a down-set $A' \subseteq [n]^d$ with $|A'| = |A|$ and $|\text{vert-bd}_{D_n^d}(A')| \leq |\text{vert-bd}_{D_n^d}(A)|$.*

Proof. For $A \subseteq [n]^d$, let $N(A) = A \cup \text{vert-bd}_{D_n^d}(A)$ denote the *closed neighborhood* of A . By the above remarks it suffices to prove that $|N(C_i(A))| \leq |N(A)|$ for all $A \subseteq [n]^d$ and all $i \in [d]$.

Let us write B for $C_i(A)$. The case $d = 1$ is easy; we have $|N(A)| \geq |A| + 1 = |N(B)|$ for all $A \subseteq [n]$ apart from $A = \emptyset$ and $A = [n]$. For $A = \emptyset$ and $A = [n]$ we have $|N(A)| = |A| = |B| = |N(B)|$.

For $d > 1$ it is sufficient to show that for each $x \in [n]^{d-1}$ we have

$$|(N(B))^i(x)| \leq |(N(A))^i(x)|.$$

Fixing an arbitrary $x \in [n]^{d-1}$, we have

$$(N(A))^i(x) = \bigcup_{y \in [n]^{d-1} : \|y-x\|_\infty \leq 1} N(A^i(y))$$

and

$$(N(B))^i(x) = \bigcup_{y \in [n]^{d-1} : \|y-x\|_\infty \leq 1} N(B^i(y)).$$

The sets $N(B^i(y))$ are initial segments of $[n]$, hence they are nested, and thus

$$|(N(B))^i(x)| \leq \max_{y \in [n]^{d-1} : \|y-x\|_\infty \leq 1} |N(B^i(y))|.$$

From the 1-dimensional case we know

$$|N(B^i(y))| \leq |N(A^i(y))|,$$

and therefore,

$$|(N(B))^i(x)| \leq |(N(A))^i(x)|. \quad \square$$

Proposition 5.5 now follows from Theorem 5.6, the previous lemma, and the next lemma.

LEMMA 5.8. *Let $A \subseteq [n]^d$ be a down-set. Then $|\text{vert-bd}_{D_n^d}(A)| \geq |\text{edge-bd}_{G_n^d}(A)| - d^2 n^{d-2}$.*

To prove this lemma we need the following claim.

CLAIM 5.9. *Let $A \subseteq [n]^d$ be a down-set. Then $|\text{edge-bd}_{G_n^d}(A)| \leq dn^{d-1}$.*

Proof. Since A is a down-set, for each $k \in [d]$ and $x \in [n]^{d-1}$ there is at most one $t \in [n-1]$ such that $\{u, v\} \in \text{edge-bd}_{G_n^d}(A)$, where $u = (x_1, \dots, x_{k-1}, t, x_{k+1}, \dots, x_d)$ and $v = (x_1, \dots, x_{k-1}, t+1, x_{k+1}, \dots, x_d)$. Hence there can be at most dn^{d-1} edges in $\text{edge-bd}_{G_n^d}(A)$. $\square$

Proof of Lemma 5.8. Let

$$E = \text{edge-bd}_{G_n^d}(A) \setminus \{\{u, v\} \in E(G_n^d) : u_i = v_i = n \text{ for some } i\}.$$

We show that E can be injectively mapped into $\text{vert-bd}_{D_n^d}(A)$. Let us construct a mapping $f: E \rightarrow [n]^d$ in the following way: For arbitrary edge $\{u, v\} \in E$, let i be the number of the component in which u and v differ by 1, and without loss of generality, let $u_i \leq v_i$. Then we put

$$f(\{u, v\}) = u + \underbrace{(1, 1, \dots, 1, 0, \dots, 0)}_{i \text{ times } 1}.$$

It is easy to check that $u_j < n$ for $j \leq i$, and hence $z = f(\{u, v\}) \in [n]^d$ indeed. We have $u \in A$ and $v \notin A$. If we had $z \in A$, then since A is a down-set and $v_i \leq z_i$ for all $i \in [d]$, it would follow that $v \in A$ as well. Thus $z \notin A$, and since $\|u - z\|_\infty = 1$, we arrive at $f(\{u, v\}) \in \text{vert-bd}_{D_n^d}(A)$.

Now we suppose for contradiction that there exist two edges $\{u_1, v_1\}, \{u_2, v_2\} \in E$ that are mapped by f to the same vertex z . Let i_1 be the number of the components where u_1 and v_1 differ by 1, and let i_2 be the number of the component where u_2 and v_2 do so. Without loss of generality, let $i_1 < i_2$. Then

$$v_2 = (z_1 - 1, \dots, z_{i_2-1} - 1, z_{i_2}, z_{i_2+1}, \dots, z_d)$$

is in all components less than or equal to

$$u_1 = (z_1 - 1, \dots, z_{i_1-1} - 1, z_{i_1} - 1, z_{i_1+1}, \dots, z_d).$$

Since A is a down-set, $u_1 \in A$ implies $v_2 \in A$. This contradiction shows that f is injective as claimed.

It remains to estimate the size of E . For $i \in [d]$ let us put $F_i = \{\{u, v\} \in \text{edge-bd}_{G_n^d}(A) : u_i = v_i = n\}$. It is easy to see that

$$|F_i| = |\text{edge-bd}_{G_n^{d-1}}(A_i)|,$$

where

$$A_i = \{(x_1, \dots, x_{d-1}) \in [n]^{d-1} : (x_1, \dots, x_{i-1}, n, x_i, \dots, x_{d-1}) \in A\}.$$

Since A_i is a down-set, Claim 5.9 gives $|F_i| \leq (d-1)n^{d-2}$, and thus

$$|E| \geq |\text{edge-bd}_{G_n^d}(A)| - \sum_{i=1}^d |F_i| \geq |\text{edge-bd}_{G_n^d}(A)| - d(d-1)n^{d-2},$$

which is even slightly better than the bound claimed in the lemma. $\square$

6. Proofs of Theorems 1.1 and 1.2. *Monochromatic connected separators again.* We will need the following standard consequence of Proposition 3.1.

COROLLARY 6.1. *Let G be the graph of a simply connected simplicial complex, and let $V(G)$ be colored red and blue. Then there exists a partition of $V(G)$ into three disjoint sets A , B , and S such that*

- *no edge connects a vertex of A to a vertex of B ,*
- *S is a monochromatic connected subgraph of G , and*
- *$|A|, |B| \leq \frac{2}{3}|V(G)|$.*

Proof. Let S be as in Proposition 3.1, and let $V_1, V_2, \dots, V_m$ be the vertex sets of the components of $G \setminus S$ ordered by size in the descending order. Let i be the largest index with $|V_1 \cup V_2 \cup \dots \cup V_i| \leq \frac{2}{3}|V(G)|$. Then $|V_1| + \dots + |V_i| \leq \frac{2}{3}|V(G)|$

and $|V_{i+2}| + \dots + |V_m| \leq \frac{1}{3}|V(G)|$. Let us put $A = V_1 \cup V_2 \cup \dots \cup V_i$ and $B = V \setminus A \setminus S$. Obviously, $|A| \leq \frac{2}{3}|V(G)|$, and since $|V(G)| - |B| \geq |A| \geq |V_{i+1}|$ and $|B| = |V_{i+1}| + |V_{i+2}| + \dots + |V_m| \leq |V_{i+1}| + \frac{1}{3}|V(G)|$, we arrive at $|B| \leq \frac{2}{3}|V(G)|$ too. $\square$

Proof of Theorem 1.1. We consider a two-colored graph T_n^d as in the theorem and a partition $V(T_n^d) = A \cup B \cup S$ as in Corollary 6.1 (S is a monochromatic connected separator). We may assume $|S| \leq \frac{1}{3}n^d$, for otherwise we would be done. Then $|A|$ or $|B|$ is between $\frac{1}{3}n^d$ and $\frac{2}{3}n^d$; let us fix the notation A, B so that $\frac{1}{3}n^d \leq |A| \leq \frac{2}{3}n^d$.

We now consider the grid graph G_n^d as a subgraph of T_n^d . We have $S \supseteq \text{ver-bd}_{T_n^d}(A) \supseteq \text{ver-bd}_{G_n^d}(A)$, and $|\text{ver-bd}_{G_n^d}(A)| \geq |L_n^d(i_{2/3})| - |L_n^{d-1}(i_{1/2})|$ by Proposition 5.4. The required estimates for $|L_n^d(i_{2/3})|$ and $|L_n^{d-1}(i_{1/2})|$ are established in the appendix. Theorem 1.1 is proved. $\square$

Proof of Theorem 1.2. We consider a two-coloring of the graph D_n^d . Lemma 4.1 tells us that D_n^d is the graph of a simply connected simplicial complex, and so using Corollary 6.1 we again find a partition $V(D_n^d) = A \cup B \cup S$. We may assume $|S| \leq \frac{1}{3}n^d$, for otherwise we would be done. We can again assume $\frac{1}{3}n^d \leq |A| \leq \frac{2}{3}n^d$, and using Proposition 5.5 we get $|S| \geq |\text{ver-bd}_{D_n^d}(A)| \geq n^{d-1} - d^2n^{d-2}$. Theorem 1.2 is proved. $\square$

7. An upper bound for certain triangulated grids. In this section we show that there exists a triangulated grid T_n^d in which every edge connects either two vertices from the same diagonal layer $L_n^d(i)$ or vertices from two consecutive diagonal layers $L_n^d(i)$ and $L_n^d(i+1)$. Thus, each monochromatic connected component in the diagonal-layer coloring of such T_n^d is contained in some $L_n^d(i)$.

To see that such triangulation exists, we consider the arrangement $\mathcal{A}$ of the following hyperplanes: $H_j^i = \{x : x_i = j\}$ for $j \in [n]$, $i = 0, \dots, d$, and $S_j = \{x : \sum x_k = j\}$ for $j = d, \dots, nd$. The vertices of this arrangement are of two types: the vertices of the first type are formed as the intersection of d hyperplanes among H_j^i , and the vertices of the second type are formed as the intersection of one hyperplane S_j and $d-1$ of the hyperplanes H_j^i .

The vertices of the first type are the grid points of $[n]^d$. An intersection of $d-1$ hyperplanes H_j^i is either empty or a line parallel to one of the coordinate axes, and if we intersect this line by some hyperplane S_j we get a vertex with integral coordinates. Thus, if we restrict ourselves to the hypercube $[1, n]^d$, the only possible vertices of our arrangement are the grid points of $[n]^d$.

The arrangement $\mathcal{A}$ is a polyhedral complex. Let $\mathcal{C}$ be the subcomplex of $\mathcal{A}$ consisting of all faces contained in $[1, n]^d$. The union of all faces of $\mathcal{C}$ is exactly $[1, n]^d$, $\mathcal{C}$ refines the subdivision of $[1, n]^d$ into unit cubes, and it remains to refine $\mathcal{C}$ to a triangulation. This can be done, for example, using the *bottom-vertex triangulation* (or *canonical triangulation*) of Clarkson [Cla88].

For each face F of $\mathcal{C}$, which is a convex polytope, we define the *bottom vertex* v as the vertex of F with the lexicographically smallest coordinate vector.

The bottom-vertex triangulation of $\mathcal{C}$ is defined inductively. For $k \geq 1$, we assume that all faces of $\mathcal{C}$ dimension at most k have already been triangulated (for $k = 1$ this is satisfied automatically, since the 1-dimensional faces are segments). Given a $(k+1)$ -dimensional face F , we let v be the bottom vertex of F , and we triangulate F as follows. For each simplex σ from the union of the triangulations of all proper faces of F , we consider the cone over σ with apex v , that is, the simplex $\text{conv}(\sigma \cup \{v\})$. Taking all of these simplices plus all simplices in the triangulation of all proper faces of F yields the bottom-vertex triangulation of F .

It is not hard to verify that this indeed defines a simplicial complex refining $\mathcal{C}$, and thus the desired triangulated grid T_n^d .

8. Open problems. We have considered colorings of the vertices of $[n]^d$ by two colors. The case of d colors is essentially solved by the d -dimensional HEX lemma. A natural and, in our opinion, very interesting question is, What happens for colorings by k colors for $3 \leq k \leq d-1$? An obvious conjecture is that the minimum possible size of the largest monochromatic connected subgraph should be about n^{d+1-k} in this case (for d and k fixed and $n \rightarrow \infty$), but at present we have nothing nontrivial to say about this question.

A natural bound one would expect in Theorem 1.1 (for a triangulated grid) would be the size of the “middle diagonal layer” $L_n^d(i_{1/2})$. However, our method yields only $|L_n^d(i_{2/3})|$. The improved bound of $|L_n^d(i_{1/2})|$ would follow immediately by our method from the following conjecture: *If $A \subseteq [n]^d$ is such that removing it from the grid G_n^d leaves no connected component larger than $\frac{1}{2}n^d$, then $|A| \geq |L_n^d(i_{1/2})|$.* This does not seem to follow from the vertex-isoperimetric inequality for G_n^d , since, in principle, for example, we might have a situation where $G_n^d \setminus A$ has three connected components of size about $\frac{1}{3}n^d$ each, and A is a common boundary of all three. While this or similar situations seem unlikely, we currently do not have a proof of the conjecture. We note that Kleitman [Kle86] proved a similar conjecture for the hypercube (that is, for the grid G_2^d), but adapting his method to larger grids seems nontrivial.

In the appendix below, we derive asymptotic bounds for the size of the diagonal layers $|L_n^d(i_{1/2})|$ and $|L_n^d(i_{2/3})|$, using a quantitative version of the central limit theorem. However, the estimated quantity is of a quite basic nature and it might well be that considerably more precise results are known.

Appendix. Bounding the size of diagonal layers. Let us recall that $L_n^d(i) = \{x \in [n]^d : \sum_{j=1}^d x_j = i\}$ and $i_\alpha = \min\{i : \sum_{j \leq i} |L_n^d(j)| > \alpha n^d\}$. Here we prove the following estimates.

PROPOSITION A.1. *There exist constants $C \geq c > 0$ such that for all $n > 1$ and all d we have*

$$\frac{c}{\sqrt{d}} n^{d-1} \leq |L_n^d(i_{2/3})| \leq |L_n^d(i_{1/2})| \leq \frac{C}{\sqrt{d}} n^{d-1}.$$

We use the following quantitative version of the central limit theorem by Feller [Fel43].

THEOREM A.2. *Let $X_1, \dots, X_d$ be independent random variables and suppose that their second moments $\sigma_1^2, \dots, \sigma_d^2$ exist. Let $S_i = X_1 + \dots + X_i$ and $s_i^2 = \sigma_1^2 + \dots + \sigma_i^2$ for $i = 1, \dots, d$, let $F_d(t) = \text{Prob}[S_d \leq t]$ be the distribution function of S_d , and let $\Phi(u)$ be the distribution function of the standard normal distribution. Let there exist a nonincreasing sequence $\lambda_1, \dots, \lambda_d$, such that $|X_k| \leq \lambda_k s_k$ for $k = 1, \dots, d$.*

If $0 < \lambda_d u < (3 - \sqrt{5})/4$, then

$$1 - F_d(us_d) = e^{-(1/2)u^2 Q_d(u)} \left((1 - \Phi(u)) + \theta \lambda_d e^{-(1/2)u^2} \right).$$

Here θ has an upper bound independent of the sequence $\{X_k\}$ and of d ; furthermore,

$$Q_d(u) = \sum_{\nu=1}^{\infty} q_{d,\nu} u^\nu,$$

where coefficient $q_{d,\nu}$ depends only on the ν first moments of $X_1, \dots, X_d$.

If, more particularly,

$$0 < \lambda_d u < 1/12,$$

then

$$|\theta| < 9$$

and

$$|q_{d,\nu}| < (1/7)(12\lambda_d)^\nu.$$

Moreover, for any $0 < i < j \leq d$ we have

$$|Q_j(u) - Q_i(u)| \leq (1/2)(s_j^2 - s_i^2)/s_n^2.$$

Let us assume that n is odd, $n = 2k + 1$; the case of even n is similar. Let us consider the sequence of independent random variables $X_1, \dots, X_d$, where each X_i is uniformly distributed on the set of integers $\{-k, -k + 1, \dots, k\}$. For this sequence $X_1, \dots, X_d$ we define S_d, F_d, s_d as in Feller's theorem. It is easy to see that $n^d F_d(t - \frac{nd+d}{2}) = \sum_{i=0}^t |L_n^d(i)|$ for all $t \in \mathbf{N}$.

In our case $\frac{\sqrt{dk}}{3} \leq s_d \leq \sqrt{dk}$, so we can put $\lambda_d = \frac{3}{\sqrt{d}}$. For $0 \leq u \leq 1$ and $d > 10^6$ it is seen that $\lambda_d u < \frac{1}{300}$, and thus

$$\begin{aligned} |Q_d(u)| &\leq \sum_{\nu=1}^{\infty} |q_{d,\nu} u^\nu| \leq \sum_{\nu=1}^{\infty} |q_{d,\nu}| \leq \sum_{\nu=1}^{\infty} \frac{1}{7} (12\lambda_d)^\nu \\ &\leq \frac{12\lambda_d}{7} \sum_{\nu=0}^{\infty} (12\lambda_d)^\nu \leq \frac{36}{7\sqrt{d}} \sum_{\nu=0}^{\infty} \left(\frac{12}{300}\right)^\nu \leq \frac{6}{\sqrt{d}}. \end{aligned}$$

Therefore we arrive at

$$\begin{aligned} e^{-3d^{-1/2}} &\leq e^{-\frac{1}{2}u^2 Q_d(u)} \leq e^{3d^{-1/2}} \\ 1 - 3d^{-1/2} &\leq e^{-\frac{1}{2}u^2 Q_d(u)} \leq 1 + 4d^{-1/2}. \end{aligned}$$

From

$$1 - F_d(us_d) = e^{-(1/2)u^2 Q_d(u)} \left((1 - \Phi(u)) + \theta \lambda_d e^{-(1/2)u^2} \right)$$

we derive

$$\begin{aligned} 1 - F_d(us_d) &\leq \left(1 + 4d^{-1/2}\right) \left(1 - \Phi(u) + 27d^{-1/2}\right) \\ F_d(us_d) &\geq \Phi(u) - 32d^{-1/2} \end{aligned}$$

and also

$$\begin{aligned} 1 - F_d(us_d) &\geq \left(1 - 3d^{-1/2}\right) \left(1 - \Phi(u) - 27d^{-1/2}\right) \\ F_d(us_d) &\leq \Phi(u) + 32d^{-1/2}. \end{aligned}$$

Let $q_1 = \Phi^{-1}(0.7)$ and $q_2 = \Phi^{-1}(0.8)$; we have $0 < q_1, q_2 < 1$. Then the i -layers for $i \in [q_1 s_d + \frac{nd+d}{2}, q_2 s_d + \frac{nd+d}{2}]$ contain together at least $\frac{1}{30}n^d$ vertices, since

$$\begin{aligned} F_d(q_2 s_d) - F_d(q_1 s_d) &\geq \Phi(q_2) - 32d^{-1/2} - \Phi(q_1) - 32d^{-1/2} \\ &= 1/10 - 64d^{-1/2} \geq 1/30. \end{aligned}$$

There are at most $(q_2 s_d - q_1 s_d) \leq kd^{1/2}$ such layers, so the largest of them, $L_n^d(q_1 s_d + \frac{nd+d}{2})$, contains at least $\frac{1}{15d^{1/2}}n^{d-1}$ vertices. Now the i th layer for $i = q_1 s_d + \frac{nd+d}{2}$ cuts off a set of size at least $\frac{2}{3}n^d$ from the grid, since

$$\begin{aligned} F_d(q_1 s_d) &\geq \Phi(q_1) - 32d^{-1/2} \\ &= 7/10 - 32d^{-1/2} \geq 0.668 > 2/3. \end{aligned}$$

Thus, $|L_n^d(i_{2/3})| > |L_n^d(q_1 s_d + \frac{nd+d}{2})|$ and, therefore, $|L_n^d(i_{2/3})| \geq \frac{1}{15d^{1/2}}n^{d-1}$. Let us remind that during all the previous steps we still assume that $d > 10^6$.

In the following lemma we show the upper bound on the size of the middle layer. To prove this, we need to show that two consecutive layers cannot differ too much in size.

LEMMA A.3. *For any $n, d, i \in \mathbf{N}$ the following hold:*

- (a) $|L_n^d(i_{1/2})| \leq 1000 \frac{n^{d-1}}{\sqrt{d}}$,
- (b) $|L_n^d(i)| - |L_n^d(i+1)| \leq |L_n^{d-1}(i-n)| \leq 2000 \frac{n^{d-2}}{\sqrt{d}}$.

Proof. We prove both parts together by induction on d . First, we show that $|L_n^d(i)| - |L_n^d(i+1)| \leq |L_n^{d-1}(i-n)| \leq |L_n^{d-1}(i_{1/2})|$. Let L' be $L_n^d(i)$ shifted by 1 in the first component, i.e., $L' = L_n^d(i) + (1, 0, \dots, 0)$. Since $(L' \cap [n]^d) \subseteq L_n^d(i+1)$, we arrive at $|L_n^d(i)| - |L_n^d(i+1)| \leq |L' \setminus [n]^d|$. It is easy to check that $L' \setminus [n]^d$ is an exact copy of the set $L_n^{d-1}(i-n)$, which always has a smaller size than the middle layer $L_n^{d-1}(i_{1/2})$. The rest of (b) follows from (a).

To prove (a) for $d \leq 10^6$, we note that the size of any layer $L_n^d(i)$ is bounded above by n^{d-1} and for $d \leq 10^6$ we have $n^{d-1} \leq 1000 \frac{n^{d-1}}{\sqrt{d}}$. To prove (a) for $d > 10^6$, we estimate the size of the $\frac{n}{6}$ consecutive layers $L_n^d(i)$ for $i \in [i_{1/2}, i_{1/2} + \frac{n}{6}]$:

$$\begin{aligned} F_n^d\left(i_{1/2} + \frac{n}{6}\right) - F_n^d\left(i_{1/2}\right) &\leq \left(\Phi\left(\frac{n}{6s_d}\right) + \frac{32}{\sqrt{d}}\right) - \left(\Phi(0) - \frac{32}{\sqrt{d}}\right) \\ &\leq \Phi\left(\frac{1}{\sqrt{d}}\right) - \Phi(0) + \frac{64}{\sqrt{d}} \\ &\leq \frac{1}{\sqrt{2\pi}} \left(\int_0^{d^{-1/2}} e^{-(1/2)y^2} dy \right) + \frac{64}{\sqrt{d}} \\ &\leq 65d^{-1/2}. \end{aligned}$$

Among the $n/6$ layers the size of the smallest layer $L_n^d(i_{1/2} + \frac{n}{6})$ is at most

$$\left| L_n^d\left(i_{1/2} + \frac{n}{6}\right) \right| \leq \frac{65d^{-1/2}n^d}{n/6} \leq 400 \frac{n^{d-1}}{\sqrt{d}}.$$

Using (b) we estimate the size of the middle layer as

$$\begin{aligned} |L_n^d(i_{1/2})| &\leq \left| L_n^d\left(i_{1/2} + \frac{n}{6}\right) \right| + \frac{n}{6} \left(2000 \frac{n^{d-2}}{\sqrt{d}} \right) \\ &\leq 1000 \frac{n^{d-1}}{\sqrt{d}}. \quad \square \end{aligned}$$

Acknowledgments. We would like to thank Nati Linial for discussions that were the starting point of our research reported in this paper, and Imre Leader for useful advice concerning isoperimetric inequalities. We also thank two anonymous referees for many useful comments and suggestions.

REFERENCES

- [ADO+03] N. ALON, G. DING, B. OPOROWSKI, AND D. VERTIGAN, *Partitioning into graphs with only small components*, J. Combin. Theory Ser. B, 87 (2003), pp. 231–243.
- [BL91a] B. BOLLOBÁS AND I. LEADER, *Compressions and isoperimetric inequalities*, J. Combin. Theory Ser. A, 56 (1991), pp. 47–62.
- [BL91b] B. BOLLOBÁS AND I. LEADER, *Edge-isoperimetric inequalities in the grid*, Combinatorica, 11 (1991), pp. 299–314.
- [BS05] R. BERKE AND T. SZABÓ, *Relaxed two-coloring of cubic graphs*, J. Combin. Theory Ser. B, 97 (2007), pp. 652–668.
- [Cla88] K. L. CLARKSON, *A randomized algorithm for closest-point queries*, SIAM J. Comput., 17 (1988), pp. 830–847.
- [DOS+96] G. DING, B. OPOROWSKI, D. SANDERS, AND D. VERTIGAN, *Partitioning into Two Graphs with Only Small Components*, Louisiana State University, Baton Rouge, LA, 1996; available online at <http://citeseer.ist.psu.edu/205996.html>.
- [Fel43] W. FELLER, *Generalization of a probability limit theorem of Cramér*, Trans. Amer. Math. Soc., 54 (1943), pp. 361–372.
- [G79] D. GALE, *The game of Hex and the Brouwer fixed-point theorem*, Amer. Math. Monthly, 86 (1979), pp. 818–827.
- [Hat01] A. HATCHER, *Algebraic Topology*, Cambridge University Press, Cambridge, UK, 2001; also available online at <http://math.cornell.edu/~hatcher#AT1>.
- [HST03] P. HAXELL, T. SZABÓ, AND G. TARDOS, *Bounded size components—partitions and transversals*, J. Combin. Theory Ser. B, 88 (2003), pp. 281–297.
- [JW96] B. JACKSON AND N. C. WORMALD, *On the linear k -arboricity of cubic graphs*, Discrete Math, 162 (1996), pp. 293–297.
- [Kle86] D. KLEITMAN, *On a problem of Yuzvinsky on separating the n -cube*, Discrete Math., 60 (1986), pp. 207–213.
- [KMR+97] J. KLEINBERG, R. MOTWANI, P. RAGHAVAN, AND S. VENKATASUBRAMANIAN, *Storage management for evolving databases*, in Proceedings of the 38th IEEE Symposium on Foundations of Computer Science, Miami Beach, FL, 1997.
- [LMST07] N. LINIAL, J. MATOUŠEK, O. SHEFFET, AND G. TARDOS, *Graph coloring with no large monochromatic components*, submitted; extended abstract in Proc. EUROCOMB 2007, Electron. Notes Discrete Math., 29C (2007), pp. 115–122.
- [LS93] N. LINIAL AND M. SAKS, *Low diameter graph decompositions*, Combinatorica, 13 (1993), pp. 441–454.
- [MN98] J. MATOUŠEK AND J. NEŠETŘIL, *Invitation to Discrete Mathematics*, Oxford University Press, Oxford, UK, 1998.
- [Mun84] J. R. MUNKRES, *Elements of Algebraic Topology*, Addison-Wesley, Menlo Park, CA, 1984.
- [Tho99] C. THOMASSEN, *Two-coloring the edges of a cubic graph such that each monochromatic component is a path of length at most 5*, J. Combin. Theory Ser. B, 75 (1999), pp. 100–109.

MAXIMIZING SUPERMODULAR FUNCTIONS ON PRODUCT LATTICES, WITH APPLICATION TO MAXIMUM CONSTRAINT SATISFACTION*

ANDREI KROKHIN[†] AND BENOIT LAROSE[‡]

Abstract. Recently, a strong link has been discovered between supermodularity on lattices and tractability of optimization problems known as maximum constraint satisfaction problems. This paper strengthens this link. We study the problem of maximizing a supermodular function which is defined on a product of n copies of a fixed finite lattice and given by an oracle. We exhibit a large class of finite lattices for which this problem can be solved in oracle-polynomial time in n . We also obtain new large classes of tractable maximum constraint satisfaction problems.

Key words. supermodular function, lattices, optimization, tractability, constraint satisfaction

AMS subject classifications. 90C27, 68Q25, 68T20

DOI. 10.1137/060669565

1. Introduction. Sub- and supermodular set functions are special real-valued functions defined on the powerset of a set. They are well studied in combinatorics (see bibliographical survey [10]) and have numerous applications in combinatorial optimization [9, 11] and elsewhere (see, e.g., [24, 28]). Minimization of submodular set functions is one of the most well-known tractable problems in combinatorial optimization [11, 27]. Examples of other combinatorial problems that can be solved by using submodular set function minimization include the minimum s - t cut problem in networks and finding the largest common independent set in two matroids. The submodular set function minimization problem was also considered for various relaxations of submodular functions such as intersecting or crossing submodular functions. Such relaxations are either defined on some family of subsets of a set, or use a slightly modified form of submodularity [27].

A more general form of sub- and supermodularity is the one where functions are defined on general (i.e., algebraic) lattices [28]. This form of supermodularity is very popular in financial and actuarial mathematics (see, e.g., supermodular games [28] and supermodular order on multivariate distributions [23, 25]), but it is relatively little studied in combinatorics and optimization (see, e.g., section 60.3a of [27]), except for the special case when the lattice is a chain, i.e., a totally ordered set. Sub- and supermodular functions on finite chains can be alternatively represented by matrices and arrays which are called Monge and anti-Monge, respectively, matrices and arrays. Such matrices and arrays are used to identify tractable cases of hard optimization problems such as TSP [1].

*Received by the editors September 11, 2006; accepted for publication (in revised form) November 4, 2007; published electronically February 27, 2008. A preliminary version of some parts of this paper appears in Proceedings of CP'05, Sitges, Spain 2005.

<http://www.siam.org/journals/sidma/22-1/66956.html>

[†]Department of Computer Science, Durham University, Durham, DH1 3LE, UK (andrei.krokhin@durham.ac.uk). This author's research is partially supported by the UK EPSRC grant EP/C543831/1.

[‡]Department of Mathematics and Statistics, Concordia University, Montréal H3G 1M8, QC, Canada (larose@mathstat.concordia.ca). This author's research is partially supported by grants from NSERC, FQRNT, and CRM.

Recently, the general form of supermodularity (on lattices) has been applied to classify the complexity of maximum constraint satisfaction problems which are very actively studied optimization problems in artificial intelligence [6] and complexity theory [5]. In a maximum constraint satisfaction problem, informally speaking, one is given a finite collection of constraints on overlapping sets of variables, and the goal is to find an assignment of values to the variables with a maximum number (or total weight) of satisfied constraints. Some recent papers [3, 8, 19] discovered that there is a strong link between supermodularity on lattices and tractability of maximum constraint satisfaction problems with a restricted set of allowed constraints, and developing this connection is the main aim of this paper.

2. Sub- and supermodularity, and lattices. In this section we describe the problem of maximizing a supermodular function on a product lattice.

DEFINITION 2.1. *Let A be finite set. A function $f : 2^A \rightarrow \mathbb{R}$ is called a supermodular set function (on A) if the inequality*

$$f(X) + f(Y) \leq f(X \cap Y) + f(X \cup Y)$$

holds for all $X, Y \subseteq A$, and it is called submodular if the inverse inequality holds for all X, Y .

The *submodular function minimization problem* is, given a submodular function f on A , to find a subset $X \subseteq A$ with minimum $f(X)$. It is known [11, 17, 16, 26, 27] that a submodular function on a set A can be minimized in polynomial time (in $|A|$) provided getting a value of f is a primitive operation.

A partial order on a set D is called a *lattice order* if, for every $x, y \in D$, there exists a greatest lower bound $x \sqcap y$ and a least upper bound $x \sqcup y$. The corresponding algebra $\mathcal{L} = (D, \sqcap, \sqcup)$ is called a *lattice*. A subset of D closed under the operations $\sqcap$ and $\sqcup$ is called a *sublattice* of $\mathcal{L}$. If $\mathcal{L}_i$ is a lattice on D_i , $i = 1, \dots, n$, then the *product lattice* $\mathcal{L}_1 \times \dots \times \mathcal{L}_n$ is a lattice with base set $D_1 \times \dots \times D_n$ and operations acting componentwise. The lattice $\mathcal{L}^n$ is the (direct) product of n copies of $\mathcal{L}$, and it is known as the *n th power* of $\mathcal{L}$. For more information about lattices, see [13].

DEFINITION 2.2. *Let $\mathcal{L}$ be a lattice on D . A function $f : D^n \rightarrow \mathbb{R}$ is called supermodular on $\mathcal{L}$ if*

$$f(\mathbf{a}) + f(\mathbf{b}) \leq f(\mathbf{a} \sqcap \mathbf{b}) + f(\mathbf{a} \sqcup \mathbf{b}) \text{ for all } \mathbf{a}, \mathbf{b} \in \mathcal{L}^n,$$

and f is called submodular on $\mathcal{L}$ if the inverse inequality holds.

If A is a set with $|A| = n$, then, by identifying subsets of A with 0-1 n -tuples, one can easily check that the submodular (set) functions on A are simply the n -ary submodular functions on a lattice on $\{0, 1\}$ with order $0 < 1$. Therefore, it is natural to consider the problem of minimizing the submodular functions on a given fixed finite lattice $\mathcal{L}$, in the following form:

INSTANCE: A number $n \geq 1$ and a submodular function f on $\mathcal{L}^n$.

GOAL: Find an element $\mathbf{a} \in \mathcal{L}^n$ such that $f(\mathbf{a}) = \min \{f(\mathbf{b}) \mid \mathbf{b} \in \mathcal{L}^n\}$.

We will denote this problem by $\text{SFM}(\mathcal{L})$. We will say that $\text{SFM}(\mathcal{L})$ is *oracle-tractable* if it can be solved in polynomial time in n (provided getting the value of f on a tuple is a primitive operation). It was mentioned in [3] as an open question whether or not $\text{SFM}(\mathcal{L})$ is oracle-tractable for any fixed lattice $\mathcal{L}$. This question was motivated in [3] by its applications in constraint satisfaction which we will describe in the next section.

Note that a function f is submodular if and only if $-f$ is supermodular. Therefore, the $\text{SFM}(\mathcal{L})$ problem can also be understood as the *supermodular function maximization problem*. We will always use this reformulation because we will later apply algorithms for $\text{SFM}(\mathcal{L})$ to solve certain maximization problems.

Recall that a lattice is called *distributive* if it can be represented by subsets of a set A , where the operations $\sqcap$ and $\sqcup$ are interpreted as set-theoretic intersection and union, respectively. Note that, in some earlier papers on submodular functions, a family of subsets closed under intersection and union is simply called a lattice family (or a ring family). The following result is proved in section 49.3 of [27] (see also [26]).

THEOREM 2.3. *A submodular function defined on a lattice family L on a set A can be minimized in polynomial time in $|A|$, provided we can compute in polynomial time the largest and the smallest sets in the family, and the preorder $\preceq$ on A defined as follows: $u \preceq v$ if and only if each set $U \in L$ containing v also contains u .*

Assume that we fix a finite distributive lattice $\mathcal{L}$. It is well known that $\mathcal{L}$ can be represented by subsets of a set A such that $|A| \leq |\mathcal{L}|$. Clearly, we can compute in constant time the sets (as 0-1 $|A|$ -tuples) representing the largest and the smallest elements of $\mathcal{L}$, and we can also compute the preorder $\preceq$ in constant time. Obviously, for any n , the lattice $\mathcal{L}^n$ can be represented by subsets of a set B of cardinality $n|A|$ (since an element of $\mathcal{L}$ is represented by using $|A|$ bits), while the representations for the largest and the smallest elements of $\mathcal{L}^n$ and the preorder for $\mathcal{L}^n$ can be trivially obtained from those for $\mathcal{L}$. It follows that we can solve the problem $\text{SFM}(\mathcal{L})$ in polynomial time in n (in fact, in $n|\mathcal{L}|$, but $|\mathcal{L}|$ is a constant).

To the best of our knowledge, there was up to now not a single nondistributive finite lattice $\mathcal{L}$ for which the problem $\text{SFM}(\mathcal{L})$ is known to be oracle-tractable. We will provide such examples in this paper.

The problem of minimizing submodular functions on nondistributive lattices was mentioned in [17]. However, it was not clear in that paper what the parameter should be in this case, since, in combinatorics, submodular functions are traditionally considered to be defined on (some or all) subsets of a set, and the standard parameter in such situations was always the cardinality of the set. We believe that our formulation of the problem $\text{SFM}(\mathcal{L})$ is an appropriate form of generalization of the standard SFM problem to the case of arbitrary lattices.

We will generalize this problem even further, by considering classes of finite lattices. Let $\mathbf{C}$ be a fixed finite class of finite lattices. Define the optimization problem $\text{SFM}(\mathbf{C})$ as follows:

INSTANCE: A lattice $\mathcal{L}' = \mathcal{L}_1 \times \dots \times \mathcal{L}_n$ such that $\mathcal{L}_i \in \mathbf{C}$ for all $1 \leq i \leq n$, and a supermodular function f on $\mathcal{L}'$.

GOAL: Find an element $\mathbf{a} \in \mathcal{L}'$ such that $f(\mathbf{a}) = \max \{f(\mathbf{b}) \mid \mathbf{b} \in \mathcal{L}'\}$.

One can extend the notion of oracle-tractability to the problems $\text{SFM}(\mathbf{C})$ in a natural way, assuming that an instance is given by an n -tuple of names of lattices in the product and by an oracle for the function f . For any finite class of finite distributive lattices one can follow the same procedure as for a single distributive lattice, so, clearly, the following statement holds.

PROPOSITION 2.4. *$\text{SFM}(\mathbf{C})$ is oracle-tractable for any finite class $\mathbf{C}$ of finite distributive lattices.*

3. Maximum constraint satisfaction. Maximum constraint satisfaction problems are well-studied combinatorial optimization problems. The standard example of such problems are MAX k -CUT and MAX k -SAT.

Let D denote a finite set with $|D| > 1$. Let $R_D^{(m)}$ denote the set of all m -ary predicates over D , that is, functions from D^m to $\{0, 1\}$, and let $R_D = \bigcup_{m=1}^{\infty} R_D^{(m)}$. Also, let $\mathbb{Z}^+$ denote the set of all nonnegative integers.

DEFINITION 3.1. A constraint over a set of variables $V = \{x_1, x_2, \dots, x_n\}$ is an expression of the form $f(\mathbf{x})$, where

- $f \in R_D^{(m)}$ is called the constraint predicate, and
- $\mathbf{x} = (x_{i_1}, \dots, x_{i_m})$ is called the constraint scope.

The constraint f is said to be satisfied on a tuple $\mathbf{a} = (a_{i_1}, \dots, a_{i_m}) \in D^m$ if $f(\mathbf{a}) = 1$.

DEFINITION 3.2. For a finite $\mathcal{F} \subseteq R_D$, an instance of MAX CSP($\mathcal{F}$) is a pair (V, C) , where

- $V = \{x_1, \dots, x_n\}$ is a set of variables taking their values from the set D ,
- C is a collection of constraints $f_1(\mathbf{x}_1), \dots, f_q(\mathbf{x}_q)$ over V , where $f_i \in \mathcal{F}$ for all $1 \leq i \leq q$.

The goal is to find an assignment $\varphi : V \rightarrow D$ that maximizes the number of satisfied constraints, that is, to maximize the function $f : D^n \rightarrow \mathbb{Z}^+$, defined by $f(x_1, \dots, x_n) = \sum_{i=1}^q f_i(\mathbf{x}_i)$. If the constraints have (positive integral) weights ϱ_i , $1 \leq i \leq q$, then the goal is to maximize the total weight of satisfied constraints, that is, maximize the function $f : D^n \rightarrow \mathbb{Z}^+$, defined by $f(x_1, \dots, x_n) = \sum_{i=1}^q \varrho_i \cdot f_i(\mathbf{x}_i)$.

Example 3.3 (MAX k -CUT). In the MAX k -CUT problem, one is given an undirected graph $G = (V, E)$ with weighted edges, and the goal is to find a partition of V into k parts, $V = V_0 \cup V_1 \cup \dots \cup V_{k-1}$, maximizing the total weight of edges with endpoints in different parts. This problem is exactly the MAX CSP($\{\neq_k\}$) problem where $\neq_k$ is the binary disequality predicate on $\{0, 1, \dots, k-1\}$. To see this, think of vertices of a given graph as variables, and apply the predicate to every pair of variables x, y such that (x, y) is an edge in the graph, while keeping all weights the same.

Since predicates are functions, one can consider supermodular predicates on a lattice. For a finite lattice $\mathcal{L}$, we will denote by $\text{Spmod}_{\mathcal{L}}$ the set of all predicates that are supermodular on $\mathcal{L}$.

It is easy to see that if $f(x_1, \dots, x_n) = \sum_{i=1}^q \varrho_i \cdot f_i(\mathbf{x}_i)$ and, for some lattice $\mathcal{L}$, every f_i is supermodular on $\mathcal{L}$, then f is also supermodular on $\mathcal{L}$. Moreover, it is clear that one can compute the value of f on a given tuple in linear time in the size of the instance. Hence, we immediately obtain the following lemma.

LEMMA 3.4. Let $\mathcal{L}$ be a finite lattice such that the problem SFM($\mathcal{L}$) is oracle-tractable. Then, for any finite set $\mathcal{F} \subseteq \text{Spmod}_{\mathcal{L}}$, the problem MAX CSP($\mathcal{F}$) is tractable.

It is intriguing that all known tractable problems MAX CSP($\mathcal{F}$) are essentially (i.e., possibly, after removing redundant elements from D) of this form (i.e., with $\mathcal{F} \subseteq \text{Spmod}_{\mathcal{L}}$ for some lattice $\mathcal{L}$ on D). In particular, it is known [3] (and follows from Proposition 2.4 and Lemma 3.4) that MAX CSP($\mathcal{F}$) is tractable whenever $\mathcal{F}$ consists of supermodular predicates on some distributive lattice.

In the rest of this section, we present evidence that supermodularity on lattices probably is the right tool for studying the complexity of problems MAX CSP($\mathcal{F}$).

First, we will consider a form of supermodular constraints that can be defined on any lattice.

DEFINITION 3.5. A predicate $f \in R_D^{(n)}$ will be called 2-monotone¹ on a lattice $\mathcal{L}$

¹In [3], such predicates are called *generalized 2-monotone*.

on D if it can be expressed as follows:

$$(3.1) \quad f(\mathbf{x}) = 1 \Leftrightarrow ((x_{i_1} \sqsubseteq a_{i_1}) \wedge \dots \wedge (x_{i_s} \sqsubseteq a_{i_s})) \vee ((x_{j_1} \sqsupseteq b_{j_1}) \wedge \dots \wedge (x_{j_t} \sqsupseteq b_{j_t})),$$

where $\mathbf{x} = (x_1, \dots, x_n)$, $a_{i_1}, \dots, a_{i_s}, b_{j_1}, \dots, b_{j_t} \in D$, and either of the two disjuncts may be empty (i.e., the value of s or t may be zero).

It is straightforward to check that every 2-monotone predicate on a lattice is supermodular on it. The next theorem is, to the best of our knowledge, the only one available on the complexity of supermodular constraints on arbitrary lattices.

THEOREM 3.6 (see [3]). *Let $\mathcal{L}$ be a lattice on a finite set D . If $\mathcal{F}$ consists of 2-monotone predicates on $\mathcal{L}$, then $\text{MAX CSP}(\mathcal{F})$ is tractable.*

Note that 2-monotone predicates can be defined on any poset, since the definition does not use the property of the order to be lattice. However, it was shown in [20] that if $\mathcal{F}$ consists of all binary 2-monotone predicates on a *nonlattice* poset, then $\text{MAX CSP}(\mathcal{F})$ is **NP**-hard.

An *endomorphism* of $\mathcal{F}$ is a unary operation π on D such that, for all $f \in \mathcal{F}$ and all $(a_1, \dots, a_m) \in D^m$, we have $f(a_1, \dots, a_m) = 1 \Rightarrow f(\pi(a_1), \dots, \pi(a_m)) = 1$. We say that $\mathcal{F}$ is a *core* if every endomorphism of $\mathcal{F}$ is injective (i.e. a permutation). The intuition here is that if $\mathcal{F}$ is not a core, then it has a noninjective endomorphism π , which implies that, for every assignment φ , there is another assignment $\pi\varphi$ that satisfies all constraints satisfied by φ and uses only a restricted set of values, so the problem $\text{MAX CSP}(\mathcal{F})$ can be reduced to a similar problem over this smaller set.

THEOREM 3.7 (see [5, 3, 19]). *Let $|D| \leq 3$, and let $\mathcal{F} \subseteq R_D$ be a core. If there is a chain $\mathcal{C}$ on D such that $\mathcal{F} \subseteq \text{Spmod}_{\mathcal{C}}$, then $\text{MAX CSP}(\mathcal{F})$ is tractable. Otherwise, $\text{MAX CSP}(\mathcal{F})$ is **NP**-hard.*

For an element $d \in D$, define the unary predicate u_d so that $u_d(x) = 1 \Leftrightarrow x = d$. Let $C_D = \{u_d \mid d \in D\}$.

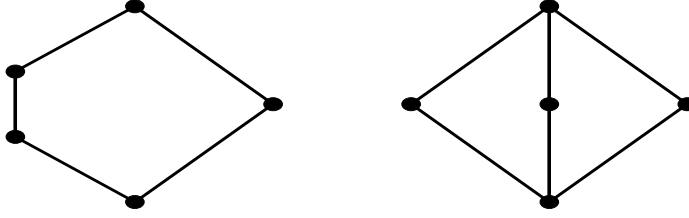
THEOREM 3.8 (see [7, 8]). *Let D be any finite set, and assume that $C_D \subseteq \mathcal{F} \subseteq R_D$. If there is a chain $\mathcal{C}$ on D such that $\mathcal{F} \subseteq \text{Spmod}_{\mathcal{C}}$, then $\text{MAX CSP}(\mathcal{F})$ is tractable. Otherwise, $\text{MAX CSP}(\mathcal{F})$ is **NP**-hard.*

Note that (assuming that supermodularity is the right tool) chains are the only lattices that could possibly appear in Theorems 3.7 and 3.8 because, as is easy to check, every lattice with at most three elements is a chain and every predicate of the form u_d is supermodular on a lattice if and only if the lattice is a chain (the latter assertion is essentially Lemma 5.1 of [3]).

However, it is known that classes of supermodular predicates on (essentially) different lattices are pairwise incomparable. More precisely, for any lattice $\mathcal{L}$, let $\mathcal{L}^\partial$ denote the dual lattice of $\mathcal{L}$, i.e., the one obtained from $\mathcal{L}$ by reversing the order (or by swapping the lattice operations, which is the same). It is obvious from the definition that the classes of supermodular functions on $\mathcal{L}$ and on $\mathcal{L}^\partial$ coincide. It was shown in [20] that, for any finite lattice $\mathcal{L}'$ (on the same set as $\mathcal{L}$) such that $\mathcal{L}'$ is neither $\mathcal{L}$ nor $\mathcal{L}^\partial$, there exists a predicate which is 2-monotone (and hence supermodular) on $\mathcal{L}$, but not supermodular on $\mathcal{L}'$. Hence, essentially, one cannot exclude *any* lattice from these considerations.

It follows that the problem $\text{SFM}(\mathcal{L})$ restricted to supermodular functions on a nondistributive lattice $\mathcal{L}$, such as the functions that can appear in instances of $\text{MAX CSP}(\mathcal{F})$, is of special interest.

It is a basic fact in lattice theory (see, e.g., [13]) that a lattice is distributive if it does not contain, as a sublattice, one of the two minimal nondistributive lattices: the pentagon $\mathcal{N}_5$ and the diamond $\mathcal{M}_3$. These two lattices are depicted in Figure 3.1.

FIG. 3.1. The pentagon $\mathcal{N}_5$ and the diamond $\mathcal{M}_3$.

It will follow from the results in this paper that $\text{MAX CSP}(\mathcal{F})$ is tractable if $\mathcal{F} \subseteq \text{Spmod}_{\mathcal{N}_5}$ or $\mathcal{F} \subseteq \text{Spmod}_{\mathcal{M}_3}$.

4. SFM and constructions on lattices. In this section we show that tractability of $\text{SFM}(\mathcal{L})$ is preserved under certain constructions on lattices, and exhibit a large class of nondistributive lattices for which $\text{SFM}(\mathcal{L})$ is tractable.

4.1. General constructions. A *congruence* on a lattice $\mathcal{L}$ is an equivalence relation θ such that, for all $a, b, c, d \in \mathcal{L}$, the conditions $a\theta b$ and $c\theta d$ imply that both $(a \sqcap c)\theta(b \sqcap d)$ and $(a \sqcup c)\theta(b \sqcup d)$ hold.

If θ is a congruence on $\mathcal{L}$ and $a \in \mathcal{L}$, then let $a[\theta]$ denote the θ -class containing a . It is well known that every θ -class is a sublattice of $\mathcal{L}$. It is also well known that the family of all θ -classes forms a lattice, called a *factor-lattice* of $\mathcal{L}$ and denoted $\mathcal{L}/\theta$, with operations defined as follows: $a[\theta] \sqcap a'[\theta] = (a \sqcap a')[\theta]$ and $a[\theta] \sqcup a'[\theta] = (a \sqcup a')[\theta]$.

We will now introduce a certain notion of a product of classes of lattices that was intensively studied in lattice theory (see, e.g., [14] or pp. 489–490 of [13]).

DEFINITION 4.1. *If $\mathbf{V}$ and $\mathbf{W}$ are classes of lattices, then their Mal'tsev product, denoted $\mathbf{V} \circ \mathbf{W}$, consists of all lattices $\mathcal{L}$ such that there is a congruence θ on $\mathcal{L}$ with the following properties:*

1. *the lattice $\mathcal{L}/\theta$ belongs to $\mathbf{W}$,*
2. *every θ -class is a lattice from $\mathbf{V}$.*

Let $\mathbf{D}$ denote the class of all distributive lattices, let $\mathbf{D}_k$ denote the class of all distributive lattices with at most k elements, and let $\mathbf{D}_{fin}$ denote the class of all finite lattices from $\mathbf{D}$.

Example 4.2. The lattice $\mathcal{N}_5$ belongs to $\mathbf{D}_3 \circ \mathbf{D}_2$. It is easy to check that the equivalence relation θ whose two classes are within the ovals in Figure 4.1 is a congruence. The classes of the congruence are distributive lattices (chains), and the lattice $\mathcal{N}_5/\theta$ is a distributive lattice (a two-element chain).

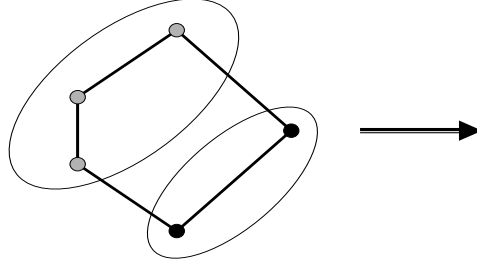
THEOREM 4.3. *Suppose that $\mathbf{V}, \mathbf{W}$ are finite classes of finite lattices. If $\text{SFM}(\mathbf{V})$ and $\text{SFM}(\mathbf{W})$ are both oracle-tractable, then $\text{SFM}(\mathbf{V} \circ \mathbf{W})$ is oracle-tractable as well.*

Proof. Let $\mathcal{L}' = \mathcal{L}_1 \times \cdots \times \mathcal{L}_n$ such that $\mathcal{L}_i \in \mathbf{V} \circ \mathbf{W}$ for all $1 \leq i \leq n$. Then, for every $1 \leq i \leq n$, there exist congruences θ_i such that the lattices $\mathcal{K}_i = \mathcal{L}_i/\theta_i$ belong to $\mathbf{W}$ and every θ_i -class belongs to $\mathbf{V}$.

Define a function f' on $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_n$ by letting

$$f'(a_1[\theta_1], \dots, a_n[\theta_n]) = \max f|_{a_1[\theta_1] \times \cdots \times a_n[\theta_n]}.$$

Let us check that f' is a supermodular function on $\mathcal{K}$. Take two arbitrary elements in $\mathcal{K}$, say $(a_1[\theta_1], \dots, a_n[\theta_n])$ and $(b_1[\theta_1], \dots, b_n[\theta_n])$. Choose a'_i, b'_i , $1 \leq i \leq n$, so that $a_i\theta_i a'_i$ and $b_i\theta_i b'_i$ for all $1 \leq i \leq n$, and $f'(a_1[\theta_1], \dots, a_n[\theta_n]) = f(a'_1, \dots, a'_n)$ and

FIG. 4.1. The pentagon $\mathcal{N}_5$ is in $\mathbf{D}_3 \circ \mathbf{D}_2$.

$f'(b_1[\theta_1], \dots, b_n[\theta_n]) = f(b'_1, \dots, b'_n)$. Since each θ_i is a congruence, it follows that

$$\begin{aligned}
 f'(a_1[\theta_1], \dots, a_n[\theta_n]) + f'(b_1[\theta_1], \dots, b_n[\theta_n]) &= f(a'_1, \dots, a'_n) + f(b'_1, \dots, b'_n) \\
 &\leq f(a'_1 \sqcap b'_1, \dots, a'_n \sqcap b'_n) + f(a'_1 \sqcup b'_1, \dots, a'_n \sqcup b'_n) \\
 &\leq f'((a'_1 \sqcap b'_1)[\theta_1], \dots, (a'_n \sqcap b'_n)[\theta_n]) + f'((a'_1 \sqcup b'_1)[\theta_1], \dots, (a'_n \sqcup b'_n)[\theta_n]) \\
 &= f'((a_1 \sqcap b_1)[\theta_1], \dots, (a_n \sqcap b_n)[\theta_n]) + f'((a_1 \sqcup b_1)[\theta_1], \dots, (a_n \sqcup b_n)[\theta_n]) \\
 &= f'(a_1[\theta_1] \sqcap b_1[\theta_1], \dots, a_n[\theta_n] \sqcap b_n[\theta_n]) + f'(a_1[\theta_1] \sqcup b_1[\theta_1], \dots, a_n[\theta_n] \sqcup b_n[\theta_n]).
 \end{aligned}$$

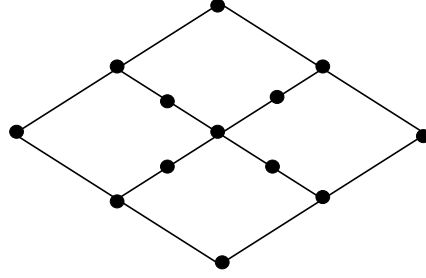
Since $\mathcal{K}$ is a direct product of lattices from $\mathbf{W}$, we infer that f' can be maximized in polynomial time if the evaluation of f' on a given tuple is a primitive operation. That is, f' can be maximized in at most $p_1(n)$ number of steps, where p_1 is a fixed polynomial, and some of the steps are evaluations of f' on a given element of $\mathcal{K}$.

Assume that $\text{SFM}(\mathbf{V})$ can be solved in $p_2(n)$ steps, some of which are function evaluations. Now, to prove the theorem, it suffices to show that f' can be evaluated on any given element of $\mathcal{K}$ in $p_2(n)$ steps (assuming that evaluating f on a given element of $\mathcal{L}'$ is a primitive operation). Fix an element $(a_1[\theta_1], \dots, a_n[\theta_n])$ of $\mathcal{K}$. The goal now is to maximize f on $a_1[\theta_1] \times \dots \times a_n[\theta_n]$. Every $a_i[\theta_i]$ is a lattice from $\mathbf{V}$, so evaluating $f'(a_1[\theta_1], \dots, a_n[\theta_n])$ can be done in $p_2(n)$ steps by assumption of the theorem. Hence, f can be maximized in $p_1(p_2(n))$ steps, some of which are evaluations of f on a given tuple. $\square$

COROLLARY 4.4. *If $\mathcal{L}_1$ and $\mathcal{L}_2$ are finite lattices such that $\text{SFM}(\mathcal{L}_i)$ is oracle-tractable for $i = 1, 2$, then $\text{SFM}(\mathcal{L}_1 \times \mathcal{L}_2)$ is oracle-tractable as well.*

Proof. Let $\mathbf{V} = \{\mathcal{L}_1\}$ and $\mathbf{W} = \{\mathcal{L}_2\}$. It is immediate that $\text{SFM}(\mathbf{V})$ and $\text{SFM}(\mathbf{W})$ are oracle-tractable. The lattice $\mathcal{L}' = \mathcal{L}_1 \times \mathcal{L}_2$ belongs to $\mathbf{V} \circ \mathbf{W}$. Indeed, the relation θ on $\mathcal{L}'$ defined so that $(a_1, b_1)\theta(a_2, b_2)$ if and only if $b_1 = b_2$ is a congruence such that $\mathcal{L}'/\theta$ is isomorphic to $\mathcal{L}_2$ while every θ -class is isomorphic to $\mathcal{L}_1$. Clearly, when maximizing an n -ary supermodular function on $\mathcal{L}_1 \times \mathcal{L}_2$, one can identify $\mathcal{L}'/\theta$ with $\mathcal{L}_2$ and every θ -class with $\mathcal{L}_1$. The result now follows from Theorem 4.3. $\square$

DEFINITION 4.5. *A mapping φ from a lattice $\mathcal{L}_1$ to a lattice $\mathcal{L}_2$ is called a homomorphism if, for all $a, b \in \mathcal{L}_1$, it holds that $\varphi(a \sqcap b) = \varphi(a) \sqcap \varphi(b)$ and $\varphi(a \sqcup b) = \varphi(a) \sqcup \varphi(b)$. If such a mapping φ is onto, then $\mathcal{L}_2$ is said to be a homomorphic image of $\mathcal{L}_1$.*

FIG. 4.2. The lattice $\hat{\mathcal{L}}$.

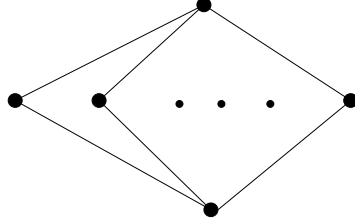
THEOREM 4.6. *Fix a finite lattice $\mathcal{L}$. If $\mathcal{L}$ is a homomorphic image of some finite lattice $\mathcal{L}_1$ such that $\text{SFM}(\mathcal{L}_1)$ is oracle-tractable, then $\text{SFM}(\mathcal{L})$ is oracle-tractable as well.*

Proof. Every supermodular function f on $\mathcal{L}^n$ can be transformed into a supermodular function f_1 on $\mathcal{L}_1^n$ by letting $f_1(a_1, \dots, a_n) = f(\varphi(a_1), \dots, \varphi(a_n))$ where φ is a surjective homomorphism from $\mathcal{L}_1$ to $\mathcal{L}$. It is straightforward to check that f_1 is indeed supermodular on $\mathcal{L}_1$. Since φ is surjective, the maximum of f_1 coincides with the maximum of f . By assumption, one can find a tuple $\mathbf{a}$ maximizing f_1 in polynomial time in n , and the tuple maximizing f is obtained from $\mathbf{a}$ by applying φ componentwise. $\square$

It is clear, say, since the pentagon belongs to $\mathbf{D}_3 \circ \mathbf{D}_2$, that Theorem 4.3 extends (compared to $\mathbf{D}_{fin}$) the class of lattices $\mathcal{L}$ for which $\text{SFM}(\mathcal{L})$ is proved to be oracle-tractable. We note that Theorem 4.6 further extends this class. Consider, for example, the lattice $\hat{\mathcal{L}}$ shown in Figure 4.2. It is shown in the proof of Theorem 1 of [14] that this lattice is a homomorphic image of a lattice belonging to $(\mathbf{D}_4 \circ \mathbf{D}_4) \circ \mathbf{D}_4$, so $\text{SFM}(\hat{\mathcal{L}})$ is oracle-tractable. It is easy to verify that this lattice is *simple*, that is, it has no congruences except for the equality relation and the full binary relation. Lemma 1 of [14] states that if a simple lattice belongs to Mal'tsev product of two classes of lattices, then it belongs to one of these classes. Hence, since $\hat{\mathcal{L}}$ is simple and nondistributive, it does not belong to any class obtainable from $\mathbf{D}_{fin}$ by using Mal'tsev product.

It follows from Proposition 2.4 and Theorems 4.3 and 4.6 that if a finite lattice $\mathcal{L}$ belongs to a class obtained from finite sets of distributive lattices by (repeatedly) using Mal'tsev product and also taking homomorphic images, then $\text{SFM}(\mathcal{L})$ is oracle-tractable. What is the family $\mathbf{F}$ of finite lattices $\mathcal{L}$ which can be obtained as described above? Unfortunately, it seems quite difficult to give a precise characterization of this family because, by [22], the process of repeatedly applying Mal'tsev product to $\mathbf{D}$ results in different classes of lattices for different orders of applying the operation. However, in the next subsection we describe a well-understood and rich subclass of $\mathbf{F}$, which can be obtained by repeatedly applying Mal'tsev product only to $\mathbf{D}_2$. The subclass consists of the so-called *bounded* finite lattices.

We will now describe some lattices that definitely do not belong to the family $\mathbf{F}$. For $t \geq 3$, a t -diamond (or simply a diamond), denoted $\mathcal{M}_t$, is a lattice on an $(t+2)$ -element set such that $0_{\mathcal{M}_t}, 1_{\mathcal{M}_t}$ are the least and the greatest element, respectively, and all t elements in $\mathcal{M}_t \setminus \{0_{\mathcal{M}_t}, 1_{\mathcal{M}_t}\}$ are pairwise incomparable. The Hasse diagram of $\mathcal{M}_t$ is given in Figure 4.3. Note that $\mathcal{M}_3$ is often referred to as *the* diamond. It is well known and easy to check that every diamond is a simple nondistributive lattice. In addition, every diamond has the property that, for every finite lattice $\mathcal{L}$ having

FIG. 4.3. A diamond lattice $\mathcal{M}_t$.

$\mathcal{M}_t$ as a homomorphic image, $\mathcal{L}$ also contains $\mathcal{M}_t$ as a sublattice (see Lemma 6.21 of [15] or the more general Theorem 2.47 of [18]). Hence, in order to show that no diamond belongs to $\mathbf{F}$, it is sufficient to prove that if a finite lattice $\mathcal{L}$ containing $\mathcal{M}_t$ as a sublattice belongs to $\mathbf{X} \circ \mathbf{Y}$, then at least one of these two classes contains a lattice with the same property. So assume that $\mathcal{L} \in \mathbf{X} \circ \mathbf{Y}$. By definition, $\mathcal{L}$ has a congruence θ such that $\mathcal{L}/\theta \in \mathbf{Y}$ and every θ -class is in $\mathbf{X}$. It is well known and easy to show that the restriction θ' of θ on $\mathcal{M}_t$ is a congruence of $\mathcal{M}_t$. By simplicity of $\mathcal{M}_t$, all the elements of $\mathcal{M}_t$ belong either to a single θ' -class or to pairwise different θ' -classes. In the former case, $\mathcal{M}_t$ is entirely contained in some θ -class, so this class is the required lattice in $\mathbf{X}$. In the latter case, $\mathcal{L}/\theta \in \mathbf{Y}$ is the required lattice. This argument can be easily generalized to show that, in fact, no lattice containing $\mathcal{M}_3$ as a sublattice belongs to $\mathbf{F}$.

Even though we have been unable to prove that $\text{SFM}(\mathcal{M}_t)$ is oracle-tractable, we will show in section 5 that the problems $\text{MAX CSP}(\mathcal{F})$ with $\mathcal{F} \subseteq \text{Spmod}_{\mathcal{M}_t}$ are tractable.

4.2. Finite bounded lattices. Let $\mathcal{L}$ be a lattice and $u \sqsubseteq v$ two comparable (and not necessarily distinct) elements of $\mathcal{L}$. Let I denote the interval $[u, v] = \{x \in \mathcal{L} \mid u \sqsubseteq x \sqsubseteq v\}$ in $\mathcal{L}$.

DEFINITION 4.7. *The lattice $\mathcal{L}[I]$ is said to be obtained from $\mathcal{L}$ by doubling the interval I if the base set of $\mathcal{L}[I]$ is $(\mathcal{L} \setminus I) \cup (I \times \{0, 1\})$ and $x \sqsubseteq y$ holds in $\mathcal{L}[I]$ if and only if*

- $x \sqsubseteq y$ in $\mathcal{L}$ and $x, y \notin I$, or
- $x \in \mathcal{L} \setminus I$, $y = (b, j)$ and $x \sqsubseteq b$ in $\mathcal{L}$, or
- $x = (a, i)$, $y \in \mathcal{L} \setminus I$ and $a \sqsubseteq y$ in $\mathcal{L}$, or
- $x = (a, i)$, $y = (b, j)$ such that $a, b \in I$ and $a \sqsubseteq b$ in $\mathcal{L}$, and $i \leq j$.

Essentially, doubling an interval I in $\mathcal{L}$ means replacing it with $I \times \{0, 1\}$.

DEFINITION 4.8. *A finite lattice $\mathcal{L}$ is called bounded² if there is a sequence $\mathcal{L}_1, \dots, \mathcal{L}_n$ such that $\mathcal{L}_1$ is one-element, $\mathcal{L}_n$ is isomorphic to $\mathcal{L}$, and, for $j = 1, \dots, n-1$, there exists an interval I_j in $\mathcal{L}_j$ such that $\mathcal{L}_{j+1}$ is isomorphic to $\mathcal{L}_j[I_j]$.*

In other words, a finite lattice is called bounded if it can be obtained from the one-element lattice by successive doubling of intervals. For example, Figure 4.4 demonstrates that the pentagon is a bounded lattice, while it is not hard to check the lattice $\hat{\mathcal{L}}$ from Figure 4.2 is not bounded. Bounded lattices play an important role in lattice theory [18]. The name “bounded” comes from an equivalent characterization of such lattices, which is usually used as the definition (but does not play any role in this

²Not to be confused with posets that have both the least and the greatest element.

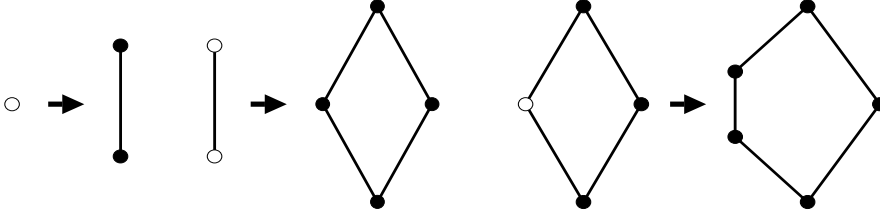


FIG. 4.4. Obtaining the pentagon by doubling intervals (in three steps). In each step, the white elements indicate the interval to be doubled.

paper). Theorem 2.46 of [18] states that a finite lattice $\mathcal{L}$ is bounded if and only if there exists a free lattice $\mathcal{FL}$ and a congruence α on $\mathcal{FL}$ such that $\mathcal{L}$ is isomorphic to $\mathcal{FL}/\alpha$, and every α -class in $\mathcal{FL}$ has both the least and the greatest element.

It is easy to check that for any lattice $\mathcal{L}$ and any interval I in $\mathcal{L}$, the binary relation θ on $\mathcal{L}[I]$ defined by the rule

$$(x, y) \in \theta \Leftrightarrow \text{either } x = y \notin I \text{ or } x = (a, i), y = (a, j) \text{ for some } a \in I$$

is a congruence of $\mathcal{L}[I]$. Moreover, the lattice $\mathcal{L}[I]/\theta$ is isomorphic to $\mathcal{L}$, and every θ -class is either a one-element lattice $\mathcal{C}_1$ or a two-element chain $\mathcal{C}_2$. Therefore, we have that $\mathcal{L}[I]$ belongs to the Mal'tsev product $\{\mathcal{C}_1, \mathcal{C}_2\} \circ \{\mathcal{L}\}$. So, if we inductively define $\mathbf{D}_2^1 = \mathbf{D}_2$ and $\mathbf{D}_2^n = \mathbf{D}_2 \circ \mathbf{D}_2^{n-1}$ for $n > 1$, then we have that every finite bounded lattice belongs to $\mathbf{D}_2^n$ for a suitable n . Hence, Theorem 4.3 (together with Lemma 3.4) implies the following statement.

PROPOSITION 4.9. *Let $\mathcal{L}$ be a fixed finite bounded lattice. Then $\text{SFM}(\mathcal{L})$ is oracle-tractable and, for any finite set $\mathcal{F} \subseteq \text{Spmod}_{\mathcal{L}}$, $\text{MAX CSP}(\mathcal{F})$ is tractable.*

It is known (see Lemma 2.40 of [18]) that the class of finite bounded lattices is a pseudovariety, that is, it is closed under taking homomorphic images, sublattices, and finite direct products. Since the two-element chain is a bounded lattice, it follows that every finite distributive lattice is bounded. Interestingly, if one allows not only finite, but arbitrary direct products (i.e., direct products of infinitely many lattices), then one can generate *all* lattices from finite bounded lattices (this fact immediately follows from Theorems 2.25 and 2.44 of [18]). In other words, every (not necessarily finite) lattice is a homomorphic image of a subalgebra of a direct product of (possibly infinitely many) finite bounded lattices. Unfortunately, this fact seems to be useless for the algorithmic aspects which we are interested in.

5. MAX CSP on diamonds. In this section, we consider problems $\text{MAX CSP}(\mathcal{F})$ where $\mathcal{F}$ consists of supermodular predicates on a diamond $\mathcal{M}_t$ (see Figure 4.3). The middle elements of $\mathcal{M}_t$ are called *atoms*. Note that, for every pair of distinct atoms a and b , we have $a \sqcap b = 0_{\mathcal{M}_t}$ and $a \sqcup b = 1_{\mathcal{M}_t}$. For simplicity, let $\mathcal{L}$ denote an arbitrary (fixed) t -diamond, $t \geq 3$, throughout this section.

5.1. The structure of supermodular predicates on diamonds. In this subsection, we describe the structure of supermodular predicates on $\mathcal{L}$ by representing them as logical formulas involving constants (elements of $\mathcal{L}$) and the order relation $\sqsubseteq$ of $\mathcal{L}$.

For a subset $D' \subseteq D$, let $u_{D'}$ denote the predicate such that $u_{D'}(x) = 1 \Leftrightarrow x \in D'$. The following lemma can be easily derived directly from the definition of supermodularity.

LEMMA 5.1. *A unary predicate $u_{D'}$ is in $\text{Spmod}_{\mathcal{L}}$ if and only if either both $0_{\mathcal{L}}, 1_{\mathcal{L}} \in D'$ or else $|D'| \leq 2$ and at least one of $0_{\mathcal{L}}, 1_{\mathcal{L}}$ is in D' .*

For a sequence $\mathbf{y} = (x_{i_1}, \dots, x_{i_m})$ of variables and a sequence $\mathbf{c} = (c_{i_1}, \dots, c_{i_m})$ of elements of $\mathcal{L}$, we write $\mathbf{y} \sqsubseteq \mathbf{c}$ to denote $\bigwedge_{1 \leq s \leq m} (x_{i_s} \sqsubseteq c_{i_s})$, and the condition $\mathbf{y} \sqsupseteq \mathbf{c}$ is defined dually.

THEOREM 5.2. *Every predicate $f(x_1, \dots, x_n)$ in $\text{Spmod}_{\mathcal{L}}$, such that f takes both values 0 and 1, can be represented as one of the following logical implications:*

1. $[(x_i \sqsubseteq a_1) \vee \dots \vee (x_i \sqsubseteq a_l)] \implies (x_i \sqsubseteq 0_{\mathcal{L}})$, where the a_j 's are atoms;
2. $\neg(\mathbf{y} \sqsupseteq \mathbf{c}) \implies (\mathbf{z} \sqsubseteq \mathbf{d})$, where $\mathbf{y}$ and $\mathbf{z}$ are some subsequences of $(x_1, \dots, x_n)$, and $\mathbf{c}, \mathbf{d}$ are tuples of elements of $\mathcal{L}$ (of corresponding length) such that $\mathbf{c}$ contains no $0_{\mathcal{L}}$ and $\mathbf{d}$ no $1_{\mathcal{L}}$;
3. $[(x_i \sqsubseteq b_1) \vee \dots \vee (x_i \sqsubseteq b_k) \vee \neg(\mathbf{y} \sqsupseteq \mathbf{c})] \implies (x_i \sqsubseteq a)$, where the b_j 's are atoms, $\mathbf{y}$ does not contain x_i , and $a \neq 1_{\mathcal{L}}$;
4. $\neg(x_i \sqsupseteq b) \implies [\neg(x_i \sqsupseteq a_1) \wedge \dots \wedge \neg(x_i \sqsupseteq a_l) \wedge (\mathbf{y} \sqsubseteq \mathbf{c})]$, where the a_j 's are atoms, $\mathbf{y}$ does not contain x_i , and $b \neq 0_{\mathcal{L}}$;
5. $\neg(\mathbf{y} \sqsupseteq \mathbf{c}) \implies \text{false}$, where $\mathbf{y}$ is a subsequence of $(x_1, \dots, x_n)$ and $\mathbf{c}$ contains no $0_{\mathcal{L}}$;
6. $\text{true} \implies (\mathbf{y} \sqsubseteq \mathbf{c})$, where $\mathbf{y}$ is a subsequence of $(x_1, \dots, x_n)$ and $\mathbf{c}$ contains no $1_{\mathcal{L}}$.

Conversely, every predicate that can be represented in one of the above forms belongs to $\text{Spmod}_{\mathcal{L}}$.

Example 5.3. The unary predicate of type (1) above is the same as $u_{D'}$ where $D' = D \setminus \{a_1, \dots, a_l\}$. The predicates $u_{D'} \in \text{Spmod}_{\mathcal{L}}$ with $|D'| \leq 2$ are the unary predicates of types (5) and (6).

Remark 5.4. Note that constraints of types (2), (5), and (6) are 2-monotone on $\mathcal{L}$, while constraints of types (3) and (4) (and most of those of type (1)) are not.

Proof of Theorem 5.2. It is straightforward to verify that all of the predicates in the list are actually supermodular. Now we prove the converse. Consider first the case where the predicate f is *essentially unary*, i.e., there is a variable x_i such that $f(x_1, \dots, x_n) = u_{D'}(x_i)$ for some $D' \subsetneq D$. If $D' = \{x : x \sqsubseteq a\}$ or $D' = \{x : x \sqsupseteq a\}$ for some atom a , then f is of the form (5) or (6); otherwise, both $0_{\mathcal{L}}$ and $1_{\mathcal{L}}$ are in D' by Lemma 5.1, and if $a_1, \dots, a_l$ denote the atoms of the lattice that are not in D' , then it is clear that f is described by the implication (1).

Now we may assume that f is not essentially unary. If it is 2-monotone, then it is easy to see that f must be described by an implication of type (2), (5), or (6). So now we assume that f is not essentially unary and it is not 2-monotone; we prove that it is described by an implication of type (3) or (4). We require a few claims:

Claim 0. *The set X of all tuples $\mathbf{u}$ such that $f(\mathbf{u}) = 1$ is a sublattice of $\mathcal{L}^n$, i.e., is closed under join and meet.*

This follows immediately from the supermodularity of f .

Claim 1. *There exist indices $1 \leq i_1, \dots, i_k, j_1, \dots, j_l \leq n$, atoms $e_1, \dots, e_k$ and $b_1, \dots, b_l$ of $\mathcal{L}$ such that $f(\mathbf{x}) = 1$ if and only if*

$$[\neg(x_{i_1} \sqsupseteq e_1) \wedge \dots \wedge \neg(x_{i_k} \sqsupseteq e_k)] \bigvee [\neg(x_{j_1} \sqsubseteq b_1) \wedge \dots \wedge \neg(x_{j_l} \sqsubseteq b_l)].$$

Notice first that the set Z of tuples $\mathbf{u}$ such that $f(\mathbf{u}) = 0$ is *convex* in $\mathcal{L}^n$; i.e., if $\mathbf{u} \sqsubseteq \mathbf{v} \sqsubseteq \mathbf{w}$ with $f(\mathbf{u}) = f(\mathbf{w}) = 0$, then $f(\mathbf{v}) = 0$. To show this we construct a tuple $\mathbf{v}'$ as follows: for each coordinate i it is easy to find an element v'_i such that $v_i \sqcap v'_i = u_i$ and $v_i \sqcup v'_i = w_i$. Hence $\mathbf{v} \sqcap \mathbf{v}' = \mathbf{u}$ and $\mathbf{v} \sqcup \mathbf{v}' = \mathbf{w}$ so by supermodularity of f neither $\mathbf{v}$ nor $\mathbf{v}'$ is in X . It follows in particular that neither $0_{\mathcal{L}^n}$ nor $1_{\mathcal{L}^n}$ is in

Z ; indeed, if $0_{\mathcal{L}^n} \in Z$, let $\mathbf{a}$ be the smallest element in X (the meet of all elements in X), which exists by Claim 0. Since Z is convex it follows that every element above $\mathbf{a}$ is in X so f is 2-monotone, a contradiction. The argument for $1_{\mathcal{L}^n}$ is identical.

Now let $\mathbf{w} \in Z$ be minimal, and let $\mathbf{v} \sqsubseteq \mathbf{w}$. As above we can find a tuple $\mathbf{v}'$ such that $\mathbf{v} \sqcup \mathbf{v}' = \mathbf{w}$ and $\mathbf{v} \sqcap \mathbf{v}' = 0_{\mathcal{L}^n}$; by supermodularity of f it follows that $\mathbf{v} = \mathbf{w}$ or $\mathbf{v}' = \mathbf{w}$. It is easy to deduce from this that there exists a coordinate s such that w_s is an atom of $\mathcal{L}$ and $w_t = 0_{\mathcal{L}}$ for all $t \neq s$. A similar argument shows that every maximal element of Z has a unique coordinate which is an atom and all others are equal to $1_{\mathcal{L}}$.

Since Z is convex, we have that $f(\mathbf{x}) = 0$ if and only if $\mathbf{x}$ is above some minimal element of Z and below some maximal element of Z ; Claim 1 then follows immediately.

For each index $i \in \{i_1, \dots, i_k\}$ that appears in the expression in Claim 1, there is a corresponding condition of the form

$$\neg(x_i \sqsupseteq e_{s_1}) \wedge \dots \wedge \neg(x_i \sqsupseteq e_{s_r});$$

let I_i denote the set of elements of $\mathcal{L}$ that satisfy this condition. Obviously it cannot contain $1_{\mathcal{L}}$ and must contain $0_{\mathcal{L}}$. Similarly, define for each index $j \in \{j_1, \dots, j_l\}$ the set F_j of all elements of $\mathcal{L}$ that satisfy the corresponding condition of the form

$$\neg(x_j \sqsubseteq b_{t_1}) \wedge \dots \wedge \neg(x_j \sqsubseteq b_{t_q});$$

it is clear that $0_{\mathcal{L}} \notin F_j$ and $1_{\mathcal{L}} \in F_j$.

The condition of Claim 1 can now be rephrased as follows: $f(\mathbf{x}) = 1$ if and only if $x_i \in I_i$ for all $i \in \{i_1, \dots, i_k\}$ or $x_j \in F_j$ for all $j \in \{j_1, \dots, j_l\}$. It is straightforward to verify that since f is not 2-monotone, one of the I_i or one of the F_j must contain 2 distinct atoms. We consider the first case, and we show that the predicate f is of type (4). The case where some F_j contains 2 atoms is dual and will yield type (3).

Claim 2. Suppose that I_i contains distinct atoms c and d for some $i \in \{i_1, \dots, i_k\}$. Then (a) i is the only index with this property, (b) $\{j_1, \dots, j_l\} = \{i\}$, and (c) F_i does not contain 2 distinct atoms.

We prove (b) first. We have that

$$f(0_{\mathcal{L}}, \dots, 0_{\mathcal{L}}, c, 0_{\mathcal{L}}, \dots, 0_{\mathcal{L}}) = f(0_{\mathcal{L}}, \dots, 0_{\mathcal{L}}, d, 0_{\mathcal{L}}, \dots, 0_{\mathcal{L}}) = 1$$

(where c and d appear in the i th position) and by supermodularity it follows that we also have $f(0_{\mathcal{L}}, \dots, 0_{\mathcal{L}}, 1_{\mathcal{L}}, 0_{\mathcal{L}}, \dots, 0_{\mathcal{L}}) = 1$. Since I_i does not contain $1_{\mathcal{L}}$, we have that $x_j \in F_j$ for each $j \in \{j_1, \dots, j_l\}$; since F_j never contains $0_{\mathcal{L}}$, (b) follows immediately. Since $\{j_1, \dots, j_l\}$ is nonempty, (a) follows immediately from (b). Finally, if F_i contained distinct atoms, then by dualizing the preceding argument we would obtain that $\{i_1, \dots, i_k\} = \{i\}$ from which it would follow that f would be essentially unary, contrary to our assumption. This concludes the proof of the claim.

Let b denote the minimal element in F_i , and for each index $s \in \{i_1, \dots, i_k\}$ different from i let c_s denote the (unique) maximal element of I_s ; then we can describe f as follows: $f(\mathbf{x}) = 1$ if and only if

$$[x_i \in I_i \wedge (\mathbf{y} \sqsubseteq \mathbf{c})] \vee (x_i \sqsupseteq b),$$

where $\mathbf{y}$ is a tuple of variables different from x_i and $\mathbf{c}$ is the tuple whose entries are the c_s defined previously. It remains to rewrite the condition $x_i \in I_i$. Suppose first

that there exists at least one atom of $\mathcal{L}$ outside I_i , and let $a_1, \dots, a_l$ denote the atoms outside I_i . Then it is clear that $x_i \in I_i$ if and only if $\neg(x_i \sqsupseteq a_1) \vee \dots \vee \neg(x_i \sqsupseteq a_l)$ holds, so the predicate f is of type (4) (simply restate the disjunction as an implication). Now for the last possibility, where I_i contains all of D except $1_{\mathcal{L}}$; then it is easy to see that f can be described by the following:

$$[\neg(x_i \sqsupseteq b) \wedge (\mathbf{y} \sqsubseteq \mathbf{c})] \vee (x_i \sqsupseteq b),$$

and this completes the proof of the theorem. $\square$

We remark that the preceding theorem can be extended to give a similar characterization of the supermodular constraints on lattices in the larger class of so-called *relatively complemented lattices* [21].

5.2. Supermodular constraints on diamonds are tractable.

THEOREM 5.5. *If $\mathcal{F} \subseteq \text{Spmod}_{\mathcal{L}}$, then $\text{MAX CSP}(\mathcal{F})$ can be solved (to optimality) in $O(n^3 \cdot |\mathcal{L}|^3 + q^3)$ time where n is the number of variables and q is the number of constraints in an instance.*

Proof. We will show how the problem can be reduced to the well-known tractable problem MIN CUT.

Let $\mathcal{I} = \{\rho_1 \cdot f_1(\mathbf{x}_1), \dots, \rho_q \cdot f_q(\mathbf{x}_q)\}$, $q \geq 1$, be an instance of weighted MAX CSP ($\mathcal{F}$), over a set of variables $V = \{x_1, \dots, x_n\}$.

Construction.

Let ∞ denote an integer greater than $\sum \rho_i$. For each constraint f_i , fix a representation as described in Theorem 5.2. In the following construction, we will refer to the type of f_i which will be a number from 1 to 6 according to the type of representation. Every condition of the form $(\mathbf{y} \sqsubseteq \mathbf{c})$ will be read as $\bigwedge (x_{i_s} \sqsubseteq c_{i_s})$, and every condition of the form $\neg(\mathbf{y} \sqsupseteq \mathbf{c})$ as $\bigvee \neg(x_{i_s} \sqsupseteq c_{i_s})$, where i_s runs through the indices of variables in $\mathbf{y}$. Moreover, we replace every (sub)formula of the form $\neg(x \sqsupseteq 1_{\mathcal{L}})$ by $\bigvee_{i=1}^n \neg(x \sqsupseteq a_i)$, where $a_1, \dots, a_n$ are the atoms of $\mathcal{L}$.

Next, we construct a digraph $G_{\mathcal{I}}$.

- The vertices of $G_{\mathcal{I}}$ are as follows:
 - $\{T, F\} \cup \{x_d \mid x \in V, d \in \mathcal{L}\} \cup \{\bar{x}_d \mid x \in V, d \in \mathcal{L} \text{ is an atom}\} \cup \{e_i, \bar{e}_i \mid i = 1, 2, \dots, q\}$.³

For each f_i of type (5), we identify the vertex e_i with F . Similarly, for each f_i of type (6), we identify the vertex $\bar{e}_i$ with T .

- The arcs of $G_{\mathcal{I}}$ are defined as follows:
 - For each atom c in $\mathcal{L}$ and for each $x \in V$, there is an arc from $x_{0_{\mathcal{L}}}$ to x_c with weight ∞ , and an arc from $\bar{x}_c$ to $x_{1_{\mathcal{L}}}$ with weight ∞ .
 - For each pair of distinct atoms c, d in $\mathcal{L}$ and for each $x \in V$, there is an arc from x_c to $\bar{x}_d$ with weight ∞ .
 - For each f_i , there is an arc from $\bar{e}_i$ to e_i with weight ρ_i .
 - For each f_i of types (1–4), and each subformula of the form $(x \sqsubseteq a)$ or $\neg(x \sqsupseteq a)$ in the consequent of f_i , there is an arc from e_i to x_a or $\bar{x}_a$, respectively, with weight ∞ .
 - For each f_i of types (1–4), and each subformula of the form $(x \sqsubseteq a)$ or $\neg(x \sqsupseteq a)$ in the antecedent of f_i , there is an arc from x_a or $\bar{x}_a$, respectively, to $\bar{e}_i$, with weight ∞ .
 - For each f_i of type (5), and each subformula of the form $\neg(x \sqsupseteq a)$ in it, there is an arc from $\bar{x}_a$ to $\bar{e}_i$ with weight ∞ .

³The vertices x_d will correspond to the expressions $x \sqsubseteq d$ and $\bar{x}_d$ to $\neg(x \sqsupseteq d)$.

- For each f_i of type (6), and each subformula of the form $(x \sqsubseteq a)$ in it, there is an arc from e_i to x_a with weight ∞ .

Arcs with weight less than ∞ will be called *constraint arcs*.

It is easy to see that $G_{\mathcal{I}}$ is a digraph with source T (corresponding to **true**) and sink F (corresponding to **false**). Note that paths of nonconstraint arcs between vertices corresponding to any given variable $x \in V$ precisely correspond to logical implications that hold between the corresponding assertions. Throughout the proof, we say “a cut in $G_{\mathcal{I}}$ ” meaning a (T, F) -cut.

Define the *deficiency* of an assignment φ as the difference between $\sum_{i=1}^q \rho_i$ and the evaluation of φ on $\mathcal{I}$. In other words, the deficiency of φ is the total weight of constraints not satisfied by φ . We will prove that minimal cuts in $G_{\mathcal{I}}$ exactly correspond to optimal assignments to $\mathcal{I}$. More precisely, we will show that, for each minimal cut in $G_{\mathcal{I}}$ with weight ρ , there is an assignment for $\mathcal{I}$ with deficiency at most ρ , and, for each assignment to $\mathcal{I}$ with deficiency ρ' , there is a cut in $G_{\mathcal{I}}$ with weight ρ' .

The semantics of the construction of $G_{\mathcal{I}}$ will be as follows: the vertices of the form x_a or $\bar{x}_a$ correspond to assertions of the form $x \sqsubseteq a$ or $\neg(x \sqsupseteq a)$, respectively, and arcs denote implications about these assertions. Given a minimal cut in $G_{\mathcal{I}}$, we will call a vertex x_a *reaching* if F can be reached from it without crossing the cut. Furthermore, if a vertex x_a is reaching, then this will designate that the corresponding assertion is false, and otherwise the corresponding assertion is true. A constraint is not satisfied if and only if the corresponding constraint arc crosses the cut.

Let C be a minimal cut in $G_{\mathcal{I}}$. Obviously, C contains only constraint arcs. First we show that, for every variable $x \in V$, there is a unique minimal element $a \in \mathcal{L}$ such that x_a is nonreaching. All we need to show is the following: if c, d are distinct atoms such that both x_c and x_d are both nonreaching, then so is x_{0_c} . Assume that, on the contrary, x_{0_c} is reaching. Then there is a path from x_{0_c} to F not crossing the cut. It is easy to notice that such a path has to go through a vertex $\bar{x}_a$ for some atom $a \in \mathcal{L}$, since the second vertex on this path must be of the form x_b for some atom b , and it is followed either by a vertex $\bar{x}_a$ or else by three vertices $\bar{e}_i, e_i, x_{b'}$ for some $1 \leq i \leq q$ and some atom b' . However, we have an arc from at least one of the vertices x_c, x_d to $\bar{x}_a$, and hence at least one of these vertices would have a path to F not crossing the cut, a contradiction.

Note that, for every $x \in V$, there are no arcs coming out of x_{1_c} . Hence, for every $x \in V$, there is a unique minimal element $v \in \mathcal{L}$ such that F cannot be reached from x_v without crossing the cut.

Define an assignment φ_C as follows:

$\varphi_C(x)$ is the unique minimal element a such that x_a is nonreaching.

We now make some observations. Note that, for all $x \in V$ and $a \in \mathcal{L}$, we have that $\varphi_C(x) \sqsubseteq a$ if and only if x_a is nonreaching. Moreover, if $\bar{x}_a$ is reaching, then, for each atom $b \neq a$, we have an arc from x_b to $\bar{x}_a$ meaning that $\varphi_C(x) \not\sqsubseteq b$, and hence $\varphi_C(x) \sqsupseteq a$. Furthermore, if $\bar{x}_a$ is nonreaching, then $\varphi_C(x) \neq a$. Indeed, if $\varphi_C(x) = a$, then x_b is reaching for all atoms $b \neq a$, and, since every path from x_b to F has to go through a vertex $\bar{x}_c$ for some c , we have that $\bar{x}_c$ is reaching. Then $c \neq a$, and there is an arc from x_a to $\bar{x}_c$, so x_a is reaching, a contradiction. To summarize, we have the following:

- If a node of the form x_a or $\bar{x}_a$ is reaching, then the corresponding assertion is falsified by the assignment φ_C .
- If a node of the form x_a is nonreaching, then $\varphi_C(x) \sqsubseteq a$.

- If a node of the form $\bar{x}_a$ is nonreaching, then the truth value of the corresponding assertion is undecided.

Suppose that a constraint arc corresponding to a constraint f_i is not in the cut. We claim that f_i is satisfied by the assignment φ_C . To show this, we will go through the possible types of f_i .

If f_i is of type (1), (2), (5), or (6), then the claim is straightforward. For example, let f_i be of type (1). If the node $x_{0_{\mathcal{L}}}$ corresponding to the consequent is reaching, then so are all nodes corresponding to the antecedent. Hence, all atomic formulas are falsified by the assignment φ_C , and the implication is true. If $x_{0_{\mathcal{L}}}$ is nonreaching, then $\varphi_C(x) = 0_{\mathcal{L}}$, and the constraint is clearly satisfied. The argument for types (2), (5), and (6) is very similar.

Let f_i be of type (3). Then, if the node corresponding to the consequent is nonreaching, then the consequent is satisfied by φ_C , and so the constraint is satisfied. If this node is reaching, then every node corresponding to the disjuncts in the antecedent is reaching. Then both antecedent and consequent are falsified by φ_C , and the constraint is satisfied.

Let f_i be of type (4), that is, of the form

$$\neg(x_i \sqsupseteq b) \implies [\neg(x_i \sqsupseteq a_1) \wedge \cdots \wedge \neg(x_i \sqsupseteq a_l) \wedge (\mathbf{y} \sqsubseteq \mathbf{c})].$$

If a node corresponding to some conjunct in the consequent is reaching, then the node corresponding to the antecedent is also reaching. So $\varphi_C(x_i) \sqsupseteq b$, and the constraint is satisfied. More generally, if the node corresponding to the antecedent is reaching, then the constraint is satisfied regardless of what happens with the consequent. Assume that all nodes corresponding to conjuncts in the consequent and in the antecedent are nonreaching. Then the conjunct $(\mathbf{y} \sqsubseteq \mathbf{c})$ is satisfied by φ_C . Furthermore, we know (see the observations above) that $\varphi_C(x_i) \neq b$, and also that $\varphi_C(x_i) \neq a_s$ for $1 \leq s \leq l$. If $\varphi_C(x_i) = 1_{\mathcal{L}}$, then both the antecedent and the consequent of f_i are false, and hence f_i is satisfied. Otherwise, $\varphi_C(x_i) \not\sqsupseteq b$ and $\varphi_C(x_i) \not\sqsupseteq a_s$ for $1 \leq s \leq l$, so f_i is satisfied anyway.

Conversely, let φ be an assignment to $\mathcal{I}$, and let K be the set of constraints in $\mathcal{I}$ that are not satisfied by φ . Consider any path from T to F . It is clear that if all constraints corresponding to constraint arcs on this path are satisfied, then we have a chain of valid implications starting from **true** and finishing at **false**. Since this is impossible, at least one constraint corresponding to such an arc is not satisfied by φ . Hence, the constraint arcs corresponding to constraints in K form a cut in $G_{\mathcal{I}}$. Furthermore, by the choice of K , the weight of this cut is equal to the deficiency of φ .

It follows that the standard algorithm [12] for the MIN CUT problem can be used to find an optimal assignment for any instance of MAX CSP($\mathcal{F}$). This algorithm runs in $O(k^3)$, where k is the number of vertices in the graph. Since the number of vertices in $G_{\mathcal{I}}$ is at most $2(1 + n \cdot |D| + q)$, the result follows. $\square$

We remark that a partial converse to Theorem 5.5 was proved in [20] where it is shown that if $\mathcal{F}$ contains all 2-monotone predicates on a diamond and any predicate which is not supermodular on that diamond, then MAX CSP($\mathcal{F}$) is NP-hard.

6. Conclusion. We have described a large class of lattices on which the SFM problem is oracle-tractable, and an even larger class of lattices $\mathcal{L}$ such that MAX CSP($\mathcal{F}$) is tractable whenever all predicates in $\mathcal{F}$ are supermodular on $\mathcal{L}$. We believe that more progress in the study of maximum constraint satisfaction can be achieved by further blending supermodular optimization and algebraic lattice theory.

The three most standard constructions on algebras (and on lattices in particular) are the forming of homomorphic images, subalgebras, and direct products. We showed (Corollary 4.4 and Theorem 4.6) that two of them preserve oracle-tractability of the problems $\text{SFM}(\mathcal{L})$. However, it is an open problem whether, for any finite lattice $\mathcal{L}_1$, oracle-tractability of $\text{SFM}(\mathcal{L}_1)$ implies oracle-tractability of $\text{SFM}(\mathcal{L}_2)$ for any sublattice $\mathcal{L}_2$ of $\mathcal{L}_1$.

This paper explores two ways of obtaining tractability results for MAX CSP problems—one is via oracle-tractability of the SFM problem, and the other via explicit description of supermodular predicates. One other way that remains to be explored has to do with implicit methods (see [2, 4]) of showing that some predicates can be simulated (or “strictly implemented” [8, 19]) by other predicates. One interesting question in this connection is whether any supermodular predicate on a lattice can be simulated, i.e., strictly implemented, by 2-monotone predicates on that lattice. A positive answer to this question would, together with Theorem 3.6, imply that $\text{MAX CSP}(\mathcal{F})$ is tractable for arbitrary sets $\mathcal{F}$ of supermodular predicates on a lattice.

Finally, we hope that the study of sub- and supermodular functions on lattices will find more applications in combinatorial optimization.

REFERENCES

- [1] R. E. BURKARD, B. KLINZ, AND R. RUDOLF, *Perspectives of Monge properties in optimization*, Discrete Appl. Math., 70 (1996), pp. 95–161.
- [2] D. COHEN, M. COOPER, AND P. JEAUVONS, *An algebraic characterization of complexity for valued constraints*, Lecture Notes in Comput. Sci. 4204, Springer-Verlag, Berlin, 2006, pp. 107–121.
- [3] D. COHEN, M. COOPER, P. JEAUVONS, AND A. KROKHIN, *Supermodular functions and the complexity of MAX CSP*, Discrete Appl. Math., 149 (2005), pp. 53–72.
- [4] D. COHEN, M. COOPER, P. JEAUVONS, AND A. KROKHIN, *The complexity of soft constraint satisfaction*, Artificial Intelligence, 170 (2006), pp. 983–1016.
- [5] N. CREIGNOU, S. KHANNA, AND M. SUDAN, *Complexity Classifications of Boolean Constraint Satisfaction Problems*, SIAM Monogr. Discrete Math. Appl. 7, SIAM, Philadelphia, 2001.
- [6] R. DECHTER, *Constraint Processing*, Morgan Kaufmann, San Francisco, CA, 2003.
- [7] V. DEINEKO, P. JONSSON, M. KLASSON, AND A. KROKHIN, *Supermodularity on chains and complexity of maximum constraint satisfaction*, in Proceedings of Eurocomb’05, Vol. AE, DMTCS, Berlin, Germany, 2005, pp. 51–56.
- [8] V. DEINEKO, P. JONSSON, M. KLASSON, AND A. KROKHIN, *The Approximability of MAX CSP with Fixed-Value Constraints*, Technical report CoRR cs.CC/0602075, 2006.
- [9] A. FRANK, *Applications of submodular functions*, in Surveys in Combinatorics, London Math. Soc. Lecture Note Ser. 187, K. Walker, ed., Cambridge University Press, Cambridge, UK, 1993, pp. 85–136.
- [10] A. FRANK, *Matroids and submodular functions*, in Annotated Bibliographies in Combinatorial Optimization, M. Dell’Amico, F. Maffoli, and S. Martello, eds., John Wiley, Chichester, UK, 1997, pp. 65–80.
- [11] S. FUJISHIGE, *Submodular Functions and Optimization*, Ann. Discrete Math. 58, 2nd ed., Elsevier, Amsterdam, 2005.
- [12] A. GOLDBERG AND R. E. TARJAN, *A new approach to the maximum-flow problem*, J. Assoc. Comput. Mach., 35 (1988), pp. 921–940.
- [13] G. GRÄTZER, *General Lattice Theory*, 2nd ed., Birkhäuser-Verlag, Basel, Switzerland, 2002.
- [14] G. GRÄTZER AND D. KELLY, *Products of lattice varieties*, Algebra Universalis, 21 (1985), pp. 33–45.
- [15] D. HOBBY AND R. N. MCKENZIE, *The Structure of Finite Algebras*, Contemp. Math. 76, AMS, Providence, RI, 1988.
- [16] S. IWATA, *A fully combinatorial algorithm for submodular function minimization*, J. Combin. Theory Ser. B, 84 (2002), pp. 203–212.
- [17] S. IWATA, L. FLEISCHER, AND S. FUJISHIGE, *A combinatorial strongly polynomial algorithm for minimizing submodular functions*, J. ACM, 48 (2001), pp. 761–777.

- [18] P. JIPSEN AND H. ROSE, *Varieties of Lattices*, Lecture Notes in Math. 1533, Springer-Verlag, Berlin, 1992.
- [19] P. JONSSON, M. KLASSON, AND A. KROKHIN, *The approximability of three-valued MAX CSP*, SIAM J. Comput., 35 (2006), pp. 1329–1349.
- [20] A. KROKHIN AND B. LAROSE, *Maximum Constraint Satisfaction on Diamonds*, Technical report CS-RR-408, Department of Computer Science, University of Warwick, Warwick, UK, 2004.
- [21] A. KROKHIN AND B. LAROSE, *Supermodular Sublattices in Products of Relatively Complemented Lattices*, manuscript, 2006.
- [22] V. B. LENDER, *The groupoid of prevarieties of lattices*, Siberian Math. J., 16 (1975), pp. 1214–1223 (in Russian).
- [23] A. MÜLLER AND D. STOYAN, *Comparison Methods for Stochastic Models and Risks*, Wiley Ser. Prob. Stat., Wiley & Sons, Chichester, UK, 2002.
- [24] H. NARAYANAN, *Submodular Functions and Electrical Networks*, Ann. Discrete Math. 54, North-Holland, Amsterdam, 1997.
- [25] S. D. PROMISLOW AND V. R. YOUNG, *Supermodular functions on finite lattices*, Order, 22 (2005), pp. 389–413.
- [26] A. SCHRIJVER, *A combinatorial algorithm minimizing submodular functions in polynomial time*, J. Combin. Theory Ser. B, 80 (2000), pp. 346–355.
- [27] A. SCHRIJVER, *Combinatorial Optimization: Polyhedra and Efficiency, Vol. B, Matroids, Trees, Stable Sets*, Springer-Verlag, Berlin, 2003.
- [28] D. TOPKIS, *Supermodularity and Complementarity*, Frontiers Econom. Res., Princeton University Press, Princeton, NJ, 1998.

ON PLANAR QUASI-PARITY GRAPHS*

CLÁUDIA LINHARES SALES[†], FRÉDÉRIC MAFFRAY[‡], AND BRUCE REED[§]

Abstract. A graph G is strict quasi parity (SQP) if every induced subgraph of G that is not a clique contains a pair of vertices with no odd chordless path between them (an even pair). Hougardy conjectured that the minimal forbidden subgraphs for the class of SQP graphs are the odd chordless cycles, the complements of odd or even chordless cycles, and some line-graphs of bipartite graphs. Here we prove this conjecture for planar graphs. We also give a constructive characterization of all the planar minimal forbidden subgraphs for the class of SQP graphs.

Key words. perfect graphs, even pairs, planar graphs

AMS subject classifications. 68R10, 68Q25, 05C85, 05C17, 90C27

DOI. 10.1137/060672649

1. Introduction. A graph G is *perfect* if, for every induced subgraph H of G , the chromatic number of H is equal to the size of its largest clique. A *hole* is any chordless cycle of length at least five. A hole is odd or even according to its length. The most famous problem in the context of perfect graphs was Berge’s perfect graph conjecture: *A graph G is perfect if and only if neither G nor $\bar{G}$ contains an odd hole.* This was proved by Chudnovsky et al. [1]. However, other problems concerning perfect graphs remain open, some of which are related to the concept of an *even pair*. Two nonadjacent vertices of a graph G form an *even pair* (resp., *odd pair*) if every chordless path between them has an even (resp., odd) number of edges. A graph is *strict quasi parity* (SQP) if each of its induced subgraphs either contains an even pair or is a clique. Fonlupt and Uhry [3] introduced even pairs and showed that the graph obtained by contracting an even pair in a perfect graph is perfect. Meyniel also showed that no minimal imperfect graph contains an even pair [11]. He proposed the definition of SQP graphs, which, by his result, are perfect. The main interest of this class lies in the fact that many classical families of perfect graphs are SQP, as proved by Meyniel (see also the more recent survey [2]). Meyniel also suspected that the minimal non-SQP graphs must have a simple structural characterization. More formally, Hougardy [5] proposed the following conjecture.

CONJECTURE 1. *Every minimal non-SQP graph is either an odd hole, or the complement of a hole, or the line graph of a bipartite graph.*

Let us call *obstruction* any graph G which is not a clique, has no even pair, and is such that every proper induced subgraph of G has an even pair or is a clique. We will prove Hougardy’s conjecture for planar graphs; namely, we will prove the following theorem.

*Received by the editors October 18, 2006; accepted for publication (in revised form) August 28, 2007; published electronically February 29, 2008.

<http://www.siam.org/journals/sidma/22-1/67264.html>

[†]DC-LIA, Universidade Federal do Ceará, Fortaleza, CE, Brazil (linhares@lia.ufc.br). This author was partially supported by the Brazilian agency CNPq. This work was done when this author was a Ph.D. student at LSD2-IMAG, Grenoble, France.

[‡]CNRS, Laboratoire G-SCOP, 46 Avenue Félix Viallet, 38031 Grenoble Cedex, France (frederic.maffray@g-scop.inpg.fr).

[§]School of Computer Science, McGill University, 3480 University, Montreal, QC, H3A 2A7 Canada (breed@cs.mcgill.ca).

THEOREM 1 (main theorem). *Every planar obstruction is an odd hole or the line-graph of a bipartite graph.*

We say that the neighborhood of a vertex is *correct* if it consists of either (a) three vertices with exactly one edge between them, or (b) four vertices with exactly two, nonincident, edges between them. To prove the main theorem it suffices to establish the following.

THEOREM 2. *In a planar obstruction that is not an odd hole, every neighborhood is correct.*

Let us recall a classical result which we will use several times. Call *diamond* the graph that consists of a clique of size four minus one edge. Call *claw* the graph on four vertices with one vertex of degree three and three vertices of degree one.

THEOREM 3 (see [4]). *A graph G is the line-graph of a simple bipartite graph if and only if G contains no odd hole, no diamond, and no claw.*

Proof of Theorem 1 from Theorem 2. Let G be a planar obstruction different from an odd hole. By Theorem 2, every vertex belongs to exactly two maximal cliques, one of size three and the other of size three or two. Hence G contains no claw and no diamond. Thus G is the line-graph of a bipartite graph by Theorem 3.

The proof of Theorem 2 is the object of the next sections. It will be convenient to first prove it for 3-connected obstructions (Theorem 4), and then to derive it for all obstructions via some decomposition technique. In section 4 we will give a more precise description of planar obstructions.

For an introduction to planar graphs, see [12]. We finish this section with a few definitions and some simple lemmas.

We will frequently use “sees” and “misses” instead of “is adjacent to” and “is not adjacent to.” The subgraph of a graph G induced by a set $X \subseteq V(G)$ is denoted by $G[X]$. For vertices $x, y \in V(G)$, an x, y -path is a path whose endvertices are x and y . If P is a path and u, v are vertices in P , we let $P[u, v]$ denote the subpath of P between u and v . If two paths P, Q have a common endvertex, we use $P + Q$ to denote the path obtained by their concatenation. For any vertex x we let $N(x)$ denote the neighborhood of x . For a set $X \subseteq V(G)$, we let $N(X) = \{u \in V(G) - X \mid u \text{ sees some } x \in X\}$.

LEMMA 1. *Let G be a graph that contains no odd hole. Then*

1. *if P is a chordless path in G and x is a vertex of $G - P$ that sees the two endvertices of P , then P has even length if and only if it contains an even number of edges of $N(x)$.*
2. *if H is a hole in G and x is a vertex of $G - H$ that sees at least two nonadjacent vertices of H , then H contains an even number of edges of $N(x)$.*

Proof. If 1 fails, it is a routine matter to check that $P + x$ contains an odd hole. To prove 2, consider two nonadjacent vertices u, v of H that are neighbors of x and apply 1 to the two u, v -paths of H . $\square$

LEMMA 2. *An obstruction contains no clique cutset.*

Proof. Suppose that G is an obstruction which has a clique cutset C . First, assume that for some component B of $G - C$ the induced subgraph $G[B \cup C]$ is not a clique. So $G[B \cup C]$ has an even pair a, b . No chordless path between a and b can go through another component of $G - C$ because C is a clique; it follows that a, b is an even pair of G , a contradiction. Now assume that $G[B \cup C]$ is a clique for every component B of $G - C$. Pick any a, b in two different components of $G - C$: clearly a, b is an even pair of G , a contradiction. $\square$

LEMMA 3. *Let x, y be nonadjacent vertices in a minimal cutset C in a graph that contains no odd hole. Then all chordless x, y -paths in $G - (C - \{x, y\})$ have the same parity.*

Proof. Since C is a minimal cutset, every component D of $G - C$ contains a neighbor of x and a neighbor of y , and so there is a chordless x, y -path with interior in D . If the lemma fails, there must exist an even chordless x, y -path P whose interior lies in one component of $G - C$ and an odd chordless x, y -path Q whose interior lies in another component of $G - C$, and then $P + Q$ is an odd hole, a contradiction. $\square$

LEMMA 4. *In an obstruction different from an odd hole, every vertex has degree at least three and lies in a triangle.*

Proof. Let x be any vertex in such an obstruction G . If x either has degree one, or has degree two and its neighbors are adjacent, then $N(x)$ is a clique cutset, contradicting Lemma 2. Suppose x has degree at least two and its neighbors are pairwise nonadjacent. Let u, v be any two neighbors of x . If P is an odd chordless u, v -path in G , then $P + x$ contains an odd hole by Lemma 1. So there is no such path; but then u, v form an even pair, a contradiction. $\square$

2. 3-connected planar obstructions. Here we prove Theorem 2 for 3-connected planar obstructions.

THEOREM 4. *In a 3-connected planar obstruction every neighborhood is correct.*

Proof. Let G be a 3-connected planar obstruction, and let us fix one planar representation of G . We shall arrive at the desired conclusion of Theorem 4 through a succession of lemmas that gradually restrict the possible types of neighborhood in G . $\square$

LEMMA 5. *In G the neighborhood of every vertex induces either a chordless cycle or a collection of vertex-disjoint paths.*

Proof. Let x be any vertex of G . If $N(x)$ contains a vertex y of degree three in $N(x)$, then planarity implies that x and y are in a triangle cutset of G , contradicting Lemma 2. Thus the maximum degree in $N(x)$ is two. Moreover, if $N(x)$ properly contains a set S of vertices which induce a cycle, then planarity implies that $S \cup \{x\}$ contains a triangle cutset separating $N(x) - S$ from $V(G) - (N(x) \cup \{x\})$, again a contradiction. Lemma 5 follows. $\square$

LEMMA 6. *G has no vertex x such that $N(x) \cup \{x\}$ induces a diamond.*

Proof. Suppose on the contrary that there exists such a vertex x , and let y, z be the nonadjacent vertices in $N(x)$. Since G is an obstruction, there is an odd chordless path P between y and z . But then $P + x$ is an odd hole, a contradiction. $\square$

A *star-cutset* is a cutset C such that some vertex $x \in C$ sees all vertices of $C - x$.

LEMMA 7. *G has no star-cutset.*

Proof. Suppose that G contains a star*-cutset C , and let x be a vertex of C that sees all of $C - x$. We claim that each component K of $G - C$ has a nonneighbor of x . Suppose the contrary, that is, $K \subseteq N(x)$. By Lemma 5, K induces a cycle or a path. Let y be any vertex of K . Then $N(y) \cup \{y\}$ induces a diamond, a contradiction to Lemma 6. So our claim holds, and consequently we can assume that $C = \{x\} \cup N(x)$. As G is 3-connected, each component of $G - C$ is adjacent to at least three vertices of C and hence three vertices of $N(x)$. We can enumerate the neighbors of x as $x_1, \dots, x_d$ clockwise around x and choose a component K of $G - C$ with $x_1 \in N(K), x_j \in N(K), N(K) \subseteq \{x_1, \dots, x_j\}$, and such that j is as small as possible over all clockwise enumerations and all choices of K . This choice, the fact that $|N(K)| \geq 3$, and planarity imply that any other component L of $G - C$ satisfies $N_L \subseteq \{x_j, x_{j+1}, \dots, x_d, x_1\}$. It follows that $\{x, x_1, x_j\}$ is a cutset of G . Then $x_1 x_j$ is not an edge of G , for otherwise G has a clique cutset $\{x, x_1, x_j\}$. Furthermore, $\{x, x_1, x_j\}$ is a minimal cutset since G is 3-connected. Thus for every component L of $G - \{x, x_1, x_j\}$ there is a chordless x_1, x_j -path in $L \cup \{x_1, x_j\}$. Such paths must

all have the same parity, for otherwise we could find two paths of different parity in different components of $G - \{x, x_1, x_j\}$, and their union would be an odd hole in G . Actually all these paths are odd, for otherwise $\{x_1, x_j\}$ would be an even pair. There is a chordless x_1, x_j -path P in K since $\{x_1, x_j\} \subseteq N(K)$, and by the above remark, P is odd. Now, $P + x$ is an odd hole, a contradiction. This completes the proof of Lemma 7. $\square$

LEMMA 8. *If G has a cutset C of size three which induces a subgraph with just one edge, then there is a vertex w with $N(w) = C$.*

Proof. Assume that $C = \{x, y, z\}$ with $xy \in E(G)$ and $yz, xz \notin E(G)$. Since G is 3-connected, C is a minimal cutset. By Lemma 3, all chordless x, z -paths in $G - y$ have the same parity p_{xz} and all chordless y, z -paths in $G - x$ have the same parity p_{yz} . If p_{xz} is even and p_{yz} is odd (or vice versa), then x, z (or y, z) form an even pair, a contradiction. So p_{xz} and p_{yz} are the same. Since G has no star-cutset, there is a chordless x, z -path P_1 in $G - (N(y) \cup \{y\})$. The interior of P_1 is in some component K_1 of $G - C$. Similarly, there is a y, z -path P_2 in $G - (N(x) \cup \{x\})$. The interior of P_2 is also in K_1 , for otherwise $P_1 + xy + P_2$ forms an odd hole. Let K be any component of $G - C - K_1$, and let P be any chordless y, z -path with interior in K . Call w the neighbor of y along P . Let w' be the last neighbor of x along P , starting from y . If $w' \neq w$, then one of the paths $xw' + P[w', z] + P_1$ and $xw' + P[w', z] + P_2 + yx$ is chordless and odd, a contradiction. It follows that x sees w and no other vertex of $P - y$. Thus the set $W = N(x) \cap N(y) \cap K$ separates x and y from z in $K \cup \{x, y, z\}$. Suppose that W has two vertices w, w' . Let P_3 (resp., P_4) be a w, z -path (w', z -path) in $G \setminus (K \cup \{x, y\})$. If w, w' are adjacent, then it is easy to see that the subgraph induced by $P_1 \cup P_2 \cup P_3 \cup P_4$ contains a subdivision of K_5 , which contradicts the planarity. If w, w' are not adjacent, then, since they do not form an even pair, there is an odd w, w' -path Q in G , and any such path must be in $G - (K \cup \{x, y\})$. But then it is easy to see that the subgraph induced by $P_1 \cup P_2 \cup P_3 \cup P_4 \cup Q$ contains a subdivision of K_5 or of $K_{3,3}$, which contradicts planarity. So $|W| \leq 1$, say $W = \{w\}$. Now $K = \{w\}$; otherwise $\{w, z\}$ would be a cutset of size two in G . So $N(w) = \{x, y, z\}$, which completes the proof of Lemma 8. $\square$

LEMMA 9. *For every vertex x of degree three, the neighborhood of x contains exactly one edge.*

Proof. Clearly, $N(x)$ does not induce a triangle K_3 , for otherwise either $N(x)$ is a clique cutset or G is K_4 , both of which are impossible. Lemma 6 implies that $N(x)$ does not induce a path of length two. On the other hand, $N(x)$ must contain an edge, for otherwise any two vertices in $N(x)$ would form an even pair. This proves the lemma. $\square$

LEMMA 10. *Every face of G is a hole.*

Proof. Every face of G is a cycle because G is 2-connected. If a face F has a chord xy , planarity implies that $\{x, y\}$ is a clique cutset of G that separates one component of $F - \{x, y\}$ from the other, a contradiction to the fact that G is 3-connected. $\square$

LEMMA 11. *Any two intersecting faces in G intersect in an edge or a vertex.*

Proof. Suppose that two faces F, F' intersect in at least two nonadjacent vertices. Then it is easy to see that there are two vertices of $F \cap F'$ that form a cutset, a contradiction to the fact that G is 3-connected. $\square$

For any vertex x of G , we denote by F_x the face of $G - x$ such that x lies in the interior of F_x . Note that F_x is a cycle since $G - x$ is 2-connected.

LEMMA 12. *For every vertex x , if yz is an edge between two vertices of F_x that are not consecutive along F_x , then x misses y and z and there is a vertex w of F_x such that $N(w) = \{x, y, z\}$.*

Proof. In the opposite case, $\{x, y, z\}$ would be either a clique cutset or a cutset, contradicting Lemma 8. $\square$

LEMMA 13. *For every vertex x of degree at least four, F_x is a hole.*

Proof. Consider any such vertex x and assume that F_x is not chordless. By Lemma 12, there is a vertex w on F_x with $N(w) = \{x, y, z\}$, $y, z \in F_x$, edges $yz, yw, wz \in E(G)$, and $xy, xz \notin E(G)$. Let F_1 and F_2 be the two faces of G containing the edge wx . By Lemma 10, F_1 and F_2 both induce holes of G . Since each of F_1 and F_2 contains one of y or z , neither is a triangle and hence both are even. By Lemma 11, $F_1 \cap F_2 = \{w, x\}$. By Lemma 12 and the fact that $|N(x)| \geq 4$, there is no edge from $F_1 - \{w, x\}$ to $F_2 - \{w, x\}$. But now $F_1 + F_2 - w$ is an odd hole in G , a contradiction. $\square$

LEMMA 14. *For every vertex x of degree at least four in G , $N(x)$ induces a graph with an even number of edges.*

Proof. By Lemma 13, F_x is a hole. If $N(x)$ induces a disjoint set of paths with an odd number of edges, then $F_x \cup \{x\}$ contains an odd hole by Lemma 1, a contradiction. $\square$

LEMMA 15. *If G has a diamond on vertices $\{a, b, c, d\}$ such that $ad \notin E(G)$, then there is no path from a to d in $G - [N(b) \cup N(c) - \{a, d\}]$.*

Proof. Since $\{a, d\}$ is not an even pair there exists an odd chordless a, d -path in $G - b - c$. Choose such a path P_1 with $P_1 \cap (N(b) \cup N(c))$ minimal over all such paths. Suppose indirectly that there exists a chordless a, d -path P_2 in $G - [N(b) \cup N(c) - \{a, d\}]$. Let $x \in P_1 - a - d$ be such that $xb \in E(G)$; such an x exists or else $P_1 + b$ induces an odd hole. Let P_3, P_4 be the two subpaths of $P_1 - x$ such that $a \in P_3$ and $d \in P_4$. Clearly there is a chordless path from P_3 to P_4 whose interior vertices are in $P_2 - P_1$. Choose a shortest such path P_5 so that its endpoint f in P_3 is as close as possible to a along P_3 and its endpoint g in P_4 is as close as possible to d along P_4 . Then the path $P_6 = P_3[a, f] + P_5 + P_4[g, d]$ is a chordless path. Now, as $P_5 - f - g \subseteq P_2$ we know that $(P_5 - f - g) \cap (N(b) \cup N(c)) = \emptyset$. Observe that P_6 contains strictly fewer vertices of $N(b) \cup N(c)$ than P_1 . Thus, by the minimality in our choice of P_1 , P_6 is even. It follows that the subpath $P_1[f, g]$ is of different parity from P_5 . Note that by planarity $(P_1[f, g] - f - g) \cap N(c) = \emptyset$. Let r (resp., s) be the neighbor of c closest to f (resp., to g) along P_3 (resp., P_4). Then either $c + P_1[r, s]$ or $c + P_1[r, s] - P_1[f, g] + P_5$ is an odd hole. This contradiction completes the proof of Lemma 15. $\square$

LEMMA 16. *G contains no diamond.*

Proof. Suppose on the contrary that $\{a, b, c, d\}$ is a diamond in G , with $ad \notin E(G)$.

First, suppose that $N(a) \subseteq N(b) \cup N(c)$. Let $y_1, \dots, y_k$ ($k \geq 3$) be the neighbors of a clockwise in the planar representation of G , with $b = y_1$. If c is none of y_2, y_k , then $\{a, b, c\}$ is a triangle cutset, a contradiction to Lemma 2. So we may assume that $c = y_k$. Then b misses every y_i with $3 \leq i \leq k - 1$, for otherwise $\{a, b, y_i\}$ is a triangle cutset, again a contradiction. Likewise c misses y_j with $2 \leq j \leq k - 2$. It follows that $k \leq 4$. If $k = 3$, then Lemma 9 is contradicted by a . So $k = 4$, b sees y_2 , c sees y_3 , and by Lemma 14, $N(a)$ is the 4-cycle $b-y_2-y_3-c$. Since y_2, y_3 must not form an even pair, there is an odd path P between them; but then $P + a$ is an odd hole, a contradiction. Therefore $N(a) \not\subseteq N(b) \cup N(c)$.

Consider any $y \in N(a) - (N(b) \cup N(c))$, and let K be the component of $G - (N(b) \cup N(c))$ that contains y . By Lemma 15, no vertex of K sees d . By Lemma 7, $N(c) \cup \{c\}$ is not a cutset, and so there is a vertex in $N(b) - N(c)$ adjacent to K . Likewise there is a vertex in $N(c) - N(b)$ adjacent to K . By planarity, we can choose $e \in N(b) - N(c)$ and $f \in N(c) - N(b)$ such that both e and f see elements of K

and $\{e, b, c, f\}$ is a cutset separating K from d . (More precisely, if $z_1, \dots, z_k$ are the neighbors of b clockwise in the planar representation of G , with $z_1 = d, z_2 = c, z_3 = a$, then e is the vertex z_j that has a neighbor in K and such that j is the largest such integer. Vertex f is defined similarly.) There is a chordless e, f -path P_1 with interior in K by definition. This path must be odd or else $P_1 + b + c$ would be an odd hole.

We claim that $S = \{e, b, c, f\}$ is a minimal cutset. Suppose that some proper subset S' of S is a cutset. We have $|S'| = 3$ since G is 3-connected. Lemma 7 implies that S' is not $S - e$ or $S - f$. Thus, and by symmetry, S' is $S - b = \{e, c, f\}$. Then Lemmas 7 and 8 imply that e does not see f and that there is a vertex w with $N(w) = \{e, c, f\}$. But then $P_1 + w$ form an odd hole. So the claim holds.

Now, by Lemma 3, all chordless e, f -paths in $G - b - c$ have the same parity, which is odd since P_1 is one such path. By the same lemma, all chordless b, f -paths in $G - c - e$ have the same parity, which is odd, for otherwise $\{b, f\}$ would be an even pair. Similarly all chordless c, e -paths in $G - b - f$ are odd.

Suppose that e and f have no common neighbor in $G - K - a$. Since G is 3-connected, there exist paths from d to $\{e, f\}$ in $G - (K \cup \{a, b, c\})$. Consider a shortest one P_2 ; it can be assumed by symmetry that P_2 is from d to e . Note that there are no neighbors of f in $P_2 - e$ by the minimality of P_2 and the fact that e and f have no common neighbor in $G - K - a$. The path $P_2 + c$ contains a shortest c, e -path P'_2 , and P'_2 is odd as mentioned above. But now $P_1 + P'_2$ induces an odd hole (even if $P_1 = ef$), a contradiction. So e and f have a common neighbor in $G - K - a$. It follows that $ef \in E(G)$ since every chordless e, f -path in $G - b - c$ must be odd.

Let g be the common neighbor of e and f such that the cycle $gebef$ bounds a region of the plane containing d and no other common neighbor of e and f . If there exists a path from d to $\{e, f\}$ in $G - \{b, c, g\}$, then, as in the above case where e and f had no common neighbor, we obtain a contradiction. So $\{b, c, g\}$ is a cutset, and Lemma 8 implies $N(d) = \{b, c, g\}$. But then b, c, d, e, f, g induce a proper subgraph of G which is a $\bar{C}_6$, a contradiction. This completes the proof of Lemma 16. $\square$

LEMMA 17. *G contains no vertex of degree five or greater.*

Proof. Suppose that x is a vertex of degree at least five in G . By Lemmas 13, 14, and 16, F_x is an even hole, it contains an even number of edges of $N(x)$, and no three consecutive vertices of F_x are neighbors of x . We claim that

- (1) *no vertex of $G - x - F_x$ sees two nonadjacent vertices of F_x .*

Assume on the contrary that there exists a triple $\{a, s, t\}$ such that a is in $G - x - F_x$ and s, t are nonadjacent neighbors of a in F_x . Let P_1 and P_2 be the two chordless s, t -paths in F_x , where the length of P_1 is less than or equal to the length of P_2 . We choose a triple $\{a, s, t\}$ that minimizes the length of P_1 over all appropriate choices. Path P_1 has even length, for otherwise either $P_1 + a$ would be an odd hole or the minimality of P_1 would be contradicted. Then P_2 is also even since F_x is an even hole. Note that $\{a, s, t, x\}$ is a cutset that separates $P_1 - s - t$ from $P_2 - s - t$. For $i = 1, 2$, call K_i the component of $G - \{a, s, t, x\}$ that contains $P_i - s - t$. There is a neighbor of x on $P_1 - s - t$, or else $\{a, s, t\}$ would be a cutset contradicting Lemma 7. Call u (resp., v) the neighbor of x in $P_1 - s - t$ that is closest to s (resp., to t). Consider the chordless paths $P_1^s = P_1[s, u] + ux$ and $P_1^t = P_1[t, v] + vx$. Note that these two paths have their interiors in K_1 . Likewise there is a neighbor of x on $P_2 - s - t$. Call y (resp., z) the neighbor of x that is closest to s (resp., t) on $P_2 - s - t$. Consider the chordless paths $P_2^s = P_2[s, y] + yx$ and $P_2^t = P_2[t, z] + zx$. Note that these two paths have their interior in K_2 .

First suppose that either (a) $u = v$ or (b) u, v are different and not adjacent. In case (a), clearly P_1^s and P_1^t have the same parity p since P_1 has even length. In case (b), $P_1^s + sa + at + P_1^t$ is a hole, which is even, so P_1^s and P_1^t again have the same parity p . In either case, both P_2^s, P_2^t also have parity p , for otherwise either $P_2^s + P_1^s$ or $P_2^t + P_1^t$ would be an odd hole. Since $\{s, t\}$ is not an even pair, there is an odd chordless s, t -path Q in G . Clearly Q does not use a . If Q also does not use x , then its interior lies in some component of $G - \{a, s, t, x\}$ and so one of $Q + P_1$ or $Q + P_2$ is an odd hole, a contradiction. Thus Q contains x . So Q consists of a chordless s, x -path Q' in $G - a - t$ plus a chordless x, t -path Q'' in $G - a - s$, and the paths Q', Q'' have different parities. Note that each of Q', Q'' has its interior in some component of $G - \{a, s, t, x\}$. Up to symmetry, Q' is odd and Q'' is even, and so x does not see t . If the parity p is even, then x does not see s , and at least one of $Q' + P_1^s$ or $Q' + P_2^s$ is an odd hole, a contradiction. If p is odd, then at least one of $Q'' + P_1^t$ or $Q'' + P_2^t$ is an odd hole, a contradiction.

Now suppose that u, v are different and adjacent. Then P_1^s and P_1^t have different parities. By Lemma 14, P_2 has an odd number of edges of $N(x)$. Call s' and t' , respectively, the neighbor of s and t on P_2 . Vertex a sees no vertex of $P_2 - \{s, s', t, t'\}$, for otherwise there would exist a chordless a, x -path R in $a + x + P_2 - \{s, s', t, t'\}$ and then either $R + as + P_1^s$ or $R + at + P_1^t$ would be an odd hole. Moreover, if a sees one of s', t' , then it sees both, or else $F_x + a$ would contain an odd hole by Lemma 1 applied to P_2 and a . Suppose that a sees both s', t' . Let $F'_x = F_x + a - P_1$. So F'_x is a hole and must be even. Since G contains no diamond, x misses both s', t' . So x sees at least three vertices of F'_x (since x has degree at least five), and F'_x contains an odd number of edges of $N(x)$, so $x + F'_x$ contains an odd hole by Lemma 1, a contradiction. Thus a misses both s', t' . Call P'_1 the interior of P_1 and let $F'_x = F_x + a - P'_1$. A similar argument shows that $x + F'_x$ contains an odd hole, a contradiction. Thus (1) holds.

Let $E_x = \{e_0, \dots, e_{k-1}\}$ be the set of edges in $N(x)$, with $e_i = u_i v_i$. Lemmas 1, 4, 14, and 16 imply that E_x is a matching of even size $k \geq 2$, so we can enumerate the components of $G[F_x] - E_x$ as $C_0, \dots, C_{k-1}$ so that for $0 \leq i \leq k-1$ the endpoints of C_i are v_{i-1} and u_i (modulo k). By Lemma 1, each C_i is an even path, so it has a nonempty interior. Let W be the set of vertices of $G - x - F_x$ that have a neighbor in F_x . By Lemma 16 and (1), for every vertex $w \in W$ the set $N(w) \cap F_x$ consists of one or two consecutive vertices of some C_i . Thus W can be partitioned into $W_0, \dots, W_{k-1}$, where $W_i = \{w \in W \mid N(w) \cap F_x \subset C_i\}$. Let $W_i^* = \{w \in W_i \mid w \text{ has a neighbor in the interior of } C_i\}$, and $X_i = W_i - W_i^*$ (so every vertex $w \in X_i$ satisfies $N(w) \cap F_x = \{v_{i-1}\}$ or $\{u_i\}$). For $0 \leq i, j \leq k-1$ with $i \neq j$, call (i, j) -connection any path $P = f \cdots g$ in $G - x - F_x$ such that $f \in W_i, g \in W_j$, there are nonadjacent vertices $y \in C_i \cap N(f)$ and $z \in C_j \cap N(g)$, and the interior vertices of P are in $G - (W_0^* \cup \dots \cup W_{k-1}^*)$. Say that an (i, j) -connection is *pure* if its interior vertices are in $G - (W - X_i - X_j)$. We claim that

(2) *there exists a pure $(i, i+1)$ -connection for some i .*

By Lemma 7, the graph $G - \{x, u_0, v_0, \dots, u_{k-1}, v_{k-1}\}$ is connected, which implies that for each $i = 0, \dots, k-1$ there exists an (i, j) -connection for some j .

We remark that, for every $i \neq j$, every (i, j) -connection $P = f \cdots g$ contains a pure (i', j') -connection for some i', j' . We prove this remark by induction on the length of P . Suppose that P itself is not pure. So there is a vertex $h \in (P - f - g) \cap X_l$ for some $l \neq i, j$. Pick any $y \in F_x \cap N(f), z \in F_x \cap N(g), t \in F_x \cap N(h)$. It cannot be that t sees both y, z , so either $P[f, h]$ or $P[h, g]$ is a connection, which, by induction,

contains a pure connection. So the remark is proved, and in particular there exists a pure connection.

Let $P = f \cdots g$ be a pure (i, j) -connection. Let P' be any path of F_x between vertices $y \in F_x \cap N(f)$, $z \in F_x \cap N(g)$, and let R be the region of the plane that is bounded by $P \cup P'$ and does not contain x . We choose i, j, P, P' such that the length of P' is minimized. If $|j - i| = 1$, we are done. Otherwise, there is some l such that $e_{l-1} \cup C_l \cup e_l \subseteq P'$. We know that there exists an (l, m) -connection P'' for some m . By the planarity of G , either P'' lies entirely inside R , and then every pure connection contained in P'' contradicts the choice of P' ; or P'' intersects P , and then $P \cup P''$ contains an (l, i) - or an (l, j) -connection, which contains a pure connection that contradicts the minimality of P' . So (2) holds.

By (2) and up to symmetry we can assume that there exists a pure $(1, 2)$ -connection. We choose a shortest one $P = f \cdots g$, with $f \in W_1$ and $g \in W_2$. Let y (resp., y') be the neighbor of f on C_1 closest to v_0 (resp., to u_1). Recall that y, y' are either equal or adjacent. Likewise let z (resp., z') be the neighbor of g along C_2 closest to u_2 (resp., to v_1). Let Q be the y, z -path of F_x that does not contain $u_1 v_1$, and let Q' be the y', z' -path of F_x that contains $u_1 v_1$. Suppose $|E_x| \geq 4$. Then no interior vertex h of P sees v_0 or u_2 , for otherwise $P[h, g]$ or $P[f, h]$ would contradict the choice of P . Thus $P \cup Q$ induces a hole, which contains $|E_x| - 1$ edges of E_x ; so, by Lemma 1, G contains an odd hole, a contradiction. So $|E_x| = 2$. Now Q and Q' play a symmetric role. If no interior vertex of P sees any of u_0, v_0, u_1, v_1 , then $P \cup Q$ and $P \cup Q'$ induce two holes, each of them contains one edge of E_x , and at least one of them contains three neighbors of x ; so, by Lemma 1, G contains an odd hole, a contradiction. So we may assume, up to symmetry, that some interior vertex h of P is adjacent to u_1 . Then $z = v_1$, for otherwise $P[h, g]$ would contradict the choice of P ; no vertex g' of $P - g$ is adjacent to v_1 , for otherwise $P[f, g']$ would contradict the choice of P ; no vertex f' of $P - f$ is adjacent to v_0 , for otherwise $P[f', g]$ would contradict the choice of P ; and no vertex d of P is adjacent to u_0 , for otherwise $P[d, h]$ would contradict the choice of P . Thus $P \cup Q$ induces a hole, which contains one edge of E_x and at least three neighbors (v_1, u_0, v_0) of x , and so by Lemma 1, G contains an odd hole, a contradiction. This completes the proof of Lemma 17. $\square$

Now, since G is 3-connected, it has minimum degree three. This and the above lemmas yield that every vertex of G has a correct neighborhood, and Theorem 4 is established.

3. Obstructions that have a 2-cutset. In this section we consider obstructions that are not necessarily 3-connected. For this purpose we need some new definitions.

DEFINITION 1 (candidate). *A candidate is any planar graph G such that*

- (c1) *G is the line-graph of a bipartite graph;*
- (c2) *every neighborhood is correct;*
- (c3) *G is 2-connected;*
- (c4) *every 2-cutset is an odd pair.*

LEMMA 18. *Any candidate satisfies the following additional properties:*

- (c5) *If $\{a, b\}$ is a 2-cutset in a candidate G , then $G - \{a, b\}$ has exactly two components;*
- (c6) *a candidate has no clique-cutset;*
- (c7) *in a 3-connected candidate, every 3-cutset C contains at most one edge, and if it contains one edge, then C is the neighborhood of a vertex.*

Proof. If (c5) fails, then the neighborhood of a or b is incorrect.

To prove (c6), assume C is a clique cutset in a candidate G . By (c3) and (c4), C has at least three vertices. By (c2), G has no clique of size four, so $|C| = 3$. Let a be a vertex of C , and let B_1 and B_2 be components of $G - C$. Vertex a must have a neighbor in B_1 and a neighbor in B_2 , or else $C - \{a\}$ would be a 2-cutset violating (c4). But then the neighborhood of a is incorrect, a contradiction.

To prove (c7), consider a 3-cutset $C = \{a, b, c\}$ in a 3-connected candidate G . Let B_1 and B_2 be two components of $G - C$. By (c6) we can assume that a, c are not adjacent. Suppose ab is an edge. Vertex a must have a neighbor x in B_1 , as well as a neighbor y in B_2 , or else $\{b, c\}$ would be a cutset. Since the neighborhood of a is correct, we may assume by symmetry that b sees x and misses y . Vertex b must have a neighbor z in B_2 , or else $\{a, c\}$ would be a cutset separating $B_1 \cup \{b\}$ from B_2 . Then b has no neighbor in $B_1 - \{x\}$, or else the neighborhood of b would be incorrect. The same holds for a . Now if $B_1 - \{x\}$ is not empty, then $\{c, x\}$ is a cutset, a contradiction. So, $B_1 = \{x\}$, and then neighborhood correctness implies that x sees c and c misses b . This proves (c7). $\square$

An edge is called *flat* if it does not lie in a triangle. It is easy to see that an edge ab is flat in a graph G containing no odd hole if and only if the pair $\{a, b\}$ is an odd pair in the graph $G - ab$.

DEFINITION 2 (glueing). Let $M_1 = (V_1, E_1), \dots, M_k = (V_k, E_k)$ be some graphs. Assume that there exist two vertices a, b such that $V_i \cap V_j = \{a, b\}$ for all i, j with $1 \leq i < j \leq k$, where ab is an edge of each M_i . We will say that the graphs are *glueable (along ab)*. Given such graphs as above, we can build a graph with vertex-set $V_1 \cup \dots \cup V_k$ and edge-set $E_1 \cup \dots \cup E_k - \{ab\}$. This construction will be called *glueing the graphs $M_1, \dots, M_k$ along edge ab* .

DEFINITION 3 (unglueing). Let G be a graph with a 2-cutset $\{a, b\}$, where a and b are nonadjacent, and let $B_1, \dots, B_k$ be the components of $G - \{a, b\}$. For each i , we can build a graph with vertex-set $B_i \cup \{a, b\}$ and with all the edges of $G[B_i \cup \{a, b\}]$ plus the edge ab . This will be called *unglueing the graph G along ab* .

It is easy to see that the operations of glueing and unglueing preserve planarity, and that unglueing preserves the absence of odd holes. Conversely, if $M_1, \dots, M_k$ contain no odd hole and ab is a flat edge in each of them, then glueing them along ab produces a graph with no odd hole. Hence they are perfection-preserving for planar graphs. (The fact that they are perfection-preserving for general graphs is also true and is an easy consequence of a lemma of Tucker [14].)

LEMMA 19 (the glueing lemma for candidates).

1. Let M_1 and M_2 be two graphs that are glueable along an edge ab which is a flat edge in each of them. If M_1 and M_2 are candidates, then the graph obtained from glueing them along ab is a candidate.
2. Conversely, let M be a candidate which admits a 2-cutset $\{a, b\}$. Let M_1, M_2 be the two graphs obtained by unglueing M along this cutset. Then each of M_1, M_2 is a candidate.

Proof. To prove the first part, let H_1, H_2 be vertex-disjoint bipartite graphs such that $M_1 = L(H_1)$ and $M_2 = L(H_2)$. Since ab is an edge of M_i ($i = 1, 2$), the edges a and b in H_i are adjacent. We call z_i their common endpoint and a_i, b_i , respectively, the other endpoint of a and b . Vertex z_i has degree two in H_i because ab is a flat edge of M_i . Let H be the bipartite graph obtained from the union of H_1 and H_2 by removing z_1 and z_2 and adding edges a_1a_2 and b_1b_2 . It is a routine matter to check that $M = L(H)$ and that every neighborhood is correct in M . Also, M has no clique cutset Q , for otherwise Q would obviously be a clique cutset of either M_1 or M_2 . Thus (c1), (c2), (c3) hold for M . To check (c4), suppose on the contrary that

$\{x_1, x_2\}$ is an even 2-cutset of G . If x_1, x_2 both lie in some M_i , then $\{x_1, x_2\}$ would be an even 2-cutset of M_i , which is impossible. So we may assume $x_1 \in M_1 - \{a, b\}$ and $x_2 \in M_2 - \{a, b\}$. Let P_i (resp., Q_i) be any chordless x_i, a -path (resp., x_i, b -path) in $M_i - b$ (resp., in $M_i - a$). Note that P_1 and P_2 have the same parity since $P_1 + P_2$ is a chordless x_1, x_2 -path. Likewise Q_1 and Q_2 have the same parity. If P_1 is even and Q_1 is odd, then every chordless x_1, a -path in M_1 is even, which is impossible. It follows that P_1 and Q_1 have the same parity. There must be a chord from P_1 and Q_1 , for otherwise $P_1 + Q_1$ would be an even chordless a, b -path, which is impossible. This chord implies that $(P_1 + P_2 + Q_1 + Q_2) - \{x_1, x_2\}$ is a connected subgraph of $G - \{x_1, x_2\}$ that contains a and b . Thus a, b lie in the same component of $G - \{x_1, x_2\}$, which in turn implies that x_1, x_2 lie in the same component of $G - \{a, b\}$, a possibility that was excluded earlier.

To prove the second part, let M be a candidate which admits a 2-cutset $\{a, b\}$. Suppose that a has degree three and call a' the vertex such that aa' is a flat edge; then both $\{a, b\}$ and $\{a', b\}$ are 2-cutsets of M , so one of them is not an odd pair, contradicting (c3). Thus a and, similarly, b have degree four. It follows that in M_1 as well as in M_2 both a and b have a correct neighborhood (of degree three); note that any other vertex of M_1 and M_2 has the same neighborhood as in M ; so M_1 and M_2 have property (c2). Since the ungluing operation is perfection-preserving, M_1 and M_2 contain no odd hole. It follows that they are line-graphs of bipartite graphs by Theorem 3; i.e., (c1) holds for both of them. Suppose now that M_1 has a cut vertex x . Thus, x is one of a, b (say $x = a$), for otherwise x would be a cut vertex of M , a contradiction. There is a component C of $M_1 - a$ such that $b \notin C$ and there is no path in $M_1 - a$ joining b to any vertex of C . Hence a is also a cut vertex in M (separating b from C), a contradiction to M being a candidate. The same argument works for M_2 . So (c3) holds for M_1 and M_2 . Finally, suppose that $\{x, y\}$ is any 2-cutset of M_1 . Thus $\{x, y\}$ is also a 2-cutset of M , and by (c4) all chordless x, y -paths in M have odd length. Consider a chordless x, y -path P in M_1 . If P does not use the edge ab , then P is a path in M and it is odd by the argument in the preceding sentence. So assume P uses the edge ab . Suppose that x, y do not lie in the same component of $M - \{a, b\}$. This implies that any two vertices from $\{a, b, x, y\}$ form a 2-cutset of M . It is easy to see that at least one of these six cutsets would violate (c4) in M . So x, y must lie in the same component of $M - \{a, b\}$. Let Q be a chordless a, b -path of M whose interior vertices are in the component of $M - \{a, b\}$ that does not contain x and y ; then the path $P - ab + Q$ must be odd, and since Q is odd, this implies that P is odd. So M_1 and M_2 satisfy (c4), and the lemma is proved. $\square$

THEOREM 5. *Every candidate is even pair-free.*

Proof. This theorem is proved by induction on the number of vertices of the candidate. We distinguish between the candidates that are 3-connected and those that are not. For 3-connected planar candidates, the desired result is a consequence of a result of Hsu [8, Theorem 9.2] which we formulate as follows.

LEMMA 20 (Hsu [8]). *Let G be a 3-inseparable planar perfect graph such that the neighborhood of each vertex is correct. Then between any two nonadjacent vertices there are chordless paths of both parities.*

In this lemma, “3-inseparable” is a weaker condition than 3-connected, and it is easy to check that every 3-connected candidate is 3-inseparable. So Lemma 20 gives immediately the even pair-freeness of every 3-connected candidate. Now we look at candidates that are not 3-connected. A property of candidates will be useful.

LEMMA 21. *Let x, y be two nonadjacent vertices in a graph G that satisfies (c2). Let $v_0v_1 \cdots v_kv_{k+1}$ and $w_0w_1 \cdots w_lw_{l+1}$ be internally disjoint chordless paths from*

$x = v_0 = w_0$ to $y = v_{k+1} = w_{l+1}$. Then the only possible chords between these two paths are v_1w_1 and v_kw_l .

Proof. Let v_iw_j be any chord with $1 \leq i \leq k$, $1 \leq j \leq l$. Neighborhood correctness for v_i implies that there is exactly one edge induced by $\{v_{i-1}, v_{i+1}, w_j\}$; since there is no edge $v_{i-1}v_{i+1}$, by symmetry we can assume that w_j sees v_{i-1} and misses v_{i+1} . Now if $j \geq 2$, then neighborhood correctness is violated because w_j has four neighbors and there is no edge $w_{j-1}w_{j+1}$. If $j \leq 1$, the same holds unless $w_{j-1} = v_{i-1}$, i.e., $i = 1$, and so the chord is v_1w_1 . $\square$

LEMMA 22. *Any candidate G that admits a 2-cutset is even pair-free.*

Proof. Let $\{a, b\}$ be a 2-cutset of G . Recall from (c4) that $\{a, b\}$ is an odd pair and from (c5) that $G - \{a, b\}$ has exactly two components B_1, B_2 . Let G_1, G_2 be the two graphs obtained by ungluing G along ab . By the second part of the glueing lemma, G_1 and G_2 are candidates. By the induction hypothesis, G_1 and G_2 are even pair-free.

Consider any pair of nonadjacent vertices x, y in G . If x, y are both in one of G_1, G_2 , say in G_1 , then by the induction hypothesis there exists an odd chordless x, y -path P in G_1 . If P does not use the edge ab , then P is the desired odd chordless x, y -path in G . If P uses the edge ab , then replace in P the edge ab by any chordless a, b -path lying in $B_2 \cup \{a, b\}$. Since a, b is an odd pair the parity is not changed by this replacement and we get an odd chordless x, y -path in G .

Now, assume that x is in B_1 and y is in B_2 . Since G is 2-connected there exist two internally vertex-disjoint paths Q_1 and Q_2 from y to a and b , respectively. Clearly, Q_1 and Q_2 are entirely in $B_2 \cup \{a, b\}$. Call y_1 and y_2 the neighbors of y along Q_1 and Q_2 , respectively. Let R be any (odd) chordless a, b -path entirely in $B_1 \cup \{a, b\}$. By Lemma 21, there is no chord between the chordless y, a -paths Q_1 and $Q_2 + R$ except maybe for the chord y_1y_2 . We claim that we can assume that y_1y_2 is not an edge. Indeed, suppose that y_1 and y_2 are adjacent. By (c4), the set $\{y_1, y_2\}$ is not a cutset of G , so there exists a shortest path Q from y to $Q_1 \cup Q_2 - \{y, y_1, y_2\}$. By symmetry we can assume that the other endpoint of Q is on $Q_1 - \{y, y_1\}$. Then $Q \cup Q_1 - \{y_1\}$ connects y to a and is disjoint from Q_2 , so there exists a shortest path Q' in $Q \cup Q_1$ which connects y to a and is disjoint from Q_2 . The neighbor of y on Q' is not adjacent to y_2 , or else the neighborhood of y would be incorrect. Then we replace the pair of paths Q_1, Q_2 by the pair Q', Q_2 . So the claim holds.

Now, since y_1, y_2 are not adjacent, $Q_1 + Q_2 + R$ is a hole. Since R is odd, we may assume by symmetry that Q_1 is odd and Q_2 is even. By the induction hypothesis, G_1 contains an odd path P from x to b . If P does not use vertex a , then $P + Q_2$ is the desired odd chordless path from x to y . If P uses a , then $P - b$ is an even chordless x, a -path, and so $P - b + Q_1$ is the desired odd path. $\square$

Now the proof of Theorem 5 is complete. $\square$

THEOREM 6. *Every planar obstruction is an odd hole or a candidate.*

Proof. Let G be a planar obstruction different from an odd hole. We prove Theorem 6 by induction on the number of vertices of G . If G is 3-connected, the result holds by Theorem 4. Now assume G is not 3-connected. By Lemma 2, G is 2-connected. Let $\{a, b\}$ be a 2-cutset of G . By Lemma 3 and since G has no even pair, every chordless a, b -path in G is odd. Let $B_1, \dots, B_k$ be the connected components of $G - \{a, b\}$, and put $G_i = G[B_i \cup \{a, b\}] + ab$. Note that G_i is not a clique (because a, b have no common neighbor in G). Let Q_i be any chordless a, b -path in $B_i \cup \{a, b\}$.

We first claim that each G_i has no even pair. Indeed, let x, y be any two nonadjacent vertices in G_i . In G there exists an odd chordless x, y -path P from x to y . If P does not use both a, b , then P lies entirely in G_i and we are done. If P uses both

a, b , then $P[a, b]$ is odd; so $P[x, a] + ab + P[b, y]$ is an odd chordless x, y -path in G_i and we are done again.

Since each G_i is even pair-free, it contains an obstruction G'_i . For each i , both a, b lie in G'_i , or else G'_i would be a proper induced subgraph of G , which is impossible. Moreover, G'_i is not an odd hole, for otherwise replacing the edge ab in this hole by Q_j with $j \neq i$ would produce an odd hole in G , which is impossible. The induction hypothesis (and the fact that each G'_i has strictly fewer vertices than G) implies that each G'_i is a candidate. Note that the edge ab is flat in each G'_i , because a, b have no common neighbor in G . Let G' be the graph obtained by glueing G'_1 and G'_2 along the edge ab . The glueing lemma for candidates implies that G' is a candidate. Theorem 5 implies that G' has no even pair, and so it contains an obstruction G'' . Since G itself is an obstruction, we must actually have $G'' = G' = G$. In particular, G is a candidate. $\square$

Proof of Theorem 2. Let G be a planar obstruction different from an odd hole. Theorem 6 implies that G is a candidate. Thus every neighborhood is correct, and Theorem 2 is established.

LEMMA 23 (the glueing lemma for obstructions).

1. Let M_1 and M_2 be two obstructions that are glueable along an edge ab such that ab is a flat edge in each M_i . Then the graph M obtained by glueing them along ab is an obstruction.

2. Conversely, let M be an obstruction which admits a 2-cutset. Let M_1, M_2 be the two graphs obtained by unglueing M along this cutset. Then each of M_1, M_2 is an obstruction.

Proof. To prove the first part of the lemma, suppose that M is not an obstruction. We know by the glueing lemma for candidates that M is a candidate. So M contains strictly an obstruction M' . Note that M' cannot be entirely in $M_1 - ab$ or in $M_2 - ab$, for otherwise M' would be entirely in either $M_1 - \{a, b\}$ or $M_2 - \{a, b\}$, contradicting the fact that M_1 and M_2 are obstructions, or a or b would be an incorrect vertex in M' . Now $\{a, b\} \cap M'$ is a cutset of M' , and hence $\{a, b\} \subset M'$. Let M'_1 and M'_2 be the graphs obtained by unglueing M' along $\{a, b\}$. We know that M'_1 and M'_2 are candidates. Since M' is a proper subgraph of M , and by symmetry, we may assume that M'_1 is a proper subgraph of M_1 . But this is a contradiction to the fact that M_1 is an obstruction.

For the second part of the lemma, suppose that M_1 is not an obstruction. We know that M_1 is a candidate. So M_1 contains strictly an obstruction M'_1 . If M'_1 does not contain the edge ab , then M'_1 is a subgraph of M , contradicting the fact that M is an obstruction. If M'_1 contains the edge ab , then we can glue M'_1 with M_2 . We obtain a candidate which is a proper subgraph of M , again a contradiction. $\square$

Now we have proved that every planar obstruction is an odd hole or a candidate and that every candidate contains an obstruction, and thus *a planar perfect graph is an SQP graph if and only if it contains no candidate*. However, it is not true that every candidate is an obstruction: a counterexample is the line-graph of $K_{3,3}$. In the next section we study the difference between obstructions and mere candidates.

4. Distinguishing obstructions from candidates. Here we show how to find a subset of vertices of a candidate whose removal yields an obstruction.

DEFINITION 4 (danger). *In a candidate G , we denote by danger any proper subset of vertices D such that each component of $G - D$ is a candidate.*

For example, in the line-graph of $K_{3,3} - e$, two adjacent vertices of degree three form a danger.

The definition entails the following lemma, which is immediate.

LEMMA 24. *A candidate is an obstruction if and only if it contains no danger.*

The question now is how to find a danger in a graph. Given a vertex x , we can define iteratively a subset $D(x)$ as follows. First, remove x from G and let $H = G - x$. Then apply the following procedure:

Step 1. If H has an incorrect vertex y , then remove y and iterate with each component of $H - y$.

Step 2. If H has a cut vertex y , then remove y and iterate with each component of $H - y$.

Step 3. If H has a 2-cutset C that violates (c4) in H , then remove C and iterate with each component of $H - C$.

Step 2 (resp., Step 3) is applied only when there is no possibility of applying Step 1 (resp., Step 2). This procedure is applied as long as possible. We let $D(x)$ denote the set of all removed vertices at the end of the procedure. Clearly, this definition entails immediately that if $D(x)$ does not contain all vertices of G , then $D(x)$ is a danger. Furthermore, we have the next lemma.

LEMMA 25. *Let G be a candidate containing a danger D , and let x be any vertex of D . Then $D(x) \subseteq D$ and $D(x)$ is a danger.*

Proof. We prove that $D(x) \subseteq D$. This will imply that $D(x)$ does not consist of all vertices of G , and so $D(x)$ will be a danger. Consider any subset R of D . We claim that

If $G - R$ has an incorrect vertex y , then y is in D .

For otherwise, y would be an incorrect vertex in $G - D$, contradicting the fact that each component of $G - D$ is a candidate.

If a component H of $G - R$ has only correct vertices and has a cut vertex y , then y must be in D .

For otherwise, let B_1 and B_2 be the components of $H - y$, and let H' be the component of $G - D$ containing y . Clearly H' is a subgraph of H . By neighborhood correctness and by symmetry, we may assume that y has two neighbors in B_1 and one or two neighbors in B_2 . At least one vertex from $N(y) \cap B_1$ and one vertex from $N(y) \cap B_2$ must be in H' , or else y would have an incorrect neighborhood in H' . But then y is a cut vertex in H' , a contradiction.

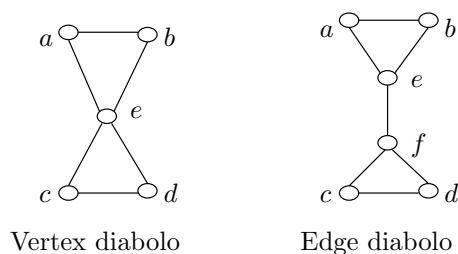
If a component H of $G - R$ has only correct vertices, has no cut vertex, and has a cutset C violating (c4) in H , then $C \subseteq D$.

For otherwise, call B_1 and B_2 the components of $H - C$, assume that some vertex y in C is in $G - D$, and call H' the component of $G - D$ containing y . As above, at least one vertex from $N(y) \cap B_1$ and one vertex from $N(y) \cap B_2$ must be in H' , or else either y would be incorrect or $C - \{y\}$ would be a 1-cutset in H . But then $C \cap H'$ is either a 1-cutset or a 2-cutset violating (c4) in H' , a contradiction.

Now the fact that $D(x) \subseteq D$ follows from the definition of $D(x)$ and from the three points above, applied iteratively, starting with $R = \{x\}$. $\square$

The following lemma is now obvious.

LEMMA 26. *A candidate G is an obstruction if and only if $D(x) = V(G)$ for every vertex x of G .*

FIG. 1. *The two types of diabolo.*

As a consequence, we can decide if a candidate contains a danger, and if it does, we can find a danger efficiently; we have only to build $D(x)$ for each vertex until we find one such that $D(x)$ does not contain all vertices. Notice that the construction of $D(x)$ is polynomial, as it involves only degree and connectivity tests.

Finally, this allows us to find an obstruction in any given candidate G . If G contains no danger D , then G is an obstruction; if not, then find such a danger D and continue with the components of $G - D$, which is a candidate.

4.1. Constructing all candidates. We are going to describe a method for constructing a family $\mathcal{F}$ of graphs starting from $\bar{C}_6$. We will show that $\mathcal{F}$ is exactly the class of all candidates.

Let G be the line-graph of a connected bipartite graph H , where the vertices are colored blue or red along the natural bipartition. Using the one-to-one correspondence between the vertices of H and the maximal cliques of G , we can consider the edges of G colored blue or red accordingly.

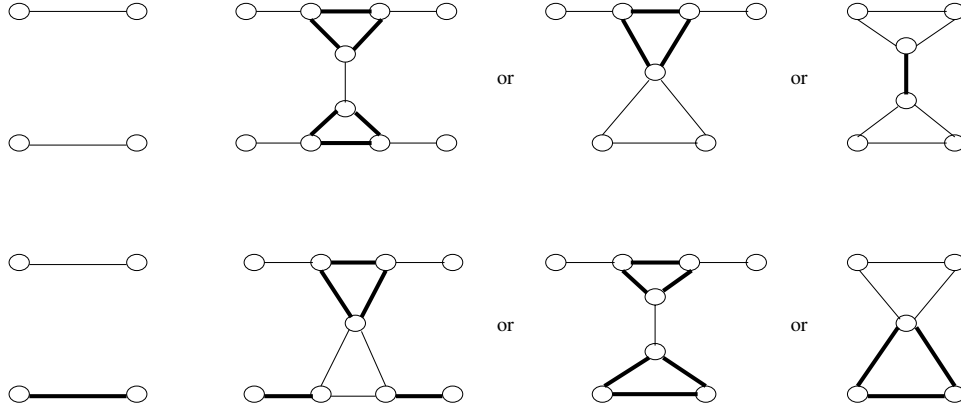
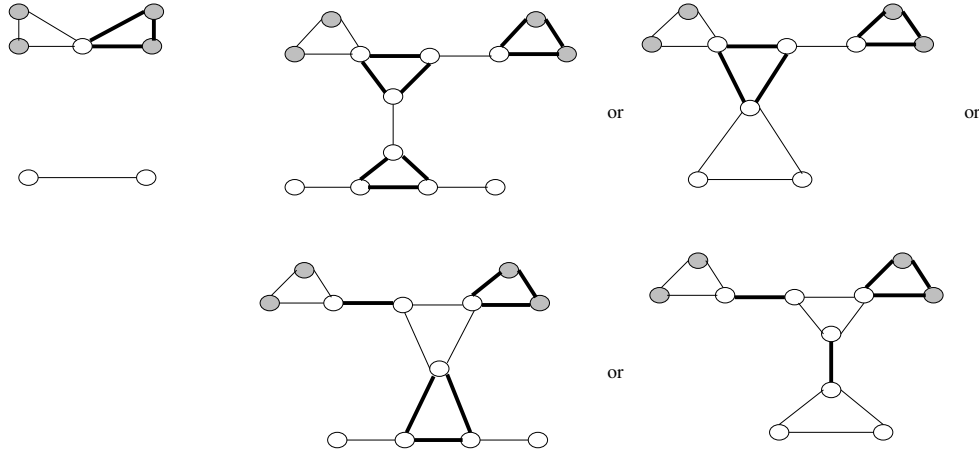
DEFINITION 5 (diabolo). *The left and right graphs depicted in Figure 1 will be called, respectively, vertex-diabolo and edge-diabolo. Edges ab and cd are the bases of the diabolo, and vertex e (and f) is (are) the center(s) of the diabolo.*

DEFINITION 6 (splitting elements). *Let G be a planar graph without an odd hole and where every neighborhood is correct (so G is the line-graph of a bipartite graph). We call a splitting element of G any flat edge and any vertex of degree four. Then a splitting operation is the following:*

1. *Choose two splitting elements on the same face.*
2. *If one of the chosen elements is a vertex f , with neighbors x_1, x_2, x_3, x_4 such that x_1x_2 and x_3x_4 are edges, then replace it by a chordless path $f_1f_2f_3$ such that f_2 has degree two, f_1 is linked to x_1 and x_2 , and f_3 is linked to x_3 and x_4 . Edges f_1f_2 and f_2f_3 are called new.*
3. *If one of the chosen elements is an edge, then either call it new, or subdivide it into a path of length three (in this case, only the middle edge is called new).*
4. *A diabolo is introduced whose two bases are two new edges, one from each splitting element. The type of diabolo (edge or vertex) will be chosen so as to respect the blue/red bipartition of maximal cliques.*

Figures 2–4 illustrate the splitting operations, which allow us to create new planar graphs from old.

We can also view the splitting operations on the bipartite graph H , of which G is the line-graph, rather than on G . For example, in the first row in Figure 2, the hypothesis means that in H there exists a vertex x_{ab} of degree two incident with the edges a and b , there exists a vertex x_{cd} of degree two incident with the edges c and d , and the vertices x_{ab}, x_{cd} are on the same side of the bipartition of H . (In the

FIG. 2. *Edge-edge operations.*FIG. 3. *Edge-vertex operations.*

second row of Figure 2, the vertices x_{ab}, x_{cd} are on different sides of the bipartition.) For the first transformation illustrated in the first row of Figure 2, in H split the vertex x_{ab} into two vertices x_a and x_b to be the new endpoints of the edges a and b , respectively; likewise split x_{cd} into x_c and x_d . Add new vertices y_{ab}, y_{cd}, x and edges $x_a y_{ab}, x_b y_{ab}, x_c y_{cd}, x_d y_{cd}, x y_{ab}, x y_{cd}$. It is easy to check that this transformation on H corresponds exactly to the transformation on G . A similar transformation on H can be defined for each splitting operation illustrated in Figures 2–4.

DEFINITION 7 (class $\mathcal{F}$). *We define the class $\mathcal{F}$ of graphs by the following inductive scheme:*

- $\bar{C}_6 \in \mathcal{F}$;
- if a graph G is in $\mathcal{F}$, then any graph obtained from G by splitting is in $\mathcal{F}$;
- if two graphs of $\mathcal{F}$ are glueable along a flat edge, then the graph resulting from the corresponding glueing is in $\mathcal{F}$.

See Figure 6.

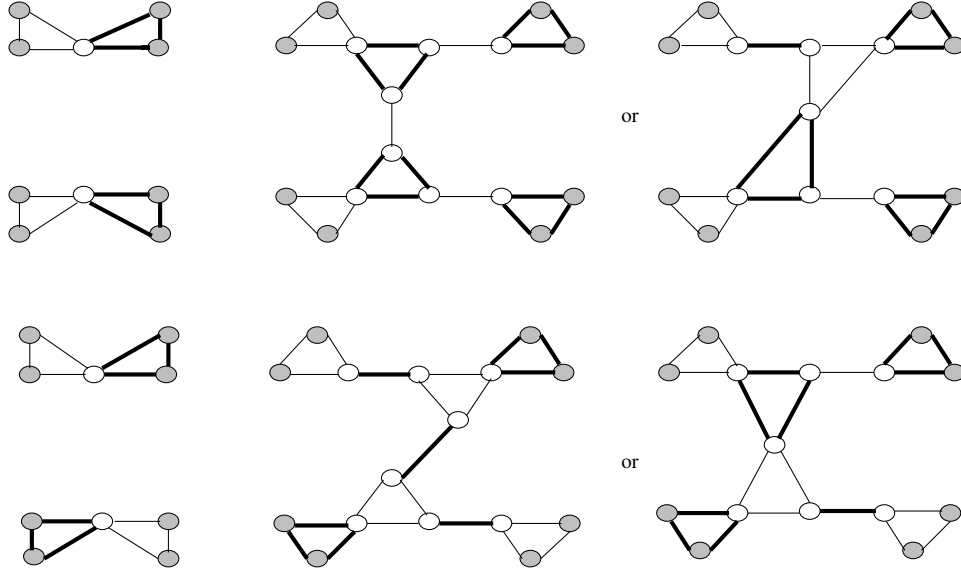


FIG. 4. Vertex-vertex operations.

THEOREM 7. *All graphs in $\mathcal{F}$ are candidates.*

Proof. The proof is by induction on the number of vertices. The theorem is true for $\bar{C}_6$. Now let us consider a graph $G' \neq \bar{C}_6$ in $\mathcal{F}$ and assume that the theorem is true for every graph smaller than G' . We know that G' is obtained either by splitting a graph G from $\mathcal{F}$ or by glueing two graphs G_1, G_2 from $\mathcal{F}$ along a common flat edge. If G' results from glueing, the desired conclusion is immediate from the glueing lemma and the induction hypothesis. So we may assume that G' results from splitting G . By the induction hypothesis, G is a candidate. Let H be the bipartite graph such that $G = L(H)$, with the vertices of H and the edges of G colored blue or red according to the bipartition. It is easy to see that G' is also the line-graph of a bipartite graph H' , because the splitting operation preserves the bipartition of the maximum cliques, and to see that G' is 2-connected. Clearly, splitting also preserves the fact that each neighborhood is correct. Also, it is a routine matter to check that the splitting operation maintains all path parities of G in G' ; i.e., if there exists a path between two vertices x, y of G , then G' contains a path of the same parity between the corresponding vertices. So we have established properties (c1)–(c3) for G' . Finally, suppose that G' has a cutset C which violates property (c4). Then C is necessarily a cutset of G , which contradicts the fact that G is a candidate. $\square$

It remains to prove that the operations generate all planar candidates. Thus we must prove that for each candidate we can, by performing inverse operations, obtain a bunch of $\bar{C}_6$'s.

We will say that G' arises from G by a *reduction* if G arises from G' via a splitting.

Recall that each neighborhood in a candidate must be correct. This implies that every vertex x and every flat edge e of a candidate corresponds to a possible reduction. Unfortunately, this reduction may fail to yield a proper smaller candidate, as we may create a graph that violates (c2) or (c4). We are going to show that these problems can be avoided by an appropriate choice of x or e . That is, we will show that if G is a 3-connected candidate, we can choose a reduction that yields a 3-connected candidate. If G is not 3-connected, then G arises from smaller candidates by glueing.

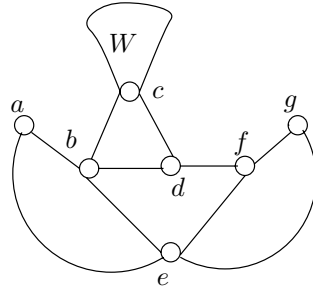


FIG. 5. A diabolo $\{b, c, d, W\}$ whose reduction generates a diamond

Putting these results together will give the desired result that every planar candidate is in $\mathcal{F}$.

Call *trivial cutset* a cutset C such that one component of $G - C$ is a vertex or an edge or a triangle. Let C be a trivial 3-cutset. Any diabolo associated with a vertex of C defines a reduction which does not change the connectivity of G .

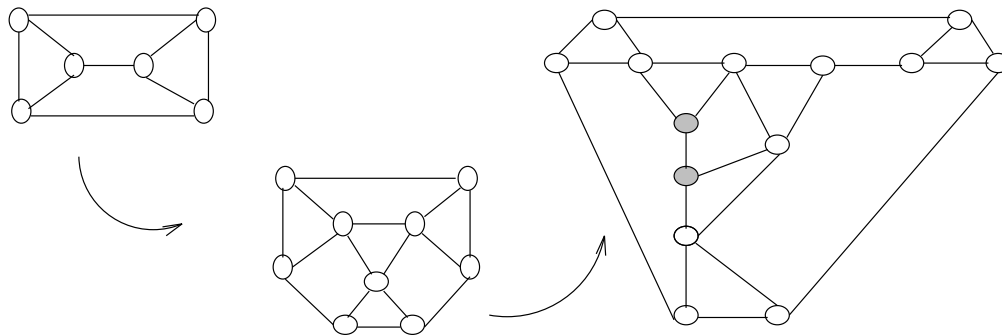
Observe that only the vertex-edge and vertex-vertex reverse operations can produce an incorrect neighborhood for a candidate, by creating a diamond, and that they do so only if the vertices to be identified belong to a face of length four. We can see this situation in Figure 5.

LEMMA 27. *Let G be a 3-connected candidate. Let W be a diabolo whose reduction would generate a diamond. Then, there is another diabolo whose reduction will generate a 3-connected candidate.*

Proof. Let us label the vertices as in Figure 5. We must examine two cases. First, suppose that the face containing c, d, f, g is not of length four. Since the identification of d and g cannot produce a diamond, consider the diabolo W' having vertex e as the center. We claim that the reduction corresponding to W' generates a 3-connected candidate. The nontriangular face containing a and c cannot be of length four with a of degree three, since g and a neighbor of c in the other nontriangular face will form a 2-cutset which is an even pair. Also, if e is in a 3-cutset, this cutset is trivial (by a tedious but easy analysis) and so the graph obtained by the reduction remains 3-connected and is a candidate. Now, if the face containing c, d, f, g has length four, then cg must be an edge. Note that cg cannot be a flat edge, for otherwise $\{c, a\}$ would be a 2-cutset. So the reduction corresponding to the diabolo centered on df generates a 3-connected and diamond-free graph (it is easy to see that, in this case, all 3-cutsets containing f are trivial). $\square$

LEMMA 28. *If a vertex x of a 3-connected candidate G belongs to a nontrivial cutset, then one of the two nontriangular faces containing x has a vertex y such that any 3-cutset containing y is trivial.*

Proof. Let F_1 and F_2 be the nontriangular faces containing x . Let y' and z be the vertices of F_1 and F_2 , respectively, such that $G - \{x, y', z\}$ has the smallest component between all nontrivial 3-cutsets. Let y be a vertex of F_1 in $G - \{x, y', z\}$ such that y is a neighbor of y' and such that the edge $y'y$ belongs to a triangle. The vertex y is contained in a diabolo W which contains the intersection of F_1 and another nontriangular face F'_2 . We claim that every 3-cutset containing y is trivial. Otherwise, there exists a 3-cutset $\{y, y'', x'\}$ with y'' in F_1 and x' in F'_2 such that there is no component included in W . Therefore, either y' is not in the external face or $\{y'', y'\}$ or $\{x', x''\}$ is a 2-cutset, where x'' is the other vertex of W' that is on the external face, which in both cases is a contradiction. $\square$

FIG. 6. A graph G in $\mathcal{F}$. The gray vertices form a danger.

LEMMA 29. *In every 3-connected candidate, there exists a diablo whose reduction generates a candidate.*

Proof. Consider a candidate G . Choose some vertex x of G . If the reduction on the diablo defined by this vertex produces a 2-connected graph G' in which every neighborhood is correct, and in which all 2-cutsets form odd pairs, then clearly G' is a candidate and the lemma holds. Otherwise, by Lemma 28, there is a diablo in G centered in a vertex y such that the corresponding reduction produces a 3-connected graph. By Lemma 27, this diablo may be replaced by one whose corresponding reduction produces a graph with a correct neighborhood. This lemma also guarantees that this reduction will keep the 3-connectivity. So, the lemma follows. $\square$

THEOREM 8. *All candidates belong to $\mathcal{F}$.*

Proof. This proof is by induction on the number of vertices. The theorem is trivially true for the smallest candidate $\bar{C}_6$. Suppose it is true for candidates with less than n vertices and consider a candidate G with n vertices. If G has a 2-cutset, let G_1 and G_2 be the graphs obtained by ungluing G along this 2-cutset. Clearly, G_1 and G_2 are smaller than G and are candidates. So G is contained in $\mathcal{F}$. If G is 3-connected, by Lemma 29, there is an inverse operation that generates a candidate G' smaller than G . By the induction hypothesis, G' can be reduced to a set of $\bar{C}_6$'s. So, G can be obtained from G' by an operation of Definition 7 and then belongs to $\mathcal{F}$. $\square$

LEMMA 30. *A graph G in $\mathcal{F}$ is an obstruction if and only if G has no dangers.*

This is a mere reformulation of Lemma 24.

Let $\mathcal{F}^*$ be the subclass of $\mathcal{F}$ made of the graphs containing no danger. The following is a corollary of all the results in this section.

THEOREM 9. *The class $\mathcal{F}^*$ consists exactly of all obstructions.*

Proof. It is clear that every graph from $\mathcal{F}^*$ is an obstruction. By Theorem 8, we can reduce any candidate to a collection of $\bar{C}_6$'s, and so from this collection we can obtain the original graph by performing in reverse order the same operations. So every obstruction is in $\mathcal{F}^*$. $\square$

5. Conclusion and comments. We proved that Hougardy's conjecture is true for planar graphs. A related problem is to find a polynomial-time algorithm that decides if a graph is a planar SQP graph. We solved this problem in [9] using a decomposition technique due to Hsu [6], which is very different from the results presented here. Another related problem is to find an even pair (or decide that there is none) in a planar perfect graph. This problem was solved by Hsu and Porto [7] and by Linhares Sales and Sampaio [10]. Let us also recall that coloring can be done in polynomial time for all planar perfect graphs with an algorithm due to Hsu [6].

Acknowledgments. We are grateful to Claude Benzaken and Myriam Preissmann for their suggestions, and to the anonymous referees for useful comments.

REFERENCES

- [1] M. CHUDNOVSKY, N. ROBERTSON, P. SEYMOUR, AND R. THOMAS, *The strong perfect graph theorem*, Ann. of Math., 164 (2006), pp. 51–229.
- [2] H. EVERETT, C. M. H. DE FIGUEIREDO, C. LINHARES SALES, F. MAFFRAY, O. PORTO, AND B. A. REED, *Even pairs*, in Perfect Graphs, J. L. Ramírez-Alfonsín and B. A. Reed, eds., Wiley Interscience, New York, 2001, pp. 67–92.
- [3] J. FONLUPUT AND J. P. UHRY, *Transformations which preserve perfectness and h -perfectness of graphs*, Ann. Discrete Math., 16 (1982), pp. 83–85.
- [4] F. HARARY AND C. HOLZMANN, *Line graphs of bipartite graphs*, Rev. Soc. Mat. Chile, 1 (1974), pp. 19–22.
- [5] S. HOUGARDY, *Perfekte Graphen*, Ph.D. thesis, Institut für Ökonometrie und Operations Research, Universität Bonn, Bonn, Germany, 1991.
- [6] W.-L. HSU, *Coloring planar perfect graphs by decomposition*, Combinatorica, 6 (1986), pp. 381–385.
- [7] W.-L. HSU AND O. PORTO, *A Polynomial Algorithm to Test an Even Pair in a Planar Perfect Graph*, manuscript, 1994.
- [8] W.L. HSU, *Recognizing planar perfect graphs*, J. ACM, 34 (1987), pp. 255–288.
- [9] C. LINHARES SALES, F. MAFFRAY, AND B. A. REED, *Recognizing planar strict quasi-parity graphs*, Graphs Combin., 17 (2001), pp. 745–757.
- [10] C. LINHARES SALES AND R. M. SAMPAIO, *On the complexity of finding even pairs in planar perfect graphs*, Electron Notes Discrete Math., 7 (2001), pp. 186–189.
- [11] H. MEYNIEL, *A new property of critical imperfect graphs and some consequences*, European J. Combin., 8 (1987), pp. 313–316.
- [12] T. NISHIZEKI AND N. CHIBA, *Planar Graphs: Theory and Algorithms*, North-Holland Math. Stud. 14, North-Holland, Amsterdam, 1988.
- [13] A. TUCKER, *The validity of the perfect graph conjecture for K_4 -free graphs*, in Topics on Perfect Graphs, C. Berge and V. Chvátal, eds., North-Holland, Amsterdam, 1984, pp. 149–157.
- [14] A. TUCKER, *Coloring perfect $(K_4 - e)$ -free graphs*, J. Combin. Theory Ser. B, 43 (1987), pp. 313–318.

AN EXPLICIT BOUND ON DOUBLE EXPONENTIAL SUMS RELATED TO DIFFIE–HELLMAN DISTRIBUTIONS*

MEI-CHU CHANG[†] AND CHUI ZHI YAO[†]

Abstract. Let p be a prime and V an integer of order t in the multiplicative group modulo p . In this paper, we give an explicit bound on the double exponential sums. For $\varepsilon > 0$, we have $\tilde{S}_{a,b,c}(t) =$

$$\sum_{x,y=1}^t e_p(aV^x + bV^y + cV^{xy}) \ll t^{2-r(\varepsilon)}, \text{ where } r(\varepsilon) = \begin{cases} \frac{1}{4440} & \text{if } p^{\frac{1}{2}} \leq t \leq p^{\frac{2}{3}}, \\ \frac{1}{20}\varepsilon^3(\log \frac{1}{\varepsilon})^{-6} & \text{if } p^{\frac{1}{4}+\varepsilon} \leq t < p^{\frac{1}{2}}. \end{cases}$$

Key words. Diffie–Hellman triples, Diffie–Hellman distributions, exponential sums

AMS subject classifications. Primary, 05C38, 15A15; Secondary, 05A15, 15A18

DOI. 10.1137/070699767

1. Introduction. The Diffie–Hellman key exchange algorithm [6] is the first practical public key cryptosystem published and it remains one of the cornerstones of modern cryptography to date. The algorithm is a simple but ingenious way for two parties to establish a common secret key over an insecure channel. The security of this algorithm is based on the assumption that certain desirable properties are possessed by the Diffie–Hellman triples (V^x, V^y, V^{xy}) , where V is an integer of multiplicative order t modulo $p \geq 3$ that is $V^x \not\equiv 1 \pmod{p}$, $x = 1, \dots, t-1$, $V^t \equiv 1 \pmod{p}$. It has been shown in [5] and then improved in [4] that such triples are uniformly distributed in the sense of Weyl when $x, y = 0, \dots, t$ (see [17] for details on this notion of uniformity). Although such results do not guarantee the security of the Diffie–Hellman key, they are nevertheless very desirable since it provides evidence that the Diffie–Hellman cryptosystem can withstand statistic-based attacks. On the other hand, studying the distribution of these triples via finding bounds on double exponential sums is a very natural and attractive number theoretic question. Various other applications and generalizations of the exponential sums bounds and related results of [5] and [4] can be found in [1], [7], [8], [2], [10], [9], [11], and [16].

In this paper we continue the study of double exponential sums related to the Diffie–Hellman triples (V^x, V^y, V^{xy}) as initiated in [5]. For integers a, b, c , we define the following exponential sums:

$$\tilde{S}_{a,b,c}(t) = \sum_{x,y=1}^t e_p(aV^x + bV^y + cV^{xy}),$$

where

$$e_p(\theta) = \exp\left(\frac{2\pi i\theta}{p}\right).$$

*Received by the editors August 8, 2007; accepted for publication October 25, 2007; published electronically February 29, 2008.

<http://www.siam.org/journals/sidma/22-1/69976.html>

[†]Department of Mathematics, University of California, Riverside, Riverside, CA 92521 (mcc@math.ucr.edu, zhi@math.ucr.edu).

We obtain an explicit bound on sums $\tilde{S}_{a,b,c}(t)$ by studying slightly different sums

$$S_{a,c}(t) = \sum_{y=1}^t \left| \sum_{x=1}^t e_p(aV^x + cV^{xy}) \right|$$

for which obviously $|\tilde{S}_{a,b,c}(t)| \leq S_{a,c}(t)$. An estimate of the sums $S_{a,c}(t)$ obtained in [5] was improved and generalized by [4]. It had been shown in [4] that $S_{a,c}(t) \ll t^{5/3}p^{1/4}$ for $t > p^{3/4+\varepsilon}$ and the authors posted an open question to find an estimate of the sum for t in a lower range. In response to their open question, Bourgain [3] obtained a bound of the sum that was nontrivial for $t \geq p^\varepsilon$. Although his bound $S_{a,c}(t) \ll t^{2-\delta}$ was not explicit in the sense that there was no clear relationship between ε and δ , his estimates remained nontrivial over remarkably short intervals. At about the same time, Garaev [12] was not only able to improve the bound obtained in [4], but more importantly he was able to extend the range. His bound $S_{a,c}(t) \ll t^{7/4}p^{1/8+\varepsilon}$ was nontrivial beginning with $t > p^{1/2+\varepsilon}$. Later on, Garaev and Karatsuba [13] obtained the bound $S_{a,c}(t) \ll tp^{7/8+o(1)}$ by introducing a new ingredient to the argument of [12]. Their bound includes the one in [12] and improves it for any thin set β where $|\beta| < p^{1-d}$ for some $d > 0$. In this paper, we are able to give an explicit bound on the range beginning with $t > p^{1/4+\varepsilon}$. Although our bound is not as sharp as the Garaev–Karatsuba estimate contained in [13], nor our applicable range as wide as Bourgain’s in [3], we are able to give an explicit bound that will work in the range that was not covered by any explicit bound so far.

2. Lemmas. Our bound on the exponential sum $S_{a,c}(t)$ relies on the following lemmas. Each lemma is used to prove the subsequent lemmas or to prove our main result. Our first lemma is a simple identity which is a basic tool for using exponential sums in the study of different problems module m , but we state the lemma here only for case p as a prime.

LEMMA 1. *For any integer u ,*

$$\sum_{\lambda=0}^{p-1} e_p(\lambda u) = \begin{cases} p & \text{if } u \equiv 0 \pmod{p}, \\ 0 & \text{if } u \not\equiv 0 \pmod{p}. \end{cases}$$

The next lemma is a generalization of Lemma 1 to elements in finite fields $\mathbb{F}_p^d$. For $d \geq 2$ and $\mathbb{F}_p^d = \mathbb{F}_p \times \cdots \times \mathbb{F}_p$, we identify $\mathbb{F}_p$ with $\{0, 1, \dots, p-1\}$. Furthermore, for $y = (y_1, \dots, y_d)$ and $x = (x_1, \dots, x_d) \in \mathbb{F}_p^d$, we denote

$$x \cdot y = \sum_{i=1}^d x_i y_i.$$

LEMMA 2. *With the above notation and for an arbitrarily fixed element $\hat{y}$ of $\mathbb{F}_p^d$, we have*

$$\sum_{x \in \mathbb{F}_p^d} e_p(x \cdot \hat{y}) = \begin{cases} p^d & \text{if } y_i \equiv 0 \pmod{p} \ \forall i, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Note that if $y_i \equiv 0 \pmod{p}$ for each i , then $x \cdot \hat{y} \equiv 0 \pmod{p}$ and

$$\sum_{x \in \mathbb{F}_p^d} e_p(x \cdot \hat{y}) = \sum_{x \in \mathbb{F}_p^d} 1 = p^d.$$

Otherwise, there exists an i such that $y_i \neq 0$.

Without loss of generality, we may assume $y_1 \neq 0$; then

$$\begin{aligned} \sum_{x \in \mathbb{F}_p^d} e_p(x \cdot \hat{y}) &= \sum_{x \in \mathbb{F}_p^d} e_p \left(\sum_{i=1}^d x_i y_i \right) \\ &= \sum_{x \in \mathbb{F}_p^d} e_p(x_1 y_1) \cdots e_p(x_d y_d) \\ &= \sum_{x_d=0}^{p-1} e_p(x_d y_d) \cdots \sum_{x_1=0}^{p-1} e_p(x_1 y_1). \end{aligned}$$

Applying Lemma 1 to the quantity $\sum_{x_1=0}^{p-1} e_p(x_1 y_1)$ gives us the desired result. $\square$

We let $h : \mathbb{F}_p^d \longrightarrow \mathbb{C}$, and denote

$$(1) \quad \|h\|_2 = \left(\sum_{x \in \mathbb{F}_p^d} |h(x)|^2 \right)^{\frac{1}{2}}.$$

LEMMA 3. *With the above notation and if $f, g : \mathbb{F}_p^d \longrightarrow \mathbb{C}$, then*

$$\left| \sum_{x, y \in \mathbb{F}_p^d} f(x) g(y) e_p(x \cdot y) \right| \leq p^{\frac{d}{2}} \|f\|_2 \|g\|_2.$$

Proof. Note that $|\sum_{x, y \in \mathbb{F}_p^d} f(x) g(y) e_p(x \cdot y)| = |\sum_{x \in \mathbb{F}_p^d} f(x) \sum_{y \in \mathbb{F}_p^d} g(y) e_p(x \cdot y)|$.

To prove the lemma, we first apply the triangle inequality and then apply the Cauchy-Schwarz inequality

$$\begin{aligned} &\left| \sum_{x, y \in \mathbb{F}_p^d} f(x) g(y) e_p(x \cdot y) \right| \\ &\leq \sum_{x \in \mathbb{F}_p^d} |f(x)| \left| \sum_{y \in \mathbb{F}_p^d} g(y) e_p(x \cdot y) \right| \\ &\leq \left(\sum_{x \in \mathbb{F}_p^d} |f(x)|^2 \right)^{\frac{1}{2}} \left(\sum_{x \in \mathbb{F}_p^d} \left| \sum_{y \in \mathbb{F}_p^d} g(y) e_p(x \cdot y) \right|^2 \right)^{\frac{1}{2}} \\ &= \|f\|_2 \left(\sum_{y_1, y_2 \in \mathbb{F}_p^d} g(y_1) \overline{g(y_2)} \sum_{x \in \mathbb{F}_p^d} e_p(x \cdot (y_1 - y_2)) \right)^{\frac{1}{2}}. \end{aligned}$$

Finally, we apply Lemma 2 to the quantity $\sum_{x \in \mathbb{F}_p^d} e_p(x \cdot (y_1 - y_2))$ to obtain

$$\begin{aligned} \left| \sum_{x, y \in \mathbb{F}_p^d} f(x) g(y) e_p(x \cdot y) \right| &\leq \|f\|_2 \left(p^d \sum_{y \in \mathbb{F}_p^d} |g(y)|^2 \right)^{\frac{1}{2}} \\ &= p^{\frac{d}{2}} \|f\|_2 \|g\|_2, \end{aligned}$$

which is the desired result. $\square$

The next lemma is a generalization of Lemma 2 to the multiplicative subgroup of $(\mathbb{F}_p^*)^d$; but first we define the following notations.

We let $k \in \mathbb{N}$, $\tilde{H} < (\mathbb{F}_p^*)^d$ with $|\tilde{H}| = H$, and denote

$$(2) \quad u_{4k} = \left| \left\{ (x_1, \dots, x_{4k}) \in \tilde{H}^{4k} : \sum_{i=1}^k x_i - \sum_{i=k+1}^{2k} x_i = \sum_{i=2k+1}^{3k} x_i - \sum_{i=3k+1}^{4k} x_i \right\} \right|.$$

Moreover, we let $x = (x_1, \dots, x_d), y = (y_1, \dots, y_d), z = (z_1, \dots, z_d) \in \tilde{H}$, and denote

$$x \cdot y \cdot z = \sum_{i=1}^d x_i y_i z_i.$$

LEMMA 4. *With the above notation and if $x \in \tilde{H}$ and $b \in (F_p^*)^d$, then*

$$\left| \sum_{x \in \tilde{H}} e_p(b \cdot x) \right| \leq H^{1-\frac{1}{k}} p^{\frac{d}{8k^2}} (u_{4k})^{\frac{1}{4k^2}}.$$

Proof. Note that for any fixed element y of $\tilde{H}$,

$$\sum_{x \in \tilde{H}} e_p(b \cdot x) = \sum_{x \in \tilde{H}} e_p(b \cdot x \cdot y).$$

Define

$$S = \left| \sum_{x \in \tilde{H}} e_p(b \cdot x) \right|.$$

Then,

$$S = \frac{1}{H} \sum_{y \in \tilde{H}} \left| \sum_{x \in \tilde{H}} e_p(b \cdot x \cdot y) \right|.$$

To prove the lemma, we first apply the Hölder's inequality to the right-hand side

with $q = 2k$ to obtain

$$\begin{aligned}
 S &\leq \frac{1}{H} \cdot H^{1-\frac{1}{2k}} \left(\sum_{y \in \tilde{H}} \left| \sum_{x \in \tilde{H}} e_p(b \cdot x \cdot y) \right|^{2k} \right)^{\frac{1}{2k}} \\
 (3) \quad &= H^{-\frac{1}{2k}} \left(\sum_{y \in \tilde{H}} \left| \sum_{x \in \tilde{H}} e_p(b \cdot x \cdot y) \right|^{2k} \right)^{\frac{1}{2k}}.
 \end{aligned}$$

Next, we will bound the quantity $\sum_{y \in \tilde{H}} \left| \sum_{x \in \tilde{H}} e_p(b \cdot x \cdot y) \right|^{2k}$ by again applying Hölder's inequality with $q = 2k$.

Note,

$$\begin{aligned}
 &\sum_{y \in \tilde{H}} \left| \sum_{x \in \tilde{H}} e_p(b \cdot x \cdot y) \right|^{2k} \\
 &= \sum_{y \in \tilde{H}} \sum_{x_1, \dots, x_{2k} \in \tilde{H}} e_p(b \cdot y \cdot (x_1 + \dots + x_k - x_{k+1} - \dots - x_{2k})) \\
 &\leq \sum_{x_1, \dots, x_{2k} \in \tilde{H}} \left| \sum_{y \in \tilde{H}} e_p \left(b \cdot y \cdot \left(\sum_{i=1}^k x_i - \sum_{j=k+1}^{2k} x_j \right) \right) \right| \\
 &\leq \left((H^{2k})^{1-\frac{1}{2k}} \right) \left(\sum_{x_1, \dots, x_{2k} \in \tilde{H}} \left| \sum_{y \in \tilde{H}} e_p \left(b \cdot y \cdot \left(\sum_{i=1}^k x_i - \sum_{j=k+1}^{2k} x_j \right) \right) \right|^{2k} \right)^{\frac{1}{2k}} \\
 &= H^{2k-1} \left(\sum_{\substack{x_1, \dots, x_{2k} \in \tilde{H} \\ y_1, \dots, y_{2k}}} e_p \left(b \cdot \left(\sum_{i=1}^k y_i - \sum_{j=k+1}^{2k} y_j \right) \cdot \left(\sum_{i=1}^k x_i - \sum_{j=k+1}^{2k} x_j \right) \right) \right)^{\frac{1}{2k}}.
 \end{aligned}$$

Now, we combine the above bound with (3) to obtain

$$(4) \quad S \leq H^{1-\frac{1}{k}} \left(\sum_{\substack{x_1, \dots, x_{2k} \in \tilde{H} \\ y_1, \dots, y_{2k}}} e_p \left(b \cdot \left(\sum_{i=1}^k y_i - \sum_{j=k+1}^{2k} y_j \right) \cdot \left(\sum_{i=1}^k x_i - \sum_{j=k+1}^{2k} x_j \right) \right) \right)^{\frac{1}{4k^2}}.$$

Furthermore, we define

$$f(x) = \left| \left\{ (x_1, \dots, x_{2k}) \in \tilde{H}^{2k} : \sum_{i=1}^k x_i - \sum_{j=k+1}^{2k} x_j = x \right\} \right|$$

and

$$f(by) = \left| \left\{ (y_1, \dots, y_{2k}) \in \tilde{H}^{2k} : \sum_{i=1}^k b \cdot y_i - \sum_{j=k+1}^{2k} b \cdot y_j = b \cdot y \right\} \right|.$$

Also, by (1) and (2), we have

$$\|f\|_2^2 = \sum_{x \in \mathbb{F}_p^d} |f(x)|^2 = u_{4k}.$$

With the above notation we can express (4) as follows;

$$S \leq H^{1-\frac{1}{k}} \left(\sum_{x, y \in \mathbb{F}_p^d} f(x) f(b \cdot y) e_p(b \cdot x \cdot y) \right)^{\frac{1}{4k^2}}.$$

Finally, we apply Lemma 3 to the right-hand side of the above equality to obtain

$$\begin{aligned} S &\leq H^{1-\frac{1}{k}} \left(p^{\frac{d}{2}} \|f\|_2^2 \right)^{\frac{1}{4k^2}} \\ &= H^{1-\frac{1}{k}} p^{\frac{d}{8k^2}} \left(\|f\|_2^2 \right)^{\frac{1}{4k^2}} \\ &= H^{1-\frac{1}{k}} p^{\frac{d}{8k^2}} (u_{4k})^{\frac{1}{4k^2}} \end{aligned}$$

which is the desired result. $\square$

Let m be a natural number and G a multiplicative subgroup of $\mathbb{Z}_p^*$ with $\text{ord}(G) = t$; then we define $T_m(G)$ to be the number of solutions of

$$x_1 + \dots + x_k = y_1 + \dots + y_k,$$

where $x_1, \dots, x_m, y_1, \dots, y_m$ are element of G .

Next, with the above notation, we will state a lemma due to Konyagin (see Lemma 5 in [14]) and a lemma due to Konyagin and Shparlinski (see [15] or Lemma 3 in [14]).

LEMMA 5 (see Konyagin [14]). *For every natural number m , there exists a constant $C(m)$ such that if $t = |G| < p^{\frac{2}{3}}$ and $m = 2$, or if $t < p^{\frac{1}{2}}$ and $m > 2$, then*

$$T_m(G) \leq C(m) t^{2m-2+\frac{1}{2m-1}}.$$

LEMMA 6 (see Konyagin and Shparlinski [15]). *Let $a, \theta \in \mathbb{F}_p^*$ with $\text{ord}(\theta) = t$. If k and l are elements of natural numbers, then*

$$\max_{a \in \mathbb{F}_p^*} \left| \sum_{x=1}^t e_p(a\theta^x) \right| \leq p^{\frac{1}{2l^m}} T_l^{\frac{1}{2l^m}} T_m^{\frac{1}{2l^m}} t^{1-\frac{1}{l}-\frac{1}{m}}.$$

The above sum has numerous applications in additive problems modulo p , pseudo-random generators, coding theory, and other problems. We will use its bound to prove

our result. Note that for ε in the interval $(0, \frac{1}{4})$ and $p^{\frac{1}{4}+\varepsilon} \leq t < p^{\frac{1}{2}}$, we can combine Lemma 5 and Lemma 6 by taking $l = m$ to obtain the bound

$$(5) \quad \max_{a \in \mathbb{F}_p^*} \left| \sum_{x=1}^t e_p(a\theta^x) \right| \leq tp^{-\varepsilon(\log \frac{1}{\varepsilon})^{-2}}.$$

Next, we will state another lemma due to Konyagin and Shparlinski (see [15]), which gives us the bound for the above sum for $t \geq p^{\frac{1}{3}}$.

LEMMA 7 (see Konyagin and Shparlinski [15]). *Let $a, \theta \in \mathbb{F}_p^*$ with $\text{ord}(\theta) = t$; then*

$$\max_{a \in \mathbb{F}_p^*} \left| \sum_{x=1}^t e_p(a\theta^x) \right| \ll \begin{cases} p^{\frac{1}{2}} & \text{if } t > p^{\frac{2}{3}}, \\ p^{\frac{1}{4}} t^{\frac{3}{8}} & \text{if } p^{\frac{1}{2}} \leq t \leq p^{\frac{2}{3}}, \\ p^{\frac{1}{8}} t^{\frac{5}{8}} & \text{if } p^{\frac{1}{3}} \leq t \leq p^{\frac{1}{2}}. \end{cases}$$

3. Main theorem.

THEOREM 1. *Let a, c be integers that are coprime to p . Let $\varepsilon > 0$ and $\theta \in \mathbb{F}_p^*$ with $\text{ord}(\theta) = t$.*

If $p^{\frac{1}{4}+\varepsilon} \leq t < p^{\frac{1}{2}}$ or $p^{\frac{1}{2}} \leq t \leq p^{\frac{2}{3}}$, then there exists a positive constant $r(\varepsilon)$ such that

$$S_{a,c}(t) = \sum_{y=1}^t \left| \sum_{x=1}^t e_p(a\theta^x + c\theta^{xy}) \right| \ll t^{2-r(\varepsilon)},$$

where

$$r(\varepsilon) = \begin{cases} \frac{1}{4440} & \text{if } p^{\frac{1}{2}} \leq t \leq p^{\frac{2}{3}}, \\ \frac{1}{20}\varepsilon^3 \left(\log \frac{1}{\varepsilon}\right)^{-6} & \text{if } p^{\frac{1}{4}+\varepsilon} \leq t < p^{\frac{1}{2}}. \end{cases}$$

Proof. Case 1. $p^{\frac{1}{4}+\varepsilon} \leq t < p^{\frac{1}{2}}$.

Let $\hat{H} = \hat{H}_y = \{(\theta^x, \theta^{xy}) : x = 1, \dots, t\} < (\mathbb{F}_p^*)^2$.

As before, we denote $(a, c) \cdot (\theta^x, \theta^{xy}) = a\theta^x + c\theta^{xy}$.

To prove the first case, we apply Lemma 4 to the quantity $|\sum_{x=1}^t e_p(a\theta^x + c\theta^{xy})|$ with $d = 2$, $r = 2k$, $b = (a, c)$, and $x = (\theta^x, \theta^{xy})$.

We obtain

$$\left| \sum_{x=1}^t e_p(a\theta^x + c\theta^{xy}) \right| \leq t^{1-\frac{2}{r}} p^{\frac{1}{r^2}} (u_{2r}(y))^{\frac{1}{r^2}},$$

where

$$u_{2r}(y) = \left| \left\{ (x_1, \dots, x_{2r}) \in \{1, \dots, t\}^{2r} : \theta^{x_1} + \dots + \theta^{x_r} = \theta^{x_{r+1}} + \dots + \theta^{x_{2r}} \right. \right. \\ \left. \left. \text{and } \theta^{yx_1} + \dots + \theta^{yx_r} = \theta^{yx_{r+1}} + \dots + \theta^{yx_{2r}} \right\} \right|.$$

Furthermore, we note that for $r \in 2\mathbb{N}$, $p^{-1}t^{2r-\frac{r}{20-20r^2}} \geq 0$. We let

$$(6) \quad \lambda = \frac{r}{20-20r^2};$$

then

$$\begin{aligned}
 S_{a,c}(t) &= \sum_{y=1}^t \left| \sum_{x=1}^t e_p(a\theta^x + c\theta^{xy}) \right| \\
 &\leq \sum_{\substack{1 \leq y \leq t \\ u_{2r}(y) \leq t^{2r-\lambda} p^{-1}}} t^{1-\frac{2}{r}} p^{\frac{1}{r^2}} (u_{2r}(y))^{\frac{1}{r^2}} + \sum_{\substack{1 \leq y \leq t \\ u_{2r}(y) > t^{2r-\lambda} p^{-1}}} \left| \sum_{x=1}^t e_p(a\theta^x + c\theta^{xy}) \right| \\
 &\leq \sum_{y=1}^t t^{1-\frac{2}{r}} p^{\frac{1}{r^2}} (t^{2r-\lambda} p^{-1})^{\frac{1}{r^2}} + vt \\
 (7) \quad &= t^{2-\frac{\lambda}{r^2}} + vt,
 \end{aligned}$$

where

$$v = |\{y : u_{2r}(y) > t^{2r-\lambda} p^{-1}\}|.$$

Next, we define

$$T = \left| \left\{ \begin{array}{l} (x_1, \dots, x_{2r}, y) \in \{1, \dots, t\}^{2r+1} : \theta^{x_1} + \dots + \theta^{x_r} = \theta^{x_{r+1}} + \dots + \theta^{x_{2r}} (*) \\ \text{and } \theta^{yx_1} + \dots + \theta^{yx_r} = \theta^{yx_{r+1}} + \dots + \theta^{yx_{2r}} (**) \end{array} \right\} \right|.$$

Then since $T \geq vt^{2r-\lambda} p^{-1}$, we have $v \leq Tt^{\lambda-2r} p$.

Furthermore, we write

$$T = \sum_{\bar{x} \in \Omega} T_{\bar{x}},$$

where

$$\Omega = \left\{ \bar{x} = (x_1, \dots, x_{2r}) \in \{1, \dots, t\}^{2r} : \bar{x} \text{ satisfy } (*) \right\},$$

$$T_{\bar{x}} = |\{y = 1, \dots, t\} : y \text{ satisfy } (**)|.$$

Next, note that for each ε in $(0, \frac{1}{4})$, we can find an even positive integer $r \geq \frac{(\log \frac{1}{\varepsilon})^2}{2\varepsilon}$ such that if $p^{\frac{1}{4}+\varepsilon} \leq t$ and $r \geq \frac{(\log \frac{1}{\varepsilon})^2}{2\varepsilon}$, then

$$(8) \quad \frac{t^{2r}}{p^{2r\varepsilon(\log \frac{1}{\varepsilon})^{-2}}} \leq \frac{t^{2r}}{p}.$$

Now we will bound $|\Omega|$ by applying the bound (5) and (8).

Let $r \geq \frac{(\log \frac{1}{\varepsilon})^2}{2\varepsilon}$ and recall that $p^{\frac{1}{4}+\varepsilon} \leq t < p^{\frac{1}{2}}$ is part of our assumption.

Note that,

$$\begin{aligned}
|\Omega| &= \frac{1}{p} \sum_{j=0}^{p-1} \left| \sum_{x=1}^t e_p(j\theta^x) \right|^{2r} \\
&\leq \frac{t^{2r}}{p} + \max_{j \in \mathbb{F}_p^*} \left| \sum_{x=1}^t e_p(j\theta^x) \right|^{2r} \\
&< \frac{t^{2r}}{p} + \left(tp^{-\varepsilon \left(\log \frac{1}{\varepsilon}\right)^{-2}} \right)^{2r} \\
&= \frac{t^{2r}}{p} + \frac{t^{2r}}{p^{2r\varepsilon \left(\log \frac{1}{\varepsilon}\right)^{-2}}} \\
&\ll \frac{t^{2r}}{p}.
\end{aligned}$$

Next, we will bound $T_{\bar{x}}$ by following the same argument as in [4]. Let $\theta = g^m$, where $m = \frac{p-1}{t}$ and g is a primitive root of $\mathbb{F}_p^*$. To bound $T_{\bar{x}}$, we need to bound the solutions in $y = 1, \dots, t$ of

$$g^{yk_1} + \dots + g^{yk_r} = g^{yk_{r+1}} + \dots + g^{yk_{2r}} \text{ with } k_i = mx_i.$$

The number of these solutions is equal to $\frac{N}{m}$, where N is the number of solutions in $z \in \mathbb{F}_p^*$ of

$$z^{k_1} + \dots + z^{k_r} = z^{k_{r+1}} + \dots + z^{k_{2r}}.$$

Apply Lemma 7 of [4] with $n = 2r, \tau_i = k_i$, and $a_1 = \dots = a_r = 1 = -a_{r+1} = \dots = -a_{2r}$. We have $N \ll p^{1-\frac{1}{2r-1}} D^{\frac{1}{2r-1}}$, where $D = \min_{1 \leq i \leq 2r} \max_{j \neq i} (k_j - k_i, p-1) = m \min_{1 \leq i \leq 2r} \max_{j \neq i} (x_j - x_i, p-1)$. Hence, $T_{\bar{x}} \leq t^{1-\frac{1}{2r-1}} D(x_1, \dots, x_{2r})^{\frac{1}{2r-1}}$ with $D(x_1, \dots, x_{2r}) = \min_i \max_{j \neq i} (x_j - x_i, t)$. Thus,

$$T = \sum_{\bar{x} \in \Omega} T_{\bar{x}} \ll \sum_{\bar{x} \in \Omega} t^{1-\frac{1}{2r-1}} D(\bar{x})^{\frac{1}{2r-1}} = t^{1-\frac{1}{2r-1}} \sum_{\bar{x} \in \Omega} D(\bar{x})^{\frac{1}{2r-1}}.$$

Then, we follow the same argument as in [3] to bound T .

Let $\Omega = \bar{\Omega} \cup \tilde{\Omega}$, where $\bar{\Omega} = \{\bar{x} : D(\bar{x}) \leq t^{\frac{9}{10}}\}$ and $\tilde{\Omega} = \Omega \setminus \bar{\Omega}$; then

$$\begin{aligned}
T &= \sum_{\bar{x} \in \bar{\Omega}} T_{\bar{x}} + \sum_{\bar{x} \in \tilde{\Omega}} T_{\bar{x}} \\
&\leq |\Omega| t^{1-\frac{1}{2r-1}} \left(t^{\frac{9}{10}} \right)^{\frac{1}{2r-1}} + |\tilde{\Omega}| t \\
&= |\Omega| t^{1-\frac{1}{10(2r-1)}} + |\tilde{\Omega}| t.
\end{aligned}$$

Since $|\Omega| < 2\frac{t^{2r}}{p}$, we have

$$(9) \quad \begin{aligned} T &< 2\frac{t^{2r}}{p}t^{1-\frac{1}{10(2r-1)}} + \left|\tilde{\Omega}\right|t \\ &< 2t^{2r+1-\frac{1}{20r}}p^{-1} + t\left|\tilde{\Omega}\right|. \end{aligned}$$

Next, we will apply Bourgain's bound on $|\tilde{\Omega}|$ (see [3]) for which $D(\bar{x}) > t^{\frac{9}{10}}$. With $r = Q$ in Bourgain's bound, we have

$$(10) \quad \left|\tilde{\Omega}\right| < t^{\frac{8r}{5}}.$$

Next, note that for ε in $(0, \frac{1}{4})$ and $r \geq \frac{(\log \frac{1}{\varepsilon})^2}{2\varepsilon}$, we have $\varepsilon \geq \frac{80r-8r^2+1}{4(8r^2-1)}$. Also, one can easily check that if $p^{\frac{1}{4}+\varepsilon} \leq t$ and $\varepsilon \geq \frac{80r-8r^2+1}{4(8r^2-1)}$, then

$$(11) \quad t^{2r+1-\frac{1}{20r}}p^{-1} \geq t^{\frac{8r}{5}+1}.$$

Let $\varepsilon \geq \frac{80r-8r^2+1}{4(8r^2-1)}$ and recall that $p^{\frac{1}{4}+\varepsilon} \leq t \leq p^{\frac{1}{2}}$ is part of our assumption. Now we combine the bounds (9), (10), and (11) to bound T .

$$(12) \quad T \ll t^{2r+1-\frac{1}{20r}}p^{-1} + t\left|\tilde{\Omega}\right| \ll t^{2r+1-\frac{1}{20r}}p^{-1} + t^{\frac{8r}{5}+1} \ll t^{2r+1-\frac{1}{20r}}p^{-1}.$$

Next, we apply (12) to the inequality $v \leq Tt^{\lambda-2r}p$ to obtain a bound for v .

$$v \leq Tt^{\lambda-2r}p \ll \left(t^{2r+1-\frac{1}{20r}}p^{-1}\right)t^{\lambda-2r}p = t^{1+\lambda-\frac{1}{20r}}.$$

Then we apply the bound of v to (7); we therefore have

$$\begin{aligned} S_{a,c}(t) &= \sum_{y=1}^t \left| \sum_{x=1}^t e_p(a\theta^x + c\theta^{xy}) \right| \\ &\leq t^{2-\frac{\lambda}{r^2}} + t \cdot v \\ &\ll t^{2-\frac{\lambda}{r^2}} + t \cdot t^{1+\lambda-\frac{1}{20r}} \\ &= t^{2-\frac{\lambda}{r^2}} + t^{2+\lambda-\frac{1}{20r}} \\ &\ll t^{2-\frac{\lambda}{r^2}}. \end{aligned}$$

Substituting (6) to the above equation gives us

$$(13) \quad S_{a,c}(t) \ll t^{2-\frac{1}{20r+20r^3}}.$$

Finally, note that $2\lceil \frac{(\log \frac{1}{\varepsilon})^2}{4\varepsilon} \rceil$ is the least even positive integer greater than or equal to $\frac{(\log \frac{1}{\varepsilon})^2}{2\varepsilon}$ and $2\lceil \frac{(\log \frac{1}{\varepsilon})^2}{4\varepsilon} \rceil < \frac{(\log \frac{1}{\varepsilon})^2}{2\varepsilon} + 1$. Thus, we can substitute $r = \frac{(\log \frac{1}{\varepsilon})^2 + 2\varepsilon}{2\varepsilon}$ to

(13). Since

$$\frac{1}{20 \left(2 \left\lceil \frac{(\log \frac{1}{\varepsilon})^2}{4\varepsilon} \right\rceil \right) + 20 \left(2 \left\lceil \frac{(\log \frac{1}{\varepsilon})^2}{4\varepsilon} \right\rceil \right)^3} \geq \frac{1}{20} \varepsilon^3 \left(\log \frac{1}{\varepsilon} \right)^{-6},$$

we obtain the desired result.

Case 2. $p^{\frac{1}{2}} \leq t \leq p^{\frac{2}{3}}$.

The argument is similar to that of Case 1, with a few adjustments. First, we replace the bound (5) with the Konyagin–Shparlinski bound from Lemma 7. Next, we replace the conditions for (8) and (11) by assuming $r \geq 6$. Finally, we substitute $r = 6$ to (13) to obtain the desired result. $\square$

Remark 1. For $\varepsilon \geq 0$ and $p^{\frac{1}{3}+\varepsilon} \leq t \leq p^{\frac{1}{2}}$, we can obtain the bound $S_{a,c}(t) \ll t^{2-\frac{\varepsilon^3}{50}}$ by a similar argument presented above with a few adjustments. First, we replace the bound (5) with the Konyagin–Shparlinski bound from Lemma 7. Then, we substitute the conditions for the bounds (8) and (11) with the following assumptions:

$$(8') \quad \varepsilon \geq \frac{4}{3r},$$

$$(11') \quad \varepsilon \geq \frac{60r - 8r^2 + 1}{3(8r^2 - 1)}.$$

Finally, note that it is clear that for each ε there exists r which satisfies both inequalities; we may substitute $r = \frac{4}{3\varepsilon}$ to (13) to obtain the result. Also, note that $\frac{\varepsilon^3}{50} > \frac{1}{20}\varepsilon^3 \left(\log \frac{1}{\varepsilon} \right)^{-6}$ for $\varepsilon \geq \frac{1}{28}$. Thus, $t^{2-\frac{\varepsilon^3}{50}}$ is a better bound for $p^{\frac{29}{84}} \leq t \leq p^{\frac{1}{2}}$.

Theorem 1 and the inequality $|\tilde{S}_{a,b,c}(t)| \leq S_{a,c}(t)$ imply the following corollary.

COROLLARY 1. *Given $p \geq 3$ and $V \in \mathbb{Z}$ of multiplicative order $t \pmod{p}$, we have*

$$\max_{\gcd(a,b,c,p)=1} \left| \tilde{S}_{a,b,c}(t) = \sum_{x,y=1}^t e_p(aV^x + bV^y + cV^{xy}) \right| \ll t^{2-r(\varepsilon)},$$

where

$$r(\varepsilon) = \begin{cases} \frac{1}{4440} & \text{if } p^{\frac{1}{2}} \leq t \leq p^{\frac{2}{3}}, \\ \frac{1}{20}\varepsilon^3 \left(\log \frac{1}{\varepsilon} \right)^{-6} & \text{if } p^{\frac{1}{4}+\varepsilon} \leq t < p^{\frac{1}{2}}. \end{cases}$$

Acknowledgment. The authors are thankful to Professor I. E. Shparlinski for his useful comments, suggestions, and references.

REFERENCES

- [1] W. BANKS, A. CONFLITTI, J. FRIEDLANDER, AND I. SHPARLINSKI, *Exponential sums over Mersenne numbers*, *Compositio. Math.*, 140 (2004), pp. 15–30.
- [2] W. BANKS, J. FRIEDLANDER, S. KONYAGIN, AND I. SHPARLINSKI, *Incomplete exponential sums and Diffie-Hellman triples*, *Math. Proc. Cambridge Philos. Soc.*, 140 (2006), pp. 193–206.
- [3] J. BOURGAIN, *Estimates on exponential sums related to the Diffie-Hellman distributions*, *Geom. Funct. Anal.*, 15 (2005), pp. 1–34.

- [4] R. CANETTI, J. FRIEDLANDER, S. KONYAGIN, M. LARSEN, D. LIEMAN, AND I. SHPARLINSKI, *On the statistical properties of Diffie-Hellman distributions*, Israel J. Math., 120 (2000), pp. 23–46.
- [5] R. CANETTI, J. FRIEDLANDER, AND I. SHPARLINSKI, *On certain exponential sums and the distribution of Diffie-Hellman triples*, J. London Math. Soc., (2) 59 (1999), pp. 799–812.
- [6] W. DIFFIE AND M. HELLMAN, *New directions in cryptography*, IEEE Trans. Inform. Theory, 22 (1976), pp. 644–654.
- [7] J. FRIEDLANDER, J. HANSEN, AND I. SHPARLINSKI, *On the distribution of the power generator modulo a prime power*, in Unusual Applications of Number Theory, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., Amer. Math. Soc., Providence, RI, 2004, pp. 71–79.
- [8] J. FRIEDLANDER, S. KONYAGIN, AND I. SHPARLINSKI, *Some doubly exponential sums over $\mathbb{Z}_m$* , Acta Arith., 105 (2002), pp. 349–370.
- [9] J. FRIEDLANDER, M. LARSEN, D. LIEMAN, AND I. SHPARLINSKI, *On the correlation of binary M -sequences*, Des. Codes Cryptogr., 16 (1999), pp. 249–256.
- [10] J. FRIEDLANDER, D. LIEMAN, AND I. SHPARLINSKI, *On the distribution of the RSA generator*, in Sequences and their Applications, (Singapore 1998), Springer Set. Discrete Math. Theor. Comput. Sci., Springer-Verlag, London, 1999, pp. 205–212.
- [11] J. FRIEDLANDER AND I. SHPARLINSKI, *On the distribution of Diffie-Hellman triples with sparse exponents*, SIAM J. Discrete Math., 14 (2001), pp. 162–169.
- [12] M. GARAIEV, *Double exponential sums related to Diffie-Hellman distributions*, Int. Math. Res. Not., 17 (2005), pp. 1005–1014.
- [13] M. GARAIEV AND A. KARATSUBA, *New estimates of double trigonometric sums with exponential functions*, Arch. Math. (Basel), 87 (2006), pp. 33–40.
- [14] S. KONYAGIN, *Estimates of trigonometric sums over subgroups and Gaussian sums*, in The Fourth International Conference on “Modern Problems of Number Theory and its Applications,” dedicated to the 180th Anniversary of P. L. Chebyshev and the 110th Anniversary of I. M. Vinogradov, Topical Problems, Part 3, Department of Mechanics and Mathematics, Moscow Lomonosov State University, Moscow, Russia, 2002, pp. 86–114.
- [15] S. KONYAGIN AND I. SHPARLINSKI, *Character Sums with Exponential Functions and their Applications*, Cambridge Tracts in Math. 136, Cambridge University Press, Cambridge, UK, 1999.
- [16] I. SHPARLINSKI, *Communication complexity and Fourier coefficients of the Diffie-Hellman key*, Lecture Notes in Comput. Sci., 1776 (2000), pp. 259–268.
- [17] H. WEYL, *Über die Gleichverteilung von Zahlen mod Eins*, Math. Ann., 77 (1916), pp. 313–352.

CHARACTERIZING IDEAL WEIGHTED THRESHOLD SECRET SHARING*

AMOS BEIMEL[†], TAMIR TASSA[‡], AND ENAV WEINREB[†]

Abstract. Weighted threshold secret sharing was introduced by Shamir in his seminal work on secret sharing. In such settings, there is a set of users where each user is assigned a positive weight. A dealer wishes to distribute a secret among those users so that a subset of users may reconstruct the secret if and only if the sum of weights of its users exceeds a certain threshold. On one hand, there are nontrivial weighted threshold access structures that have an ideal scheme—a scheme in which the size of the domain of shares of each user is the same as the size of the domain of possible secrets (this is the smallest possible size for the domain of shares). On the other hand, other weighted threshold access structures are not ideal. In this work we characterize all weighted threshold access structures that are ideal. We show that a weighted threshold access structure is ideal if and only if it is a hierarchical threshold access structure (as introduced by Simmons), or a tripartite access structure (these structures generalize the concept of bipartite access structures due to Padró and Sáez), or a composition of two ideal weighted threshold access structures that are defined on smaller sets of users. We further show that in all those cases the weighted threshold access structure may be realized by a linear ideal secret sharing scheme. The proof of our characterization relies heavily on the strong connection between ideal secret sharing schemes and matroids, as proved by Brickell and Davenport.

Key words. ideal secret sharing, matroids, hierarchical access structures, tripartite access structures

AMS subject classifications. 94A60, 94A62

DOI. 10.1137/S0895480104445654

1. Introduction. A *threshold secret sharing scheme* enables a dealer to distribute a secret among a set of users, by giving each user a piece of information called a *share*, such that only large sets of users will be able to reconstruct the secret from the shares that they got, while smaller sets gain no information on the secret. Threshold secret sharing schemes were introduced and efficiently implemented, independently, by Blakley [5] and Shamir [28]. Efficient threshold secret sharing schemes were used in many cryptographic applications, e.g., Byzantine agreement [26], secure multiparty computations [3, 10], and threshold cryptography [13].

In this paper we deal with *weighted* threshold secret sharing schemes. In these schemes, considered already by Shamir [28], the users are not of the same status. That is, each user is assigned a positive weight, and a set can reconstruct the secret if the sum of weights assigned to its users exceeds a certain threshold. As a motivation, consider sharing a secret among the shareholders of some company, each holding a different number of shares. Such settings are closely related to the concept of *weighted threshold functions*, which play an important role in complexity theory and learning theory.¹

*Received by the editors August 12, 2006; accepted for publication (in revised form) October 18, 2007; published electronically March 7, 2008.

<http://www.siam.org/journals/sidma/22-1/44565.html>

[†]Department of Computer Science, Ben-Gurion University, Beer-Sheva 84105, Israel (beimel@cs.bgu.ac.il, www.cs.bgu.ac.il/~beimel; weinrebe@cs.bgu.ac.il).

[‡]Division of Computer Science, The Open University, Ra'anana, Israel, and Department of Computer Science, Ben-Gurion University, Beer-Sheva 84105, Israel (tamirta@openu.ac.il).

¹A weighted threshold function is a Boolean function $f : \{0, 1\}^n \rightarrow \{0, 1\}$, where each variable is assigned a positive weight and $f(x_1, \dots, x_n) = 1$ if and only if the sum of weights that are assigned to the variables whose value is 1 exceeds a given threshold.

Ito, Saito, and Nishizeki [16] generalized the notion of secret sharing such that there is an arbitrary monotone collection of authorized sets, called the *access structure*. The requirements are that only sets in the access structure are allowed to reconstruct the secret, while sets that are not in the access structure should gain no information on the secret. A simple argument shows that in every secret sharing scheme the domain of possible shares for each user is at least as large as the domain of possible secrets (see [19]). Shamir's threshold secret sharing scheme is *ideal* in the sense that the domain of shares of each user coincides with the domain of possible secrets. Ideal secret sharing schemes are the most space-efficient schemes. However, some access structures do not have an ideal secret sharing scheme that realizes them [4]. Namely, some access structures demand share domains that are larger than the domain of secrets. Access structures that may be realized by an ideal secret sharing scheme are called ideal. Ideal secret sharing schemes and ideal access structures have been studied in, e.g., [1, 6, 7, 17, 20, 21, 23, 25, 27, 32, 35, 36]. Ideal access structures are known to have certain combinatorial properties. In particular, there is a strong relation between ideal access structures and matroids [7].

While threshold access structures are ideal, weighted threshold access structures are not necessarily so. For example, the access structure on four users with weights 1, 1, 1, and 2, and a threshold of 3, has no ideal secret sharing scheme (see Example 4.9 for a proof). Namely, in any perfect secret sharing scheme that realizes this access structure, the share domain of at least one user is larger than the domain of secrets. On the other hand, there exist ideal weighted threshold access structures other than the trivial threshold ones. For example, consider the access structure on nine users, where the weights are 16, 16, 17, 18, 19, 24, 24, 24, and 24 and the threshold is 92. Even though this access structure seems more complicated than the previous one, it has an ideal secret sharing scheme (see Example 7.5). Another example of an ideal weighted threshold access structure is the one having weights 1, 1, 1, 1, 1, 3, 3, and 3 and threshold 6 (see Example 8.10).

In this paper we give a combinatorial characterization of ideal weighted threshold access structures. We show that if a weighted threshold access structure is ideal, then one of the following statements holds:

1. It is a *multilevel* or *hierarchical* threshold access structure. This type of access structure was introduced by Simmons in [29]. In such settings, the users are divided into a hierarchy of levels, and each level has an associated threshold. A set of users is authorized if it has a subset whose size is at least the threshold that corresponds to the level of the lowest member in that subset.
2. It is a *tripartite* access structure. In such access structures, the set of users is partitioned into three classes, and a given subset is authorized if and only if the sizes of its intersection with each of the three classes satisfy some threshold conditions (see section 7 for details).
3. The set of users is composed of *strong* users (users with larger weights) and *weak* users (users with smaller weights), and the access structure is a combinatorial composition of two ideal weighted threshold access structures, one that is defined on the subset of strong users and another that is defined on the subset of weak users.

(Formal definitions are given in section 3.)

Most of the access structures that play a part in our characterization are well known [25, 29, 35]. Tripartite access structures were considered by Collins [11], who gave a necessary condition for a tripartite access structure to be ideal. We consider a

specific family of tripartite access structures and design ideal secret sharing schemes for this family. Following our work, Herranz and Sáez [15] constructed ideal secret sharing schemes for a wider family of tripartite access structures, and Farràs, Martí-Farré, and Padró [14] exactly characterizes the tripartite access structures that are ideal.

The present study generalizes the work of Morillo et al. [23], who characterized the ideal weighted threshold access structures in which all the minimal authorized sets have at most two users. The proof of our characterization relies heavily on the strong connection between ideal secret sharing schemes and matroids, as presented in [7]. We utilize results regarding the structure of matroids in order to understand the structure of ideal weighted threshold access structures. An important tool in our analysis is composition of ideal access structures. When composing two access structures, defined on two disjoint sets, one gets an access structure on a larger set of users. The resulting access structure is ideal if and only if the original two are ideal. This enables us to characterize ideal weighted threshold access structures in a recursive manner, as described above. Composition of access structures was studied, e.g., in [1, 4, 8, 12, 22, 34].

Related work. Secret sharing schemes for general access structures were defined by Ito, Saito, and Nishizeki in [16]. More efficient schemes were presented in, e.g., [4, 6, 18, 31]. We refer the reader to [30, 33] for extensive surveys on secret sharing schemes. However, for most access structures the known secret sharing schemes are highly inefficient; that is, the size of the shares is exponential in n , the number of users. It is not known whether better schemes exist. For weighted threshold access structures the situation is somewhat better. Shamir [28] proposed a weighted threshold secret sharing scheme that is based on his ideal threshold secret sharing scheme. In that scheme, the ratio between the share size of each user and the size of the secret equals the weight assigned to that user, which may be exponential in the number of users n . In a recent work [2], secret sharing schemes were constructed for arbitrary weighted threshold access structures in which the shares are of size $O(n^{\log n})$. Furthermore, under reasonable computational assumptions, a secret sharing scheme with computational security was constructed in [2] for every weighted threshold access structure with a polynomial share size.

Organization. We begin in section 2 by supplying the necessary definitions. Then, in section 3, we state our characterization theorem and outline its proof. We proceed to describe in section 4 the connection between matroids and ideal secret sharing, and then prove, in section 5, several properties of matroids that are associated with weighted threshold access structures. Thereafter, we discuss the connection between ideal weighted threshold access structures and two families of access structures: hierarchical threshold access structures in section 6, and tripartite access structures in section 7. Finally, in section 8 we complete the proof of the characterization theorem by proving that if an ideal weighted threshold access structure is neither hierarchical nor tripartite, then it is a composition of two ideal weighted threshold access structures on smaller sets of users.

2. Definitions and notations.

2.1. Secret sharing. We next define secret sharing schemes and ideal secret sharing schemes. We start by defining an access structure which is the collection of authorized sets that can reconstruct the secret.

DEFINITION 2.1 (access structure). Let $U = \{u_1, \dots, u_n\}$ be a set of users. A collection $\Gamma \subseteq 2^U$ is monotone if $B \in \Gamma$ and $B \subseteq C$ imply that $C \in \Gamma$. An access structure is a monotone collection $\Gamma \subseteq 2^U$ of nonempty subsets of U . Sets in Γ are called authorized, and sets not in Γ are called unauthorized. A set B is called a minterm of Γ if $B \in \Gamma$ and for every $C \subsetneq B$ the set C is unauthorized. A user u is called self-sufficient if $\{u\} \in \Gamma$. A user is called redundant if there is no minterm that contains it. An access structure is called connected if it has no redundant users.

DEFINITION 2.2 (secret sharing scheme). Let S be a finite set of secrets, where $|S| \geq 2$. An n -user secret sharing scheme Π with domain of secrets S is a randomized mapping from S to a set of n -tuples $\prod_{i=1}^n S_i$, where S_i is called the share-domain of u_i . A dealer shares a secret $s \in S$ among the n users of some set U according to Π by first sampling (according to some predefined probability distribution) a vector of shares $\Pi(s) = (s_1, \dots, s_n) \in \prod_{i=1}^n S_i$, and then privately communicating each share s_i to the user u_i . We say that Π realizes an access structure $\Gamma \subseteq 2^U$ if the following two requirements hold:

1. *Correctness.* The secret s can be reconstructed by any authorized set of users. That is, for any set $B \in \Gamma$ (where $B = \{u_{i_1}, \dots, u_{i_{|B|}}\}$), there exists a reconstruction function² $\text{RECON}_B : S_{i_1} \times \dots \times S_{i_{|B|}} \rightarrow S$ such that for every $s \in S$ and for every possible value of $\Pi_B(s)$ the restriction of $\Pi(s)$ to its B -entries,

$$\text{RECON}_B(\Pi_B(s)) = s.$$

2. *Privacy.* Every unauthorized set can learn nothing about the secret (in the information theoretic sense) from the shares of the users in the set. Formally, for any set $C \notin \Gamma$, for every two secrets $a, b \in S$, and for every possible $|C|$ -tuple of shares $\langle s_i \rangle_{u_i \in C}$:

$$\Pr[\Pi_C(a) = \langle s_i \rangle_{u_i \in C}] = \Pr[\Pi_C(b) = \langle s_i \rangle_{u_i \in C}].$$

In every secret sharing scheme, the size of the domain of shares of each user is at least the size of the domain of the secrets [19], namely $|S_i| \geq |S|$ for all $i \in [n]$. This motivates the next definition.

DEFINITION 2.3 (ideal access structure). A secret sharing scheme with domain of secrets S is ideal if the domain of shares of each user is S . An access structure Γ is ideal if for some finite domain of secrets S there exists an ideal secret sharing scheme realizing it.

Most previously known secret sharing schemes are *linear*. The concept of linear secret sharing schemes was introduced by Brickell [6] in the ideal setting and was later generalized to nonideal schemes. Linear schemes are equivalent to monotone span programs [18]. For simplicity we define only ideal linear schemes.

DEFINITION 2.4 (ideal linear secret sharing scheme). Let $\mathbb{F}$ be a finite field. An ideal linear secret sharing scheme over $\mathbb{F}$ takes the following form. The domain of secrets and shares is $S = \mathbb{F}$. The scheme is specified by $n + 1$ vectors in $\mathbb{F}^t$ for some integer t : a vector $\mathbf{T}_i$ per user $u_i \in U$, where $i \in [n]$, and a so-called target vector $\mathbf{T}$. To share a secret $s \in \mathbb{F}$, the dealer chooses a random vector $\mathbf{R} \in \mathbb{F}^t$ such that $\mathbf{R} \cdot \mathbf{T} = s$, and then the share of user u_i is $\mathbf{R} \cdot \mathbf{T}_i$.

²We stress that there is no requirement that RECON_B can be efficiently implemented. However, this is the case for all the secret sharing schemes designed in this paper.

The next theorem characterizes the access structure that is realized by a linear secret sharing scheme.

THEOREM 2.5 (see [6, 18]). *A linear secret sharing scheme with vectors $\{\mathbf{T}_i\}_{i \in [n]}$ and $\mathbf{T}$ realizes the following access structure Γ :*

$$A \in \Gamma \text{ if and only if } \mathbf{T} \in \text{span}\{\mathbf{T}_i : u_i \in A\}.$$

2.2. Weighted threshold access structures. In this paper we concentrate on special access structures, so-called weighted threshold access structures, which were already introduced in [28].

DEFINITION 2.6 (weighted threshold access structure—WTAS). *Let $w : U \rightarrow \mathbb{N}$ be a weight function on U and $T \in \mathbb{N}$ be a threshold. Define $w(A) := \sum_{u \in A} w(u)$ and $\Gamma = \{A \subseteq U : w(A) \geq T\}$. Then Γ is called a weighted threshold access structure (WTAS) on U .*

It is easy to see that Definition 2.6 does not restrict generality when assuming that the weights and the threshold are integers. In other words, given a WTAS with real positive weights and threshold, there exist integer weights and threshold that induce the very same access structure.

2.3. Terminology and notation. First, for a positive integer n , we denote by $[n]$ the set $\{1, 2, \dots, n\}$. Throughout this paper we assume that the users are ordered in a nondecreasing order according to their weights, i.e.,

$$0 < w(u_1) \leq w(u_2) \leq \dots \leq w(u_n).$$

Let $A = \{u_{i_j}\}_{1 \leq j \leq k}$ be an ordered subset of U , where $1 \leq i_1 < \dots < i_k \leq n$. In order to avoid two-leveled indices, we will denote the users in such a subset with the corresponding lowercase letter, namely, $A = \{a_j\}_{1 \leq j \leq k}$. We denote the first (lightest) and last (heaviest) users of A by $A_{\min} = a_1$ and $A_{\max} = a_k$, respectively. For an arbitrary ordered subset A we let $A_{s,t} = \{a_j\}_{s \leq j \leq t}$ denote a run-subset. If $s > t$, then $A_{s,t} = \emptyset$. Two types of runs that we shall meet frequently are prefixes and suffixes. A prefix of a subset A is a run-subset of the form $A_{1,\ell}$, while a suffix takes the form $A_{\ell,k}$, where $1 \leq \ell \leq k$. A suffix $A_{\ell,k}$ is a proper suffix of $A_{1,k}$ if $\ell > 1$.

We conclude this section by introducing the precedence relation $\prec$. When applied to users, $u_i \prec u_j$ indicates that $i < j$ (and, in particular, $w(u_i) \leq w(u_j)$). This relation induces the following lexicographic order on subsets of U :

- $\emptyset \prec A$ for all $A \subseteq U$ such that $A \neq \emptyset$.
- If $a_1 \prec b_1$, then $A \prec B$; if $b_1 \prec a_1$, then $B \prec A$; otherwise, $A \prec B$ if and only if $(A \setminus \{a_1\}) \prec (B \setminus \{b_1\})$.

3. Characterizing ideal WTASs. The main result of this paper is a combinatorial characterization of ideal WTASs. We define in Definitions 3.1–3.4 the building blocks that play an essential role in this characterization. Using these definitions, we state Theorem 3.5, our main result, that characterizes ideal WTASs. We then outline the proof of that theorem, where the full proof is given in the subsequent sections.

3.1. Building blocks.

DEFINITION 3.1 (hierarchical threshold access structure—HTAS). *Let m be an integer, U be a set of users, and $\{L_i\}_{1 \leq i \leq m}$ be a partition of U into a hierarchy of m disjoint levels, i.e., $U = \bigcup_{i=1}^m L_i$. Furthermore, let $\{k_i\}_{1 \leq i \leq m}$ be a sequence of decreasing thresholds, $k_1 > k_2 > \dots > k_m$. This hierarchy and sequence of thresholds*

induces a hierarchical threshold access structure (HTAS) on U :

$$\Gamma_H = \left\{ A \subseteq U : \text{There exists } i \in [m] \text{ such that } \left| A \cap \bigcup_{j=i}^m L_j \right| \geq k_i \right\}.$$

That is, a set $A \subseteq U$ is in Γ_H if and only if it contains at least k_i users from the i th level and above, for some $i \in [m]$.

The family of HTASs was introduced by Simmons in [29] and further studied by Brickell, who proved their ideality [6]. An explicit ideal scheme for these access structures was constructed in [35].

Remark 3.2. Without loss of generality, we can assume that $|L_i| > k_i - k_{i+1}$ for every $i \in [m-1]$, and $|L_m| \geq k_m$. Indeed, if $|L_i| \leq k_i - k_{i+1}$ for some $i \in [m-1]$, then the i th threshold condition in the HTAS definition implies the $(i+1)$ th threshold condition, and, consequently, the $(i+1)$ th condition is redundant.

We next present tripartite access structures (TPASs), which are a generalization of bipartite access structures that were presented in [25]. In a *bipartite* access structure, the set of users is partitioned into two classes, and a subset is authorized if the sizes of its intersection with each of these classes satisfy some predefined conditions. In a *tripartite* access structure, the set of users is partitioned into three classes, and, similarly to bipartite access structures, a given subset is authorized if the sizes of its intersection with each of these classes satisfy some conditions. In the next definition we describe two special types of TPASs. We would like to stress that there exist other types of TPASs, but herein we concentrate only on those two types, since only they are relevant for the characterization of ideal WTASs. Later on we prove that these types of TPASs are ideal by constructing a linear ideal secret sharing scheme that realizes them. That scheme is a generalization of a scheme from [25] for bipartite access structures. Following our work, Farràs, Martí-Farré, and Padró [14] exactly characterized the TPASs that are ideal.

DEFINITION 3.3 (tripartite access structure—TPAS). *Let U be a set of n users such that $U = A \cup B \cup C$, where A , B , and C are disjoint and A and C are nonempty. Let m, d, t be positive integers such that $m \geq t$. Then the following is a tripartite access structure (a TPAS) on U :*

$$\Delta_1 = \{X \subseteq U : (|X| \geq m \text{ and } |X \cap (B \cup C)| \geq m - d) \text{ or } |X \cap C| \geq t\}.$$

Namely, a set X is in Δ_1 if either it has at least m users, $(m-d)$ of which are from $B \cup C$, or it has at least t users from C . If $|B| \leq d + t - m$, then the following is another type of TPAS:

$$\Delta_2 = \{X \subseteq U : (|X| \geq m \text{ and } |X \cap C| \geq m - d) \text{ or } |X \cap (B \cup C)| \geq t\}.$$

That is, a set X is in Δ_2 if either it has at least m users, $(m-d)$ of which are from C , or it has at least t users from $B \cup C$.

DEFINITION 3.4 (composition of access structures). *Let U_1 and U_2 be disjoint sets of users, and let Γ_1 and Γ_2 be access structures on U_1 and U_2 , respectively. Let $u_1 \in U_1$, and set $U = U_1 \cup U_2 \setminus \{u_1\}$. Then the composition of Γ_1 and Γ_2 via u_1 is*

$$\Gamma = \left\{ X \subseteq U : \begin{array}{l} X_1 \in \Gamma_1 \text{ or } (X_2 \in \Gamma_2 \text{ and } X_1 \cup \{u_1\} \in \Gamma_1), \\ \text{where } X_1 = X \cap U_1 \text{ and } X_2 = X \cap U_2 \end{array} \right\};$$

namely, $X \subseteq U$ is authorized in this access structure if either $X_1 = X \cap U_1$ is authorized in Γ_1 , or $X_1 \cup \{u_1\}$ is authorized in Γ_1 and $X_2 = X \cap U_2$ is authorized in Γ_2 .

The motivation behind this definition is as follows. We start with an access structure Γ_1 on U_1 . Then we allow the replacement of one of the users in U_1 , denoted u_1 , by substitute users of a disjoint set U_2 , where Γ_2 specifies the subsets of U_2 that are legitimate substitutes for u_1 . Hence, in the set $U = U_1 \cup U_2 \setminus \{u_1\}$ that is obtained from U_1 after replacing u_1 with U_2 , a subset is authorized if either it is a subset in Γ_1 that does not include u_1 or it was obtained from a subset in Γ_1 that did include u_1 , by replacing u_1 with an authorized subset of substitutes from Γ_2 . In Lemma 8.1 we show that Γ is ideal if and only if both Γ_1 and Γ_2 are ideal.

3.2. The characterization. Following the notation convention in section 2.3, the set U is viewed as a sequence which is ordered in a monotonic nondecreasing order according to the weights. Let M be the lexicographically minimal minterm of Γ (that is, $M \in \Gamma$ is a minterm, and $M \prec M'$ for all other minterms $M' \in \Gamma$). It turns out that the form of M plays a significant role in the characterization of an ideal WTAS Γ .

If M is a prefix of U , namely $M = U_{1,k} = \{u_1, \dots, u_k\}$ for some $k \in [n]$, then, as we prove in section 6.3, the access structure is an HTAS of at most three levels. If M is a lacunary prefix, in the sense that $M = U_{1,k} \setminus \{u_\ell\}$ for $1 \leq \ell < k \leq n$, then, as we prove in section 7, the access structure is a TPAS. Otherwise, if M is neither a prefix nor a lacunary prefix, the access structure is a composition of two WTASs on smaller sets. More specifically, we identify a prefix $U_{1,k}$, where $1 < k < n$, that could be replaced by a single substitute user u , and then show that Γ is a composition of a WTAS on $U_{1,k}$ and another WTAS on $U_{k+1,n} \cup \{u\}$. Since Γ is ideal, so are the two smaller WTASs, as implied by Lemma 8.1. Hence, this result, which we prove in section 8, completes the characterization of ideal WTASs in a recursive manner.

Hence, our main result in this paper is as follows.

THEOREM 3.5 (the characterization theorem). *Let U be a set of users, $w : U \rightarrow \mathbb{N}$ be a weight function, T be a threshold, and Γ be the corresponding WTAS. Then Γ is ideal if and only if one of the following three conditions holds:*

- *The access structure Γ is an HTAS of at most three levels.*
- *The access structure Γ is a TPAS.*
- *The access structure Γ is a composition of Γ_1 and Γ_2 , where Γ_1 and Γ_2 are ideal WTASs defined on sets of users smaller than U .*

In particular, if Γ is an ideal WTAS, then there exists a linear ideal secret sharing scheme that realizes it.

Remark 3.6. In the ideal schemes for WTASs that we construct in this paper, the size of the domain of secrets is at most $2^{\text{poly}(n)}$. This is the field size needed for the HTAS constructions of [35] and for the TPAS construction introduced in section 7. As the composition of ideal access structures does not change the size of the underlying field, we get that the field size for our entire construction is bounded by $2^{\text{poly}(n)}$. This is a comfortable bound since it implies that both the secret and the shares may be represented by $\text{poly}(n)$ bits.

4. Matroids and ideal secret sharing schemes. Matroids and ideal secret sharing schemes are strongly related. If an access structure is ideal, then there is a matroid that reflects its structure. On the other hand, every matroid that is representable over some finite field is the reflection of some ideal access structure. In this

section we review some basic results from the theory of matroids and describe their relation to ideal secret sharing schemes. For more background on matroid theory the reader is referred to [24].

Matroids are a combinatorial structure that generalizes both linear spaces and the set of circuits in an undirected graph. They are a useful tool in several fields of theoretical computer science, e.g., optimization algorithms. A matroid $\mathcal{M} = \langle V, \mathcal{I} \rangle$ is a finite set V and a collection $\mathcal{I}$ of subsets of V that satisfy the following three axioms:

- (I1) $\emptyset \in \mathcal{I}$.
- (I2) If $X \in \mathcal{I}$ and $Y \subseteq X$, then $Y \in \mathcal{I}$.
- (I3) If X and Y are members of $\mathcal{I}$ with $|X| = |Y| + 1$, then there exists an element $x \in X \setminus Y$ such that $Y \cup \{x\} \in \mathcal{I}$.

The elements of V are called the *points* of the matroid, and the sets in $\mathcal{I}$ are called the *independent* sets of the matroid. Axiom (I1) assures that there is at least one independent set in $\mathcal{I}$. Axiom (I2) asserts that the collection $\mathcal{I}$ is closed under containment. Finally, axiom (I3) enables the expansion of every small independent set in $\mathcal{I}$. A *dependent set* of the matroid is any subset of V that is not independent. The minimal dependent sets are called *circuits*. A matroid is said to be *connected* if for any two points there exists a circuit that contains both of them.

We now discuss the relations between ideal secret sharing schemes and matroids. Let Γ be an access structure on a set of users $U = \{u_1, \dots, u_n\}$. If Γ is ideal, then, by the results of [7, 21], there exists a matroid $\mathcal{M}$ corresponding to Γ . The points of $\mathcal{M}$ are the users in U together with an additional point, denoted u_0 , that could be thought of as representing the dealer. We denote hereinafter by

$$\mathcal{C}_0 = \{X \cup \{u_0\} : X \text{ is a minterm of } \Gamma\}$$

the set of all Γ -minterms, supplemented by u_0 .

THEOREM 4.1 (see [7, 21]). *Let Γ be a connected ideal access structure on U . Then there exists a connected matroid $\mathcal{M}$ on $U \cup \{u_0\}$ such that $\mathcal{C}_0$ is exactly the set of circuits of $\mathcal{M}$ containing u_0 .*

The next result implies the uniqueness of the matroid $\mathcal{M}$ that corresponds to a given connected ideal access structure, as discussed in Theorem 4.1, and, additionally, it provides means to identify *all* the circuits of that matroid.

LEMMA 4.2 (see [24, Theorem 4.3.2]). *Let e be an element of a connected matroid $\mathcal{M}$, and let $\mathcal{C}_e$ be the set of circuits of $\mathcal{M}$ that contain e . Then all of the circuits of $\mathcal{M}$ that do not contain e are the minimal sets of the form*

$$(C_1 \cup C_2) \setminus \bigcap \{C_3 : C_3 \in \mathcal{C}_e, C_3 \subseteq C_1 \cup C_2\},$$

where C_1 and C_2 are distinct circuits in $\mathcal{C}_e$.

The unique matroid whose existence and uniqueness are guaranteed by Theorem 4.1 and Lemma 4.2 is referred to as *the matroid corresponding to Γ* . This next definition will enable us to explicitly define the matroid corresponding to Γ using the authorized sets in Γ .

DEFINITION 4.3 (critical user). *Let M_1 and M_2 be distinct minterms of Γ . A user $x \in M_1 \cup M_2$ is critical for $M_1 \cup M_2$ if the set $M_1 \cup M_2 \setminus \{x\}$ is unauthorized. In addition, we define*

$$D(M_1, M_2) = (M_1 \cup M_2) \setminus \{x \in M_1 \cup M_2 : x \text{ is critical for } M_1 \cup M_2\}.$$

With this definition, the following corollary is a straightforward result of Theorem 4.1 and Lemma 4.2.

COROLLARY 4.4. *Let Γ be a connected ideal access structure on U . Then there exists a unique connected matroid $\mathcal{M}$ on $U \cup \{u_0\}$ such that $\mathcal{C}_0$ is exactly the set of circuits of $\mathcal{M}$ containing u_0 . Furthermore, the minimal sets in*

$$\{D(M_1, M_2) : M_1, M_2 \text{ are distinct minterms of } \Gamma\}$$

are the circuits of the matroid that do not contain u_0 .

In what follows we use the following simple consequence of Corollary 4.4.

COROLLARY 4.5. *Let M_1 and M_2 be two distinct minterms of Γ . Then $D(M_1, M_2)$ is a dependent set of $\mathcal{M}$.*

Note that $D(M_1, M_2)$ is a dependent set of $\mathcal{M}$, but is not necessarily a circuit of $\mathcal{M}$. The following basic result from matroid theory provides another way of establishing the dependency of a set in a matroid.

LEMMA 4.6 (see [24, p. 16]). *Let C_1 and C_2 be two distinct circuits in a matroid, $e \in C_1 \cap C_2$, and $d \in C_1 \setminus C_2$ such that $d \neq e$. Then there exists a circuit $C_3 \subseteq (C_1 \cup C_2) \setminus \{e\}$, where $d \in C_3$.*

In most applications of the lemma we will ignore d and the fact that $d \in C_3$. Since every minterm of Γ is properly contained in a circuit of $\mathcal{M}$, it forms an independent set of $\mathcal{M}$. However, these are not necessarily the only independent sets in $\mathcal{M}$. The following claim is a simple corollary of axiom (I3) in the matroid definition.

CLAIM 4.7. *Let $\mathcal{M}$ be a matroid and I be an independent set of size m . Then any independent set of size $\ell < m$ can be extended to an independent set of size m .*

Finally, the next lemma is applicable when adding an element to an independent set results in a dependent set.

LEMMA 4.8. *Let I be an independent set in a matroid $\mathcal{M}$, and let e be an element of $\mathcal{M}$ such that $I \cup \{e\}$ is dependent. Then $\mathcal{M}$ has a unique circuit contained in $I \cup \{e\}$, and that circuit contains e .*

To prove that an access structure is not ideal we use the above results that define the dependent sets in the matroid corresponding to Γ . In particular, we use the basic fact that an independent set of $\mathcal{M}$ cannot contain a dependent set (see axiom (I2)). By Theorem 4.1, for every minterm $M \in \Gamma$, the set $M \cup \{u_0\}$ is a circuit of $\mathcal{M}$. Thus, the set M is an independent set of $\mathcal{M}$. Therefore, if D is a dependent set of $\mathcal{M}$, it cannot be contained in any minterm of Γ . The next example shows how to use the above statements in order to demonstrate that a given access structure is not ideal.

Example 4.9. Consider the WTAS Γ on the set $U = \{u_1, u_2, u_3, u_4\}$ with weights $w(u_1) = w(u_2) = w(u_3) = 1$ and $w(u_4) = 2$ and threshold $T = 3$. The minterms of Γ are $\{u_1, u_2, u_3\}$, $\{u_1, u_4\}$, $\{u_2, u_4\}$, and $\{u_3, u_4\}$. It follows from Benaloh and Leichter [4] that this access structure is not ideal.³ We first give a simple proof, based on Corollary 4.5, showing that Γ is not ideal. Assume that it is ideal and consider the minterms $M_1 = \{u_1, u_4\}$ and $M_2 = \{u_2, u_4\}$. The set $\{u_1, u_2\}$ is unauthorized, and thus u_4 is critical for $M_1 \cup M_2$. On the other hand, the users u_1 and u_2 are not critical for $M_1 \cup M_2$. Therefore, by Corollary 4.5, the set $D(M_1, M_2) = \{u_1, u_2\}$ is a dependent set of $\mathcal{M}$, the matroid corresponding to Γ . However, the set $\{u_1, u_2, u_3\}$ is a minterm of Γ and, consequently, it is independent in $\mathcal{M}$. Since $\{u_1, u_2\} \subset \{u_1, u_2, u_3\}$, we

³In [9] it was shown that if the domain of secrets is S , the size of the domain of shares of at least one user in that access structure must be at least $|S|^{1.5}$. That result improved upon previous bounds that were derived in [8].

arrive at the absurd conclusion that a dependent set is contained in an independent set. This contradiction implies that Γ is not ideal.

We proceed to sketch an alternative proof for the nonideality of Γ , similar to the proof that was given in [4]. We present this proof for two reasons. First, we want to show that using matroids simplifies and shortens the proofs substantially. Matroid-based proofs, however, tend to be technical and, consequently, not very intuitive. The alternative proof that we present herein reveals some of the intuition that is sometimes obfuscated in matroid-based proofs.

Assume towards a contradiction that Γ has an ideal secret sharing scheme over some domain of secrets S . Letting s_i denote the share of user u_i for $1 \leq i \leq 4$, we will show that s_2 may be computed from s_1 . (This statement is equivalent to the statement in the first proof that the set $\{u_1, u_2\}$ is dependent.) Having shown that, we will conclude that $\{u_1, u_3\}$ can reconstruct the secret (since $\{u_1, u_2, u_3\}$ is an authorized set and u_2 is unnecessary in the presence of u_1), in contradiction to the definition of Γ .

We need to show that s_2 is uniquely determined by s_1 . Assume, by contradiction, that there exists $s_1 \in S_1$ and $s_2, s'_2 \in S_2$, where $s_2 \neq s'_2$, such that both assignments of shares to $\{u_1, u_2\}$, namely $\langle s_1, s_2 \rangle$ and $\langle s_1, s'_2 \rangle$, are possible. Since $\{u_1, u_2\}$ is unauthorized, that set cannot learn any information on the secret from its shares. Hence, given the pair of shares $\langle s_1, s_2 \rangle$, every secret is possible. Hence, for every $s \in S$, there exists at least one share $s_4(s)$ for u_4 such that $\langle s_1, s_2, s_4(s) \rangle$ is a possible assignment of shares to $\{u_1, u_2, u_4\}$, given the secret s . On the other hand, since the shares of $\{u_1, u_2, u_4\}$ determine the secret, $s_4(s)$ must be unique, since otherwise the domain of shares of u_4 would be larger than the domain of secrets, in a contradiction with ideality. Similarly, for every $s \in S$ there exists a unique share $s'_4(s)$ such that $\langle s_1, s'_2, s'_4(s) \rangle$ is another possible assignment of shares to $\{u_1, u_2, u_4\}$, given the secret s .

In view of the above, the sets $\{s_4(s) : s \in S\}$ and $\{s'_4(s) : s \in S\}$ are of size $|S|$. However, since they are both subsets of S_4 and, by ideality, $|S_4| = |S|$, they both equal S_4 . One conclusion is that each secret $s \in S$ is consistent with each share $s_4 \in S_4$. Another conclusion is that there are secrets $a, b \in S$ such that $s_4(a) = s'_4(b) := s_4$. If $a \neq b$, then the pair of shares $\langle s_1, s_4 \rangle$ is insufficient to determine whether the secret is a or b , in contradiction to the correctness requirement that $\{u_1, u_4\} \in \Gamma$ should be able to reconstruct the secret. Thus, $a = b$, whence both pairs of shares $\langle s_2, s_4 \rangle$ and $\langle s'_2, s_4 \rangle$ are possible assignments of shares to $\{u_2, u_4\}$, given that the secret is a . But this implies that the size of S_2 is at least $|S| + 1$ (by going over all secrets $s \in S$ and looking for a corresponding assignment of shares where the share of u_4 is s_4), thus contradicting ideality.

4.1. Restrictions.

DEFINITION 4.10 (restriction). *Let $Y, X \subseteq U$ be two disjoint subsets of users. The restriction of Γ that is induced by Y on X is defined as the following access structure:*

$$\Gamma_{Y,X} = \{Z \subseteq X : Z \cup Y \in \Gamma\}.$$

In other words, $\Gamma_{Y,X}$ consists of all subsets of X that complete Y to an authorized set in Γ . Since $\Gamma_{Y,X}$ is defined on a smaller set of users, restrictions can be helpful in recursively characterizing the structure of Γ . The following known result assures us that if Γ is ideal, $\Gamma_{Y,X}$ is ideal as well. For completeness, we give a proof herein.

LEMMA 4.11. *Let Γ be an access structure on a set of users U . Let $Y, X \subseteq U$ be sets such that $X \cap Y = \emptyset$. If Γ is ideal, then $\Gamma_{Y,X}$ is ideal.*

Proof. Let $Y = \{u_{i_1}, \dots, u_{i_{|Y|}}\}$. We construct an ideal secret sharing scheme Π' for $\Gamma_{Y,X}$ using an ideal scheme Π for Γ . Fix an arbitrary secret $s' \in S$, and let $s_Y = \langle s'_{i_1}, \dots, s'_{i_{|Y|}} \rangle = \Pi_Y(s')$ be a possible vector of shares for the users of Y . Let $s \in S$ be a secret to be distributed by Π' . The scheme Π' randomly chooses an n -tuple $\Pi(s) = \langle s_1, \dots, s_n \rangle$ such that $s_{i_j} = s'_{i_j}$ for all $1 \leq j \leq |Y|$ (such a selection is possible because Y is unauthorized, whence every assignment of shares to Y is consistent with every secret). Every user $u_j \in X$ is then assigned by Π' the share s_j . (The shares of Y , namely $\langle s'_{i_1}, \dots, s'_{i_{|Y|}} \rangle$, are considered “publicly known.”)

The scheme is correct since if $Z \in \Gamma_{Y,X}$, then $Y \cup Z \in \Gamma$, and thus the members of Z , who know the shares s_Y of Y , can reconstruct the secret as in the scheme Π for Γ . Similarly, the scheme is private, since if $Z \notin \Gamma_{Y,X}$, then $Y \cup Z \notin \Gamma$, and thus the users of Z , who know the shares of Y , will not have any information regarding the secret. The scheme is ideal since the share-domain of all the users is the same as in the original ideal scheme. $\square$

The following result gives a relation between the matroids corresponding to Γ and $\Gamma_{Y,X}$.

LEMMA 4.12. *Let Γ be an ideal access structure on a set of users U . Let $Y, X \subseteq U$ be sets such that $X \cap Y = \emptyset$ and Y is independent in the matroid corresponding to Γ . Then if a set $I \subseteq X$ is independent in the matroid corresponding to $\Gamma_{Y,X}$, it is also independent in the matroid corresponding to Γ .*

Proof. Let $\mathcal{M} = \langle U \cup \{u_0\}, \mathcal{I} \rangle$ be the matroid associated with Γ . We define the matroid $\mathcal{M}' = \langle X \cup \{u_0\}, \mathcal{I}' \rangle$ as follows: A set $Z \subseteq X \cup \{u_0\}$ is in $\mathcal{I}'$ if and only if $Y \cup Z \in \mathcal{I}$. This is indeed a matroid: axiom (I1) holds because $Y \in \mathcal{I}$ so that $\emptyset \in \mathcal{I}'$. As for axiom (I2), $Y \cup B \in \mathcal{I}$ for every $B \in \mathcal{I}'$; therefore, if $A \subseteq B$, then $Y \cup A \in \mathcal{I}$, whence $A \in \mathcal{I}'$. Axiom (I3) holds as well. Let $A, B \in \mathcal{I}'$ be such that $|A| > |B|$. Since $Y \cup A \in \mathcal{I}$ and $Y \cup B \in \mathcal{I}$ while $|Y \cup A| > |Y \cup B|$, there must be an element $x \in A \setminus B$ such that $Y \cup (B \cup \{x\}) \in \mathcal{I}$. Therefore, $B \cup \{x\} \in \mathcal{I}'$, as required.

We claim that $\mathcal{M}'$ is the unique matroid corresponding to $\Gamma_{Y,X}$. That is, for every $M \subseteq X$, the set $M \cup \{u_0\}$ is a circuit of $\mathcal{M}'$ if and only if M is a minterm of $\Gamma_{Y,X}$. Proving that, we get that if $I \in \mathcal{I}'$, then $Y \cup I \in \mathcal{I}$, and therefore I itself is independent in $\mathcal{M}$.

Suppose that $M \cup \{u_0\}$ is a circuit of $\mathcal{M}'$. Then $Y \cup M \cup \{u_0\}$ is dependent in $\mathcal{M}$ and consequently must contain a circuit of $\mathcal{M}$. That circuit must contain $M \cup \{u_0\}$ because otherwise there would be a proper subset $B \subset M \cup \{u_0\}$ such that $Y \cup B$ is dependent in $\mathcal{M}$, thus implying that B is dependent also in $\mathcal{M}'$, in a contradiction with our assumption that $M \cup \{u_0\}$ is a circuit of $\mathcal{M}'$. Let $Y' \cup M \cup \{u_0\}$ be that circuit of $\mathcal{M}$, where $Y' \subseteq Y$. Hence, by the characterization of the matroid that corresponds to a given access structure, $Y' \cup M$ is a minterm of Γ . Hence, $Y \cup M$ is in Γ , and, consequently, $M \in \Gamma_{Y,X}$. It is left to show that M is a minterm of $\Gamma_{Y,X}$. Let M' be a proper subset of M . Since $M \cup \{u_0\}$ is a circuit of $\mathcal{M}'$, the set $M' \cup \{u_0\}$ is independent in $\mathcal{M}'$. This implies that $Y \cup M' \cup \{u_0\}$ is independent in $\mathcal{M}$. Therefore $Y \cup M' \notin \Gamma$, and thus $M' \notin \Gamma_{Y,X}$. This completes the proof of the first direction.

For the second direction, assume that M is a minterm of $\Gamma_{Y,X}$. We next prove that $M \cup \{u_0\}$ is dependent in $\mathcal{M}'$. The set $Y \cup M$ is authorized in Γ . Let $Y' \subseteq Y$ be such that $Y' \cup M \in \Gamma$ and $|Y'|$ is minimal (it is possible that $Y' = \emptyset$). We claim that $Y' \cup M$ is a minterm of Γ : Indeed, every user in Y' is critical in $Y' \cup M$, as implied by the minimality of $|Y'|$; on the other hand, every user in M is critical, for otherwise

M would have a proper subset $M' \subset M$ such that $Y \cup M' \in \Gamma$, thus contradicting our assumption that M is a minterm of $\Gamma_{Y,X}$. It follows that $Y' \cup M \cup \{u_0\}$ is a circuit of $\mathcal{M}$. Therefore, the set $Y \cup M \cup \{u_0\}$ is dependent in $\mathcal{M}$. This implies that $M \cup \{u_0\}$ is dependent in $\mathcal{M}'$.

To complete the proof we need to prove that $M \cup \{u_0\}$ is a circuit in $\mathcal{M}'$. Assume towards a contradiction that $M \cup \{u_0\}$ properly contains a circuit of $\mathcal{M}'$. If this circuit is of the form $M' \cup \{u_0\}$, for some $M' \subset M$, then by the proof of the first direction, M' is a minterm of $\Gamma_{Y,X}$, contradicting the fact that M is a minterm of $\Gamma_{Y,X}$. Therefore, the circuit is some $M' \subseteq M$. This implies that $Y \cup M'$ is dependent in $\mathcal{M}$. Hence, there exists $Y'' \subseteq Y$ such that $Y'' \cup M'$ is a circuit of $\mathcal{M}$. However, as shown earlier, $Y' \cup M \cup \{u_0\}$ is a circuit of $\mathcal{M}$ as well, for some $Y' \subseteq Y$. Applying Lemma 4.6 for these two circuits, with e being an arbitrary user in M' , and $d = u_0$, we get that there is a circuit contained in $(M \setminus \{e\}) \cup Y \cup \{u_0\}$ that contains u_0 . This implies that $(M \setminus \{e\}) \cup Y \in \Gamma$, whence $(M \setminus \{e\}) \in \Gamma_{Y,X}$, in contradiction to our assumption that M is a minterm of $\Gamma_{Y,X}$. $\square$

A special family of restrictions of Γ are restrictions in which $Y = \emptyset$, i.e., $\Gamma_{\emptyset,X}$. Such a restriction consists of all subsets of X that are authorized in Γ . We shall refer to such a restriction simply as *the restriction of Γ to X* . We conclude with a trivial claim regarding restrictions of WTASs.

CLAIM 4.13. *Let $X, Y \in U$ such that $X \cap Y = \emptyset$. If Γ is a WTAS, then $\Gamma_{Y,X}$ is a WTAS.*

5. WTASs and matroids. In this section we prove several properties of matroids that are associated with ideal WTASs. These properties will serve us later in characterizing ideal WTASs. Let Γ be an ideal WTAS on $U = \{u_1, \dots, u_n\}$ corresponding to a weight function $w : U \rightarrow \mathbb{N}$ and a threshold T . Let $\mathcal{M}$ be the matroid corresponding to Γ .

An authorized set of Γ may contain many different minterms. The following simple lemma shows that one of these minterms is a suffix of the set.

LEMMA 5.1. *Any $X \in \Gamma$ contains a suffix minterm. Namely, if $X = \{x_1, \dots, x_k\}$ and $X \in \Gamma$, then there exists $i \in [k]$ such that $X_{i,k} = \{x_i, \dots, x_k\}$ is a minterm.*

Proof. Let $s \in [k]$ be the largest index such that $X_{s,k}$ is authorized. We claim that $X_{s,k}$ is a minterm. Indeed, x_s is critical for $X_{s,k}$ since, removing it from $X_{s,k}$, we get $X_{s+1,k}$, which is unauthorized by the choice of s . Consequently, every user $x_i \in X_{s,k}$ is critical because x_s has the smallest weight in $X_{s,k}$. Therefore, $X_{s,k}$ is a minterm. $\square$

We saw that minterms of Γ are independent sets of $\mathcal{M}$. Different minterms of Γ may be of different sizes. By Claim 4.7, if M and M' are minterms and $|M| < |M'|$, we can add (at least $|M'| - |M|$) users to M such that the resulting set is still an independent set of $\mathcal{M}$. The next lemma asserts that the users that may be added to M must have a smaller weight than the weight of every member of M .

LEMMA 5.2. *Let M be a minterm of Γ . Let $y \in U \setminus M$ be a user such that $w(M_{\min}) \leq w(y)$. Then $M \cup \{y\}$ is a dependent set of $\mathcal{M}$.*

Proof. Let $X = (M \setminus \{M_{\min}\}) \cup \{y\}$ be the set that is obtained by replacing the minimal element of M with y . Since $w(X) \geq w(M)$, the set X is authorized, and thus it contains a minterm M' . Moreover, $M \neq M'$ since $M_{\min} \in M \setminus M'$. Therefore, by Corollary 4.5, the set $M \cup M' = M \cup \{y\}$ is dependent in $\mathcal{M}$. $\square$

While Γ may have minterms of different cardinalities, we show in the next lemma that whenever two minterms have the same minimal member they must be of the same size.

LEMMA 5.3. *Let X and Y be minterms of the access structure Γ such that $X_{\min} = Y_{\min}$. Then $|X| = |Y|$.*

Proof. As X and Y are minterms, they are independent sets of the matroid $\mathcal{M}$. Assume that $|X| < |Y|$. Then, by axiom (I3) of the matroid definition, there exists $y \in Y \setminus X$ such that the set $X \cup \{y\}$ is independent. However, as $X_{\min} = Y_{\min}$ is a user with the minimal weight in both X and Y , we have that $w(X_{\min}) \leq w(y)$. Consequently, in view of Lemma 5.2, the set $X \cup \{y\}$ is dependent. This contradiction implies that $|X| \geq |Y|$. Arguing along the same lines, we conclude that also $|X| \leq |Y|$, whence both minterms are of the same size. $\square$

In the matroid $\mathcal{M}$ corresponding to an access structure Γ , each minterm M of Γ defined a circuit $M \cup \{u_0\}$ in $\mathcal{M}$. This may be viewed as a *local* relation between the minterms of Γ and the dependent sets of $\mathcal{M}$. However, when dealing with WTASs, the *global* structure of Γ enables us to prove the dependency of various sets in $\mathcal{M}$. It turns out that the lexicographic order on the minterms of Γ , with respect to the relation $\prec$ (defined in section 2.3), is strongly related to dependence in $\mathcal{M}$. This is demonstrated through the following definition and lemmas.

DEFINITION 5.4 (canonical complement). *Let P be a prefix of some minterm of Γ . Let $Y \subseteq U$ be the lexicographically minimal set such that*

- $P_{\max} \prec Y_{\min}$, and
- the set $P \cup Y$ is a minterm of Γ .

Then the set Y is called the canonical complement of P .

The following lemma shows that replacing the canonical complement by a user that precedes the first element of the canonical complement results in a dependent set.

LEMMA 5.5. *Let P be a prefix of some minterm of Γ . Let $Y = \{y_1, \dots, y_t\}$ be the canonical complement of P , and let b be a user such that $P_{\max} \prec b \prec y_1$. Then $P \cup \{b\}$ is dependent. Furthermore, the set $P \cup \{b\}$ includes a unique circuit, and that circuit contains b .*

Proof. If $P = \emptyset$, then, since Γ is connected, there exists a minterm that starts with u_1 , whence $y_1 = u_1$. Therefore, it cannot be that $b \prec y_1$, and thus the claim is trivially true. Otherwise, if $P \neq \emptyset$, denote by M_1 the minterm $M_1 = P \cup Y$. Let $X_2 = (M_1 \setminus \{P_{\max}\}) \cup \{b\}$ be the set resulting from replacing $P_{\max}$ with b in M_1 . Since $w(P_{\max}) \leq w(b)$, the set X_2 is authorized (though not necessarily a minterm). Let M_2 be the suffix minterm contained in X_2 (such a minterm exists in view of Lemma 5.1). It must be that $b \in M_2$, since otherwise $M_2 \subseteq Y$, where Y —the canonical complement of P —is a proper subset of a minterm and thus is unauthorized.

Let $A = M_1 \cup M_2 = P \cup \{b\} \cup Y$. We proceed to show that every user in Y is critical for A . This will show that $D(M_1, M_2) \subseteq (M_1 \cup M_2) \setminus Y = P \cup \{b\}$. But since, by Corollary 4.5, the set $D(M_1, M_2)$ is dependent, this will imply that also $P \cup \{b\}$ is dependent. We also observe that it is sufficient to show that $Y_{\min} = y_1$ is critical for A ; this will imply that also all other members of Y , having weight that is no smaller than $w(y_1)$, are also critical for A .

In view of the above, it suffices to show that y_1 is critical for A . Suppose this is not the case; namely, the set $A \setminus \{y_1\}$ is authorized. Since $A \setminus \{y_1\}$ results from M_1 by replacing y_1 by b , where $w(b) \leq w(y_1)$, and since M_1 is a minterm, it must be that $A \setminus \{y_1\}$ is also a minterm. But this is a contradiction to the choice of y_1 as the first element in the canonical complement of P . Hence, all the elements of Y are critical for A , and consequently $P \cup \{b\}$ is dependent.

Since P is part of a minterm, it must be that P is independent. Thus, by

Claim 4.8, the set $P \cup \{b\}$ must contain a unique circuit, and that circuit contains b . $\square$

The next lemma is a generalization of Lemma 5.5. Given a prefix P of some minterm, it gives a sufficient condition for a set $B \subseteq U$ so that $P \cup B$ is dependent in $\mathcal{M}$. It plays an essential role in our study of the characterization of ideal WTASs.

LEMMA 5.6. *Let P be a prefix of some minterm of Γ . Let $Y = \{y_1, \dots, y_t\}$ be the canonical complement of P , and $B = \{b_1, \dots, b_j\}$ be a set such that $P_{\max} \prec B_{\min}$ and $b_j \prec y_j$. Then, the set $P \cup B$ is dependent.*

Proof. The proof is by induction on j . The case $j = 1$ is handled by Lemma 5.5. We assume that the lemma holds for all $i \in [j - 1]$ and proceed to prove it for $i = j$. If $b_i \prec y_i$ for some $i \in [j - 1]$, then the statement of the lemma holds by induction (in fact, in that case we conclude that $P \cup B_{1,i}$ is dependent, which is a stronger claim than what we need to show). Therefore, we may assume hereinafter that $y_i = b_i$ or $y_i \prec b_i$ for all $i \in [j - 1]$.

Assume towards a contradiction that $P \cup B$ is independent. Denote $D = P \cup \{y_1\} \cup B$ and $P' = P \cup \{y_1\}$. The canonical complement of P' is $Y_{2,t}$. On the other hand, since $P'_{\max} = y_1 \preceq b_1 \prec b_2$, and $b_j \prec y_j$, the induction hypothesis, applied to P' and $B_{2,j}$, implies that $P' \cup B_{2,j}$ is dependent. Since $P' \cup B_{2,j} \subseteq D$, the set D is also dependent. Therefore, since we assumed that $P \cup B$ is independent, Claim 4.8 implies that D contains a unique circuit C . The uniqueness of the circuit C in D and the fact that $P' \cup B_{2,j}$ is a dependent subset of D imply that $C \subseteq P' \cup B_{2,j}$. This, in turn, implies that $b_1 \notin C$, unless $y_1 = b_1$. This will allow us to derive a contradiction.

There are three possible cases regarding the position of b_1 with respect to the members of Y : $y_1 = b_1$, $y_1 \prec b_1 \prec y_2$, or $y_2 \preceq b_1$. The first case is trivial: If $y_1 = b_1$, then $P \cup B = P' \cup B_{2,j}$, and thus it is dependent, in contradiction to our assumption.

If $y_1 \prec b_1 \prec y_2$, then by Lemma 5.5, applied to P' and b_1 , the set $P' \cup \{b_1\}$ is dependent and contains a circuit C' that contains b_1 . As $b_1 \in C'$ and $b_1 \notin C$, we conclude that $C \neq C'$. However, as $C' \subseteq D$, this contradicts the uniqueness of the circuit C in D .

The only case left is $y_2 \preceq b_1$. As P' is independent, there must be an index $s \in \{2, \dots, j\}$ such that $b_s \in C$ (for, otherwise, as $b_1 \notin C$, $C \cap B = \emptyset$, and, consequently, $C \subseteq P'$). Consider the set $B^s = B \setminus \{b_s\}$. Since $|B^s| = j - 1$, $P'_{\max} = y_1 \prec b_1 = B_{\min}^s$, and the $(j - 1)$ th element of B^s (which is either b_j , if $s < j$, or b_{j-1}) precedes the $(j - 1)$ th element of the canonical complement of P' (which is y_j), we conclude, by the induction hypothesis, that $P' \cup B^s$ is dependent. Therefore, it contains a circuit C'' that does not contain b_s . Since $b_s \in C$ but $b_s \notin C''$, we infer that $C \neq C''$. However, as $C'' \subseteq D$, this contradicts the uniqueness of C in D . This completes the proof that the set $P \cup B$ is dependent. $\square$

6. WTASs and HTASs. In this section we discuss the family of HTASs, from Definition 3.1, and their relation to WTASs. In section 6.1 we prove certain properties of the matroids associated with such access structures. In section 6.2 we characterize the intersection between HTASs and WTASs. Finally, in section 6.3, we show that if an ideal WTAS has a minterm in the form of a prefix of U , then it is in fact an HTAS.

When discussing an HTAS on some set $U = \{u_1, \dots, u_n\}$, we shall assume that the users in U are ordered according to their position in the hierarchy, from the lowest level to the highest. Namely, that

$$(6.1) \quad L_i = U_{\ell_i, \ell_{i+1}-1} = \{u_{\ell_i}, \dots, u_{\ell_{i+1}-1}\} \quad \forall i \in [m]$$

for some sequence $\ell_1 = 1 < \ell_2 < \dots < \ell_m < \ell_{m+1} = n + 1$. Given a nonempty subset $A \subseteq U$, if $A_{\min} \in L_i$, then A is said to be of level i and is denoted by $L(A) = i$.

6.1. The matroid associated with HTASs. Since any HTAS Γ_H is ideal [6, 35], Theorem 4.1 implies that there exists a matroid $\mathcal{M}$ that is associated with it. We derive herein some properties of such matroids that will be used later in our study of the structure of matroids corresponding to ideal WTASs.

CLAIM 6.1. *Let Γ_H and $\mathcal{M}$ be an HTAS and its associated matroid. Then every circuit of $\mathcal{M}$ not containing u_0 is a union of two distinct minterms.*

Proof. By Corollary 4.4, the circuits of the matroid $\mathcal{M}$ that do not contain u_0 are the union of two minterms of Γ_H , excluding users that are critical to that union. All that needs to be shown is that if M_1, M_2 are two minterms in Γ_H , no user in their union, $u \in M_1 \cup M_2$, is critical. Assume, without loss of generality, that $L(M_1) = i \leq L(M_2)$. Then, as M_1 is a minterm, we conclude that it satisfies the i th threshold condition with an equality; namely, M_1 has k_i users from $\bigcup_{j=i}^m L_j$. Hence, since $M_2 \subseteq \bigcup_{j=i}^m L_j$ and $M_2 \neq \emptyset$, the union $M_1 \cup M_2$ has at least $k_i + 1$ users from $\bigcup_{j=i}^m L_j$. Therefore, we can remove any user from $M_1 \cup M_2$ and still have an authorized set. Hence, no user $u \in M_1 \cup M_2$ is critical for the union. $\square$

CLAIM 6.2. *Let Γ_H and $\mathcal{M}$ be an HTAS and its associated matroid. Then U_{1,k_1+1} is a circuit of $\mathcal{M}$.*

Proof. Consider the two sets U_{1,k_1} and U_{2,k_1+1} . The first one is a minterm of Γ_H . Regarding the second, there are two options: either U_{2,k_1+1} is a minterm (satisfying the k_1 -threshold condition), or it satisfies already the k_2 -threshold condition (this may happen only if $|L_1| = \ell_2 - 1 = k_1 - k_2 + 1$, which is the minimal possible value for the size of that level). In both cases, U_{2,k_1+1} contains a minterm that includes u_{k_1+1} . Hence, U_{1,k_1+1} is a union of two distinct minterms, and therefore, by Corollary 4.5, it is dependent. We claim that it is actually a circuit. If not, then it must contain a circuit C of size smaller than $k_1 + 1$. By Claim 6.1, this circuit is a union of two minterms. Since its size is smaller than $k_1 + 1$, it must be a union of two minterms from levels L_2 and above. However, by Remark 3.2, $|L_1| \geq k_1 - k_2 + 1$, or, equivalently, $|U_{1,k_1+1}| - |L_1| \leq k_2$. This implies that the set U_{1,k_1+1} contains at most k_2 users from levels L_2 and above. Since any minterm from levels L_2 and above must be of size at least k_2 , we infer that U_{1,k_1+1} can contain at most one minterm from those levels. This contradicts our earlier conclusion that it is a union of two minterms from those levels. We therefore conclude that C must be a circuit of $\mathcal{M}$. $\square$

6.2. The intersection of WTASs and HTASs. Let U be a set of users and let Γ be a monotone access structure on U that is both a WTAS and an HTAS. Namely, on one hand, there exist a weight function $w : U \rightarrow \mathbb{N}$ and a threshold $T \in \mathbb{N}$ such that Γ is the corresponding WTAS on U and, on the other hand, there exist a hierarchy in U , where $U = \bigcup_{i=1}^m L_i$, and thresholds $k_1 > k_2 > \dots > k_m$ such that Γ is also the corresponding HTAS on U . Our goal herein is to characterize such access structures that belong to those two classes of access structures.

It should be noted that in the following discussion we cannot assume that the levels of U in the HTAS are nicely organized in a monotonic order as in (6.1).

We proceed with some notation. Let A_i , where $i \in [m - 1]$, denote a subset of $k_i - k_{i+1}$ users of largest weights in L_i , and let A_m be a subset of $k_m - 1$ users of largest weights in L_m . We also let a_i denote a user of minimal weight in L_i , for all $i \in [m]$. Note that, by Remark 3.2, we have that $a_i \in L_i \setminus A_i$ for every $i \in [m]$.

We set $A = \bigcup_{i=1}^m A_i$ and $B = A_m \cup \{a_m\}$. The set A is unauthorized since it fails to satisfy each of the threshold conditions by exactly one user, while B is authorized since it is composed of k_m users of level L_m . Therefore $w(A) < T$ and $w(B) \geq T$. But as $B = A \cup \{a_m\} \setminus (A \setminus A_m)$, we infer that

$$(6.2) \quad w(A \setminus A_m) < w(a_m).$$

We refer to the last level L_m in an HTAS as trivial if $k_m = 1$ (namely, all users of such a level are self-sufficient). We proceed to characterize access structures that are both WTAS and HTAS.

LEMMA 6.3. *If the last level is not trivial, then the number of levels m in the HTAS is at most 2.*

Proof. Assume that $m \geq 3$ and that the last level is not trivial, i.e., $k_m \geq 2$. Let Z_m be a subset of $k_m - 2 \geq 0$ users of largest weight in L_m . Define $Z_{m-1} = A_{m-1} \cup \{a_{m-1}\}$ and $Z_{m-2} = A_{m-2} \cup \{a_{m-2}\}$. Their union $Z = Z_{m-2} \cup Z_{m-1} \cup Z_m$ is authorized since its cardinality is $|Z| = |A_{m-2}| + |A_{m-1}| + k_m = k_{m-2}$, and consequently, it satisfies the $(m-2)$ th threshold condition.

Next, consider the set $B = Z \cup \{a_m\} \setminus (A_{m-1} \cup A_{m-2}) = Z_m \cup \{a_{m-2}, a_{m-1}, a_m\}$. Since, by (6.2), $w(A_{m-1} \cup A_{m-2}) < w(a_m)$, the definition of B implies that $w(B) > w(Z)$. Therefore, as Z is authorized, so is B . However, while B is authorized in the WTAS, it is unauthorized in the HTAS, as implied by the following inequalities: $|B \cap L_m| = |Z_m \cup \{a_m\}| = k_m - 1 < k_m$, $|B \cap (L_{m-1} \cup L_m)| = k_m < k_{m-1}$, and $|B| = k_m + 1 < k_{m-2}$. This contradiction implies that either $m \leq 2$ or the last level is trivial. $\square$

According to Lemma 6.3, access structures that are both WTAS and HTAS and have no self-sufficient users have a very simple hierarchical structure: Either there is one level (in which case the access structure is a simple threshold access structure) or two levels. In the latter case, we show below that if the difference between k_1 and k_2 is larger than 1, the size of the first level is minimal (see Remark 3.2).

LEMMA 6.4. *If the last level is not trivial, $m = 2$, and $k_1 - k_2 > 1$, then $|L_1| = k_1 - k_2 + 1$.*

Proof. Assume towards a contradiction that $|L_1| > k_1 - k_2 + 1$. Let Z_2 be a subset with $k_2 - 2$ users of largest weight in L_2 and $Z_1 = A_1 \cup \{b_1, b_2\}$, where b_1 and b_2 are two users in $L_1 \setminus A_1$ (note that our assumptions imply that Z_1 and Z_2 are well defined). Their union $Z = Z_1 \cup Z_2$ is authorized since $|Z| = k_1$. Next, consider the set $B = Z \cup \{a_2\} \setminus A_1 = Z_2 \cup \{a_2, b_1, b_2\}$. Since, by (6.2), $w(A_1) < w(a_2)$, we conclude that B is also authorized. Hence, it must be also authorized in the HTAS. However, as $|B \cap L_2| = k_2 - 1$ and $|B| = k_2 + 1 < k_1$, the set B is unauthorized in the HTAS. $\square$

Lemmas 6.3 and 6.4 imply that if Γ is an access structure on U that is both a WTAS and an HTAS, and if U has no self-sufficient users (i.e., the last level in the HTAS is not trivial), then either $m = 1$ or $m = 2$, where in the latter case either $k_1 = k_2 + 1$ or $|L_1| = k_1 - k_2 + 1$. If, on the other hand, the access structure has self-sufficient users, i.e., the last level in the HTAS is trivial, then its restriction to the first $m-1$ levels is still a WTAS as well as an HTAS, and, in addition, that restriction has no self-sufficient users. This implies that the structure of that restriction is as described above. This is summarized as follows.

CLAIM 6.5. *Let Γ be both a WTAS and an HTAS. Then the HTAS-parameters of Γ satisfy one of the following conditions:*

1. $m = 1$.

2. $m = 2$ and $k_1 = k_2 + 1$.
3. $m = 2$ and $|L_1| = k_1 - k_2 + 1$.
4. $m \in \{2, 3\}$, the level L_m is trivial, and the restriction of Γ_H to the first $m - 1$ levels is of the form that is described in cases 1–3.

To complete the characterization we proceed to prove the converse. Namely, that any HTAS with parameters as described in Claim 6.5 is also a WTAS. This is shown by constructing the appropriate weight function and threshold in each case.

CLAIM 6.6. *Let Γ_H be an HTAS that satisfies one of the conditions in Claim 6.5. Then there exist a weight function and a threshold such that Γ_H coincides with the corresponding WTAS.*

Proof. Case 1 is a simple threshold, so we can assign each user the weight 1 and set the threshold to k_1 . In case 2, the weight assigned to every user in L_1 is k_2 , the weight assigned to users in L_2 is k_1 , and the threshold is $k_1 \cdot k_2$. In case 3, the weight assigned to users in L_1 and L_2 is 1 and $k_1 - k_2 + 1$, respectively, while the threshold is $k_2(k_1 - k_2 + 1)$. In case 4, we assign weights to the users in the first $m - 1$ levels and set an appropriate threshold according to the condition that is satisfied by the restriction of Γ_H to the first $m - 1$ levels and the rules that were described above for the first three cases; in addition, we assign to every user in the trivial level a weight that equals the threshold. In all four cases it is a straightforward matter to verify that the induced WTAS coincides with the original HTAS. $\square$

6.3. Ideal WTASs with a prefix minterm are HTASs. In this section we make the first step towards proving our main result, Theorem 3.5. Let Γ be an ideal WTAS on a set U of n users, corresponding to a weight function $w : U \rightarrow \mathbb{N}$ and a threshold T . Assume that U possesses a prefix minterm $U_{1,k}$ for some $k \in [n]$ (namely, there exists $k \in [n]$ such that the k users of smallest weights form a minterm). We claim that Γ is an HTAS. We first describe the partition of U into levels and determine the corresponding thresholds. Denoting the resulting HTAS by Γ_H , we proceed to prove that $\Gamma = \Gamma_H$.

The decomposition of U to levels will respect the order of users according to their weights. Namely, each level will be a run of U , and our goal is to determine the transition points between one level and the subsequent one. Since $U_{1,k}$ is a minterm, $U_{1,i}$ is authorized for every $i \in \{k, \dots, n\}$. By Lemma 5.1, for every such i there exists a run-minterm ending at u_i . Let us denote the length of that run-minterm by μ_i . By the nondecreasing monotonicity of the weights, we infer that the sequence of lengths $\boldsymbol{\mu} = (\mu_i)_{k \leq i \leq n}$ is monotonically nonincreasing. Denote by m the number of distinct values assumed by the sequence $\boldsymbol{\mu}$, and let us denote those values by $k_1 > \dots > k_m$. Then the HTAS Γ_H is defined as follows: m is the number of levels, and k_i is the i th threshold. As for the levels, we denote by ℓ_i , where $i \in [m]$, the index of the first user in the first run-minterm of length k_i (e.g., $\ell_1 = 1$ since $U_{1,k}$ is the first run-minterm of length $k = k_1$ and its first user is u_1); then the i th level in the hierarchy is $L_i = U_{\ell_i, \ell_{i+1}-1}$, where $\ell_{m+1} = n + 1$.

In addition, we denote by U_{s_i, t_i} the right-most run-minterm whose length is k_i , where $i \in [m]$, and consider the set $A_i = U_{s_i+1, \ell_{i+1}-1}$. As U_{s_i, t_i} is the last minterm that contains k_i users and U_{ℓ_{i+1}, t_i+1} is the first minterm that contains k_{i+1} users, the set A_i consists of the last $k_i - k_{i+1}$ users in L_i , where $k_{m+1} = 1$. (Note that if u is the last element of the *last* k_i -long minterm and u' is the last element in the *first* k_{i+1} -long minterm, then u and u' are consecutive elements.) An important observation regarding those subsets is that, given $i \in [m]$ and $u_h \in A_i$, there is no run-minterm of the WTAS Γ that starts with u_h ; indeed, if $U_{h,j}$ were a run-minterm, then it would

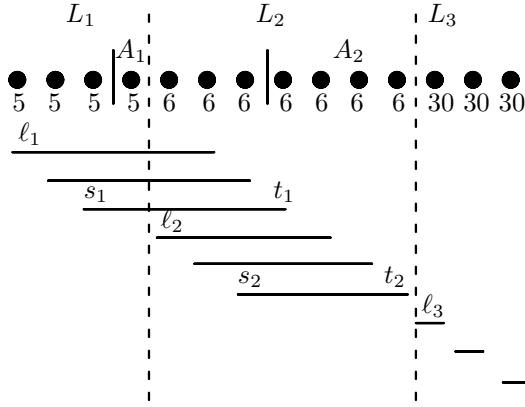


FIG. 6.1. A WTAS that is also an HTAS.

be a proper subset of the minterm U_{s_i, t_i} if $j \leq t_i$, or a proper superset of the minterm U_{ℓ_{i+1}, t_i+1} if $j \geq t_i + 1$.

An illustration of the construction of the levels of the HTAS appears in Figure 6.1. The example in the illustration is of a WTAS with 14 users of weights 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 30, 30, 30 and threshold $T = 30$. The vertical dashed lines indicate the three levels in the corresponding HTAS (the third one being a trivial one), and the horizontal lines indicate all of the run-minterms in that access structure.

Next, we prove that the WTAS Γ coincides with the HTAS Γ_H described above.

THEOREM 6.7. *Let Γ be an ideal WTAS on U that has a prefix minterm. Then Γ is an HTAS.*

Proof. Let Γ_H be the HTAS as described above. We will prove that $\Gamma = \Gamma_H$, thus showing that Γ is an HTAS. We start by proving that $\Gamma_H \subseteq \Gamma$. Let $X \in \Gamma_H$. Then for some $i \in [m]$ the set X has at least k_i users from $\bigcup_{j=i}^m L_j$. Letting $B_i = U_{\ell_i, \ell_i + k_i - 1}$ denote the set of the first k_i users from $\bigcup_{j=i}^m L_j$, the nondecreasing monotonicity of the weights implies that $w(X) \geq w(B_i)$. By the construction of levels in Γ_H , the set B_i is a minterm of Γ , whence $w(B_i) \geq T$. Therefore, $w(X) \geq T$ and, consequently, $X \in \Gamma$.

Conversely, assume that $X \notin \Gamma_H$. Then X has at most $k_i - 1$ users from $\bigcup_{j=i}^m L_j$, for every $i \in [m]$. Consider the set $A = \bigcup_{i=1}^m A_i$. By the definition of A , it has exactly $k_i - 1$ users from $\bigcup_{j=i}^m L_j$, for every $i \in [m]$. Moreover, A is the set with the maximal weight among the sets that are unauthorized in the HTAS, and thus $w(X) \leq w(A)$. Therefore, it suffices to show that $A \notin \Gamma$ in order to conclude that $X \notin \Gamma$ and, thus, complete the proof.

To this end, assume that $A \in \Gamma$. Then A contains some minterm $M \in \Gamma$. Assume that M is of level i , $L(M) = i$; namely, i is the lowest level for which $M \cap L_i \neq \emptyset$. Then $M \cap A_i$ is a prefix of M . In order to arrive at a contradiction, we proceed to show that there can be no minterm of Γ that has a prefix which is a subset of A_i for some $i \in [m]$.

Assume, by contradiction, that there are such minterms, and let M' be the lexicographically minimal minterm of Γ of that sort. Let $u_h = M'_{\min}$, and let j be the maximal index such that $M' = U_{h,j} \cup Z$ for some $Z \subset U$. Since $u_h \in A_i$ for some $i \in [m]$, and we observed earlier that no run-minterm of Γ starts in A_i , we conclude that $Z \neq \emptyset$ (and $u_{j+1} \prec Z_{\min}$ because of the maximality of j). We claim that

$j < t_i$. Indeed, if $j \geq t_i$, then M' is a proper superset of $\hat{M} = U_{\ell_{i+1}, t_i} \cup \{Z_{\min}\}$. As $w(\hat{M}) \geq w(U_{\ell_{i+1}, t_i+1}) \geq T$, we conclude that $\hat{M} \in \Gamma$. But then we get a contradiction since a minterm M' cannot be a proper superset of an authorized set $\hat{M}$.

Next, define $Q = M' \cup \{u_{j+1}\} \setminus \{u_j\}$. The set Q is authorized in Γ , and, by Lemma 5.1, it contains a suffix minterm M'' that must contain u_{j+1} , for otherwise it would be a proper subset of M' . Therefore, $M' \cup M'' = M' \cup \{u_{j+1}\} = U_{h,j} \cup \{u_{j+1}\} \cup Z$. We claim that all members of Z are critical for this union. Assume, by contradiction, that $M^* = M' \cup \{u_{j+1}\} \setminus \{z\}$ is authorized, for some $z \in Z$. Since $w(u_{j+1}) \leq w(z)$ and M' was a minterm of Γ , also M^* is a minterm of Γ . But M^* is a minterm of Γ that starts within A_i and $M^* \prec M'$, thus contradicting our choice of M' . Hence, by Corollary 4.5, the set $U_{h,j+1}$ is dependent in $\mathcal{M}$. However, since $s_i < h$ and $j+1 \leq t_i$, this dependent set is properly contained in the minterm U_{s_i, t_i} , leading to a contradiction. Hence, $A \notin \Gamma$, and the proof is thus complete. $\square$

7. Ideal WTASs and TPASs. In the previous section we dealt with the case where the lexicographically minimal minterm of Γ is a prefix; of U . Here, we handle the case where the lexicographically minimal minterm of Γ is a lacunary prefix; namely, it takes the form $M = U_{1,d} \cup U_{d+2,k}$ for some $1 \leq d \leq k-2$ and $k \leq n$. Throughout this section we assume that there is at least one minterm starting with the user u_2 , and that there are no self-sufficient users. If this is not the case, then Γ is a simple composition of access structures, as shown in section 8.2. The section is organized as follows. First we show that $U_{2,k}$ must be a minterm of Γ . Consequently, we show that Γ is one of the two TPASs that were defined in Definition 3.3. Finally, in section 7.1, we design linear ideal secret sharing schemes that realize TPASs.

Consider the set $U_{2,k}$. As $U_{2,k} = M \cup \{u_{d+1}\} \setminus \{u_1\}$ and $w(u_{d+1}) \geq w(u_1)$, the set $U_{2,k}$ is authorized. We proceed to show that it is in fact a minterm of Γ .

CLAIM 7.1. *Let Γ be an ideal WTAS with $M = U_{1,d} \cup U_{d+2,k}$ being its lexicographical minimal minterm for some $1 \leq d \leq k-2$ and $k \leq n$. Then if there exists a minterm that has u_2 as its minimal user, $U_{2,k}$ is a minterm of Γ .*

Proof. Let M' be the lexicographically minimal minterm starting with u_2 . Assume towards a contradiction that the authorized set $U_{2,k}$ is not a minterm. The minterm M' cannot be a run-minterm since $U_{2,k-1}$ is unauthorized (as its weight does not exceed that of $M \setminus \{u_1\}$, which is a proper subset of the minterm M) and $U_{2,k}$ is already an authorized set (since $w(U_{2,k}) \geq w(M)$) that is not a minterm. Let s be the maximal index such that $U_{2,s} \subset M'$. In that case, as M' is not a run-minterm, $M' = U_{2,s} \cup Z$, where $Z \neq \emptyset$ and $u_{s+1} \prec Z_{\min}$. We claim that $s \leq k-2$. Assume, towards a contradiction, that $s \geq k-1$. Then $w(M') \geq w(U_{2,k-1}) + w(Z) \geq w(U_{2,k})$. However, as explained earlier, $U_{2,k}$ is an authorized subset that is not a minterm. Hence, if we remove from $U_{2,k}$ its minimal user— u_2 —we are left with a set that is still authorized. But M' weighs more than $U_{2,k}$ and it also contains u_2 , whence also $M' \setminus \{u_2\}$ is authorized. But this contradicts the fact that M' is a minterm. We conclude that M' —the lexicographically minimal minterm starting with u_2 —takes the form $M' = U_{2,s} \cup Z$, where $s \leq k-2$ and Z is a nonempty set with $Z_{\min} \succ u_{s+1}$.

As Z is the canonical complement of $U_{2,s}$, Lemma 5.5 implies that the set $U_{2,s+1}$ is dependent. Hence, it contains a circuit C_1 of the corresponding matroid $\mathcal{M}$. The circuit C_1 must include the user u_{d+1} (the user that is missing in the minterm M), since, otherwise, as $s+1 \leq k-1$, the circuit C_1 would be a subset of the minterm $M = U_{1,d} \cup U_{d+2,k}$. We claim, and prove in the next paragraph, that $U_{1,d+1}$ is also a circuit of $\mathcal{M}$. Since $u_1 \in U_{1,d+1}$ but $u_1 \notin C_1$, the sets C_1 and $U_{1,d+1}$ are two distinct circuits. In addition, both circuits include u_{d+1} . Therefore, by Lemma 4.6,

the set $C_1 \cup U_{1,d+1} \setminus \{u_{d+1}\}$ is dependent. However, this set is contained in the minterm $M = U_{1,d} \cup U_{d+2,k}$, which is an independent set of the matroid $\mathcal{M}$. This contradiction implies that $U_{2,k}$ is a minterm in Γ .

Hence, we need only to show that $U_{1,d+1}$ is a circuit of $\mathcal{M}$. Let $Y = U_{d+2,k}$, $X = U_{1,d+1}$ and consider $\Gamma_{Y,X}$, namely, the restriction of Γ induced by Y on X . By Lemma 4.11, $\Gamma_{Y,X}$ is an ideal WTAS, and $U_{1,d}$ is a prefix minterm in that access structure. Therefore, by Theorem 6.7, the access structure $\Gamma_{Y,X}$ is an HTAS, and, by Claim 6.2, the set $U_{1,d+1}$ is a circuit in the matroid corresponding to $\Gamma_{Y,X}$. By Lemma 5.5, applied with $P = U_{1,d}$, the set $U_{d+2,k}$ as the canonical complement of P , and $b = u_{d+1}$, the set $P \cup b = U_{1,d+1}$ is dependent in $\mathcal{M}$. Assume, by contradiction that it is not a circuit. Therefore, it possesses a proper subset $C' \subset U_{1,d+1}$ that is a circuit. By Lemma 4.12, the set C' would have to be dependent in the matroid corresponding to $\Gamma_{Y,X}$. But as $U_{1,d+1}$ is a circuit in that matroid, this is impossible. Hence, $U_{1,d+1}$ is also a circuit in $\mathcal{M}$. This completes the proof. $\square$

In view of the above, as we assume herein that there are minterms that start with u_2 , we conclude that $U_{2,k}$ is a minterm. This is a prefix minterm in the restriction of Γ to $U_{2,n}$, denoted hereinafter Γ' . Since Γ' is ideal by Lemma 4.11, and a WTAS by Claim 4.13, and has a prefix minterm, Theorem 6.7 implies that it is also an HTAS. Since we assumed that Γ has no self-sufficient users, then by Claim 6.5, the HTAS Γ' has at most two levels.

We separate our discussion into two cases, according to the number of levels in Γ' . First, we consider the case where Γ' is a simple threshold access structure.

CLAIM 7.2. *If Γ' is an HTAS of one level (i.e., it is a threshold access structure), then a set $X \subseteq U$ is in Γ if and only if $|X| \geq k - 1$ and $|X \cap U_{d+2,n}| \geq k - 1 - d$.*

Proof. Let us begin with sets X that do not include the user u_1 , namely, $X \subseteq U_{2,n}$. For such sets, the second threshold condition ($|X \cap U_{d+2,n}| \geq k - 1 - d$) is implied by the first one ($|X| \geq k - 1$). Since we assumed that Γ' , which is the restriction of Γ to $U_{2,n}$, is a simple threshold access structure, and the value of the threshold equals $|U_{2,k}| = k - 1$, the statement is straightforward in this case.

Next, consider sets X such that $u_1 \in X$. If X satisfies the two conditions in the claim, then its weight is no less than the weight of $M = U_{1,d} \cup U_{d+2,k}$. As $w(M) \geq T$, we conclude that $X \in \Gamma$. Conversely, assume that $X \in \Gamma$ but fails to satisfy one of the two conditions in the claim. Without loss of generality, we may assume that X is a minterm of Γ . Thus, by Lemma 5.3, $|X| = |M| = k - 1$. Therefore, X may violate only the second condition in the claim, namely, $|X \cap U_{d+2,n}| < k - 1 - d$. But then $U_{1,d+1} \subset X$, and that cannot be since then $X \prec M = U_{1,d} \cup U_{d+2,k}$ and M was assumed to be the lexicographically minimal minterm in Γ . $\square$

The access structure that is described in Claim 7.2 is a TPAS of type Δ_1 (see Definition 3.3) with

$$A = U_{1,d+1}, \quad B = U_{d+2,n-1}, \quad C = U_n, \quad m = k - 1, \quad t = 2.$$

It should be noted that a set X belongs to Δ_1 if it satisfies at least one of two threshold conditions. The second condition is $|X \cap C| \geq t$, which, in our case, is an empty condition (as $|C| = 1 < t = 2$). Hence, that access structure is in fact *bipartite*; namely, the question of whether X is authorized or not is determined only by the two numbers $|X \cap A|$ and $|X \cap (B \cup C)|$.

We are left with the case where Γ' is an HTAS of exactly two levels. Assume that $X \in \Gamma'$ if it has at least k_1 users from $U_{2,n}$ or at least k_2 users from $U_{r+1,n}$ for some $3 \leq r \leq n - 2$. As $U_{2,k}$ is a minterm, $k_1 = k - 1$.

CLAIM 7.3. *Under the above assumptions, a set $X \subseteq U$ is in Γ if and only if $|X \cap U_{r+1,n}| \geq k_2$ or both $|X| \geq k_1$ and $|X \cap U_{d+2,n}| \geq k_1 - d$.*

Proof. Assume first that $X \subseteq U$ satisfies one of the above two threshold conditions. If $|X \cap U_{r+1,n}| \geq k_2$, then $X \in \Gamma'$ and thus $X \in \Gamma$. If X satisfies the second condition, then it also belongs to Γ since the set of minimal weight X that satisfies both $|X| \geq k_1$ and $|X \cap U_{d+2,n}| \geq k_1 - d$ is $X = M = U_{1,d} \cup U_{d+2,k}$, and $M \in \Gamma$.

Assume next that $X \in \Gamma$. Without loss of generality, we may assume that X is a minterm and prove that it satisfies one of the two threshold conditions in the claim. We separate the discussion into two cases. If $u_1 \notin X$, then $X \in \Gamma'$. Thus either $|X \cap U_{r+1,n}| \geq k_2$ or $|X| \geq k_1$. If $|X| \geq k_1$, then, as $u_1 \notin X$, it must be that $|X \cap U_{d+2,n}| \geq k_1 - d$. If, on the other hand, $u_1 \in X$, then by Lemma 5.3, we get that $|X| = k_1$. But then, since no minterm contains $U_{1,d+1}$, it must be that $|X \cap U_{d+2,n}| \geq k_1 - d$. $\square$

The access structure Γ , as appears in Claim 7.3, coincides with one of the two TPASs in Definition 3.3. In case $r \geq d+1$, we set $A = U_{1,d+1}$, $B = U_{d+2,r}$, $C = U_{r+1,n}$, and then Γ is of type Δ_1 ,

$$\Gamma = \{X \subseteq U : (|X| \geq k_1 \text{ and } |X \cap (B \cup C)| \geq k_1 - d) \text{ or } |X \cap C| \geq k_2\}.$$

If, on the other hand, $r < d+1$, we claim that Γ is of type Δ_2 . To that end, we set $A = U_{1,r}$, $B = U_{r+1,d+1}$, and $C = U_{d+2,n}$. By Claim 7.3, the structure of Γ is

$$\Gamma = \{X \subseteq U : (|X| \geq k_1 \text{ and } |X \cap C| \geq k_1 - d) \text{ or } |X \cap (B \cup C)| \geq k_2\}.$$

It remains to show that the size of the set B is bounded, in accord with Definition 3.3.

CLAIM 7.4. *If $r < d+1$, then $|B| < d + k_2 - k_1$.*

Proof. The access structure Γ' , the restriction of Γ to $U_{2,n}$, is an HTAS of two levels. A set X is in Γ' if it has at least k_1 users from $U_{2,n}$ or at least k_2 users from $U_{r+1,n}$. Therefore, $U_{2,k} = U_{2,k_1+1}$ is a minterm and so is $U_{r+1,r+k_2}$. Consequently, $r + k_2 > k_1 + 1$ (since otherwise the second minterm would have been a proper subset of the first one). As $|B| = |U_{r+1,d+1}| = d + 1 - r$, we infer that $|B| < d + k_2 - k_1$. $\square$

Example 7.5. Consider the set $U = \{u_1, \dots, u_9\}$, and let Γ be a WTAS where the weights are 16, 16, 17, 18, 19, 24, 24, 24, and 24 and the threshold is 92. First note that there is no prefix minterm as $w(U_{1,5}) = 16 + 16 + 17 + 18 + 19 = 86 < 92$ and $w(U_{2,6}) = 16 + 17 + 18 + 19 + 24 = 94 \geq 92$, and thus $U_{1,6}$ is not a minterm. The lexicographically minimal minterm is $U_{1,3} \cup U_{5,6}$, and so $d = 3$ and $k = 6$. The HTAS structure resulting from the restriction of Γ to $U_{2,9}$ is composed of two levels $U_{2,5}$ and $U_{6,9}$, that is $r = 5$, with thresholds $k_1 = 5$ and $k_2 = 4$. Since $r > d + 1$, the access structure is of type Δ_1 . Setting $A = U_{1,4}$, $B = \{u_5\}$, and $C = U_{6,9}$, a set is authorized if it has at least four users from C , or if it has at least five users, at least two of whom are from $B \cup C$.

We summarize this section in the next theorem.

THEOREM 7.6. *Let Γ be an ideal WTAS with $M = U_{1,d} \cup U_{d+2,k}$ being its lexicographical minimal minterm for some $1 \leq d \leq k - 2$ and $k \leq n$. If there exists a minterm in Γ that has u_2 as its minimal member and Γ has no self-sufficient users, then Γ is a TPAS.*

7.1. Ideal linear secret sharing schemes for TPASs. We now turn to describe ideal linear secret sharing schemes that implement the TPASs Δ_1 and Δ_2 in Definition 3.3. Those schemes extend the ideas of the schemes that were presented

in [25] for bipartite access structures. In fact, if we take $B = \emptyset$, then our schemes coincide with the schemes of [25].

Let us first briefly recall the definition of those access structures. Assume that $|U| = n$ and $U = A \cup B \cup C$, where A , B , and C are disjoint, and $A, C \neq \emptyset$. Let $m, d, t \in \mathbb{N}$ be positive integers such that $m \geq t$. Then

$$\Delta_1 = \{X \subseteq U : (|X| \geq m \text{ and } |X \cap (B \cup C)| \geq m - d) \text{ or } |X \cap C| \geq t\}$$

and

$$\Delta_2 = \{X \subseteq U : (|X| \geq m \text{ and } |X \cap C| \geq m - d) \text{ or } |X \cap (B \cup C)| \geq t\},$$

where in Δ_2 the size of B is restricted to $|B| \leq d + t - m$.

Let $\mathbb{F}$ be a sufficiently large finite field. For a $\lambda \in \mathbb{F}$ and an integer s define the Vandermonde vector $\mathbf{V}_s(\lambda) := (1, \lambda, \lambda^2, \dots, \lambda^{s-1})$ (for clarity, we denote hereinafter all vectors with bold-faced letters). Recall that for any set $A \subseteq \mathbb{F}$ if $|A| \leq s$, then the vectors $\{\mathbf{V}_s(\lambda) : \lambda \in A\}$ are linearly independent. Thus, for any set $A \subseteq \mathbb{F}$ if $|A| \geq s$, the vectors $\{\mathbf{V}_s(\lambda) : \lambda \in A\}$ span the entire space $\mathbb{F}^s$. Let $E = \mathbb{F}^m$ and E_1 and E_2 be subspaces of E of dimensions d and t , respectively, such that if $d + t \geq m$, then $E = E_1 + E_2$; otherwise $E_1 \cap E_2 = \{\vec{0}\}$. Denote $r = \dim(E_1 \cap E_2) = \max(0, d + t - m)$. If $r > 0$, let $\Lambda = \{\lambda_1, \dots, \lambda_r\}$ be a set of elements from $\mathbb{F}$; otherwise $\Lambda = \emptyset$. Let

$$\varphi_1 : \mathbb{F}^d \rightarrow E_1 \quad \text{and} \quad \varphi_2 : \mathbb{F}^t \rightarrow E_2$$

be two isomorphisms such that

$$(7.1) \quad \varphi_1(\mathbf{V}_d(\lambda_i)) = \varphi_2(\mathbf{V}_t(\lambda_i)) \quad \forall \lambda_i \in \Lambda.$$

Finally, we define the functions $\mathbf{a} : \mathbb{F} \rightarrow E_1$ and $\mathbf{b} : \mathbb{F} \rightarrow E_2$ as follows:

$$\mathbf{a}(\lambda) = \varphi_1(\mathbf{V}_d(\lambda)) \quad \text{and} \quad \mathbf{b}(\lambda) = \varphi_2(\mathbf{V}_t(\lambda)) \quad \forall \lambda \in \mathbb{F}.$$

The target vector in the ideal linear secret sharing schemes that we construct below is $\mathbf{b}(\lambda_T)$ for some $\lambda_T \in \mathbb{F} \setminus \Lambda$. Note that the target vector is in $E_2 \setminus E_1$.

Next, we describe the function $\boldsymbol{\eta} : U \rightarrow E$ that assigns vectors from E to users in U , so that the vectors of authorized subsets span the target vector, while the vectors of unauthorized subsets do not; i.e., $\mathbf{b}(\lambda_T) \in \text{span}\{\boldsymbol{\eta}(S)\}$ if and only if $S \in \Gamma$. By Theorem 2.5, this implies that the scheme realizes Γ . The technique we use is similar for both types of access structures, $\Gamma = \Delta_1$ and $\Gamma = \Delta_2$. We assign the vectors to users one at a time and prove correctness by induction. Assume that we already assigned vectors to all users in $U' \subset U$ and that we wish to find a proper assignment of a vector to an additional user $u \in U \setminus U'$. We shall denote this stage in the assignment by $\langle U', u \rangle$. The goal is to find an assignment $\boldsymbol{\eta}(u)$ of a vector so that the following conditions hold:

- C1. If $S \subseteq U'$, $S \notin \Gamma$ but $S' := S \cup \{u\} \in \Gamma$, then $\boldsymbol{\eta}(u)$ should be such that $\mathbf{b}(\lambda_T) \in \text{span}\{\boldsymbol{\eta}(S')\}$.
- C2. If $S \subseteq U'$, $S \notin \Gamma$ and $S' := S \cup \{u\} \notin \Gamma$, then $\boldsymbol{\eta}(u)$ should be such that $\mathbf{b}(\lambda_T) \notin \text{span}\{\boldsymbol{\eta}(S')\}$.

We shall always look for the vector $\boldsymbol{\eta}(u)$ in a single-parameterized family of vectors $\{\boldsymbol{\zeta}(\lambda) \in E : \lambda \in \mathbb{F}\}$ (in the constructions that we present here, $\boldsymbol{\zeta}(\lambda)$ is one of $\mathbf{a}(\lambda)$, $\mathbf{b}(\lambda)$, and $\mathbf{V}_m(\lambda)$). Using Lemma 7.7 below, we shall show that at each stage of the assignment, $\langle U', u \rangle$, and for every $S \subseteq U'$, where $S \notin \Gamma$, there are finitely many bad

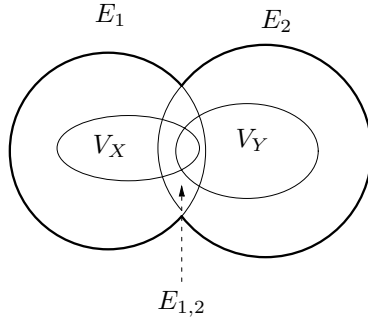


FIG. 7.1. The linear spaces in the proof of Lemma 7.7.

choices of the parameter λ that give rise to vectors $\eta(u) = \zeta(\lambda)$ that would violate the relevant condition above. Hence, since $|2^{U'} \setminus \Gamma|$ is bounded, there exists a good choice of λ such that $\eta(u) = \zeta(\lambda)$ satisfies all necessary conditions, provided that the field is sufficiently large. Thus, in the rest of the proof we bound the number of bad values for a specific set S .

Similarly to [25], it can be verified that the size of the field that our schemes require, which is the size of the domain of secrets and shares, is $2^{O(n)}$.

LEMMA 7.7. *Let $X \subseteq \mathbb{F}$ and $Y \subseteq \mathbb{F} \setminus \Lambda$ be such that $|X| < d$ and $|X| + |Y| < m$. Let $V = \text{span}\{\mathbf{a}(x), \mathbf{b}(y) : x \in X, y \in Y\}$. Then $E_1 \setminus V \neq \emptyset$.*

Proof. Hereinafter, let $V_X = \text{span}\{\mathbf{a}(x) : x \in X\}$, $V_Y = \text{span}\{\mathbf{b}(y) : y \in Y\}$, and $E_{1,2} = E_1 \cap E_2$ (see Figure 7.1 for an illustration of those linear spaces). Let us begin with the fairly easy case when $m \geq d + t$. In that case, since $E_1 \cap E_2 = \{\vec{0}\}$, we claim that $V \cap E_1 = V_X$, whence $\dim(E_1 \setminus V) = \dim E_1 - \dim V_X = d - |X| > 0$. Indeed, let $\mathbf{u}$ be an arbitrary vector from $V \cap E_1$. Then, on one hand, $\mathbf{u} = \mathbf{v}_x + \mathbf{v}_y$ for some $\mathbf{v}_x \in V_X$ and $\mathbf{v}_y \in V_Y$, and on the other hand, $\mathbf{u} \in E_1$. Therefore, $\mathbf{v}_y = \mathbf{u} - \mathbf{v}_x \in E_1 - V_X = E_1$. But since $\mathbf{v}_y \in V_Y \subseteq E_2$ and $E_1 \cap E_2 = \{\vec{0}\}$, we infer that $\mathbf{v}_y = \vec{0}$. This implies that $\mathbf{u} = \mathbf{v}_x \in V_X$.

Next, we prove the claim for the case where $r = d + t - m > 0$. Without loss of generality, we may assume that $|Y| = t - k$ for some $0 \leq k \leq r$. Indeed, if $|Y| > t$, it may be reduced to $Y^- \subset Y$ of size $|Y^-| = t$ without affecting V . If, on the other hand, $|Y| < t - r$, it may be extended to a superset $Y^+ \supset Y$ of size $|Y^+| = t - r$. Since $|X| + |Y^+| < d + t - r = m$, we may prove the statement for X and Y^+ and then infer that it also holds for X and Y .

Assume, towards a contradiction, that $E_1 \subseteq V$, whence $\dim(V \cap E_1) = \dim E_1 = d$. We claim that this assumption implies the following two statements:

$$(7.2) \quad \dim(V_Y \cap E_{1,2}) = r - k \quad \text{and} \quad \dim(V_X \cap E_{1,2}) < k.$$

After proving these two statements, we may derive the sought-after contradiction as follows. Since we assumed that $E_1 \subseteq V = V_X + V_Y$, all vectors $e \in E_{1,2}$ may be expressed as $e = x + y$, where $x \in V_X$ and $y \in V_Y$. But since $\dim E_{1,2} = r > \dim(V_X \cap E_{1,2}) + \dim(V_Y \cap E_{1,2})$, as implied by the two statements in (7.2), there exists at least one vector $e \in E_{1,2}$ such that $e = x + y$, for which either $x \in E_1 \setminus E_2$ or $y \in E_2 \setminus E_1$. Without loss of generality, we assume that $x \in E_1 \setminus E_2$. But then $x = e - y \in E_{1,2} + V_Y \subseteq E_2$, as opposed to our assumption that $x \notin E_2$.

Next, we prove the statements in (7.2). Let us begin with the equality in (7.2). The dimension of $E_{1,2} + V_Y$ is t since the $t - k$ base vectors of V_Y , namely $\{\mathbf{b}(y) : y \in Y\}$, may be extended to a full basis of E_2 by any k base vectors of $E_{1,2}$ of the form $\mathbf{b}(\lambda)$, $\lambda \in \Lambda$ (recall that $|\Lambda| = r \geq k$ and that $Y \subset \mathbb{F} \setminus \Lambda$). Therefore,

$$\dim(V_Y \cap E_{1,2}) = \dim V_Y + \dim E_{1,2} - \dim(E_{1,2} + V_Y) = (t - k) + r - t = r - k.$$

In order to prove the inequality in (7.2), we first prove that

$$(7.3) \quad \dim(V_X + E_{1,2}) = d.$$

Since $V_X + E_{1,2} \subseteq E_1$ and $\dim E_1 = d$, all we need to show is that $E_1 \subseteq V_X + E_{1,2}$. We assumed that $E_1 \subseteq V = V_X + V_Y$. Hence, for each $e \in E_1$ there exist $x \in V_X$ and $y \in V_Y$ such that $e = x + y$. But then $y = e - x \in E_1 + V_X = E_1$. Since $y \in V_Y \subset E_2$ we conclude that $y \in E_{1,2}$. This implies that $e \in V_X + E_{1,2}$, whence we arrive at the required conclusion that $E_1 \subseteq V_X + E_{1,2}$. Finally, the inequality in (7.2) is a straightforward consequence of (7.3). Indeed,

$$\dim(V_X \cap E_{1,2}) = \dim V_X + \dim E_{1,2} - \dim(V_X + E_{1,2}) = |X| + r - d;$$

recalling that $|X| < m - |Y| = m - t + k$ and $r = t + d - m$, we infer that $\dim(V_X \cap E_{1,2}) < k$. That concludes the proof. $\square$

7.1.1. Realizing Δ_1 . Here, Γ is an access structure of type Δ_1 . We assign the vectors in three stages. We first describe the assignment $\boldsymbol{\eta}(u)$ for all $u \in C$, then for $u \in A$, and finally for $u \in B$.

Step 1: Assigning vectors to users in C . Let $C = \{c_i\}_{1 \leq i \leq |C|}$, and let γ_i , where $1 \leq i \leq |C|$, be distinct elements of $\mathbb{F} \setminus \Lambda$, all different from λ_T . Then

$$\boldsymbol{\eta}(c_i) = \mathbf{b}(\gamma_i), \quad 1 \leq i \leq |C|.$$

This is essentially the Shamir t -out-of- $|C|$ scheme. Therefore, every t users in C can reconstruct $\mathbf{b}(\lambda_T)$, while no $t - 1$ users can.

Step 2: Assigning vectors to users in A . Let $A = \{a_i\}_{1 \leq i \leq |A|}$. Our claim is that, assuming $\mathbb{F}$ is large enough, we may find $\alpha_i \in \mathbb{F}$, where $1 \leq i \leq |A|$, so that the assignment $\boldsymbol{\eta}(a_i) = \mathbf{a}(\alpha_i)$ is consistent with Γ , restricted to $A \cup C$, and, in addition, for any $S \subseteq A \cup C$, where $S \notin \Gamma$, the following equality holds:

$$(7.4) \quad \dim \text{span}\{\boldsymbol{\eta}(S)\} = |S \cap (B \cup C)| + \min(|S \cap A|, d).$$

(Herein, we could have replaced $S \cap (B \cup C)$ with $S \cap C$, since S does not include users from B .) Namely, whenever we augment an unauthorized set $S \subseteq A \cup C$ with an additional user from A , the dimension of the corresponding subspace that is spanned by the vectors owned by S increases by 1, up to a limit of d users from A (i.e., all users from A beyond the first d ones are redundant). Using (7.4), we guarantee that if $S \in \Gamma$ and $|S \cap C| < t$, then the vectors assigned to S span E , and in particular, span the target vector $\mathbf{b}(\lambda_T)$. We shall prove this claim by induction. Assume that we already assigned vectors to $A_{1,j} := \{a_1, \dots, a_j\}$ for some $0 \leq j < |A|$, so that Γ is respected and (7.4) holds for all $S \in 2^{U'} \setminus \Gamma$, where $U' = A_{1,j} \cup C$. This assumption is obviously true for $j = 0$. We proceed to look for an assignment $\boldsymbol{\eta}(a_{j+1}) = \mathbf{a}(\alpha_{j+1})$ for the next user from A so that conditions C1 and C2 hold for all $S \in 2^{U'} \setminus \Gamma$ (in order to keep respecting Γ), and equality (7.4) holds for all $S \in 2^{U' \cup \{a_{j+1}\}} \setminus \Gamma$. As explained

earlier, it suffices to concentrate on a single set $S \in 2^{U'} \setminus \Gamma$ and bound the number of bad choices of $\alpha_{j+1} \in \mathbb{F}$ for that specific set.

Let $S \in 2^{U'} \setminus \Gamma$ be such that $S' := S \cup \{a_{j+1}\} \in \Gamma$. Since $|S'| = |S| + 1$ but $|S' \cap (B \cup C)| = |S \cap (B \cup C)|$, the fact that $S \notin \Gamma$ and $S' \in \Gamma$ implies that $|S| = m - 1$ and $|S \cap (B \cup C)| \geq m - d$. This, in turn, implies that $|S \cap A_{1,j}| < d$. Therefore, by Lemma 7.7, applied to $X = \{\alpha_i\}_{1 \leq i \leq j}$ and $Y = \{\gamma_i\}_{1 \leq i \leq |C|}$, the vectors owned by S span a subspace of E_1 of dimension $d - 1$ at the most. Hence, for all but at most $d - 1$ values of $\alpha_{j+1} \in \mathbb{F}$, we have $\dim \text{span}\{\boldsymbol{\eta}(S')\} = \dim \text{span}\{\boldsymbol{\eta}(S)\} + 1$. Now, by the induction hypothesis, the set S satisfies (7.4). Therefore, since $|S \cap A| = |S \cap A_{1,j}| < d$, we get that

$$\dim \text{span}\{\boldsymbol{\eta}(S)\} = |S \cap C| + \min(|S \cap A|, d) = |S \cap C| + |S \cap A| = |S| = m - 1.$$

This implies that $\dim \text{span}\{\boldsymbol{\eta}(S')\} = m = \dim E$ so that $\mathbf{b}(\lambda_T) \in \text{span}\{\boldsymbol{\eta}(S')\}$ and C1 holds.

Let $S \in 2^{U'} \setminus \Gamma$ be such that $S' := S \cup \{a_{j+1}\} \notin \Gamma$. If $|S \cap A_{1,j}| \geq d$, then $\text{span}\{\boldsymbol{\eta}(S')\} = \text{span}\{\boldsymbol{\eta}(S)\}$ for any choice of α_{j+1} ; hence, no matter how we chose α_{j+1} in this case, the augmented set S' will remain incapable of spanning the target vector. Also, since any d vectors of the form $\mathbf{a}(\alpha_i)$ span the same subspace E_1 , any choice of α_{j+1} will result with an assignment that keeps respecting (7.4).

Hence, we restrict our attention to subsets $S \in 2^{U'} \setminus \Gamma$ for which $|S \cap A_{1,j}| < d$. We claim that in this case $|S| \leq m - 2$. Assume, by contradiction, that $|S| \geq m - 1$. As $|S \cap A_{1,j}| < d$, we infer that $|S \cap C| \geq m - d$. Since $|S' \cap C| = |S \cap C| \geq m - d$ and $|S'| = |S| + 1 \geq m$, we infer that $S' \in \Gamma$, as opposed to our assumption. Hence, $|S| \leq m - 2$. Our goal is to find α_{j+1} so that the target vector is not in $\text{span}\{\boldsymbol{\eta}(S')\}$ (in order to keep respecting Γ) and that (7.4) still holds for all $S \in 2^{U'} \setminus \Gamma$. Concentrating on the first part of our mission, we want to find α_{j+1} so that $\mathbf{a}(\alpha_{j+1}) \notin \text{span}\{\boldsymbol{\eta}(S) \cup \{\mathbf{b}(\lambda_T)\}\}$. By Lemma 7.7, for $X = \{\alpha_i\}_{1 \leq i \leq j}$ and $Y = \{\gamma_i\}_{1 \leq i \leq |C|} \cup \{\lambda_T\}$, the space $\text{span}\{\boldsymbol{\eta}(S) \cup \{\mathbf{b}(\lambda_T)\}\} \cap E_1$ is a subspace of E_1 of dimension $d - 1$ at the most. We infer that for all but possibly $d - 1$ values of $\alpha_{j+1} \in \mathbb{F}$, the vector $\mathbf{a}(\alpha_{j+1})$ increases the dimension of $\text{span}\{\boldsymbol{\eta}(S) \cup \{\mathbf{b}(\lambda_T)\}\}$ by 1. It is easy to see that for all such selections, we get an assignment that will also respect (7.4). This completes the proof that a proper assignment of vectors exists for all users in $A \cup C$. We are ready to proceed to the third and final stage in the assignment.

Step 3: Assigning vectors to users in B. Here we assign vectors to the users in B (note that B may be empty). Let $B = \{b_i\}_{1 \leq i \leq |B|}$. As before, we claim that, provided $\mathbb{F}$ is large enough, there exist $\beta_i \in \mathbb{F}$, where $1 \leq i \leq |B|$, so that the assignment $\boldsymbol{\eta}(b_i) = \mathbf{V}_m(\beta_i)$ is consistent with Γ , and, in addition, for any $S \notin \Gamma$, (7.4) holds. Namely, whenever we augment an unauthorized set S with an additional user from B , the corresponding subspace that is spanned by the vectors owned by S increases by 1. We shall prove this claim by induction. Assume that we already assigned vectors to $B_{1,j} := \{b_1, \dots, b_j\}$ for some $0 \leq j < |B|$, so that Γ is respected and (7.4) holds for all $S \in 2^{U'} \setminus \Gamma$, where $U' = B_{1,j} \cup A \cup C$. This assumption is obviously true for $j = 0$. Next, we look for an assignment for the next user from B , that is, $\boldsymbol{\eta}(b_{j+1}) = \mathbf{V}_m(\beta_{j+1})$, so that conditions C1 and C2 hold for all $S \in 2^{U'} \setminus \Gamma$ (in order to keep respecting Γ), and (7.4) holds for all $S \in 2^{U' \cup \{b_{j+1}\}} \setminus \Gamma$. Again, we concentrate on a single set $S \in 2^{U'} \setminus \Gamma$ and bound the number of bad choices of $\beta_{j+1} \in \mathbb{F}$ for that specific set.

Let $S \in 2^{U'} \setminus \Gamma$ be such that $S' := S \cup \{b_{j+1}\} \in \Gamma$. Since $S \notin \Gamma$ but $S' \in \Gamma$, and the added user b_{j+1} is from B , we conclude that while S fails to satisfy the first of the two

threshold conditions in the definition of Δ_1 , i.e., $|S| < m$ or $|S \cap (B \cup C)| < m - d$, the set S' does satisfy that condition, namely, $|S'| \geq m$ and $|S' \cap (B \cup C)| \geq m - d$. There are two cases to consider here, according to which of the two threshold conditions were violated by S :

1. $|S| = m - 1$.
2. $|S| \geq m$ but $|S \cap (B \cup C)| = m - d - 1$.

In the first case, we claim that $|S \cap A| \leq d$. Indeed, since $|S' \cap (B \cup C)| \geq m - d$ and $|S' \cap (B \cup C)| = |S \cap (B \cup C)| + 1$, we have $|S \cap (B \cup C)| \geq m - d - 1$, whence $|S \cap A| = |S| - |S \cap (B \cup C)| \leq (m - 1) - (m - d - 1) = d$. Applying the induction hypothesis to S , we get by (7.4) that

$$\dim \text{span}\{\boldsymbol{\eta}(S)\} = |S \cap (B \cup C)| + \min(|S \cap A|, d) = |S \cap (B \cup C)| + |S \cap A| = |S| = m - 1.$$

Therefore, for all but $m - 1$ values of β_{j+1} ,

$$\dim \text{span}\{\boldsymbol{\eta}(S')\} = \dim \text{span}\{\boldsymbol{\eta}(S)\} + 1 = m,$$

whence $\mathbf{b}(\lambda_T) \in \text{span}\{\boldsymbol{\eta}(S')\}$.

Next, we show how to reduce the second case to the first case. In the second case $|S \cap A| = |S| - |S \cap (B \cup C)| \geq m - (m - d - 1) = d + 1$. Hence, S includes too many users from A , since for all $a \in A$ the vector $\boldsymbol{\eta}(a)$ is chosen from the d -dimensional space E_1 . Let $a \in S \cap A$, and define $S_1 = S \setminus \{a\}$ and $S'_1 = S_1 \cup \{b_{j+1}\}$. Obviously, $S_1 \notin \Gamma$ and, as can be easily verified, $S'_1 \in \Gamma$. Since $\text{span}\{\boldsymbol{\eta}(S)\} = \text{span}\{\boldsymbol{\eta}(S_1)\}$ and $\text{span}\{\boldsymbol{\eta}(S')\} = \text{span}\{\boldsymbol{\eta}(S'_1)\}$, if β_{j+1} is chosen properly for S_1 , it will also be an appropriate choice for S . Hence, we may repeat this reduction stage until we reach a set S_1 of size $|S_1| = m - 1$, and then apply the arguments of the first case to conclude that all but $m - 1$ values of β_{j+1} are appropriate choices.

Let $S \in 2^{U'} \setminus \Gamma$ be such that $S' := S \cup \{b_{j+1}\} \notin \Gamma$. Arguing along the same lines as above, we may assume, without loss of generality, that $|S \cap A| \leq d$, since any d members of A span E_1 . We claim that $|S| \leq m - 2$. Indeed, if $|S| \geq m - 1$, then $|S \cap (B \cup C)| \geq m - 1 - d$, as $|S \cap A| \leq d$. This implies that $|S' \cap (B \cup C)| \geq m - d$ and $|S'| \geq m$, whence $S' \in \Gamma$, in contradiction to our assumption. Therefore, the dimension of $\text{span}\{\boldsymbol{\eta}(S) \cup \{\mathbf{b}(\lambda_T)\}\}$ is at most $m - 2 + 1$. Consequently, for all but possibly $m - 1$ values of $\beta_{j+1} \in \mathbb{F}$, we have $\mathbf{V}_m(\beta_{j+1}) \notin \text{span}\{\boldsymbol{\eta}(S) \cup \{\mathbf{b}(\lambda_T)\}\}$. Hence, for all but possibly $m - 1$ values of $\beta_{j+1} \in \mathbb{F}$, we get that $\mathbf{b}(\lambda_T) \notin \text{span}\{\boldsymbol{\eta}(S) \cup \{\mathbf{V}_m(\beta_{j+1})\}\}$. Clearly, all such selections yield assignments that respect (7.4).

This concludes the description of the assignment $\boldsymbol{\eta} : U \rightarrow E$ and the proof that it yields an ideal linear sharing scheme that realizes a given access structure Γ of type Δ_1 .

7.1.2. Realizing Δ_2 . Here we propose an ideal linear secret sharing scheme for access structures Γ that are TPASs of type Δ_2 . Recall that, by Definition 3.3, in such access structures $|B| \leq |\Lambda| = r = d + t - m$. We concentrate on cases where $r > 0$ (if $r = 0$, then $B = \emptyset$, and then we get a bipartite access structure).

Step 1: Assigning vectors to users in $B \cup C$. The users in B and C are assigned distinct vectors of the form $\mathbf{b}(\beta_i)$ and $\mathbf{b}(\gamma_i)$, respectively, where $\{\beta_i\}_{1 \leq i \leq |B|} \subset \Lambda$ and $\{\gamma_i\}_{1 \leq i \leq |C|} \subset \mathbb{F} \setminus (\Lambda \cup \{\lambda_T\})$. Clearly, any t users from $B \cup C$ may reconstruct $\mathbf{b}(\lambda_T)$ while no $t - 1$ users can.

Step 2: Assigning vectors to users in A . Let $A = \{a_i\}_{1 \leq i \leq |A|}$. Our claim is that, assuming $\mathbb{F}$ is large enough, we may find $\alpha_i \in \mathbb{F}$, where $1 \leq i \leq |A|$, so that the assignment $\boldsymbol{\eta}(a_i) = \mathbf{a}(\alpha_i)$ is consistent with Γ , and, in addition, for any $S \subseteq U$ where

$S \notin \Gamma$, the following equality holds:

$$(7.5) \quad \dim \text{span}\{\boldsymbol{\eta}(S)\} = |S \cap C| + \min(|S \cap (A \cup B)|, d).$$

We shall prove this claim by induction. Assume that we already assigned vectors to $A_{1,j} := \{a_1, \dots, a_j\}$ for some $0 \leq j < |A|$, so that Γ is respected and (7.5) holds for all $S \in 2^{U'} \setminus \Gamma$, where $U' = A_{1,j} \cup B \cup C$. This assumption is true for $j = 0$ since then, on one hand, $\dim \text{span}\{\boldsymbol{\eta}(S)\} = |S|$, while, on the other hand, $S \cap A = \emptyset$ and $|S \cap B| \leq |B| \leq d+t-m \leq d$. We proceed to look for an assignment $\boldsymbol{\eta}(a_{j+1}) = \mathbf{a}(\alpha_{j+1})$ for the next user from A , so that conditions C1 and C2 hold for all $S \in 2^{U'} \setminus \Gamma$ and (7.5) holds for all $S \in 2^{U' \cup \{a_{j+1}\}} \setminus \Gamma$.

Let $S \in 2^{U'} \setminus \Gamma$ be such that $S' := S \cup \{a_{j+1}\} \in \Gamma$. We infer that $|S| = m - 1$ and $|S \cap C| \geq m - d$. This implies that $|S \cap (A \cup B)| = |S| - |S \cap C| \leq d - 1$. By Lemma 7.7, applied to $X = \{\alpha_i\}_{1 \leq i \leq j} \cup \{\beta_i\}_{1 \leq i \leq |B|}$ and $Y = \{\gamma_i\}_{1 \leq i \leq |C|}$ (here we use the fact that $\beta_i \in \Lambda$ so that, in view of (7.1), $\mathbf{a}(\beta_i) = \mathbf{b}(\beta_i)$), the vectors owned by S span a subspace of E_1 of dimension $d - 1$ at the most. Hence, for all but $d - 1$ values of $\alpha_{j+1} \in \mathbb{F}$, $\dim \text{span}\{\boldsymbol{\eta}(S')\} = \dim \text{span}\{\boldsymbol{\eta}(S)\} + 1$. Now, by the induction hypothesis, the set S satisfies (7.5). Therefore, since $|S \cap (A \cup B)| < d$, we get that

$$\begin{aligned} \dim \text{span}\{\boldsymbol{\eta}(S)\} &= |S \cap C| + \min(|S \cap (A \cup B)|, d) \\ &= |S \cap C| + |S \cap (A \cup B)| = |S| = m - 1. \end{aligned}$$

This implies that $\dim \text{span}\{\boldsymbol{\eta}(S')\} = m$ so that $\mathbf{b}(\lambda_T) \in \text{span}\{\boldsymbol{\eta}(S')\}$.

Let $S \in 2^{U'} \setminus \Gamma$ be such that $S' := S \cup \{a_{j+1}\} \notin \Gamma$. There are two cases to consider here. If $|S \cap (A \cup B)| \geq d$, then any choice of α_{j+1} is good. This is because, in view of (7.1), the vectors given to users in B are $\mathbf{b}(\beta_i) = \mathbf{a}(\beta_i)$, where $\beta_i \in \Lambda$, and consequently, if $|S \cap (A \cup B)| \geq d$, the vectors held by the users of S already span all of E_1 . Hence, for any choice of α_{j+1} , $\text{span}\{\boldsymbol{\eta}(S')\} = \text{span}\{\boldsymbol{\eta}(S)\}$ and S' will still satisfy (7.5).

This leaves us with the case $|S \cap (A \cup B)| < d$. In that case $|S| \leq m - 2$ (since if $|S| \geq m - 1$, then $|S \cap C| = |S| - |S \cap (A \cup B)| \geq m - d$ and $|S'| = |S| + 1 \geq m$, and consequently, S' would have been authorized). Our goal is to find α_{j+1} so that the target vector is not in $\text{span}\{\boldsymbol{\eta}(S')\}$ (in order to keep respecting Γ) and that (7.5) still holds. Concentrating on the first part of our mission, we want to find α_{j+1} so that $\mathbf{a}(\alpha_{j+1}) \notin \text{span}\{\boldsymbol{\eta}(S) \cup \{\mathbf{b}(\lambda_T)\}\}$. By Lemma 7.7, for $X = \{\alpha_i\}_{1 \leq i \leq j} \cup \{\beta_i\}_{1 \leq i \leq |B|}$ and $Y = \{\gamma_i\}_{1 \leq i \leq |C|} \cup \{\lambda_T\}$, the space $\text{span}\{\boldsymbol{\eta}(S) \cup \{\mathbf{b}(\lambda_T)\}\} \cap E_1$ is a subspace of E_1 of dimension $d - 1$ at the most. Hence, for all but possibly $d - 1$ values of $\alpha_{j+1} \in \mathbb{F}$, the vector $\mathbf{a}(\alpha_{j+1})$ increases the dimension of $\text{span}\{\boldsymbol{\eta}(S) \cup \{\mathbf{b}(\lambda_T)\}\}$ by 1. It is easy to see that for all such selections we get an assignment that will also respect (7.5).

8. A recursive characterization of ideal WTASs by means of composition. In this section we show that if Γ is an ideal WTAS that is not one of the structures that we identified in sections 6.3 and 7, namely, an HTAS or a TPAS, then it is a composition of two ideal WTASs that are defined on smaller sets of users. Such a result provides a recursive characterization of ideal WTASs.

8.1. WTASs and composition of access structures. We begin with the following fundamental lemma that asserts that a composition of two access structures is ideal if and only if those two access structures are ideal.

LEMMA 8.1. *Let U_1 and U_2 be disjoint sets. Let $u_1 \in U_1$, and define $U = U_1 \cup U_2 \setminus \{u_1\}$. Suppose that Γ_1 and Γ_2 are access structures on U_1 and U_2 , respectively,*

such that u_1 is not redundant in Γ_1 and $\Gamma_2 \neq \emptyset$. Furthermore, let Γ be the composition of Γ_1 and Γ_2 via u_1 . Then Γ is ideal if and only if both Γ_1 and Γ_2 are ideal. Moreover, if both Γ_1 and Γ_2 have an ideal linear secret sharing scheme, then Γ has an ideal linear secret sharing scheme.

Proof. Assume that Γ_1 and Γ_2 are ideal. We show that Γ is ideal by describing an ideal secret sharing scheme that realizes it. Given a secret s , we share it among the users of U_1 using an ideal secret sharing scheme for Γ_1 . Then, if s_1 is the share of u_1 in that scheme, we share it among the users of U_2 using an ideal secret sharing scheme for Γ_2 . The resulting scheme is clearly ideal, and its correctness and perfect security stem from the correctness and perfect security of the two schemes for Γ_1 and Γ_2 . It is easy to verify that if both secret sharing schemes for Γ_1 and for Γ_2 are linear, then so is the resulting scheme for Γ .

Conversely, suppose that Γ is ideal. Let $M \subset U$ be a minterm of Γ , and let $M = M_1 \cup M_2$, where $M_1 = M \cap (U_1 \setminus \{u_1\})$ and $M_2 = M \cap U_2$. We choose M for which $M_2 \neq \emptyset$ (since u_1 is not redundant in Γ_1 and $\Gamma_2 \neq \emptyset$ such a minterm exists). By Definition 3.4, $M \in \Gamma$ if and only if $M_1 \cup \{u_1\} \in \Gamma_1$ and $M_2 \in \Gamma_2$. Furthermore,

$$(8.1) \quad \begin{aligned} M_1 \cup M'_2 &\in \Gamma \text{ if and only if } M'_2 \in \Gamma_2 \text{ and} \\ M'_1 \cup M_2 &\in \Gamma \text{ if and only if } M'_1 \cup \{u_1\} \in \Gamma_1. \end{aligned}$$

From the first observation above we conclude that Γ_2 coincides with Γ_{M_1, U_2} —the restriction of Γ that M_1 induces on U_2 . Since by Lemma 4.11 any restriction of an ideal access structure is ideal, we infer that Γ_2 is ideal. As for Γ_1 , let x be an arbitrary element of M_2 , let $U' = U_1 \setminus \{u_1\} \cup \{x\}$, and consider the restriction

$$\Gamma' := \Gamma_{M_2 \setminus \{x\}, U'} = \{Q \subseteq U' : (M_2 \setminus \{x\}) \cup Q \in \Gamma\}.$$

We claim, and prove below, that Γ' is isomorphic to Γ_1 under the natural bijection from 2^{U_1} to $2^{U'}$ (i.e., $A \subseteq U_1$ is mapped to $A' := A \setminus \{u_1\} \cup \{x\} \in U'$ if $u_1 \in A$, and to $A' = A$ otherwise). Therefore, since Γ' is a restriction of Γ , it is ideal, as implied by Lemma 4.11, and, consequently, so is Γ_1 .

We conclude with a proof of the isomorphism. Assume that $A \in \Gamma_1$. If $u_1 \notin A$, then $A \in \Gamma$. Hence, also $A' = A \in \Gamma$ and, in particular, $A' \in \Gamma'$. If $u_1 \in A$, then $(A \setminus \{u_1\}) \cup M_2 \in \Gamma$ (this is implied by the second observation in (8.1)); hence, $A' \cup (M_2 \setminus \{x\}) \in \Gamma$, and therefore $A' \in \Gamma'$. Conversely, assume that $A' \in \Gamma'$, namely, $(M_2 \setminus \{x\}) \cup A' \in \Gamma$. If $x \notin A'$, then $A' \subseteq U_1 \setminus \{u_1\}$. As M_2 is a minterm of Γ_2 (this is implied by the first observation in (8.1)), $M_2 \setminus \{x\} \notin \Gamma_2$. Hence, by Definition 3.4, we conclude that $A' \in \Gamma_1$. But in that case $A = A'$ so that $A \in \Gamma_1$. If $x \in A'$, then $(A' \setminus \{x\}) \cup M_2 \in \Gamma$, from which it follows that $A = (A' \setminus \{x\}) \cup \{u_1\} \in \Gamma_1$. $\square$

The recursive characterization of ideal WTASs will be obtained by distinguishing between two types of users. Specifically, we shall identify a subset of so-called strong users that takes the form of a suffix, $S = U_{k,n}$, where $k \geq 3$, and then the complement subset will be thought of as the subset of weak users, $W = U_{1,k-1}$. A subset of strong users will be called S -cooperative if it is unauthorized, but it may become authorized if we add to it some weak users.

DEFINITION 8.2 (cooperative set). *Given $Y \subseteq S$, if $Y \notin \Gamma$ but $W \cup Y \in \Gamma$, then Y is called an S -cooperative set.*

By Claim 4.13, the access structure $\Gamma_{Y,W}$, the restriction of Γ induced by Y on W , is a WTAS for any partition $U = W \cup S$ and $Y \subseteq S$. We proceed to define a condition on the set S such that if it is satisfied for some suffix $S = U_{k,n}$, where $k \geq 3$,

the access structure Γ is a composition of two ideal WTASs that are defined on sets smaller than U .

DEFINITION 8.3 (strong set). *If for any two S -cooperative sets $Y_1, Y_2 \subseteq S$ the corresponding restrictions of Γ to W coincide, i.e., $\Gamma_{Y_1, W} = \Gamma_{Y_2, W}$, the set S is called a strong set of users.*

If S is a strong set of users, there exists an access structure on W , denoted Γ_W , such that $\Gamma_W = \Gamma_{Y, W}$ for all cooperative subsets $Y \subseteq S$. In that case, every minterm $M \in \Gamma$ is either contained in S or, otherwise, $M \cap W \in \Gamma_{M \cap S, W}$. In the second case we infer that $M \cap S$ is S -cooperative, and since S is strong we get that $M \cap W \in \Gamma_W$. The following theorem shows that if S is a strong set of users, Γ is a composition of two ideal WTASs.

THEOREM 8.4. *Let Γ be an ideal WTAS on U . Suppose that $S = U_{k, n}$, for some $k \geq 3$, is a strong set of users. Then Γ is a composition of two ideal WTASs, each of which is defined on a smaller set than U (one is defined on W , and the other is defined on S supplemented by an additional user).*

Proof. Let Y be an arbitrary S -cooperative set, $Q \subseteq W$ be a minterm of $\Gamma_{Y, W}$, and x be an arbitrary element of Q . Both $\Gamma_{Y, W}$ and $\Gamma_{Q \setminus \{x\}, S \cup \{x\}}$ are restrictions of Γ . Thus, by Lemma 4.11, they are both ideal, and by Claim 4.13, they are both WTASs. Let Δ be the composition of $\Gamma_{Q \setminus \{x\}, S \cup \{x\}}$ and $\Gamma_{Y, W}$ via x . (Since x belongs to the domains of both access structures, W and $S \cup \{x\}$, we may define the composition by first replacing the user x in $\Gamma_{Q \setminus \{x\}, S \cup \{x\}}$ with a copy x' and then compose $\Gamma_{Q \setminus \{x\}, S \cup \{x\}}$ and $\Gamma_{Y, W}$ via x' .) We note that Δ , like Γ , is defined on $U = W \cup S$. We proceed to show that $\Gamma = \Delta$, thus proving the lemma. (Note that $\Gamma_{Y, W}$ is defined on W , whose size is $k - 1 < n$, and Γ' is defined on $S \cup \{x\}$, whose size is $n - k + 2 < n$.)

Let M be a minterm of Γ . Then either $M \subseteq S$ or $M \cap W \neq \emptyset$. If $M \subseteq S$, then $M \in \Gamma_{Q \setminus \{x\}, S \cup \{x\}}$. Since $x \notin M$ we conclude, by Definition 3.4, that $M \in \Delta$. If, on the other hand, $M \cap W \neq \emptyset$, define $M_W = M \cap W$ and $M_S = M \cap S$. Therefore, M_S is an S -cooperative set. Since S is a strong set of users, we infer that $\Gamma_{M_S, W} = \Gamma_{Y, W}$, whence $M_W \in \Gamma_{Y, W}$. In addition, $Q \in \Gamma_{Y, W} = \Gamma_{M_S, W}$. This implies that $M_S \cup Q \in \Gamma$ and, consequently, that $M_S \cup \{x\} \in \Gamma_{Q \setminus \{x\}, S \cup \{x\}}$. Therefore, by Definition 3.4, $M = M_W \cup M_S \in \Delta$.

For the other direction, let M be a minterm of Δ , and, as before, $M_S = M \cap S$ and $M_W = M \cap W$. Then either $M_S \in \Gamma_{Q \setminus \{x\}, S \cup \{x\}}$ or both $M_W \in \Gamma_{Y, W}$ and $M_S \cup \{x\} \in \Gamma_{Q \setminus \{x\}, S \cup \{x\}}$. We separate the discussion into the following two cases.

If $M_S \in \Gamma_{Q \setminus \{x\}, S \cup \{x\}}$, then $M_S \cup (Q \setminus \{x\}) \in \Gamma$. That means that $Q \setminus \{x\} \in \Gamma_{M_S, W}$. On the other hand, as Q is a minterm of $\Gamma_{Y, W}$, it must hold that $Q \setminus \{x\} \notin \Gamma_{Y, W}$. Hence, $\Gamma_{M_S, W} \neq \Gamma_{Y, W}$. This implies that M_S is not S -cooperative. But since $M_S \cup (Q \setminus \{x\}) \in \Gamma$, the only way M_S is not S -cooperative is if M_S is authorized by itself in Γ . However, $M = M_W \cup M_S$ was a minterm of Δ . Therefore, M_W must be empty, whence $M = M_S$. This brings us to the sought-after conclusion that $M \in \Gamma$.

Now suppose that $M_S \notin \Gamma_{Q \setminus \{x\}, S \cup \{x\}}$ but $M_S \cup \{x\} \in \Gamma_{Q \setminus \{x\}, S \cup \{x\}}$ and $M_W \in \Gamma_{Y, W}$. Since $M_S \cup \{x\} \in \Gamma_{Q \setminus \{x\}, S \cup \{x\}}$ we get that $Q \cup M_S \in \Gamma$. On the other hand, $M_S \notin \Gamma_{Q \setminus \{x\}, S \cup \{x\}}$, and thus $M_S \notin \Gamma$. Therefore, M_S is S -cooperative. Since $M_W \in \Gamma_{Y, W}$ and S is a strong set of users, this implies that $M_W \in \Gamma_{M_S, W}$. Therefore $M_W \cup M_S \in \Gamma$, and thus $M \in \Gamma$. $\square$

8.2. Simple compositions. In this section we show that in two simple cases an ideal WTAS is a composition of two ideal WTASs defined on sets smaller than U . The first case is when there are self-sufficient users, and the second is where u_2 starts no minterm of Γ .

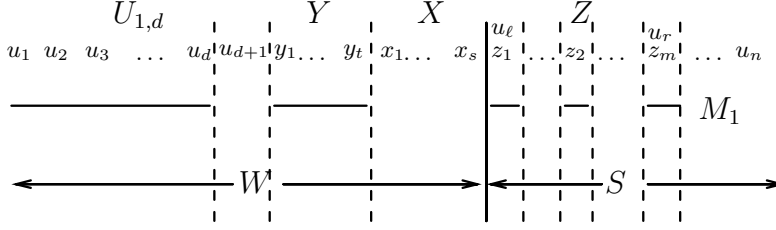


FIG. 8.1. Notations for the composition.

Suppose that u_n is a self-sufficient user. Let $\Gamma' = \Gamma_{\emptyset, U_{1,n-1}}$ be the restriction of Γ to $U_{1,n-1}$ (namely, all authorized sets that do not include u_n). By Lemma 4.11, the access structure Γ' is ideal, and by Claim 4.13, it is an ideal WTAS. Let $\Gamma_\vee$ be the simple 1-out-of-2 access structure on the set $\{u', u_n\}$, where u' is an additional dummy user. $\Gamma_\vee$ is clearly an ideal WTAS. It is easy to see that Γ is the composition of Γ' and $\Gamma_\vee$ via u' .

If Γ has no minterm that starts with u_2 , then every minterm that contains u_1 must contain also u_2 (otherwise, we could have replaced u_1 by u_2 in order to get a minterm that starts with u_2). Hence, for every $U_{3,n}$ -cooperative set, $Y \subseteq U_{3,n}$, the access structure that Y induces on $U_{1,2}$ is the same, $\Gamma_{Y, U_{1,2}} = \{U_{1,2}\}$. Therefore, $U_{3,n}$ is a strong set of users in this case. Hence, by Theorem 8.4, the access structure Γ is a composition of two ideal WTASs that are defined on sets smaller than U .

To conclude, we have proved the following lemma.

LEMMA 8.5. *Let Γ be an ideal WTAS on U . If Γ has self-sufficient users, or u_2 starts no minterm of Γ , then Γ is a composition of two ideal WTASs that are defined on sets smaller than U .*

8.3. Identifying composition structures. In this section we show that if Γ is an ideal WTAS, but is not one of the access structures that were characterized in Theorems 6.7 (HTAS) and 7.6 (TPAS), then it is a composition of two ideal WTASs as described in section 8.1. In view of Lemma 8.5, we assume hereinafter that u_2 is the minimal user in some minterm of Γ .

Let M_1 be the lexicographically minimal minterm in Γ . Let u_r be the maximal user in M_1 . Since Γ is neither an HTAS nor a TPAS, then, by Theorems 6.7 and 7.6, there must be at least two users in $U_{1,r-1}$ that are not in M_1 . Let u_ℓ be the *minimal* user in M_1 such that at least two users in $U_{1,\ell-1}$ are missing from M_1 , and let u_d be the *maximal* user in M_1 such that $U_{1,d} \subset M_1$. We denote the users in $M_1 \cap U_{d+1,\ell-1}$, if there are any, by $Y = \{y_1, \dots, y_t\}$. Note that if Y is not empty, then Y is a run of U , and $y_1 = u_{d+2}$. Next, if $Y \neq \emptyset$, we denote the set of users of $U \setminus M_1$ between y_t and u_ℓ (excluding those two users) by $X = \{x_1, \dots, x_s\}$. (Note that $X = U_{d+t+2,\ell-1}$.) Otherwise, we denote the set $U_{d+2,\ell-1}$ as $X = \{x_1, \dots, x_s\}$. Finally, we denote the users of $M_1 \cap U_{\ell,n}$ by $Z = \{z_1, \dots, z_m\}$. Note that the sets X and Z are never empty, and that $z_1 = u_\ell$. The above notations are depicted in Figure 8.1.

We claim that either $U_{\ell,n}$ or $U_{d+2,n}$ is a strong set of users. We start by partitioning U into $W = U_{1,\ell-1}$ and $S = U_{\ell,n}$. We show that if all the S -cooperative sets are of the same size, then $\Gamma_{Y_1, W} = \Gamma_{Y_2, W}$ for every two cooperative sets $Y_1, Y_2 \subseteq S$; namely, S is a set of strong users. If, however, that condition does not hold, we shall show that $U_{d+2,n}$ is a strong set.

An important result in our study is that all S -cooperative sets are at least as large as Z ; namely, their size is at least m .

CLAIM 8.6. *Let $W = U_{1,\ell-1}$, $S = U_{\ell,n}$, and M be a minterm of the access structure Γ . If $M \cap W \neq \emptyset$, then M contains at least m users from S .*

Proof. Assume towards a contradiction that M is a minterm that intersects W such that $|M \cap S| < m$. Recall that every minterm of Γ is an independent set of $\mathcal{M}$, the matroid associated with Γ . Hence, the minterm M_1 is an independent set of size $d + t + m$ of $\mathcal{M}$, and thus, by Claim 4.7, every independent set of $\mathcal{M}$ that is smaller than $d + t + m$ can be expanded to an independent set of size $d + t + m$. Therefore, if $|M| < d + t + m$, the minterm M can be expanded to an independent set I of size $d + t + m$; otherwise, we set $I = M$. By Lemma 5.2, this expansion can be done only by adding to M users that precede $M_{\min}$. As M intersects W , the users in $I \setminus M$ are all from W . Hence, $|I \cap S| = |M \cap S| \leq m - 1$. Therefore, $|I \cap W| = |I| - |I \cap S| \geq d + t + m - (m - 1) = d + t + 1$. Next, we view M_1 as the canonical complement of the empty set (see Definition 5.4). Its $(d + t + 1)$ th element is z_1 . By Lemma 5.6 for $P = \emptyset$, $Y = M_1$, and $j = d + t + 1$, any $d + t + 1$ members of W form a dependent set. Hence I , which was assumed to be independent, contains a dependent set. This contradiction implies that $|M \cap S| \geq m$. $\square$

8.3.1. When all S -cooperative sets are of the same size. Here we show that if all S -cooperative sets are of the same size, namely $|Z| = m$, the set S is a strong set. We accomplish this by showing that all the S -cooperative sets of size m induce the same access structure on W , which is the access structure induced by the S -cooperative set Z . We begin by showing that the weight of every S -cooperative set is at least the weight of Z .

CLAIM 8.7. *Let $V \subseteq S$ be an S -cooperative set. Then $w(V) \geq w(Z)$.*

Proof. Assume towards a contradiction that $w(V) < w(Z)$, and consider the set $W \cup Z$. Since Z is S -cooperative, the set $W \cup Z$ is authorized. Thus, by Lemma 5.1, it must contain a suffix minterm of the form $B \cup Z$, where B is a suffix of W . There are two possible cases: either $B \cup V$ is authorized, or not.

If $B \cup V$ is authorized, then, since $w(V) < w(Z)$, the set $B \cup V$ is a minterm. Hence, as $B \cup V$ and $B \cup Z$ are two minterms that have the same minimal user, $B_{\min}$, Lemma 5.3 implies that $|V| = |Z| = m$. The set $B \cup V$ is independent in $\mathcal{M}$. If $|B \cup V| < d + t + m$, Claim 4.7 implies that $B \cup V$ can be expanded to an independent set I of size $d + t + m$; if $|B \cup V| \geq d + t + m$, we set $I = B \cup V$. By Lemma 5.2, all users in $I \setminus (B \cup V)$ must be from W . Hence, I includes at least $d + t$ users from W . On the other hand, since $|V| = |Z| = m$ and $w(V) < w(Z)$, there must be an index j such that $v_j \prec z_j$. Since M_1 is the canonical complement of the empty set $\emptyset$, we get from Lemma 5.6, applied to $P = \emptyset$, $Y = M_1$, and $B = I_{1,d+t+j}$, that the latter set is dependent. This is impossible since I is independent. Therefore, $B \cup V$ cannot be authorized.

If $B \cup V$ is unauthorized, we let Q be the canonical complement of B . Since $B \cup Z$ is a minterm, we get from Lemma 5.3 that $|Q| = |Z| = m$. On the other hand, by Claim 8.6, since V is S -cooperative, $|V| \geq m$. Since $B \cup V \notin \Gamma$ but $B \cup Q \in \Gamma$, we infer that $w(V) < w(Q)$. Therefore, there must be an index $j \in [m]$ such that $v_j \prec q_j$. Thus, by Lemma 5.6, we get that $B \cup V$ is dependent. On the other hand, since V is S -cooperative and B is a suffix of W , the set $B \cup V$ may be expanded to an authorized superset by adding to it users that precede $B_{\min}$, one by one, until the first time that we get an authorized set. This construction, where in each stage we add a new user that is smaller than all current users in the set, guarantees that we end up with a minterm. But a minterm of Γ cannot contain a dependent set. Therefore, this case is not possible either. We are lead to the conclusion that $w(V) \geq w(Z)$. $\square$

We are now ready to prove that all the cooperative sets of size m induce the same access structure on W .

CLAIM 8.8. *Let V be an S -cooperative set of size m . Then $\Gamma_{V,W} = \Gamma_{Z,W}$.*

Proof. By Claim 8.7, $w(V) \geq w(Z)$. If $w(V) = w(Z)$, the claim is trivial, since Γ is a WTAS. Therefore, we assume that $w(V) > w(Z)$. We first show that $U_{1,d} \cup Y \cup V$ is a minterm of Γ . Since $M_1 = U_{1,d} \cup Y \cup Z$ is authorized, the set $U_{1,d} \cup Y \cup V$ is authorized as well. Assume it is not a minterm. Then, by Lemma 5.1 it contains a suffix minterm of the form $B \cup V$, where B is a suffix of $U_{2,d} \cup Y$. Let Q be the canonical complement of B . Since $B \cup V$ is a minterm, by Lemma 5.3, it must be that $|Q| = |V| = m = |Z|$. Since $B \cup Q$ is authorized and $B \cup Z$ is unauthorized, $w(Z) < w(Q)$. Therefore, there must be an index $j \in [m]$ for which $z_j \prec q_j$. Thus, by Lemma 5.6, the set $B \cup Z$ is dependent. However, this set is contained in $U_{1,d} \cup Y \cup Z$, which is a minterm. This contradiction implies that $U_{1,d} \cup Y \cup V$ is a minterm of Γ . Consequently, since $U_{1,d} \cup Y \cup Z$ is a minterm and $U_{2,d} \cup Y \cup V$ is unauthorized (being a proper subset of a minterm), we infer that $w(Z) + w(u_1) > w(V)$.

We are now ready to prove that $\Gamma_{Z,W} = \Gamma_{V,W}$. Since we deal with the case where $w(Z) < w(V)$, the inclusion $\Gamma_{Z,W} \subseteq \Gamma_{V,W}$ is obvious. For the opposite inclusion, it is sufficient to concentrate on minterms of $\Gamma_{V,W}$. Let M be a minterm of $\Gamma_{V,W}$. Thus, $M \cup V \in \Gamma$, and since $M \cup V \setminus M_{\min} \notin \Gamma$, the set $M \cup V$ is a minterm in Γ . There are two possible cases: If $u_1 \in M$, the minterm $M \cup V$ must be of the same size as M_1 , by Lemma 5.3. Since $|M_1| = d + t + m$ and $|V| = m$, we get that $|M| = d + t$. As $M_1 = U_{1,d} \cup Y \cup Z$ is the minimal minterm in Γ in terms of the precedence order $\prec$, the weight of $U_{1,d} \cup Y$ is minimal among all sets of size $d + t$ that are contained in a minterm. This implies that $w(M) \geq w(U_{1,d} \cup Y)$. This, in turn, implies that $M \cup Z \in \Gamma$ and thus $M \in \Gamma_{Z,W}$.

The second case is when $u_1 \notin M$. Assume, towards a contradiction, that $M \cup Z \notin \Gamma$. Let Q be the canonical complement of M . By Lemma 5.3, $|Q| = |V| = m = |Z|$. Since $M \cup Z \notin \Gamma$ and $M \cup Q \in \Gamma$, $w(Z) < w(Q)$. Therefore, there must be an index $j \in [m]$ such that $z_j \prec q_j$. Thus, by Lemma 5.6, $M \cup Z$ is dependent. However, since $M \cup V \in \Gamma$ and $w(Z) + w(u_1) > w(V)$, we get that $\{u_1\} \cup M \cup Z \in \Gamma$. Moreover, it must be a minterm since any proper subset of $\{u_1\} \cup M \cup Z$ is of weight that does not exceed that of the unauthorized set $M \cup Z$. We have arrived at the absurd situation where the dependent set $M \cup Z$ is contained in a minterm. This contradiction implies that $M \cup Z$ is authorized, so $M \in \Gamma_{Z,W}$. $\square$

Combining Claims 8.6 and 8.8, we arrive at the following corollary.

COROLLARY 8.9. *If there are no S -cooperative sets of size larger than m , then S is a strong set.*

Example 8.10. Consider the set $U = \{u_1, \dots, u_8\}$, and let Γ be a WTAS where the weights are 1, 1, 1, 1, 1, 3, 3, and 3 and the threshold is 6. The lexicographically minimal minterm is $\{u_1, u_2, u_3, u_6\}$, and so there is no prefix minterm and no lacunary minterm. In this example $W = U_{1,5}$ and $S = U_{6,8}$, and the access structure is a composition of a 2-out-of-4 threshold access structure on $S \cup \{u'\}$ and a 3-out-of-5 threshold access structure on the weak side W , where u' is an additional dummy user.

8.3.2. When large S -cooperative sets exist. The conclusion from the previous section is that whenever all S -cooperative sets for $S = U_{\ell,n}$ are of the same size (i.e., $|Z| = m$), the set S is a strong set, and hence, by Theorem 8.4, the access structure Γ is a composition of ideal WTASs that are defined on two smaller sets. Here, we continue to deal with the case where there are S -cooperative sets of size larger than m . In that case we identify another strong set. Specifically, we show that $U_{d+2,n}$ is a strong set of users.

Recall that we assume that there are minterms that start with u_2 . In order to prove that $U_{d+2,n}$ is a strong set we show that for every $U_{d+2,n}$ -cooperative set, V , the access structure that it induces on $U_{1,d+1}$, namely $\Gamma_{V,U_{1,d+1}}$, is a d -out-of- $(d+1)$ threshold access structure. The proof has two stages:

- First we show that if M is a minterm of Γ such that $|M \cap S| > m$, then $M \cap U_{1,d+1} = \emptyset$ (Claim 8.15). This immediately implies that V , being a $U_{d+2,n}$ -cooperative set, intersects S in exactly m users (Corollary 8.16).
- In the second stage, we analyze the restricted access structure that Z induces on W , namely $\Gamma_{Z,W}$, and conclude that $\Gamma_{V,U_{1,d+1}}$ is a d -out-of- $(d+1)$ access structure.

We start with two technical claims. The following claim shows that the weight of z_1 is much larger than the weights of the users preceding it.

CLAIM 8.11. $w(u_{d+1}) + w(x_1) < w(z_1)$.

Proof. Suppose the claim is false and $w(u_{d+1}) + w(x_1) \geq w(z_1)$. Then we can replace z_1 by u_{d+1} and x_1 in the minterm M_1 and get an authorized set M'_1 . By Lemma 5.1, this set contains a suffix minterm. This minterm must intersect W , since Z by itself is unauthorized. But then M'_1 would be a minterm that contains only $m-1$ elements of S , in contradiction to Claim 8.6. $\square$

CLAIM 8.12. *The sets $U_{1,d} \cup Y \cup \{x_1\}$ and $U_{1,d+1}$ are circuits of $\mathcal{M}$ —the matroid that corresponds to Γ .*

Proof. We start with the set $V := U_{1,d} \cup Y \cup \{x_1\}$. Denote by Γ' the restriction $\Gamma_{Z,W \setminus \{u_{d+1}\}}$. By Lemma 4.11 and Claim 4.13, Γ' is an ideal WTAS. Moreover, the set $U_{1,d} \cup Y$ is a prefix minterm of this access structure. Therefore, by Theorem 6.7, Γ' is an HTAS. Thus, by Claim 6.2, V is a circuit of the matroid that corresponds to Γ' , and its size is $|V| = d + t + 1$. As M_1 is the canonic complement of the empty set, z_1 is its $(d+t+1)$ th element, and $x_1 \prec z_1$, we conclude by Lemma 5.6 that V is dependent in $\mathcal{M}$. It must be a circuit in $\mathcal{M}$, for otherwise, by Lemma 4.12, we would get a contradiction to the fact that it is a circuit in the matroid that corresponds to the restriction Γ' .

The proof for $U_{1,d+1}$ is similar. Let $\Gamma'' = \Gamma_{Y \cup Z, U_{1,d+1}}$ be the restriction that $Y \cup Z$ induces on $U_{1,d+1}$. It is an ideal WTAS, and $U_{1,d}$ is a prefix minterm in it. Thus, by Claim 6.2, $U_{1,d+1}$ is a circuit of the matroid that corresponds to Γ'' . As u_{d+1} , which is the $(d+1)$ th user in $U_{1,d+1}$, precedes the $(d+1)$ th user in the canonical complement of the empty set (which is y_1 , or z_1 if $Y = \emptyset$), we infer by Lemma 5.6 that $U_{1,d+1}$ is dependent in $\mathcal{M}$. Finally, by Lemma 4.12, the set $U_{1,d+1}$ must be a circuit of $\mathcal{M}$. $\square$

We now turn to showing that large S -cooperative sets are not contained in minterms that intersect $U_{1,d+1}$. This is done in two steps. First we show that if V is an S -cooperative set of size larger than m , there exists a suffix B of Y such that $B \cup V$ is a minterm. Then, relying upon this result, we show that no minterm of $\Gamma_{V,W}$ intersects $U_{1,d+1}$. This completes the first stage of the proof outlined above.

CLAIM 8.13. *If there exists an S -cooperative set V of size larger than m , there exists a suffix B of Y such that $B \cup V$ is a minterm.*

Proof. Consider the set $A := U_{1,d} \cup Y \cup V$. Since $|V| > |Z|$, the set A cannot be a minterm, as implied by Lemma 5.3. By Claim 8.7, $w(V) \geq w(Z)$. Thus, the set A is authorized and, consequently, it contains a suffix minterm of the form $B \cup V$, where B is a proper suffix of $U_{1,d} \cup Y$.

In order to prove the claim, we need to show that $B \subseteq Y$. Assume, towards a contradiction, that this is not true, namely, that $B_{\min} \in U_{1,d}$. Let $Q = B \setminus \{B_{\min}\} \cup \{x_1\}$. Since $B_{\min} \prec x_1$, the weight of Q is at least the weight of B , and therefore

$Q \cup V$ is an authorized set. Hence, it has a suffix minterm that must contain $\{x_1\} \cup V$. The union of this minterm and the minterm $B \cup V$ is $B \cup \{x_1\} \cup V$. We claim that all the users in V are critical for $B \cup \{x_1\} \cup V$, in the sense of Definition 4.3. Let $v_i \in V$, and assume that it is not critical; i.e., $B \cup \{x_1\} \cup V \setminus \{v_i\} \in \Gamma$. Since $B \cup V \setminus \{B_{\min}\} \notin \Gamma$ (because $B \cup V$ is a minterm), we conclude that $w(B \cup V \setminus \{B_{\min}\}) < w(B \cup \{x_1\} \cup V \setminus \{v_i\})$, whence $w(v_i) < w(B_{\min}) + w(x_1)$. Therefore, as $w(z_1) \leq w(v_i)$ and $w(B_{\min}) \leq w(u_{d+1})$, we arrive at the conclusion that $w(z_1) < w(u_{d+1}) + w(x_1)$, in contradiction to Claim 8.11. This implies that all of the users of V are critical for $B \cup \{x_1\} \cup V$. Thus, by Corollary 4.5, the set $B \cup \{x_1\}$ is dependent. However, $B \cup \{x_1\}$ is properly contained in $U_{1,d} \cup Y \cup \{x_1\}$, which, by Claim 8.12, is a circuit of $\mathcal{M}$. This contradiction implies that B does not intersect $U_{1,d}$ and is therefore a suffix of Y . $\square$

Remark 8.14. If there is an S -cooperative set V of size larger than m , the set Y is not empty. Otherwise, the set B in Claim 8.13 would be empty, whence V is a minterm, in contradiction to our assumption that V is S -cooperative. Therefore, in the rest of this section we assume that Y is not empty.

CLAIM 8.15. *If V is an S -cooperative set of size larger than m and A is a minterm of $\Gamma_{V,W}$, then $A \cap U_{1,d+1} = \emptyset$.*

Proof. Suppose that there is a minterm $A \in \Gamma_{V,W}$ such that $A \cap U_{1,d+1} \neq \emptyset$. We pick a minterm A of that sort having the property that $A \cap (Y \cup X) \preceq M \cap (Y \cup X)$ for all minterms $M \in \Gamma_{V,W}$ such that $M \cap U_{1,d+1} \neq \emptyset$. By Claim 8.13, there is a suffix B of Y such that $B \in \Gamma_{V,W}$. Therefore, as A is a minterm of $\Gamma_{V,W}$, it must be that $Y \not\subseteq A$ (otherwise, if $Y \subseteq A$, the minterm $A \cup V$ would have been a proper superset of the minterm $B \cup V$ since A includes in addition users from $U_{1,d+1}$). Let j be the smallest index such that $y_j \notin A$ (recall that $Y \neq \emptyset$), and consider the set $Q = A \setminus \{A_{\min}\} \cup \{y_j\}$. Since $A_{\min} \in U_{1,d+1}$, we conclude that $w(A_{\min}) \leq w(y_j)$. Thus $Q \cup V$ is authorized in Γ and must contain a suffix minterm M' that contains y_j . The union of $A \cup V$ and M' is $A \cup \{y_j\} \cup V$. The discussion is now separated into two cases: either A is a run of U or it is not.

If A is a run of U , then we first claim that every user of V is critical for $A \cup \{y_j\} \cup V$. Let $v_i \in V$, and assume that it is not critical for $A \cup \{y_j\} \cup V$; i.e., $A \cup \{y_j\} \cup V \setminus \{v_i\} \in \Gamma$. Since $A \cup V \setminus \{A_{\min}\} \notin \Gamma$ (as $A \cup V$ is a minterm), we conclude, by comparing the weights of the two latter sets, that $w(v_i) < w(A_{\min}) + w(y_j)$. However, since $w(z_1) \leq w(v_i)$, $w(y_j) \leq w(x_1)$, and $w(A_{\min}) \leq w(u_{d+1})$, we get that $w(z_1) < w(u_{d+1}) + w(x_1)$, in contradiction to Claim 8.11. Therefore, all users in V are critical for $A \cup \{y_j\} \cup V$. Thus, by Corollary 4.5, the set $A \cup \{y_j\}$ is dependent. Since we are now dealing with the case where A is a run of U , the set $A \cup \{y_j\}$ must be contained in $U_{1,d+1} \cup Y$. Since $A \cup \{y_j\}$ is dependent, it includes a circuit C . That circuit must include the user u_{d+1} , for otherwise C would have been a subset of the independent set $U_{1,d} \cup Y$. Moreover, the circuit C must also contain y_j , for otherwise $C \subseteq A$, while A is contained in a minterm and thus is independent. On the other hand, by Claim 8.12, the set $U_{1,d+1}$ is a circuit of $\mathcal{M}$ that does not contain y_j . Therefore, u_{d+1} is in the intersection of the two distinct circuits $U_{1,d+1}$ and C . Hence, by Lemma 4.6, the set $U_{1,d+1} \cup C \setminus \{u_{d+1}\}$ is dependent. But that latter set is contained in the independent set $U_{1,d} \cup Y$. This contradiction settles the statement whenever A is a run of U .

If A is not a run of U , we claim that every user $a_i \in A$ such that $y_j \prec a_i$ is critical for $A \cup \{y_j\} \cup V$. Otherwise we could replace the user a_i in the minterm $A \cup V$ with the user y_j and still get a minterm that intersects $U_{1,d+1}$. But this would contradict our choice of A as a minterm that is lexicographically minimal in $Y \cup X$. This implies that all users of $A \cup \{y_j\} \cup V$ that have weight greater than or equal to y_j are critical

for that set. In particular, the subset $(A \cap X) \cup V$ consists of only critical users, and consequently, by Claim 4.5, the set $A \setminus X \cup \{y_j\}$ is dependent. Since A is contained in a minterm, this dependent set must contain a circuit C that contains y_j . Since $U_{1,d} \cup Y$ is an independent set of $\mathcal{M}$, it must be that $u_{d+1} \in C$; otherwise C is properly contained in an independent set. However, since $U_{1,d+1}$ is a circuit of $\mathcal{M}$, we get that u_{d+1} is in the intersection of two distinct circuits, C and $U_{1,d+1}$ (note that $y_j \in C$ but $y_j \notin U_{1,d+1}$). Therefore, by Lemma 4.6, the set $C \cup U_{1,d+1} \setminus \{u_{d+1}\}$ is dependent. However, this set is properly contained in $U_{1,d} \cup Y$, which is an independent set of $\mathcal{M}$, and thus we arrive at a similar contradiction as in the previous case. $\square$

COROLLARY 8.16. *Every $U_{d+2,n}$ -cooperative set intersects S in exactly m users.*

Proof. Let B be a $U_{d+2,n}$ -cooperative set. Then there exists a nonempty set $A \subseteq U_{1,d+1}$ such that $M := A \cup B$ is a minterm of Γ . Denote $M_W = M \cap W$ and $M_S = M \cap S$, so that $M = A \cup B = M_W \cup M_S$. As $U_{1,d+1} \subset W$ and $U_{d+2,n} \supset S$, we infer that $M_W \cap U_{1,d+1} = A \neq \emptyset$ and $M_S = B \cap S$.

On the one hand, Claim 8.6 implies that $|M_S| \geq m$, since $M_W \neq \emptyset$. On the other hand, by Claim 8.15, if $|M_S| > m$, then $M_W \cap U_{1,d+1} = \emptyset$. But as $M_W \cap U_{1,d+1} \neq \emptyset$, we conclude that $|M_S| = |B \cap S| = m$. $\square$

We now turn to the second stage of showing that $U_{d+2,n}$ is a strong set. We will show that the existence of large S -cooperative sets and the existence of minterms that start with u_2 imply that $U_{d+2,n}$ is a strong set. By Claim 8.15, every $U_{d+2,n}$ -cooperative set intersects S in exactly m users. By Claim 8.8, if V is an S -cooperative set of size m , then $\Gamma_{V,W} = \Gamma_{Z,W}$. Therefore, it is enough to study the minterms in $\Gamma_{Z,W}$ that intersect $U_{1,d+1}$. Somehow, surprisingly, the existence of a large S -cooperative set affects the structure of $\Gamma_{Z,W}$. The next two claims will serve us in characterizing the structure of $\Gamma_{Z,W}$ in Claim 8.19.

CLAIM 8.17. *If there is an S -cooperative set V of size larger than m , there exists a nonempty suffix B of Y such that $B \cup \{x_1\} \in \Gamma_{Z,W}$.*

Proof. By Claim 8.13, there exists a suffix B of Y such that $B \cup V$ is a minterm of Γ . We next prove that $B \cup \{x_1\} \cup V \setminus \{v_1\} \in \Gamma$. Replacing $B_{\min}$ with x_1 in $B \cup V$, we get an authorized set. Denote by M' the suffix minterm contained in that set, and note that the union of M' and $B \cup V$ is exactly $B \cup \{x_1\} \cup V$. We claim that v_1 cannot be critical for $B \cup \{x_1\} \cup V$. If it was critical, then every user in V would also be critical for that set, and thus, by Corollary 4.5, the set $B \cup \{x_1\}$ would be dependent in $\mathcal{M}$. However, the set $B \cup \{x_1\}$ is properly contained in $U_{1,d} \cup Y \cup \{x_1\}$, which is a circuit by Claim 8.12. Therefore, $B \cup \{x_1\}$ is independent, and thus v_1 cannot be critical for $B \cup \{x_1\} \cup V$. Hence, the set $B \cup \{x_1\} \cup V \setminus \{v_1\}$ is authorized. But since it is obtained from the minterm $B \cup V$ by replacing v_1 with $x_1 \prec v_1$, it must be also a minterm.

We now prove the claim by induction on the size of V . If $|V| = m + 1$, then since $B \cup \{x_1\} \cup V \setminus \{v_1\}$ is a minterm, we conclude that $B \cup \{x_1\}$ is in $\Gamma_{V',W}$, where $V' = V \setminus \{v_1\}$ is an S -cooperative set of size m . But since, in view of Claim 8.8, $\Gamma_{V',W} = \Gamma_{Z,W}$, we arrive at the sought-after conclusion that $B \cup \{x_1\} \in \Gamma_{Z,W}$. If $|V| > m + 1$, then V' is an S -cooperative set of size $|V| - 1$. Hence, we may apply the induction hypothesis to conclude that there exists a suffix B of Y such that $B \cup \{x_1\} \in \Gamma_{Z,W}$. $\square$

The other assumption that affects the structure of $\Gamma_{Z,W}$ is that there are minterms of Γ that start with u_2 .

CLAIM 8.18. *If u_2 is the minimal user in some minterm of Γ , then $U_{2,d+1} \cup Y$ is a minterm of $\Gamma_{Z,W}$.*

Proof. The access structure $\Gamma_{Z,W}$ is an ideal WTAS and its lexicographically

minimal minterm is $U_{1,d} \cup Y$. Let M be a minterm of Γ that starts with u_2 . In particular, M intersects W . By Claims 8.6 and 8.15, the minterm M contains exactly m members of S . Thus, By Claim 8.8, $\Gamma_{M \cap S, W} = \Gamma_{Z, W}$. This implies that $M \cap W$ is a minterm of $\Gamma_{Z, W}$ that starts with u_2 . Recall that, by Remark 8.14, $Y \neq \emptyset$. Hence, Claim 7.1 applies to the access structure $\Gamma_{Z, W}$, and the set $U_{2,d+1} \cup Y$ is a minterm of $\Gamma_{Z, W}$. $\square$

We may now describe the structure of $\Gamma_{Z, W}$. Since we assume that there exists a minterm of Γ that starts with u_2 , the set $U_{2,d+1} \cup Y$ is a minterm of $\Gamma_{Z, W}$, as implied by Claim 8.18. As $U_{2,d+1} \cup Y$ is a run, and it is a prefix in $W \setminus \{u_1\}$, we get that the restriction $\Gamma_{Z, W \setminus \{u_1\}}$ has a prefix minterm. Therefore, by Theorem 6.7, it is an HTAS. The threshold of the first level is $|U_{2,d+1} \cup Y| = d + t$. By Claim 8.17, the run minterm of this HTAS that ends with x_1 starts with an element of Y . Therefore, the transition between the first and the second level in that HTAS (if it exists) occurs within Y . This enables us to prove that $U_{d+2,n}$ is a strong set of users.

CLAIM 8.19. *Suppose that there is an S -cooperative set of size larger than m , and that there is a minterm of Γ that starts with u_2 . Then $U_{d+2,n}$ is a strong set of users.*

Proof. In order to show that $U_{d+2,n}$ is a strong set of users, we show that for any $U_{d+2,n}$ -cooperative set V , the corresponding restriction induced by V on $U_{1,d+1}$, namely $\Gamma_{V, U_{1,d+1}}$, is a d -out-of- $(d+1)$ threshold access structure. Let V be a $U_{d+2,n}$ -cooperative set. Hence, there exists a nonempty set $A \subseteq U_{1,d+1}$ such that $A \cup V$ is a minterm of Γ . Denote $V_W = V \cap W$ and $V_S = V \cap S$. Note that the set V_S is S -cooperative. Moreover, since $A \cup V$ is a minterm of Γ , then $A \cup V_W$ is a minterm of $\Gamma_{V_S, W}$ that intersects $U_{1,d+1}$. Therefore, by Claim 8.15, we conclude that $|V_S| = m$. Thus, by Claim 8.8, $\Gamma_{V_S, W} = \Gamma_{Z, W}$. Consequently, since $A \cup V_W$ is a minterm of $\Gamma_{V_S, W}$, it is also a minterm of $\Gamma_{Z, W}$.

We claim that the size of that minterm is $|A \cup V_W| = d + t$. Indeed, if the minimal user in $A \cup V_W$ is u_1 , then, by Lemma 5.3, $|A \cup V_W \cup V_S| = |M_1| = d + t + m$, and so, as $|V_S| = m$, we get that $|A \cup V_W| = d + t$. Otherwise, $A \cup V_W \subseteq W \setminus \{u_1\}$, and consequently, $A \cup V_W \in \Gamma_{Z, W \setminus \{u_1\}}$. But the latter access structure was shown above to be an HTAS; furthermore, we showed that the first level in that HTAS includes all of $U_{2,d+1}$ and that the corresponding threshold is $d + t$. Since $A \cup V_W$ starts within $U_{2,d+1}$, it is of size $d + t$.

First, we show that for any $D \subseteq U_{1,d+1}$ of size d , $D \cup V \in \Gamma$. Since $w(D) \geq w(U_{1,d})$, we need only to show that $U_{1,d} \cup V \in \Gamma$. We know that $M_1 = U_{1,d} \cup Y \cup Z \in \Gamma$. Since $A \subseteq U_{1,d+1}$ it must be that $|A| \leq d$ (otherwise $A = U_{1,d+1}$, and then $A \cup V$ would be a minterm that precedes M_1 with respect to the lexicographical order). Hence, $|V_W| \geq t$. Since Y consists of the first t users in $U_{d+2,n}$, we conclude that $w(V_W) \geq w(Y)$. Therefore V_W can replace Y in M_1 , and so $U_{1,d} \cup V_W \in \Gamma_{Z, W} = \Gamma_{V_S, W}$. This implies that $U_{1,d} \cup V_W \cup V_S = U_{1,d} \cup V \in \Gamma$, as required. Hence, every subset of d users from $U_{1,d+1}$ completes V to an authorized set.

Finally, we need to show that no smaller subset of $U_{1,d+1}$ completes V to an authorized set. To that end, let B be a minterm of $\Gamma_{V, U_{1,d+1}}$. We proceed to show that $|B| \geq d$. Using the above notation, $B \cup V_W \cup V_S$ is a minterm of Γ , or, equivalently, $B \cup V_W$ is a minterm of $\Gamma_{V_S, W} = \Gamma_{Z, W}$. Arguing along the same lines as we did for $A \cup V_W$, we conclude that $|B \cup V_W| = d + t$. Assume towards a contradiction that $|B| < d$. Then $|V_W| > t$, and consequently, as $V_W \subset U_{d+2,n}$, we get that $w(V_W) \geq w(Y \cup \{x_1\})$. But according to Claim 8.17, $Y \cup \{x_1\} \in \Gamma_{Z, W}$. Hence, since $\Gamma_{Z, W} = \Gamma_{V_S, W}$, we infer that $Y \cup \{x_1\} \cup V_S \in \Gamma$. This implies that also $V_W \cup V_S = V$

is authorized, in a contradiction to our assumption that V is a $U_{d+2,n}$ -cooperative set. That completes the proof. $\square$

8.4. Proof of Theorem 3.5—The characterization theorem. Let Γ be an ideal WTAS defined on a set of users U , and let M_1 be its lexicographically minimal minterm. If either Γ has self-sufficient users or u_2 starts no minterm of Γ , then, by Lemma 8.5, the access structure Γ is a composition of two ideal WTASs on smaller sets of users.

If M_1 is a prefix, then, by Theorem 6.7, the access structure Γ is an HTAS. If M_1 is a lacunary prefix, namely, $M_1 = U_{1,d} \cup U_{d+2,k}$ for some $1 \leq d \leq k-2$ and $k \leq n$, then, by Theorem 7.6, the access structure Γ is a TPAS. Otherwise, by Corollary 8.9 and Claim 8.19, there exists within U a subset of strong users. In this case, we conclude by Theorem 8.4 that the access structure Γ is a composition of two ideal WTASs that are defined on sets smaller than U .

As for the other direction, HTASs are ideal and may be realized by linear secret sharing schemes, as shown in [6, 35]. TPASs are also ideal and may be realized by linear secret sharing schemes, as shown here in section 7.1. Finally, given two ideal access structures, we showed in Lemma 8.1 how to construct an ideal secret sharing scheme for their composition. Hence, the composition is also ideal. Furthermore, by Lemma 8.1, if the secret sharing schemes for the two basic access structures are linear, so is the resulting scheme for the composition of the two access structures. This completes the proof of the characterization theorem. $\square$

REFERENCES

- [1] A. BEIMEL AND B. CHOR, *Universally ideal secret sharing schemes*, IEEE Trans. Inform. Theory, 40 (1994), pp. 786–794.
- [2] A. BEIMEL AND E. WEINREB, *Monotone circuits for monotone weighted threshold functions*, Inform. Process. Lett., 97 (2006), pp. 12–18.
- [3] M. BEN-OR, S. GOLDWASSER, AND A. WIGDERSON, *Completeness theorems for noncryptographic fault-tolerant distributed computations*, in Proceedings of the 20th ACM Symposium on the Theory of Computing, Chicago, IL, 1988, ACM, New York, pp. 1–10.
- [4] J. BENALOH AND J. LEICHTER, *Generalized secret sharing and monotone functions*, in Advances in Cryptology—CRYPTO ’88, S. Goldwasser, ed., Lecture Notes in Comput. Sci. 403, Springer-Verlag, New York, 1990, pp. 27–35.
- [5] G. R. BLAKLEY, *Safeguarding cryptographic keys*, in Proceedings of the 1979 AFIPS National Computer Conference, R. E. Merwin, J. T. Zanca, and M. Smith, eds., vol. 48 of AFIPS Conference proceedings, AFIPS (American Federation of Information Processing Societies) Press, 1979, pp. 313–317.
- [6] E. F. BRICKELL, *Some ideal secret sharing schemes*, J. Combin. Math. Combin. Comput., 6 (1989), pp. 105–113.
- [7] E. F. BRICKELL AND D. M. DAVENPORT, *On the classification of ideal secret sharing schemes*, J. Cryptology, 4 (1991), pp. 123–134.
- [8] E. F. BRICKELL AND D. R. STINSON, *Some improved bounds on the information rate of perfect secret sharing schemes*, J. Cryptology, 5 (1992), pp. 153–166.
- [9] R. M. CAPOCELLI, A. DE SANTIS, L. GARGANO, AND U. VACCARO, *On the size of shares for secret sharing schemes*, J. Cryptology, 6 (1993), pp. 157–168.
- [10] D. CHAUM, C. CRÉPEAU, AND I. DAMGÅRD, *Multiparty unconditionally secure protocols*, in Proceedings of the 20th ACM Symposium on the Theory of Computing, Chicago, IL, 1988, ACM, New York, pp. 11–19.
- [11] M. J. COLLINS, *A Note on Ideal Tripartite Access Structures*, Technical Report 2002/193, Cryptology ePrint Archive, 2002, online at <http://eprint.iacr.org/>.
- [12] R. CRAMER, I. DAMGÅRD, AND U. MAURER, *General secure multi-party computation from any linear secret-sharing scheme*, in Advances in Cryptology—EUROCRYPT 2000, B. Preneel, ed., Lecture Notes in Comput. Sci. 1807, Springer-Verlag, New York, 2000, pp. 316–334.

- [13] Y. DESMEDT AND Y. FRANKEL, *Shared generation of authenticators and signatures*, in Advances in Cryptology—CRYPTO '91, J. Feigenbaum, ed., Lecture Notes in Comput. Sci. 576, Springer-Verlag, New York, 1992, pp. 457–469.
- [14] O. FARRÀS, J. MARTÍ-FARRÉ, AND C. PADRÓ, *Ideal multipartite secret sharing schemes*, in Advances in Cryptology—EUROCRYPT 2007, M. Naor, ed., Lecture Notes in Comput. Sci. 4515, Springer-Verlag, New York, 2007, pp. 448–465.
- [15] J. HERRANZ AND G. SÁEZ, *New Results on Multipartite Access Structures*, Technical Report 2006/048, Cryptology ePrint Archive, 2006, online at <http://eprint.iacr.org/>.
- [16] M. ITO, A. SAITO, AND T. NISHIZEKI, *Secret sharing schemes realizing general access structure*, in Proceedings of the IEEE Global Telecommunication Conference, Globecom 87, 1987, IEEE Press, Piscataway, NJ, pp. 99–102.
- [17] W. JACKSON, K. M. MARTIN, AND C. M. O'KEEFE, *Ideal secret sharing schemes with multiple secrets*, J. Cryptology, 9 (1996), pp. 233–250.
- [18] M. KARCHMER AND A. WIGDERSON, *On span programs*, in Proceedings of the 8th IEEE Conference on Structure in Complexity Theory, Boston, 1993, IEEE Press, Piscataway, NJ, pp. 102–111.
- [19] E. D. KARNIN, J. W. GREENE, AND M. E. HELLMAN, *On secret sharing systems*, IEEE Trans. Inform. Theory, 29 (1983), pp. 35–41.
- [20] J. MARTÍ-FARRÉ AND C. PADRÓ, *Secret sharing schemes on access structures with intersection number equal to one*, Discrete Appl. Math., 154 (2006), pp. 552–563.
- [21] K. M. MARTIN, *Discrete Structures in the Theory of Secret Sharing*, Ph.D. thesis, Applied Mathematics and Computation, University of London, London, 1991.
- [22] K. M. MARTIN, *New secret sharing schemes from old*, J. Combin. Math. Combin. Comput., 14 (1993), pp. 65–77.
- [23] P. MORILLO, C. PADRÓ, G. SÁEZ, AND J. L. VILLAR, *Weighted threshold secret sharing schemes*, Inform. Process. Lett., 70 (1999), pp. 211–216.
- [24] J. G. OXLEY, *Matroid Theory*, Oxford University Press, London, 1992.
- [25] C. PADRÓ AND G. SÁEZ, *Secret sharing schemes with bipartite access structure*, IEEE Trans. Inform. Theory, 46 (2000), pp. 2596–2605.
- [26] M. O. RABIN, *Randomized Byzantine generals*, in Proceedings of the 24th IEEE Symposium on Foundations of Computer Science, Tucson, AZ, 1983, IEEE Press, Piscataway, NJ, pp. 403–409.
- [27] P. D. SEYMOUR, *On secret-sharing matroids*, J. Combin. Theory Ser. B, 56 (1992), pp. 69–73.
- [28] A. SHAMIR, *How to share a secret*, Comm. ACM, 22 (1979), pp. 612–613.
- [29] G. J. SIMMONS, *How to (really) share a secret*, in Advances in Cryptology—CRYPTO '88, S. Goldwasser, ed., Lecture Notes in Comput. Sci. 403, Springer-Verlag, New York, 1990, pp. 390–448.
- [30] G. J. SIMMONS, *An introduction to shared secret and/or shared control and their application*, in Contemporary Cryptology, The Science of Information Integrity, G. J. Simmons, ed., IEEE Press, Piscataway, NJ, 1992, pp. 441–497.
- [31] G. J. SIMMONS, W. JACKSON, AND K. M. MARTIN, *The geometry of shared secret schemes*, Bulletin of the Institute of Combinatorics and Its Applications, 1 (1991), pp. 71–88.
- [32] J. SIMONIS AND A. ASHIKHMIN, *Almost affine codes*, Designs, Codes and Cryptography, 14 (1998), pp. 179–197.
- [33] D. R. STINSON, *An explication of secret sharing schemes*, Designs, Codes and Cryptography, 2 (1992), pp. 357–390.
- [34] D. R. STINSON, *New general lower bounds on the information rate of secret sharing schemes*, in Advances in Cryptology—CRYPTO '92, E. F. Brickell, ed., Lecture Notes in Comput. Sci. 740, Springer-Verlag, New York, 1993, pp. 168–182.
- [35] T. TASSA, *Hierarchical threshold secret sharing*, J. Cryptology, 20 (2007), pp. 237–264.
- [36] T. TASSA AND N. DYN, *Multipartite secret sharing by bivariate interpolation.*, in Proceedings of the 33rd International Colloquium on Automata, Languages and Programming, M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener, eds., Lecture Notes in Comput. Sci. 4052, Springer-Verlag, New York, 2006, pp. 288–299.

ADAPTED LIST COLORING OF GRAPHS AND HYPERGRAPHS*

A. V. KOSTOCHKA[†] AND XUDING ZHU[‡]

Abstract. We introduce and study *adapted list coloring* of graphs and hypergraphs. This is a generalization of ordinary list coloring and adapted coloring, and has more applications than these. We prove that the upper bounds on the adaptable choosability of graphs and uniform hypergraphs in terms of maximum degree are sufficiently stronger than those on the ordinary choosability, while the bounds in terms of degeneracy are the same. We also characterize simple graphs with adaptable choosability 2.

Key words. adapted coloring, list coloring, degree conditions

AMS subject classification. 05C15

DOI. 10.1137/070698385

1. Introduction. Hell and Zhu [6] introduced and studied the notion of *adapted coloring* of graphs.

DEFINITION 1.1. Let H be a multihypergraph (i.e., parallel hyperedges are allowed), and let $F : E(H) \rightarrow \{1, 2, \dots, k\}$ be an edge k -coloring of H (not required to be proper). A vertex k -coloring $c : V \rightarrow \{1, 2, \dots, k\}$ of H is adapted to F if for every color j , the preimage $c^{-1}(j)$ does not contain any hyperedge of color j . In other words, if an edge e of H has color j , then it cannot have all its vertices colored with j .

While the multiplicity of edges does not matter in the ordinary vertex coloring, it does in the adapted coloring. If every (hyper)edge of G has multiplicity k and F colors these edges with all distinct k colors, then any vertex k -coloring of G adapted to this edge coloring is an ordinary proper k -coloring. But choosing different edge colorings yields different problems that have more applications than the ordinary vertex coloring. In particular, adaptable vertex coloring has applications to matrix partitions of graphs, trigraph homomorphisms, and full constraint satisfaction problems [3, 4, 5, 7].

The adaptable chromatic number is defined in a natural way.

DEFINITION 1.2. Let G be a multihypergraph. The adaptable chromatic number $\chi_{ad}(G)$ of G is the minimum integer k such that for every edge k -coloring F of G , there is a vertex k -coloring of G adapted to F .

Hell and Zhu [6] proved that a nonempty graph G has $\chi_{ad}(G) \leq 2$ if and only if G has an edge e such that $G - e$ is bipartite. Thus graphs with $\chi_{ad}(G) \leq 2$ can be determined in polynomial time. However, for $k \geq 3$, it is NP-hard to determine whether a graph G has $\chi_{ad}(G) \leq k$. It was also shown in [6] that for any positive integers k and g , there are graphs G of girth at least g such that $\chi_{ad}(G) = \chi(G) = k$.

In this paper, we introduce and study a logical generalization of the adapted coloring—*adapted list coloring*. We will show a rather general application and derive

*Received by the editors July 25, 2007; accepted for publication (in revised form) November 9, 2007; published electronically March 7, 2008.

<http://www.siam.org/journals/sidma/22-1/69838.html>

[†]Department of Mathematics, University of Illinois, Urbana, IL, 61801 and Institute of Mathematics, Novosibirsk, 630090, Russia (kostochk@math.uiuc.edu). This author's research was supported in part by NSF grant DMS-06-50784 and by grant 05-01-00816 of the Russian Foundation for Basic Research.

[‡]Department of Applied Mathematics, National Sun Yat-Sen University, Kaohsiung, Taiwan and the National Center for Theoretical Sciences (zhu@math.nsysu.edu.tw). This author's research was partially supported by the National Science Council under grant NSC95-2115-M-110-013-MY3.

several properties of such coloring.

DEFINITION 1.3. *Let H be a multihypergraph and $F : E(H) \rightarrow \mathbf{N}$ be any edge coloring of H (not required to be proper). Let a list assignment $L : V(H) \rightarrow 2^{\mathbf{N}}$ be a function that assigns a list $L(v) \subset \mathbf{N}$ to every vertex v of H . A vertex coloring c of $V(H)$ is an L -coloring adapted to F if for every $v \in V(H)$, $c(v) \in L(v)$, and for every color j , the preimage $c^{-1}(j)$ does not contain any hyperedge of color j . An adapted list coloring of H is an L -coloring of H adapted to F for a list L and an edge coloring F .*

Adapted list coloring can model the following rather general problem. There are jobs $J_1, \dots, J_n$ to be made and people (or machines) $P_1, \dots, P_m$ who can perform these jobs. Some people can do several jobs, but everybody has his/her own limits. For every job J_i , there is the list $L(i)$ of people who can perform this job. For every person P_l , there are inclusion minimal sets $e_{l,1}, \dots, e_{l,k_l}$ such that P_l cannot perform all jobs in the set $e_{l,i}$ at the same time. The question is whether we can assign to each P_l a set $c(l)$ of jobs not containing any of the sets $e_{l,1}, \dots, e_{l,k_l}$ so that each job is assigned to somebody who can perform it. This is exactly a list adapted coloring question: Let H be the hypergraph with $V(H) = \{J_1, \dots, J_n\}$ and $E(H) = \bigcup_{l=1}^m \{e_{l,1}, \dots, e_{l,k_l}\}$. The people are the colors, and each vertex J_i is given a list $L(i)$ of admissible colors. Let F be the edge coloring of H such that $F(e_{l,j}) = l$ for every $l = 1, \dots, m$ and $j = 1, \dots, k_l$. By definition, H has a coloring c adapted to F with $c(J_i) \in L(i)$ exactly when there is a proper job assignment in the original problem.

There are other natural applications of adapted list coloring, but in this paper we will concentrate on the properties of this coloring.

DEFINITION 1.4. *We say that a (multi)hypergraph G is adaptably k -choosable if for every edge coloring F and for every list assignment L with $|L(v)| = k$ for all v , there exists an L -coloring of G adapted to F . The minimum integer k , for which G is adaptably k -choosable, is called the adaptable choosability of G and is denoted by $\text{ch}_{ad}(G)$.*

It follows from the definition that if G has an L -coloring adapted to F , then every subgraph H of G also has an L -coloring adapted to F . Let $\text{ch}(G)$ denote the (ordinary) choosability of G . Then we have $\chi_{ad}(G) \leq \text{ch}_{ad}(G) \leq \text{ch}(G)$.

One may expect that for most graphs, the adaptable choosability is significantly less than the ordinary choosability. This is indeed the case. Moreover, for every multigraph G , $\text{ch}_{ad}(G) \leq \lceil \sqrt{8\Delta(G)} \rceil$. For large $\Delta(G)$ this bound is much better than the Brooks' bound for ordinary coloring. However, the situation with degeneracy is quite different. Our first main result says that there are many d -degenerate graphs G with $\chi_{ad}(G) = d + 1$. Every such graph G must have $\chi(G) = \text{ch}(G) = \text{ch}_{ad}(G) = \chi_{ad}(G)$. Furthermore, all this holds for hypergraphs.

THEOREM 1.5. *For any integers $g \geq 3$, $r \geq 2$, and $d \geq 1$, there is an r -uniform hypergraph H of girth at least g that is d -degenerate and $\text{ch}_{ad}(H) = \chi_{ad}(H) = d + 1$.*

Graphs G with $\text{ch}(G) \leq 2$ were characterized by Erdős, Rubin, and Taylor [2]. Graphs G with $\chi_{ad}(G) \leq 2$ were characterized by Hell and Zhu [6]. Our second main result is a characterization of adaptably 2-choosable simple graphs. Of course, it is enough to describe adaptably 2-choosable multigraphs without vertices of degree 0 and 1.

THEOREM 1.6. *A simple connected graph G with minimum degree at least 2 is adaptably 2-choosable if and only if G consists of two or three internally disjoint paths connecting two distinct vertices of G .*

Theorem 1.6 yields a polynomial time algorithm that decides whether a given graph is adaptably 2-choosable.

We use standard notation. For a vertex x of a multihypergraph G , $E_G(x)$ (or $E(x)$ if G is clear from the context) is the set of hyperedges of G containing x , and $d(x) = |E(x)|$ is the degree of x . By $\Delta(G)$ we denote the maximum degree of G .

For brevity, we will often say “a (k, F) -coloring” instead of “a vertex k -coloring adapted to F ” and “an (L, F) -coloring” instead of “an L -coloring adapted to F .”

In the next section we give upper bounds on adaptable choosability in terms of maximum degree, in section 3 we construct many d -degenerate uniform hypergraphs with adaptable choosability $d + 1$, and in the last section we characterize adaptably 2-choosable graphs.

2. Adaptable choosability and maximum degree. Erdős and Lovász used their famous Local Lemma to prove the following bound.

THEOREM 2.1 (see [1, Theorem 2]). *A $(k + 1)$ -chromatic r -uniform hypergraph contains an edge which is intersected by at least $k^{r-1}/4$ other edges. Thus, the degree of at least one vertex is $> k^{r-1}/4r$.*

Practically repeating their proof word by word for adaptable list coloring, one gets the following result.

PROPOSITION 2.2. *If G is an r -uniform (multi)hypergraph and $\Delta(G) \leq k^r/4r$, then $\text{ch}_{ad}(G) \leq k$. In particular, $\text{ch}_{ad}(G) \leq \lceil \sqrt{8\Delta(G)} \rceil$ for every (multi)graph G .*

For every fixed r , the order of magnitude of this bound is close to the best possible, for example, for complete r -uniform hypergraphs, $K_n^{(r)}$.

PROPOSITION 2.3. *If n is large in comparison with r and*

$$(1) \quad \Delta(K_n^{(r)}) \geq 2r k^r \ln k,$$

then $\text{ch}_{ad}(K_n^{(r)}) \geq \chi_{ad}(K_n^{(r)}) > k$. In particular, $\text{ch}_{ad}(K_n) \geq \chi_{ad}(K_n) > 0.5\sqrt{n/\ln n}$.

Proof. Let $\Delta = \Delta(K_n^{(r)})$. By definition,

$$(2) \quad \Delta = \binom{n-1}{r-1} \leq n^{r-1}/(r-1)!.$$

Let k be the largest integer satisfying (1), and let $m = \lceil n/k \rceil$. Consider a random k -coloring F of the edges of $K_n^{(r)}$, where each of the edges is colored uniformly and independently from all the other edges with one of the colors $1, \dots, k$. Let G_i be the hypergraph formed by the edges of color i , $i = 1, \dots, k$. We will prove that

(3) *with positive probability, the independence number of every G_i is less than m .*

That would imply that for some edge k -coloring F of $K_n^{(r)}$, in any vertex k -coloring of $K_n^{(r)}$ adapted to F , each color can appear at most $m - 1$ times, and hence $K_n^{(r)}$ simply has no such coloring. So, we are now proving (3).

For each i , the probability that G_i has an independent set of size m is at most $\binom{n}{m}(1 - \frac{1}{k})^{\binom{m}{r}}$. Hence, (3) will hold if

$$(4) \quad \binom{n}{m} \left(1 - \frac{1}{k}\right)^{\binom{m}{r}} < \frac{1}{k}.$$

Since $m \geq n/k$, the left-hand side of (4) is at most

$$\left(\frac{ne}{m}\right)^m \exp\left\{\frac{-1}{k} \binom{m}{r}\right\} \leq \exp\left\{m \left(\ln ek - \frac{1}{rk} \binom{m-1}{r-1}\right)\right\}.$$

Therefore, (4) will hold for large n if

$$(5) \quad 0.1 + \ln ek < \frac{1}{rk} \binom{m-1}{r-1}.$$

By (2), for n (and hence m and k) large in comparison with r ,

$$\frac{1}{rk} \binom{m-1}{r-1} \geq 0.9 \frac{1}{rk} \frac{m^{r-1}}{(r-1)!} \geq \frac{0.9}{rk} \frac{n^{r-1}}{(r-1)!k^{r-1}} \geq \frac{0.9}{rk^r} \Delta.$$

Thus (5) will hold for large n if

$$(6) \quad 0.9\Delta > rk^r(0.1 + \ln ek).$$

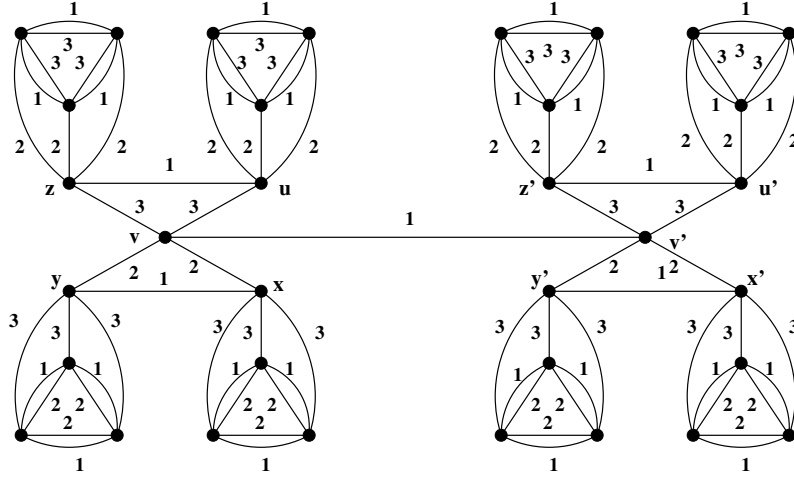
By the definition of k for large n , $0.1 + \ln ek < 1.1 \ln k$. Hence, to satisfy (6), it is enough that $0.9\Delta > 1.1rk^r \ln k$, which follows from (1). $\square$

For a positive integer k , let $f(k)$ (respectively, $f_m(k)$) be the largest Δ such that $\text{ch}_{ad}(G) \leq k$ for every simple graph (respectively, multigraph) G with maximum degree at most Δ . It follows from Propositions 2.2 and 2.3 that $k^2/8 \leq f_m(k) \leq f(k) \leq 4k^2 \ln k$. It is also interesting to find better estimates of $f_m(k)$ and $f(k)$ for small k . Theorem 1.6 yields $f_m(2) = f(2) = 2$. Theorem 2.4 below provides another upper bound for $\text{ch}_{ad}(G)$ when k is small.

THEOREM 2.4. *If G is a multigraph with maximum degree Δ , then $\text{ch}_{ad}(G) \leq \lceil \Delta/2 \rceil + 1$.*

Proof. Let F be an arbitrary edge coloring of G . We shall prove that if each vertex x of G is given a list $L(x)$ of at least $\lceil d(x)/2 \rceil + 1$ permissible colors, then G has an (L, F) -coloring. (Recall that “ (L, F) -coloring” means “ L -coloring adapted to F ”.) We prove this by induction on the number of edges of G . Without loss of generality, we assume that G is connected. If G has a vertex of degree 1, then $|L(x)| \geq 2$. Therefore, any (L, F) -coloring of $G - x$ can be extended to an (L, F) -coloring of G . If G is a cycle, then the conclusion follows from Theorem 1.6 in section 4. If G has a vertex x of degree 3, then $|L(x)| \geq 3$. Assume $e \in E(x)$ and $F(e) = j$. Remove color j from $L(x)$ to obtain a list assignment L' of $G - e$. By induction hypothesis, $G - e$ has an (L', F) -coloring c . Since $c(x) \neq F(e)$, c is also an (L, F) -coloring of G . Assume that G has a vertex x of degree $d(x) \geq 4$ with $|L(x)| = \lceil d(x)/2 \rceil + 1 < d(x)$. If for some edge $e \in E(x)$, $F(e) \notin L(x)$, then an (L, F) -coloring of $G - e$ is also an (L, F) -coloring of G . Thus we assume that $F(e) \in L(x)$ for all $e \in E(x)$. Then there are two edges $e, e' \in E(x)$ with $F(e) = F(e') = j$. Remove color j from $L(x)$ to obtain a list assignment L' of $G - \{e, e'\}$. By induction hypothesis, $G - \{e, e'\}$ has an (L', F) -coloring c . Since $c(x) \neq F(e) = F(e')$, c is also an (L, F) -coloring of G . $\square$

Theorem 2.4 implies that $f_m(3) \geq 4$. Figure 1 shows an edge 3-colored multigraph G with $\Delta(G) = 5$ that has no vertex 3-coloring adaptable to the given edge coloring. Indeed, assume it has such a 3-coloring c . Then either $c(v) \neq 1$ or $c(v') \neq 1$. By the left-right symmetry, we may assume that $c(v) \neq 1$. Now by the up-down symmetry, we may assume that $c(v) = 2$. Then $c(x) \neq 2$ and $c(y) \neq 2$. Since edge xy has color 1, at least one of x and y (say, x) should be colored with 3. Thus all three vertices

FIG. 1. A multigraph G with $\text{ch}_{ad}(G) \geq 4$.

below x can be colored with only 1 or 2, but the double edges between these vertices forbid any two of them to have the same color.

This example shows that $f_m(3) = 4$. We also have an example showing that $f(3) \leq 5$, but were not able to determine whether $f(3)$ is 4 or 5.

3. Adaptable choosability of d -degenerate graphs. While the upper bounds on adaptable choosability of graphs in terms of maximum degree are significantly better than those on ordinary choosability, this is not the case for bounds in terms of degeneracy. Recall that a multihypergraph G is d -degenerate if each subhypergraph G' of G has a vertex of degree at most d . It is obvious that $\text{ch}(G) \leq d + 1$ for every d -degenerate G . Since $\text{ch}_{ad}(G) \leq \text{ch}(G)$, we have $\text{ch}_{ad}(G) \leq d + 1$. In [6], examples of d -degenerate graphs G with $\chi_{ad}(G) = d + 1$ are given. It is also shown in [6] that for any positive integer d , there are graphs of large girth with $\chi_{ad}(G) = \chi(G) = d + 1$. However, these graphs are far from being d -degenerate. In this section, we prove Theorem 1.5, which is a common strengthening and generalization of these results in [6]. But first we prove a weaker statement.

PROPOSITION 3.1. *For any integers $g \geq 3$, $r \geq 2$, and $k \geq 1$, there is an r -uniform hypergraph G with girth at least g and $\text{ch}_{ad}(G) \geq k$.*

Proof. Let n be large in comparison with g, r , and k . Consider the random r -uniform n -vertex hypergraph $\mathbf{H}(n, r, k)$ with colored edges (and labeled vertices) obtained in the following way: Every r -element subset e of $\{1, \dots, n\}$ independently of all other such r -tuples with probability $1 - (\ln \ln n)/n^{r-1}$ is not in $\mathbf{H}(n, r, k)$, and for each $i = 1, \dots, k$, with probability $\ln \ln n / kn^{r-1}$, e is an edge in $\mathbf{H}(n, r, k)$ colored with color i .

Let F denote the obtained edge coloring, and let H_i be the subhypergraph of $\mathbf{H}(n, r, k)$ induced by the edges of color i , $i = 1, \dots, k$. We will prove using a standard argument that with positive probability both (i) and (ii) below hold:

- (i) there is a set S of at most $n/3$ vertices such that $\mathbf{H}(n, r, k) - S$ has no cycles of length g or less;
- (ii) independence number of every H_i is at most $n/2k$.

If we prove this, then there is a hypergraph H' for which both (i) and (ii) hold. Hence, the hypergraph $H'' = H' - S$ has at least $2n/3$ vertices and in any (k, F) -

coloring of H'' each color can be used for at most $n/2k$ vertices. Thus H'' has no such coloring.

First, consider property (i). For given two vertices v and w , the probability that $\mathbf{H}(n, r, k)$ has at least two edges containing both v and w is at most

$$\sum_{t=2}^{\binom{n-2}{r-2}} \left(\frac{\ln \ln n}{n^{r-1}} \right)^t \binom{\binom{n-2}{r-2}}{t} \leq \sum_{t=2}^{\binom{n-2}{r-2}} \frac{1}{t!} \left(\frac{\ln \ln n}{(r-2)!n} \right)^t \leq \left(\frac{\ln \ln n}{n} \right)^2.$$

Since there are $\binom{n}{2}$ such pairs $\{v, w\}$, the expectation E_2 of the number of the cycles of length two in $\mathbf{H}(n, r, k)$ is at most $0.5(\ln \ln n)^2$. Similarly, for $3 \leq j \leq g$, the expectation E_j of the number of the cycles of length j in $\mathbf{H}(n, r, k)$ is at most $(\ln \ln n)^j/j$. Therefore, the expectation E_0 of the number of the cycles of all lengths from 2 to g in $\mathbf{H}(n, r, k)$ is at most $(\ln \ln n)^g$. Hence for n large, (i) holds with probability at least $1 - 3(\ln \ln n)^g/n > 0.9$.

Similar to the proof of Proposition 2.3, for each $1 \leq i \leq k$, the probability that H_i has an independent set of size $m = \lceil n/2k \rceil$ is at most

$$\binom{n}{m} \left(1 - \frac{\ln \ln n}{kn^{r-1}} \right)^{\binom{m}{r}} \leq \left(\frac{ne}{m} \exp \left\{ -\frac{\ln \ln n}{kn^{r-1}} \binom{m-1}{r-1} / r \right\} \right)^m.$$

For fixed r and k and large n , the m th root of the right-hand side is at most

$$2ke \exp \left\{ -0.5 \frac{\ln \ln n}{kn^{r-1}} \frac{n^{r-1}}{r!k^{r-1}} \right\} \leq 6k \frac{(kr)^r}{\ln n} < \frac{1}{2k}.$$

It follows that with probability at least $0.9 - \frac{k}{2k} = 0.4$, both (i) and (ii) hold for $\mathbf{H}(n, r, k)$. $\square$

Now we are ready to prove Theorem 1.5. For convenience, we restate it here.

THEOREM 1.5. *For any integers $g \geq 3$, $r \geq 2$, and $d \geq 1$, there is an r -uniform d -degenerate hypergraph G with girth at least g and $\text{ch}_{ad}(G) = \chi_{ad}(G) = d + 1$.*

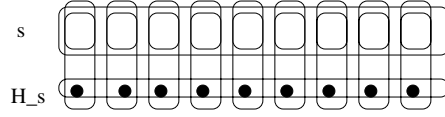
Proof. As observed above, $\chi_{ad}(G) \leq \text{ch}_{ad}(G) \leq \text{ch}(G) \leq d + 1$ for every d -degenerate G . So, it suffices to show that there is an r -uniform d -degenerate hypergraph G' of girth at least g with $\chi_{ad}(G') > d$, i.e., G' has a edge d -coloring F' such that G' has no vertex (d, F') -coloring.

We prove this by induction on d . When $d = 1$, for each r , let G_r be the hypergraph that contains one edge of size r and nothing else. Clearly, $\chi_{ad}(G_r) = 2$.

Assume that the conclusion is true for degeneracy $d - 1$ and all r and g . Given positive integers r and g , let H be an r -uniform $(d - 1)$ -degenerate hypergraph of girth at least g with $\chi_{ad}(H) = d$. Let $n = |V(H)|$ and $r' = n(r - 1)$. By Proposition 3.1, there exists an r' -uniform hypergraph $G = (V, E)$ of girth at least g with $\chi_{ad}(G) \geq d + 1$. Let G' be the hypergraph constructed as follows. For each edge $s = \{v_1, v_2, \dots, v_{r'}\}$ of G , add a disjoint copy of H , called H_s (including the vertices and edges). The vertices of H_s are $\{u_{1,s}, u_{2,s}, \dots, u_{n,s}\}$. Partition these vertices in s into n parts, $X_1, X_2, \dots, X_n$, each of cardinality $r - 1$. Add the edges $\{v_{j,s}\} \cup X_j$ ($j = 1, 2, \dots, n$). We denote the subhypergraph of G' induced by $s \cup V(H_s)$ by G_s (see Figure 2).

Since $\chi_{ad}(H) = d$, there is a $(d - 1)$ -edge coloring F of H such that H has no $(d - 1, F)$ -coloring. Since $\chi_{ad}(G) \geq d + 1$, there is an edge d -coloring Q of G such that G has no (d, Q) -coloring.

Define edge d -coloring F' of G' as follows. Let e be a hyperedge of G' . Then e is contained in G_s for some hyperedge s of G . Assume that $Q(s) = j$. If e is a hyperedge

FIG. 2. The part of G' corresponding to an edge s of G .

of H_s , then let

$$F'(e) = \begin{cases} F(e) & \text{if } F(e) \neq j, \\ d & \text{if } F(e) = j. \end{cases}$$

If e is of the form $\{v_{i,e}\} \cup X_i$, then let $F'(e) = i$. We shall show that G' has no (d, F') -coloring. Assume to the contrary that $c' : V(G') \rightarrow \{1, 2, \dots, d\}$ is such a coloring. Since $\chi_{ad}(G) = d + 1$, there is a hyperedge $s = \{v_1, v_2, \dots, v_{r'}\}$ of G such that $c'(v_1) = c'(v_2) = \dots = c'(v_{r'}) = Q(s)$. If $c(v_{i,s}) = Q(s)$ for any $i \in \{1, 2, \dots, n\}$, then $\{v_{i,s}\} \cup X_i$ is a hyperedge of G' of color $Q(s)$, and all the vertices of this hyperedge are colored by $Q(s)$, which is a contradiction. If $c'(v_{i,s}) \neq Q(s)$ for all $i \in \{1, 2, \dots, n\}$, then the restriction of c' to the copy G_s of G is a $(d - 1, F)$ -coloring, contrary to our assumption.

By construction, G' is d -degenerate and has girth at least g . This proves Theorem 1.5. $\square$

4. Adaptably 2-choosable graphs. The next observation will be quite helpful.

PROPOSITION 4.1. *Let G be a multihypergraph, L be a list assignment for the vertices of G , and F be an edge coloring of G . If G has no (L, F) -coloring, but after deleting any edge or vertex the resulting multihypergraph G' has, then*

- (a) *for every $e \in E(G)$, color $F(e)$ is present in the list of every $v \in e$;*
- (b) *for every $v \in V(G)$, every color in $L(v)$ is a color of some edge containing v .*

Proof. If $F(e)$ is not present in the list of some vertex $v \in e$ and $G - e$ has an (L, F) -coloring, then this coloring is also an (L, F) -coloring of G . If $L(v)$ contains a color α that is not a color of any edge containing v , and $G - v$ has an (L, F) -coloring c , then letting $c(v) = \alpha$ we get an (L, F) -coloring of G . $\square$

Given a coloring F of the edges of a hypergraph G , there is a unique list assignment L such that both (a) and (b) in Proposition 4.1 hold. We will denote this list assignment by $L_{G,F}$ and will say that it *respects* F . Furthermore, instead of “ $(L_{G,F}, F)$ -coloring of G ” we will simply say “ F -coloring of G .” For a vertex x of G , let $F(x) = \{F(e) : e \in E(x)\}$.

LEMMA 4.2. *Let $P = (v_0, v_1, \dots, v_n)$ be a path and F be any edge coloring of P . Then P has no F -coloring.*

Proof. Let $L = L_{P,F}$. For $i = 1, \dots, n$, let $e_i = v_{i-1}v_i$. Assume that c is an F -coloring of P . Then $L(v_0) = \{F(e_1)\}$ and hence $c(v_0) = F(e_1)$. This forces $L(v_1) \neq \{F(e_1)\}$. Since L respects F , we must have $F(e_2) \neq F(e_1)$ and $c(v_1) = F(e_2)$. Repeating this argument, we conclude that $c(v_i) = F(e_{i+1})$ for $i = 0, 1, \dots, n-1$. But $L(v_n) = \{F(e_n)\}$ and hence $c(v_n) = F(e_n) = c(v_{n-1})$, contrary to our assumption that c is an F -coloring of P . $\square$

COROLLARY 4.3. *Let C be a cycle and F be an edge coloring of P . If for some vertex $x \in V(C)$, $|L_{C,F}(x)| = 1$, then C has no F -coloring.*

Proof. Let $L = L_{C,F}$. Split the vertex x of C into two vertices x', x'' with each of x', x'' incident to one edge in $E(x)$, and let $L(x') = L(x'') = L(x)$. The resulting

graph is a path P , and F can be viewed as an edge coloring of P . Any F -coloring of C is also an F -coloring of P . Thus the conclusion follows from Lemma 4.2. $\square$

LEMMA 4.4. *If G is a subdivision of K_4 , then G is not adaptably 2-choosable.*

Proof. Consider first the case that no two adjacent edges are subdivided, i.e., that the subdivided edges form a matching of K_4 . Then the set of nonsubdivided edges of K_4 contains a 4-cycle C_4 . Let F be an edge coloring of G such that the edges of this 4-cycle are colored by color 1. The other edges of G are colored arbitrarily so that each vertex is incident to edges of two distinct colors. For example, the first two graphs in Figure 3 are colored in this way. Let $L = L_{G,F}$. Assume that G has an F -coloring c .

Let the four vertices of the 4-cycle C_4 be v_1, v_2, v_3, v_4 (these four vertices are also the only vertices of G of degree 3) in this order. Then either neither of v_1 and v_3 is colored by color 1, or neither of v_2 and v_4 is colored by 1. Without loss of generality, assume that neither of v_1 and v_3 is colored by 1. Then the restriction of c to the path $P_{1,3}$ of G connecting v_1 and v_3 is an (L', F) -coloring of $P_{1,3}$, where $L' = L_{P_{1,3}, F}$. This contradicts Lemma 4.2.

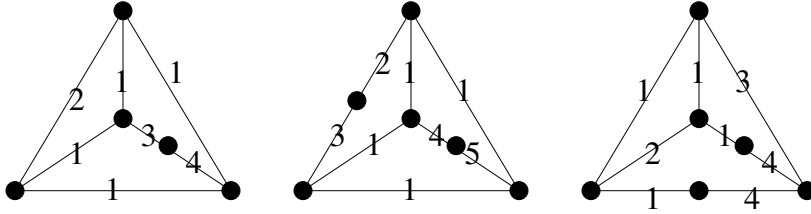


FIG. 3. Some subdivisions of K_4 .

Next, assume that G has two adjacent edges that are subdivided. Let H be the third graph in Figure 3. Then G is a subdivision of H .

Let F be an edge coloring F of a multigraph Q and $L = L_{G,F}$. We say that F is *difficult* if the following hold:

- (i) $|L(v)| = 2$ for each $v \in V(Q)$.
- (ii) For each color j , the edges of color j induce a tree T_j .
- (iii) Q has no (L, F) -coloring.

The edge coloring F of graph H in Figure 3 is difficult: (i) and (ii) are obvious. To see that (iii) holds, observe that if the top vertex is colored by 1, then the vertex at the center and the leftmost vertex must both be colored by 2, which is a contradiction. If the top vertex is colored by 3, then the rightmost vertex must be colored by 4, which forces both degree 2 vertices of H be colored by 1, which in turn forces the vertex at the center and the leftmost vertex be both colored by 2; which is again a contradiction.

CLAIM 4.1. *Suppose that a multigraph Q has a difficult coloring and Q' is obtained from Q by subdividing one edge. Then Q' has a difficult coloring.*

Proof. Assume Q' is obtained from Q by inserting a vertex w on an edge $e^* = uv$ of Q . In other words, e^* is subdivided into two edges $e' = uw$ and $e'' = wv$. Assume that e^* has color j . Let T_j be the tree induced by the edges of color j . Let T'_j be the tree obtained from T_j by subdividing the edge e^* into $e' = uw$ and $e'' = wv$. Vertex w separates $E(T'_j)$ into two parts, say E_1 (containing e') and E_2 (containing e''). Let j' be a new color (a color not used on any edge of Q). Let F' be the edge coloring of

Q' defined as follows:

$$F'(e) = \begin{cases} j & \text{if } e \in E_1, \\ j' & \text{if } e \in E_2, \\ F(e) & \text{otherwise.} \end{cases}$$

Now we show that F' is a difficult coloring of Q' . It is obvious that (i) and (ii) hold. Assume to the contrary that c' is an $L_{Q', F'}$ -coloring of Q' adaptable to F' . Let c be the L -coloring of Q defined as

$$c(x) = \begin{cases} j & \text{if } c'(x) = j \text{ or } j', \\ c'(x) & \text{otherwise.} \end{cases}$$

Observe that $c'(w) \in \{j, j'\}$, and hence either $c'(u) \neq j$ or $c'(v) \neq j'$. In either case, the ends u and v of the edge e^* cannot both have color j . This implies that c is an (L, F) -coloring of Q , contrary to our assumption that Q is not F -colorable. $\square$

By repeatedly applying Claim 4.1, we derive that each subdivision of H has a difficult edge coloring, and hence is not adaptably 2-choosable. This completes the proof of Lemma 4.4. $\square$

It follows from Lemma 4.4 that adaptably 2-choosable graphs are K_4 -minor free, i.e., series-parallel graphs.

LEMMA 4.5. *If a connected multigraph G has two cycles of length at least 3 that have at most one vertex in common, then G is not adaptably 2-choosable.*

Proof. Assume G has two cycles C_1 and C_2 of length at least 3 that have at most one vertex in common. Let P be a shortest path connecting C_1 and C_2 . Let $\{v\} = V(P) \cap V(C_1)$ and $\{v'\} = V(P) \cap V(C_2)$. This includes the case that $v' = v$.

Let G' be the subgraph of G induced by $V(C_1 \cup C_2 \cup P)$. Let F be an edge coloring of G' such that (1) the two edges of C_1 incident with v are colored with 1, (2) the two edges of C_2 incident with v' are colored with 2, and (3) the remaining edges are colored so that each vertex is incident to edges of two distinct colors (for example, all other edges may have different colors). We shall show that there is no F -coloring of G' . Assume to the contrary that c is such a coloring. If $c(v) = 1$, then the restriction of c to C_1 is an F -coloring of C_1 , contrary to Corollary 4.3. Similarly, if $c(v') = 2$, then the restriction of c to C_2 is an F -coloring of C_2 . If $c(v) \neq 1$ and $c(v') \neq 2$, then the restriction of c to the path P is an F -coloring of P , a contradiction to Lemma 4.2. $\square$

LEMMA 4.6. *If a simple graph G consists of four internally vertex-disjoint paths P_1, P_2, P_3 , and P_4 connecting two vertices u, v , then G is not adaptably 2-choosable.*

Proof. If P is a path not of length 2, then there is a proper edge coloring of P with at most three colors, such that the two end edges (which can be the same edge if P has length 1) are colored with the same color. If P is a path not of length 1, then there is a proper edge coloring of P with at most three colors, such that the two end edges are colored with two given distinct colors.

Case 1. At least two of the paths are not of length 2. Since at most one of $P_1, \dots, P_4$ has length 1, we can renumber our paths so that P_1 and P_2 are not of length 1, and P_3 and P_4 are not of length 2. Color properly the edges of P_1 (respectively, P_2) so that the edge incident to u is colored by 1 (respectively, by 2), and the edge incident to v is colored by 2 (respectively, by 1). Color the edges of P_3 (respectively, P_4) so that both end edges are colored by 1 (respectively, by 2). We have $L(u) = L(v) = \{1, 2\}$. If c is an F -coloring of G , then $c(u), c(v) \in \{1, 2\}$. If $c(u) = c(v) = 1$, then the restriction of c to P_3 is an F -coloring of P_3 , contrary to Lemma 4.2. If $c(u) = c(v) = 2$, then we

get a contradiction by considering path P_4 . If $c(u) = 1$ and $c(v) = 2$, then we obtain a contradiction with the edge coloring of P_1 , and if $c(u) = 2$ and $c(v) = 1$, then we obtain a contradiction with the edge coloring of P_2 .

Case 2. Each of P_1 , P_2 , and P_3 has length 2 and the length of P_4 is not 1. By above we can color the edges of G so that (a) the edges incident to u in P_1 and P_2 are colored with 1, and in P_3 and P_4 are colored with 2; (b) the edges incident to v in P_1 and P_3 are colored with 3, and in P_2 and P_4 are colored with 4. Now $L(u) = \{1, 2\}$ and $L(v) = \{3, 4\}$. Similarly to Case 1, each of the four paths forbids one of the four possibilities $(1, 3)$, $(1, 4)$, $(2, 3)$, and $(2, 4)$ for the pair $(c(u), c(v))$.

Case 3. $|E(P_1)| = |E(P_2)| = |E(P_3)| = 2$ and $|E(P_4)| = 1$. For $1 \leq i \leq 3$, let w_i be the central vertex of the path P_i . We let $F(uv) = F(uw_1) = F(w_2v) = 1$, $F(uw_2) = F(uw_3) = 2$ and $F(w_1v) = F(w_3v) = 3$. It is easy to check that G has no F -coloring for this F . $\square$

Now we are ready to prove Theorem 1.6. For convenience, we restate it here.

THEOREM 1.6. *A simple connected graph G with minimum degree at least 2 is adaptably 2-choosable if and only if G consists of two or three internally disjoint paths connecting two distinct vertices of G .*

Proof. Assume first that a simple connected graph G with minimum degree at least 2 is adaptably 2-choosable. By Lemma 4.4 G is series-parallel. Furthermore, by Lemma 4.5, G is 2-connected and any two cycles of G have at least two vertices in common. This implies that G has two vertices u and v such that for some $k \geq 2$, G consists of k internally vertex disjoint paths connecting u and v . By Lemma 4.6, $k \leq 3$. This proves the “only if” part of the theorem.

Suppose now that the “if” part of the theorem does not hold. Then there is an inclusion minimal subgraph G of a graph G' consisting of three internally disjoint paths connecting some vertices u and v that has no L -coloring adapted to an edge-coloring F , where $|L(w)| = 2$ for every $w \in V(G)$. By Proposition 4.1, G has no isolated or pendant vertices, and it is enough to consider F -colorings.

Let F be an edge coloring of G such that $|F(x)| = 2$ for each vertex x of G . Assume $F(u) = \{a, b\}$ and $F(v) = \{c, d\}$ (where $\{a, b\}$ and $\{c, d\}$ need not be disjoint).

There are four distinct ways to color u and v , namely $(c(u), c(v)) \in \{(a, c), (a, d), (b, c), (b, d)\}$. For each P_i , there is exactly one choice of the pair $(c(u), c(v))$ that cannot be extended to a coloring of P_i adapted to F . Thus, if $2 \leq k \leq 3$, then at least one of the four choices of $(c(u), c(v))$ can be extended to a coloring adapted to F for each of the u, v -paths. $\square$

Remark. Using the proof of the “only if” part of the theorem, one can characterize also all adaptably 2-choosable multigraphs. The remaining argument is a simple case analysis, so we just give a preliminary idea and state the characterization.

We need the following notion: a *double x, y -path* is a multigraph whose underlying simple graph is an x, y -path and in which every edge has multiplicity two. If G has a vertex x with only one neighbor y (does not matter how many edges connect x and y) and $G - x$ is adaptably 2-choosable, then G is also adaptably 2-choosable. Thus it is enough to characterize adaptably 2-choosable multigraphs whose underlying simple graphs have minimum degree at least two. Theorem 1.6 tells us that these underlying graphs have very limited structure.

THEOREM 4.7. *A connected multigraph G whose underlying simple graph G_s has minimum degree at least 2 is adaptably 2-choosable if and only if all of the following holds:*

- (a) G_s consists of two or three internally disjoint paths connecting two distinct vertices x and y of G ;

- (b) all double x, y -paths have the same parity;
- (c) if there are three internally disjoint x, y -paths P_1, P_2 , and P_3 such that P_1 is a double x, y -path of an odd length, then one of P_2 and P_3 has length 1;
- (d) if there are three internally disjoint x, y -paths P_1, P_2 , and P_3 such that P_1 is a double x, y -path of an even length, then at least one of P_2 and P_3 has length 2.

This description is somewhat more sophisticated than Theorem 1.6, but it yields a polynomial-time algorithm for recognition of adaptably 2-choosable multigraphs.

Acknowledgment. We thank a referee for helpful comments.

REFERENCES

- [1] P. ERDŐS AND L. LOVÁSZ, *Problems and results on 3-chromatic hypergraphs and some related questions*, in Infinite and Finite Sets, A. Hajnal et. al., eds., Colloq. Math. Soc. János Bolyai 10, North-Holland, Amsterdam, 1975, pp. 609–627.
- [2] P. ERDŐS, A. RUBIN, AND H. TAYLOR, *Choosability in graphs*, in Proceedings of the West Coast Conference on Combinatorial Graph Theory and Computing, Arcata, CA, 1979, Congress. Numer., 26 1979, Utilitas Math. Winnipeg, Manitoba, 1980, pp. 125–157.
- [3] T. FEDER AND P. HELL, *Full constraint satisfaction problems*, SIAM J. Comput., 36 (2006), pp. 230–246.
- [4] T. FEDER, P. HELL, S. KLEIN, AND R. MOTWANI, *List partitions*, SIAM J. Discrete Math., 16 (2003), pp. 449–478.
- [5] T. FEDER, P. HELL, D. KRAL, AND J. SGALL, *Two algorithms for general list matrix partition*, in Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), Vancouver, BC, 2005, pp. 870–876.
- [6] P. HELL AND X. ZHU, *Adaptable chromatic number of graphs*, European J. Combin., to appear.
- [7] P. HELL AND J. NEŠETŘIL, *Graphs and homomorphisms*, Oxford University Press, Oxford, UK, 2004.

UNIT CIRCULAR-ARC GRAPH REPRESENTATIONS AND FEASIBLE CIRCULATIONS*

MIN CHIH LIN[†] AND JAYME L. SZWARCFITER[‡]

Abstract. In a recent paper, Durán et al. [*J. Algorithms*, 58 (2006), pp. 67–78] described an algorithm of complexity $O(n^2)$ for recognizing whether a graph G with n vertices and m edges is a unit circular-arc (UCA) graph. Furthermore, the following open questions were posed in the above paper: (i) Is it possible to construct a UCA model for G in polynomial time? (ii) Is it possible to construct a UCA model, whose extremes of the arcs correspond to integers of polynomial size? (iii) If (ii) is true, could such a model be constructed in polynomial time? In the present paper, we describe a characterization of UCA graphs, based on network circulations. The characterization leads to a different recognition algorithm and to answering these questions in the affirmative. We construct a UCA model whose extremes of the arcs correspond to integers of size $O(n)$. The proposed algorithms, for recognizing UCA graphs and constructing UCA models, have complexities $O(n + m)$. Furthermore, the complexities reduce to $O(n)$, if a proper circular-arc (PCA) model of G is already given as the input, provided the extremes of the arcs are ordered. We remark that a PCA model of G can be constructed in $O(n + m)$ time, using the algorithm by Deng, Hell, and Huang [*SIAM J. Comput.*, 25 (1996), pp. 390–403]. Finally, we also describe a linear time algorithm for finding feasible circulations in networks with nonnegative lower capacities and unbounded upper capacities. Such an algorithm is employed in the model construction for UCA graphs.

Key words. algorithm, circular-arc graph, circular-arc model, circulations, networks, proper circular-arc graph, unit circular-arc graph

AMS subject classifications. 05C62, 05C85, 05C05, 68R10, 90B10

DOI. 10.1137/060650805

1. Introduction. In a recent paper, Durán et al. [2] described an algorithm of complexity $O(n^2)$ for recognizing whether a graph G , with n vertices and m edges, is a unit circular-arc (UCA) graph. Furthermore, the following open questions were posed in [2]:

- (i) Is it possible to construct a UCA model for G in polynomial time?
- (ii) Is it possible to construct a UCA model, whose extreme points of the arcs correspond to integers of polynomial size?
- (iii) If (ii) is true, could such a model be constructed in polynomial time?

Problems (ii) and (iii) were also proposed in the book by Spinrad [10]. As for problem (i), the proof of the characterization of UCA graphs in terms of forbidden subgraphs by Tucker [13], actually contains an algorithm for constructing a UCA model. However, due to the possible manipulation of large integers, the complexity of this algorithm is unknown, and so far the construction of a UCA model in polynomial time remains unsolved [2, 10].

*Received by the editors January 24, 2006; accepted for publication (in revised form) November 12, 2007; published electronically March 7, 2008.

<http://www.siam.org/journals/sidma/22-1/65080.html>

[†]Departamento de Computación, Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Buenos Aires 1428, Argentina (oscarlin@dc.uba.ar). This author's research was partially supported by UBACyT grants X184 and X212, PICT ANPCyT grant 11-09112, and CNPq under PROSUL project Proc. 490333/2004-4.

[‡]Universidade Federal do Rio de Janeiro, Instituto de Matemática, NCE and COPPE, Caixa Postal 2324, 20001-970 Rio de Janeiro, Brazil (jayme@nce.ufrj.br). This author's research was partially supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, and Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro, FAPERJ, Brazil.

In the present paper, we propose a characterization for UCA graphs, which leads to a different recognition algorithm and to answering in affirmative the previous three questions. An extended abstract of the present paper appeared in [7]. The proposed algorithm recognizes UCA graphs and constructs UCA models whose extremes of the arcs correspond to integers of $O(n)$ size, in overall time $O(n + m)$. Furthermore, if the graph G is already given by a proper circular-arc (PCA) model of it, then the complexity of the proposed algorithm reduces to $O(n)$, provided the extremes of the arcs are ordered. Observe that a PCA model for G , when given by its vertices and edges, can be constructed in $O(n + m)$ time, employing the algorithm by Deng, Hell, and Huang [1]. Recall that UCA graphs are PCA graphs.

Our method employs network circulations. We formulate an algorithm finding a feasible circulation in a network with nonnegative lower capacities and unbounded upper capacities. The algorithm has linear time complexity in the size of the network.

Denote by G an undirected graph, with vertex set $V(G)$ and edge set $E(G)$. For an edge $e \in E(G)$, write $e = uv$, where u and v are the *extremes* of e . Denote by $d_G(v)$ for the *degree* of v in G . We may also simply write $d(v)$, instead. Use a similar notation for a digraph D . For a directed edge $e = uv \in E(D)$, say that u is the *start* and v the *end* of e . Write $d_G^-(v)$ and $d_G^+(v)$ for the *indegree* and *outdegree* of vertex v , respectively. Again, also write $d^-(v)$ and $d^+(v)$, simply. Moreover, $E^-(v)$, $E^+(v) \subseteq E(D)$, represent the set of edges of D entering and leaving v , respectively. Also, D is *connected* when its underlying (undirected) graph is connected. Say that D is *strongly connected* when D contains paths from u to v and from v to u , for any $u, v \in V(D)$. Finally, the *bridge* of D is an edge contained in no (directed) cycle of D .

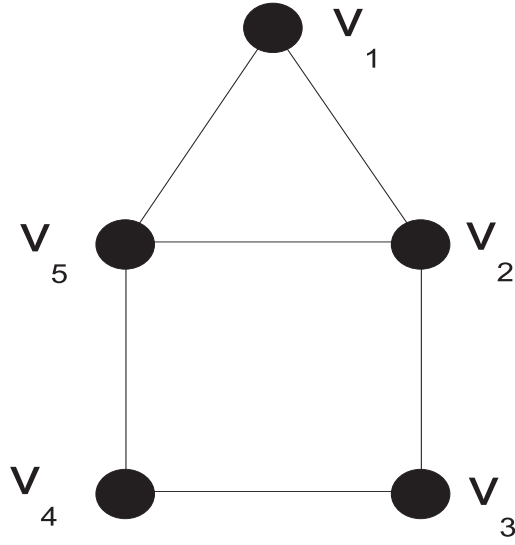
An *out-arborescence* (*in-arborescence*) T of D is a spanning connected subdigraph of D such that there is a distinguished vertex $v^* \in V(T)$, called the *root* of T , satisfying $d_T^-(v^*) = 0$ ($d_T^+(v^*) = 0$), while $d_T^-(v) = 1$ ($d_T^+(v) = 1$) for all of the remaining vertices $v \neq v^*$ of T . Clearly, a strongly connected digraph admits both an out-arborescence and in-arborescence, with root at any arbitrary vertex. When $d_T^+(v) = 0$ ($d_T^-(v) = 0$), call v a *leaf* of T . For $u, v \in V(T)$, if T contains a path from u to v , then u is an *ancestor* of v , and v a *descendant* of u . A *leaf-root ordering* of T is a sequence $v_1, \dots, v_n$ of its vertices, such that $i < j$ implies v_i is not an ancestor (descendant) of v_j in T .

A *network* is a digraph D , having real values b_j, c_j associated to each edge $e_j \in E(D)$. Call b_j the *lower capacity* of e_j , while c_j is the *upper capacity* of e_j . A *circulation* of D is a function W assigning a real number w_j to each edge e_j of D , called the *flow* of e_j . Denote $w^-(v_i) = \sum_{w_j \in E^-(v_i)} w_j$ and $w^+ = \sum_{w_j \in E^+(v_i)} w_j$, for each $v_i \in V(D)$. A circulation is *feasible* when it satisfies

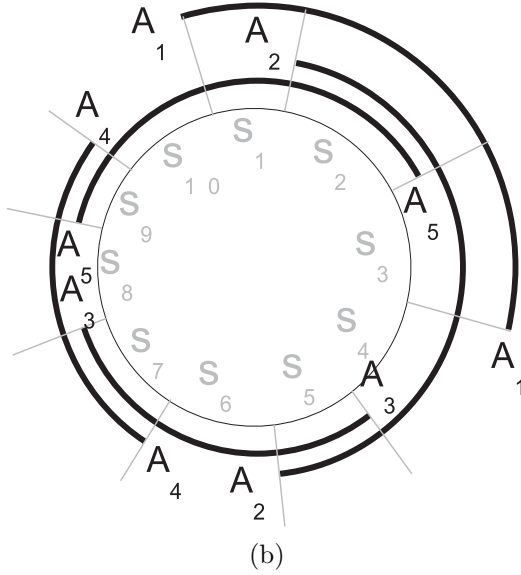
$$w^-(v_i) = w^+(v_i), \text{ for each } v_i \in V(D), \text{ and } b_j \leq w_j \leq c_j, \text{ for each } e_j \in E(D).$$

For our purposes, we only consider networks with finite nonnegative lower capacities and flow values, and unbounded the upper capacities. In this case, the condition $w_j \leq c_j$ is always satisfied.

A *circular-arc* (CA) model for a graph G is a pair $(C, \mathcal{A})$, where C is a circle and $\mathcal{A}$ is a collection of arcs of C , such that each arc $A_i \in \mathcal{A}$ corresponds to a vertex $v_i \in V(G)$, and any A_i, A_j intersects precisely when v_i, v_j are adjacent in G , $1 \leq i, j \leq n$, and $i \neq j$. A CA graph is one admitting a CA model. When no arc of $\mathcal{A}$ properly contains another arc of $\mathcal{A}$ then $(C, \mathcal{A})$ is a PCA model, while if all arcs of $\mathcal{A}$ have the same length, then $(C, \mathcal{A})$ is a UCA model. A PCA (UCA) graph is one admitting a PCA (UCA) model. A *normal* model is a PCA model where no two



(a)



(b)

FIG. 1. A PCA graph and a normal model of it.

arcs cover the circle. Figure 1(a) depicts a PCA graph and a normal model of it is in Figure 1(b). When traversing the circle C , always choose the clockwise direction. If s, t are points of C , then write (s, t) to mean the arc of C defined by traversing the circle from s to t . Call s, t the *extremes* of (s, t) , while s is the *start* and t the *end* of the arc. The *extremes* of $\mathcal{A}$ are the extremes of $A_i \in \mathcal{A}$. For $A_i \in \mathcal{A}$, write $A_i = (s_i, t_i)$. We assume the labelling $A_1, \dots, A_n$ of the arcs is such that the sequence of the start points $s_1, \dots, s_n$ is in the circular ordering of C . For $p, q \in A_i$, if $(s_i, p) \subseteq (s_i, q)$, then p *precedes* q , and q *succeeds* p in A_i . For $p, q, t \in C$, write $\max\{p, q\}_t$ to represent the

point, p or q , which is farthest from t , in the circular ordering of C . If p is an extreme of $\mathcal{A}$, then denote by $PRED(p)$ and $SUC(p)$, the extreme of $\mathcal{A}$, which immediately precedes and succeeds p , in the circular ordering, respectively. We assume that all arcs of C are open, no two extremes of distinct arcs of $\mathcal{A}$ coincide, and no single arc covers C . An arc of C defined by two consecutive extremes of $\mathcal{A}$ is a *segment* of $(C, \mathcal{A})$. Clearly, C is the union of the $2n$ segments of $(C, \mathcal{A})$ and the extreme points. Figure 1(b) shows the ten segments of the corresponding CA model.

The first characterization of CA graphs leading to a polynomial time recognition algorithm is due to Tucker [12, 14]. Subsequently, faster algorithms have been described by Hsu [6] and by Eschen and Spinrad [3]. More recently, a linear time recognition algorithm for CA graphs has been formulated by McConnell [8]. Deng, Hell, and Huang [1] described an algorithm for recognizing PCA graphs in linear time. This algorithm also constructs a corresponding PCA model for the graph. As for UCA graphs, the first polynomial time recognition algorithm is that by Durán et al. [2].

Given a graph G , the algorithm [2] initially employs algorithm [1] to construct a PCA model $(C, \mathcal{A})$ of G , if any. Clearly, G is not UCA if such a model does not exist. Afterwards, this algorithm proceeds with two main phases: The first of them is to transform $(C, \mathcal{A})$ into a normal model. The last phase is the actual algorithm for deciding if G is a UCA graph, employing the normal model constructed in the first phase. The complexities of the algorithms corresponding to these two phases are both $O(n^2)$. The method proposed in the present paper is also composed by similar phases. However, the complexities of the two corresponding algorithms are $O(n)$.

In section 2, we describe the characterizations in which are based the proposed algorithms. The characterization of UCA graphs relates them to circulations in special networks. Section 3 describes the algorithm for finding feasible circulations in general networks, having nonnegative lower capacities and unbounded upper capacities. Section 4 presents the algorithm for constructing normal models. Section 5 contains the actual algorithm for recognizing UCA graphs and constructing UCA models. Further comments form the last section.

2. A characterization for UCA graphs. In this section, we characterize UCA graphs in terms of circulations of a special network. The following two theorems are basic to our purposes.

THEOREM 1 (see Golumbic [4] and Tucker [13]). *Let G be a PCA graph. Then G admits a normal model.*

THEOREM 2 (see Tucker [13]). *Let G be a UCA graph and $(C, \mathcal{A})$ a normal model of it. Then G admits a UCA model, such that its extremes are in the same circular ordering as those of $(C, \mathcal{A})$.*

Let G be a CA graph and $(C, \mathcal{A})$ a CA model of it. Denote by $S_1, \dots, S_{2n}$ the segments of $(C, \mathcal{A})$ in circular ordering, where the start of S_1 and that of A_1 coincide. Denote by l_j the length of S_j . Clearly, any arc $A_i \in \mathcal{A}$ may be decomposed into the segments which form it. The length of A_i equals $\sum_{S_j \subseteq A_i} l_j$. The decomposition allows one to represent relations between lengths of arcs by relations between lengths of the corresponding segments. Consequently, the condition of equality between any two arc lengths required by a UCA model can be expressed by a system of $n - 1$ linear equations q_i , together with $2n$ inequalities, called the *full system* of $(C, \mathcal{A})$:

$$(2.1) \quad q_i : \sum_{S_j \subseteq A_i} l_j = \sum_{S_j \subseteq A_{i+1}} l_j, \quad i = 1, \dots, n - 1,$$

$$(2.2) \quad l_j > 0, \quad j = 1, \dots, 2n.$$

Clearly, equation q_i corresponds to the condition that the length of arc A_i equals that of A_{i+1} . The segment lengths l_j are the variables of such a system of equations. Our aim is to find a solution of the full system, if existing. The values of l_j , obtained from the solution of the system, would define a UCA model for G . In equation q_i , call the term $\sum_{S_j \subseteq A_i} l_j$ the *left side* of q_i , while $\sum_{S_j \subseteq A_{i+1}} l_j$ is its *right side*. Each equation q_i can possibly be simplified if the length l_j of a same segment appears in both the left and right sides of q_i . In this case, subtract l_j from both sides of q_i . The equations so obtained form the *reduced system* of $(C, \mathcal{A})$. The following lemma describes a useful property of these systems.

LEMMA 3. *Let G be a PCA graph, $(C, \mathcal{A})$ a normal model of it, and R the reduced system of $(C, \mathcal{A})$. Then each segment length l_j of $(C, \mathcal{A})$ appears at least once in R , unless S_j is a segment contained in all arcs of $\mathcal{A}$ or in none of them. Furthermore, l_j appears at most twice in R . In the latter situation, the two instances of l_j in R occur in different sides of their equations.*

Proof. Let S_j be any segment of the normal model $(C, \mathcal{A})$. Denote by $\mathcal{A}'$ the set of arcs of $\mathcal{A}$ covering S_j . If $\mathcal{A}' = \emptyset$, then no arc of $\mathcal{A}$ contains S_j , implying that l_j does not appear in R , and the lemma is satisfied. Consider $\mathcal{A}' \neq \emptyset$. We know that $\mathcal{A}'$ does not entirely cover C , otherwise there would be two arcs of $\mathcal{A}$ which would do so, contradicting $(C, \mathcal{A})$ to be normal. Consequently, the circular ordering of the extremes of $(C, \mathcal{A})$, when restricted to the arcs of $\mathcal{A}'$, is a linear ordering. Let A_i and A_k be the first and last arcs of $\mathcal{A}'$, respectively, in the considered linear ordering.

First, let $i \leq k$. Observe that l_j is not on the right side of any of the equations $q_i, \dots, q_k$, of R . Because if l_j is on the right side of q_t , then $k \geq t \geq i$ in the full system. However, l_j is also on the left side of q_t , hence it gets simplified in R . However, l_j is on the right side of q_{i-1} in R provided $i > 1$. Similarly, l_j is not on the left side of any $q_i, \dots, q_{k-1}$ but it is at the left of q_k , provided $k < n$. On the other hand, $i = 1$ and $k = n$ imply that S_j is common to all arcs of $\mathcal{A}$. Consequently, each l_j appears at least once in R , unless S_j is contained in all arcs of $\mathcal{A}$, or in none of them.

Consider the alternative $i > k$. If $i - 1 > k$, then l_j appears on the left of q_{i-1} and on the right of q_k . There are no other occurrences of l_j . Finally, when $i - 1 = k$ it follows that S_j is common to all arcs of $\mathcal{A}$. Again, each l_j always appears at least once in R , except when S_j is contained in all arcs of $\mathcal{A}$. Furthermore, it appears at most twice in R . In the latter situation, l_j occurs on different sides of the corresponding equations. $\square$

Aiming to solve R efficiently, we describe a graph theoretical model for it. Define the *segment digraph* D , as follows. There is a vertex $v_i \in V(D)$ for each equation q_i of R , and one edge e_j for each segment S_j of $(C, \mathcal{A})$. In addition, there is a distinguished vertex $v_0 \in V(D)$. Each edge $e_j \in E(D)$ is directed according to:

- (i) If l_j appears at the left of some equation q_i of R , then e_j starts at v_i ; otherwise it starts at v_0 .
- (ii) If l_j appears at the right of some equation q_k , then e_j ends at v_k ; otherwise it ends at v_0 .

The description of D is complete. Clearly, D has n vertices and $2n$ edges and no loops, except possibly at v_0 .

By adding to each edge $e_j \in E(D)$ a lower capacity $b_j = 1$ and an unbounded upper capacity c_j , we obtain the *segment network* of D .

As an example, the reduced system of the PCA model of Figure 1(b) is illustrated in Figure 2(a), while the corresponding segment digraph is in Figure 2(b). Similarly,

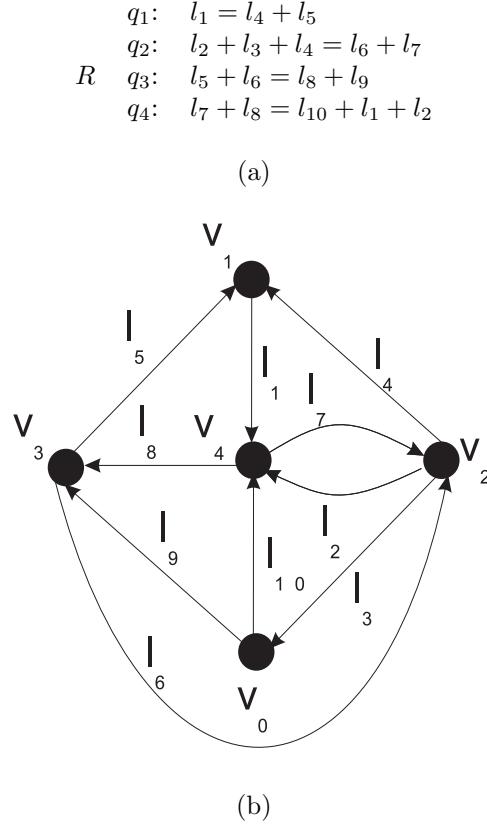


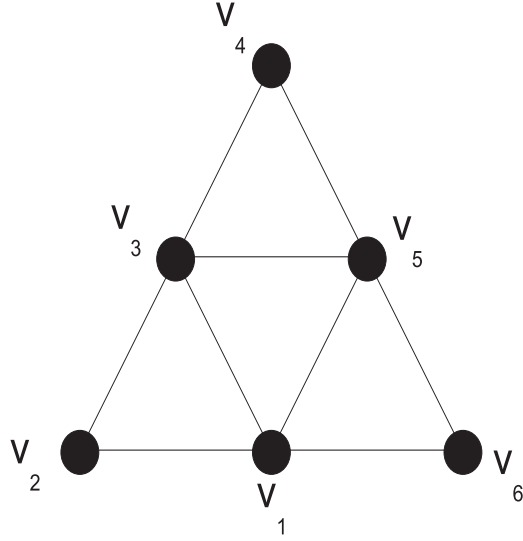
FIG. 2. The reduced system of Figure 1(b) and its segment digraph.

the reduced system and segment digraph corresponding to the PCA model of Figure 3(b) appear in Figures 4(a) and 4(b), respectively.

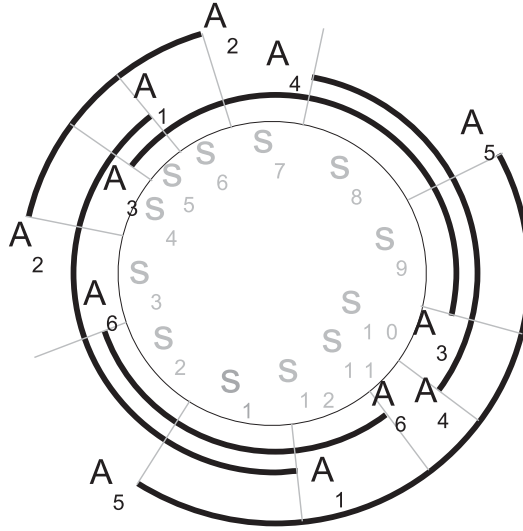
The following theorem characterizes UCA graphs in terms of feasible circulations.

THEOREM 4. *Let G be a PCA graph, $(C, \mathcal{A})$ a normal model of it, and D its segment network. Then G is a UCA graph if and only if D admits a feasible circulation.*

Proof. Let G be a UCA graph, and $(C, \mathcal{A})$ a normal model of G . By Theorem 2, there exists a UCA model $(C', \mathcal{A}')$ of G such that the endpoints of the arcs of $\mathcal{A}$ and $\mathcal{A}'$ are in the same ordering. That is, $(C, \mathcal{A})$ and $(C', \mathcal{A}')$ differ only by the lengths of the corresponding segments. Let R be the reduced system of $(C, \mathcal{A})$ and D be the segment network. Define a circulation W by assigning a flow w_j to each edge e_j of D , equal to the length l'_j of the segment of $(C', \mathcal{A}')$, corresponding to S_j . We show that such an assignment W is a circulation of D . Let $v_i \in V(D)$, $1 \leq i \leq n-1$, and q_i the corresponding equation in R . Then $w^-(v_i) = \sum_{S_j \subseteq A_i \setminus A_{i+1}} l'_j$. Similarly, $w^+(v_i) = \sum_{S_j \subseteq A_{i+1} \setminus A_i} l'_j$. Because $(C', \mathcal{A}')$ is a UCA model, the lengths of arcs A'_i and A'_{i+1} of $\mathcal{A}'$ are the same, implying that $w^-(v_i) = w^+(v_i)$. It remains to show that this equality also holds for v_0 . This assertion follows from the fact that $\sum_{0 \leq i \leq n-1} w^-(v_i) = \sum_{0 \leq i \leq n-1} w^+(v_i)$, because every edge e_j counts l'_j units both in $\sum_{0 \leq i \leq n-1} w^-(v_i)$ and $\sum_{0 \leq i \leq n-1} w^+(v_i)$. Since $w^-(v_i) = w^+(v_i)$, for each $1 \leq i \leq n-1$, it follows that $w^-(v_0) = w^+(v_0)$. Consequently, W is indeed a feasible circulation of D .



(a)



(b)

FIG. 3. Another PCA graph with its normal model.

Conversely, by hypothesis, D admits a feasible circulation W . We prove that G is UCA by constructing a UCA model $(C', \mathcal{A}')$ for G . First, consider the situation where no segment S_j of $(C, \mathcal{A})$ is contained in all arcs of $\mathcal{A}$ or in none of them. Let w_j be the flow of edge e_j of D . Construct the model $(C', \mathcal{A}')$ of G , by maintaining the endpoints of the arcs of $\mathcal{A}'$ in the same circular ordering as those in $\mathcal{A}$, while possibly modifying the lengths of the segments. The length of C' is defined as $\sum_{e_j \in E(D)} w_j$, while the length of the segment of $(C', \mathcal{A}')$ corresponding to S_j is set to the flow w_j of e_j . We show that the model so constructed is a UCA model.

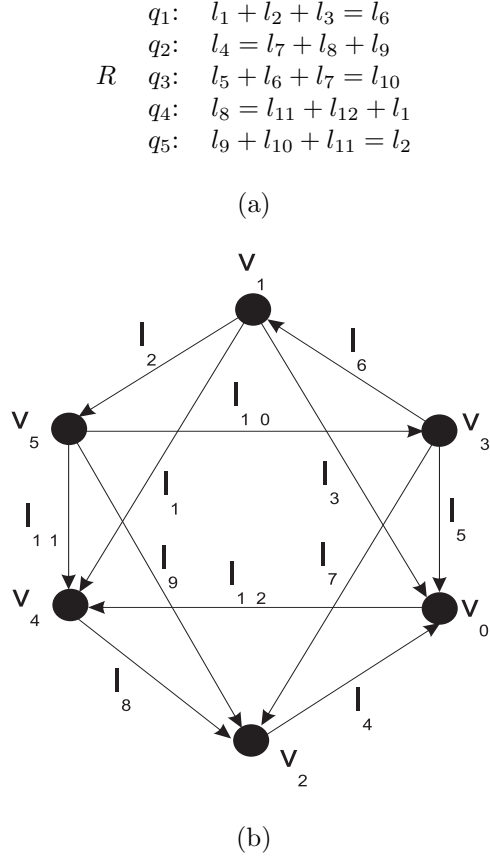


FIG. 4. The reduced system of Figure 3(b) and its segment digraph.

Let R be the reduced system of $(C, \mathcal{A})$. We show that the assignment $l_j := w_j$ for each segment S_j of $(C, \mathcal{A})$ is a solution to R . Let q_i be an equation of R and v_i the corresponding vertex of D . By the construction of D , the left side of q_i corresponds to the flow values of the edges of D ending at v_i , while the right side of q_i corresponds to the flow values of the edges, starting at v_i . Because D is feasible, $w^-(v_i) = w^+(v_i)$, implying that $\sum_{S_j \subseteq A_i} l_j = \sum_{S_j \subseteq A_{i+1}} l_j$. Moreover, all l_j are positive, because the lower capacities are satisfied. Consequently, W is a solution to R . Clearly, any solution to R is also a solution to the full system. On the other hand, because no segment is contained in all arcs of $\mathcal{A}$ or in none of them, Lemma 3 implies that all segment lengths appearing in the full system also appears in the reduced system, and consequently have been assigned a required value. The latter implies that all arc lengths of $\mathcal{A}'$ are equal, meaning that $(C', \mathcal{A}')$ is a UCA model.

Finally, examine the situation where there is a segment S_j contained in all arcs of $\mathcal{A}$ or in none of them. Construct $(C', \mathcal{A}')$ as above, and assign any positive length to S_j . Clearly, all arc lengths of $\mathcal{A}'$ are equal and $(C', \mathcal{A}')$ is a UCA model. $\square$

The above theorem implies that the problem of recognizing if a given graph G is UCA is equivalent to deciding if the corresponding segment network admits a feasible circulation. We describe a characterization for the existence of circulations in general networks, with arbitrary real nonnegative lower capacities and unbounded upper ca-

capacities. This is a special case of Hoffman's circulation theorem, as below. See [9]. For a subdigraph B of D , write $b^-(B) = \sum_{e_j \in E^-(B)} b_j$ and $c^+(B) = \sum_{e_j \in E^+(B)} c_j$.

THEOREM 5 (see Hoffman [5]). *Let D be a network having an arbitrary real lower capacity b_j and upper capacity c_j , for each edge $e_j \in E(D)$. Then D admits a feasible circulation W if and only if $b^-(B) \leq c^+(B)$, for each subdigraph B of D . Moreover, if b and c are integers, then the flow values can be taken as integers.*

COROLLARY 6. *Let D be a network with real nonnegative lower capacities and unbounded upper capacities. Then D admits a feasible circulation if and only if all bridges of D have lower capacity zero. In this case, a circulation with integer flow values always exists.*

Proof. Suppose D admits a feasible circulation W . To the contrary, assume that D has a bridge e_j with positive lower capacity. Then $e_j \in E^+(B), E^-(B')$ for some distinct strongly connected components B, B' of D . In this situation, we know that $w(E^-(B')) > 0$ and $w(E^+(B')) > 0$. The latter implies that there exists some edge e_k leaving B' and enters a strongly connected component $B'' \neq B, B'$, such that $w_k > 0$. Repeating this argument over and over leads to the contradiction that D is not finite. Consequently, no such e_j may exist.

Conversely, by hypothesis all bridges of D have lower capacity equal to zero. Let B be any strongly connected component of D . Then any proper subdigraph B' of B satisfies $E^-(B'), E^+(B') \neq \emptyset$. Since the upper capacities of D are unbounded and the lower capacities are not, it follows that $b^-(B') < c^+(B')$. Using Theorem 5, we conclude that B admits a feasible circulation, then repeat this argument for the remaining strongly connected components of D . Finally, assign the flow value equal to zero to all bridges of D . We have so obtained a feasible circulation. By simply replacing each lower capacity b_j by the value $\lceil b_j \rceil$, we obtain a feasible circulation with integer flow values. $\square$

As examples, observe that the segment digraph of Figure 2(b) is strongly connected, while that of Figure 4(b) is not. Consequently, the graph of Figure 1(a) is a UCA graph, while that of Figure 3(a) is not.

3. Finding feasible circulations. Let D be a network having finite real nonnegative lower capacities and unbounded upper capacities. We describe an algorithm which finds a feasible circulation W of D , or reports that such circulation does not exist. For each edge e_j of D , let b_j be the given nonnegative lower capacity. The algorithm below computes integer flow values w_j of the feasible circulation W of D .

1. Find the strongly connected components and the bridges of D . If any bridge has a positive lower capacity, then report that no feasible circulation exists and stop. Otherwise, set the flow value $w_j := 0$ for each bridge edge e_j of D . Then perform step 2 for each strongly connected component B of D . At termination, stop.
2. Assign the initial flow values $w_j := \lceil b_j \rceil$, for each edge $e_j \in E(B)$. Then compute $w^-(v_i) = \sum_{e_j \in E^-(v_i)} w_j$ and $w^+(v_i) = \sum_{e_j \in E^+(v_i)} w_j$ for each $v_i \in V(B)$. Finally, run *OUT PROCEDURE* and afterwards run *IN PROCEDURE*.

OUT PROCEDURE. Choose an arbitrary vertex $v^* \in V(B)$ and find an out-arborescence T of B , with root v^* . Repeat the following operations, for each vertex $v \in V(T)$, $v \neq v^*$, in leaf-root ordering: if $w^-(v) < w^+(v)$ then set $w^+(u) := w^+(u) + w^+(v) - w^-(v)$, $w_j := w_j + w^+(v) - w^-(v)$ and $w^-(v) := w^+(v)$, where $e_j = uv$ is the edge of T ending at v .

IN PROCEDURE. Construct an in-arborescence T of B , with the same root v^* , as chosen in the computation of the *OUT PROCEDURE* for B . Repeat the following operations, for each vertex $v \in V(T)$, $v \neq v^*$, in leaf-root ordering: if $w^+(v) < w^-(v)$, then set $w^-(u) := w^+(u) + w^-(v) - w^+(v)$, $w_j := w_j + w^-(v) - w^+(v)$, and $w^+(v) := w^-(v)$, where $e_j = vu$ is the edge of T starting at v .

Next, we assert the correctness of the algorithm.

THEOREM 7. *The above algorithm finds a feasible circulation for D .*

Proof. From Corollary 6, it follows that the algorithm correctly reports the nonexistence of a circulation, whenever there exists a bridge having a positive lower capacity. Otherwise, let B be any strongly connected component of D . We show that the flow values assigned by the algorithm to the edges of B define a feasible circulation for B . Furthermore, since the bridges are assigned a zero flow, the latter implies that the circulation of D so obtained is also feasible.

First, observe that since B is strongly connected it admits both an out-arborescence and an in-arborescence, rooted at the same arbitrary vertex $v^* \in V(B)$. Classify the vertices v_i of D into three types, according to the initial values of $w^-(v_i)$ and $w^+(v_i)$. Say that a vertex v is *in-deficient* when $w^-(v) < w^+(v)$, *out-deficient* when $w^+(v) < w^-(v)$, and *balanced* when $w^-(v) = w^+(v)$.

Consider the computation of *OUT PROCEDURE* and examine the vertices of the chosen out-arborescence T of B , in the ordering chosen by the algorithm. Let $v \neq v^*$ be the vertex next to be visited, in the considered ordering. If v is in-deficient, then the algorithm chooses the edge e_j of T ending at v and modify its weight from w_j to the value $w_j + w^+(v) - w^-(v)$. It is clear that v becomes balanced. Moreover, since T contains exactly one edge e_j entering v , we can assure that v remains balanced until the end of *OUT PROCEDURE*. Consequently, B does not contain in-deficient vertices at the completion of the procedure, except possibly v^* . Similarly, after the completion of *IN PROCEDURE*, B does not contain out-deficient vertices, except possibly v^* . Furthermore, the computation of *IN PROCEDURE* cannot create in-deficient vertices. Consequently, after completing the visits to the second tree, we can assure that all vertices are balanced, except possibly v^* . However, $w_j = 0$, for every bridge e_j of D . Consequently, $\sum_{v \in V(B)} w^-(v) = \sum_{v \in V(B)} w^+(v) = \sum_{e_j \in E(B)} w_j$. That is, $w^-(v^*) = \sum_{v \in V(B)} w^-(v) - \sum_{v \neq v^*} w^-(v) = \sum_{v \in V(B)} w^+(v) - \sum_{v \neq v^*} w^+(v) = w^+(v^*)$, meaning that v^* is automatically balanced. Finally, the vertices of B' outside B are clearly balanced, because all edges entering and leaving them have flow value zero. Therefore W is indeed a feasible circulation. $\square$

It is simple to determine the complexity of the algorithm. Finding strongly connected components can be done in $O(n + m)$ time [11], so as for the initial flow value assignments. The computation of *OUT PROCEDURE* and *IN PROCEDURE* require $O(n)$ time. Consequently, the overall time is $O(n + m)$.

Finally, observe that the flow values of the feasible circulation obtained by the algorithm are all integers of size $w_j < 2 \sum_{e_j \in E(D)} \lceil b_j \rceil$.

4. Constructing normal models. Let G be a PCA graph and $(C, \mathcal{A})$ a PCA model of it. We describe an algorithm for transforming $(C, \mathcal{A})$ into a normal model. The algorithm is based on [2, 4, 13]. For each $A_i \in \mathcal{A}$, the idea is to traverse C , searching for an arc A_j of $\mathcal{A}$ which together with A_i would cover C . If such an arc is found, then the arc A_j is conveniently shrunk with the purpose that A_i, A_j would no longer cover C . Clearly, it has to be assured that the new model so obtained is still a PCA model for G . Recall from Theorem 1 that any PCA graph admits a normal

model.

The proposed method considers the arcs A_i in the circular ordering $A_1, \dots, A_n$. For each i , only a fraction of the arcs of $\mathcal{A}$ would be examined in general, with the above goal of searching for an arc A_j , which together with A_i , cover C . The algorithm consists of the computation of a procedure $NORMAL(i)$, for $i = 1, \dots, n$. Each computation $NORMAL(i)$ examines the extreme points $s'_i, \dots, t'_i \in A_i \cup \{s_i, t_i\}$, consecutive in the circular ordering, where s'_i is the *first point* and t'_i the *last point* of $NORMAL(i)$, respectively. Each iteration inside $NORMAL(i)$ considers one of the extreme points $s'_i, \dots, t'_i$, in circular ordering. Let p be the current point under examination by $NORMAL(i)$. Clearly, p is either a start point or an end point. The following is the action taken, according to these alternatives. If p is a start point, then nothing is done and the examination of p is terminated. Otherwise, $p = t_j$ for some j , and check whether A_i and A_j cover C . In the affirmative case, shrink A_j in such a way that A_i and A_j would no longer cover C , and terminate the examination of p . The actual shrinking operation consists of moving t_j counterclockwise in C , so that it becomes an interior point of the segment $(PRED(s_i), s_i)$. When A_i and A_j do not cover C , then terminate the computation $NORMAL(i)$, disregarding all possible extreme points of $(C, \mathcal{A})$ contained in A_i which lie after $p = t_j$, when traversing A_i in the circular ordering. Consequently, such an extreme point is the last point t'_i of $NORMAL(i)$. The first point s'_i of $NORMAL(i)$ is determined at the moment $NORMAL(i)$ starts, as follows. When $i = 1$, define $s'_1 = s_1$. For $i > 1$, let $s'_i = \max\{s_i, t'_{i-1}\}_{s_{i-1}}$. Similarly as above, the algorithm disregards all extreme points contained in A_i , which lie before s'_i in A_i .

The procedure is formulated below. If p and q are points of C , then denote by $POINT(p, q)$ an arbitrarily chosen point of the (open) arc (p, q) . The external calls are $NORMAL(i)$, for $i = 1, \dots, n$.

procedure $NORMAL(i)$
 if $i = 1$ **then** $p := s_1$ **else** $p := \max\{p, s_i\}_{s_{i-1}}$
 repeat
 if p is an end point t_j **then**
 if $A_i \cup A_j = C$ **then** $t_j := POINT(PRED(s_i), s_i)$
 else return
 $p := SUC(p)$

The following theorem assures that $NORMAL(i)$ terminates, and $(C, \mathcal{A})$ is a normal model after the computation of $NORMAL(n)$.

THEOREM 8. *The algorithm is correct.*

Proof. The algorithm consists of the computation of $NORMAL(i)$, for $i = 1, \dots, n$. We prove that after completing $NORMAL(i)$, no two arcs A_k and A_j cover C , with $1 \leq k \leq i$. Furthermore, we ought to prove that the shrinking operations maintain $(C, \mathcal{A})$ as a PCA model for G .

Clearly, if there is an arc A_j such that A_i and A_j cover C , then $s_j, t_j \in A_i$, and t_j precedes s_j in A_i . Suppose the point $p = t_j$ is reached during $NORMAL(i)$. Then $A_i \cup A_j = C$ implies that t_j will be moved to $POINT(PRED(s_i), s_i)$. Let A_j^* be the new arc A_j , after shrinking. We know that A_i, A_j^* no longer cover C . We show that the configuration of $(C, \mathcal{A})$ still reflects the adjacencies of G . Clearly, A_i and A_j^* still intersect, otherwise A_i, A_j would not cover C . We discuss the intersections between A_j^* and any other arc $A_k \neq A_i$. We can restrict to arcs A_k which intersect A_i and have an extreme point, s_k or t_k , inside (s_i, t_j) , since the intersections of A_j^* and the remaining arcs are not affected by the shrinking of A_j . The following are the possible

alternatives.

- Case 1.* $s_k \in (s_i, t_j)$. If $t_k \in (s_i, t_i)$, then t_k precedes s_k in A_i , otherwise $(C, \mathcal{A})$ is not a PCA model. Consequently, A_k and A_j^* also intersect. When $t_k \notin (s_i, t_i)$ it follows that A_j^* and A_k again intersect, because $s_j \in (s_i, t_i)$.
- Case 2.* $s_k \notin (s_i, t_j)$. By the initial assumption, $t_k \in (s_i, t_j)$. Examine the possibilities for s_k . The situation $s_k \in (s_i, t_j)$ is in Case 1. Let $s_k \notin (s_i, t_j)$. Then $s_k \in (t_j, s_j)$, otherwise A_j contains A_k . Consequently, A_j^* and A_k again intersect.

Consequently, shrinking A_j in the manner $NORMAL(i)$ does actually preserve adjacencies. In what follows, we show that the modified model is still a PCA model. It is sufficient to prove that A_j^* cannot be contained in any other arc of $\mathcal{A}$, except for A_j which has been removed. Suppose the contrary, that is, A_j^* is contained in some arc A_k . In this case, the circular ordering of the considered extreme points must be $t_j^*, s_i, t_k, t_j, s_k, s_j, t_i$; otherwise A_k also contains A_j , or A_k does not contain A_j^* . Consequently, $A_k \cup A_i = C$. Since t_k precedes t_j in A_i , it follows that A_k would have been shrunk before A_j in $NORMAL(i)$. The latter contradicts A_k to contain A_j^* .

It remains to show that any arc A_k , which together with A_i covers C , is detected by the algorithm. Suppose the contrary. That is, $NORMAL(i)$ terminates and does not shrink A_k . Then $t_k \notin (s'_i, t'_i)$. Because $t_k \in (s_i, t_i)$, there are two possibilities.

- Case 1.* $t_k \in (s_i, s'_i)$. Then $s_i \neq s'_i$, which implies $i > 1$ and $t'_{i-1} \in A_i$. We also know that $s'_i \equiv t'_{i-1}$ is an end point t_j satisfying $A_{i-1} \cup A_j \neq C$. Observe that A_{i-1} and A_j may (or may not) coincide. Because $(C, \mathcal{A})$ is a PCA model, the sequence $s_j, s_{i-1}, s_i, t_k, t_j, t_{i-1}, t_i$ is in the circular ordering. Try to locate s_k . Because $A_i \cup A_k = C$, $s_k \in (t_k, t_i)$. However, $s_k \notin (t_k, t_{i-1})$ because A_k and A_{i-1} would also cover C , implying that $NORMAL(i-1)$ would also have moved t_k from the considered position, by an inductive argument. On the other hand, $s_k \notin (t_{i-1}, t_i)$ because in this case $A_{i-1} \cup A_k \neq C$, implying that t_k would have been reached by $NORMAL(i-1)$ before t_j . In the latter situation, $NORMAL(i-1)$ would have terminated at t_k , contradicting the assumption.
- Case 2.* $t_k \in (t'_i, t_i)$. We know there is an arc A_j , such that $t_j = t'_i$ and $A_i \cup A_j \neq C$. Because $A_i \cup A_k = C$, it follows that $s_k \in (t_k, t_i)$. In this situation, A_k contains A_j , contradicting $(C, \mathcal{A})$ to be a PCA model. Hence, Case 2 also does not occur.

Therefore every arc, which together with A_i covers C , is shrunk by the algorithm, completing the proof of correctness. $\square$

THEOREM 9. *The algorithm terminates within $O(n)$ time. Moreover, there are less than $6n$ iterations of the repeat block during the entire process.*

Proof. The number of steps performed by the algorithm corresponds to the number of iterations of the *repeat* block of the described formulation. Such a number is equivalent to the number of assignments of extreme points to p during the entire process. We will prove that the total number of such assignments is less than $6n$. The first value of p is s_1 and the last one, in the worst case, is t_n . We know that p never moves counterclockwise in C . First, assume for a while that p does not stay stationary in two consecutive iterations, and also disregard the shrinking of arcs. Any of the $2n$ extreme points can be assigned to p . If $t_n \notin A_1$, then the assignments would correspond at most to a complete turn around the circle, because the model is PCA. That is, $2n$ assignments. On the other hand, if $t_n \in A_1$, then any extreme lying between s_1 and t_n may be assigned again to p , but at most once more, besides

its assignment in *NORMAL*(1). This corresponds to an additional n units in our account.

Next, examine the case when p stays stationary in two consecutive iterations. Such a situation may only occur, possibly at the first iteration within *NORMAL*(i), when $s'_i = t'_{i-1}$, for $i > 1$. Overall, it counts to additional $n - 1$ units, at most, in the total number of assignments to p .

Finally, examine the shrinking operations. Each time an arc A_j is shrunk, an extreme point t_j^* is placed in a different position of $(C, \mathcal{A})$, offering the possibility to t_j^* itself to be assigned to p . Hence, the shrinking operations may contribute to our account. In fact, if an arc gets shrunk k times during the algorithm, this may represent additional k assignments to p , in the worst case. Furthermore, there might be $O(n)$ distinct arcs, each of them covering C , together with A_j . However, the next assertion assures that even in this case, the number of times A_j gets shrunk is low.

Claim. Any arc of $(C, \mathcal{A})$ may be shrunk at most twice during the entire algorithm.

Proof. Let A_j be an arc of $(C, \mathcal{A})$. Clearly, if A_j together with some other arc does not cover C , then A_j is not shrunk at all. Otherwise, consider the model $(C, \mathcal{A})$ at the beginning of the computation *NORMAL*(k), where A_j gets shrunk for the first time, if any. Denote by $A_{j_1}, \dots, A_{j_l}$ the arcs of $(C, \mathcal{A})$, each of them covers C , together with A_j . This implies that all extreme points of the arcs $A_{j_1}, \dots, A_{j_l}$ belong to (s_j, t_j) . Furthermore, s_{j_k} precedes t_{j_k} in A_j , $1 \leq k \leq l$. Without loss of generality, the starting points $s_{j_1}, \dots, s_{j_l}$ are in the circular ordering. Clearly, the arc of least index among $A_j, A_{j_1}, \dots, A_{j_l}$ corresponds to one of this collection which is first handled by the algorithm. Let $i = \min\{j, j_1, \dots, j_l\}$. The following are the alternatives for i :

- Case 1.* $i = j$. We know that *NORMAL*(j) does not shrink A_j . Furthermore, Theorem 8 implies that *NORMAL*(j) would shrink all arcs $A_{j_1}, \dots, A_{j_l}$. Consequently, A_j never gets shrunk.
- Case 2.* $i = j_1$. Because A_j and A_{j_1} cover C , during *NORMAL*(j_1) the assignment $p := t_j$ necessarily occurs, since Theorem 8 assures that *NORMAL*(j_1) shrinks A_j . Then t_j moves to the position t_j^* , becoming the predecessor of s_1 . The arc A_j^* so obtained does not intersect any of $A_{j_1}, \dots, A_{j_l}$, and consequently, it cannot get shrunk. Consequently, A_j is shrunk only once during the process.
- Case 3.* $i = j_k$, for some $1 < k \leq l$. Then *NORMAL*(j_k) moves t_j to t_j^* , where t_j^* , this time, is the new predecessor of s_{j_k} . Following the circular ordering of $(C, \mathcal{A})$, the first subsequent computation which can possibly shrink A_j is *NORMAL*(j_1). By Case 2, the latter would shrink A_j , exactly once more. Consequently, A_j gets shrunk twice.

The conclusion is that any arc may get shrunk at most twice during the process. $\square$

Consequently, the shrinking operations contribute at most $2n$ units to our account. Therefore the algorithm performs less than $6n$ iterations of the *repeat* block. Each of these iterations requires no more than constant time. That is, the overall complexity is $O(n)$. $\square$

5. Recognition and model construction. In this section, we describe the algorithms for recognizing UCA graphs and for constructing UCA models. Also, we determine the size of the obtained model by showing that its extremes correspond to integers of size $O(n)$.

The algorithm is as follows. Let G be the given graph.

1. Verify if G is a PCA graph, using the algorithm [1]. In the affirmative case, let $(C, \mathcal{A})$ be the PCA model so obtained. In the negative case, G is not UCA.
2. Transform $(C, \mathcal{A})$ into a normal model $(C', \mathcal{A}')$, using the algorithm of section 4.
3. Construct the segment digraph D of $(C', \mathcal{A}')$. Find the connected components of D . Verify if all connected components are strongly connected. In the negative case, report that G is not UCA and stop. Otherwise, report that G is UCA and construct a UCA model $(C'', \mathcal{A}'')$ of it by running Step 4.
4. Construct the segment network of $(C', \mathcal{A}')$ by assigning to each edge $e_j \in E(D)$ a unit lower capacity and an unbounded upper capacity. Then find the flow values of a feasible circulation W of D , using the algorithm of section 3. Let $p'_1, \dots, p'_{2n}$ be the extremes of $(C', \mathcal{A}')$, in circular ordering, with p_1 corresponding to the start point of the first arc of the model. Then construct $(C'', \mathcal{A}'')$ as follows. The length of C'' is $\sum_{e_j \in E(D)} w_j$. The extremes $p''_1, \dots, p''_{2n}$ of $(C'', \mathcal{A}'')$ are defined by the following conditions: p''_1 is an arbitrary point of C'' , and the segment (p''_j, p''_{j+1}) has length w_j , $1 \leq j < 2n$.

The correctness follows basically from Theorems 4, 7, and 8, and Corollary 6. We evaluate the complexity of the recognition algorithm. Step 1 requires $O(n + m)$ time [1]. Step 2 takes $O(n)$ time by Theorem 9. As for Step 3, note that for constructing the segment digraph, we need to construct the full and reduced systems of $(C', \mathcal{A}')$. The first of them is a family of $n - 1$ equations, whose variables are segment lengths. In each side of the equations, the segment lengths are consecutive. Consequently, we need $4n - 4$ indices to represent these equations, overall. On the other hand, in the reduced system, each segment length appears at most twice. That is, it requires at most $4n$ indices. The segment digraph has n vertices and at most $2n$ edges. There is no difficulty to construct D in $O(n)$ time. The connected and strongly connected components can be constructed in linear time [11]. Step 4 also requires linear time in the size of D , according to section 3. Consequently, the overall complexity of algorithm is $O(n + m)$. Furthermore, Steps 2, 3, and 4 require $O(n)$ time, meaning that the complexity of the algorithm reduces to $O(n)$, if the input is a PCA model of G , with ordered extremes of the arcs.

Finally, we verify the size of the UCA model obtained by the algorithm.

COROLLARY 10. *Let $(C, \mathcal{A})$ be a UCA model constructed by the algorithm. Then the extremes of $(C, \mathcal{A})$ correspond to integers of size $< 4n$.*

Proof. Let $p_1, \dots, p_{2n}$ be the extremes of $(C, \mathcal{A})$ in circular ordering. We associate the segment length l_j to the extreme p_{j+1} , $1 \leq j < 2n$. From Step 4 of the algorithm, we know that the length l_j of each segment of $(C, \mathcal{A})$ equals the flow value w_j of the corresponding edge e_j of D . On the other hand, we know from section 3 that $w_j < 2m$. Since $m < 2n$, it follows that $l_j < 4n$. $\square$

6. Conclusion. The proposed algorithms for recognizing UCA graphs and constructing UCA models terminate within $O(n + m)$ time. Furthermore, the complexity reduces to $O(n)$ when the input is a PCA model. It should be noted that the algorithms require the extreme points of the given model to be ordered. If the input is a graph given by its vertices and edges, then there is a preprocessing step, which consists of running the algorithm by Deng, Hell, and Huang [1], for constructing a PCA model of the graph. The algorithm [1] can be implemented so as to construct a PCA model with ordered extreme points within $O(n + m)$ time. On the other hand,

if the input is already a PCA model and its extreme points are not ordered, then we need an additional $O(n \log n)$ steps for ordering.

The UCA model constructed by the proposed algorithm is such that its extreme points correspond to integers of size $O(n)$. Regarding this feature, we leave the four related questions below.

Given a UCA graph G , find a UCA model whose extremes of the arcs correspond to integers, satisfying

- (i) the maximum segment length is minimized.
- (ii) the circle length is minimized.

More generally, for given n , find the least l , such that any UCA graph with n vertices admits a UCA model whose extremes of the arcs correspond to integers, such that

- (iii) the maximum segment length is $\leq l$.
- (iv) the circle length is $\leq l$.

Corollary 10 implies that $l < 4n$ for Question (iii), and that $l < 4n^2$ for Question (iv).

Acknowledgments. We are grateful to Professor Pavol Hell and an anonymous referee for their most valuable suggestions and comments, which lead to improvements in different parts of this paper.

REFERENCES

- [1] X. DENG, P. HELL, AND J. HUANG, *Linear time representation algorithms for proper circular-arc graphs and proper interval graphs*, SIAM J. Comput., 25 (1996), pp. 390–403.
- [2] G. DURÁN, A. GRAVANO, R. M. MCCONNELL, J. P. SPINRAD, AND A. TUCKER, *Polynomial time recognition of unit circular-arc graphs*, J. Algorithms, 58 (2006), pp. 67–78.
- [3] E. ESCHEN AND J. P. SPINRAD, *An $O(n^2)$ algorithm for circular-arc graph recognition*, in Proceedings of the 4th Annual ACM-SIAM Symposium on Discrete Algorithms, Austin, TX, ACM, New York, SIAM, Philadelphia, 1993, pp. 128–137.
- [4] M. C. GOLUMBIC, *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, New York, 1980.
- [5] A. J. HOFFMAN, *Some recent applications of the theory of linear inequalities to extremal combinatorial analysis*, in Proceedings of the Symposia in Applied Mathematics, American Mathematical Society, Providence, RI, 1960, pp. 113–117.
- [6] W. L. HSU, *$O(mn)$ algorithms for the recognition and isomorphism problems on circular-arc graphs*, SIAM J. Comput., 24 (1995), pp. 411–439.
- [7] M. C. LIN AND J. L. SZWARCFITER, *Efficient construction of unit circular-arc models*, in Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms, Miami, FL, 2006, pp. 309–315.
- [8] R. M. MCCONNELL, *Linear-time recognition of circular-arc graphs*, Algorithmica, 37 (2003), pp. 93–147.
- [9] A. SCHRIJVER, *Combinatorial Optimization, Vol A: Polyhedra and Efficiency*, Springer-Verlag, Berlin, 2003.
- [10] J. P. SPINRAD, *Efficient Graph Representations*, Fields Institute Monographs 19, AMS, Providence, RI, 2003.
- [11] R. E. TARJAN, *Depth-first search and linear graph algorithms*, SIAM J. Comput., 1 (1972), pp. 146–160.
- [12] A. TUCKER, *Characterizing circular-arc graphs*, Bull. Amer. Math. Soc., 76 (1970), pp. 1257–1260.
- [13] A. TUCKER, *Structure theorems for some circular-arc graphs*, Discrete Math., 7 (1974), pp. 167–195.
- [14] A. TUCKER, *An efficient test for circular-arc graphs*, SIAM J. Comput., 9 (1980), pp. 1–24.

ASYMPTOTIC DETERMINATION OF EDGE-BANDWIDTH OF MULTIDIMENSIONAL GRIDS AND HAMMING GRAPHS*

REZA AKHTAR[†], TAO JIANG[†], AND ZEVI MILLER[†]

Abstract. The edge-bandwidth $B'(G)$ of a graph G is the bandwidth of the line graph of G . More specifically, for any bijection $f : E(G) \rightarrow \{1, 2, \dots, |E(G)|\}$, let $B'(f, G) = \max\{|f(e_1) - f(e_2)| : e_1 \text{ and } e_2 \text{ are incident edges of } G\}$, and let $B'(G) = \min_f B'(f, G)$. We determine asymptotically the edge-bandwidth of d -dimensional grids P_n^d and of the Hamming graph K_n^d , the d -fold Cartesian product of K_n . Our results are as follows. (i) For fixed d and $n \rightarrow \infty$, $B'(P_n^d) = c(d)dn^{d-1} + O(n^{d-\frac{3}{2}})$, where $c(d)$ is a constant depending on d , which we determine explicitly. (ii) For fixed even n and $d \rightarrow \infty$, $B'(K_n^d) = (1 + o(1))\sqrt{\frac{d}{2\pi}}n^d(n-1)$. Our results extend recent results by Balogh, Mubayi, and Pluhár [*Theoret. Comput. Sci.*, 359 (2006), pp. 43–57], who determined $B'(P_n^2)$ asymptotically as a function of n and $B'(K_2^d)$ asymptotically as a function of d .

Key words. bandwidth, edge-bandwidth, Hamming graph, grid, boundary, isoperimetric problem

AMS subject classification. 05C78

DOI. 10.1137/060660291

1. Introduction. Let $G = (V(G), E(G))$ be a simple graph on n vertices. A labeling f is a bijection of $V(G)$ to $\{1, \dots, n\}$. When there is no ambiguity, we will simply write $B(f)$ for $B(f, G)$. The *bandwidth* of f is

$$B(f, G) := \max\{|f(u) - f(v)| : uv \in E(G)\}.$$

The *bandwidth* $B(G)$ of G is

$$B(G) := \min_f \{B(f, G)\}.$$

The notion was introduced by Harper in his influential paper [13], in which he determined the bandwidth of the d -dimensional hypercube by solving the corresponding vertex isoperimetric problem in the hypercube. There are several motivations for studying the bandwidth problem: sparse matrix computations, representing data structures by linear arrays, VLSI layouts, mutual simulations of interconnection networks, and minimizing the effects of noise in the multichannel communication of data (see [10, 11, 24, 7]). The bandwidth problem is NP-hard and is inapproximable by any multiplicative constant even for trees [28]. Bandwidths are known only for a few families of graphs including hypercubes [13], multidimensional grids [8, 22, 23], complete trees [16], and various meshlike graphs (see [16, 17, 20]).

The edge-bandwidth was introduced by Hwang and Lagarias [18]. Here we label the edges instead of the vertices, and the *bandwidth* of an edge-labeling f of a graph

*Received by the editors May 18, 2006; accepted for publication (in revised form) August 17, 2007; published electronically March 20, 2008. The U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. Copyright is owned by SIAM to the extent not limited by these rights.

<http://www.siam.org/journals/sidma/22-2/66029.html>

[†]Department of Mathematics and Statistics, Miami University, Oxford, OH 45056 (reza@calico.mth.muohio.edu, jiangt@muohio.edu, millerz@muohio.edu). The research of the second author was supported in part by the National Security Agency under grant H98230-07-1-0027.

G is

$$B'(f, G) := \max\{|f(uv) - f(vw)| : uv, vw \in E(G)\}.$$

In other words it is the maximum difference of labels between a pair of incident edges. When there is no ambiguity, we will write $B'(f)$ for $B(f, G)$. The edge-bandwidth of a graph G is

$$B'(G) := \min_f B'(f, G).$$

Naturally, $B'(G) = B(L(G))$, where $L(G)$ is the line graph of G . In [19], Jiang et al. reintroduced the notion of edge-bandwidth and studied the relationship between $B(G)$ and $B'(G)$. They determined the edge-bandwidth of caterpillars, the complete graph K_n , and the balanced complete bipartite graph $K_{n,n}$. Gupta [12] pointed out that the inequality $B(T) \leq B'(T) \leq 2B(T)$ for a tree T obtained in [19] together with Unger's inapproximation result [28] for bandwidth imply that determining the edge-bandwidth is also NP-hard.

Recently, there has been an increase of interest in the study of edge-bandwidth. Calamoneri, Massini, and Vrto [9] obtained tight bounds on the edge-bandwidth of complete k -ary trees and bounds on the edge-bandwidth of the hypercube and butterfly graphs. Balogh, Mubayi, and Pluhár [5] subsequently obtained asymptotically tight bounds on the edge-bandwidth of two-dimensional grids and tori, the Cartesian product of two cliques, and the hypercube. Sharpening the result of Balogh, Mubayi, and Pluhár [5] on the two-dimensional grids and tori while confirming a conjecture of Calamoneri, Massini, and Vrto [9], Pikhurko and Wojciechowski [27] showed that the edge-bandwidth of an m by n grid, where $m \geq n$, is $2n - 1$. They also showed that the edge-bandwidth of an m by n torus, where $m \geq n$, is between $4n - 5$ and $4n - 1$. In an unpublished manuscript, Akhtar, Jiang, and Pritikin have independently shown that the edge-bandwidth of an m by n grid, where $m \geq n$, is between $2n - 2$ and $2n - 1$ and that the edge-bandwidth of an m by n torus, where $m \geq n$, is between $4n - 5$ and $4n - 1$.

In this paper, we determine the edge-bandwidth of the d -dimensional grids P_n^d asymptotically when d is fixed and $n \rightarrow \infty$, and we obtain lower and upper bounds on the edge-bandwidth of the Hamming graph K_n^d . When n is a fixed positive even integer and $d \rightarrow \infty$, our lower and upper bounds match asymptotically.

The Cartesian product of graphs G and H , denoted by $G \square H$, is the graph with vertex set $V(G \square H) = \{(u, v) : u \in V(G), v \in V(H)\}$ specified by putting (u, v) adjacent to (u', v') if and only if (i) $u = u'$ and $vv' \in e(H)$ or (ii) $v = v'$ and $uu' \in E(G)$. The d -fold Cartesian product $G \square G \square \cdots \square G$ is denoted by G^d . Let P_n denote a path on n vertices. The d -fold Cartesian product P_n , denoted by P_n^d , is also known as the d -dimensional grid. Equivalently, we can view P_n^d as a graph whose vertices are vectors $\langle u_1, u_2, \dots, u_d \rangle$ of length d where, for all i , $u_i \in \{0, 1, \dots, n-1\}$ and any two vertices $\langle x_1, x_2, \dots, x_d \rangle$ and $\langle y_1, y_2, \dots, y_d \rangle$ are adjacent if and only if there exists $i \in \{1, 2, \dots, d\}$ such that $|x_i - y_i| = 1$ and $x_j = y_j$ for all $j \neq i$. Given a vertex $x = \langle x_1, x_2, \dots, x_d \rangle$, we define the weight of x by $wt(x) = x_1 + x_2 + \cdots + x_d$. For each $r \in \{1, 2, \dots, n\}$, let $L(n, d, r) = \{x \in V(P_n^d) : wt(x) = r\}$. Let $l(n, d, r) = |L(n, d, r)|$ and $l^*(n, d) = \max\{l(n, d, r) : 0 \leq r \leq (n-1)d\}$. We determine $B'(P_n^d)$ asymptotically as a function of n .

THEOREM 1.1. *Let d be a fixed positive integer. We have*

$$B'(P_n^d) = (1 + o(1))dl^*(n, d) = c(d)dn^{d-1} + O(n^{d-\frac{3}{2}}),$$

where $c(d)$ is a constant depending on d , given by $c(d) = \sum_{j=0}^{\lfloor d/2 \rfloor} (-1)^j \binom{d}{j} \frac{(d-2j)^{d-1}}{2^{d-1}(d-1)!}$. Furthermore, $\frac{1}{2\sqrt{d}} \leq c(d) \leq \frac{2\sqrt{11}}{\sqrt{d}}$.

A simple calculation shows that $c(2) = 1$. Hence $B'(P_n^2) = (1+o(1))2n$, which was obtained by Balogh, Mubayi, and Pluhár [5]. For another example, $c(3) = \frac{3}{4}$, yielding $B'(P_n^3) = (1+o(1))\frac{9}{4}n^2$. Theorem 1.1 determines asymptotically the edge-bandwidth of grids of arbitrary dimension.

The d -fold Cartesian product of the complete graph K_n , denoted by K_n^d , is also called the Hamming graph. We obtain lower and upper bounds on $B'(K_n^d)$ for fixed n as a function of d . When n is even, our lower and upper bounds match asymptotically, yielding the following theorem.

THEOREM 1.2. *Let n be a fixed positive even integer. We have*

$$B'(K_n^d) = (1 + o(1)) \frac{\sqrt{d}}{\sqrt{2\pi}} n^d (n-1).$$

Throughout the paper we will use standard enumeration estimates, as found in [25].

2. General bounds. The standard techniques for obtaining lower bounds on bandwidth use isoperimetric inequalities. Many vertex and edge isoperimetric problems have been considered in the literature. Given a graph G and a set $S \subseteq V(G)$, let

$$\partial(S) = \{v \in V(G) - S : \exists u \in S \text{ such that } uv \in E(G)\}.$$

We call $\partial(S)$ the (vertex) *boundary* of S . In other words, $\partial(S) = N_G(S) - S$. Given an optimal numbering f of $V(G)$, let S be the set of vertices receiving labels $1, 2, \dots, k$. Then the highest label assigned to a vertex in $\partial(S)$ is at least $k + |\partial(S)|$. Let v be the vertex with the highest label in $\partial(S)$. It has a neighbor u in S , whose label is at most k . So $|f(v) - f(u)| \geq |\partial(S)|$, which implies $B(G) = B(f, G) \geq |\partial(S)|$. Similarly, consider a vertex x in $\partial(V - S)$ with the smallest label. Its label is at most $k - |\partial(V - S)| + 1$. It has a neighbor y in $V - S$, whose assigned label is at least $k + 1$. So $|f(y) - f(x)| \geq |\partial(V - S)|$. Thus, $B(G) = B(f, G) \geq |\partial(V - S)|$. Therefore, we have $B(G) \geq \max\{|\partial(S)|, |\partial(V - S)|\}$. This yields the following proposition.

PROPOSITION 2.1 (see [13]). *Let G be a graph and k an integer, where $0 \leq k \leq |V(G)|$. Then*

$$B(G) \geq \min_{S \subseteq V(G), |S|=k} \max\{|\partial(S)|, |\partial(V - S)|\}.$$

For each k , let $L_k(G) = \min_{S \subseteq V(G), |S|=k} |\partial(S)|$. By Proposition 2.1, $B(G) \geq L_k(G)$. Since this holds for each k with $0 \leq k \leq |V(G)|$, we have $B(G) \geq \max_k L_k(G)$. This lower bound $\max_k L_k(G)$ for $B(G)$ is often referred to as the *Harper bound*. In general, the Harper bound need not be sharp, and calculating it is difficult (NP-hard).

When the Harper bound is not very useful, it is sometimes useful to consider the iterated boundary (shadow) instead. Given a nonnegative integer q , let

$$\partial^{(\leq q)}(S) = \{v \in V(G) - S : v \text{ is at distance at most } q \text{ from some vertex in } S\}.$$

Hence, in particular, $\partial(S) = \partial^{(\leq 1)}(S)$. Consider an optimal numbering f of $V(G)$. Let S be the set of vertices receiving labels $1, 2, \dots, k$. Let q be any integer such that $1 \leq q \leq n$. Let v be a vertex in $\partial^{(\leq q)}(S)$ with the highest label. Then $f(v) \geq$

$k + |\partial^{\leq q}(S)|$. Vertex v is at distance at most q from some vertex u in S . Note that $f(u) \leq k$. Hence for some edge on a shortest u, v -path, the difference between the f -labels of its two end points is at least $|f(u) - f(v)|/q \geq |\partial^{\leq q}(S)|/q$. This yields the following proposition.

PROPOSITION 2.2. *Let G be a graph and k an integer, where $0 \leq k \leq |V(G)|$. Then*

$$B(G) \geq \min_{S \subseteq V(G), |S|=k} \max_{1 \leq q \leq n} \left| \frac{\partial^{(\leq q)}(S)}{q} \right|.$$

Our discussions above apply similarly to a set of edges. We define the boundary of a set $S' \subseteq E(G)$ of edges in G by

$$\partial(S') = \{e \in E(G) - S' : \exists e' \in S' \text{ such that } e \text{ and } e' \text{ are incident}\}.$$

The iterated boundary for S' is then given, for $q \geq 1$, by

$$\partial^{(\leq q)}(S') = \{e \in E(G) - S' : e \text{ is at distance at most } q \text{ from some edge of } S'\}.$$

We see that $\partial(S') = \partial^{(\leq 1)}(S)$ and $\partial^{(\leq q)}(S') = \partial^{(\leq q-1)}(S) \cup \partial(\partial^{(\leq q-1)}(S'))$. The edge analogues of Propositions 2.1 and 2.2 are then obtained by replacing $B(G)$ by $B'(G)$ and $V(G)$ by $E(G)$.

3. The weight function in multidimensional grids. Recall that $V(P_n^d) = \{\langle x_1, x_2, \dots, x_d \rangle : x_i \in \{0, 1, \dots, n-1\} \text{ for all } i = 1, 2, \dots, d\}$. Two vertices $\langle x_1, x_2, \dots, x_d \rangle$ and $\langle y_1, y_2, \dots, y_d \rangle$ are adjacent if they differ by 1 in one coordinate and agree in all other coordinates. Again, the weight $wt(x)$ of a vertex $x = \langle x_1, x_2, \dots, x_d \rangle$ is defined by $wt(x) = x_1 + x_2 + \dots + x_d$. Given positive integers n, d and an integer r , with $0 \leq r \leq (n-1)d$, let $L(n, d, r) = \{x \in V(P_n^d) : wt(x) = r\}$. Let $l(n, d, r) = |L(n, d, r)|$. Let $l^*(n, d) = \max\{l(n, d, r) : 0 \leq r \leq (n-1)d\}$. It is easy to see that $l(n, d, r)$ is the number of integer solutions to the equation $x_1 + x_2 + \dots + x_d = r$, where $0 \leq x_i \leq n-1$ for each $i \in [d]$. By considering the value of x_1 one can easily derive the following recurrence relation on $l(n, d, r)$, which also appeared in [26].

PROPOSITION 3.1. *Let n, d be positive integers and r an integer. We have $l(n, 1, r) = 1$ if $0 \leq r \leq n-1$ and $l(n, 1, r) = 0$ otherwise. For all d and r with $d \geq 2$ and $0 \leq r \leq d(n-1)$,*

$$l(n, d, r) = \sum_{j=0}^{n-1} l(n, d-1, r-j),$$

and for other values of r we have $l(n, d, r) = 0$.

It is straightforward to derive an exact formula for $l(n, d, r)$ using a generating function. This was done in [26], and we include a short proof for completeness.

PROPOSITION 3.2. *Let n, d, r be positive integers. We have*

$$l(n, d, r) = \sum_{j=0}^{\lfloor \frac{r}{d} \rfloor} (-1)^j \binom{d}{j} \binom{r-jn+d-1}{d-1}.$$

Proof. By prior discussion, $l(n, d, r)$ is the number of integer solutions to $x_1 + x_2 + \dots + x_d = r$, with $0 \leq x_i \leq n-1$ for each $i \in [d]$. For fixed n, d , the generating function for $l(n, d, r)$ is

$$g(x) = (1 + \dots + x^{n-1})^d = \left(\frac{1-x^n}{1-x} \right)^d = (1-x^n)^d \left(\frac{1}{1-x} \right)^d.$$

As $l(n, d, r)$ equals the coefficient of x^r in the above expansion, we have

$$\begin{aligned} l(n, d, r) &= \sum_{j=0}^{\infty} (-1)^j \binom{d}{j} \binom{r - jn + d - 1}{d - 1} \\ &= \sum_{j=0}^{\lfloor \frac{r}{n} \rfloor} (-1)^j \binom{d}{j} \binom{r - jn + d - 1}{d - 1}. \quad \square \end{aligned}$$

The function $l(n, d, r)$ is well-studied in the theory of posets as the *rank number* in the poset of divisors of a number. Consider a positive integer m with prime factorization $m = p_1^{k_1} p_2^{k_2} \dots p_d^{k_d}$, where the p_i are distinct primes and $k_i \geq 1$ for each i . The *rank* of m is $K = k_1 + \dots + k_d$, and $N_r(m)$ is the number of divisors of m of rank r . It is easy to see that, in the case $k_1 = k_2 = \dots = k_d = n - 1$, $N_r(m)$ is precisely $l(n, d, r)$. Chapter 4 of Anderson [3] gives a detailed discussion about the function $N_r(m)$. In particular, we have the following proposition.

PROPOSITION 3.3 (see [3, Chapter 4]). *Let $K = (n - 1)d$. Then the following apply:*

1. $l(n, d, i) = l(n, d, K - i)$ for each i with $0 \leq i \leq K$.
2. For fixed n and d , $l(n, d, i)$ is strictly increasing in i for $i \leq \lfloor \frac{K}{2} \rfloor$ and strictly decreasing in i for $i \geq \lceil \frac{K}{2} \rceil$.

By Proposition 3.3, for fixed n and d , $l(n, d, r)$ is a symmetric and unimodal function of r on $\{0, 1, \dots, (n - 1)d\}$ with maximum value at $r = \lfloor \frac{(n-1)d}{2} \rfloor$ and at $r = \lceil \frac{(n-1)d}{2} \rceil$. Thus, $l^*(n, d) = l(n, d, \lfloor \frac{(n-1)d}{2} \rfloor)$. Setting $r = \lfloor \frac{(n-1)d}{2} \rfloor$ in the formula in Proposition 3.2 for $l(n, d, r)$ then gives us a formula for $l^*(n, d)$. When d is fixed and n tends to infinity, the leading term is a multiple of n^{d-1} whose leading coefficient can be expressed exactly as a sum. We then give a closed form estimate of this leading coefficient using earlier results of Anderson [1, 2] given below.

THEOREM 3.4 (see [2]). *Let $k_1, k_2, \dots, k_d$ be nonnegative integers. Let $K = k_1 + \dots + k_d$. Let s denote the number of integer solutions to the equation*

$$x_1 + x_2 + \dots + x_d = \frac{K}{2}, \text{ where } \forall i \ 0 \leq x_i \leq k_i.$$

Let $A = \frac{1}{3} \sum_{i=1}^d k_i(k_i + 2)$ and $\tau = \prod_{i=1}^d (1 + k_i)$. Then there exist positive constants C_1 and C_2 such that

$$C_1 \frac{\tau}{\sqrt{A}} \leq s \leq C_2 \frac{\tau}{\sqrt{A}}.$$

Furthermore, we can take $C_2 = \sqrt{11}$, and for any small $\epsilon > 0$ we can take $C_1 = \frac{1}{\sqrt{3}} - \epsilon$ when K is sufficiently large.

COROLLARY 3.5. *Let n, d be positive integers, where $n \geq 2$. There exist positive constants C_1 and C_2 such that*

$$C_1 \frac{n^{d-1}}{\sqrt{d}} \leq l^*(n, d) \leq C_2 \frac{n^{d-1}}{\sqrt{d}}.$$

Furthermore, we can take $C_2 = 2\sqrt{11}$, and for any small $\epsilon > 0$ we can take $C_1 = 1 - \epsilon$ when n is sufficiently large.

Proof. By our earlier discussion, $l^*(n, d) = l(n, d, \lfloor \frac{(n-1)d}{2} \rfloor)$ is the number of integer solutions to the equation $x_1 + x_2 + \cdots + x_d = \lfloor \frac{(n-1)d}{2} \rfloor$, where $0 \leq x_i \leq n-1$ for each i . We apply Theorem 3.4. Here $k_i = n-1$ for each $i \in [d]$. So $A = \frac{d}{3}(n^2-1)$ and $\tau = n^d$. Also, $s = l^*(n, d)$. The claim follows. $\square$

We now give the exact formula for $l^*(n, d)$ together with an asymptotic formula for it. There is an exact formula for the coefficient of the leading term in the form of a summation; this coefficient can be bounded using Corollary 3.5. Before we proceed, we need the following routine estimation of binomial coefficients. Recall that if x is a real number and k is an integer, then $\binom{x}{k} := \frac{x(x-1)\cdots(x-k+1)}{k!}$. Also, it is straightforward to check that $e^{-x} \geq 1-x$ for all real x and $1-x \geq e^{-2x}$ for x with $0 < x < \frac{1}{2}$.

LEMMA 3.6. *Let k be a positive integer and N, m nonnegative real numbers such that $N > \max\{2k, mk\}$. Then $\frac{N^k}{k!}(1 - \frac{2k^2}{N}) \leq \binom{N+m}{k} \leq \frac{N^k}{k!}(1 + \frac{2mk}{N})$. So $|\binom{N+m}{k} - \frac{N^k}{k!}| \leq 2(m+2)N^{k-1}$.*

Proof. For the lower bound, we have $\binom{N+m}{k} / \frac{N^k}{k!} > \frac{(N+m-k)^k}{k!} / \frac{N^k}{k!} > \frac{(N-k)^k}{k!} / \frac{N^k}{k!} = (1 - \frac{k}{N})^k$. Since $0 < \frac{k}{N} < \frac{1}{2}$, $(1 - \frac{k}{N})^k \geq e^{-2\frac{k}{N}k} \geq 1 - \frac{2k^2}{N}$.

For the upper bound, we have $\binom{N+m}{k} / \frac{N^k}{k!} < \frac{(N+m)^k}{k!} / \frac{N^k}{k!} = (\frac{N+m}{N})^k = (1 + \frac{m}{N})^k$. Since $(1 + \frac{1}{x})^x < e$ when $x > 0$, we have $(1 + \frac{m}{N})^{\frac{N}{m}} < e$ and hence $(1 + \frac{m}{N})^k < e^{\frac{mk}{N}}$. Since $\frac{mk}{N} < 1$ and $e^x < 1 + 2x$ when $0 < x < 1$, we have $e^{\frac{mk}{N}} < 1 + \frac{2mk}{N}$. Therefore, $\binom{N+m}{k} / \frac{N^k}{k!} < 1 + \frac{2mk}{N}$. The last statement follows readily from the lower and upper bounds. $\square$

THEOREM 3.7. *Let n, d be positive integers. Then*

$$l^*(n, d) = \sum_{j=0}^{\lfloor \frac{d}{2} \rfloor} (-1)^j \binom{d}{j} \binom{\lfloor \frac{(d-2j)n}{2} + \frac{d}{2} \rfloor - 1}{d-1}.$$

We have $l^*(n, 1) = 1, l^*(n, 2) = 2$. For fixed $d \geq 3$, as $n \rightarrow \infty$, we have

$$l^*(n, d) = c(d)n^{d-1} + O(n^{d-2}),$$

where $c(d)$ is a constant depending on d , given by

$$c(d) = \sum_{j=0}^{d/2} (-1)^j \binom{d}{j} \frac{(d-2j)^{d-1}}{2^{d-1}(d-1)!}.$$

Also, $\frac{1}{2\sqrt{d}} \leq c(d) \leq \frac{2\sqrt{11}}{\sqrt{d}}$.

Proof. The exact formula for $l^*(n, d)$ given above is obtained by setting $r = \lfloor \frac{(n-1)d}{2} \rfloor$ in the formula for $l(n, d, r)$ in Proposition 3.2. For fixed d , we derive an asymptotic formula for $l^*(n, d)$ as a function of n . Fix j with $d-2j > 0$. Let $N = \frac{(d-2j)n}{2}$ and $m = (\lfloor \frac{(d-2j)n}{2} + \frac{d}{2} \rfloor - 1) - N$. Then $m \leq d/2$. For sufficiently large n , since d is fixed, we have $N > \max\{2(d-1), m(d-1)\}$. By applying Lemma 3.6, we have

$$\left| \binom{\lfloor \frac{(d-2j)n}{2} + \frac{d}{2} \rfloor - 1}{d-1} - \frac{(d-2j)^{d-1}}{2^{d-1}(d-1)!} \cdot n^{d-1} \right| \leq 2(m+2)n^{d-2} \leq (d+4)n^{d-2}.$$

So

$$\begin{aligned} l^* &= \sum_{j=0}^d (-1)^j \binom{d}{j} \binom{\lfloor \frac{dn-2jn}{2} + \frac{d}{2} \rfloor - 1}{d-1} \\ &= \sum_{j=0}^{\lfloor d/2 \rfloor} (-1)^j \binom{d}{j} \cdot \frac{(d-2j)^{d-1}}{2^{d-1}(d-1)!} \cdot n^{d-1} + O(n^{d-2}). \end{aligned}$$

Let $c(d) = \sum_{j=0}^{\lfloor d/2 \rfloor} (-1)^j \binom{d}{j} \frac{(d-2j)^{d-1}}{2^{d-1}(d-1)!}$. We have $l^*(n, d) = c(d)n^{d-1} + O(n^{d-2})$. Note that $c(d)$ is a constant depending only on d , and, by Corollary 3.5, $\frac{1}{2\sqrt{d}} \leq c(d) \leq \frac{2\sqrt{11}}{\sqrt{d}}$. $\square$

Remark 3.8. When n is fixed and $d \rightarrow \infty$, by using Laplace's method one can show that $l^*(n, d) = \sqrt{\frac{6}{\pi d} \left(\frac{n^2}{n^2-1}\right)} \cdot n^{d-1} + o(n^{d-1})$. See the concluding remarks section for further discussion.

Next, we show that, for all r relatively close to $\lfloor \frac{(n-1)d}{2} \rfloor$ and $\lceil \frac{(n-1)d}{2} \rceil$, $l(n, d, r)$ is close to the maximum value $l^*(n, d)$. This property is important to establishing our lower bound on the edge-bandwidth of P_n^d in the next section.

LEMMA 3.9. *Let n, d, r be integers such that $n, d \geq 2$. If $|r - \lfloor \frac{(n-1)d}{2} \rfloor| = t$, where $0 \leq t \leq \lfloor \frac{n}{2} \rfloor - 1$, then $l(n, d, r) \geq (1 - \frac{t}{n-1})l^*(n, d)$.*

Proof. Let $M = \lfloor \frac{(n-1)d}{2} \rfloor$. By Proposition 3.3, $l^*(n, d) = l(n, d, M)$. For convenience, let $f(j) = l(n, d-1, j)$. By Proposition 3.3, $f(j)$ is symmetric on $[0, (n-1)(d-1)]$, and it is unimodal with peak(s) at $M' = \lfloor \frac{(n-1)(d-1)}{2} \rfloor$ and $M'' = \lceil \frac{(n-1)(d-1)}{2} \rceil$. By Proposition 3.1,

$$l^*(n, d) = l(n, d, M) = \sum_{i=M-(n-1)}^M f(i).$$

Consider the interval of integers $I = [M - (n-1), M]$. Assume first that n is odd. In this case $M' = M'' = \frac{(n-1)(d-1)}{2}$ is the only element at the center of I . The function $f(j)$ is symmetric and unimodal on I with a single peak at $\frac{(n-1)(d-1)}{2}$. Let A denote the sum of $f(j)$ over the first t elements of I , B the sum of $f(j)$ over the last t elements of I , and C the sum of $f(j)$ over the middle $n-2t$ elements of I . The symmetry and unimodality of $f(j)$ imply that $A = B$ and $\frac{C}{A} \geq \frac{n-2t}{t}$. So $l^*(n, d) = A + B + C \geq (2 + \frac{n-2t}{t})A$, which implies $A = B \leq \frac{t}{n} l^*(n, d)$. Now, by Proposition 3.1, $l(n, d, M+t)$ is the sum of $f(j)$ over $I+t$, the translation of I by t . So clearly, $l(n, d, M+t) \geq l(n, d, M) - A \geq (1 - \frac{t}{n})l^*(n, d)$. Similarly, $l(n, d, M-t) \geq l(n, d, M) - B \geq (1 - \frac{t}{n})l^*(n, d)$.

When n is even, we consider the subcases depending on whether d is even or odd. In each subcase, similar analysis shows that $A, B \geq (1 - \frac{t}{n-1})l^*(n, d)$, and hence $l(n, d, M+t) \geq (1 - \frac{t}{n-1})l^*(n, d)$ and $l(n, d, M-t) \geq (1 - \frac{t}{n-1})l^*(n, d)$. $\square$

4. The edge-bandwidth of a multidimensional grid. Bollobás and Leader [8] solved the vertex isoperimetric problem in grids. We will use their result to obtain asymptotically tight bounds on the edge-bandwidth of a multidimensional grids. Our result extends that of Balogh, Mubayi, and Pluhár on two-dimensional grids to grids of any dimension.

DEFINITION 4.1. The simplicial order on $V(P_n^d)$ is defined by $x < y$ if either $wt(x) < wt(y)$ or $wt(x) = wt(y)$ and $x_s > y_s$, where $s = \min\{t : x_t \neq y_t\}$.

Bollobás and Leader [8] showed that for any k , with $0 \leq k \leq |V(P_n^d)|$, the initial segment of length k in the simplicial order has the smallest boundary among all sets of k vertices. For each r with $0 \leq r \leq (n-1)d$, let $B(n, d, r) = \{x \in V(P_n^d) : wt(x) \leq r\}$. In other words, $B(n, d, r) = \bigcup_{j=0}^r L(n, d, j)$. Note that each edge in P_n^d has one end point in $L(n, d, r)$ and the other end point in $L(n, d, r+1)$ for some r .

In the theorem below, let $N(A) = A \cup \partial(A)$, $N(C) = C \cup \partial(C)$, and $N^{(\leq q)}(A) = A \cup \partial^{(\leq q)}(A)$.

THEOREM 4.2 (see [8] (vertex isoperimetric inequality in the grid)). Let $A \subseteq V(P_n^d)$, and let C be the initial segment of length $|A|$ in the simplicial order on $V(P_n^d)$. Then $|N(A)| \geq |N(C)|$. In particular, if $|A| \geq |B(n, d, r)|$, then $|N(A)| \geq |B(n, d, r+1)|$. Also, for all $q \geq 1$ we have $|N^{(\leq q)}(A)| \geq |B(n, d, r+q)|$.

DEFINITION 4.3. Let G be a graph on n vertices. Let $\sigma : x_1 < x_2 < \dots < x_n$ be a linear order on $V(G)$. The labeling f of $V(G)$ satisfying $f(x_i) = i$ is the vertex labeling of $V(G)$ induced by σ .

LEMMA 4.4. Let $G = P_n^d$. Let f be the labeling of $V(G)$ induced by the simplicial order on G . For every $uv \in E(G)$, we have $|f(u) - f(v)| \leq l^*(n, d) + 2l^*(n, d-1) - 1$.

Proof. Without loss of generality, suppose $wt(u) = r$ and $wt(v) = r+1$. Let $A = \{x \in L(n, d, r) : f(x) > f(u)\}$, $B = \{x \in L(n, d, r+1) : f(x) < f(v)\}$, and $C = \{x \in L(n, d, r+1) : f(x) > f(v)\}$. Let $a = |A|$, $b = |B|$, $c = |C|$. Then

$$|f(u) - f(v)| = a + b + 1 \quad \text{and} \quad b + c + 1 = l(n, d, r+1).$$

Suppose $u = \langle u_1, u_2, \dots, u_d \rangle$ and $v = \langle v_1, v_2, \dots, v_d \rangle$. Since $uv \in E(G)$ and $wt(v) = wt(u) + 1$, there exists $j \in [d]$ such that $v_j = u_j + 1$ and $v_i = u_i$ for all $i \in [d] - j$. By our definition of A , vertices in A all have their first coordinate at most u_1 . Let $A_1 = \{x \in A : x_1 = u_1 \text{ or } u_1 - 1\}$, and let $A_2 = \{x \in A : x_1 \leq u_1 - 2\}$. We have $A = A_1 \cup A_2$. Since there are at most $l^*(n, d-1) - 1$ vertices other than u that have u_1 in the first coordinate and there are at most $l^*(n, d-1)$ vertices that have $u_1 - 1$ in the first coordinate, $|A_1| \leq 2l^*(n, d-1) - 1$.

For each $x \in A_2$, let $g(x) = \langle x_1 + 1, x_2, \dots, x_d \rangle$. It is easy to see that g is an injection of A_2 into $L(n, d, r+1)$. Furthermore, for each $x \in A_2$, since $x_1 + 1 < u_1 \leq v_1$, we have $f(g(x)) > f(v)$ by the definition of the simplicial order. So $g(x) \in C$. So g is an injection of A_2 into C . Thus, $|A_2| \leq c$ and $a = |A| = |A_1| + |A_2| \leq 2l^*(n, d-1) - 1 + c$. Now we have

$$\begin{aligned} |f(u) - f(v)| &= a + b + 1 \leq 2l^*(n, d-1) - 1 + c + b + 1 \\ &= 2l^*(n, d-1) - 1 + l(n, d, r+1) \leq l^*(n, d) + 2l^*(n, d-1) - 1. \quad \square \end{aligned}$$

THEOREM 4.5. Let d be a fixed positive integer. For all positive integers n we have

$$B'(P_n^d) \leq d[l^*(n, d) + 2l^*(n, d-1)] = c(d)dn^{d-1} + O(n^{d-2}),$$

where $c(d)$ is defined as in Theorem 3.7.

Proof. First, we define a digraph H from $G = P_n^d$ by orienting each edge xy from x to y if $x < y$ in the simplicial order. For each vertex x , let $E^+(x)$ denote the set of out edges from x . Note that $|E^+(x)| \leq d$ for each x . We define a labeling g of $E(H)$ (and of $E(G)$) using $1, 2, \dots, |E(H)|$ as follows. Suppose the vertices are $u_1, u_2, \dots$, where $u_1 < u_2 < \dots$ in the simplicial order. Starting with 1 we assign the

first $|E^+(u_1)|$ consecutive labels to $E^+(u_1)$, then the next $|E^+(u_2)|$ consecutive labels to $E^+(u_2)$, and so on. Let $e = u_i u_j$ and $e' = u_j u_k$ be two incident edges in G at the vertex u_j . We consider three cases depending on how e and e' are oriented.

Case 1. $i < j < k$ or $i > j > k$. By symmetry, we may assume $i < j < k$. Then we have $wt(u_i) = r - 1, wt(u_j) = r, wt(u_k) = r + 1$ for some r . By Lemma 4.4, $j - i \leq l^*(n, d) + 2l^*(n, d - 1) - 1$. Note that $e \in E^+(u_i)$ and $e' \in E^+(u_j)$. By our definition of g , we have $|g(e') - g(e)| \leq |\bigcup_{t=i}^j E^+(u_t)| \leq d(j - i + 1) \leq d(l^*(n, d) + 2l^*(n, d - 1))$.

Case 2. $i < j$ and $j > k$. In this case, we have $wt(u_i) = wt(u_k) = r - 1$ and $wt(u_j) = r$ for some r . In particular, $|k - i| \leq l(n, d, r - 1) - 1 \leq l^*(n, d) - 1$. Also, $e \in E^+(u_i)$ and $e' \in E^+(u_k)$. Without loss of generality, suppose $i < k$. By our definition of g , we have $|g(e') - g(e)| \leq |\bigcup_{t=i}^k E^+(u_t)| \leq d|k - i + 1| \leq dl^*(n, d) \leq d(l^*(n, d) + 2l^*(n, d - 1))$.

Case 3. $i > j$ and $k > j$. In this case, $e, e' \in E^+(u_j)$, and $|g(e) - g(e')| \leq d < d(l^*(n, d) + 2l^*(n, d - 1))$.

We have shown that $|g(e) - g(e')| \leq d(l^*(n, d) + 2l^*(n, d - 1))$ for every pair of incident edges e and e' in G . This yields $B'(G) \leq B'(g, G) \leq d(l^*(n, d) + 2l^*(n, d - 1))$. By using $l^*(n, d) = c(d)n^{d-1} + O(n^{d-2})$, we get $B'(G) \leq c(d)dn^{d-1} + O(n^{d-2})$. $\square$

We now derive a lower bound on $B'(P_n^d)$ that matches the upper bound in Theorem 4.5 asymptotically when d is fixed and $n \rightarrow \infty$. Our proof is based on the method used by Calamoneri, Massini, and Vrto [9] and Balogh, Mubayi, and Pluhár [5]. We need an easy lemma.

LEMMA 4.6. *Let n and d be positive integers. Let $A \subseteq V(P_n^d)$. Then there are at least $d|A| - dn^{d-1}$ edges incident to A in P_n^d . Also, $|E(P_n^d)| = dn^d - dn^{d-1}$.*

Proof. The graph P_n^d is a spanning subgraph of C_n^d , the d -fold Cartesian product of C_n . We can think of obtaining C_n^d from P_n^d by adding edges of the form uv , where $u = \langle u_1, u_2, \dots, u_d \rangle$ and $v = \langle v_1, v_2, \dots, v_d \rangle$ satisfy that $u_i = 0, v_i = n - 1$ or $u_i = n - 1, v_i = 0$ for some $i \in [d]$ and $u_j = v_j$ for all $j \in [d] - \{i\}$. It is easy to see that there are dn^{d-1} such edges. Since C_n^d is $2d$ -regular, we have $E(P_n^d) = E(C_n^d) - dn^{d-1} = dn^d - dn^{d-1}$.

In C_n^d , since each vertex has degree $2d$, there are at least $2d|A|/2 = d|A|$ edges incident to A . So in P_n^d , there are at least $d|A| - dn^{d-1}$ edges incident to A . $\square$

THEOREM 4.7. *Let $d \geq 2$ be a fixed positive integer. Then we have as $n \rightarrow \infty$*

$$B'(P_n^d) \geq dl^*(n, d)(1 - o(1)) = c(d)dn^{d-1} + \Omega(n^{d-\frac{3}{2}}).$$

Proof. Throughout the proof, whenever necessary, we assume that n is sufficiently large. Let g be an edge labeling of $G = P_n^d$, with $B'(g, G) = B'(G)$. Let S denote the set of edges receiving labels $1, 2, \dots, |E(G)|/2$. We color the edges in S red and the rest of the edges white.

Let us call a vertex *red* if all of its incident edges are red, a vertex *white* if all of its incident edges are white, and a vertex *mixed* if it is incident to both red edges and white edges. Let R denote the set of red vertices, W the set of white vertices, and M the set of mixed vertices. We consider two cases. For convenience, let $l^* = l^*(n, d)$.

Case 1. $|M| \geq 5l^*$. Call a vertex v *bad* if $d(v) \leq 2d - 3$. Let D denote the set of bad vertices in G . If $v = \langle v_1, v_2, \dots, v_d \rangle$ is bad, then it has a 0 or $n - 1$ in at least three of the d coordinates. So $|D| \leq \binom{d}{3} 2^3 n^{d-3} < 2d^3 n^{d-3}$. By Corollary 3.5, when n is sufficiently large, we have $l^* \geq \frac{1}{2\sqrt{d}} n^{d-1} > 2d^3 n^{d-3} > |D|$. Hence, $|M - D| \geq 5l^* - l^* = 4l^*$. Each vertex in $M - D$ has degree at least $2d - 2$. So the total number of edges incident to M is at least $(2d - 2)|M - D|/2 = (d - 1)|M - D| \geq 4(d - 1)l^* \geq 2dl^*$, since $d \geq 2$.

Let $E(M)$ denote the set of edges in G incident to M , and let $E = E(G)$. We have $|E(M)| \geq 2dl^*$. So either $|E(M) \cap (E - S)| \geq dl^*$ or $|E(M) \cap S| \geq dl^*$. Note

that $E(M) \cap (E - S) \subseteq \partial(S)$ and $E(M) \cap S \subseteq \partial(E - S)$. Hence, we have either $|\partial(S)| \geq dl^*$ or $|\partial(E - S)| \geq dl^*$. By Proposition 2.1, we have $B'(g, G) \geq dl^* = c(d)dn^{d-1} + O(n^{d-2})$.

Case 2. $|M| < 5l^*$. We have $|R| + |M| + |W| = n^d$. Hence, either $|R| < n^d/2$ or $|W| < n^d/2$. Without loss of generality, we may assume that $|R| < n^d/2$; otherwise, we switch S with $E - S$ and hence R with W . Let $H = G[R \cup M]$ denote the subgraph of G induced by R and M . Note that $S \subseteq E(H)$. Hence, $|S| \leq |E(H)| \leq 2d(|R| + |M|)/2 = d(|R| + |M|)$. On the other hand, $|S| = |E(G)|/2 \geq (1/2)(dn^d - dn^{d-1})$. So we have

$$d(|R| + |M|) \geq |S| \geq \frac{1}{2}dn^d - \frac{1}{2}dn^{d-1}.$$

From this we get

$$|R| + |M| \geq \frac{1}{2}n^d - \frac{1}{2}n^{d-1}.$$

So

$$|R| \geq \frac{1}{2}n^d - \frac{1}{2}n^{d-1} - |M| \geq \frac{1}{2}n^d - \frac{1}{2}n^{d-1} - 5l^*.$$

By Corollary 3.5, for sufficiently large n we have $l^* \geq \frac{1}{2\sqrt{d}}n^{d-1}$. Hence, $\frac{1}{2}n^{d-1} \leq \sqrt{dl^*}$ and $|R| \geq \frac{1}{2}n^d - \sqrt{dl^*} - 5l^* \geq \frac{1}{2}n^d - (d+5)l^*$. We have

$$(1) \quad \frac{1}{2}n^{d-1} - (d+5)l^* < |R| < \frac{1}{2}n^{d-1}.$$

Since $l(n, d, j)$ is symmetric on $[0, d(n-1)]$ and $\sum_{j=0}^{d(n-1)} l(n, d, j) = |V(G)| = n^d$, we have

$$(2) \quad \sum_{j=0}^{\lfloor \frac{(n-1)d}{2} \rfloor - 1} l(n, d, j) \leq \frac{1}{2}n^d \leq \sum_{j=0}^{\lfloor \frac{(n-1)d}{2} \rfloor} l(n, d, j).$$

For convenience, let $M = \lfloor \frac{(n-1)d}{2} \rfloor$. By Lemma 3.9, $l(n, d, M-t) \geq (1 - \frac{t}{n-1})l^*$ for $t \in [0, \lfloor \frac{n}{2} \rfloor - 1]$. Let n be sufficiently large so that $n > \binom{d+7}{2}$. We have

$$(3) \quad \sum_{t=1}^{d+6} l(n, d, M-t) \geq \sum_{t=1}^{d+6} \left(1 - \frac{t}{n-1}\right) l^* = (d+6)l^* - \frac{\binom{d+7}{2}}{n-1} l^* \geq (d+5)l^*.$$

By (2) and (3), we have $\sum_{j=0}^{M-d-7} l(n, d, j) \leq |R| \leq \sum_{j=0}^M l(n, d, j)$. In other words,

$$(4) \quad |B(n, d, M-d-7)| < |R| < |B(n, d, M+1)|.$$

Therefore, we have $|B(n, d, r)| \leq |R| < |B(n, d, r+1)|$ for some $r \in [M-d-7, M+1]$. Let $q = \lfloor \sqrt{n} \rfloor - 1$. We may assume that n is sufficiently large so that $q \geq d+7$. Let A be a subset of R , with $|A| = |B(n, d, r)|$. By Theorem 4.2, $|N^{(\leq q)}(A)| \geq |B(n, d, r+q)|$. Since $N^{(\leq q)}(A) \subseteq N^{(\leq q)}(R)$, we have $|N^{(\leq q)}(R)| \geq |B(n, d, r+q)|$. Hence,

$$(5) \quad \begin{aligned} |\partial^{(\leq q)}(R)| &= |N^{(\leq q)}(R)| - |R| \geq |B(n, d, r+q)| \\ &\quad - |B(n, d, r+1)| = \sum_{j=2}^q l(n, d, r+j). \end{aligned}$$

For each $j \in [r+2, r+q]$, we have $j \in [M-d-7, M+q+1]$. Since $|j-M| \leq q+1$, by Lemma 3.9,

$$l(n, d, j) \geq \left(1 - \frac{q+1}{n-1}\right) l^* \geq \left(1 - \frac{\lfloor \sqrt{n} \rfloor}{n-1}\right) l^* \geq \left(1 - \frac{2}{\sqrt{n}}\right) l^*.$$

Hence,

$$|\partial^{(\leq q)}(R)| \geq \sum_{j=2}^q l(n, d, r+j) \geq (q-1) \left(1 - \frac{2}{\sqrt{n}}\right) l^*.$$

Let $E(\partial^{(\leq q)}(R))$ denote the set of edges incident to $\partial^{(\leq q)}(R)$. By Lemma 4.6,

$$|E(\partial^{(\leq q)}(R))| \geq d|\partial^{(\leq q)}(R)| - dn^{d-1}.$$

Finally, note that $E(\partial^{(\leq q)}(R)) \subseteq \partial^{(\leq q+1)}(S)$. By applying Proposition 2.2 to the line graph $L(G)$, we have

$$\begin{aligned} B'(g, G) &\geq |\partial^{(\leq q+1)}(S)|/(q+1) \\ &\geq |E(\partial^{(\leq q)}(R))|/(q+1) \\ &\geq (d|\partial^{(\leq q)}(R)| - dn^{d-1})/(q+1) \\ &\geq d \left[(q-1) \left(1 - \frac{2}{\sqrt{n}}\right) \right] l^*/(q+1) - dn^{d-1}/(q+1) \\ &= dl^* \left(\frac{\sqrt{n}-2}{\sqrt{n}} \right) \left(1 - \frac{2}{\sqrt{n}}\right) + O(n^{d-\frac{3}{2}}) \quad (\text{using } q = \sqrt{n} - 1) \\ &= c(d)dn^{d-1} + O(n^{d-\frac{3}{2}}) \quad (\text{by Theorem 3.7}). \quad \square \end{aligned}$$

5. Edge-bandwidth of the Hamming graph I: Upper bound. In this section we derive an upper bound on $B'(K_n^d)$. We view any vertex x in K_n^d as an n -ary d -string $x = \langle x_1, x_2, \dots, x_d \rangle$, where, for all i , $x_i \in \{0, 1, \dots, n-1\}$. Two vertices $x = \langle x_1, x_2, \dots, x_d \rangle$ and $y = \langle y_1, y_2, \dots, y_d \rangle$ are adjacent in K_n^d if and only if the two strings differ in precisely one coordinate. As in section 4, for each $x = \langle x_1, x_2, \dots, x_d \rangle$, we define the *weight* of x as $wt(x) = x_1 + x_2 + \dots + x_d$. Note that, when $n = 2$, K_n^d is just the d -dimensional hypercube Q_d , and, for each vertex x , $wt(x)$ is the number of 1's in the binary string that represents x . The edge-bandwidth $B'(Q_d)$ was asymptotically determined by Balogh, Mubayi, and Pluhár [5], while the vertex bandwidth $B(Q_d)$ was completely determined by Harper in his paper [13]. We will combine the labelings used in these results to design a labeling that yields an upper bound on $B'(K_n^d)$. Let us recall the labelings used in [5, 13].

DEFINITION 5.1. *The vertex-Hales numbering of $V(Q_d)$ is a bijection $h : V(Q_d) \rightarrow \{1, 2, \dots, 2^d\}$ such that $h(x) < h(y)$ if either $wt(x) < wt(y)$ or $wt(x) = wt(y)$ and $x_s > y_s$, where $s = \min\{t : x_t \neq y_t\}$.*

Note that, in section 4, we called a similar ordering on $V(P_n^d)$ the simplicial order. Harper showed that the vertex-Hales numbering achieves the vertex bandwidth for Q_d .

THEOREM 5.2 (see [13]). *Let h be the vertex-Hales numbering of Q_d . Then*

$$B(Q_d) = B(h) = \sum_{k=0}^{d-1} \binom{k}{\lfloor k/2 \rfloor} = (1 + o(1)) \binom{d}{\lfloor d/2 \rfloor}.$$

The last equality used standard estimates of binomial coefficients. Next we describe the labeling used by Balogh, Mubayi, and Pluhár in establishing an asymptotically tight upper bound on $B'(Q_d)$. For convenience, we will call this numbering the *edge-Hales numbering*. Our description of the numbering is equivalent to the one used in [5].

DEFINITION 5.3 (see [5]). *The edge-Hales numbering of Q_d is a bijection $f : E(Q_d) \rightarrow \{1, 2, \dots, d2^{d-1}\}$ such that for any two edges vw and xy , where $wt(w) = wt(v) + 1$ and $wt(y) = wt(x) + 1$, we have $f(vw) < f(xy)$ if either (i) $h(v) < h(x)$ or (ii) $v = x$ and $h(w) < h(y)$.*

THEOREM 5.4 (see [5]). *Let f denote the edge-Hales labeling of Q_d . Then*

$$\left(\frac{d}{2} + o(d)\right) \binom{d}{\lfloor d/2 \rfloor} \leq B'(Q_d) \leq B'(f) \leq \left(\frac{d}{2} + o(d)\right) \binom{d}{\lfloor d/2 \rfloor}.$$

Now we combine the two numberings mentioned above to obtain a total numbering on $V(Q_d) \cup E(Q_d)$, which we will call the *mixed-Hales numbering* of Q_d . This numbering is produced by the following algorithm.

ALGORITHM 5.5 (the mixed-Hales numbering m of Q_d).

Input: The d -dimensional hypercube Q_d .

Output: A bijection $m : V(Q_d) \cup E(Q_d) \rightarrow \{1, 2, \dots, 2^d + d2^{d-1}\}$.

Initialization:

(1) Denote the edges of Q_d by e_i , $1 \leq i \leq d2^{d-1}$, according to the edge-Hales numbering f of Q_d . That is, e_i is the edge e with $f(e) = i$.

(2) Let $m(0^d) = 1$, where 0^d denotes the all-0 string of length d .

(3) For all i , $1 \leq i \leq d2^{d-1}$, let $m(e_i) = i + 1$.

(4) Set $i = 1$.

Iteration:

(5) Suppose $e_i = xy$, where $wt(y) = wt(x) + 1$. If $m(y)$ is not yet defined, then

(5a) let $m(y) = m(e_i) + 1$;

(5b) for all $j > i$, let $m(e_j) = m(e_i) + 1$.

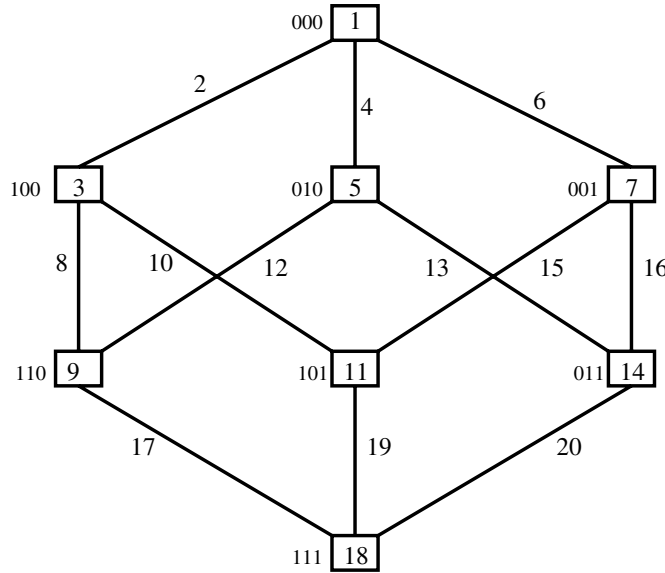
(6) Let $i = i + 1$.

(7) If $i = d2^{d-1} + 1$, terminate; otherwise, go to step (5).

Intuitively speaking, to obtain the mixed-Hales numbering we process the edges one by one in increasing order of edge-Hales label. The algorithm gives a vertex y an m -label at the earliest opportunity, as soon as we process the first edge in the edge-Hales numbering incident to y . See Figure 1 for the mixed-Hales labeling of Q_3 .

Next, we summarize some useful facts about mixed-Hales numbering in the following proposition. In particular, we see that the ordering on $V(Q_d)$ and the ordering on $E(Q_d)$ inherited from the mixed-Hales numbering m of Q_d are precisely the vertex-Hales numbering h and the edge-Hales numbering f of Q_d , respectively.

PROPOSITION 5.6. *Let h, f, m denote the vertex-Hales, edge-Hales, and mixed-Hales numberings of Q_d , respectively. Let 0^d denote the all-0 string of length d and 1^d the all-1 string of length d . For each vertex $x \in V(Q^d) - \{0^d, 1^d\}$, let x^- denote the neighbor of x with the smallest h label and x^+ the neighbor of x with the largest h label. Then the following apply:*

FIG. 1. The mixed-Hales numbering of Q_3 .

1. For each vertex $x \in V(Q_d) - \{0^d, 1^d\}$, the string representing x^- is obtained from the string for x by flipping the rightmost 1 to a 0, and the string representing x^+ is obtained from the string for x by flipping the rightmost 0 to a 1.
2. For each vertex $x \in V(Q_d) - \{0^d, 1^d\}$, among the edges incident to x , xx^- has the smallest m label and xx^+ has the largest m label. Hence, in particular, $m(x) = m(xx^-) + 1$.
3. For any two edges e, e' in Q_d , if $f(e) < f(e')$, then $m(e) < m(e')$.
4. For any two edges vw, xy in Q_d , where $wt(w) = wt(v) + 1$ and $wt(y) = wt(x) + 1$, $m(vw) < m(xy)$ if and only if either $h(v) < h(x)$ or $v = x$ and $h(w) < h(y)$.
5. For any two vertices $x, y \in V(Q_d) - \{0^d, 1^d\}$, if $h(x) < h(y)$, then $h(x^-) \leq h(y^-)$ and $h(x^+) \leq h(y^+)$.
6. For any two vertices, x, y in Q_d , if $h(x) < h(y)$, then $m(x) < m(y)$.

Proof. Part 1 follows immediately from the definition of the vertex-Hales numbering (Definition 5.1). Parts 2 and 3 follow immediately from Algorithm 5.5. Part 4 follows from the definition of the edge-Hales numbering f (Definition 5.3) and part 3.

To prove part 5, suppose $h(x) < h(y)$. Since $h(x) < h(y)$, we have $wt(x) \leq wt(y)$. If $wt(x) < wt(y)$, then $wt(x^-) < wt(y^-)$ and $h(x^-) < h(y^-)$ hold trivially. So we may assume that $wt(x) = wt(y)$. By part 1, the string representing x^- is obtained from the string representing x by flipping the rightmost 1 in x to 0, and the string for y^- is obtained from the string for y by flipping the rightmost 1 in y to 0. Let j denote the smallest coordinate in which x and y differ. Since $h(x) < h(y)$, we have $x_j = 1$ and $y_j = 0$. Since $wt(x) = wt(y)$, if x has k many 1's in coordinates $j+1, j+2, \dots, d$, then y should have exactly $k+1$ many 1's in coordinates $j+1, j+2, \dots, d$. If $k \geq 1$, then clearly $h(x^-) < h(y^-)$. If $k = 0$, then $x^- = y^-$, and hence $h(x^-) = h(y^-)$. By a very similar argument, we have $h(x^+) \leq h(y^+)$.

To prove part 6, we may assume without loss of generality that $x, y \notin \{0^1, 1^d\}$. Suppose $h(x) < h(y)$. Since $m(x) = m(xx^-) + 1$ and $m(y) = m(yy^-) + 1$, to prove $m(x) < m(y)$ it suffices to prove that $m(xx^-) < m(yy^-)$. By part 5, $h(x^-) \leq h(y^-)$. Thus, we have either $h(x^-) < h(y^-)$ or $x^- = y^-$ and $h(x) < h(y)$. By part 4, we have $m(xx^-) < m(yy^-)$. This completes the proof. $\square$

In the next proposition, we bound the number of vertices and the number of edges whose m -labels lie between the m -labels of two incident edges.

PROPOSITION 5.7. *Let h, f, m denote the vertex-Hales, edge-Hales, and mixed-Hales numberings of Q_d , respectively. Let e, e' be two incident edges in Q_d , where $m(e) < m(e')$. Then there are at most $B(h) + 1$ vertices z with $m(e) \leq m(z) \leq m(e')$, and there are at most $B'(f) + 1$ edges e^* with $m(e) \leq m(e^*) \leq m(e')$.*

Proof. By Proposition 5.6 part 2, an edge e^* satisfies $m(e) \leq m(e^*) \leq m(e')$ if and only if $f(e) \leq f(e^*) \leq f(e')$. Hence, since e and e' are incident, there are at most $|f(e') - f(e)| + 1 \leq B'(f) + 1$ such edges e^* . Next, suppose e and e' are both incident to x . Let z be a vertex with $m(e) \leq m(z) \leq m(e')$. If $x = 0^d$ or 1^d , then it is easy to see that there are at most $d < B(h)$ such vertices z . Hence, we may assume that $x \notin \{0^d, 1^d\}$. By Proposition 5.6 part 2, $m(xx^-) \leq m(z) \leq m(xx^+)$. Since $m(x) = m(xx^-) + 1$ and $m(z) > m(xx^-)$, we have $m(z) \geq m(x)$. We show that also $m(z) \leq m(x^+)$. Suppose first that $h(z^-) > h(x)$. Then by Proposition 5.6 part 4, $m(zz^-) > m(xx^+)$, and hence $m(z) = m(zz^-) + 1 > m(xx^+)$, a contradiction. So we must have $h(z^-) \leq h(x)$. By Proposition 5.6 part 5, $h((z^-)^+) \leq h(x^+)$. Since $h(z) \leq h((z^-)^+)$, we have $h(z) \leq h(x^+)$, and thus $m(z) \leq m(x^+)$. So any vertex z satisfying $m(e) \leq m(z) \leq m(e')$ must satisfy $h(x) \leq h(z) \leq h(x^+)$. There are at most $|h(x^+) - h(x)| + 1 \leq B(h) + 1$ such vertices z . $\square$

LEMMA 5.8. *Let p, q, t be positive integers. Let G be a graph obtained from Q_d by replacing each vertex x of Q_d by a t -vertex graph G_x having p edges and each edge xy of Q_d by a set of q cross edges between $V(G_x)$ and $V(G_y)$. Then $B'(G) \leq p(B(h) + 1) + q(B'(f) + 1)$, where h, f denote the vertex-Hales and edge-Hales numberings of Q_d , respectively.*

Proof. Apply Algorithm 5.5 to obtain the mixed-Hales numbering m of Q_d . List elements of $V(Q_d) \cup E(Q_d)$ in the order determined by m , and call this list L . We produce a labeling g of $E(G)$ as follows. Start with label 1. As we scan L , each time we encounter a vertex x , we allocate the next p consecutive labels to the edges of G_x , and each time we encounter an edge $e = xy$, we allocate the next q consecutive labels to the set of q cross edges between $V(G_x)$ and $V(G_y)$.

Consider any pair of incident edges e and e' in G with $g(e) < g(e')$. Suppose both are incident to vertex w , and w lies in G_x . If $x \in \{0^d, 1^d\}$, then it is easy to see that $|g(e) - g(e')| \leq p + dq < p(B(h) + 1) + q(B'(f) + 1)$. Hence we may assume that $x \notin \{0^d, 1^d\}$. By our labeling scheme $|g(e) - g(e')|$ is maximum when e is among the set of q edges of G associated with edge xx^- in Q_d and e' is among the set of q edges of G associated with edge xx^+ in Q_d . By Proposition 5.7 and the definition of g , $|g(e) - g(e')| \leq p(B(h) + 1) + q(B'(f) + 1)$. $\square$

Now we apply Lemma 5.8 to get an upper bound on $B'(K_n^d)$. For convenience, we consider only even n . For odd n , we can upper bound $B'(K_n^d)$ by $B'(K_{n+1}^d)$. We can view K_n^d as being obtained from Q_d by replacing each vertex of Q_d with a copy of $K_{n/2}^d$ and replacing each edge of Q_d by the set of edges between two neighboring copies of $K_{n/2}^d$ in K_n^d . More specifically, for each $x = \langle x_1, \dots, x_d \rangle \in V(Q_d)$, let $O(x)$ denote the subgraph of K_n^d induced by the set of vertices $\{w = \langle w_1, \dots, w_d \rangle : 0 \leq w_i \leq n/2 - 1 \text{ if } x_i = 0, n/2 \leq w_i \leq n - 1 \text{ if } x_i = 1\}$. Then each $O(x)$ is a copy of

$K_{n/2}^d$. For each edge $xy \in E(Q_d)$, let $E(O(x), O(y))$ denote the set of edges in K_n^d having one end point in $O(x)$ and the other end point in $O(y)$. It is easy to see that $|E(O(x), O(y))| = (\frac{n}{2})^2 (\frac{n}{2})^{d-1} = (\frac{n}{2})^{d+1}$ for all $xy \in E(Q_d)$. We denote this quantity by $q(n, d)$. By applying Lemma 5.8 with $p = e(K_{n/2}^d)$ and $q = q(n, d)$, we have the following.

LEMMA 5.9. *Let d be a positive integer and n a positive even integer. We have*

$$B'(K_n^d) \leq e(K_n^d)(B(h) + 1) + q(n, d)(B'(f) + 1).$$

THEOREM 5.10. *Let n be a fixed positive even integer. Let d be a positive integer. We have*

$$B'(K_n^d) \leq (d + o(d)) \binom{d}{\lfloor d/2 \rfloor} \left(\frac{n}{2}\right)^d (n-1) = (1 + o(1)) \frac{\sqrt{d}}{\sqrt{2\pi}} n^d (n-1),$$

as $d \rightarrow \infty$.

Proof. Using Lemma 5.9, Theorems 5.2 and 5.4, $e(K_{n/2}^d) = d(\frac{n}{2})^{d-1} \binom{n/2}{2}$, $q(n, d) = (\frac{n}{2})^{d+1}$, and $\sum_{k=0}^{d-1} \binom{k}{\lfloor k/2 \rfloor} = (1 + o(1)) \binom{d}{\lfloor d/2 \rfloor} = (1 + o(1)) \frac{\sqrt{2}}{\sqrt{\pi d}} \cdot 2^d$, we have

$$\begin{aligned} B'(K_n^d) &\leq \left[d \left(\frac{n}{2}\right)^{d-1} \binom{n/2}{2} \right] \cdot \left[1 + \sum_{k=0}^{d-1} \binom{k}{\lfloor k/2 \rfloor} \right] \\ &\quad + \left(\frac{n}{2}\right)^{d+1} \cdot \left[\left(\frac{d}{2} + o(d)\right) \binom{d}{\lfloor d/2 \rfloor} \right] \\ &= (1 + o(1)) \binom{d}{\lfloor d/2 \rfloor} \left[d \left(\frac{n}{2}\right)^{d-1} \cdot \binom{n/2}{2} \right] \\ &\quad + (1 + o(1)) \cdot \frac{d}{2} \cdot \binom{d}{\lfloor d/2 \rfloor} \cdot \left(\frac{n}{2}\right)^{d+1} \\ &= (1 + o(1)) \binom{d}{\lfloor d/2 \rfloor} \cdot d \cdot \left(\frac{n}{2}\right)^{d-1} \cdot \left(\binom{n/2}{2} + \frac{1}{2} \left(\frac{n}{2}\right)^2 \right) \\ &= (1 + o(1)) \binom{d}{\lfloor d/2 \rfloor} \cdot \left(\frac{1}{2}\right)^{d+1} d n^d (n-1) \\ &= (1 + o(1)) \frac{\sqrt{d}}{\sqrt{2\pi}} n^d (n-1). \quad \square \end{aligned}$$

On a historic note, the idea of blowing up each vertex of Q_d to an “orthant” subgraph $O(x)$ of K_n^d was used by Harper in [15] in his constructive upper bound for $B(K_n^d)$. It was independently discovered by one of us while supervising an M.A. thesis [4]. To develop a constructive upper bound on $B'(K_n^d)$, we used this idea by labeling edges internal to orthants in blocks in the order of the vertex-Hales numbering of Q_d which achieved $B(Q_d)$ [13]. Our next step was to label cross edges between neighboring orthants in blocks in the order of the edge-Hales labeling of Q_d achieving $B'(Q_d)$ asymptotically, described earlier and originating in [5]. In this way both classes of edges, either internal to orthants or crossing between orthants, are labeled. The final step was to merge these two labelings efficiently. The key idea here is to notice that the edge-Hales labeling of Q_d is in some natural sense induced by the vertex-Hales numbering.

6. Edge-bandwidth of the Hamming graph II: Lower bound. In this section we establish a lower bound for $B'(K_n^d)$ which matches the upper bound of the previous section asymptotically when n is even. Our technique employs a theorem of Harper giving a solution to the isoperimetric problem in $[0, 1]^d$.

The approach is to look at K_n^d geometrically as a d -dimensional box having side length 1 and so containing n^d many d -dimensional cells of length $1/n^d$ in each dimension. More specifically, we consider the following mapping from $V(K_n^d)$ to $[0, 1]^d$. For each $i \in \{0, 1, \dots, n-1\}$, let I_i denote the interval $[\frac{i}{n}, \frac{i+1}{n}]$ of real numbers. For each vertex $x = \langle x_1, x_2, \dots, x_d \rangle$ in K_n^d , where each $x_i \in \{0, 1, \dots, n-1\}$, let $g(x) = I_{x_1} \times I_{x_2} \times \dots \times I_{x_d}$. Thus $g(x)$ is a d -dimensional cell of length $1/n^d$ in each dimension. Under this mapping, a subset W of k vertices in $V(K_n^d)$ then corresponds to a collection S of k of these cells. S is a set of measure k/n^d in $[0, 1]^d$, the d -fold product of the unit interval, equipped with the Lebesgue measure.

Given two points $x = \langle x_1, x_2, \dots, x_d \rangle$ and $y = \langle y_1, y_2, \dots, y_d \rangle$ in $[0, 1]^d$, we say that x and y are *neighbors* if there exists $k \in \{1, 2, \dots, d\}$ such that $x_k \neq y_k$ but $x_i = y_i$ for all $i \neq k$. We write $x \leftrightarrow y$ if x and y are neighbors. Given a measurable subset S of $[0, 1]^d$, we define the shadow $\Phi(S)$ to be

$$\Phi(S) = \{y \in [0, 1]^d - S : \exists x \in S \ x \leftrightarrow y\}.$$

We define iterated shadows of S recursively as follows. Let $\Phi^{(1)}(S) = \Phi(S)$, and inductively define $\Phi^{(l+1)}(S) = \Phi(\Phi^{(l)}(S))$.

For a measurable subset S of $[0, 1]^d$, let $|S|$ denote its measure. The following proposition is clear from our definitions and discussions above.

PROPOSITION 6.1. *Let g be the mapping from $V(K_n^d)$ to $[0, 1]^d$ defined above. Let W be a subset of $V(K_n^d)$ and l a positive integer with $1 \leq l \leq d$. We have $g(\partial^{(\leq l)}(W)) = \Phi^{(\leq l)}(g(W))$. Hence, $|\partial^{(\leq l)}(W)| = |\Phi^{(\leq l)}(g(W))| \cdot n^d$.*

To establish a lower bound on $B'(K_n^d)$ we will need to establish a lower bound on $|\partial(W)|$ for any subset W of $V(K_n^d)$ of a given size k . By Proposition 6.1, it suffices to consider the problem of minimizing $|\Phi(S)|$ over all measurable subsets S of $[0, 1]^d$ of a given measure. In [14], Harper showed that the problem of minimizing $|\Phi(S)|$ over all measurable subsets S of $[0, 1]^d$ of a given measure reduces to that of minimizing $|\Phi(S)|$ over all suitably “compressed” such subsets. He then showed by using variational methods that for any given v , $0 \leq v \leq 1$, the smallest value of $|\Phi(S)|$ over all such subsets of measure v is achieved by a “Hamming ball,” which we define as follows.

DEFINITION 6.2. *Let t be real number with $0 \leq t \leq 1$. For any binary d -tuple x of Q_d , let $K(x, t)$ be the subset of $[0, 1]^d$ defined by*

$$K(x, t) = \{y \in [0, 1]^d : 0 \leq y_i \leq t \text{ if } x_i = 1, \text{ and } t < y_i \leq 1 \text{ if } x_i = 0, 1 \leq i \leq d\}.$$

Define the subset $HB(d, k, t)$ of $[0, 1]^d$ called a Hamming ball by

$$HB(d, k, t) = \bigcup_{wt(x) \leq k} K(x, t).$$

Note that if x has weight i , then the measure of $K(x, t)$ is $t^i(1-t)^{d-i}$. So the measure of $HB(d, k, t)$ is $v = v(d, k, t) = \sum_{i=0}^k \binom{d}{i} t^i(1-t)^{d-i}$, and the size of its shadow is $|\Phi(HB(d, k, t))| = \binom{d}{k+1} t^{k+1}(1-t)^{d-k-1}$. For convenience, we use $\alpha(d, m, t)$ to denote $\binom{d}{m} t^m(1-t)^{d-m-1}$. Then $|\Phi(HB(d, k, t))| = \alpha(d, k+1, t)$. We can now state Harper’s theorem.

THEOREM 6.3 (see [14, Theorem 1]). *For any given v with $0 \leq v \leq 1$, the minimum of $|\Phi(S)|$ over all subsets S of $[0, 1]^d$ of measure v is achieved by some $HB(d, k, t)$ for suitable k and t .*

To make effective use of Theorem 6.3, we need to analyze $|\Phi(HB(d, k, t))| = \alpha(d, k + 1, t)$. Observe that if X is a random variable drawn from the binomial distribution $BIN(d, t)$, then $v(d, k, t) = \sum_{i=0}^k \binom{d}{i} t^i (1-t)^{d-i} = \Pr(X \leq k)$ and $\alpha(d, k + 1, t) = \binom{d}{k+1} t^{k+1} (1-t)^{d-k-1} = \Pr(X = k + 1)$.

Our general approach is again the one used in [9, 5], and the proof of Theorem 4.7. Given an optimal edge labeling of $E(K_n^d)$, we consider the set $S' \subseteq E(K_n^d)$ of size about $|E(K_n^d)|/2$ receiving the smallest labels. We then lower bound $|\partial(S')|$ or $|\partial^{(\leq q)}(S')|$ for an adequate q . By an argument similar to the one used in the proof of Theorem 4.7, lower bounding $|\partial(S')|$ is reduced to lower bounding $|\partial(S)|$ for a corresponding set $S \subseteq V(K_n^d)$, with $|S|$ near $\frac{1}{2}n^d$. This then reduces to lower bounding $|\Phi(S)|$, where $S \subseteq [0, 1]^d$, with S having measure near $\frac{1}{2}$. By Theorem 6.3, we need to estimate $|\Phi(HB(d, k, t))| = \alpha(d, k + 1, t)$ when $v(d, k, t)$ is near $\frac{1}{2}$. In light of our earlier observation, this means lower bounding $\Pr(X = k + 1)$ when $\Pr(X \leq k)$ is close to $\frac{1}{2}$, where X is a random variable drawn from $BIN(d, t)$. One could in principle obtain such a lower bound by approximating $BIN(d, t)$ using a normal distribution. But the error analysis in such an approximation is quite involved. Further, when the expected value dt is either too close to 0 or too close to d , a normal distribution approximation becomes infeasible.

Here we use a self-contained and completely combinatorial approach to obtain our estimates. We think our approach is of independent interest. We need a lemma from [21].

LEMMA 6.4 (see [21, Lemma B.7]). *Let n be a positive integer and p a real number such that $0 \leq p \leq 1$. Let X be a random variable drawn from the binomial distribution $B(n, p)$ (where n is the number of independent trials and p is the probability of success of each trial). Then*

$$\Pr(X \leq \lfloor np \rfloor - 1) \leq 1/2 \leq \Pr(X \leq \lceil np \rceil).$$

Lemma 6.4 suggests that if $\Pr(X \leq k)$ is close to $\frac{1}{2}$, then k is close to the expected value np . Thus, we need to lower bound $\Pr(X = k + 1)$ for those k close to np .

LEMMA 6.5. *If z is a real number with $0 < z < \frac{1}{2}$, then $1 - z \geq e^{-2z}$. Let x, y, a be positive real numbers such that $x, y \geq 2a$. Then*

$$\left(1 + \frac{a}{x}\right)^x \left(1 - \frac{a}{y}\right)^y \geq e^{-\frac{a^2}{x} - \frac{a^2}{y}}.$$

Proof. For any real number w , with $0 < w \leq \frac{1}{2}$, we have

$$\begin{aligned} \ln(1 + w) &= w - \frac{w^2}{2} + \frac{w^3}{3} - \frac{w^4}{4} + \cdots \geq w - \frac{w^2}{2} > w - w^2, \\ \ln(1 - w) &= -w - \frac{w^2}{2} - \frac{w^3}{3} - \frac{w^4}{4} - \cdots \geq -w - w^2. \end{aligned}$$

If $0 < z < \frac{1}{2}$, then $\ln(1 - z) \geq -z - z^2 \geq -2z$. So $1 - z \geq e^{-2z}$. Since $x, y \geq 2a$, $0 < \frac{a}{x}, \frac{a}{y} \leq \frac{1}{2}$. Let $K = \left(1 + \frac{a}{x}\right)^x \left(1 - \frac{a}{y}\right)^y$. It suffices to show that $\ln K \geq -\frac{a^2}{x} - \frac{a^2}{y}$. Indeed, we have

$$\ln K = x \ln \left(1 + \frac{a}{x}\right) + y \ln \left(1 - \frac{a}{y}\right) \geq x \left(\frac{a}{x} - \frac{a^2}{x^2}\right) + y \left(-\frac{a}{y} - \frac{a^2}{y^2}\right) = -\frac{a^2}{x} - \frac{a^2}{y}. \quad \square$$

Our next lemma says that if X is a random variable drawn from $BIN(n, p)$, where np is not too small or too large and k is near np , then $Pr(X = k)$ is lower bounded by $\sqrt{\frac{2}{\pi n}}(1 - o(1))$ as $n \rightarrow \infty$. We postpone its somewhat technical proof to the appendix.

LEMMA 6.6. *Let n be a positive integer. Let p be a real number such that $0 < p < 1$. Suppose $5\sqrt{\ln n} \leq np \leq n - 5\sqrt{\ln n}$. Let k be a nonnegative integer. Let X be a random variable drawn from $BIN(n, p)$.*

If $|k - np| \leq \sqrt{\ln n}$, then $Pr(X = k) \geq \sqrt{\frac{2}{\pi n}} \cdot e^{-\frac{1}{\sqrt{n}}}$ when n is sufficiently large.

Now we use Lemma 6.6 to give lower bounds on $|\Phi(S)|$, for subsets S of $[0, 1]^d$ having measure near $\frac{1}{2}$.

THEOREM 6.7. *Let d be a sufficiently large positive integer. Let S be a subset of $[0, 1]^d$ with measure $v = |S|$.*

If $\left| |S| - \frac{1}{2} \right| \leq \sqrt{\frac{2}{\pi d}} \cdot \sqrt{\ln d} \cdot e^{-\frac{10}{\sqrt{\ln d}}}$, then $|\Phi(S)| \geq \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{1}{\sqrt{d}}}$.

Proof. By Theorem 6.3, $|\Phi(S)|$ is minimum when S is a Hamming ball. Hence, we may assume $S = HB(d, k, t)$ for some k and t , where k is an integer with $1 \leq k \leq d - 1$ and t is a real number with $0 < t < 1$. We have $|S| = v(d, k, t) = \sum_{i=0}^k \alpha(d, i, t)$ and $|\Phi(S)| = \alpha(d, k + 1, t)$. As observed earlier, if X is a random variable drawn from $BIN(d, t)$, then $|S| = Pr(X \leq k)$ and $|\Phi(S)| = Pr(X = k + 1)$. By our assumption,

$$(6) \quad \left| Pr(X \leq k) - \frac{1}{2} \right| \leq \sqrt{\frac{2}{\pi d}} \cdot \sqrt{\ln d} \cdot e^{-\frac{10}{\sqrt{\ln d}}}.$$

By Lemma 6.4,

$$(7) \quad Pr(X \leq \lfloor dt \rfloor - 1) \leq \frac{1}{2} \leq Pr(X \leq \lceil dt \rceil).$$

We consider several cases.

Case 1. $5\sqrt{\ln d} \leq dt \leq d - 5\sqrt{\ln d}$. By Lemma 6.6, for each integer m with $|m - dt| \leq \sqrt{\ln d}$, we have $Pr(X = m) \geq \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{1}{\sqrt{d}}}$, when d is large. Hence, for sufficiently large d , we have

$$\begin{aligned} Pr(X \leq \lceil dt \rceil + \lceil \sqrt{\ln d} - 3 \rceil) &\geq \frac{1}{2} + (\sqrt{\ln d} - 3) \cdot \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{1}{\sqrt{d}}} \\ &\geq \frac{1}{2} + \sqrt{\frac{2}{\pi d}} \cdot \sqrt{\ln d} \cdot e^{-\frac{10}{\sqrt{\ln d}}} \geq Pr(X \leq k), \end{aligned}$$

where the last inequality follows from (6). So $k \leq \lceil dt \rceil + \lceil \sqrt{\ln d} \rceil - 3$. By a similar argument, we have $k \geq \lfloor dt \rfloor - 1 - (\lceil \sqrt{\ln d} \rceil - 3)$. It follows that $|(k + 1) - dt| \leq \sqrt{\ln d}$. Since $5\sqrt{\ln d} \leq dt \leq d - 5\sqrt{\ln d}$, by Lemma 6.6, we have

$$|\Phi(S)| = Pr(X = k + 1) \geq \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{1}{\sqrt{d}}}.$$

Case 2. $dt < 5\sqrt{\ln d}$. In this case, we have $t < \frac{5\sqrt{\ln d}}{d}$. Suppose $t \leq \frac{1}{4d}$. Then

$$Pr(X \leq k) \geq Pr(X = 0) = (1 - t)^d \geq \left(1 - \frac{1}{4d}\right)^d \geq e^{-2 \cdot \frac{1}{4d} \cdot d} \geq e^{-\frac{1}{2}} > 0.6,$$

contradicting (6) for large d . So $t > \frac{1}{4d}$.

If $k > 3dt$, then by Markov's inequality we have $Pr(X > k) \leq \frac{1}{3}$ and hence $Pr(X \leq k) > \frac{2}{3}$, contradicting (6) for large d . So we have $k \leq 3dt \leq 15\sqrt{\ln d}$. Let $m = k + 1$. Then $m \leq 16\sqrt{\ln d}$. We have (recalling that $\frac{1}{4} \leq dt \leq 5\sqrt{\ln d}$)

$$\begin{aligned} |\Phi(S)| &= Pr(X = k + 1) = Pr(X = m) = \binom{d}{m} t^m (1 - t)^{d-m} \\ &\geq \left(\frac{d}{m}\right)^m \cdot t^m \cdot (1 - t)^d \geq \left(\frac{dt}{m}\right)^m \cdot e^{-2td} \geq \left(\frac{1}{4m}\right)^m \cdot e^{-10\sqrt{\ln d}} \\ &= e^{-m \ln 4m - 10\sqrt{\ln d}} \geq e^{-\frac{1}{4} \ln d} \geq \frac{1}{d^{\frac{1}{4}}} \geq \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{1}{\sqrt{d}}} \quad \text{for large } d. \end{aligned}$$

Case 3. $dt > d - 5\sqrt{\ln d}$. In this case, we apply almost identical reasoning as in Case 2 but with $1 - t$ playing the role of t . First we show $1 - t \geq \frac{1}{4d}$. Then we apply Markov's inequality to $d - X$ to show that $d - k \leq 15\sqrt{\ln d}$ and therefore $d - m \leq 16\sqrt{\ln d}$. Then we establish the lower bound on $|\Phi(S)|$ as in Case 2, with $1 - t$ playing the role of t and $d - m$ playing the role of m . We omit the details. $\square$

COROLLARY 6.8. *Let d be a sufficiently large positive integer. Let S be a subset of $[0, 1]^d$ with measure $v = |S|$. Suppose*

$$\frac{1}{2} - \sqrt{\frac{2}{\pi d}} \cdot \sqrt{\ln d} \cdot e^{-\frac{10}{\sqrt{\ln d}}} \leq |S| \leq \frac{1}{2}.$$

Then either

$$|\Phi(S)| \geq 2\sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{10}{\sqrt{\ln d}}}$$

or there exists an integer $l \geq 2$ such that

$$\frac{|\Phi^{(\leq l)}(S) - \Phi(S)|}{l + 1} \geq \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{10}{\sqrt{\ln d}}}.$$

Proof. For each i , let $S_i = \Phi^{(i)}(S)$. Let l denote the smallest m such that $|S_1 \cup S_2 \cup \dots \cup S_m| \geq \sqrt{\frac{2}{\pi d}} \cdot \sqrt{\ln d} \cdot e^{-\frac{10}{\sqrt{\ln d}}}$. Let $i \in \{1, \dots, l\}$. Let $S' = S \cup S_1 \cup \dots \cup S_{i-1}$. Then

$$|S'| = |S| + |S_1 \cup \dots \cup S_{i-1}| \leq \frac{1}{2} + \sqrt{\frac{2}{\pi d}} \cdot \sqrt{\ln d} \cdot e^{-\frac{10}{\sqrt{\ln d}}}.$$

By Theorem 6.7, we have

$$|S_i| = |\Phi(S')| \geq \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{1}{\sqrt{d}}}.$$

Since this holds for all $i \leq l$, we have

$$|S_1 \cup \dots \cup S_{l-1}| \geq (l - 1) \cdot \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{1}{\sqrt{d}}}.$$

If $l > \sqrt{\ln d} - 3$, then

$$|S_1 \cup \dots \cup S_{l-1}| > (\sqrt{\ln d} - 4) \cdot \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{1}{\sqrt{d}}} > \sqrt{\frac{2}{\pi d}} \cdot \sqrt{\ln d} \cdot e^{-\frac{10}{\sqrt{\ln d}}},$$

contradicting our choice of l . So we have $l \leq \sqrt{\ln d} - 3$.

Now, if $|S_1| = |\Phi(S)| \geq 2\sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{10}{\sqrt{\ln d}}}$, then we are done. Otherwise, we have

$$|\Phi^{(\leq l)}(S) - \Phi(S)| = |S_1 \cup \dots \cup S_l| - |S_1| \geq \sqrt{\frac{2}{\pi d}} \cdot (\sqrt{\ln d} - 2) \cdot e^{-\frac{10}{\sqrt{\ln d}}}.$$

Thus since $l + 1 \leq \sqrt{\ln d} - 3$, we get

$$\frac{|\Phi^{(\leq l)}(S) - \Phi(S)|}{l + 1} \geq \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{10}{\sqrt{\ln d}}}. \quad \square$$

The two corollaries which follow are just the equivalents of Theorem 6.7 and Corollary 6.8 for K_n^d , under the correspondence between $[0, 1]^d$ and K_n^d described in Proposition 6.1.

COROLLARY 6.9. *Let n and d be positive integers, where d is sufficiently large. Let S be a subset of $V(K_n^d)$.*

$$\text{If } \left| |S| - \frac{n^d}{2} \right| \leq \sqrt{\frac{2}{\pi d}} \cdot \sqrt{\ln d} \cdot e^{-\frac{10}{\sqrt{\ln d}}} \cdot n^d, \text{ then } |\Phi(S)| \geq \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{1}{\sqrt{d}}} \cdot n^d.$$

COROLLARY 6.10. *Let n and d be positive integers, where d is sufficiently large. Let S be a subset of $V(K_n^d)$. Suppose*

$$\left(\frac{1}{2} - \sqrt{\frac{2}{\pi d}} \cdot \sqrt{\ln d} \cdot e^{-\frac{10}{\sqrt{\ln d}}} \right) \cdot n^d \leq |S| \leq \frac{1}{2} \cdot n^d.$$

Then either

$$|\partial(S)| \geq 2\sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{10}{\sqrt{\ln d}}} \cdot n^d$$

or there exists an integer $l \geq 2$ such that

$$\frac{|\partial^{(\leq l)}(S) - \partial(S)|}{l + 1} \geq \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{10}{\sqrt{\ln d}}} \cdot n^d.$$

Corollary 6.9 readily yields the following lower bound on the vertex bandwidth $B(K_n^d)$.

THEOREM 6.11. *Let n be a fixed positive integer. For all positive integers d we have $B(K_n^d) \geq (1 - o(1))\sqrt{\frac{2}{\pi d}} n^d$.*

Proof. Consider a labeling f of $V(K_n^d)$ that achieves the bandwidth. Let S denote the set of vertices receiving labels $1, 2, \dots, \lfloor n^d/2 \rfloor$. By Corollary 6.9, $|\partial(S)| \geq (1 - o(1))\sqrt{\frac{2}{\pi d}} n^d$. Thus, $B(K_n^d) = B(f) \geq |\partial(S)| \geq (1 - o(1))\sqrt{\frac{2}{\pi d}} n^d$. $\square$

In [15], Harper obtained the same lower bound on $B(K_n^d)$ when both n and d go to ∞ . For even n , the lower bound for $B(K_n^d)$ in Theorem 6.11 matches asymptotically the upper bound for $B(K_n^d)$ given in [15] as $d \rightarrow \infty$. Recently, Balogh et al. [6] obtained improved upper and lower bounds on $B(K_n^3)$.

We now establish a lower bound on $B'(K_n^d)$ that asymptotically matches the upper bound given in Theorem 5.10 when n is even. We follow the approach used in [9, 5].

THEOREM 6.12. *Let n be a fixed positive integer. We have*

$$B'(K_n^d) \geq (1 - o(1)) \sqrt{\frac{d}{2\pi}} n^d (n-1) \quad \text{as } d \rightarrow \infty.$$

Proof. Let f be an optimal labeling of $E(K_n^d)$ using labels $1, \dots, \binom{n}{2} dn^{d-1}$. Let S denote the set of edges receiving the first half of the labels. That is, S is the set of the edges receiving labels $1, 2, \dots, \frac{1}{2} \binom{n}{2} dn^{d-1}$. Let us call the edges in S red and the rest of the edges white.

For a vertex $x \in V(K_n^d)$, let $E(x)$ denote the set of edges incident to x . A vertex x is called *red* if all of the edges in $E(x)$ are red and *white* if all of the edges in $E(x)$ are white; otherwise, it is called *mixed*. Let R, W , and M denote the set of red, white, and mixed vertices, respectively.

We have

$$(8) \quad |R| + |W| + |M| = n^d.$$

For $x \in M$, let $r(x)$ denote the number of red edges in $E(x)$. Hence $1 \leq r(x) \leq d(n-1)$. By (double) counting the red edges and the white edges, we have

$$(9) \quad |R| \cdot d(n-1) + \sum_{x \in M} r(x) = \binom{n}{2} dn^{d-1} = |W| \cdot d(n-1) + \sum_{x \in M} (d(n-1) - r(x)).$$

It readily follows from (9) that $|R|, |W| < \frac{1}{2} n^d$. Note that the white edges incident to M belong to $\partial(S)$. Similarly, the red edges incident to M belong to $\partial(E(K_n^d) - S)$. Therefore, we have

$$(10) \quad |\partial(S)| \geq \frac{1}{2} \sum_{x \in M} (d(n-1) - r(x)) \quad \text{and} \quad |\partial(E(K_n^d) - S)| \geq \frac{1}{2} \sum_{x \in M} r(x).$$

By combining these two inequalities we obtain

$$(11) \quad B'(K_n^d) \geq \max\{|\partial(S)|, |\partial(E(K_n^d) - S)|\} \geq \frac{|M| \cdot (n-1)d}{4}.$$

If $|M| \geq 2\sqrt{\frac{2}{\pi d}} \cdot n^d \cdot e^{-\frac{10}{\sqrt{\ln d}}}$, then by (11) we have

$$B'(K_n^d) \geq \frac{|M| \cdot (n-1)d}{4} \geq \sqrt{\frac{d}{2\pi}} n^d (n-1) \cdot e^{-\frac{10}{\sqrt{\ln d}}} = (1 - o(1)) \sqrt{\frac{d}{2\pi}} n^d (n-1),$$

and we are done. Hence, we may assume that

$$(12) \quad |M| < 2\sqrt{\frac{2}{\pi d}} \cdot n^d \cdot e^{-\frac{10}{\sqrt{\ln d}}}.$$

Either

$$\sum_{x \in M} r(x) \leq |M| \cdot d(n-1)/2 \quad \text{or} \quad \sum_{x \in M} (d(n-1) - r(x)) \leq |M| \cdot d(n-1)/2.$$

Without loss of generality, let us assume that the first inequality holds, since otherwise we could switch the roles of red and white vertices. By combining this with (9) and (12), we have

$$(13) \quad \binom{n}{2} dn^{d-1} \leq |R| \cdot d(n-1) + |M| \cdot d(n-1)/2 \leq |R| \cdot d(n-1) + \frac{\sqrt{2}}{\sqrt{\pi d}} n^d \cdot d(n-1).$$

This yields the lower bound

$$\left(\frac{1}{2} - \sqrt{\frac{2}{\pi d}}\right) n^d \leq |R| < \frac{1}{2} n^d.$$

By Corollary 6.10, we have either

$$(14) \quad |\partial(R)| \geq 2 \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{10}{\sqrt{\ln d}}} \cdot n^d$$

or there exists an integer $l \geq 2$ such that

$$(15) \quad \frac{|\partial^{(\leq l)}(R) - \partial(R)|}{l+1} \geq \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{10}{\sqrt{\ln d}}} \cdot n^d.$$

Now clearly $\partial(R) \subseteq M$. So if (14) holds, then we have

$$|M| \geq 2 \sqrt{\frac{2}{\pi d}} \cdot n^d \cdot e^{-\frac{10}{\sqrt{\ln d}}},$$

contradicting (12). Hence, we may assume that (15) holds instead. Let $C = \partial^{(\leq l)}(R) - \partial(R)$. Let $E(C)$ denote the set of edges incident to C . Recalling that S is the set of red edges, we have $E(C) \subseteq \partial^{(\leq l+1)}(S)$. Note that $|E(C)| \geq |C| \cdot d(n-1)/2$. By applying Proposition 2.2 to the line graph $L(K_n^d)$, we have

$$\begin{aligned} B'(K_n^d) &\geq B'(f) \geq \frac{|\partial^{(\leq l+1)}(S)|}{(l+1)} \geq \frac{|E(C)|}{(l+1)} \geq \frac{|C| \cdot d(n-1)}{2(l+1)} \\ &= \frac{|\partial^{(\leq l)}(R) - \partial(R)| \cdot d(n-1)}{2(l+1)} \\ &\geq \sqrt{\frac{2}{\pi d}} \cdot e^{-\frac{10}{\sqrt{\ln d}}} \cdot n^d \cdot \frac{d(n-1)}{2} \quad (\text{by (15)}) \\ &= (1 - o(1)) \sqrt{\frac{d}{2\pi}} \cdot n^d (n-1). \quad \square \end{aligned}$$

7. Concluding remarks. 1. We thank one of the referees for pointing us to the work of Pellegrini [26]. There the weight function $l(n, d, r)$ is expressed as an exponential sum, using the recurrence in Proposition 3.1 and standard results in Fourier series. This led to [26, Proposition 3]

$$l(n, d, r) = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \cos(x(2r - d(n-1))) \left(\frac{\sin(nx)}{\sin x} \right)^d dx.$$

Recall that $l^*(n, d) = l(n, d, \lfloor \frac{n(d-1)}{2} \rfloor)$. So by letting $r = \lfloor \frac{n(d-1)}{2} \rfloor$ and adapting Pellegrini's application of Laplace's method to estimate this integral for fixed n and $d \rightarrow \infty$, we obtain

$$(16) \quad l^*(n, d) = \sqrt{\frac{6}{\pi d} \left(\frac{n^2}{n^2 - 1} \right)} n^{d-1} + o(n^{d-1}), \quad \text{when } n \text{ is fixed and } d \rightarrow \infty.$$

We note that in Pellegrini's application it is important to have n fixed and $d \rightarrow \infty$. Since we are concerned with the case when d is fixed and $n \rightarrow \infty$, we cannot use (16) in our result. Formula (16) will be useful when one studies $B'(P_n^d)$ when n is fixed and $d \rightarrow \infty$.

2. We have asymptotically determined $B'(K_n^d)$ for fixed even n and $d \rightarrow \infty$, showing that $B' \sim \sqrt{\frac{d}{2\pi}} n^d (n-1)$. In the upper bound labeling, the evenness of n made it possible to partition K_n^d into 2^d many orthants, each isomorphic to $K_{n/2}^d$.

When $n = 2m + 1$ is odd, a similar partition is possible. Here an orthant corresponding to a weight k vertex of Q_d is the Cartesian product of k copies of K_{m+1} and $d - k$ copies of K_m . By applying the same labeling scheme, and noting that a maximum dilation occurs at a pair of cross edges incident to a vertex in an orthant associated with a hypercube node of weight $\lfloor \frac{(m+1)d}{2m+1} \rfloor$, we obtain an upper bound

$$B'(K_n^d) \leq \frac{2m+1}{2m} \sqrt{\frac{m}{m+1}} \sqrt{\frac{d}{2\pi}} n^d (n-1) (1 + o(1)),$$

when $n = 2m + 1$ is fixed and $d \rightarrow \infty$.

This upper bound is within a factor of $\frac{2m+1}{2m} \sqrt{\frac{m}{m+1}}$ from the lower bound in Theorem 6.12. Note that this factor is close to 1 and tends to 1 as $m \rightarrow \infty$. It would be interesting to determine $B'(K_n^d)$ asymptotically when n is odd and $d \rightarrow \infty$. We suspect that the upper bound given here is closer to the truth.

Appendix. Proof of Lemma 6.6.

Proof. We recall Stirling's formula that, for all integers $n \geq 1$, $n! = \sqrt{2\pi n} \cdot (n/e)^n \cdot e^{\frac{\theta}{12n}}$ for some $\theta = \theta(n)$, with $0 < \theta < 1$. We have for some $\theta_1, \theta_2, \theta_3 \in (0, 1)$

$$\begin{aligned} Pr(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &= \frac{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n}{\sqrt{2\pi k} \left(\frac{k}{e}\right)^k \sqrt{2\pi(n-k)} \left(\frac{n-k}{e}\right)^{n-k}} \cdot p^k \cdot (1-p)^{n-k} \cdot e^{\frac{\theta_1}{12n} - \frac{\theta_2}{12k} - \frac{\theta_3}{12(n-k)}} \\ (17) \quad &\geq \frac{1}{2\pi} \cdot \sqrt{\frac{n}{k(n-k)}} \cdot \left(\frac{np}{k}\right)^k \cdot \left(\frac{n-np}{n-k}\right)^{n-k} \cdot e^{-\frac{1}{12k} - \frac{1}{12(n-k)}}. \end{aligned}$$

Let $a = np - k$. Then $np = k + a$ and $|a| \leq \sqrt{\ln n}$ by our assumption. We have

$$\begin{aligned} (18) \quad \left(\frac{np}{k}\right)^k \cdot \left(\frac{n-np}{n-k}\right)^{n-k} &= \left(1 + \frac{a}{k}\right)^k \left(1 - \frac{a}{n-k}\right)^{n-k} \geq e^{-\frac{a^2}{k} - \frac{a^2}{n-k}} \\ &\geq e^{-\ln n \left(\frac{1}{k} + \frac{1}{n-k}\right)}, \end{aligned}$$

where the first inequality follows from Lemma 6.5. By (17) and (18), we have for large n

$$\begin{aligned} (19) \quad Pr(X = k) &\geq \frac{1}{\sqrt{2\pi}} \cdot \sqrt{\frac{n}{k(n-k)}} \cdot e^{-(\ln n + \frac{1}{12}) \left(\frac{1}{k} + \frac{1}{n-k}\right)} \\ &\geq \frac{1}{\sqrt{2\pi}} \cdot \sqrt{\frac{n}{k(n-k)}} \cdot e^{-2 \ln n \left(\frac{1}{k} + \frac{1}{n-k}\right)}. \end{aligned}$$

Note that $k(n-k)$ is unimodal and is maximized at $k = n/2$. We consider two cases.

Case 1. $k \leq n^{\frac{3}{4}}$ or $n - k \leq n^{\frac{3}{4}}$. In this case, $k(n - k) \leq n^{\frac{3}{4}} \cdot n$. So $\frac{n}{k(n-k)} \geq \frac{1}{n^{\frac{3}{4}}}$. Also, $4\sqrt{n} \leq k \leq d - 4\sqrt{\ln n}$. So $\frac{n}{k(n-k)} \leq \frac{n}{4\sqrt{\ln n}(n-4\sqrt{\ln n})} \leq \frac{1}{\sqrt{\ln n}}$. By (19), we have

$$Pr(X = k) \geq \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{n^{\frac{3}{8}}} \cdot e^{-2 \ln n \cdot \frac{1}{\sqrt{\ln n}}} = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{n^{\frac{3}{8}}} \cdot \frac{1}{d^{\frac{2}{\sqrt{\ln n}}}} \geq \sqrt{\frac{2}{\pi n}} \cdot e^{-\frac{1}{\sqrt{n}}} \quad \text{for large } n.$$

Case 2. $n^{\frac{3}{4}} \leq k \leq n - n^{\frac{3}{4}}$. In this case, we have $\frac{n}{k(n-k)} \geq \frac{n}{n^2/4} = \frac{4}{n}$ and $\frac{n}{k(n-k)} \leq \frac{n}{n^{\frac{3}{4}}(n-n^{\frac{3}{4}})} \leq \frac{2}{n^{\frac{3}{4}}}$. By (19),

$$Pr(X = k) \geq \frac{1}{\sqrt{2\pi}} \cdot \sqrt{\frac{4}{n}} \cdot e^{-2 \ln n \cdot \frac{2}{n^{\frac{3}{4}}}} \geq \sqrt{\frac{2}{\pi n}} \cdot e^{-\frac{1}{\sqrt{n}}} \quad \text{for large } n. \quad \square$$

Acknowledgments. The authors thank both referees for their careful reading and for their valuable suggestions. The authors also thank Dan Pritikin for stimulating discussions and Doron Zeilberger and Sergei Bezrukov for helpful communications and references.

REFERENCES

- [1] I. ANDERSON, *On primitive sequences*, J. London Math. Soc., 42 (1967), pp. 137–148.
- [2] I. ANDERSON, *A variance method in combinatorial number theory*, Glasg. Math. J., 10 (1969), pp. 126–129.
- [3] I. ANDERSON, *Combinatorics of Finite Sets*, Oxford University Press, New York, 1987.
- [4] E. APPELT, *On the Bandwidth of a Product of Complete Graphs*, M.A. thesis, Miami University, Oxford, OH, 2002.
- [5] J. BALOGH, D. MUBAYI, AND A. PLUHÁR, *On the edge-bandwidth of graph products*, Theoret. Comput. Sci., 359 (2006), pp. 43–57.
- [6] J. BALOGH, S. BEZRUKOV, L. HARPER, AND A. SERESS, *On the Bandwidth of 3-dimensional Hamming Graphs*, submitted.
- [7] T. BERGER-WOLF AND E. M. REINGOLD, *Index assignment for multichannel communication under failure*, IEEE Trans. Inform. Theory, 48 (2002), pp. 2656–2668.
- [8] B. BOLLOBÁS AND I. LEADER, *Compressions and isoperimetric inequalities*, J. Combin. Theory Ser. A, 56 (1991), pp. 47–62.
- [9] T. CALAMONERI, A. MASSINI, AND I. VRTO, *New results on edge-bandwidth*, Theoret. Comput. Sci., 307 (2003), pp. 503–513.
- [10] P. Z. CHINN, J. CHVÁTALOVÁ, A. K. DEWDNEY, AND N. E. GIBBS, *The bandwidth problem for graphs and matrices—A survey*, J. Graph Theory, 6 (1982), pp. 223–254.
- [11] F. R. K. CHUNG, *Labeling of graphs*, in Selected Topics in Graph Theory 3, L. W. Beineke and R. J. Wilson, eds., Academic Press, San Diego, 1988, pp. 151–168.
- [12] A. GUPTA, *private communication*.
- [13] L. H. HARPER, *Optimal numberings and isoperimetric problems on graphs*, J. Combin. Theory, 1 (1966), pp. 385–393.
- [14] L. H. HARPER, *On an isoperimetric problems for Hamming graphs*, Discrete Appl. Math., 95 (1999), pp. 285–309.
- [15] L. H. HARPER, *On the bandwidth of a Hamming graph*, Theoret. Comput. Sci., 301 (2003), pp. 491–498.
- [16] U. HENDRIK AND M. STIEBITZ, *On the bandwidth of graph products*, J. Inform. Process. Cybernet., 28 (1992), pp. 113–125.
- [17] R. HOCHBERG, C. MCDIARMID, AND M. SAKS, *On the bandwidth of triangulated triangles*, Discrete Math., 138 (1995), pp. 261–265.
- [18] F. K. HWANG AND J. C. LAGARIAS, *Minimum range sequences of all k -subsets of a set*, Discrete Math., 19 (1977), pp. 257–264.
- [19] T. JIANG, D. MUBAYI, A. SHASTRI, AND D. B. WEST, *Edge-bandwidth of graphs*, SIAM J. Discrete Math., 12 (1999), pp. 307–316.
- [20] P. LAM, W. SHIU, W. CHAN, AND Y. LIN, *On the bandwidth of convex triangulated meshes*, Discrete Math., 173 (1997), pp. 285–289.

- [21] T. LEIGHTON AND C. G. PLAXTON, *Hypercubic sorting networks*, SIAM J. Comput., 27 (1998), pp. 1–47.
- [22] H. S. MOGHADAM, *Compression Operators and a Solution to the Bandwidth Problem of the Product of n Paths*, Ph.D. thesis, University of California at Riverside, Riverside, CA, 1983.
- [23] H. S. MOGHADAM, *Bandwidth of the product of n paths*, Congr. Numer., 173 (2005), pp. 3–15.
- [24] B. MONIEN AND L. H. SUDBOROUGH, *Embedding one interconnection network in another*, in Computational Graph Theory, Comput. Suppl. 7, Springer, Vienna, 1990, pp. 257–282.
- [25] A. ODLYZKO, *Asymptotic Enumeration Methods*, Handbook of Combinatorics 1, 2, Elsevier, Amsterdam, 1995, pp. 1063–1229.
- [26] F. PELLEGRINI, *Bounds for the bandwidth of the d -ary deBruijn graph*, Parallel Process. Lett., 3 (1993), pp. 431–443.
- [27] O. PIKHURKO AND J. WOJCIECHOWSKI, *Edge-bandwidth of grids and tori*, Theoret. Comput. Sci., 369 (2006), pp. 35–43.
- [28] W. UNGER, *The complexity of the approximation of the bandwidth problem*, in Proceedings of the 1998 Annual Symposium on Foundations of Computer Science, IEEE Press, Baltimore, 1998, pp. 82–91.

POLYCHROMATIC COLORINGS OF SUBCUBES OF THE HYPERCUBE*

DAVID OFFNER†

Abstract. Alon, Krech, and Szabó [*SIAM J. Discrete Math.*, 21 (2007), pp. 66–72] called an edge-coloring of a hypercube with p colors such that every subcube of dimension d contains every color a d -polychromatic p -coloring. Denote by p_d the maximum number of colors with which it is possible to d -polychromatically color *any* hypercube. We find the exact value of p_d for all values of d .

Key words. coloring of graphs and hypergraphs, extremal problems, Ramsey theory

AMS subject classifications. 05C15, 05C35, 05D10

DOI. 10.1137/07068014X

1. Introduction. Denote by Q_n the n -dimensional hypercube, and, for fixed graphs G, H , denote by $\text{ex}(G, H)$ the maximum number of edges in a subgraph of G which does not contain a copy of H . Problems about this quantity are called Turán-type problems, after P. Turán, who gave an exact value for $\text{ex}(G, H)$ when G and H are cliques. There is a rich history of studying Turán-type problems where $G = Q_n$. To cite some well-known examples, Erdős [7] conjectured that $\text{ex}(Q_n, Q_2) = (\frac{1}{2} + o(1))e(Q_n)$. Currently it is known that

$$\frac{1}{2}(n + \sqrt{n})2^{n-1} \leq \text{ex}(Q_n, Q_2) \lesssim .623e(Q_n).$$

The lower bound is due to Brass, Harborth, and Neinborg [4] and the upper bound to Chung [5]. Bialostocki [3] proved that any subgraph of the cube not containing Q_2 as a subgraph and intersecting every Q_2 has cardinality at most $(n + \sqrt{n})2^{n-2}$. As one generalization of Erdős' problem, Erdős [7], Chung [5], Conder [6], and Alon et al. [2] also looked at the size of $\text{ex}(Q_n, C_{2k})$, where C_{2k} is a cycle of even length.

More recently, Alon, Krech, and Szabó [1] asked how large $\text{ex}(Q_n, Q_d)$ is and studied related problems. In particular, they called an edge-coloring of a hypercube with p colors such that every subcube of dimension d contains every color a d -polychromatic p -coloring. Denote by p_d the maximum number of colors with which it is possible to d -polychromatically color *any* hypercube. (Note that the value of p_d gives a lower bound on $\text{ex}(Q_n, Q_d)$; in particular, $e(Q_n)(1 - 1/p_d) \leq \text{ex}(Q_n, Q_d)$.) They proved the following.

THEOREM 1.

$$\binom{d+1}{2} \geq p_d \geq \begin{cases} \frac{(d+1)^2}{4} & \text{if } d \text{ is odd,} \\ \frac{d(d+2)}{4} & \text{if } d \text{ is even.} \end{cases}$$

The lower bound was obtained by construction and the upper bound by a surprising application of Ramsey's theorem, one version of which states that for all r, k there exists an n such that, if the edges of a d -uniform hypergraph on at least n vertices are

*Received by the editors January 15, 2007; accepted for publication (in revised form) September 6, 2007; published electronically March 20, 2008.
<http://www.siam.org/journals/sidma/22-2/68014.html>

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213 (offner@cmu.edu).

colored with k colors, there exists a complete monochromatic subgraph on r vertices. See [8] for a full treatment.

Alon, Krech, and Szabó asked whether it was possible to resolve the behavior of p_d . We will prove that for all d , p_d is equal to the lower bound in Theorem 1, i.e., the following.

THEOREM 2.

$$p_d = \begin{cases} \frac{(d+1)^2}{4} & \text{if } d \text{ is odd,} \\ \frac{d(d+2)}{4} & \text{if } d \text{ is even.} \end{cases}$$

2. Notation. We follow the notation of [1] in denoting vertices, edges, and subcubes: A vertex of Q_n is represented by a vector of n bits. An edge corresponding to two vertices whose vectors differ in one coordinate is represented by a bit vector with a $*$ in that coordinate. A d -dimensional subcube D is represented by a bit vector with d $*$'s, where the vertices of D are those vectors obtained by replacing the $*$'s with any combination of bits.

3. Proof of Theorem 1. For completeness, since the proof of Theorem 2 will use ideas from the proof of Theorem 1 in [1], and since the proof is very nice, we sketch it here.

Proof of Theorem 1 (lower bound). Assume that d is odd. The following $\frac{(d+1)^2}{4}$ -coloring of Q_n is d -polychromatic: Color the edges with elements of $\mathbf{Z}_{\frac{d+1}{2}} \times \mathbf{Z}_{\frac{d+1}{2}}$ as follows: If an edge has a 1's to the left of its $*$, and b 1's to the right of its $*$, assign it the color $(a \pmod{\frac{d+1}{2}}, b \pmod{\frac{d+1}{2}})$. For any subcube Q_d , find the $*$ with exactly $\frac{d-1}{2}$ $*$'s on each side of it. Note that, for this subcube, the set of edges with this $*$ contains all colors. The construction is the same if d is even, except a is taken mod $\frac{d}{2}$ and b is taken mod $\frac{d+2}{2}$, and for any subcube Q_d we consider the edges whose star is the $\frac{d}{2}$ th from the left. $\square$

Proof of Theorem 1 (upper bound). Suppose that we have a d -polychromatic p -coloring c of Q_n , with n huge. We will use Ramsey's theorem for d -uniform hypergraphs with $p^{d^{d-1}}$ colors. We define a $p^{d^{d-1}}$ -coloring of the d -subsets of $[n]$. Fix an arbitrary ordering of the edges of Q_d . For an arbitrary subset S of the indices, define $\text{cube}(S)$ to be the subcube whose $*$ coordinates are at the positions of S and all other coordinates are 0. Let S be a d -subset of $[n]$, and define the color of S to be the vector whose coordinates are the c -values of the edges of the d -dimensional subcube $\text{cube}(S)$ (according to our fixed ordering of the edges of Q_d). By Ramsey's theorem, if n is large enough, there is a set $T \subseteq [n]$ of $d^2 + d - 1$ coordinates such that the color-vector is the same for any d -subset of T . Fix a set S of d particular coordinates from T : those which are the (id) th elements of T for $i = 1, \dots, d$. Any two elements of S have at least $d - 1$ elements of T in between.

Call a hypercube coloring *Ramsey* if the color of an edge is determined by the number of 1's to the left of its $*$ and the number of 1's to the right of its $*$.

LEMMA 3. *The coloring of $\text{cube}(S)$ is Ramsey.*

Proof. Let e_1 and e_2 be two edges of $\text{cube}(S)$ such that they have the same number of 1's to the left of their respective $*$ and the same number of 1's to the right as well. Since there are at least $d - 1$ elements in T in between each coordinate of S , it is possible for us to find a set of d coordinates $S' \subseteq T$ and an edge e_3 of $\text{cube}(S')$ such that

- (i) e_3 is the same edge when restricted to S as e_1 and
- (ii) e_3 occupies the same position in the ordering of edges in $\text{cube}(S')$ as e_2 occupies in $\text{cube}(S)$.

Thus $c(e_1) = c(e_3) = c(e_2)$, and the lemma is proved. $\square$

To finish the proof of the upper bound in Theorem 1, note that there are exactly $1 + 2 + \cdots + d = \binom{d+1}{2}$ ways to separate at most $d - 1$ 1's by a *. By Lemma 3, a d -polychromatic edge-coloring is not possible with more colors. $\square$

4. Proof of Theorem 2.

LEMMA 4. *If a d -polychromatic p -coloring exists for any n , then one of Ramsey-type exists.*

Proof. In the proof of the upper bound of Theorem 1, we showed that, in a sufficiently large hypercube, a Ramsey-colored d -dimensional subcube exists by applying Ramsey's theorem to a $p^{\binom{d+1}{2}}$ -coloring of the d -subsets of the indices of $[n]$. To prove that, for a sufficiently large hypercube, a subcube of any dimension k exists with a Ramsey coloring, simply apply Ramsey's theorem to a $p^{\binom{k+1}{2}}$ -coloring of the k -subsets of $[n]$, defined in an analogous way, and proceed in an identical fashion. Thus for any $p \leq p_d$, one can show the existence of a Ramsey p -colored d -polychromatic hypercube of dimension k by taking any sufficiently large p -colored d -polychromatic hypercube and applying Ramsey's theorem. $\square$

Proof of Theorem 2. Fix d . By Lemma 4, we may consider a Ramsey d -polychromatic p_d -coloring on an arbitrarily large hypercube. Since the coloring is Ramsey, every edge belongs to some color class (a, b) , where there are a 1's to the left of the * and b 1's to the right. It will be convenient to arrange the color classes in a triangular array, with $(0, 0)$ in the first row, $(0, 1)$ and $(1, 0)$ in the second row, and so forth. Define the i th row to be those color classes (a, b) , with $a + b = i$, and the i th diagonal to be those classes of the form (i, j) for any j . Define an $i \times j$ rectangle to be a set of color classes of the following form: $\{(a + \alpha, b + \beta) : 0 \leq \alpha < j, 0 \leq \beta < i\}$, and say a rectangle R is located at the i th diagonal if i is the least diagonal that intersects R . Define a region to be all color classes contained in some consecutive rows and consecutive diagonals.

In this configuration, a d -cube includes the following color classes, from left to right: a $d \times 1$ rectangle, corresponding to the colors of edges using the leftmost *, followed by a $(d - 1) \times 2$ rectangle, corresponding to the colors of edges using the second-leftmost *, etc., until a $1 \times d$ rectangle corresponding to the rightmost *. All of these rectangles cover colors in the same d rows, and they may overlap. They must come in order from left to right, however, since, e.g., if there are a 1's to the left of some *, there are at least a 1's to the left of any * further to the right. Call such a sequence of rectangles a *cube sequence*, and denote them (left to right) by $r_1, r_2, \dots, r_d$. Note that any cube sequence corresponds to color classes covered by at least one cube. See the examples in Figure 1.

Choose a set of d consecutive very long rows. By our assumption that the coloring is d -polychromatic, it is not possible to find a cube sequence in these rows where every rectangle in the sequence lacks a particular color. In particular, imagine picking a color A and scanning the rows from left to right, looking for a copy of r_1 not containing A . If we find one, say, at diagonal d_1 , we note that all copies of r_1 to the left of d_1 contain A . Continuing from left to right, we might then seek a copy of r_2 not containing A , and if the first one we find is at diagonal d_2 , then we note that every copy of r_2 between d_1 and d_2 contains A . By continuing in this manner, we may partition the rows into at most d regions where, in the i th region, every copy of r_i contains A . By repeating this for the other colors, we note that each color partitions the rows into at most d regions, where, in the i th region, each copy of r_i contains the color. By taking the intersections of all of these regions, the rows are partitioned into at most $d \cdot p_d$ regions such that for each region, for every color A , there is a rectangle

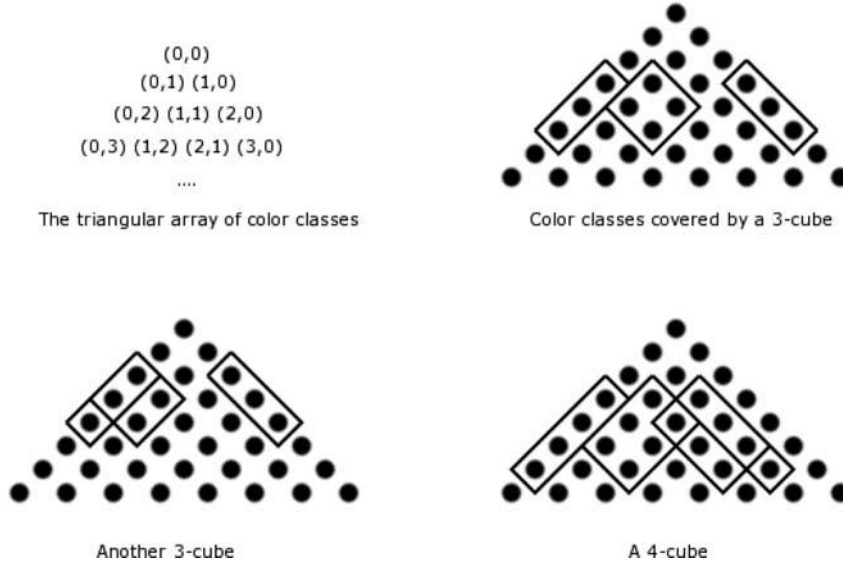


FIG. 1. The triangular array of color classes and some cube sequences.

r_A such that every copy of r_A in the region contains A . In particular, since we may choose the rows as long as we like, we can find an arbitrarily long region with this property.

Now consider such a region with m diagonals, where $m \gg d$. For each color A , assign d variables, $\{A_i\}_{i=1}^d$ corresponding to the number of times A appears in the i th row in our region. Note that, if A appears in the top or bottom row, it can be contained in at most one copy of r_A . If it appears in the second or the $d-1$ st row, it can be contained in at most two copies of r_A , and so forth. There are at least $m-d$ copies of r_A in the region, and thus

$$A_1 + 2A_2 + 3A_3 + \cdots + \left\lceil \frac{d}{2} \right\rceil A_{\lceil \frac{d}{2} \rceil} + \cdots + 2A_{d-1} + A_d \geq m - d.$$

Since $A_i + B_i + \cdots = m$, if we add up these equations for each color, we get

$$m + 2m + 3m + \cdots + \left\lceil \frac{d}{2} \right\rceil m + \cdots + 2m + m \geq p_d(m - d).$$

The left-hand side is equal to $\frac{(d+1)^2}{4}m$ if d is odd and $\frac{d(d+2)}{4}m$ if d is even. Divide both sides by m , and note that $(m-d)/m$ can be as close to 1 as desired to obtain the result. $\square$

This method may be used to immediately give other similar results. For instance, say two edges in a subcube belong to the same *layer* if their vector representation contains the same number of 1's. Then, if $q_{d,k}$ ($k \leq d$) is the maximum number such that any Q_n can be edge-colored so that the k th layer of any sub- Q_d contains all colors, then $q_{d,\lceil d/2 \rceil} = \lceil d/2 \rceil$ (e.g., $q_{3,2} = 2$).

Acknowledgments. The author thanks Oleg Pikhurko and Peter Lumsdaine for helpful discussions and Tibor Szabó and an anonymous referee for comments on the paper.

REFERENCES

- [1] N. ALON, A. KRECH, AND T. SZABÓ, *Turán's theorem in the hypercube*, SIAM J. Discrete Math., 21 (2007), pp. 66–72.
- [2] N. ALON, R. RADOIČIĆ, B. SUDAKOV, AND J. VONDRÁK, *A Ramsey-type result for the hypercube*, J. Graph Theory, 53 (2006), pp. 196–208.
- [3] A. BIALOSTOCKI, *Some Ramsey type results regarding the graph of the n -cube*, Ars Combin., 16-A (1983), pp. 39–48.
- [4] P. BRASS, H. HARBORTH, AND H. NEINBORG, *On the maximum number of edges in a C_4 -free subgraph of Q_n* , J. Graph Theory, 19 (1995), pp. 17–23.
- [5] F. CHUNG, *Subgraphs of a hypercube containing no small even cycles*, J. Graph Theory, 16 (1992), pp. 273–286.
- [6] M. CONDER, *Hexagon-free subgraphs of hypercubes*, J. Graph Theory, 17 (1993), pp. 477–479.
- [7] P. ERDŐS, *Some Problems in Graph Theory*, *Combinatorial Analysis and Combinatorial Number Theory*, in Graph Theory and Combinatorics, B. Bollobás, ed., Academic Press, New York, 1984, pp. 1–17.
- [8] R. GRAHAM, B. ROTHCHILD, AND J. SPENCER, *Ramsey Theory*, 2nd ed., John Wiley & Sons, New York, 1990.

MATROID COMPLEXITY AND NONSUCCINCT DESCRIPTIONS*

DILLON MAYHEW†

Abstract. We investigate an approach to matroid complexity that involves describing a matroid via a list of independent sets, bases, circuits, or some other family of subsets of the ground set. The computational complexity of algorithmic problems under this scheme appears to be highly dependent on the choice of input type. We define an order on the various methods of description, and we show how this order acts upon 10 types of input. We also show that under this approach several natural algorithmic problems are complete in classes thought not to be equal to P.

Key words. matroids, complexity, isomorphism, matroid minors

AMS subject classifications. 05B35, 68Q17

DOI. 10.1137/050640576

1. Introduction. The study of matroid-theoretical algorithmic problems and their complexity has been dominated by two approaches. The first approach was implicitly used by Edmonds [1] and later developed by Hausmann and Korte [4, 5]. It uses the idea of an ordinary Turing machine augmented with an oracle. Suppose that the subject of the computation is a matroid on the ground set E . When queried about a subset, X , of E , the oracle returns in unit time some information about X . That information might be the rank of X , or an answer to the question, “Is X independent?” In this case the oracle is known as the rank or independence oracle, respectively.

The second dominant approach to matroid complexity uses the standard model of a Turing machine but considers as its input only a restricted class of matroids that can be represented by some “succinct” structure, for instance, a graph, or a matrix over a field.

A third approach to the study of matroid complexity would, in some ways, be more natural. A matroid is essentially a finite set with a structured family of subsets (we shall not consider infinite matroids). An obvious way to describe a matroid to a Turing machine is to simply list the subsets that belong to this family. The advantage of this approach is that any matroid could be received as input, and not merely those matroids which belong to some restricted class.

The reason this approach has not received as much attention as the latter two is the concern that the input will be too large. A result of Knuth [8] says that if $f(n)$ is the number of nonisomorphic matroids on a set of size n , then there is a constant c such that

$$\log_2 \log_2 f(n) \geq n - \frac{3}{2} \log_2 n + c \log_2 \log_2 n$$

for sufficiently large values of n . It follows that if Σ is a finite alphabet and $\sigma : \mathcal{M} \rightarrow \Sigma^*$ is an injective encoding scheme that takes the set of all matroids to words in Σ , then there can be no polynomial function p such that $|\sigma(M)|$, the length of the word $\sigma(M)$, is bounded above by $p(|E(M)|)$ for every matroid M . Thus if we wish to use

*Received by the editors September 18, 2005; accepted for publication (in revised form) September 6, 2007; published electronically March 20, 2008.

<http://www.siam.org/journals/sidma/22-2/64057.html>

†Mathematical Institute, University of Oxford, St. Giles, OX1 3LB Oxford, UK, and School of Mathematics, Victoria University of Wellington, P.O. Box 600, Wellington 6140, New Zealand (dillon.mayhew@mcs.vuw.ac.nz).

this type of encoding function we must in some sense abandon the cardinality of the ground set of a matroid as a measure of its “size.” Historically the concern with this type of scheme has been that, because of the large size of the input, the class of problems that can be solved in polynomial time will be artificially inflated, and that therefore, in a trivial way, all algorithmic problems for matroids will be tractable.

In this paper we show that the situation is apparently more subtle than this. We show that several natural matroid problems are complete in classes thought not to be equal to P, even using an encoding scheme that works for all matroids; for instance, the scheme that describes a matroid by listing its independent sets.

A quirk of this approach to matroid complexity is that the difficulty of a computational problem appears to vary widely according to the type of input. A problem may be solvable in polynomial time if the input takes the form of a list of bases, but if the input is a list of circuits, the same problem may be NP-complete.

Before we examine the complexity of matroid-theoretical problems, we define an order on types of input, and we show how this order acts upon a set of 10 natural methods of description. This work is an analogue of that done by Hausmann and Korte [6] and Robinson and Welsh [12]. Both of these works compare different types of oracles.

Our references for basic concepts, notation, and terminology will be Oxley [11] with regards to matroids and Garey and Johnson [2] with regards to complexity theory.

2. Various types of inputs. The 10 types of input that we consider are as follows: RANK, INDEPENDENT SETS, SPANNING SETS, BASES, FLATS, CIRCUITS, HYPERPLANES, NONSPANNING CIRCUITS, DEPENDENT HYPERPLANES, and CYCLIC FLATS.

Of these, some need little explanation. A list of the independent sets, spanning sets, bases, flats, circuits, or hyperplanes of a matroid uniquely specifies that matroid. Thus the corresponding forms of input will consist simply of lists of the appropriate subsets.

The RANK input will list each subset of the ground set, along with its rank. The NONSPANNING CIRCUITS input for a matroid M will specify the rank of M and list all its nonspanning circuits. Dually, the DEPENDENT HYPERPLANES input will specify the rank of M and list its dependent hyperplanes.

A *cyclic flat* is a flat that is also a (possibly empty) union of circuits. It is known that listing the cyclic flats and specifying their ranks completely determines a matroid. Therefore the CYCLIC FLATS input will list each cyclic flat, along with its rank.

(Note that we have not specified how to describe some exceptional cases, such as a matroid M with no nonspanning circuits. In this particular case we will assume that the NONSPANNING CIRCUITS description lists only the rank of M . Other exceptional cases are easily dealt with in a similar way.)

Suppose that f and g are two functions on the positive integers. If there exist constants, c_1 and c_2 , and an integer, N , such that $c_1g(n) \leq f(n) \leq c_2g(n)$ for every positive integer $n \geq N$, then we shall write $f = \Theta(g)$. Equivalently, $f = \Theta(g)$ if and only if $f = O(g)$ and $g = O(f)$.

Suppose that M is a matroid on a ground set of size n , and that INPUT is one of the types of input discussed above. Let (M, INPUT) be a word that describes M via INPUT. We shall assume that a subset of the ground set is specified by its characteristic vector. Thus if the INPUT description involves listing i subsets of $E(M)$, then $|(M, \text{INPUT})| \geq ni$. We shall consider only “reasonable” encoding schemes

(i.e., those that do not allow, for instance, padding of words). It follows easily that $|(M, \text{INPUT})| = \Theta(ni)$.

Obviously there are many other natural ways of describing a matroid, but many are related in a fairly trivial way to one of the methods we have already discussed. For example, a matroid can be described by listing its cocircuits, but the cocircuits are exactly the complements of the hyperplanes.

3. A comparison of inputs. It is natural to ask whether one form of description is intrinsically more compact than another. In this section we attempt to answer that question.

DEFINITION 3.1. *Suppose that INPUT1 and INPUT2 are two methods for describing a matroid. Then $\text{INPUT1} \leq \text{INPUT2}$ if there exists a polynomial-time Turing machine which will produce $(M, \text{INPUT2})$ given $(M, \text{INPUT1})$ for any matroid M .*

Suppose that $\text{INPUT1} \leq \text{INPUT2}$. If a problem is in P for descriptions via INPUT2, then clearly it is in P for INPUT1. Similarly, if a problem is NP-complete for INPUT1, then the same problem is NP-hard for INPUT2.

It is clear that $\leq$ is both reflexive and transitive. The rest of this section will be devoted to proving Theorem 3.2.

THEOREM 3.2. *The 10 types of input listed in section 2 are ordered by $\leq$ according to the Hasse diagram in Figure 1.*

Theorem 3.2 will follow from the subsequent lemmas.

LEMMA 3.3. $\text{RANK} \leq \text{SPANNING SETS}$ and $\text{RANK} \leq \text{INDEPENDENT SETS}$.

Proof. To find $(M, \text{SPANNING SETS})$ given (M, RANK) , a Turing machine need only write onto its output those subsets whose rank equals $r(M)$. Similarly, to produce $(M, \text{INDEPENDENT SETS})$ the machine need only write those subsets of $E(M)$ whose rank is equal to their size. $\square$

LEMMA 3.4. $\text{SPANNING SETS} \leq \text{BASES}$ and $\text{INDEPENDENT SETS} \leq \text{BASES}$.

Proof. The bases of a matroid are exactly the minimal spanning sets and the maximal independent sets, so the result follows easily. $\square$

LEMMA 3.5. $\text{INDEPENDENT SETS} \leq \text{FLATS}$.

Proof. Given the list of independent sets of M , we can create the list of flats by finding, in turn, the closure of each independent set, and then eliminating duplications. To find the closure of the independent set I , we check each set of the form $I \cup e$, where $e \notin I$, to see if it is independent. The element e belongs to $\text{cl}(I)$ if and only if $I \cup e$ is dependent. Since $|(M, \text{INDEPENDENT SETS})| = \Theta(ni)$, where $|E(M)| = n$ and the number of independent sets is i , it is easy to see that this entire procedure can be accomplished in polynomial time. $\square$

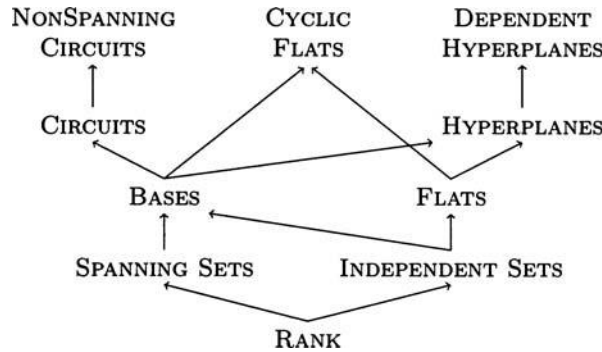


FIG. 1. An ordering of input types.

LEMMA 3.6. BASES $\leq$ CIRCUITS.

Proof. If B is a basis of a matroid M and $e \in E(M) - B$, then $B \cup e$ contains a unique circuit $C(e, B)$, known as the *fundamental circuit of e with respect to B* . Since every circuit is a fundamental circuit with respect to some basis, we can construct the list of circuits of M , given the list of bases, by creating the list of fundamental circuits with respect to each of the bases, and then eliminating duplications.

If B is a basis of M , and $e \in E(M) - B$, we find $C(e, B)$ by comparing $(B \cup e) - f$ against the list of bases for each $f \in B$. The element f is in $C(e, B)$ if and only if $(B \cup e) - f$ is a basis. $\square$

The next result follows easily using duality and Lemma 3.6.

LEMMA 3.7. BASES $\leq$ HYPERPLANES.

Let the number of bases and cyclic flats in the matroid M be denoted by $b(M)$ and $z(M)$, respectively.

PROPOSITION 3.8. *Let M be a matroid. Then $z(M) \leq b(M)$.*

Proof. The proof will be by induction on $|E(M)|$. If $|E(M)| = 0$, then $b(M) = z(M) = 1$, so the proposition holds.

Let M be a matroid such that $|E(M)| = n > 0$, and assume that the proposition holds for all matroids on ground sets of $n - 1$ elements. We may assume that there exists an element $e \in E(M)$ such that e is neither a loop nor a coloop, for otherwise $b(M) = z(M) = 1$.

Let $b_e(M)$ be the number of bases of M that contain e , and let $b_{\bar{e}}(M)$ be the number of bases of M that avoid e . Then $b_e(M) = b(M/e)$ and $b_{\bar{e}}(M) = b(M \setminus e)$. Similarly, let $z_e(M)$ be the number of cyclic flats of M that contain e , and let $z_{\bar{e}}(M)$ be the number of cyclic flats that do not contain e . It is easy to see that any cyclic flat that does not contain e is also a cyclic flat of $M \setminus e$, so $z_{\bar{e}}(M) \leq z(M \setminus e)$. Moreover, if Z is a cyclic flat of M , and $e \in Z$, then $Z - e$ is a cyclic flat of M/e . Thus $z_e(M) \leq z(M/e)$, and hence

$$z(M) = z_e(M) + z_{\bar{e}}(M) \leq z(M/e) + z(M \setminus e).$$

By the inductive hypothesis, $z(M/e) \leq b(M/e)$ and $z(M \setminus e) \leq b(M \setminus e)$, so

$$z(M) \leq b(M/e) + b(M \setminus e) = b_e(M) + b_{\bar{e}}(M) = b(M). \quad \square$$

LEMMA 3.9. BASES $\leq$ CYCLIC FLATS.

Proof. Our algorithm for generating the list of cyclic flats, given the list of bases, will start by constructing the closures of all circuits. At each repetition of the loop the algorithm will find the closure of the union of every pair of cyclic flats already on the list.

Suppose that M is a matroid on the ground set E . Note that we can check in polynomial time whether a set is independent by comparing it against the family of bases. Hence if $A \subseteq E$, we can use the greedy algorithm to find a basis I of A . The element $e \notin A$ is in $\text{cl}(A)$ if and only if $I \cup e$ is dependent. It follows that we can find $\text{cl}(A)$ in polynomial time.

From Lemma 3.6 we see that we can find the list of circuits in polynomial time, given the list of bases. Hence we can also construct the list of closures of circuits, add the closure of the empty set, and then eliminate duplications from the list in polynomial time. This completes the preprocessing that the algorithm will do before entering the loop.

Suppose that $Z_1, \dots, Z_t$ is the list of cyclic flats that has been constructed after the loop has been repeated i times. In the next repetition of the loop the algorithm

will take each pair $\{Z_j, Z_k\}$ of cyclic flats and find $\text{cl}(Z_j \cup Z_k)$. At the completion of the loop the algorithm will add these new cyclic flats to the list and then eliminate duplications. By Proposition 3.8 the length of the list at the beginning of the loop will never exceed b , the number of bases. Thus in each repetition of the loop the algorithm must find the closure of at most b^2 unions of flats. It is easy to see that after $r(M)$ repetitions of the loop the algorithm will have found every cyclic flat. Finding the rank of these flats can clearly be accomplished in polynomial time, using the greedy algorithm. Therefore the algorithm can construct $(M, \text{CYCLIC FLATS})$ in polynomial time. $\square$

LEMMA 3.10. $\text{FLATS} \leq \text{HYPERPLANES}$.

Proof. The flat F is a hyperplane of M if and only if there is no flat, F' , such that $F \subset F' \subset E(M)$. This clearly leads to a polynomial-time algorithm. $\square$

LEMMA 3.11. $\text{FLATS} \leq \text{CYCLIC FLATS}$.

Proof. A flat F fails to be a cyclic flat if and only if there is an element $e \in F$ such that $F - e$ is also a flat. This clearly leads to a polynomial-time algorithm. $\square$

LEMMA 3.12. $\text{CIRCUITS} \leq \text{NONSPANNING CIRCUITS}$.

Proof. Given $(M, \text{CIRCUITS})$ a Turing machine can check in polynomial time whether a subset $A \subseteq E(M)$ is independent. Thus such a machine could find a basis of M by using the greedy algorithm. Once the rank of M is known the rest follows easily. $\square$

The next lemma follows easily using duality and Lemma 3.12.

LEMMA 3.13. $\text{HYPERPLANES} \leq \text{DEPENDENT HYPERPLANES}$.

To complete the proof of Theorem 3.2 we must show that if INPUT1 and INPUT2 are two types of input and the preceding results do not imply that $\text{INPUT1} \leq \text{INPUT2}$, then $\text{INPUT1} \not\leq \text{INPUT2}$. It is clear that if $\text{INPUT1} \leq \text{INPUT2}$, then there must exist a polynomial p such that

$$|(M, \text{INPUT2})| \leq p(|(M, \text{INPUT1})|)$$

for any matroid M . Thus if there exists a family of matroids $M_1, M_2, M_3, \dots$ such that $\{|(M_i, \text{INPUT2})|\}_{i \geq 1}$ is not polynomially bounded by $\{|(M_i, \text{INPUT1})|\}_{i \geq 1}$, then $\text{INPUT1} \not\leq \text{INPUT2}$.

The following proposition follows immediately from transitivity.

PROPOSITION 3.14. *Suppose that $\text{INPUT1} \not\leq \text{INPUT2}$. If $\text{INPUT1} \leq \text{INPUT3}$, and $\text{INPUT4} \leq \text{INPUT2}$, then $\text{INPUT3} \not\leq \text{INPUT4}$.*

Using Proposition 3.14, it is a relatively simple matter to check that the proof of Theorem 3.2 will be completed by verifying the following cases:

1. $\text{SPANNING SETS} \not\leq \text{FLATS}$.
2. $\text{INDEPENDENT SETS} \not\leq \text{SPANNING SETS}$.
3. $\text{FLATS} \not\leq \text{NONSPANNING CIRCUITS}$.
4. $\text{CIRCUITS} \not\leq \text{DEPENDENT HYPERPLANES}$.
5. $\text{HYPERPLANES} \not\leq \text{CYCLIC FLATS}$.
6. $\text{CIRCUITS} \not\leq \text{CYCLIC FLATS}$.
7. $\text{NONSPANNING CIRCUITS} \not\leq \text{CIRCUITS}$.
8. $\text{DEPENDENT HYPERPLANES} \not\leq \text{HYPERPLANES}$.
9. $\text{CYCLIC FLATS} \not\leq \text{DEPENDENT HYPERPLANES}$.

LEMMA 3.15 (Case 1). $\text{SPANNING SETS} \not\leq \text{FLATS}$.

Proof. For $n \geq 1$ let M_n be isomorphic to $U_{n-1,n}$, the n -element uniform matroid of rank $n-1$. The number of spanning sets of M_n is $n+1$, whereas the number of flats is $2^n - n$. Thus $|(M_n, \text{SPANNING SETS})| = \Theta(n^2)$, while $|(M_n, \text{FLATS})| = \Theta(n2^n)$. $\square$

LEMMA 3.16 (Case 2). $\text{INDEPENDENT SETS} \not\leq \text{SPANNING SETS}$.

Proof. For $n \geq 1$ let M_n be isomorphic to $U_{1,n}$. The number of independent sets in M_n is $n + 1$, while the number of spanning sets is $2^n - 1$. $\square$

We denote the truncation of the matroid M by $T(M)$. If m is a positive integer, then we define $mU_{r,n}$ to be the matroid obtained by replacing each element in $U_{r,n}$ with a parallel class of size m .

LEMMA 3.17 (Case 3). FLATS $\not\leq$ NONSPANNING CIRCUITS.

Proof. For $n \geq 3$ define M_n to be $T(nU_{n-1,n} \oplus U_{2,2})$. Note that M_n contains $n^2 + 2$ elements. There are $n + 2$ parallel classes in M_n . It follows that the number of flats of M_n is at most 2^{n+2} . However, the number of nonspanning circuits of M_n is exactly $n^n + n \binom{n}{2}$. Thus $|(M_n, \text{FLATS})| = O(n^2 2^{n+2})$. On the other hand $|(M_n, \text{NONSPANNING CIRCUITS})| = \Theta(n^{n+2})$, so the result follows. $\square$

LEMMA 3.18 (Case 4). CIRCUITS $\not\leq$ DEPENDENT HYPERPLANES.

Proof. It is not difficult to see that a polynomial-time algorithm that constructs $(M, \text{DEPENDENT HYPERPLANES})$ from $(M, \text{CIRCUITS})$ for any matroid M can be used to show that $\text{HYPERPLANES} \leq \text{NONSPANNING CIRCUITS}$. This contradicts Proposition 3.14, as $\text{FLATS} \leq \text{HYPERPLANES}$ by Lemma 3.10, and $\text{FLATS} \not\leq \text{NONSPANNING CIRCUITS}$ by Lemma 3.17. $\square$

LEMMA 3.19 (Case 5). HYPERPLANES $\not\leq$ CYCLIC FLATS.

Proof. For $n \geq 3$, let M_n be isomorphic to $2U_{n-1,n}$. The hyperplanes of M_n are exactly the sets of $n - 2$ parallel classes, while any nonempty set of parallel classes is a cyclic flat as long as it does not contain exactly $n - 1$ such classes. Thus the number of hyperplanes is $(n^2 - n)/2$, while the number of cyclic flats is $2^n - n - 1$. $\square$

LEMMA 3.20 (Case 6). CIRCUITS $\not\leq$ CYCLIC FLATS.

Proof. Given $(M, \text{HYPERPLANES})$ we can certainly find $(M^*, \text{CIRCUITS})$ in polynomial time. Also, given $(M^*, \text{CYCLIC FLATS})$ we can find the cyclic flats of M in polynomial time, since the cyclic flats of M^* are the complements of the cyclic flats of M . Moreover, it is easy to see that given $(M^*, \text{CIRCUITS})$ we can find the rank of any subset in M in polynomial time.

Suppose that $\text{CIRCUITS} \leq \text{CYCLIC FLATS}$. The above discussion implies that $\text{HYPERPLANES} \leq \text{CYCLIC FLATS}$, in contradiction to Lemma 3.19. $\square$

LEMMA 3.21 (Case 7). NONSPANNING CIRCUITS $\not\leq$ CIRCUITS.

Proof. If M_n is isomorphic to $U_{n,2n}$, then M_n contains no nonspanning circuits, so by definition $|(M_n, \text{NONSPANNING CIRCUITS})| = O(n)$. On the other hand, the number of circuits is $\binom{2n}{n+1}$, which is exponential in n . $\square$

The next lemma follows using duality and Lemma 3.21.

LEMMA 3.22 (Case 8). DEPENDENT HYPERPLANES $\not\leq$ HYPERPLANES.

LEMMA 3.23 (Case 9). CYCLIC FLATS $\not\leq$ DEPENDENT HYPERPLANES.

Proof of Lemma 3.23. For $n \geq 2$ let M_n be the matroid obtained by adding a single parallel element to a member of the ground set of $U_{n,2n}$. The only cyclic flats of M_n are the empty set, the nontrivial parallel class, and the entire ground set. Thus $|(M_n, \text{CYCLIC FLATS})| = \Theta(n)$. However, any hyperplane that contains the nontrivial parallel class is dependent, so the number of such hyperplanes is $\binom{2n-1}{n-2}$. $\square$

With this lemma we have completed the proof of Theorem 3.2. $\square$

4. Matroid intersection. It is easy to see that given (M, INPUT) , it is possible to determine in polynomial time whether or not a subset of $E(M)$ is independent in M . (Henceforth we assume INPUT to be one of the types of input discussed in section 2.) Hausmann and Korte [6] and Robinson and Welsh [12] show that the standard matroid oracles can be efficiently simulated by an independence oracle. It follows

from these observations that if a computational problem can be solved by an oracle Turing machine in time that is bounded by $p(n)$ for any n -element matroid, where p is a fixed polynomial, then the same problem can be solved in polynomial time by a Turing machine which receives (M, INPUT) as its input.

The converse is not true. Consider the problem of deciding whether a matroid is uniform. Robinson and Welsh [12] note that a Turing machine equipped with an oracle cannot solve this problem in time bounded by a polynomial function of the size of the ground set. In contrast, deciding whether M is uniform given (M, INPUT) is trivial.

However, there do exist matroid-theoretical problems which are probably not solvable in polynomial time, even when the input consists of a list of some family of subsets. One of these is 3-MATROID INTERSECTION.

3-MATROID INTERSECTION.

Instance: An integer k and (M_i, INPUT) for $1 \leq i \leq 3$, where M_1 , M_2 , and M_3 are matroids with a common ground set E .

Question: Does there exist a set $A \subseteq E$ such that $|A| = k$ and A is independent in M_1 , M_2 , and M_3 ?

The fact that this problem is NP-complete was first observed by Lawler [9]. He does not specify a form of matroid input, but he remarks that the problem is NP-complete for partition matroids (direct sums of rank-one uniform matroids). It is clear from his comments that a partition matroid is to be described via the partition of its ground set into connected components. We sketch a modified version of his proof here.

THEOREM 4.1. *If $\text{CIRCUITS} \leq \text{INPUT}$ or if $\text{HYPERPLANES} \leq \text{INPUT}$, then 3-MATROID INTERSECTION is NP-complete. However, if $\text{INPUT} \leq \text{BASES}$, then 3-MATROID INTERSECTION is in P.*

Proof. Obviously the problem is in NP. It suffices to prove NP-completeness only in the case that either $\text{INPUT} = \text{CIRCUITS}$ or $\text{INPUT} = \text{HYPERPLANES}$. We provide a reduction from the following NP-complete problem.

3-DIMENSIONAL MATCHING.

Instance: A set of triples, $M \subseteq X_1 \times X_2 \times X_3$, where X_1 , X_2 , and X_3 are pairwise disjoint sets having the same cardinality.

Question: Does M contain a *matching*? (A subset $M' \subseteq M$, such that every element in $X_1 \cup X_2 \cup X_3$ is contained in exactly one triple in M' .)

Let $M \subseteq X_1 \times X_2 \times X_3$ be an instance of 3-DIMENSIONAL MATCHING, and suppose that $X_i = \{x_1^i, \dots, x_s^i\}$ for $1 \leq i \leq 3$. Suppose that M contains t triples, $p_1, \dots, p_t$. We construct three partition matroids, M_1 , M_2 , and M_3 , on the ground set $E = \{e_1, \dots, e_t\}$. Each matroid M_i contains s connected components, corresponding to the elements of X_i . The connected component corresponding to x_j^i is equal to $\{e_k \mid p_k \text{ contains } x_j^i, 1 \leq k \leq t\}$. It is clear that M contains a matching if and only if M_1 , M_2 , and M_3 contain a common independent set of size s .

It remains to show that $(M_i, \text{CIRCUITS})$ and $(M_i, \text{HYPERPLANES})$ can be constructed in polynomial time. Since the number of circuits or hyperplanes in a partition matroid is at most quadratic in the size of the ground set, this is easily done.

If $\text{INPUT} = \text{BASES}$, then we can find a common independent set of maximum size by considering the intersection of every triple of bases from the three matroids. This can clearly be done in polynomial time, so 3-MATROID INTERSECTION is in P if $\text{INPUT} \leq \text{BASES}$. $\square$

Theorem 4.1 shows that 3-MATROID INTERSECTION is either NP-complete or in P for all but two of the methods of input described in section 2: The status of the problem is open for the case that INPUT = CYCLIC FLATS or INPUT = FLATS.

5. The isomorphism problem. The following computational problem has attracted much attention.

GRAPH ISOMORPHISM.

Instance: Two graphs, G and G' .

Question: Are G and G' isomorphic?

GRAPH ISOMORPHISM is thought to be a good candidate for a problem in NP that is neither NP-complete nor in P (see [2]).

A decision problem that is polynomially equivalent to GRAPH ISOMORPHISM is *isomorphism-complete*. In this section we show that the analogous matroid problem is in general isomorphism-complete.

MATROID ISOMORPHISM.

Instance: (M, INPUT) and (M', INPUT) .

Question: Are M and M' isomorphic?

LEMMA 5.1. MATROID ISOMORPHISM is *polynomially reducible* to GRAPH ISOMORPHISM.

Proof. A proof can be found in [10]; we give an outline here. We must construct a polynomial-time computable transformation that takes descriptions of matroids to graphs in such a way that isomorphism is preserved. There are many ways in which this can be accomplished. The key idea is that a list of characteristic vectors, representing subsets of the ground set, can be seen as the rows of the vertex-adjacency matrix of a bipartite graph.

The rest of the demonstration involves refining the transformation so that, given the unlabeled bipartite graph, it is possible to reconstruct the matroid description, up to relabeling. This guarantees that the transformation preserves isomorphism. Thus we must somehow distinguish the vertices that correspond to subsets of the ground set from the vertices that correspond to elements of the ground set. In the case that INPUT relies upon ranks being assigned to sets, as is the case when INPUT = CYCLIC FLATS, we must find a method of encoding binary representations of integers in the form of graphs. Constructing a transformation that satisfies these criteria, and confirming that it is polynomial-time computable, is an easy exercise. $\square$

Next we develop a polynomial transformation from graphs to matroid descriptions. Suppose that G is a simple graph with n vertices, $\{v_1, \dots, v_n\}$, and m edges, $\{e_1, \dots, e_m\}$. Assume that $n \geq 3$. Let $X = \{x_1, \dots, x_n\}$, $X' = \{x'_1, \dots, x'_n\}$, and $Y = \{y_1, \dots, y_m\}$ be disjoint sets. The matroid $\Phi(G)$ has rank 3, and the ground set of $\Phi(G)$ is $X \cup X' \cup Y$. The nonspanning circuits of $\Phi(G)$ are sets of the form $\{x_i, x'_i\}$, where $i \in \{1, \dots, n\}$, and the sets

$$\{\{z_i, z_j, y_k\} \mid z_i \in \{x_i, x'_i\}, z_j \in \{x_j, x'_j\}, \text{ and } e_k \text{ joins } v_i \text{ to } v_j\}.$$

Thus $\Phi(G)$ can be formed by placing n parallel pairs, corresponding to the n vertices of G , in the plane in general position, and placing elements between parallel pairs that correspond to adjacent vertices, in such a way that no additional dependencies are formed.

Let G be a graph. The *cyclomatic number* of G is $|V(G)| - |E(F)|$, where F is a spanning forest of G . The *bicircular matroid* of a graph, G , denoted by $B(G)$, has the edge set of G as its ground set. The circuits of $B(G)$ are exactly the minimal connected edge sets of G with cyclomatic number two, known as *bicycles*. Thus a set

of edges is independent in $B(G)$ if and only if the subgraph of G it induces contains at most one cycle in every connected component.

Suppose that G is simple and has at least three vertices. Let $G^{\circ\circ}$ be the graph which is obtained by adding two loops at each vertex of G . Then $\Phi(G)$ is isomorphic to the matroid obtained by repeatedly truncating $B(G^{\circ\circ})$ so that its rank is reduced to three.

Suppose that G is a simple graph with $n \geq 3$ vertices and m edges. Then $|\Phi(G)| = 2n + m$. Since $m \leq n^2$, the number of independent sets in $\Phi(G)$ is $O(n^6)$. Thus $(\Phi(G), \text{INPUT})$ can be constructed in polynomial time from a description of G when $\text{INDEPENDENT SETS} \leq \text{INPUT}$.

It is easy to demonstrate that if G and G' are simple graphs on at least three vertices, then G and G' are isomorphic if and only if $\Phi(G)$ and $\Phi(G')$ are isomorphic. Since GRAPH ISOMORPHISM is generally defined in terms of simple graphs, we have proved the following result.

LEMMA 5.2. *If $\text{INDEPENDENT SETS} \leq \text{INPUT}$, then GRAPH ISOMORPHISM is polynomially reducible to $\text{MATROID ISOMORPHISM}$.*

Using duality we see that $(\Phi(G)^*, \text{SPANNING SETS})$ can also be constructed in polynomial time from a description of G , where G is any simple graph with at least three vertices. Since G and G' are isomorphic graphs if and only if $(\Phi(G))^*$ and $(\Phi(G'))^*$ are isomorphic matroids, we have proved the following.

THEOREM 5.3. *If $\text{SPANNING SETS} \leq \text{INPUT}$, or $\text{INDEPENDENT SETS} \leq \text{INPUT}$, then $\text{MATROID ISOMORPHISM}$ is isomorphism-complete.*

6. Detecting minors. It is a well-known observation of Seymour's that an oracle Turing machine cannot decide whether a matroid has a $U_{2,4}$ -minor in time that is polynomially bounded by the size of the ground set [13]. We will show that, in general, deciding whether a matroid contains a minor isomorphic to some fixed matroid can be done in polynomial time, given a list of the independent sets or some similar input. It is routine to verify the following result.

PROPOSITION 6.1. *Suppose that M is a matroid and that X and Y are disjoint subsets of $E(M)$. It is possible to construct $(M/X \setminus Y, \text{INPUT})$ in polynomial time given (M, INPUT) .*

PROPOSITION 6.2. *Let N be a matroid and let t be the number of circuits of N , where $t > 0$. If M is a matroid on the ground set E , and M has a minor isomorphic to N on the set $A \subseteq E$, then there exist circuits, $C_1, \dots, C_t$, of M such that, if $X = (C_1 \cup \dots \cup C_t) - A$ and $Y = E - (A \cup X)$, then $M/X \setminus Y \cong N$.*

Proof. Suppose that $M/X' \setminus Y'$ is isomorphic to N , where (X', Y') is a partition of $E - A$. We may assume that X' is independent. There are exactly t circuits, $C'_1, \dots, C'_t$, in $M/X' \setminus Y'$. For each circuit, C'_i , there exists a circuit C_i of M such that $C_i \subseteq C'_i \cup X'$ and $C'_i = C_i - X'$. Let X be the set $(C_1 \cup \dots \cup C_t) - A$ and let Y be $E - (A \cup X)$. It remains to show that $M/X \setminus Y \cong N$. Note that $X \subseteq X'$. If $X = X'$, then we are done, so suppose that x is an element in $X' - X$. Let $M' = M/(X' - x) \setminus (Y' \cup x)$. To complete the proof it will suffice to show that $M' = M/X' \setminus Y'$. This is elementary. $\square$

Let N be some fixed matroid. It is a consequence of Propositions 6.1 and 6.2 that the following problem is, in general, in P.

DETECTING AN N -MINOR.

Instance: (M, INPUT) .

Question: Does M have a minor isomorphic to N ?

THEOREM 6.3. *If $\text{INPUT} \leq \text{CIRCUITS}$ or $\text{INPUT} \leq \text{HYPERPLANES}$, then $\text{DETECTING AN } N\text{-MINOR}$ is in P for any fixed matroid N .*

Proof. By duality it will suffice to prove the theorem only when INPUT is equal to CIRCUITS. We may assume that we have the list of circuits of N , for we can construct it in constant time.

Suppose that the ground set of M is E , where $|E| = n$, and that M contains c circuits, so that $|(M, \text{CIRCUITS})| = \Theta(nc)$. Suppose also that $|E(N)| = s$, and that N has exactly t circuits. We may assume that $t > 0$, for otherwise the problem is trivially in P.

An algorithm to check whether M has an N -minor could simply work its way through every subset A of E such that $|A| = s$, and every set $\{C_1, \dots, C_t\}$ of t circuits of M . By Proposition 6.1 it is possible to construct $(M/X \setminus Y, \text{CIRCUITS})$, where $X = (C_1 \cup \dots \cup C_t) - A$ and $Y = E - (A \cup X)$, and then check whether $M/X \setminus Y \cong N$ in polynomial time. Proposition 6.2 guarantees that if M does have a minor isomorphic to N , then this procedure will find such a minor. Checking whether an isomorphism exists between N and $M/X \setminus Y$ will take some constant time, so the running time of the algorithm is determined by the number of subsets, $A \subseteq E$, and the number of families of circuits we must check. The first quantity is $\binom{n}{s}$, and the second is $\binom{c}{t}$, so the running time of the algorithm is $O(n^s c^t)$. Since s and t are fixed constants, the algorithm runs in polynomial time. $\square$

In contrast to Theorem 6.3, Hliněný shows that the problem of deciding whether M has a minor isomorphic N for a fixed matroid N is NP-complete when the input consists of a representation of M over the rational numbers [7]. On the other hand, Geelen, Gerards, and Whittle conjecture that the problem is in P when the input consists of a representation of M over a finite field [3].

The exponent in the running time of the algorithm described in Theorem 6.3 depends upon N . It is natural to ask whether there is a *fixed-parameter tractable* (FPT) algorithm for DETECTING AN N -MINOR, that is, an algorithm which runs in time $f(N)|(M, \text{INPUT})|^k$, where f is a function which depends only on N and k is a fixed constant. Certainly such an algorithm exists when INPUT = RANK, for we can simply consider minors of the form $M/X \setminus Y$ for all disjoint sets X and Y , and check by brute force whether any such minor is isomorphic to N . This leads to an FPT algorithm. The existence of such an algorithm when M is described via an input that lies above RANK is an open problem.

We have shown that detecting a fixed minor can be solved in polynomial time, in general. However, the problem of detecting a minor which forms part of the input is, in general, NP-complete.

MINOR ISOMORPHISM.

Instance: (M, INPUT) and (N, INPUT) .

Question: Does M have a minor isomorphic to N ?

THEOREM 6.4. *If INDEPENDENT SETS $\leq$ INPUT or SPANNING SETS $\leq$ INPUT, then MINOR ISOMORPHISM is NP-complete.*

Proof. It is easy to see that MINOR ISOMORPHISM is in NP. The following problem is well known to be NP-complete.

SUBGRAPH ISOMORPHISM.

Instance: Two graphs, G and H .

Question: Does G have a subgraph isomorphic to H ?

Suppose that the graphs G and H correspond to an instance of SUBGRAPH ISOMORPHISM. We may assume that G and H are simple and that both have at least three vertices. Since we can construct either $(\Phi(G), \text{INPUT})$ or $(\Phi(G)^*, \text{INPUT})$ in polynomial time from the description of G (and the same statement applies for H), the proof will be complete if we can demonstrate that G contains a subgraph

isomorphic to H if and only if $\Phi(G)$ contains a minor isomorphic to $\Phi(H)$. This is easily done. $\square$

Suppose that $\mathcal{M}$ is a family of matroids. A natural computational problem is to ask whether a matroid M has a minor isomorphic to a member of $\mathcal{M}$ with specified size. The proof of Theorem 6.4 can be used to show that if $\mathcal{M} = \{\Phi(K_n) \mid n \geq 1\}$, then this problem is, in general, NP-complete. We will conclude by showing that the problem is, in general, also NP-complete when $\mathcal{M} = \{U_{r,n} \mid n \geq r\}$, where r is a fixed constant.

$U_{r,n}$ -MINOR.

Instance: (M, INPUT) and an integer n .

Question: Does M have a minor isomorphic to $U_{r,n}$?

THEOREM 6.5. *If $r > 2$ is a fixed integer and $\text{INDEPENDENT SETS} \leq \text{INPUT}$, then $U_{r,n}$ -MINOR is NP-complete.*

Proof. It is easy to see that $U_{r,n}$ -MINOR is in NP. We construct a reduction from the following NP-complete problem.

INDEPENDENT SET.

Instance: An integer k and a graph G .

Question: Does G contain an independent set of k vertices?

Let $r > 2$ be a fixed integer. Suppose that the integer k and the simple graph G are an instance of INDEPENDENT SET. Let m be the number of edges in G . Let t be $\lceil (r-1)/2 \rceil$, and let tG° be the graph obtained by adding a loop to each vertex of G and then replacing each nonloop edge with a path of length t . The matroid $\Phi_r(G)$ is the bicircular matroid of tG° , repeatedly truncated so that its rank is equal to r .

Checking whether a set of edges of tG° is independent in $B(tG^\circ)$ can certainly be done in polynomial time. Since r is a fixed integer, and no independent set of $\Phi_r(G)$ exceeds r in size, it follows that $(\Phi_r(G), \text{INPUT})$ can be constructed in polynomial time as long as $\text{INDEPENDENT SETS} \leq \text{INPUT}$.

We complete the proof by showing that G has an independent set of k vertices if and only if $\Phi_r(G)$ has a rank- r uniform minor of size $k + mt$.

Every nonloop cycle of tG° contains at least $3t$ edges, so if a bicycle of tG° contains at most one loop, then it has size at least $3t+1$. This quantity is greater than r , as $r \geq 3$. If a bicycle contains two loops, then either it contains exactly $t+2$ elements, or its size is at least $2t+2$. It is straightforward to confirm that $t+2 \leq r$ and that $2t+2 > r$. This shows that the nonspanning circuits of $\Phi_r(G)$ are exactly the sets containing the t edges in a path joining two vertices of G and the two loops incident with those vertices.

Suppose that G contains an independent set of k vertices. These vertices correspond to k loops of tG° . The set containing these loops and all the nonloop edges of tG° cannot contain a nonspanning circuit of $\Phi_r(G)$, by the above discussion. Thus restricting $\Phi_r(G)$ to this set of $k + mt$ elements gives a uniform minor. We may assume that $m \geq 3$, so $k + mt$ is certainly no less than r . Therefore $\Phi_r(G)$ contains a rank- r uniform minor with $k + mt$ elements.

For the converse suppose that $\Phi_r(G)$ has a rank- r uniform minor on the set A , where $|A| = k + mt$, and assume that G has no independent set of k vertices. The number of nonloop edges in A is at most mt . Suppose that A has been chosen so that it contains as many nonloop edges as possible. Now A contains at least k loops, so there must be a pair of loops, l and l' , in A that correspond to adjacent vertices in G . Therefore one of the t edges that join l to l' in tG° does not belong to A . Let us call this edge e . Then $(A - l) \cup e$ contains no nonspanning circuits of $\Phi_r(G)$, and our assumption on A is contradicted. Therefore G has an independent set of k vertices. $\square$

If $r \leq 2$, then $U_{r,n}$ -MINOR is trivially in P. Using duality, we can show that the problem of deciding whether M has a minor isomorphic to $U_{n-r,n}$ is NP-complete for fixed values of $r > 2$ as long as SPANNING SETS $\leq$ INPUT.

7. Open problems. In this summary section we collect some open problems. 3-MATROID INTERSECTION is known to be either in P or NP-complete for all but two of the 10 types of input mentioned in section 2. The status of the problem is open when INPUT is either CYCLIC FLATS or FLATS. MATROID ISOMORPHISM is NP-complete for all forms of input, except possibly RANK. Deciding if MATROID ISOMORPHISM can be solved in polynomial time when the matroids are described via the rank of each of their subsets is an open problem. More generally, it would be interesting to know if there is any “natural” computational problem which is NP-complete for the RANK input.

DETECTING AN N -MINOR is known to be in P for all forms of input except NONSPANNING CIRCUITS, CYCLIC FLATS, and DEPENDENT HYPERPLANES. The status of the problem for these types of input is unknown. The existence or otherwise of an FPT algorithm for MINOR ISOMORPHISM is known only when INPUT = RANK. Otherwise the problem is open.

Acknowledgments. I thank my supervisor, Professor Dominic Welsh, and Professors Colin McDiarmid and James Oxley for very useful advice and discussions. I also thank the referees for their constructive comments. In particular, the question of the existence of FPT algorithms for DETECTING AN N -MINOR was raised by one of the referees.

REFERENCES

- [1] J. EDMONDS, *Submodular functions, matroids, and certain polyhedra*, in Combinatorial Structures and Their Applications (Proc. Calgary International Conference, Calgary, Alberta, 1969), Gordon and Breach, New York, 1970, pp. 69–87.
- [2] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman and Co., San Francisco, CA, 1979.
- [3] J. GEELEN, B. GERARDS, AND G. WHITTLE, *Towards a structure theory for matrices and matroids*, in International Congress of Mathematicians, Vol. III, Eur. Math. Soc., Zürich, 2006, pp. 827–842.
- [4] D. HAUSMANN AND B. KORTE, *Oracle algorithms for fixed-point problems—an axiomatic approach*, in Optimization and Operations Research (Proc. Workshop, University of Bonn, Bonn, Germany, 1977), Lecture Notes in Econom. and Math. Systems 157, Springer, Berlin, 1978, pp. 137–156.
- [5] D. HAUSMANN AND B. KORTE, *Worst-case behaviour of polynomial bounded algorithms for independence systems*, Z. Angew. Math. Mech., 58 (1978), pp. T477–T479.
- [6] D. HAUSMANN AND B. KORTE, *Algorithmic versus axiomatic definitions of matroids*, Math. Programming Stud., (1981), pp. 98–111.
- [7] P. HLINĚNÝ, *On matroid representability and minor problems*, in Mathematical Foundations of Computer Science 2006, Lecture Notes in Comput. Sci. 4162, Springer, Berlin, 2006, pp. 505–516.
- [8] D. E. KNUTH, *The asymptotic number of geometries*, J. Combin. Theory Ser. A, 16 (1974), pp. 398–400.
- [9] E. L. LAWLER, *Polynomial-bounded and (apparently) nonpolynomial-bounded matroid computations*, in Combinatorial Algorithms (Courant Computer Science Symposia 9, New York University, New York, 1972), Algorithmics Press, New York, 1973, pp. 49–57.
- [10] D. MAYHEW, *Matroids and Complexity*, Ph.D. thesis, University of Oxford, Oxford, UK, 2005.
- [11] J. G. OXLEY, *Matroid Theory*, The Clarendon Press, Oxford University Press, New York, 1992.
- [12] G. C. ROBINSON AND D. J. A. WELSH, *The computational complexity of matroid properties*, Math. Proc. Cambridge Philos. Soc., 87 (1980), pp. 29–45.
- [13] P. D. SEYMOUR, *Recognizing graphic matroids*, Combinatorica, 1 (1981), pp. 75–78.

ON RAMSEY MINIMAL GRAPHS*

V. RÖDL[†] AND M. SIGGERS[‡]

Abstract. A graph G is r -Ramsey-minimal with respect to a graph H if every r -coloring of the edges of G yields a monochromatic copy of H , but the same is not true for any proper subgraph of G . In this paper we show that for any integer $k \geq 3$ and $r \geq 2$, there exists a constant $c > 1$ such that for large enough n , there exist at least c^{n^2} nonisomorphic graphs on at most n vertices, each of which is r -Ramsey-minimal with respect to the complete graph K_k . Furthermore, in the case $r = 2$, we give an asymmetric version of the above result.

Key words. Ramsey-minimal, Ramsey infinite, sender

AMS subject classifications. 05C35, 05C55

DOI. 10.1137/050647116

1. Introduction. A classical problem of graph theory is determining the Ramsey number for a graph H . This is the minimum number n such that any 2-coloring of the edges of the complete graph K_n on n vertices yields a monochromatic copy of H .

A variation on this problem is to look at coloring the edges of some graph G instead of K_n . Again we are interested in the minimal graphs that yield a monochromatic copy of H when 2-colored.

A graph G is *Ramsey with respect to H* if any 2-coloring of its edges yields a monochromatic copy of H . It is *Ramsey-minimal with respect to H* if it is Ramsey but none of its proper subgraphs are. For results about Ramsey-minimal graphs, see, for example, [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [12], [13], [15], [16], [18], and [19].

In [7], Burr, Erdős, and Lovász showed that for any integer $k \geq 3$ there are infinitely many nonisomorphic graphs that are Ramsey-minimal with respect to the complete graph K_k on k vertices. In [10], this result was strengthened to show that for any integer $k \geq 3$ there exists a constant $c > 1$ such that for all n large enough, there exist at least $c^{n \log n}$ nonisomorphic graphs on at most n vertices that are Ramsey-minimal with respect to K_k .

A graph G is *r -Ramsey with respect to H* if any r -coloring of its edges yields a monochromatic copy of H . It is *r -Ramsey-minimal with respect to H* if it is r -Ramsey but none of its proper subgraphs are.

In this paper we strengthen the result of [10] stated above and extend it to an arbitrary number of colors. Our main result is as follows.

THEOREM 1.1. *For any integers $r \geq 2$ and $k \geq 3$ there exist constants $c = c(r, k) > 1$ and $n_0(r, k)$ such that for all $n \geq n_0(r, k)$, there exist at least c^{n^2} nonisomorphic graphs, on at most n vertices, that are r -Ramsey-minimal with respect to K_k .*

A common generalization for Ramsey-minimal graphs is to the following nonsymmetric definition. A graph G is *Ramsey with respect to H_1 and H_2* if any 2-coloring

*Received by the editors December 8, 2005; accepted for publication (in revised form) September 25, 2007; published electronically March 20, 2008.

<http://www.siam.org/journals/sidma/22-2/64711.html>

[†]Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322 (rod@mathcs.emory.edu). This author was partially supported by NSF grant DMS 0300529.

[‡]School of Computing Sciences, Simon Fraser University, Burnaby V5A 1S6, BC, Canada (mhsiggers@gmail.com). Most of this author's work was done while at Emory University.

yields a monochromatic copy of H_1 in the first color or a monochromatic copy of H_2 in the second color.

At the end of the paper, in section 6, we prove a nonsymmetric version of the result for the case $r = 2$. That is, we observe that with similar quantifications as in Theorem 1.1 there are many graphs that are Ramsey-minimal with respect to K_ℓ and K_k .

Indeed, the results of [7] and [10] above were also proved in this more general setting.

The simplest case of Theorem 1.1 is $k = 3$ and $r = 2$. We prove this case in section 3. In section 4 we prove the theorem for $r = 2$ and all k . Finally, in section 5 we complete the proof. We begin by introducing a preliminary result, the proof of which we put off to section 7.

2. Preliminaries. Throughout the paper, we will identify a graph or hypergraph G with its edgeset $E(G)$. Given an edge e of G we will often write $G - e$ instead of $G - \{e\}$ to mean the (hyper)graph G with the edge e removed. We let $[r]$ denote the set $\{1, \dots, r\}$.

We now introduce signal senders—these are graphs that are very useful in the construction of minimal graphs. They were introduced in [7] and generalized in [10]. We will need them only in the case that $H = K_k$, but since it is not much harder, we prove their existence for a more general graph.

DEFINITION 2.1. *For a graph H , a negative signal sender $S = S^-(r, H, e, f)$ is a graph containing distinguished edges e and f with the following properties:*

- (i) *S is not r -Ramsey with respect to H .*
- (ii) *For every r -coloring of S in which there is no monochromatic copy of H , edges e and f have different colors.*

A positive signal sender $S = S^+(r, H, e, f)$ is the same but we replace the word “different” in property (ii) with “the same.”

In [7] it was proved that there exist positive and negative signal senders for complete graphs when $r = 2$. In [10] a similar proof was used to extend the result to any 3-connected graph H when $r = 2$. We extend the result of [10] to all $r \geq 2$ with the following result.

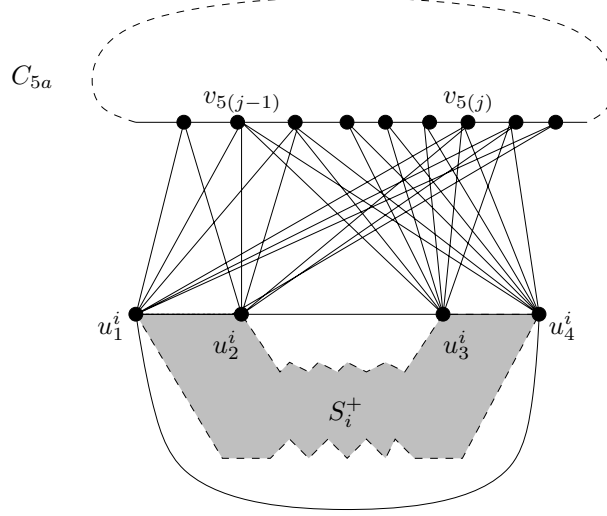
LEMMA 2.2. *For $r \geq 2$ and any 3-connected graph H there exist*

- (a) *negative signal senders $S^-(r, H, e, f)$,*
- (b) *positive signal senders $S^+(r, H, e, f)$, and*
- (c) *such negative and positive signal sender in which the distance between the edges e and f is arbitrarily large.*

The proof of this lemma is similar to the proof in [10]. We include it in section 7 for completeness. Lemma 2.2, and the constructions in section 3.1, are the main tools that we will use in the proof of Theorem 1.1.

3. Theorem 1.1: The base case. In this section we show that there exist $n_0(2, 3)$ and $c = c(2, 3) > 1$ such that for any $n \geq n_0(2, 3)$ there exist at least c^{n^2} nonisomorphic graphs on at most n vertices that are 2-Ramsey-minimal with respect to the triangle, K_3 . This will be the base case for the inductions we use in the proof of Theorem 1.1.

Let a be an odd integer. For $i = 1, \dots, a$, let $L_i \cup R_i = [a]$ be a partition. Let $\mathcal{P} = (L_i \cup R_i \mid i = 1, \dots, a)$. For each of the 2^{a^2} such choices of $\mathcal{P}$ we will construct a labeled graph $G(\mathcal{P})$ that is 2-Ramsey-minimal with respect to the triangle. We will use this construction in Theorem 3.12 to prove our main result for the case $r = 2$ and $k = 3$.

FIG. 3.1. Edges of E_i and F_i when $j-1, j+1 \in L_i$, and $j \in R_i$.

The graph $G = G(\mathcal{P})$ will be built from three components: $G^*(\mathcal{P})$, $G^\uparrow$, and $G^\downarrow$. $G^*(\mathcal{P})$ is built in section 3.1, and $G^\uparrow$ and $G^\downarrow$ are built in section 3.2. In section 3.3 these graphs are used to construct $G(\mathcal{P})$.

3.1. The graph $G^*(\mathcal{P})$. For odd a and set $\mathcal{P} = \{L_i \cup R_i \mid i = 1, \dots, a\}$ of partitions of $[a]$, we construct the labeled graph $G^*(\mathcal{P})$, as follows. (Recall that we will often identify a graph G with its edgeset $E(G)$.)

CONSTRUCTION 3.1.

(i) Let C_{5a} be a cycle of length $5a$ with vertices $v_1, \dots, v_{5a}$ and edges $\{v_j, v_{j+1}\}$ for all $j = 1, \dots, 5a$ (indices mod $5a$).

(ii) For $i = 1, \dots, a$, let S_i^+ be a copy of the positive signal sender $S^+ = S^+(2, K_3, e, f)$ with distance 3 between edges e and f (cf. Lemma 2.2). Let $\{u_1^i, u_2^i\}$ and $\{u_3^i, u_4^i\}$ be the copies of e and f , respectively, in S_i^+ . Let $U^i = \{u_1^i, u_2^i, u_3^i, u_4^i\}$.

(iii) For $i = 1, \dots, a$, add the edges $\{u_1^i, u_4^i\}$ and $\{u_2^i, u_3^i\}$. Note that this does not introduce any new triangles, because we made the assumption that there is distance 3 between e and f in S^+ . (See Figure 3.1.)

(iv) For vertex sets X and Y , let $K(X, Y)$ denote the complete bipartite graph between X and Y . For each partition $L_i \cup R_i \in \mathcal{P}$, add the following unions of complete bipartite graphs:

$$E_i = \bigcup_{j \in L_i} K(\{u_1^i, u_2^i\}, \{v_{5(j-1)}, \dots, v_{5j+1}\}), \text{ and}$$

$$F_i = \bigcup_{j \in R_i} K(\{u_3^i, u_4^i\}, \{v_{5(j-1)}, \dots, v_{5j+1}\}).$$

(See Figure 3.1.) Let $E = \bigcup_{i=1}^a E_i$ and $F = \bigcup_{i=1}^a F_i$. Observe that for $j \in [a]$ the set $\{v_{5(j-1)}, \dots, v_{5j+1}\}$ contains seven vertices, not just five. Thus the sets $\{v_{5(j-1)}, \dots, v_{5j+1}\}$ and $\{v_{5(j)}, \dots, v_{5(j+1)+1}\}$ overlap in two vertices.

Denote the resulting graph by $G^*(\mathcal{P})$. Thus

$$(3.1) \quad V(G^*(\mathcal{P})) = \{v_1, \dots, v_{5a}\} \cup \bigcup_{i=1}^a V(S_i^+), \text{ and}$$

$$G^*(\mathcal{P}) = C_{5a} \cup E \cup F \cup \bigcup_{i=1}^a S_i^+ \cup \bigcup_{i=1}^a \{\{u_1^i, u_4^i\}, \{u_2^i, u_3^i\}\}.$$

DEFINITION 3.2. We classify the vertices of C_{5a} in $G^*(\mathcal{P})$ with regard to their adjacencies to the graphs S_i^+ .

If one of the indices $j-1$ and j are in L_i , while the other one is in R_i , then the vertices v_{5j} and v_{5j+1} are adjacent to all four vertices of $U^i = \{u_1^i, u_2^i, u_3^i, u_4^i\}$. In this case, we call the vertices v_{5j} and v_{5j+1} external for U^i .

Other vertices of C_{5a} are adjacent to exactly two of the vertices of U^i . We call these vertices internal for U^i .

An edge from $v_j \in C_{5a}$ to U^i is an internal edge if the vertex v_j is internal for U^i . An internal edge of the form $e = \{u_1^i, v_{5j+3}\}$ for $i \in \{1, \dots, a\}$ and $j \in \{0, \dots, a-1\}$ is called a central edge.

LEMMA 3.3. The graph $G^*(\mathcal{P})$ has the following properties:

- (i) $|V(G^*)| = a(5 + |V(S^+)|) < 2a|V(S^+)|$.
- (ii) If $\chi : G^* \rightarrow \{1, 2\}$ is a 2-coloring of the edges of G^* such that χ is monochromatic on C_{5a} and $\chi(\{u_1^i, u_2^i\}) \neq \chi(C_{5a})$ for some $i \in \{1, \dots, a\}$, then χ has a monochromatic triangle.
- (iii) For any central edge e , $G^* - e$ has a triangle-free 2-coloring χ of its edges such that

- (a) χ is monochromatic on C_{5a} , and
- (b) $\chi(\{u_1^i, u_2^i\}) = \chi(C_{5a})$ for all edges $\{u_1^1, u_2^1\}, \dots, \{u_1^a, u_2^a\}$ except one.

Proof. Property (i) follows from (3.1) and the fact that $|V(S^+)| > 5$. The proofs of the other two properties will be considerably longer.

The proof of property (ii) relies on the observation that the only triangle-free 2-colorings χ of the complete graph K_4 on vertices a_1, a_2, b_1 , and b_2 with $\chi(\{a_1, a_2\}) \neq \chi(\{b_1, b_2\})$ have

$$(*) \quad \chi(\{a_1, b_1\}) = \chi(\{a_2, b_2\}) \neq \chi(\{a_1, b_2\}) = \chi(\{a_2, b_1\}).$$

Let χ be a triangle-free 2-coloring of G^* such that χ is monochromatic on C_{5a} , and $\chi(\{u_1^i, u_2^i\}) \neq \chi(C_{5a})$ for some i . We will assume, without loss of generality, that this i is 1, i.e., that

$$\chi(\{u_1^1, u_2^1\}) \neq \chi(C_{5a}).$$

By the properties of S_1^+ , we have that $\chi(\{u_1^1, u_2^1\}) = \chi(\{u_3^1, u_4^1\})$.

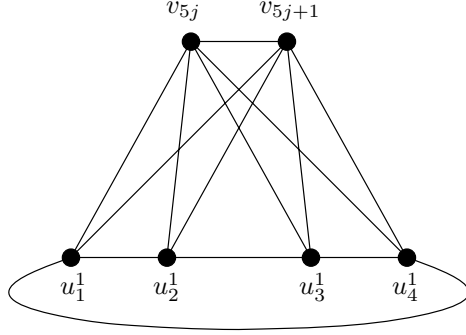
The following observations are specific cases of (*):

- (*₁) Edges from any one of the vertices of U^1 to two consecutive vertices v_i and v_{i+1} must get different colors.
- (*₂) Edges from v_i to either u_1^1 and u_2^1 , or u_3^1 and u_4^1 , must get different colors.

Indeed, if, say, u_1^1 were adjacent to both v_i and v_{i+1} , then these vertices and u_2^1 would induce a K_4 , and so observation (*) yields (*₁). Similarly, if v_i is adjacent to, say, u_1^1 and u_2^1 , then so is one of v_{i-1} and v_{i+1} . These vertices then induce a K_4 , and consequently, observation (*) yields (*₂).

The proof of property (ii) of the lemma follows from the following claim.

CLAIM 3.4. If a vertex v_i of C_{5a} is external for U^1 , then $\chi(\{u_1^1, v_i\}) = \chi(\{u_3^1, v_i\})$.

FIG. 3.2. Subgraph of G^* induced by $U^1 \cup \{v_{5j}, v_{5j+1}\}$.

Indeed, assuming this claim, we are done because every v_i in C_{5a} is adjacent to either u_1^1 or u_3^1 , and so by observation $(*_1)$,

$$\chi(\{-, v_1\}) \neq \chi(\{-, v_2\}) \neq \cdots \neq \chi(\{-, v_{5a}\}) \neq \chi(\{-, v_1\}),$$

where “ $-$ ” denotes either u_1^1 or u_3^1 . However, $5a$ is odd, and χ takes on only two values, so this is a contradiction.

To finish the proof of property (ii), we now prove Claim 3.4.

Proof. If the vertex v_i is external for U^1 , then $i = 5j$ or $5j + 1$, where $j \in L_1$ and $j - 1 \in R_1$ (or where $j \in R_1$ and $j - 1 \in L_1$, but the argument for this case is analogous.) In this situation, all possible edges between U^1 and $\{v_{5j}, v_{5j+1}\}$ are in G^* . (See Figure 3.2.) We show that $\chi(\{u_1^1, v_{5j}\}) = \chi(\{u_3^1, v_{5j}\})$. The argument that $\chi(\{u_1^1, v_{5j+1}\}) = \chi(\{u_3^1, v_{5j+1}\})$ is analogous.

Assume without loss of generality, that $\chi(\{u_1^1, v_{5j}\}) = 1$. By $(*_2)$, $\chi(\{u_2^1, v_{5j}\}) = 2$, and then by $(*_1)$, $\chi(\{u_2^1, v_{5j+1}\}) = 1$. We now have two cases as follows:

- If $\chi(\{u_2^1, u_3^1\}) = 2$, then since χ is triangle-free, we have that $\chi(\{u_3^1, v_{5j}\}) = 1$. This is the same as $\chi(\{u_1^1, v_{5j}\})$, as needed.
- If $\chi(\{u_2^1, u_3^1\}) = 1$, then since χ is triangle-free, $\chi(\{u_3^1, v_{5j+1}\}) = 2$. Thus by $(*_1)$, $\chi(\{u_3^1, v_{5j}\}) = 1$. This is the same as $\chi(\{u_1^1, v_{5j}\})$, as needed. $\square$

This completes the proof of the claim and so completes the proof of (ii).

We now prove property (iii) of the lemma. Without loss of generality, we can assume that the central edge e is $\{u_1^1, v_3\}$. Thus $1 \in L_1$, so G^* contains all edges between $\{u_1^1, u_2^1\}$ and $\{v_2, v_3, v_4\}$, and none of the edges between $\{u_3^1, u_4^1\}$ and $\{v_2, v_3, v_4\}$.

We now define a 2-coloring χ of $G^* - e$.

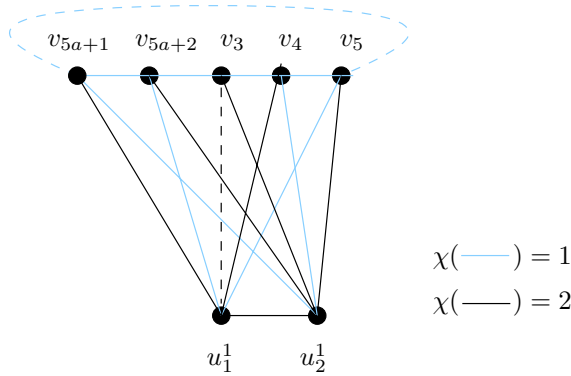
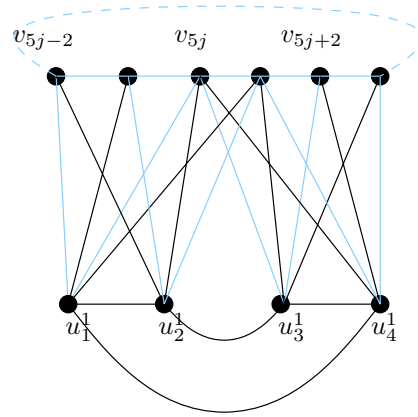
- (a) Let $\chi(f) = 1$ for all edges $f \in C_{5a}$.
 - (b) Let $\chi(\{u_1^1, u_2^1\}) = \chi(\{u_2^1, u_3^1\}) = \chi(\{u_3^1, u_4^1\}) = \chi(\{u_4^1, u_1^1\}) = 2$, and for all $i = 2, \dots, a$, let $\chi(\{u_1^i, u_2^i\}) = \chi(\{u_2^i, u_3^i\}) = \chi(\{u_3^i, u_4^i\}) = \chi(\{u_4^i, u_1^i\}) = 1$.
- Conditions (a) and (b) required by the lemma are thus satisfied. We now have to extend χ to a triangle-free 2-coloring of G^* as follows:

• For all $i \in [a]$ we have $\chi(\{u_1^i, u_2^i\}) = \chi(\{u_3^i, u_4^i\})$, so by the properties of S_i^+ we can (and do) extend χ to a triangle-free 2-coloring of S_i^+ .

• For $i = 2, \dots, a$ let χ have color 2 on all edges between U^i and $V(C_{5a})$.

Observe that so far we have no monochromatic triangles. It remains to extend χ to the edges between $V(C_{5a})$ and U^1 . This is explained below and illustrated in Figures 3.3 and 3.4.

• Relabel v_1 as v_{5a+1} , and v_2 as v_{5a+2} . (This gives v_1 an even index and v_2 an odd index.) With this relabeling in mind, for any edge of the form $f = \{v_i, u_j^1\}$, let $\chi(f) = 1$ if $i + j$ is even, or $\chi(f) = 2$ if $i + j$ is odd.

FIG. 3.3. The coloring χ .FIG. 3.4. The coloring χ when j is odd, $j-1 \in L_1$, and $j \in R_1$.

We now verify that χ is triangle-free on the subgraph of $G^* - \{u_1^1, v_3\}$, that is, induced by $V(C_{5a}) \cup U^1$. This will complete the proof. There are two types of triangles that we must consider as follows:

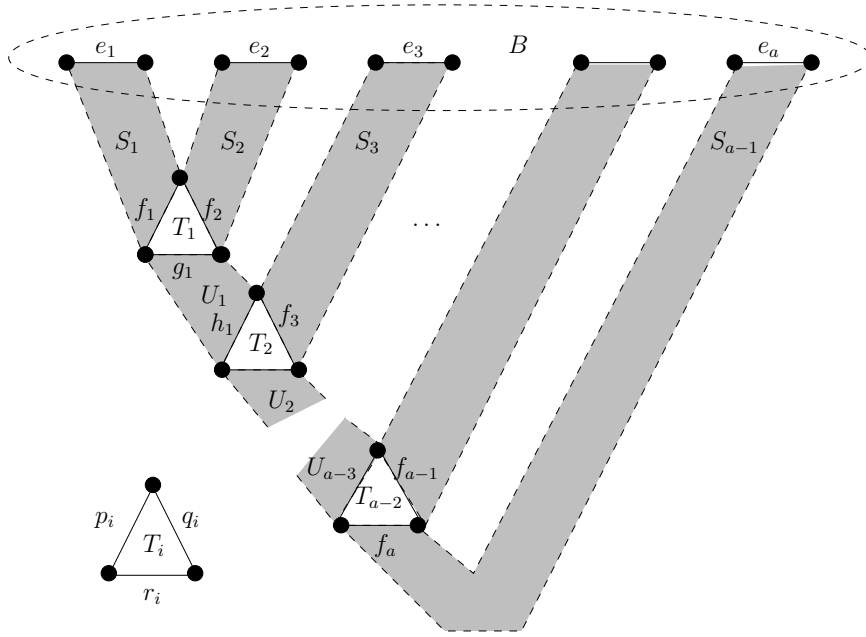
- We first consider any triangle with an edge induced by U^1 . Such a triangle must contain a vertex $v_i \in V(C_{5a}) - \{v_3\}$. However, the colors of the edges from v_i to any two adjacent vertices of U^1 are different. So such a triangle is not monochromatic.
- We then consider any triangle with an edge induced by $V(C_{5a}) - \{v_3\}$. Such a triangle must also contain a vertex $u_i^1 \in U^1$. However, the colors of the edges from u_i^1 to any two adjacent vertices of $V(C_{5a}) - \{v_3\}$ are different. So such a triangle is not monochromatic.

We have thus shown that there is a triangle-free 2-coloring χ of $G^* - \{u_1^1, v_3\}$, that satisfies properties (a) and (b). This completes the proof of the lemma. $\square$

3.2. The graphs $G^\uparrow$ and $G^\downarrow$. For positive integer $a \geq 3$ we construct $G(5a)^\uparrow$ as follows.

CONSTRUCTION 3.5.

(i) Let C_{5a} be a cycle of length $5a$ with vertices $v_1, \dots, v_{5a}$ and edges $\{v_i, v_{i+1}\}$ for all $i = 1, \dots, 5a$.

FIG. 3.5. $G^\downarrow(a)$.

- (ii) For $i = 1, \dots, 5a$, let $S_i = S^-(2, K_3, e_i, f_i)$ be a copy of the negative signal sender $S^- = S^-(2, K_3, e, f)$ with distance 2 between edges e and f .
- (iii) For $i = 1, \dots, 5a$ identify edge $\{v_i, v_{i+1}\} \in C_{5a}$ (indices mod $5a$) with edge $e_i \in S_i$.
- (iv) Identify the $5a$ edges $f_i \in S_i$ for all $i = 1, \dots, 5a$. Denote the resulting edge by f .

Unless explicitly identified, all vertices in the construction are distinct.

LEMMA 3.6. For integer $a \geq 3$, the graph $G^\uparrow = G^\uparrow(5a)$ has the following properties:

- (i) $|V(G^\uparrow)| < 5a|V(S^-)|$.
- (ii) C_{5a} is monochromatic under any triangle-free 2-coloring of $G^\uparrow$, and such a coloring exists.

Proof. The property that $|V(G^\uparrow)| < 5a|V(S^-)|$ follows from the fact that every vertex of $G^\uparrow$ is in at least one of the $5a$ copies of S^- . We now verify the second property.

Without loss of generality, let $\chi(f) = 2$. Now χ can be extended to a triangle-free coloring of each of the S_i . This done, we necessarily have that all of the edges e_i , for $i = 1, \dots, 5a$, get color 1, i.e., C_{5a} is monochromatic. By the assumption that edges e and f are distance 2 apart the only triangles in $G^\uparrow$ are entirely within one copy of S^- . Thus the coloring χ is a triangle-free coloring of $G^\uparrow$. This gives property (ii), thus completing the proof of the lemma. $\square$

For integer $a \geq 6$, we construct the graph $G^\downarrow(a)$ as follows. (See Figure 3.5.)

CONSTRUCTION 3.7.

- (i) Let B be a set of a pairwise vertex disjoint edges, $B = \{e_1, \dots, e_a\}$. (In Construction 3.9, B will be identified with the set $\{\{u_1^i, u_2^i\} \mid i = 1, \dots, a\}$.)
- (ii) For integer $i = 1, \dots, a$, let S_i be a copy of $S^- = S^-(2, K_3, e, f)$, and let e_i and f_i be the copies of e and f in S_i , respectively.

- (iii) For integer $i = 1, \dots, a-3$, let U_i be a copy of S^- , and let g_i and h_i be the copies of e and f in U_i , respectively.
- (iv) For integer $i = 1, \dots, a-2$, let T_i be a triangle with edges p_i , q_i , and r_i .
- (v) Make the following edge identifications (cf. Figure 3.5).
 - Identify f_1 with p_1 .
 - Identify f_{i+1} with q_i for $i = 1, \dots, a-2$.
 - Identify g_i with r_i for $i = 1, \dots, a-3$.
 - Identify h_i with p_{i+1} for $i = 1, \dots, a-3$.
 - Identify f_a with r_{a-2} .

LEMMA 3.8. For integer $a \geq 3$, the graph $G^\downarrow = G^\downarrow(a)$ has the following properties:

- (i) $|V(G^\downarrow)| < (2a-3)|V(S^-)|$.
- (ii) There is no triangle-free 2-coloring of $G^\downarrow$ that is monochromatic on B .
- (iii) Any 2-coloring χ of B , in which one of the color classes contains a single edge, can be extended to a triangle-free 2-coloring of $G^\downarrow$.

Proof. Property (i) follows from the fact that every vertex of $G^\downarrow$ occurs in at least one of the $2a-3$ copies S_i or U_i of S^- .

To verify property (ii), we must show that every 2-coloring of $G^\downarrow$ which is monochromatic on B must contain a monochromatic triangle.

Let $\chi : G^\downarrow \rightarrow \{1, 2\}$ be a 2-coloring of the edges of $G^\downarrow$ such that $\chi(B) = 1$. Assume that each sender S_i or U_i in $G^\downarrow$ is triangle-free under χ . We show that at least one of the triangles in the set $\mathcal{T} = \{T_1, \dots, T_{a-2}\}$ must be monochromatic by showing that $2|\mathcal{T}| + 1 = 2a-3$ of the $3(a-2)$ edges making up the triangles $\mathcal{T}$ get color 2 under χ .

Since $\chi(B) = \{1\}$, the properties of S^- give us that $\chi(f_i) = 2$ for $i = 1, \dots, a$. These f_i are distinct edges of the triangles $\mathcal{T}$. Furthermore, for each $i = 1, \dots, a-3$, U_i is a negative signal sender under a triangle-free coloring, so either $\chi(g_i) = 2$ or $\chi(h_i) = 2$. Again, these are all distinct edges of the triangles $\mathcal{T}$. Thus $a + (a-3) = 2a-3$ of the edges of the triangles $\mathcal{T}$ are color 2. One of the triangles must therefore be monochromatic, and the property is verified.

For property (iii), assume that χ has color 1 on all edges of B except for the i th one (i.e., the one in S_i). If $1 < i < a$, then χ can be extended to a triangle-free coloring of $G^\downarrow$ such that

- for $j < i-1$, $\chi(p_j) = \chi(q_j) = 2$ and $\chi(r_j) = 1$;
- for $j = i-1$, $\chi(p_j) = \chi(r_j) = 2$ and $\chi(q_j) = 1$; and
- for $j > i-1$, $\chi(p_j) = 1$ and $\chi(q_j) = \chi(r_j) = 2$.

If $i = 1$, then χ can be extended to a triangle-free coloring of $G^\downarrow$ such that for all $j = 1, \dots, a-2$, $\chi(p_j) = 1$ and $\chi(q_j) = \chi(r_j) = 2$.

The case that $i = a$ is symmetric to the case $i = 1$. $\square$

3.3. Construction of $G(\mathcal{P})$. Given a set $\mathcal{P} = \{L_i \cup R_i \mid i = 1, \dots, a\}$ of partitions of the set $\{1, \dots, a\}$, we first define the graph $G'(\mathcal{P})$. The graph $G(\mathcal{P})$ will then be a 2-Ramsey-minimal subgraph of this graph.

We construct $G'(\mathcal{P})$ as follows.

CONSTRUCTION 3.9.

- (i) Let G^* be a copy of $G^*(\mathcal{P})$.
- (ii) Let $G^\uparrow$ be a copy of $G^\uparrow(5a)$.
- (iii) Let $G^\downarrow$ be a copy of $G^\downarrow(a)$.
- (iv) Identify the cycle C_{5a} in G^* with the cycle C_{5a} in $G^\uparrow$. For $i = 1, \dots, a$, set $e_i = \{u_1^i, u_2^i\}$; i.e., identify the set $\{\{u_1^i, u_2^i\} \mid i = 1, \dots, a\}$ of pairwise vertex-disjoint edges in G^* with set $B = \{b_i \mid i = 1, \dots, a\}$ of pairwise vertex-disjoint edges in $G^\downarrow$.

Denote the resulting graph by $G'(\mathcal{P})$.

LEMMA 3.10. *The graph $G' = G'(\mathcal{P})$ has the following properties:*

- (i) *Any 2-coloring of G' has a monochromatic triangle.*
- (ii) *For any central edge e of $G^* \subset G'$ there exists a triangle-free 2-coloring of $G' - e$.*

Proof.

(i) Assume there is a triangle-free 2-coloring of G' . Then by Lemmas 3.6(ii) and 3.8(ii), C_{5a} is monochromatic and B is not. Consequently, by Lemma 3.3(ii), G^* in G' contains a monochromatic triangle.

(ii) Let e be a central edge of G^* . By Lemma 3.3(iii), $G^* - e$ can be colored with C_{5a} monochromatic of color 1, and with B getting color 1 on all edges but one. Thus Lemmas 3.6(ii) and 3.8(iii) guarantee that the rest of $G' - e$ can be colored without any monochromatic triangles. $\square$

Let $G(\mathcal{P})$ be any subgraph of $G'(\mathcal{P})$ that is minimal with respect to property (i) of Lemma 3.10. By property (ii) of Lemma 3.10, $G(\mathcal{P})$ must contain all the central edges of G^* . Thus we have the following.

FACT 3.11. *For distinct sets $\mathcal{P}$ and $\mathcal{P}'$, the labeled graphs $G(\mathcal{P})$ and $G(\mathcal{P}')$ are different.*

Let S^+ be the positive sender $S^+(2, K_3, \ell, r)$ used in the construction of $G^*(\mathcal{P})$, and let S^- be the negative sender $S^-(2, K_3, e, f)$ used in the constructions of $G^\uparrow$ and $G^\downarrow$. Let $s = \max(|V(S^+)|, |V(S^-)|)$. We then have the following $r = 2$, $k = 3$ case of Theorem 1.1.

THEOREM 3.12. *There exists a constant $c = c(2, 3) > 1$ such that for $n \geq n_0 = n_0(2, 3) = 99s$ there exist at least c^{n^2} nonisomorphic graphs, on at most n vertices, that are 2-Ramsey-minimal with respect to the triangle.*

Proof. Let a be the largest odd integer for which $a \leq \frac{n}{9s}$. Since $n \geq 99s$, we have $a \geq \frac{n}{9s} - 2 \geq 11 - 2 = 9$, which yields that

$$a \geq \frac{9}{11} \frac{n}{9s} = \frac{n}{11s}.$$

Now there are 2^{a^2} ways to choose the set $\mathcal{P}$ of a partitions $L_i \cup R_i$ of $\{1, \dots, a\}$. By Fact 3.11, each $\mathcal{P}$ yields a different graph $G(\mathcal{P})$. We therefore have 2^{a^2} different graphs $G(\mathcal{P})$. Using part (i) of Lemmas 3.3, 3.6, and 3.8, we get that each has

$$\begin{aligned} |V(G(\mathcal{P}))| &< |V(G^*(\mathcal{P}))| + |V(G^\uparrow(5a))| + |V(G^\downarrow(a))| \\ &< 2as + 5as + 2as \\ &= 9as \leq n \end{aligned}$$

vertices.

Setting $c_1 = 1/(11s)^2$, there are $2^{a^2} > 2^{c_1 n^2}$ different such labeled graphs. The isomorphism class of any one is at most $n!$, so there are at least

$$\frac{2^{c_1 n^2}}{n!} = 2^{c_1 n^2(1+o(1))} > 2^{c_2 n^2},$$

nonisomorphic graphs, for $c_2 < c_1$, and n is sufficiently large.

This is c^{n^2} , where $c = 2^{c_2}$, so the proof is complete. $\square$

4. Theorem 1.1: Induction on k . We prove Theorem 1.1 for $r = 2$ and all k , by induction on k . The induction will be done by means of the following lemma. The

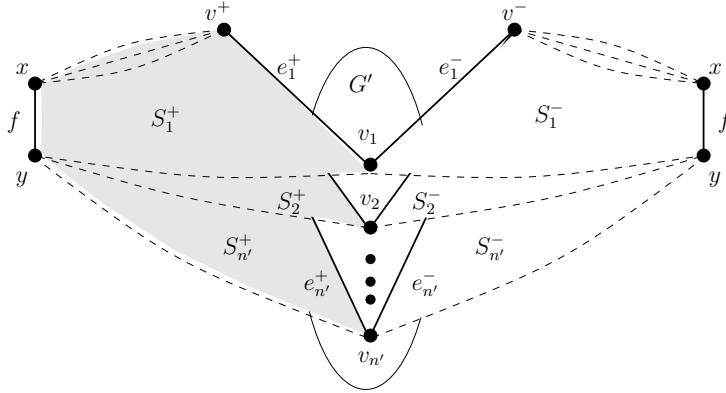


FIG. 4.1. Construction of G from Lemma 4.1. Note that the two edges f are the same edge.

proof of the lemma takes up the majority of the section. The proof of the induction on k comes only at the very end of the section.

LEMMA 4.1. *Given an integer k there exists a constant $s = s(k)$ such that for a graph G' on n' vertices that is 2-Ramsey-minimal with respect to K_{k-1} , there exists a graph G with the following properties:*

- (i) *The graph G has at most sn' vertices,*
- (ii) *G is 2-Ramsey with respect to K_k , and*
- (iii) *any 2-Ramsey-minimal subgraph of G contains G' as a subgraph.*

Proof. Let $V(G') = \{v_1, \dots, v_{n'}\}$, and construct G as follows. (See Figure 4.1.)

- For $i = 1, \dots, n'$ let $S_i^- = S^-(2, K_k, e_i^-, f_i^-)$ and $S_i^+ = S^+(2, K_k, e_i^+, f_i^+)$ be new copies of the senders $S^-(2, K_k, e, f)$ and $S^+(2, K_k, e, f)$, respectively. Assume that in all of these senders, the copies of e and f are distance at least 2 apart.

- Add new vertices v^- and v^+ . For each $i = 1, \dots, n'$, identify the endpoints of e_i^- with v^- and v_i , and the endpoints of e_i^+ with v^+ and v_i .

- Identify all of the f_i^+ and f_i^- and call this edge f . Let x and y be the endpoints of f .

Set $s = |V(S^-(2, K_k, e, f))| + |V(S^+(2, K_k, e, f))|$. We now verify that the graph G , thus constructed, has all the properties promised by the lemma.

To see property (i) notice that every vertex of G is in at least one of the n' copies of $S^-(2, K_k, e, f)$ or the n' copies of $S^+(2, K_k, e, f)$. Thus

$$|V(G)| < n'(|V(S^-(2, K_k, e, f))| + |V(S^+(2, K_k, e, f))|) = sn'.$$

Property (ii) will be verified in Claim 4.2, and property (iii) follows immediately from Claim 4.4.

CLAIM 4.2. *Any 2-coloring of G yields a monochromatic K_k .*

Proof. Let χ be a 2-coloring of G and assume, without loss of generality, that $\chi(f) = 1$. Then by the properties of the S_i^- , all edges from v^- to G' get color 2, and by the properties of the S_i^+ , all edges from v^+ to G' get color 1. Now χ induces a 2-coloring of G' , which yields a monochromatic copy of K_{k-1} in G' . If this K_{k-1} is color 1, then it makes a monochromatic K_k with v^+ . Otherwise, it makes a monochromatic K_k with v^- . $\square$

The following claim is used in Claim 4.4.

CLAIM 4.3. *Any copy K'_k of K_k in G either intersects G' in at least a K_{k-1} or occurs within S_i^- or S_i^+ for some $i = 1, \dots, n'$.*

Proof. We split the proof of the claim into three cases.

Case 1. $|V(K'_k) \cap V(G')| \geq 2$.

Without loss of generality, let v_1 and v_2 be in $|V(K'_k) \cap V(G')|$. Recall that x and y are the endpoints of f . By the assumption that the edges e and f in $S^-(2, K_k, e, f)$ and $S^+(2, K_k, e, f)$ are distance at least 2 apart, the only neighbors of v_1 are in $V(G') \cup V(S_1^+) \cup V(S_1^-) - \{x, y\}$. (Recall that x and y are the endpoints of f .) Similarly the only neighbors of v_2 are in $V(G') \cup V(S_2^+) \cup V(S_2^-) - \{x, y\}$. Thus the only common neighbors of v_1 and v_2 are in $V(G') \cup \{v^-, v^+\}$. Since the vertices v^+ and v^- are nonadjacent, only one of them belongs to $V(K'_k)$. Consequently, $|V(K'_k) \cap V(G')| \geq k - 1$, i.e., K'_k intersects G' in at least a K_{k-1} .

Case 2. $|V(K'_k) \cap V(G')| = 1$.

Without loss of generality, let $V(K'_k) \cap V(G') = \{v_1\}$. The only neighbors of v_1 are in S_1^- or S_1^+ , but there are no edges between $V(S_1^-)$ and $V(S_1^+)$ except those incident to v_1 . Consequently, K'_k occurs within S_1^- or S_1^+ .

Case 3. $|V(K'_k) \cap V(G')| = 0$.

Since the set $\{v^-, v^+, x, y\}$ induces only the one edge f , at most two of the vertices in this set can be in K'_k . Therefore there is some other vertex v of one of the S_i^- or S_i^+ in K'_k . Without loss of generality, assume that $v \in S_1^-$. All the neighbors of v are in S_1^- . Consequently K'_k occurs within S_1^- . $\square$

It follows from the following claim that any 2-Ramsey-minimal subgraph of G when restricted to $V(G')$ is G' .

CLAIM 4.4. *Removing any edge $e' \in G' \subset G$ we can 2-color G with no monochromatic K_k .*

Proof. Removing edge e' from G' , we can 2-color G' without any monochromatic K_{k-1} . Since the senders can be 2-colored without any monochromatic K_k , Claim 4.3 implies that this gives a 2-coloring of G without any monochromatic K_k . $\square$

This completes the proof of the lemma. $\square$

We now prove the induction on k , thus proving the $r = 2$ case of Theorem 1.1. By section 3, we have the base case, $k = 3$.

Proof. Assume that Theorem 1.1 is true for $r = 2$ and $k - 1$. We must provide integer $n_0(2, k)$ and constant $c(2, k) > 1$ such that for any $n \geq n_0(2, k)$ there exist at least $c(2, k)^{n^2}$ nonisomorphic graphs on at most n vertices, each of which is 2-Ramsey-minimal with respect to K_k .

Letting $s = s(k)$ be from Lemma 4.1, set

$$c(2, k) = c(2, k - 1)^{1/(8s^2)}.$$

Observe that because $c(2, k - 1) > 1$, we have that $c(2, k) > 1$. Let $n_0(2, k)$ be large enough that

$$n_0(2, k) \geq s \cdot n_0(2, k - 1),$$

and that for all $n \geq n_0(2, k)$,

$$(4.1) \quad n! \leq c(2, k)^{n^2}.$$

Given

$$n \geq n_0(2, k) \geq s \cdot n_0(2, k - 1),$$

set $n' = \lfloor \frac{n}{s} \rfloor$. Since $n_0(2, k - 1)$ is an integer, this implies

$$n' \geq n_0(2, k - 1).$$

We can therefore use the induction hypothesis to get at least $c(2, k-1)^{(n')^2}$ labeled graphs on at most n' vertices, each of which is 2-Ramsey-minimal with respect to K_{k-1} . Apply Lemma 4.1 to each of these graphs to get graphs on at most $sn' \leq n$ vertices which are 2-Ramsey with respect to K_k .

Since $n_0(2, k-1) \geq n_0(2, 3) = 99$ by Theorem 3.12, we have that $n \geq n_0(2, k) \geq 99s$, and so

$$n' = \left\lfloor \frac{n}{s} \right\rfloor > \frac{n}{s} - 1 > \frac{n}{2s}.$$

Thus there are at least

$$c(2, k-1)^{(n')^2} > c(2, k-1)^{\frac{n^2}{(2s)^2}} = c(2, k)^{2 \cdot n^2}$$

such graphs. By Lemma 4.1(iii), the 2-Ramsey-minimal subgraphs of these graphs are distinct, though not necessarily nonisomorphic. Dividing by $n!$, which is at most $c(2, k)^{n^2}$ by assumption (4.1), we get at least $c(2, k)^{n^2}$ nonisomorphic 2-Ramsey-minimal graphs for K_k , each on at most n vertices. This completes the proof of Theorem 1.1 for case $r = 2$. $\square$

5. Theorem 1.1: Induction on r . In the previous section we proved the $r = 2$ case of Theorem 1.1 by induction on k . In this section we finish the proof of the theorem by induction on r . We do the induction by means of the following lemma.

LEMMA 5.1. *Given integers r and k there exists a constant $s = s(r, k)$ such that for a graph G' on n' vertices that is $(r-1)$ -Ramsey-minimal with respect to K_k , there exists a graph G with the following properties:*

- (i) *The graph G has at most sn' vertices,*
- (ii) *G is r -Ramsey with respect to K_k , and*
- (iii) *any r -Ramsey-minimal subgraph of G contains G' as a subgraph.*

The induction on r uses Lemma 5.1 in the same way as the induction on ℓ used Lemma 4.1 in the previous section.

Let $c(r, k) = c(r-1, k)^{1/8s^2}$. For large enough n we will find at least $c(r, k)^{n^2}$ graphs on less than n vertices that are r -Ramsey minimal for K_k .

By induction, if $n' = \lfloor n/s \rfloor$ is large enough, there are at least $c(r-1, k)^{n'^2}$ graphs on at most n' vertices that are $(r-1)$ -Ramsey minimal for K_k . Applying Lemma 5.1 to each, we get

$$c(r-1, k)n'^2 > c(r-1, k)^{\frac{n^2}{(2s)^2}} = c(r, k)^{2 \cdot n^2}$$

distinct labeled graphs on at most n vertices that are r -Ramsey-minimal for K_k . Dividing by $n!$ we get at least $c(r, k)^{n^2}$ nonisomorphic graphs on at most n vertices (for n sufficiently large). This completes the induction step.

We thus finish the proof of Theorem 1.1 by proving Lemma 5.1.

Proof of Lemma 5.1. Briefly, G is as follows. Consider G' , a copy of K_{k-2} , and an edge f , which are pairwise vertex disjoint. Add edges between every vertex of G' and every vertex of K_{k-2} . Finally, for all edges e with at least one vertex in $V(K_{k-2})$, connect e to f with positive sender $S^+(r, K_k, e, f)$. (Thus G is constructed with $n' \cdot (k-2) + \binom{k-2}{2}$ copies of $S^+(r, K_k, e, f)$.) In more detail, G is as follows:

- Begin with G' . Let $V(G') = \{v_1, \dots, v_{n'}\}$.
- Add new vertices $w_1, \dots, w_{k-2}$, all edges between them, and one more vertex disjoint edge $f = \{x, y\}$.

- For $i = 1, \dots, n'$ and $j = 1, \dots, k-2$, add edge $e_{ij} = \{v_i, w_j\}$ and let $S_{ij} = S^+(r, K_k, \{v_i, w_j\}, f)$ be a copy of $S^+(r, K_k, e, f)$ in which all vertices are distinct except for v_i, w_j, x , and y . Assume that in all of these senders, the copies of e and f are distance at least 2 apart.

- For each i, j such that $1 \leq i < j \leq k-2$, let $W_{ij} = S^+(r, K_k, \{w_i, w_j\}, f)$ be a copy of $S^+(r, K_k, e, f)$ in which all vertices are distinct except for w_i, w_j, x , and y . Assume again that in all of these senders, the copies of e and f are distance at least 2 apart.

Set $s = 2k|V(S^+(r, K_k, e, f))|$. We now verify that the graph G , constructed above, has all the properties promised by the lemma.

To see property (i) notice that every vertex of G is in at least one of the $n' \cdot (k-2) + \binom{k-2}{2}$ copies of $S^+(r, K_k, e, f)$ in G . Since G' has n' vertices, and is Ramsey for K_k , $n' \geq k$. Thus $\binom{k-2}{2} < n' \cdot k$, and so

$$|V(G)| < n'(2k)|V(S^+(r, K_k, e, f))| = sn'.$$

The following claim verifies property (ii).

CLAIM 5.2. *Any r -coloring of G yields a monochromatic K_k .*

Proof. Let $\chi : V(G) \rightarrow \{1, \dots, r\}$ be an r -coloring of G and assume, without loss of generality, that $\chi(f) = r$. If there is no monochromatic K_k in any of the S_{ij} or W_{ij} , then by the properties of these senders, all edges $\{v_i, w_j\}$ and $\{w_i, w_j\}$ get color r . If any edge in G' gets colored r , then this completes a monochromatic K_k with vertices $w_1, \dots, w_{k-2}$. Thus χ must induce an $(r-1)$ -coloring on G' . But G' is $(r-1)$ -Ramsey with respect to K_k , so G contains a monochromatic K_k . $\square$

Property (iii) follows from Claim 5.4. However, for this, we first need the following claim.

CLAIM 5.3. *For any copy K'_k of K_k in G , either $V(K'_k) \subset V(G') \cup \{w_1, \dots, w_{k-2}\}$ or $V(K'_k) \subset V(S^+)$ for some sender $S^+ = S_{ij}^+$ or W_{ij}^+ .*

Proof. First observe that if there exists $v \in V(K'_k)$ such that

$$(5.1) \quad v \in V(S_{ij}^+) - \{v_i, w_j, x, y\} \quad \text{or} \quad v \in V(W_{ij}^+) - \{w_i, w_j, x, y\},$$

then all the neighbors of v are in S_{ij}^+ or W_{ij}^+ , respectively, and thus $V(K'_k) \subset V(S_{ij}^+)$ (or $V(W_{ij}^+)$), so we are done.

We may thus assume that (5.1) does not hold for any $v \in V(K'_k)$. It follows that $V(K'_k) \subset V(G') \cup \{w_1, \dots, w_{k-2}\} \cup \{x, y\}$. Since x and y have distance at least 2 from $V(G') \cup \{w_1, \dots, w_{k-2}\}$, and $k > 3$, we have that $V(K'_k) \cap \{x, y\} = \emptyset$. Thus $V(K'_k) \subset V(G') \cup \{w_1, \dots, w_{k-2}\}$, which finishes the proof of the claim. $\square$

The following claim implies that any r -Ramsey-minimal subgraph of G contains all the edges of G' , and thus property (iii) of the lemma follows.

CLAIM 5.4. *For any edge e of G' , we can r -color $G - e$ with no monochromatic K_k .*

Proof. Let e be in G' . We can $(r-1)$ -color $G' - e$ with colors $\{1, \dots, r-1\}$ without any monochromatic K_k . We extend this $(r-1)$ -coloring of $G' - e$ to an r -coloring of $G - e$. Let all edges $\{v_i, w_j\}$ and $\{w_i, w_j\}$, and the edge f , get color r . For each positive sender S_{ij}^+ (or W_{ij}^+), the copy $\{v_i, w_j\}$ (respectively, $\{w_i, w_j\}$) of e gets the same color as f , so we can extend our coloring to an r coloring of S_{ij}^+ (or W_{ij}^+) with no monochromatic K_k . By Claim 5.3, this gives a K_k -free r -coloring of $G - e$. $\square$

This completes the proof of Lemma 5.1, and by the comments at the beginning of the section, completes the proof of Theorem 1.1 $\square$

6. The nonsymmetric case. In this section we prove the following nonsymmetric case of our main theorem.

THEOREM 6.1. *For any integers $\ell \geq k \geq 3$ there exist constants $c = c(\ell, k) > 1$ and $n_0(\ell, k)$ such that for all $n \geq n_0(\ell, k)$, there exist at least c^{n^2} nonisomorphic graphs, on at most n vertices, that are Ramsey-minimal with respect to K_ℓ and K_k .*

The proof of this theorem uses the following lemma.

LEMMA 6.2. *Given integers $\ell > k \geq 3$ there exists a constant $s = s(\ell, k)$ such that for a graph G' on n' vertices that is 2-Ramsey-minimal with respect to $K_{\ell-1}$ and K_k , there exists a graph G with the following properties:*

- (i) *The graph G has at most sn' vertices,*
- (ii) *G is 2-Ramsey with respect to K_ℓ and K_k , and*
- (iii) *any 2-Ramsey-minimal subgraph of G contains G' as a subgraph.*

The proof of Theorem 6.1 follows by induction on ℓ using Lemma 6.2 in the same way as the $r = 2$ case of Theorem 1.1 followed by induction on k using Lemma 4.1 in section 4. The base case of the induction, when $\ell = k$, is given by Theorem 1.1. Assuming Lemma 6.2, and that the theorem is proved for $\ell - 1$ and k , the induction step is as follows.

Let $c(\ell, k) = c(\ell - 1, k)^{1/8s^2} > 1$ and $n_0(\ell, k) = s \cdot n_0(\ell - 1, k)$. For $n \geq n_0(\ell, k)$ we will find at least $c(\ell, k)^{n^2}$ graphs on less than n vertices that are 2-Ramsey minimal for K_ℓ and K_k .

Given $n \geq n_0(\ell, k)$, let $n' = \lfloor n/s \rfloor \geq n_0(\ell - 1, k)$. By induction, there are at least $c(\ell - 1, k)^{n'^2}$ graphs that are 2-Ramsey minimal for $K_{\ell-1}$ and K_k on at most n' vertices. Applying Lemma 6.2 to each, we get

$$c(\ell - 1, k)^{n'^2} > c(\ell - 1, k)^{\frac{n^2}{(2s)^2}} = c(\ell, k)^{2 \cdot n^2}$$

distinct labeled graphs on at most n vertices that are 2-Ramsey-minimal for K_ℓ and K_k . Dividing by $n!$ we get at least $c(\ell, k)^{n^2}$ nonisomorphic graphs on at most n vertices (for n sufficiently large). This completes the induction step.

We thus finish the proof of Theorem 6.1 by proving Lemma 6.2. First, we need a result from [10].

DEFINITION 6.3. *Given graphs H_1 and H_2 , a graph $D = D(H_1, H_2, e)$ is an (H_1, H_2) -determiner if*

- (i) *there is a 2-coloring $\chi : D \rightarrow \{1, 2\}$ of D such that there is no copy of H_1 of color 1 and no copy of H_2 of color 2, and*
- (ii) *in any such coloring χ of D , $\chi(e) = 1$.*

Theorem 2.3 of [10] states that for any two integers $\ell > k \geq 3$, there exists a (K_ℓ, K_k) -determiner $D = D(\ell, k, e)$. With this we can prove the lemma.

Proof of Lemma 6.2. Given a graph G' as described in the lemma, let $V(G') = \{v_1, \dots, v_{n'}\}$, and construct G as follows:

- For $i = 1, \dots, n'$, let $D_i = D(\ell, k, e_i)$ be a copy of the (K_ℓ, K_k) -determiner D .
- Add new vertex v_0 . For each $i = 1, \dots, n'$, identify the endpoints of e_i with v_0 and v_i .

The verification that the graph G constructed above has the necessary properties of the lemma is very similar to the proof of Lemma 4.1. $\square$

7. Signal senders. In this section, we give the proof of Lemma 2.2. The proof is a modification of the proof from [10]. Recall that this lemma guarantees, for any 3-connected graph H , and any $r \geq 2$, the existence of

- (a) a negative signal sender $S^-(r, H, e, f)$,
- (b) a positive signal sender $S^+(r, H, e, f)$, and
- (c) such negative and positive signal senders in which the distance between the edges e and f is arbitrarily large.

The section is organized as follows. In subsection 7.1, we introduce a few technical definitions and results that we will use in the proof. In subsection 7.2, we construct a negative signal sender $S^-(r, H, e, f) = S^-(r, H, \{x_0, u\}, \{x_0, u'\})$, in which the vertices u and u' are nonadjacent. This strengthening of part (a) of the lemma will help us in the later parts. In subsection 7.3, we use part (a) to construct a positive signal sender $S^+(r, H, e, f) = S^+(r, H, \{v_0, v_1\}, \{v_0, v_t\})$, in which the vertices v_1 and v_t are nonadjacent. Finally, in subsection 7.4, we use parts (a) and (b) to construct, for any integer $g \geq 2$, positive and negative signal senders $S^+(r, H, \{x_0, x_1\}, \{x_{g+1}, x_{g+2}\})$ and $S^-(r, H, \{x_0, x_1\}, \{x_{g+1}, x_{g+2}\})$ in which the edges $\{x_0, x_1\}$ and $\{x_{g+1}, x_{g+2}\}$ are distance g apart.

7.1. Technical definitions and results. We begin with a couple of definitions.

DEFINITION 7.1. *A circuit of length h in an ℓ -graph $\mathcal{H}$ is a subgraph $\{e_1, \dots, e_h\}$ of $\mathcal{H}$ such that*

$$\left| \bigcup_{i=1}^h e_i \right| \leq (\ell - 1)h.$$

The girth of an ℓ -graph is the length of its shortest circuit.

DEFINITION 7.2. *An ℓ -graph is $(r+1)$ -chromatic if $r+1$ is the minimum number of colors needed to color its vertices such that there is no edge with all vertices having the same color.*

The following is proved probabilistically in [11] and constructively in [14] and [17].

THEOREM 7.3. *For any integers $g, r \geq 2$ and $\ell \geq 3$, there exist $(r+1)$ -chromatic ℓ -graphs $\mathcal{G}$ of girth $> g$.*

We will also need the following technical definitions.

DEFINITION 7.4. *Let $r, \ell \geq 2$ be integers.*

- (i) *An oriented ℓ -graph $\mathcal{H}$ on the vertices $V = V(\mathcal{H})$ is a set of ordered ℓ element subsets, or arcs, of V .*
- (ii) *An (r, ℓ) -pattern is an ℓ -tuple of not necessarily distinct elements from $[r]$, (i.e., an element of $[r]^\ell$.)*
- (iii) *Given a set $\Phi \subseteq [r]^\ell$ of forbidden (r, ℓ) -patterns, an r -coloring $\chi : V(\mathcal{H}) \rightarrow [r]$ of the vertices of $\mathcal{H}$ is Φ -avoiding if*

$$(\chi(v_1), \dots, \chi(v_\ell)) \notin \Phi \text{ for any arc } (v_1, \dots, v_\ell) \in \mathcal{H}.$$

Observe that if $\Phi = \{(i, i, \dots, i) \mid 1 \leq i \leq r\}$ is the set of monochromatic (r, ℓ) -patterns, then a Φ -avoiding r -coloring of an oriented ℓ -graph $\mathcal{H}$ is simply an r -coloring of the vertices of $\mathcal{H}$ in which no arc is monochromatic.

7.2. The negative signal sender $S^-(r, H, \{x_0, u\}, \{x_0, u'\})$. In this subsection we prove a slightly stronger version of part (a) of the lemma by constructing, for any integer $r \geq 2$ and any 3-connected graph H , the negative signal sender $S^-(r, H, \{x_0, u\}, \{x_0, u'\})$ in which the vertices u and u' are nonadjacent.

Let $r \geq 2$ be fixed and H be a 3-connected graph.

The graph F . Let F' be a graph that is r -Ramsey-minimal with respect to H , and let F be obtained from F' by removing an edge $\{x, y\}$. We claim the following:

- (*) Under any H -free r -coloring of the edges of F , x has incident edges of every color.

Indeed, suppose that χ is an H -free r -coloring of F in which x has no incident edges of color i . By setting $\chi(\{x, y\}) = i$, we can extend χ to an r -coloring of F' in which x has only one incident edge of color i . By the r -Ramsey-minimality of F' , F' must then contain a copy of H of color i , in which the vertex x has degree 1. But since H is 3-connected, any vertex has degree at least 3. This is a contradiction.

Indeed, since H is 3-connected, each of its vertices has degree at least 2. So if χ were an H -free r -coloring of F and x had no incident edge of color i , then we could extend χ to an H -free r -coloring of F' by setting $\chi(\{x, y\}) = i$. This would be a contradiction.

The set Φ of forbidden (r, ℓ) -patterns, and the oriented ℓ -graph $\mathcal{H}$. Let F , with vertex x , be the graph constructed in the previous paragraph. Let ℓ be the degree of x in F . Fix an orientation $(x_1, \dots, x_\ell)$ of the neighborhood of x in F . Let Φ be the set of (r, ℓ) -patterns $(r_1, \dots, r_\ell)$ such that the following holds:

The r -coloring χ of the edges $\{x, x_1\}, \dots, \{x, x_\ell\}$ of F defined by $\chi(\{x, x_i\}) = r_i$, for all i , cannot be extended to an H -free r -coloring of the edges of F .

By statement (*), Φ contains all (r, ℓ) -patterns with less than r colors; in particular, it contains the monochromatic pattern $(i, i, \dots, i)$ for $i = 1, \dots, r$. Moreover, because F has an H -free r -coloring, Φ is not the set of all (r, ℓ) -patterns. The following lemma guarantees the existence of an auxiliary oriented ℓ -graph.

LEMMA 7.5. *Given integers $g, r, \ell \geq 2$, let Φ be a proper subset of the set of all (r, ℓ) -patterns, which contains the monochromatic pattern $(i, i, \dots, i)$ for each $i = 1, \dots, r$. Then there exists an oriented ℓ -graph $\mathcal{H} = \mathcal{H}(\Phi, r, \ell, g)$ with girth $> g$, having distinguished vertices u and u' , which satisfies the following:*

- (i) $\mathcal{H}$ has a Φ -avoiding r -coloring.
- (ii) The vertices u and u' do not occur together in any arc of $\mathcal{H}$.
- (iii) Under any Φ -avoiding r -coloring, u and u' get different colors.

The proof of Lemma 7.5 is given at the end of this subsection. Let $\mathcal{H}$ be the oriented ℓ -graph provided by the lemma for the set Φ .

The construction of S^- . We now construct $S^- = S^-(r, H, \{x_0, u\}, \{x_0, u'\})$. (See Figure 7.1.)

CONSTRUCTION 7.6 ($S^- = S^-(r, H, \{x_0, u\}, \{x_0, u'\})$).

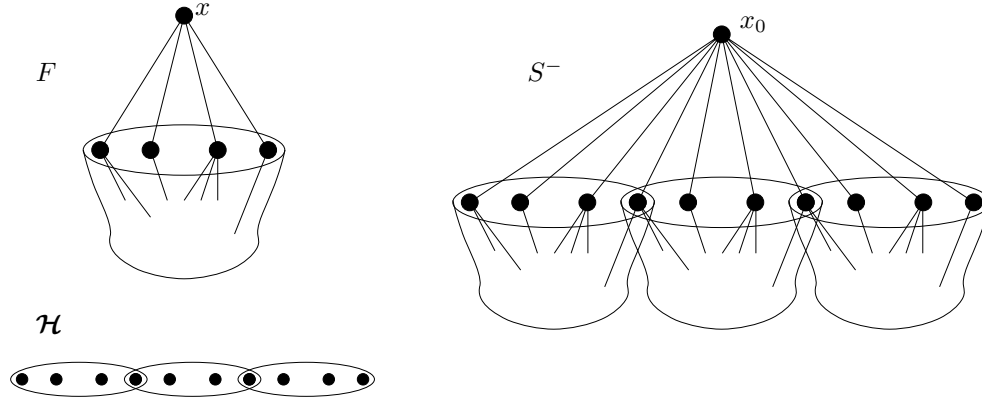
(i) For every arc $\mathbf{a} = (v_1, \dots, v_\ell)$ in $\mathcal{H}$ let $F^{\mathbf{a}}$ be a copy of F . The graph S^- is the union of these copies of F . All vertices are distinct unless explicitly identified below. Let $x^{\mathbf{a}}$ and $\{x_1^{\mathbf{a}}, \dots, x_\ell^{\mathbf{a}}\}$ denote the copies of x and its neighborhood $\{x_1, \dots, x_\ell\}$ in $F^{\mathbf{a}}$.

(ii) For every arc $\mathbf{a} = (v_1, \dots, v_\ell)$ in $\mathcal{H}$ and every $i = 1, \dots, \ell$, identify the vertex $x_i^{\mathbf{a}}$ of $F^{\mathbf{a}}$ with v_i of $\mathcal{H}$.

(iii) Identify the vertices $x^{\mathbf{a}}$ for all $\mathbf{a} \in \mathcal{H}$ and call the resulting vertex x_0 . As a result of this identification, we have that $N(x_0) = V(\mathcal{H})$ in S^- .

By property (ii) of Lemma 7.5, the vertices u and u' do not occur together in any arc of $\mathcal{H}$. Thus they are not adjacent in the graph S^- just constructed.

The proof that S^- is a negative signal sender. To finish the proof of part (a) of the lemma, we need to verify that $S^- = S^-(r, H, \{x_0, u\}, \{x_0, u'\})$ is a negative signal

FIG. 7.1. Construction of S^- from F and $\mathcal{H}$.

sender. The proof consists of the following three claims, the first being used to prove the second.

CLAIM 7.7. *Any copy of H in S^- is entirely within F^a for some $a \in \mathcal{H}$.*

Proof. Let H_0 be a copy of H in S^- . For any subgraph $\mathcal{A}$ (i.e., subset of arcs) of $\mathcal{H}$, let $F^{\mathcal{A}}$ be the subgraph of S^- induced by the vertex set $\bigcup_{a \in \mathcal{A}} V(F^a) \subset V(S^-)$. Let $\mathcal{A}^*$ be a minimum subgraph of $\mathcal{H}$ for which $H_0 \subset F^{\mathcal{A}^*}$. We will show that $|\mathcal{A}^*| = 1$. Towards a contradiction, assume that $|\mathcal{A}^*| > 1$. By the minimality of $\mathcal{A}^*$, $|\mathcal{A}^*| \leq |V(H_0)| = g$, so being a subgraph of $\mathcal{H}$, $\mathcal{A}^*$ contains no circuits. This implies that there exists $a_0 \in \mathcal{A}^*$ such that a_0 intersects the rest of $\mathcal{A}^*$ in at most one vertex, i.e., such that $|V(a_0) \cap V(\mathcal{A}^* - a_0)| \leq 1$.

If $V(a_0) \cap V(\mathcal{A}^* - a_0)$ is empty, then removing x_0 (see Construction 7.6(iii)) disconnects F^{a_0} from the rest of $F^{\mathcal{A}^*}$. But by the minimality of $F^{\mathcal{A}^*}$, the intersection of $V(H_0)$ and $V(F^{a_0}) - \{x_0\}$ is nontrivial, so this disconnects H_0 .

On the other hand, if $V(a_0) \cap V(\mathcal{A}^* - a_0)$ contains some vertex y , then y is adjacent to x_0 and the edge $\{x_0, y\}$ is in F^{a_1} for some other $a_1 \in \mathcal{A}^*$. Removing x_0 and y disconnects F^{a_0} from the rest of $F^{\mathcal{A}^*}$, but by the minimality of $\mathcal{A}^*$, the intersection of $V(H_0)$ and $V(F^{a_0}) - \{x_0, y\}$ is nontrivial, so this disconnects H_0 .

Either way we have disconnected H_0 by removing at most two vertices. This contradicts the fact that H_0 is 3-connected. Thus $|\mathcal{A}^*| = 1$ and the claim is proved. $\square$

CLAIM 7.8. *There exists an H -free r -coloring of S^- .*

Proof. Let χ' be a Φ -avoiding r -coloring of the vertices of $\mathcal{H}$. For each vertex v in $\mathcal{H}$, let $\chi(\{x_0, v\}) = \chi'(v)$. For each copy F^a of F this gives a coloring of the edges $(\{x_0, x_1^a\}, \dots, \{x_0, x_\ell^a\})$ which corresponds to a pattern not forbidden by Φ . Thus we can extend the r -coloring χ to the edges of F^a in such a way that there are no monochromatic copies of H in F^a . By Claim 7.7, there are no other copies of H in S^- , and so χ is an H -free r -coloring of S^- . $\square$

CLAIM 7.9. *Under any H -free r -coloring of S^- , the edges $\{x_0, u\}$ and $\{x_0, u'\}$ get different colors.*

Proof. Assume that we are given an H -free r -coloring χ of S^- . This induces an r -coloring χ' of $V(\mathcal{H})$ by $\chi'(v) = \chi(\{x_0, v\})$. For any arc $a = (v_1, \dots, v_\ell)$ in $\mathcal{H}$ this has the coloring

$$(\chi'(v_1), \dots, \chi'(v_\ell)) = (\chi(\{x_0, v_1\}), \dots, \chi(\{x_0, v_\ell\})) = (\chi(\{x_0, x_1^a\}), \dots, \chi(\{x_0, x_\ell^a\})),$$

which corresponds to a pattern not in Φ , since χ is an H -free r -coloring of F^a . Thus χ' is a Φ -avoiding r -coloring of $\mathcal{H}$. But any Φ -avoiding r -coloring χ' of $\mathcal{H}$ has the property that $\chi'(u) \neq \chi'(u')$. Thus $\chi(\{x_0, u\}) \neq \chi(\{x_0, u'\})$. $\square$

We have shown that S^- is a negative signal sender $S^-(r, H, \{x_0, u\}, \{x_0, u'\})$ in which the vertices u and u' are nonadjacent. The proof of part (a) of the lemma will thus be complete with the proof of Lemma 7.5. We give this now.

Proof of Lemma 7.5. Let $\mathcal{G}$ be the $(r+1)$ -chromatic ℓ -graph with girth $> g$, which is guaranteed by Theorem 7.3, and let Φ be as in the statement of the lemma. Since Φ contains all monochromatic (r, ℓ) -patterns, there is no Φ -avoiding r -coloring of the vertices of $\mathcal{G}$ under any orientation of its edges. Let $\mathcal{G}$ be $\mathcal{G}$ with some fixed orientation of the edges. Clearly $\mathcal{G}$ has no Φ -avoiding r -coloring. Let $\mathcal{H}_0$ be a subgraph of $\mathcal{G}$ that is minimal with respect to the fact that it has no Φ -avoiding r -coloring. Since Φ does not contain all (r, ℓ) -patterns there is a Φ -avoiding r -coloring of any graph with just a single arc, so $\mathcal{H}_0$ is nonempty.

We now construct ℓ different graphs $H_1, \dots, H_\ell$ and show that one of them has the properties required of $\mathcal{H}$. Let $(u_1, \dots, u_\ell)$ be an arc of $\mathcal{H}_0$, and let $v_1, \dots, v_\ell$ be new vertices. For $i = 1, \dots, \ell$, form the oriented ℓ -graph $\mathcal{H}_i$ from $\mathcal{H}_0$ by removing the arc $(u_1, \dots, u_\ell)$ and adding the arc $(v_1, \dots, v_i, u_{i+1}, \dots, u_\ell)$. Now $\mathcal{H}_\ell$ clearly has a Φ -avoiding r -coloring, while $\mathcal{H}_0$ does not. Let $i > 0$ be such that $\mathcal{H}_i$ has a Φ -avoiding r -coloring and $\mathcal{H}_{i-1}$ does not. We show that $\mathcal{H} = \mathcal{H}_i$, $u = u_i$, and $u' = v_i$ satisfy the requirements of the lemma. Certainly $\mathcal{H}$ has a Φ -avoiding r -coloring, so we have property (i) of the lemma. Since v_i is in only one arc of $\mathcal{H}$, and this arc does not contain u_i , we have property (ii) of the lemma. For property (iii), we must now show that under any Φ -avoiding r -coloring χ of $\mathcal{H}$, $\chi(u_i) \neq \chi(v_i)$. Assume that $\chi(u_i) = \chi(v_i)$; then χ is a Φ -avoiding r -coloring of $\mathcal{H}_{i-1}$. This is a contradiction, so we are done. $\square$

This completes the proof of Lemma 7.5, and thus the proof of the strengthening of part (a) of Lemma 2.2.

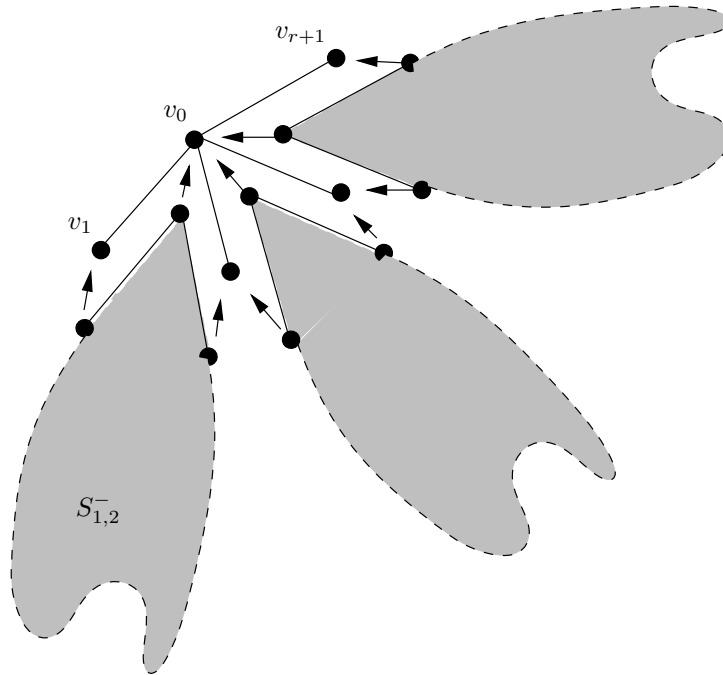
7.3. The proof of part (b) of Lemma 2.2. We first prove the case $H = K_k$ for any $k \geq 3$ by constructing a positive signal sender $S^+ = S^+(r, H, \{v_0, v_1\}, \{v_0, v_{r+1}\})$ which, like in part (a), has the added property that the vertices v_1 and v_{r+1} are not adjacent.

We then sketch the extension of the proof to the case of a general 3-connected graph H by constructing a positive signal sender $S^+ = S^+(r, H, \{v_0, v_1\}, \{v_0, v_t\})$ (t will be determined later), again having the property that the vertices v_1 and v_t are not adjacent.

In both cases, we construct S^+ , show that the vertices v_1 and v_{r+1} (or v_t) are not adjacent, and then prove that S^+ is a positive signal sender by proving analogues of Claims 7.7, 7.8, and 7.9.

The construction of S^+ in the case $H = K_k$. Take the star with root v_0 and leaves $v_1, \dots, v_{r+1}$. For each pair $1 \leq i < j \leq r+1$ except $i = 1$ and $j = r+1$, let $S_{i,j}^- = S^-(r, K_k, \{v_0, v_i\}, \{v_0, v_j\})$ be a copy of $S^-(r, K_k, \{x_0, u\}, \{x_0, u'\})$ in which we identify the copies of x_0, u , and u' with the vertices v_0, v_i , and v_j , respectively. Let the $\binom{r+1}{2} - 1$ copies of $S^-(r, K_k, \{x_0, u\}, \{x_0, u'\})$ be otherwise vertex disjoint, and let S^+ be their union. (See Figure 7.2.)

By the construction, it is clear that the vertices v_1 and v_{r+1} are not adjacent. To finish the proof of part (b) for the case $H = K_k$, we show that S^+ is a positive signal sender. We do this by proving the following claim, the parts of which are analogues to Claims 7.7, 7.8, and 7.9.

FIG. 7.2. A subgraph of S^+ .

CLAIM 7.10. *The graph S^+ has the following properties:*

- (i) *Any copy of K_k in S^+ occurs within $S_{i,j}^-$ for some $1 \leq i < j \leq r+1$.*
- (ii) *There exists a K_k -free r -coloring of S^+ .*
- (iii) *Under any K_k -free r -coloring of S^+ , the edges $\{v_0, v_1\}$ and $\{v_0, v_{r+1}\}$ get the same color.*

Proof. (i) Since each $S_{i,j}^- = S^-(r, K_k, \{v_0, v_i\}, \{v_0, v_j\})$ has the property that the vertices v_i and v_j are nonadjacent, the vertices $v_1, \dots, v_{r+1}$ of S^+ form an independent set. Thus any copy K of K_k in S^+ must contain at least $k-2$ vertices of $V(S^+) - \{v_0, v_1, \dots, v_{r+1}\}$. Let x be one such vertex. Then x is in $V(S_{i,j}^-) - \{v_0, v_i, v_j\}$ for some $1 \leq i < j \leq r+1$. But v_0, v_i, v_j are the only vertices of $S_{i,j}^-$ that have neighbors in S^+ that are outside of $S_{i,j}^-$. Thus $K \subset S_{i,j}^-$.

(ii) Define a K_k -free r -coloring χ of S^+ as follows. Let $\chi(\{v_0, v_{r+1}\}) = 1$ and $\chi(\{v_0, v_i\}) = i$ otherwise. For each of the $S_{i,j}^-$, χ gives different colors to the distinguished edges $\{v_0, v_i\}$ and $\{v_0, v_j\}$, and thus can be extended to a K_k -free r -coloring of $S_{i,j}^-$. By property (i), this yields a K_k -free coloring of S^+ .

(iii) Under any K_k -free r -coloring of S^+ , every pair of edges $\{v_0, v_i\}$ and $\{v_0, v_j\}$, for $1 \leq i < j \leq r+1$, except $\{v_0, v_1\}$ and $\{v_0, v_{r+1}\}$, gets distinct colors. Thus $\{v_0, v_1\}$ and $\{v_0, v_{r+1}\}$ get the same color. $\square$

This concludes the proof of the claim, and thus the proof of part (b) of the lemma in the particular case that $H = K_k$. Observe that this construction can be seen as taking G to be a clique on the vertices $v_1, \dots, v_{r+1}$ with the edge $\{v_1, v_{r+1}\}$ removed and letting S^+ be the “amalgam of copies of S^- along edges of G .” We now extend this proof to the general case by constructing a positive signal sender $S^+ = S^+(r, H, \{v_0, v_1\}, \{v_0, v_t\})$, for any 3-connected graph H , in which the vertices

v_1 and v_t are not adjacent. In extending the proof from the case where $H = K_k$, we have to be more careful to ensure the analogue of Claim 7.10(i). To do this, we let G be an $(r+1)$ -critical graph of large girth with an edge removed, rather than an $(r+1)$ -clique with an edge removed. The construction is otherwise the same.

Construction of S^+ in the case that H is 3-connected. Define the auxiliary graph G as follows. Begin with an $(r+1)$ -critical graph with vertices $\{v_1, \dots, v_t\}$, and having girth greater than $|V(H)|$. Without loss of generality, assume that $\{v_1, v_t\}$ is an edge of this graph, and let G be this graph with the edge $\{v_1, v_t\}$ removed. Notice that under any proper r -coloring ψ of the vertices of G , we have that $\psi(v_1) = \psi(v_t)$.

Construct S^+ as follows. Let v_0 be a vertex not in G . For every edge $\{v_i, v_j\} \in G$, let $S_{i,j}^- = S^-(r, H, \{v_0, v_i\}, \{v_0, v_j\})$ be a copy of $S^-(r, H, \{x_0, u\}, \{x_0, u'\})$. Apart from the vertices $v_0, v_1, \dots, v_t$ let these copies of $S^-(r, H, \{x_0, u\}, \{x_0, u'\})$ be vertex disjoint. That is, let $V(S_{i,j}^-) \cap V(S_{i',j'}^-)$ be $\{v_0\}$ if $i \neq i'$ and $j \neq j'$, $\{v_0, v_i\}$ if $i = i'$ and $j \neq j'$, and $\{v_0, v_j\}$ if $i \neq i'$ and $j = j'$. Let S^+ be the union of these $S_{i,j}^-$.

By the construction, the vertices v_1 and v_t are nonadjacent in S^+ . We thus have only to prove that $S^+ = S^+(r, H, \{v_0, v_1\}, \{v_0, v_t\})$ is a positive signal sender. We do so with the following claim.

CLAIM 7.11. *The graph S^+ has the following properties:*

- (i) *Any copy of H in S^+ occurs within $S_{i,j}^-$ for some $\{v_i, v_j\} \in G$.*
- (ii) *There exists an H -free r -coloring of S^+ .*
- (iii) *Under any H -free r -coloring of S^+ , the edges $\{v_0, v_1\}$ and $\{v_0, v_t\}$ get the same colors.*

Sketch of proof. Property (i) follows from the fact that we started with a graph G' of girth greater than $|V(H)|$, and from the 3-connectedness of H . The proof is a simpler version of the proof of Claim 7.7, and so we omit it.

The proofs of properties (ii) and (iii) are completely analogous to the corresponding properties in the K_k case, so again we omit them. $\square$

This completes the proof of the claim, and thus the proof of part (b) of the lemma.

7.4. The proof of part (c) of Lemma 2.2. Fix an integer $g > 0$. We now prove that there exist a positive signal sender $S^+ = S^+(r, H, \{x_0, x_1\}, \{x_{g+1}, x_{g+2}\})$, and a negative signal sender $S^- = S^-(r, H, \{x_0, x_1\}, \{x_{g+1}, x_{g+2}\})$, such that in each of them, the edges $\{x_0, x_1\}$ and $\{x_{g+1}, x_{g+2}\}$ are distance g apart. Again, the proofs are constructive.

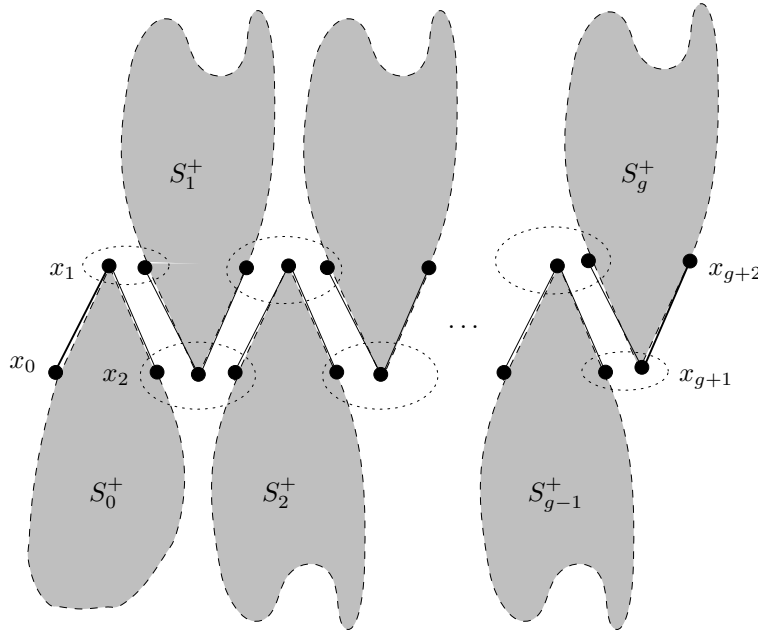
The construction of S^+ . For $i = 0, \dots, g$, let $S_i^+ = S^+(r, H, \{x_i, x_{i+1}\}, \{x_{i+1}, x_{i+2}\})$ be a copy of $S^+(r, H, \{v_0, v_1\}, \{v_0, v_t\})$ from part (b) of the lemma. Let S^+ be the union of these $g+1$ graphs, where all vertices except $x_1, \dots, x_{g+1}$ are distinct. (See Figure 7.3.)

Observe that S^+ contains the path $x_1, \dots, x_{g+1}$ of length g from $\{x_0, x_1\}$ to $\{x_{g+1}, x_{g+2}\}$. Since the vertices x_i and x_{i+2} were not adjacent in S_i^+ , for $i = 0, \dots, g$, and the only edges of S^+ are from the graphs S_i^+ , this is the shortest path between $\{x_0, x_1\}$ and $\{x_{g+1}, x_{g+2}\}$. Thus these edges are distance g apart.

To show that $S^+ = S^+(r, H, \{x_0, x_1\}, \{x_{g+1}, x_{g+2}\})$ is a positive signal sender, we prove the following claim.

CLAIM 7.12. *The graph S^+ has the following properties:*

- (i) *Any copy of H in S^+ occurs within S_i^+ for some $0 \leq i \leq g$.*
- (ii) *There exists an H -free r -coloring of S^+ .*
- (iii) *Under any H -free r -coloring of S^+ , the edges $\{x_0, x_1\}$ and $\{x_{g+1}, x_{g+2}\}$ get the same color.*

FIG. 7.3. S^+ where distinguished edges are distance g apart.

Sketch of proof. The proof of property (i) is very similar to the proof of property (i) of Claim 7.10.

To see that property (ii) is true, observe that any coloring that is constant on the edges $\{v_0, v_1\}, \{v_1, v_2\}, \dots, \{v_{g+1}, v_{g+2}\}$ can be extended to an H -free r -coloring of S^+ .

For property (iii), observe that any H -free r -coloring of S^+ is constant on the edges $\{v_0, v_1\}, \{v_1, v_2\}, \dots, \{v_{g+1}, v_{g+2}\}$. $\square$

The construction of S^- . Use exactly the same construction as for S^+ above, but replace the first positive sender, $S_0^+ = S^+(r, H, \{x_0, x_1\}, \{x_1, x_2\})$, with the negative sender $S^-(r, H, \{x_0, x_1\}, \{x_1, x_2\})$. Let S^- be the resulting graph.

The proof that S^- is a negative signal sender in which the edges $\{x_0, x_1\}$ and $\{x_{g+1}, x_{g+2}\}$ are distance g apart is very similar to the proof above for S^+ .

This completes the proof of the lemma. $\square$

REFERENCES

- [1] B. BOLLOBÁS, J. DONADELLI, Y. KOHAYAKAWA, AND R. H. SCHELP, *Ramsey minimal graphs*, J. Brazilian Comput. Soc., 7 (2002), pp. 27–37.
- [2] S. BURR, P. ERDŐS, R. FAUDREE, C. ROUSSEAU, AND R. SCHELP, *Ramsey-minimal graphs for forests*, Discrete Math., 38 (1982), pp. 23–32.
- [3] S. BURR, P. ERDŐS, R. FAUDREE, C. ROUSSEAU, AND R. SCHELP, *Ramsey-minimal graphs for star-forests*, Discrete Math., 33 (1981), pp. 227–237.
- [4] S. BURR, P. ERDŐS, R. FAUDREE, C. ROUSSEAU, AND R. SCHELP, *Ramsey-minimal graphs for matchings*, in The Theory and Applications of Graphs (Kalamazoo, MI 1980), Wiley, New York, 1981, pp. 159–168.
- [5] S. BURR, P. ERDŐS, R. FAUDREE, C. ROUSSEAU, AND R. SCHELP, *Ramsey-minimal graphs for the pair star, connected graph*, Studia Sci. Math. Hungar., 15 (1980), pp. 265–273.

- [6] S. BURR, P. ERDŐS, R. FAUDREE, C. ROUSSEAU, AND R. SCHELP, *Ramsey-minimal graphs for multiple copies*, Nederl. Akad. Wetensch. Indag. Math., 40 (1978), pp. 187–195.
- [7] S. BURR, P. ERDŐS, AND L. LOVASZ, *On graphs of Ramsey type*, Ars Combin., 1 (1976), pp. 167–190.
- [8] S. BURR, R. FAUDREE, AND R. SCHELP, *On graphs with Ramsey-infinite blocks*, European J. Combin., 6 (1985), pp. 129–132.
- [9] M. BOROWIECKI, M. HAŁUSZCZAK, AND E. SIDOROWICZ, *On Ramsey minimal graphs*, Discrete Math., 286 (2004), pp. 37–43.
- [10] S. BURR, J. NEŠETŘIL, AND V. RÖDL, *On the use of senders in generalized Ramsey theory for graphs*, Discrete Math., 54 (1985), pp. 1–13.
- [11] P. ERDŐS AND A. HAJNAL, *On chromatic numbers of graphs and set systems*, Acta Math. Acad. Sci. Hungar., 17 (1966), pp. 61–99.
- [12] R. FAUDREE, *Ramsey minimal graphs for forests*, Ars Combin., 31 (1991), pp. 117–124.
- [13] R. FAUDREE AND J. SHEEHAN, *Tree Ramsey minimal graphs*, in Proceedings of the Thirteenth Annual Southeastern Conference on Combinatorics, Graph Theory and Computing (Boca Raton, FL, 1982), Congr. Numer. 35, Utilitas Math., Winnipeg, 1982, pp. 295–315.
- [14] L. LOVÁSZ, *On chromatic numbers of finite set systems*, Acta Math. Acad. Sci. Hungar., 19 (1968), pp. 59–67.
- [15] T. ŁUCZAK, *On Ramsey minimal graphs*, Electron. J. Combin., 1 (1994), Research Paper 4 (electronic).
- [16] J. NEŠETŘIL AND V. RÖDL, *The structure of critical Ramsey graphs*, Acta Math. Acad. Sci. Hungar., 32 (1978), pp. 295–300.
- [17] J. NEŠETŘIL AND V. RÖDL, *A short proof of the existence of highly chromatic hypergraphs without short cycles*, J. Combin. Theory Ser. B, 27 (1979), pp. 225–227.
- [18] J. NEŠETŘIL AND V. RÖDL, *On Ramsey minimal graphs*, in Problèmes Combinatoires et Théorie des Graphes, Colloques Internationaux C.N.R.S. 260, Orsay, 1976, pp. 307–309.
- [19] V. RÖDL AND A. RUCIŃSKI, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc., 8 (1995), pp. 917–942.

GRAPH-DIFFERENT PERMUTATIONS*

JÁNOS KÖRNER[†], CLAUDIA MALVENUTO[†], AND GÁBOR SIMONYI[‡]

Abstract. For a finite graph G whose vertices are different natural numbers we call two infinite permutations of the natural numbers G -different if they have two adjacent vertices of G somewhere in the same position. The maximum number of pairwise G -different permutations of the naturals is always finite. We study this maximum as a graph invariant and relate it to a problem of the first two authors on colliding permutations. An improvement on the lower bound for the maximum number of pairwise colliding permutations is obtained.

Key words. extremal combinatorics, Shannon capacity of graphs, permutations

AMS subject classifications. 05D05, 05A15

DOI. 10.1137/070692686

1. Introduction. In [1] the first two authors began to investigate the following mathematical puzzle. Call two permutations of $[n] := \{1, \dots, n\}$ *colliding* if, represented by linear orderings of $[n]$, they put two consecutive elements of $[n]$ somewhere in the same position. For the maximum cardinality $\rho(n)$ of a set of pairwise colliding permutations of $[n]$ the following conjecture was formulated.

CONJECTURE 1 (see [1]). *For every $n \in \mathbb{N}$*

$$\rho(n) = \binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

It was observed in [1] that the right-hand-side expression above is actually an upper bound for $\rho(n)$, while the best lower bound given there was a somewhat deceiving

$$(1) \quad 35^{n/7-O(1)} \leq \rho(n).$$

The initial motivation for the present work was to improve on the above lower bound. However, our main purpose in this paper is to study the new graph invariant mentioned in the abstract. We will analyze its possible values for simple classes of graphs and then apply some of the results to obtain a new lower bound for $\rho(n)$.

For brevity's sake let us call a graph *natural* if its vertex set is a finite subset of $\mathbb{N}$, the set of all positive integers, and if the graph is simple (without loops and multiple edges). An infinite permutation of $\mathbb{N}$ is simply a linear ordering of all the elements of $\mathbb{N}$. (Instead of *infinite permutations of $\mathbb{N}$* we will often say simply *infinite permutations* in what follows.) For an arbitrary natural graph $G = (V(G), E(G))$ we will call the infinite permutations $\pi = (\pi(1), \pi(2), \dots, \pi(n), \dots)$ and $\sigma = (\sigma(1), \sigma(2), \dots, \sigma(n), \dots)$ G -different if there is at least one $i \in \mathbb{N}$ for which

$$\{\pi(i), \sigma(i)\} \in E(G).$$

*Received by the editors May 22, 2007; accepted for publication (in revised form) November 21, 2007; published electronically March 20, 2008.

<http://www.siam.org/journals/sidma/22-2/69268.html>

[†]Dipartimento di Informatica, Università di Roma “La Sapienza,” Via Salaria 113, Rome 00198, Italy (korn@di.uniroma1.it, claudia@di.uniroma1.it).

[‡]Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, H-1053 Budapest, Reáltanoda u. 13-15, Hungary (simonyi@renyi.hu). The research of this author was partially supported by the Hungarian Foundation for Scientific Research grants (OTKA) T046376, AT048826, and NK62321.

(We will use the same expression for a pair of finite sequences if at some coordinate they contain the two endpoints of an edge of G .) Let $\kappa(G)$ be the maximum cardinality of a set of infinite permutations any two elements of which are G -different. (It is easy to see that the finiteness of G implies that this number is finite as well; see Lemma 1 below.) Clearly, the value of κ is equal for isomorphic natural graphs. In this paper we will analyze this quantity for some elementary graphs and will apply some of the results to the problem of $\rho(n)$. We have been able to determine the value of $\kappa(G)$ only for some very small or simply structured graphs G . Thus, to further simplify matters, we ask questions about the extremal values of κ for graphs with a fixed number of edges (and, eventually, vertices). We define

$$(2) \quad K(\ell) = \max\{\kappa(G) ; |E(G)| = \ell\}$$

and

$$(3) \quad k(\ell) = \min\{\kappa(G) ; |E(G)| = \ell\}$$

as well as

$$(4) \quad K(v, \ell) = \max\{\kappa(G) ; |V(G)| = v, |E(G)| = \ell\}.$$

We also conjecture the following.

CONJECTURE 2. *For every $\ell \in \mathbb{N}$*

$$K(\ell) = 3^\ell.$$

In fact, we will show that $K(\ell)$ lies between 3^ℓ and 4^ℓ for every natural number ℓ . We will also see that $k(\ell)$ is linear in ℓ .

As we will explain, the values of $\kappa(P_r)$, where P_r is the r -vertex path, are relevant when investigating colliding permutations. Giving a lower bound on $\kappa(P_4)$, the lower bound of (1) will be improved to $10^{n/4-O(1)}$.

The concept of graph-different sequences from a fixed alphabet goes back to Shannon's classical paper on zero-error capacity [3]. As explained in the survey [2], a large body of problems in extremal combinatorics can be treated as zero-error problems in information theory. For the relationship of the present problems to zero-error information theory, we refer to [1].

2. Natural graphs and infinite permutations. Let G be a natural graph, and let again $\kappa(G)$ be the maximum cardinality of a set of infinite permutations any two elements of which are G -different, provided that this number is finite. It is easy to see that this is always the case. Let $\chi(G)$ denote the chromatic number of graph G .

LEMMA 1. *For every natural graph*

$$\kappa(G) \leq (\chi(G))^{|V(G)|}$$

holds.

Proof. Let us consider a proper coloring $c : V(G) \rightarrow \{1, \dots, \chi(G)\}$ of G . Let us write $v = |V(G)|$ and denote by $W = [\chi(G)]^{\mathbb{N}}$ the set of infinite sequences over the alphabet $\{1, \dots, \chi(G)\}$. Let μ be the uniform probability measure on W . Let us consider a set C of pairwise G -different permutations. We assign to any $\pi \in C$ the set $W(\pi)$ of all those sequences of W that for all $u \in V(G)$ have the element $c(u)$ in

the position where π contains u . By our hypothesis on C , the sets $W(\pi)$ are pairwise disjoint for the different elements $\pi \in C$, whence

$$1 = \mu(W) \geq \sum_{\pi \in C} \mu(W(\pi)) = \sum_{\pi \in C} \chi(G)^{-v} = |C| \chi(G)^{-v}. \quad \square$$

In the rest of this section we first investigate $K(\ell)$ and $k(\ell)$. Subsequently our new lower bound on $\rho(n)$ will be proved.

Let us denote by $S(G)$ the set of nonisolated vertices of the graph G . We introduce a graph transformation that increases the value of κ .

PROPOSITION 1. *Let F and G be two graphs with G obtained from F upon deleting an arbitrary edge in $E(F)$ followed by the addition of two new vertices to $V(F)$ so that the latter form an additional edge in G . Then*

$$\kappa(F) \leq \kappa(G).$$

Proof. Let us consider the $m = \kappa(F)$ pairwise F -different infinite permutations of an arbitrary optimal configuration for F . Let t be large enough for the initial prefixes of length t of these infinite sequences to be pairwise F -different, and let q be the largest integer appearing in their coordinates. By the finiteness of $\kappa(F)$ such t and q exist. Without restricting generality we can suppose that the new edge of G is $\{c, d\}$ with both c and d being strictly larger than q . We also suppose that the edge that we will delete is $\{a, b\} \in E(F)$. Let us now suffix to each of our length- t sequences a new sequence of the same length t , where the suffix to a sequence $x_1 x_2 \dots x_t$ is obtained from it by substituting every a with c and every b with d while the remaining coordinates are defined in an arbitrary manner but in such a way that the coordinates of the overall sequence of length $2t$ are all different. Clearly, the m new sequences of length $2t$ are G -different. To conclude, we can complete these new sequences to yield infinite permutations any way we like. $\square$

A straightforward consequence of the previous proposition is the following.

COROLLARY 1. $K(\ell) = \kappa(\ell K_2)$. $\square$

Thus we know that $K(\ell)$ is achieved by ℓ independent edges. It seems equally interesting to determine which graphs achieve $k(\ell)$. At first glance one might think that $S(F) \subseteq S(G)$ implies $\kappa(F) \leq \kappa(G)$, but this is false. In particular, complete graphs do not have minimum κ among graphs with the same number of edges. Yet, determining their κ value seems an interesting problem. As we will see, the right answer for what graphs achieve $k(\ell)$ turns out to be stars, at least for ℓ not too small. Below we will study the value of κ for complete graphs, stars, and paths. In particular, path graphs will take us back to the original puzzle about colliding permutations.

PROPOSITION 2. *For the complete graph K_n on n vertices*

$$\frac{(n+1)!}{2} \leq \kappa(K_n).$$

Proof. Consider the set of even permutations of $[n+1]$, and suppose $V(K_n) = [n]$. One can observe that these permutations are K_n -different. Indeed, if two arbitrary permutations of $[n+1]$ are not K_n -different, then they differ only in positions in which for some fixed $i \in [n]$ one has $n+1$ and the other has i . Thus any of these two permutations can be obtained from the other by exchanging the positions of $n+1$ and the corresponding i . But then the two permutations have different parity, and in particular they cannot both be even. In particular, the undesired relation does not

occur between even permutations, and this gives us $\frac{(n+1)!}{2}$ permutations of $[n+1]$ that are K_n -different. Next extend each of these permutations to infinite ones by suffixing the remaining natural numbers in an arbitrary order. $\square$

PROPOSITION 3. *For the graph of ℓ independent edges we have*

$$3^\ell \leq \kappa(\ell K_2) \leq 4^\ell.$$

Proof. Notice that the graph ℓK_2 has chromatic number two, and its number of vertices is 2ℓ , whence our upper bound follows by Lemma 1.

To prove the lower bound, let us denote the edge set of our graph by $E(\ell K_2) = \{\{1, 2\}, \{3, 4\}, \dots, \{2\ell - 1, 2\ell\}\}$. Consider the set of cyclic permutations $C_1 = \{(12\star), (2\star 1), (\star 12)\}$ and for every $1 < i \leq \ell$ the sets C_i obtained from C_1 by replacing 1 with $2i - 1$ and 2 with $2i$. It is clear that for every $i \in [\ell]$ any two of the three ministrings in C_i “differ” in the edge $\{2i - 1, 2i\}$ of our graph ℓK_2 , meaning that they have somewhere in the same position the two different endpoints of this edge. But this means that the 3^ℓ strings in their Cartesian product

$$C = \times_{i=1}^{\ell} C_i$$

are pairwise ℓK_2 -different as requested. Replacing the symbol $\star$ in our strings in an arbitrary order with the different numbers from $[3\ell] - [2\ell]$, we obtain 3^ℓ permutations of $[3\ell]$ that continue to be pairwise ℓK_2 -different. The extension to infinite permutations is as always. $\square$

The only infinite class of graphs for which we are able to determine κ are the stars, i.e., the complete bipartite graphs $K_{1,r}$. We have the following result.

PROPOSITION 4. *For every r*

$$\kappa(K_{1,r}) = 2r + 1.$$

Proof. By Lemma 1 we know that $\kappa(K_{1,r}) < \infty$. Let us denote its value by m . Let us consider the vertices of $K_{1,r}$ to be the elements of $[r + 1]$, and let 1 be the “central” vertex of degree r . It is obvious that in a set of m sequences (infinite permutations) achieving the maximum we are looking for, all the sequences must have the central vertex 1 in a different position. Let us consider our m sequences as vertices of a directed graph T in which $(a, b) \in E(T)$ if the sequence corresponding to a has a $j \in \{2, \dots, r + 1\}$ in the same position where the 1 of the sequence corresponding to b is placed. Then, by definition, the directed graph T must contain a tournament, implying that

$$|E(T)| \geq \binom{m}{2}.$$

On the other hand, every $a \in V(T)$ has at most r outgoing edges. This means that

$$|E(T)| \leq mr.$$

Comparing the last two inequalities, we get

$$m \leq 2r + 1.$$

To prove a matching lower bound, consider the following set of permutations of $[2r + 1]$. For every $i \in [2r + 1]$ let us define the coordinates of the i th sequence

$x_1(i)x_2(i)\dots x_{2r+1}(i)$ by $x_i(i) = 1$ and, in general, $x_{i+j}(i) = j + 1$ for any $0 \leq j \leq r$, where all the coordinate indices are considered modulo $2r + 1$. The remaining coordinates are defined in an arbitrary manner so that the resulting sequences define permutations of $[2r + 1]$. It is easily seen that this is a valid construction. In fact, observe that for any of our sequences the “useful” symbols, those of $[r + 1]$, corresponding to the vertices of the star graph, occupy $r + 1$ “cyclically” consecutive coordinates, forming cyclic intervals. Since $2(r + 1) > 2r + 1$, these intervals are pairwise intersecting, and thus for any two of them there must be a coordinate in the intersection for which the “left end” of one of the intervals is contained in the other. The resulting permutations can be considered as prefixes of infinite permutations in the usual obvious way. $\square$

Now we are ready to return to the problem of determining $k(\ell)$, at least for large enough ℓ . The following easy lemma will be needed.

LEMMA 2. *If a finite graph F contains vertex disjoint subgraphs $F_1, \dots, F_s$, then*

$$\kappa(F) \geq \prod_{i=1}^s \kappa(F_i).$$

Proof. The proof is a straightforward generalization of the construction given in the proof of Proposition 3. Let $\hat{C}_i$ be a set of $\kappa(F_i)$ infinite sequences that are obtained from $\kappa(F_i)$ pairwise F_i -different permutations of $\mathbb{N}$ by substituting all natural numbers $i \notin V(F_i)$ by a $\star$. As the sequences in $\hat{C}_i$ contain only a finite number of elements different from $\star$, we can take some finite initial segment of all these sequences that already contains all elements of $V(F_i)$. Let C_i be the set of these finite sequences. Now consider the set

$$C := \times_{i=1}^s C_i$$

of finite sequences that each contain all vertices in $\bigcup_{i=1}^s V(F_i)$ exactly once and finitely many $\star$'s. These sequences are also pairwise F_i -different for some F_i , and thus they are pairwise F -different. By construction, their number is $\prod_{i=1}^s \kappa(F_i)$. Extending them to infinite permutations of $\mathbb{N}$, the statement is proved. $\square$

It is straightforward from the previous lemma that if $F + G$ denotes the vertex disjoint union of graphs F and G , then $\kappa(F + G) \geq \kappa(F)\kappa(G)$. We do not know any example for strict inequality here. If equality was always true, that would immediately imply Conjecture 2.

Now we use Lemma 2 to prove our main result on $k(\ell)$.

PROPOSITION 5. *Let G be a natural graph with $n := |S(G)| > 20$ and $|E(G)| = \ell$. Then*

$$\kappa(G) \geq 2\ell + 1.$$

The value of $k(\ell)$ is achieved by the graph $K_{1,\ell}$ whenever $\ell > 150$.

Proof. Let G be a graph as in the statement, and let $\nu = \nu(G)$ denote the size of a largest matching in G .

First assume that $\nu \geq n/4$. Then by Proposition 3 and the obvious monotonicity of κ we have

$$\kappa(G) \geq \kappa(\nu K_2) \geq 3^\nu.$$

Since G is simple, we have $\ell \leq \binom{n}{2}$, and thus $k(\ell) \leq \kappa(K_{1, \binom{n}{2}}) = n(n - 1) + 1$. So in this case (when $\nu \geq n/4$) it is enough to prove that

$$3^{\lceil n/4 \rceil} \geq n(n - 1) + 1$$

holds. This is true if $n > 20$.

Next assume that $3 \leq \nu < n/4$. Consider a largest matching of G consisting of edges $\{u_{2i-1}, u_{2i}\}$ with $i = 1, \dots, \nu$. The set $U := \{u_1, \dots, u_{2\nu}\}$ covers all edges of G , and thus $\ell \leq \binom{2\nu}{2} + 2\nu(n - 2\nu)$. So we have

$$k(\ell) \leq \kappa(K_{1,\ell}) \leq 2 \left[\binom{2\nu}{2} + 2\nu(n - 2\nu) \right] + 1$$

in this case. On the other hand, for each vertex $a \in S(G) \setminus U$ there is an edge $\{a, u_i\}$ for some i . We also know that if a and b are two distinct vertices in $S(G) \setminus U$ and one of them is connected to u_{2j-1} (resp., u_{2j}) for some j , then the other one cannot be connected to u_{2j} (resp., u_{2j-1}), since otherwise replacing the matching edge $\{u_{2j-1}, u_{2j}\}$ with the other two edges of the path formed by the vertices a, u_{2j-1}, u_{2j}, b would result in a larger matching, a contradiction. Choosing an edge for each $a \in S(G) \setminus U$ that connects it to a vertex in U , we can form vertex disjoint star subgraphs $K_{1,\ell_1}, \dots, K_{1,\ell_\nu}$ of G , where $\ell_i \geq 1$ for all i and $\sum_{i=1}^\nu \ell_i = n - \nu$. Then by Lemma 2 and Proposition 4 we have $\kappa(G) \geq \prod_{i=1}^\nu (2\ell_i + 1)$. The latter product is minimal (with respect to the conditions on the ℓ_i 's) if one ℓ_i , say ℓ_1 , equals to $n - 2\nu + 1$ and $\ell_2 = \dots = \ell_\nu = 1$. Thus it is enough to prove that

$$3^{\nu-1}(2(n - 2\nu + 1) + 1) > 2 \left[\binom{2\nu}{2} + 2\nu(n - 2\nu) \right] + 1,$$

as the left-hand side is a lower bound on $\kappa(G)$, while the right-hand side is an upper bound on $k(\ell)$. The latter inequality would be implied by

$$3^{\nu-1}(n - 2\nu + 1) > \binom{2\nu}{2} + 2\nu(n - 2\nu) = \nu(2n - 2\nu - 1),$$

which, in turn, is equivalent to

$$\frac{3^{\nu-1}}{\nu} > \frac{2n - 2\nu - 1}{n - 2\nu + 1}.$$

The left-hand side of this last inequality is at least 3 if $\nu \geq 3$, while the right-hand side is strictly less than 3 for $\nu \leq n/4$.

The only case not yet covered is that of $\nu < 3$. For $\nu = 1$ there is nothing to prove, since then G itself is a star. If $\nu = 2$, then let a largest matching be formed by the two edges $\{u_1, u_2\}$ and $\{u_3, u_4\}$, while once again let U denote the union of their vertices. Let $a_1, \dots, a_{n-4}$ be the rest of the nonisolated vertices of G , and note that $n - 4 > 16$. Assume that some a_i is connected to both u_1 and u_2 , yielding a triangle. Then no a_j , $j \neq i$, can be connected to either of u_1 or u_2 ; otherwise we could form a larger matching. For similar reasons, if any a_j is connected to u_3 , then no a_s , $s \neq j$, can be connected to u_4 . (If some a_i forms a triangle with u_1, u_2 and some a_j with u_3 and u_4 , then the remaining vertices a_s must be isolated, implying $n \leq 6$, a contradiction.) Thus if a_i is connected to both u_1 and u_2 , then the rest of the a_j 's form a star centered at either u_3 or u_4 . Thus in this case, using again Lemma 2, Propositions 2 and 4 imply $\kappa(G) \geq 12[2(n - 4) + 1] = 24n - 84$. The foregoing also implies $\ell \leq n + 4$, thus $k(\ell) \leq 2n + 9 < 24n - 84$, whenever $n \geq 5$. Clearly, the situation is similar if we exchange the role of the two matching edges.

Assuming that no triangle is formed, we can again attach each vertex in $S(G) \setminus U$ to one of the edges $\{u_1, u_2\}$ and $\{u_3, u_4\}$, whichever it is connected to. Two vertex disjoint stars can be formed this way, establishing the lower bound $\kappa(G) \geq$

$3(2(n-3)+1) = 6n-15$. For the number of edges we now get $\ell \leq 6+2(n-4) = 2n-2$ since the graph induces at most 6 edges on U . Thus we have $k(\ell) \leq 4n-3$, which is less than $6n-15$ if $n > 6$. This completes the proof of the first statement.

If a simple graph has at most 20 vertices, then its number of edges is at most 190, so the second statement immediately follows from the first one if $\ell > 190$. If the graph contains a K_6 subgraph, then by Proposition 2 we have $\kappa(G) \geq 7!/2 > 381 = \kappa(K_{1,190}) \geq k(\ell)$ if $\ell \leq 190$. Thus we may assume $K_6 \not\subseteq G$, and this implies by Turán's theorem that $\ell \leq 160$ if $n \leq 20$. But $\kappa(K_{1,160}) = 321 \leq 6!/2$, so if the conclusion is not true, we may also assume that G has no K_5 subgraph. Applying Turán's theorem again, this gives $\ell \leq 150$ for $n \leq 20$. Thus the statement is true whenever $\ell > 150$. $\square$

Remark 1. We believe that the statement of Proposition 5 holds without any restriction on n or ℓ . Some improvement on our threshold on ℓ is easy to obtain. It seems to us, however, that proving the statement in full generality either leads to tedious case checkings or needs some new ideas.

The problem of determining κ seems interesting in itself; moreover, it helps to obtain better bounds for the original question on colliding permutations. To explain this, we introduce a notion connecting the two questions. Let $\kappa(G, n)$ be the maximum number of pairwise G -different permutations of $[n]$. Clearly,

$$(5) \quad \kappa(G) = \sup_n \kappa(G, n).$$

Notice that by the finiteness of $\kappa(G)$ the supremum above is always attained, so we could write maximum instead. Further, for the graph P_r , the path on r vertices, we have the following.

LEMMA 3. *For every $n > m > r$ the function ρ satisfies the recursion*

$$\rho(n) \geq \kappa(P_r, m)\rho(n-r).$$

Proof. We will call two arbitrary sequences of integers *colliding* if they have the same length and if somewhere in the same position they feature integers differing by 1. By the definition of $\kappa(P_r, m)$ we can construct this many sequences of length m such that in each of them every vertex of P_r appears exactly once, the other positions are occupied by the “dummy” symbol $\star$, and moreover these sequences are pairwise P_r -different. The latter implies that these sequences are pairwise colliding. Furthermore, we have, also by definition, $\rho(n-r)$ permutations of $[n-r]$ that are pairwise colliding. Let us “shift” these permutations by adding r to all of their coordinates. The new set of permutations of the set $r + [n-r] = [r+1, n]$ maintains the property that its elements are pairwise colliding. Next we execute our basic operation of “substituting” the permutations of the second set into those coordinates of any sequence $\mathbf{x}$ from the first set where the sequence $\mathbf{x}$ has a star. More precisely, consider any sequence $\mathbf{x} = x_1x_2 \dots x_m$ from our first set, and let $S(\mathbf{x}) \in \binom{[m]}{m-r}$ be the set of those coordinates which are occupied by stars. Let further $\mathbf{y} = y_1y_2 \dots y_{n-r}$ be an arbitrary sequence from our second set, i.e., a permutation of $[r+1, n]$. The sequence $\mathbf{z} = \mathbf{y} \rightarrow \mathbf{x}$ is a sequence of length n in which the first m coordinates are defined in the following manner. We have the equality $z_i = x_i$, if $i \leq m$ and $i \notin S(\mathbf{x})$. Suppose further that $S(\mathbf{x}) = \{j_1, j_2, \dots, j_{m-r}\}$. In the j_k th position we replace the symbol $\star$ by y_k . (For $i > m$ we set $z_i = y_{i-r}$.) Clearly, the resulting sequence is a permutation of $[n]$. Further, the so obtained $\kappa(P_r, m)\rho(n-r)$ permutations are pairwise colliding. $\square$

Observe next the following equality.

LEMMA 4.

$$\kappa(P_4, 5) = 10.$$

Remark 2. The existence of 10 permutations of $\{1, \dots, 5\}$ with the requested properties is implicit in [1] since the construction of the 35 colliding permutations of $\{1, \dots, 7\}$ in that paper does contain such a set in some appropriate projection of its coordinates.

Proof. Let us consider the 10 permutations of $\{1, \dots, 5\}$ obtainable by considering the cyclic configurations of $(1, 2)$, $(4, 3)$ and the single element 5. We indeed have 10 different permutations by “cutting” in all the five possible ways both of the two cyclic configurations that three building blocks can define. (So these are 12435, 24351, 43512, 35124, 51243 and similarly the five cyclic shifts of the sequence 43125.) Let us further consider the graph P_4 (or, in fact, $P_4 + K_1$) with vertex set $\{1, \dots, 5\}$ and with edge set $\{\{1, 2\}, \{2, 3\}, \{3, 4\}\}$. In other words, consecutive numbers are adjacent vertices, but 5 is isolated. It is easy to check that the 10 sequences above are P_4 -different for the natural graph we defined.

To see that 10 is an upper bound it is enough to observe that the two even elements of $\{1, \dots, 5\}$ cannot be placed in the same two positions in two permutations belonging to a set of P_4 -different permutations of $\{1, \dots, 5\}$. $\square$

The above construction gives the following improved lower bound for the exponential asymptotics of $\rho(n)$.

PROPOSITION 6.

$$\lim_{n \rightarrow \infty} \rho^{\frac{1}{n}}(n) \geq 10^{\frac{1}{4}}.$$

Proof. A simple combination of our two preceding lemmas implies

$$\rho(n) \geq 10\rho(n-4).$$

An iterated application of this inequality gives the desired result. $\square$

To close this section, let us take another look at Lemma 4. We believe that, in fact, $\kappa(P_4) = \kappa(P_4, 5) = 10$ and more generally,

$$\kappa(P_v) = \kappa(P_v, v+1) = \binom{v+1}{\lfloor \frac{v+1}{2} \rfloor}$$

for even values of v . The original conjecture (see Conjecture 1) for $\rho(v)$ would be an immediate consequence of this conjecture. To see this, suppose first that v is even. Then $\rho(v+1) = \kappa(P_{v+1}, v+1) \geq \kappa(P_v, v+1)$, and this would imply Conjecture 1 for odd values of n right away. Now, since for even n

$$\rho(n) = \kappa(P_n, n) \geq 2\kappa(P_{n-2}, n-1)$$

and likewise,

$$\binom{n}{\frac{n}{2}} = 2 \binom{n-1}{\lfloor \frac{n-1}{2} \rfloor},$$

the last two relations would lead us to settle the conjecture for n even. (The inequality above follows by putting an n to the end of each sequence in an optimal construction

for $\kappa(P_{n-2}, n-1)$ and then doubling each sequence by considering also its variant, which one obtains by exchanging in it $n-1$ and n .)

We also believe that

$$K(v, v-1) = \kappa(P_v).$$

As a combination of the two conjectures above we would arrive at the next statement.

CONJECTURE 3. *For every even $v \in \mathbb{N}$*

$$K(v, v-1) = \binom{v+1}{\lfloor \frac{v+1}{2} \rfloor}.$$

3. Related problems.

3.1. A graph covering problem. We show that determining $K(\ell)$ is equivalent to a graph covering problem introduced below. The following standard definition is needed.

DEFINITION 1. *The (undirected) line graph $L(D)$ of the directed graph $D = (V, A)$ is defined by*

$$\begin{aligned} V(L(D)) &= A, \\ E(L(D)) &= \{(a, b), (c, d)\} : b = c \text{ or } a = d\}. \end{aligned}$$

Let $\mathcal{L}$ denote the family of all finite simple graphs that are isomorphic to the line graph of some directed graph with possibly multiple edges.

Let the minimum number of graphs in $\mathcal{L}$, the edge sets of which together can cover the edges of the complete graph K_n , be denoted by $h(n)$.

PROPOSITION 7. *For any $M \in \mathbb{N}$, the minimum number ℓ for which $K(\ell) \geq M$ is equal to $h(M)$.*

Proof. Consider a construction attaining $K(\ell)$, that is, a graph G with ℓ edges and $K(\ell)$ infinite permutations that are G -different. Let this set of permutations be denoted by W , and let $\{a, b\}$ be one of the edges of G . Define a graph T_{a-b} on W as its vertex set where an edge is put between two permutations if and only if there is a position where one of them has a while the other has b . In other words, the two permutations are G -different by the edge $\{a, b\}$. Consider the graphs T_{a-b} for all edges of G . These all have the same vertex set, while the union of their edge sets clearly covers the complete graph K_M , where $M = K(\ell)$.

Next we show that all the graphs T_{a-b} belong to $\mathcal{L}$. To this end fix an edge $\{a, b\} \in V(G)$ and consider a graph D_{a-b} with its vertex set $V(D_{a-b})$ consisting of those positions where any of the permutations in W have a nonisolated vertex of G . Since G and W are finite, so is $V(D_{a-b})$. For each element of W we define an edge of D_{a-b} . For $\sigma \in W$, let i and j be the two positions where σ contains a and b , respectively. Then let σ be represented by the directed edge (i, j) in D_{a-b} . (If there is another permutation in W with a and b being in the same positions as in σ , then we have another arc (i, j) in D_{a-b} for this other permutation. Thus D_{a-b} is a directed multigraph.) Now it follows directly from the definitions that $T_{a-b} = L(D_{a-b})$, and thus T_{a-b} is indeed the line graph of a digraph. Together with the previous paragraph this proves $h(K(\ell)) \leq \ell$.

For the reverse inequality consider a covering of K_M with $h(M)$ graphs belonging to $\mathcal{L}$. Let the line graphs in this covering be $L_1, \dots, L_{h(M)}$. We may assume that $V(L_i) = [M]$ for all i by extending the smaller vertex sets through the addition of

isolated points. Let $D_1, \dots, D_{h(M)}$ be directed graphs satisfying $L_i = L(D_i)$ for all i . By $E(D_i) = V(L_i) = [M]$ we can consider the edges of all D_i 's labeled by $|E(D_i)|$ elements of $1, \dots, M$. (If L_i had some isolated vertices, then the corresponding labels are not used.) Using these digraphs, we define M permutations $\sigma_1, \dots, \sigma_M$ that are G -different for the graph $G = \ell K_2$ with $\ell = h(M)$. For all i define $t_i = |V(D_i)|$ and identify $V(D_i)$ with $[t_i]$. Consider D_1 . If D_1 has an edge labeled r and this edge is (i, j) , then put a 1 in position i of σ_r and put a 2 in position j of σ_r . Do similarly for all edges of D_1 . Then consider D_2 . If it has an edge labeled r which is (i', j') , then put a 3 in position $t_1 + i'$ of σ_r and put a 4 in position $t_1 + j'$ of σ_r . In general, if D_s has an edge labeled r which is (a, b) , then put a $2s - 1$ in position $(\sum_{k=1}^{s-1} t_k) + a$ and a $2s$ in position $(\sum_{k=1}^{s-1} t_k) + b$ of σ_r . When this is done for all edges of all D_i 's, then extend the obtained partial sequences to infinite permutations of $\mathbb{N}$ in an arbitrary manner. This way one obtains M permutations that are pairwise G -different. To see this consider two of these permutations, say, σ_q and σ_r . Look at the edge $\{q, r\}$ of our graph K_M that was covered by line graphs. Let L_i be the line graph that covered the edge $\{q, r\}$. Then D_i has an edge labeled q and another one labeled r in such a way that the head of the one is the tail of the other. This common point of these two edges defines a position of σ_q and σ_r where one of them has $2i - 1$ while the other has $2i$, making them G -different. $\square$

3.2. Fixed suborders. Here we consider a variant of the problem of the determination of $\kappa(G)$. We restrict attention to complete graphs.

Let $\kappa_{\text{id}}(K_n)$ denote the maximum number of infinite permutations of $\mathbb{N}$ that are K_n -different and contain the first n positive integers (the vertices of K_n) in their natural order.

PROPOSITION 8. *For every $n \in \mathbb{N}$*

$$\kappa_{\text{id}}(K_n) \geq C_n$$

holds, where $C_n = \frac{1}{n+1} \binom{2n}{n}$ is the n th Catalan number.

Proof. For $n = 1, 2$ we have equality: $\kappa_{\text{id}}(K_1) = 1$, $\kappa_{\text{id}}(K_2) = 2$. Set $a_0 = 1$ and $a_n := \kappa_{\text{id}}(K_n)$. It is enough to prove that the numbers a_n satisfy the inequality

$$a_{n+1} \geq \sum_{i=0}^n a_i a_{n-i}$$

that has the well-known recursion of Catalan numbers on its right-hand side (cf. [4]). We will look at our infinite permutations as infinite sequences consisting of infinitely many $\star$'s and one of each of the symbols $1, 2, \dots, n$, where the $\star$'s refer to all other symbols. Clearly, only the positions of the elements of $[n]$ are relevant with respect to the K_n -difference relation. Thus we will define the positions of the elements of $[n]$ and then let the $\star$'s be substituted by the other numbers in any way that will result in infinite permutations of $\mathbb{N}$.

Our construction is inductive. Assume that we already know that $a_k \geq C_k$ holds for $k \leq n$ and thus it suffices to prove it for $n + 1$. Fix a position of our permutations which is “far away,” meaning that it is far enough for having enough earlier positions for the following construction. Call this position j . For each $i = 0, \dots, n$ we construct $a_i a_{n-i}$ sequences having $i + 1$ at their position j . Any two of these sequences that have a different symbol at position j are K_n -different. For those sequences that have $i + 1$ at their position j do the following. Consider a construction of a_i pairwise K_i -different sequences consisting of symbols $1, \dots, i, \star$, where the symbols in $[i]$ are all

used somewhere in the first $j - 1$ positions (this is possible if j is chosen large enough). Take the first $j - 1$ coordinates of all these sequences, a_{n-i} times each, and continue each of them with an $i + 1$ at the j th position. So we have $a_i a_{n-i}$ sequences of length j with $i + 1$ at the j th position, each of these sequences are one of a_i possible types, and we have a_{n-i} copies from each type.

Now consider a_{n-i} sequences with the symbols $1, \dots, n - i, \star$ that are pairwise K_{n-i} -different, and shift each value in these sequences by $i + 1$. (The latter means that we change each value k to $k + i + 1$ in these sequences while $\star$'s remain $\star$'s.) For each type of the previous sequences take its a_{n-i} copies and suffix to each of them one of the current a_{n-i} different sequences. This way one gets $a_i a_{n-i}$ K_{n+1} -different sequences with symbol $i + 1$ at position j . Doing this for all $i = 1, \dots, n$, one obtains $\sum_{i=0}^n a_i a_{n-i}$ K_{n+1} -different sequences, proving the desired inequality. $\square$

Note added in proof. We learned from Graham Brightwell and Marianne Fairthorne that they have improved upon our lower bound in Proposition 6 using a similar approach and independently discovered some of the results (including Lemma 2 and Propositions 2 and 4) presented here.

Acknowledgments. The first author would like to thank Alexandr Kostochka for a stimulating discussion. Thanks also to Riccardo Silvestri for computer help.

REFERENCES

- [1] J. KÖRNER AND C. MALVENUTO, *Pairwise colliding permutations and the capacity of infinite graphs*, SIAM J. Discrete Math., 20 (2006), pp. 203–212.
- [2] J. KÖRNER AND A. ORLITSKY, *Zero-error information theory*, IEEE Trans. Inform. Theory, 44 (1998), pp. 2207–2229.
- [3] C. E. SHANNON, *The zero-error capacity of a noisy channel*, IRE Trans. Inform. Theory, 2 (1956), pp. 8–19.
- [4] R. P. STANLEY, *Enumerative Combinatorics, Vol. 2*, Cambridge Stud. Adv. Math. 62, Cambridge University Press, Cambridge, UK, 1999.

REALIZING DEGREE SEQUENCES WITH GRAPHS HAVING NOWHERE-ZERO 3-FLOWS*

RONG LUO[†], RUI XU[‡], WENAN ZANG[§], AND CUN-QUAN ZHANG[¶]

Abstract. The following open problem was proposed by Archdeacon: Characterize all graphical sequences π such that some realization of π admits a nowhere-zero 3-flow. The purpose of this paper is to resolve this problem and present a complete characterization: A graphical sequence $\pi = (d_1, d_2, \dots, d_n)$ with minimum degree at least two has a realization that admits a nowhere-zero 3-flow if and only if $\pi \neq (3^4, 2), (k, 3^k), (k^2, 3^{k-1})$, where k is an odd integer.

Key words. degree sequence, graph, integer flow, characterization

AMS subject classifications. 05C70, 05C38, 05C45

DOI. 10.1137/070687372

1. Introduction. Let $G = (V, E)$ be a graph and let k be a positive integer. An ordered pair (D, ϕ) is called a k -flow of G if $D = (V, A)$ is an orientation of G and $\phi : A \mapsto Z_k$ is an assignment of flows such that, for every vertex v ,

$$\sum_{e \in E^+(v)} \phi(e) \equiv \sum_{e \in E^-(v)} \phi(e) \pmod{k},$$

where Z_k is the set Z/kZ of integers modulo k , and $E^+(v)$ (resp., $E^-(v)$) is the set of all arcs in A with tail v (resp., head v). We say that (D, ϕ) is a *nowhere-zero flow* if $\phi(e) \neq 0$ for any $e \in A$. This concept was introduced by Tutte [19], and the theory of nowhere-zero flows provides an interesting way to generalize theorems about region-coloring planar graphs to general graphs; major open problems in this area are Tutte's celebrated 3-, 4-, and 5-flow conjectures. Interested readers are referred to Jaeger [8] and Seymour [17] for the main ideas of this subject and to Tutte [20] and Zhang [21] for in-depth accounts.

An integer-valued sequence $\pi = (d_1, d_2, \dots, d_n)$ is called *graphical* if there is a *simple* graph G so that the degree sequence of G is exactly the same as π ; such a graph G is called a *realization* of π . For simplicity, we shall also write a graphical sequence in terms of multiplicities, for instance, $(6, 4, 4, 3, 3, 3, 3) = (6, 4^2, 3^4)$. The problem of realizing degree sequences with graphs enjoying certain properties has been a subject of extensive research. Recently a surprising application of graph realization with 4-flows has been found in the design of critical partial Latin squares [5, 15], which leads to the proof [14] of the so-called *simultaneous edge-coloring conjecture*

*Received by the editors April 4, 2007; accepted for publication (in revised form) November 28, 2007; published electronically March 20, 2008.

<http://www.siam.org/journals/sidma/22-2/68737.html>

[†]Department of Mathematical Sciences, Middle Tennessee State University, Murfreesboro, TN 37132 (rluo@mtsu.edu). The work of this author was supported in part by the Summer Research Grant of Middle Tennessee State University 2006.

[‡]Department of Mathematics, University of West Georgia, Carrollton, GA 30118 (xu@westga.edu).

[§]Department of Mathematics, University of Hong Kong, Hong Kong, China (wzang@maths.hku.hk). This author was supported in part by the Research Grants Council of Hong Kong.

[¶]Department of Mathematics, West Virginia University, Morgantown, WV 26506-6310 (cqzhang@math.wvu.edu). This author was supported in part by the National Security Agency under grants MDA904-00-1-0061 and MDA904-01-1-0022.

by Keedwell [10, 11] and Cameron [2]. In this paper we study a closely related open problem proposed by Archdeacon.

PROBLEM 1.1 (see [1]). *Characterize all graphical sequences π such that some realization of π admits a nowhere-zero 3-flow.*

Our objective is to resolve this problem and to present a complete characterization.

THEOREM 1.2. *A graphical sequence $\pi = (d_1, d_2, \dots, d_n)$ with minimum degree at least two has a realization that admits a nowhere-zero 3-flow if and only if $\pi \neq (3^4, 2), (k, 3^k), (k^2, 3^{k-1})$, where k is an odd integer.*

It is worthwhile pointing out that the most striking difference between [14] and the present paper is not the flow number but the proof technique. We say that a graph H is a k -flow contractible configuration if for every graph G containing H as a subgraph, G admits a nowhere-zero k -flow if and only if so does G/H . Actually flow contractible configurations play important roles in both papers: For the 4-flow problem, every circuit of length at most four (cf. Seymour [16] and Catlin [3]) is contractible. The situation, however, becomes much more complicated for the 3-flow problem, and the digon is the only circuit that is 3-flow contractible [16]. So we have to appeal to other 3-flow contractible configurations. To be more precise, a graph $G = (V, E)$ is Z_k -connected [9] if for every $b : V \mapsto Z_k$ with $\sum_{v \in V} b(v) \equiv 0 \pmod{k}$, there exist an orientation $D = (V, A)$ of G and an assignment $\phi : A \mapsto \{1, 2, \dots, k\}$ such that for every vertex v ,

$$\sum_{e \in E^+(v)} \phi(e) - \sum_{e \in E^-(v)} \phi(e) \equiv b(v) \pmod{k}.$$

As shown in [4, 13], Z_3 -connected graphs contain even wheels, triangularly connected graphs with a family of well-described exceptions, etc.; it is Z_3 -connected graphs that will serve as contractible configurations in our proof.

The remainder of this paper is organized as follows. In section 2, we exhibit some basic properties concerning graphical sequences and Z_3 -connectivity. In section 3, we describe some graphical sequences with Z_3 -connected realizations. In section 4, we characterize certain graphical sequences π such that some realization of π admits a nowhere-zero 3-flow and contains nontrivial Z_3 -connected subgraphs. In section 5, we present a proof of the main theorem (Theorem 1.2), which fully characterizes all graphical sequences that can be realized to admit nowhere-zero 3-flows.

We remark that a complete characterization of graphic sequences with Z_3 -connected realizations remains an interesting problem for further study.

2. Preliminaries. Let $\pi = (d_1, d_2, \dots, d_n)$ be a graphical sequence with $d_1 \geq d_2 \geq \dots \geq d_n$. Throughout we reserve the symbol $\bar{\pi}$ for the sequence $(d_1 - 1, d_2 - 1, \dots, d_{d_n} - 1, d_{d_n+1}, \dots, d_{n-1})$, which is called the *residual sequence* obtained from π by laying off d_n . We shall frequently use the following well-known results in our proof.

LEMMA 2.1. *Let $\pi = (d_1, d_2, \dots, d_n)$ be a sequence. Then*

- (a) $\sum_{i=1}^n d_i$ is even if π is graphical;
- (b) (Hakimi [6, 7]; Kleitman and Wang [12]) π is graphical if and only if so is $\bar{\pi}$.

LEMMA 2.2 (Tutte [18]). *A cubic graph admits a nowhere-zero 3-flow if and only if it is bipartite.*

Lemma 2.2 can be further generalized in the following way.

LEMMA 2.3. *If a graph G admits a nowhere-zero 3-flow, then the subgraph of G induced by all degree-three vertices is bipartite.*

If H is a connected subgraph of a graph G , then G *contracted by* H , denoted by G/H , is the graph obtained from G by deleting all edges in H and then identifying $V(H)$ into a single vertex. The following simple observations follow instantly from the definition of Z_k -connectivity.

LEMMA 2.4 (Jaeger [8]; Seymour [16]). *Every circuit of length at most $k - 1$ is Z_k -connected.*

LEMMA 2.5 (DeVos, Xu, and Yu [4]; Lai, Xu, and Zhang [13]). *Let H be a Z_3 -connected subgraph of a graph G .*

- (a) *If G/H admits a nowhere-zero 3-flow, then so does G .*
- (b) *If G/H is Z_3 -connected, then so is G .*

In our proof these facts enable us to work on a reduced graph after a series of contractions of Z_3 -connected subgraphs. The reduced graphs/graphical sequences often enjoy much nicer properties than the original ones, and therefore are much easier to manipulate.

The following two lemmas are immediate corollaries of Lemmas 2.5 and 2.4.

LEMMA 2.6. *Let G be a Z_3 -connected graph, and let G' be obtained from G by adding a new vertex v and making it adjacent to at least two vertices of G . Then G' is Z_3 -connected.*

LEMMA 2.7. *Let $G = (V, E)$ be a Z_3 -connected graph. Then for any $u, v \in V$ with $uv \notin E$, the graph obtained from G by adding an edge uv is Z_3 -connected.*

Let $G = (V, E)$ be a graph, and let u, v, w be three vertices of G with $uv, uw \in E$. In this paper we shall use $G_{[uv, uw]}$ to stand for the graph $G \cup \{vw\} \setminus \{uv, uw\}$.

LEMMA 2.8. *Let $G = (V, E)$ be a graph, and let u, v, w be three vertices of G with degree $d(u) \geq 4$ and $uv, uw \in E$. If $G_{[uv, uw]}$ is Z_3 -connected, then so is G .*

A graph $G = (V, E)$ is *triangularly connected* if for every $e, f \in E$ there exists a sequence of circuits $C_1, C_2, \dots, C_k$ such that $e \in E(C_1)$, $f \in E(C_k)$, and $|E(C_i)| \leq 3$ for $1 \leq i \leq k$, and $E(C_j) \cap E(C_{j+1}) \neq \emptyset$ for $1 \leq j \leq k - 1$. A *wheel* W_k is the graph obtained from a k -circuit C by adding a vertex v and then making it adjacent to all vertices on C . By convention, we call v the *hub* and C the *rim* of W_k . We also call W_k *odd* if k is odd and *even* otherwise.

The following lemma gives sufficient conditions for a graph to be Z_3 -connected.

LEMMA 2.9 (DeVos, Xu, and Yu [4]). *Let G be a triangularly connected graph. Then G is Z_3 -connected, provided that one of the following conditions is satisfied:*

- (a) *G contains a nontrivial Z_3 -connected subgraph;*
- (b) *the minimum degree of G is at least four;*
- (c) *G is an even wheel W_k , with $k \geq 4$.*

The following lemma establishes the “only if” part of our main theorem (Theorem 1.2).

LEMMA 2.10. *Let k be an odd integer. Then no realization of the graphical sequences $(3^4, 2)$, $(k, 3^k)$, $(k^2, 3^{k-1})$ admits a nowhere-zero 3-flow.*

Proof. Observe the following:

- the only realization G_1 of $(3^4, 2)$ is the graph obtained from K_4 (the complete graph with four vertices) by subdividing an edge precisely once;
- a realization G_2 of $(k, 3^k)$ is *either* an odd wheel *or* several wheels sharing the hub, and at least one of these wheels is odd;
- the only realization G_3 of $(k^2, 3^{k-1})$ is $\frac{k-1}{2}$ copies of K_4 sharing a common edge. (To justify this, let u and v be the two vertices of maximum degree. Then the subgraph obtained from G_3 by deleting u and v is 1-regular.)

By Lemma 2.3, neither G_1 nor G_2 admits a nowhere-zero 3-flow. To prove the statement concerning sequence $(k^2, 3^{k-1})$, let Q_i be a copy of K_4 with vertex

set $\{w_i, x_i, y_i, z_i\}$ for $i = 1, 2, \dots, \frac{k-1}{2}$. As described above, G_3 is obtained from $Q_1, Q_2, \dots, Q_{\frac{k-1}{2}}$ by first identifying all y_i as a single vertex y and all z_i as a single vertex z , and then replacing all edges $y_i z_i$ with a single edge yz . Assume to the contrary that G_3 admits a nowhere-zero 3-flow (D, f) . Then there must exist a 3-flow (D, f_i) of each Q_i such that

- $f(e) = f_i(e)$ for every edge $e \in Q_i - \{yz\}$ (that is, $\text{supp}(f_i) \supseteq E(Q_i) - \{yz\}$) and
- $f = \sum_{i=1}^{\frac{k-1}{2}} f_i$, which implies the existence of a subscript i such that f_i is nowhere-zero on $Q_i (= K_4)$, contradicting Lemma 2.2. $\square$

3. Z_3 -connected realizations. The purpose of this section is to establish the following two theorems, which give some sufficient conditions for Z_3 -connected realizations.

THEOREM 3.1. *Let $\pi = (d_1, d_2, \dots, d_n)$ be a graphical sequence with $d_1 \geq d_2 \geq \dots \geq d_n$. If $d_n \geq 3$ and $d_{n-3} \geq 4$, then π has a Z_3 -connected realization.*

THEOREM 3.2. *Let $\pi = (d_1, d_2, \dots, d_n)$ be a graphical sequence with $n-1 = d_1 \geq d_2 \geq \dots \geq d_n \geq 3$. Then π has a Z_3 -connected realization if and only if $\pi \neq (k, 3^k), (k^2, 3^{k-1})$, where k is odd.*

Let us establish a weaker version of Theorem 3.1 before presenting a proof.

LEMMA 3.3. *Let $\pi = (d_1, d_2, \dots, d_n)$ be a graphical sequence with $d_1 \geq d_2 \geq \dots \geq d_n$. If $d_n \geq 3$ and $d_{n-2} \geq 4$, then π has a Z_3 -connected realization.*

Proof. Suppose the contrary: $\pi = (d_1, d_2, \dots, d_n)$ is a counterexample with smallest n . According to the configuration of π , we propose to consider four cases and construct a Z_3 -connected realization of π in each case, thereby reaching a contradiction. Notice that $n \geq 5$ as $n-1 \geq d_1 \geq d_{n-2} \geq 4$.

The lemma is to be proved step by step with the following observations and claims.

(1) $d_1 \geq 5$. Assume the contrary: $d_1 = 4$. Let us consider the following two cases.

Case 1. $d_1 = 4$ and $d_n = 4$. In this case $\pi = (4^n)$, the construction goes as follows: Let C be a circuit with n vertices $v_1, v_2, \dots, v_n$, and let $G^1 = C \cup \bigcup_{i=1}^n \{v_i v_{i+2}\}$, where $v_{n+1} = v_1$ and $v_{n+2} = v_2$. Clearly, G^1 is 4-regular and is triangularly connected. By Lemma 2.9(b), G^1 is Z_3 -connected.

Case 2. $d_1 = 4$ and $d_n = 3$. Since $d_1 = d_{n-2} = 4$ and $d_n = 3$, by Lemma 2.1(a) we have $\pi = (4^{n-2}, 3^2)$. Let G^1 be the graph constructed in Case 1, and let $G^2 = G^1 \setminus \{v_2 v_n\}$. Clearly G^2 is a realization of π . It remains to show that G^2 is Z_3 -connected.

From the construction of G^1 and G^2 , we see that $G^2_{[v_1 v_2, v_1 v_3]}$ is triangularly connected and contains a 2-circuit $v_2 v_3 v_2$. Since any 2-circuit is Z_3 -connected (by Lemma 2.4), Lemma 2.9(a) implies that $G^2_{[v_0 v_1, v_0 v_2]}$ is Z_3 -connected, and hence so is G^2 by Lemma 2.8. Therefore (1) holds.

(2) $d_2 = 4$. Suppose to the contrary that $d_2 \geq 5$. By (1), we have $d_1 \geq 5$ and $d_2 \geq 5$.

Since $n \geq 5$ and $d_{n-2} \geq 4$, we get $d_3 \geq 4$. Hence it can be seen that the residual sequence $\bar{\pi} = (\bar{d}_1, \bar{d}_2, \dots, \bar{d}_{n-1})$, with $\bar{d}_1 \geq \bar{d}_2 \geq \dots \geq \bar{d}_{n-1}$, satisfies $\bar{d}_{n-1} \geq 3$ and $\bar{d}_{n-3} \geq 4$. Thus Lemma 2.1(b) and the assumption on π guarantee the existence of a Z_3 -connected realization $\bar{G}$ of $\bar{\pi}$. We can then get a realization G^3 of π from $\bar{G}$ by adding a new vertex v and d_n edges joining v to the corresponding vertices in $\bar{G}$. By Lemma 2.6, G^3 is Z_3 -connected. This proves (2).

We claim that (3) $d_n = 3$. Otherwise, $d_n = 4$. By (1) and (2), we have $d_1 \geq 5$ and $d_2 = 4$. So $\pi = (d_1, 4, \dots, 4)$. By Lemma 2.1(a), d_1 is even. Thus $d_1 \geq 6$.

Set $k = \frac{d_1-4}{2}$. Let G^1 be the 4-regular triangularly connected graph exhibited in Case 1 of (1). For each $1 \leq i \leq k$, we subdivide the edge $v_{2i+2}v_{2i+3}$ once by a degree-two vertex u_i . Since $n-1 \geq d_1 = 2k+4$, we have $2k+3 \leq n-2$. Now let us identify all u_i with v_1 . Then the resulting graph G^4 is simple and is clearly a realization of π . To show that G^4 is Z_3 -connected, we replace the path $v_{2i+2}v_1v_{2i+3}$ with an edge $v_{2i+2}v_{2i+3}$ for all $1 \leq i \leq k$; the resulting graph is precisely G^1 . Since G^1 is Z_3 -connected, repeated applications of Lemma 2.8 imply that so is G^4 . Thus (3) follows.

By (1), (2), and (3), we have $d_1 \geq 5$, $d_2 = 4$, and $d_n = 3$. So π is either $(d_1, 4, \dots, 4, 3, 3)$ or $(d_1, 4, \dots, 4, 3)$.

If $d_{n-1} = 4$, then the residual sequence $\bar{\pi}$ satisfies the conditions of the theorem, so it admits a Z_3 -connected realization $\bar{G}$, and hence so does G . It remains to consider the case when $d_{n-1} = 3$. Since $d_2 = d_3 = \dots = d_{n-2} = 4$ and $d_{n-1} = d_n = 3$, we see that d_1 is even. So $d_1 \geq 6$. Set $k = \frac{d_1-4}{2}$. Let G^2 be the triangularly connected graph constructed in Case 2 of (1). For each $1 \leq i \leq k$, we subdivide the edge $v_{2i+2}v_{2i+3}$ once by a degree-two vertex u_i . Since $n-1 \geq d_n = 2k+4$, we have $2k+3 \leq n-2$. Now let us identify u_i and v_1 for each $1 \leq i \leq k$. Then the resulting graph G^5 is simple and is clearly a realization of π . To show that G^5 is Z_3 -connected, we replace the path $v_{2i+2}v_1v_{2i+3}$ with an edge $v_{2i+2}v_{2i+3}$ for all $1 \leq i \leq k$; then the resulting graph is precisely G^2 . Since G^2 is Z_3 -connected, by Lemma 2.8 so is G^5 , completing the proof of the lemma. $\square$

Proof of Theorem 3.1. Suppose the contrary: $\pi = (d_1, d_2, \dots, d_n)$ is a counterexample with smallest n . Notice that $n \geq 5$ as $n-1 \geq d_1 \geq d_{n-3} \geq 4$. By Lemma 3.3, we have

(1) $d_{n-2} = 3$.

Let us further make some simple observations.

(2) $d_2 = 4$. Otherwise, $d_2 \geq 5$ for $d_2 \geq d_{n-3} \geq 4$. So the residual sequence $\bar{\pi}$ satisfies the conditions of the theorem, and hence the assumption on π guarantees the existence of a Z_3 -connected realization $\bar{G}$ of $\bar{\pi}$. By Lemma 2.6, we can get a Z_3 -connected realization G of π from $\bar{G}$ by adding a new vertex v and d_n edges joining v to the corresponding vertices in $\bar{G}$. This contradiction implies (2).

Combining (1), (2), and the hypothesis of the theorem, we get

(3) $\pi = (d_1, 4^{n-4}, 3^3)$. So d_1 is odd by Lemma 2.1(a) and hence at least 5.

(4) $n \geq 8$. Suppose to the contrary that $n \leq 7$. Since $d_1 \geq 5$ and is odd by (3), we have $d_1 = 5$ and $n \geq 6$. So $\pi = (5, 4^{n-4}, 3^3)$.

For $n = 6$, let G be the graph obtained from the complete bipartite graph $K_{2,3}$ by adding a new vertex and then making it adjacent to each vertex in the $K_{2,3}$. Clearly each edge of G is contained in a wheel W_4 . Since W_4 is Z_3 -connected by Lemma 2.9(c), so is G by Lemma 2.5(b).

For $n = 7$, let G be the graph obtained from W_4 by adding two adjacent vertices v_1 and v_2 and then making v_1 adjacent to the hub of W_4 and a rim vertex and making v_2 adjacent to two other rim vertices. Since both W_4 and the graph obtained from G by contracting W_4 (which results in a triangle with parallel edges) are Z_3 -connected, so is G by Lemma 2.5(b); this contradiction establishes (4).

Let us distinguish between two cases according to the value of d_1 .

Case 1. $d_1 \geq n-3$. Set $k = \frac{d_1-5}{2} \geq 0$ and $m = \lfloor \frac{n-4}{2} \rfloor$. Take a wheel W_{n-4} with hub w and rim $u_1u_2u_3 \dots u_{n-4}u_1$. Let H be the graph obtained from this wheel by adding k edges u_iu_{i+m} for $i = 1, 2, \dots, k$. Then the degree sequence of H is $(n-4, 4^{2k}, 3^{n-4-2k}) = (n-4, 4^{d_1-5}, 3^{n-d_1+1})$. To get a graph G with n vertices and degree sequence $\pi = (d_1, 4^{n-4}, 3^3)$, we need to add three vertices and

$[d_1 + 4(n - 4) + 9 - (n - 4) - 4(d_1 - 5) - 3(n - d_1 + 1)]/2 = 7$ edges to H . The construction of G goes as follows: We first add a path $P = v_1v_2v_3$ to H , then we connect w and $d_1 - (n - 4)$ vertices on P , and finally add precisely one edge between each of $n - d_1 + 1$ degree-three vertices on H and P , so that there are precisely two edges between each of v_1 and v_3 and H , and there is precisely one edge between v_2 and H .

By (4), $n \geq 8$. Note that if $n = 8$, then $H = W_4$; if $n \geq 9$, then at least one edge is added to W_{n-4} , which implies that H contains an even wheel. So, by Lemma 2.9(c), H contains a Z_3 -connected subgraph (an even wheel) in either case. Clearly, G/H is triangularly connected and contains a 2-circuit. Since a 2-circuit is Z_3 -connected, so is G/H by Lemma 2.9(a). It follows from Lemma 2.5(b) that G is Z_3 -connected, a contradiction.

Case 2. $d_1 \leq n - 4$. By (3), $d_1 \geq 5$. So in this case $n \geq d_1 + 4 \geq 9$. Let us consider the sequence $\sigma = (d_1 - 1, 4^{n-7}, 3^2)$. From the construction of G^2 and G^5 of the proof of Lemma 3.3, we deduce that σ has a Z_3 -connected realization H ; let u_1, u_2, u_3 denote the vertices of H with degree three and degree $d_1 - 1$, respectively. Let G be the graph obtained from H by first adding a complete graph with four vertices v_1, v_2, v_3, v_4 , then deleting edge v_1v_3 , and finally adding a matching of size three between $\{u_1, u_2, u_3\}$ and $\{v_1, v_2, v_3\}$. Clearly, G is a realization of π . Note that G/H is a wheel W_4 with hub v_2 , so by Lemma 2.9(c) it is Z_3 -connected, and hence so is G by Lemma 2.5(b). This contradiction completes the proof of the theorem. $\square$

The proof of Theorem 3.2 is based on the following lemma.

LEMMA 3.4. *Let $\pi = (d_1, d_2, \dots, d_n)$ be a graphical sequence with $d_1 \geq d_2 \geq \dots \geq d_n \geq 2$. Then π has a connected realization G that contains an even circuit if and only if $\pi \neq (2^n), (n - 1, 2^{n-1})$, where n is odd.*

Proof. It is easy to see that when n is odd

- the unique connected realization of $\pi = (2^n)$ is an odd circuit;
- the unique connected realization of $\pi = (n - 1, 2^{n-1})$ is $\frac{n-1}{2}$ triangles sharing a common vertex.

Clearly neither of these two graphs contains an even circuit, so the “only if” part is established.

Let us proceed to the “if” part. Assume that $\pi \neq (2^n), (n - 1, 2^{n-1})$, where n is odd, but no connected realization of π contains an even circuit. We further assume that π is chosen with minimum n . Let us make some simple observations.

(1) $n \geq 5$. By the assumption on π , we have $n \neq 3$. Hence $n \geq 4$. If $n = 4$, then $\pi = (2^4)$ or $(3^2, 2^2)$ or (3^4) . In each case π has a connected realization that contains an even circuit. This contradiction yields (1).

(2) $d_n = 2$. Otherwise, $d_n \geq 3$. By (1), the residual sequence $\bar{\pi}$ satisfies the condition of the lemma. Hence it admits a connected realization H that contains an even circuit by the assumption on π . We can then get a desired realization G of π from H by adding a new vertex and making it adjacent to corresponding vertices in H , a contradiction. So (2) holds.

(3) $d_2 \geq 3$. Otherwise, by (2) we have $d_2 = 2$, and so $\pi = (d_1, 2^{n-1})$, where $d_1 = 2k$ for some integer $k \geq 1$. Let H be the graph obtained from k disjoint triangles and then gluing them at a common vertex v . Note that the number of vertices in H is $2k + 1$. According to the assumption on π , we have $n \neq 2k + 1$ and $k \geq 2$. So $2k + 1 \leq n - 1$. Let G be the graph obtained from H by inserting a degree-two vertex into the edge not containing v in the first triangle, and inserting the remaining degree-two vertices (if any) into the edge not containing v in the second triangle. Then G is a connected realization of π that contains an even circuit (of length four), a contradiction. So (3) is proved.

Consider the residual sequence $\bar{\pi}$ of π . By (3), we have $d_1 - 1 \geq d_2 - 1 \geq 2$. If $\bar{\pi}$ satisfies the conditions of the lemma, then the assumption on π guarantees a connected realization H of $\bar{\pi}$ that contains an even circuit. From H we can obviously get a desired realization of π . This contradiction implies that $\bar{\pi} = (2^{n-1})$ or $(n-2, 2^{n-2})$, where $n-1$ is odd and thus n is even.

If $\bar{\pi} = (2^{n-1})$, then $\pi = (3^2, 2^{n-2})$ by (2). We can get a desired realization of π from an n -circuit by adding a chord.

If $\bar{\pi} = (n-2, 2^{n-2})$, then the unique connected realization H of $\bar{\pi}$ is $\frac{n-2}{2}$ triangles sharing a common vertex v . Note that $\pi = (n-1, 3, 2^{n-2})$ by (2). Let G be the graph obtained from H by adding a new vertex and making it adjacent to v and one other vertex. Clearly, G is a connected realization of π and contains an even circuit. This contradiction completes the proof. $\square$

Proof of Theorem 3.2. The “only if” part follows instantly from Lemma 2.10. It remains to show the “if” part.

Consider the sequence $\sigma = (d_2 - 1, d_3 - 1, \dots, d_n - 1)$. Note that $\sigma \neq (2^{n-1})$, $(n-2, 2^{n-2})$, where $n-1$ is odd, for otherwise $\pi = (n-1, 3^{n-1})$ or $((n-1)^2, 3^{n-2})$, contradicting the hypothesis on π . By Lemma 3.4, σ has a connected realization H that contains an even circuit. Let G be the graph obtained from H by adding a new vertex and making it adjacent to each vertex of H . Clearly G is a realization of π . Since G is triangularly connected and contains an even wheel, from Lemma 2.9 we deduce that G is Z_3 -connected. $\square$

4. Partially Z_3 -connected realizations. We propose to establish the following two theorems in this section.

THEOREM 4.1. *Let $\pi = (d_1, d_2, \dots, d_n)$ be a graphical sequence with $d_1 \geq d_2 \geq \dots \geq d_n \geq 3$ and $d_3 \geq 5$. Then π has a realization G such that*

- (a) G admits a nowhere-zero 3-flow; and
- (b) G has a Z_3 -connected subgraph H that contains all vertices of G with degree at least four.

THEOREM 4.2. *Let $\pi = (d_1, d_2, 4^{n-k-2}, 3^k)$ be a graphical sequence with $n-2 \geq d_1 \geq d_2 \geq 4$, $d_1 + d_2 \geq 11$, $n-3 \geq k \geq 4$, and $n \geq 9$. Then π has a realization G such that*

- (a) G admits a nowhere-zero 3-flow; and
- (b) G has a Z_3 -connected subgraph H that contains all vertices of G with degree at least four.

Let us introduce three operations before proving these theorems, which will be used frequently in our proofs.

Let H_1 and H_2 be two disjoint graphs. A graph G obtained by adding H_2 onto H_1 via Operation A, B, or C is defined below.

Operation A. Let $u_i v_i$ for $i = 1, 2, \dots, k$ be k edges of H_2 . The graph G is obtained from the union of H_1 and H_2 by first cutting each $u_i v_i$ into two edges $u_i x_i$ and $y_i v_i$ and then identifying each of x_i and y_i with a vertex of H_1 .

Operation B. Let $u_i v_i$ for $i = 1, 2, \dots, k$ be k edges of H_2 . The graph G is obtained from the union of H_1 and H_2 by inserting a degree-two vertex x_i into each $u_i v_i$ and then identifying each x_i with a vertex in H_1 .

Operation C. Let u be a vertex in H_2 with $d(u) = k \geq 2$, and let $u_1, u_2, \dots, u_t$ be t neighbors of u . The graph G is obtained from the union of H_1 and H_2 by splitting u into $t+1$ vertices $u'_1, u'_2, \dots, u'_t, u'$ such that u_i is the only neighbor of u'_i and that $d(u') = d(u) - t$, and then identifying each of those $t+1$ vertices with a vertex in H_1 .

LEMMA 4.3. *Let H_1 be a Z_3 -connected graph, and let H_2 be a cubic bipartite graph with at least four vertices. (H_2 is simple if it has at least six vertices or contains precisely two 2-circuits otherwise.) Then the graph G obtained by adding H_2 onto H_1 via Operation A, B, and/or C admits a nowhere-zero 3-flow.*

Remark. Obviously, the new graph G obtained via Operation C is simple as long as H_1 and H_2 are simple. Let H_3 be the graph obtained from H_2 by cutting all $u_i v_i$. Then the new graph G obtained via Operation A is simple if, first, H_3 and H_1 are simple; second, the edges in H_3 joining the same vertex of H_1 form a matching in H_3 . If G is obtained via Operation B, then G is simple if both H_1 and the graph obtained from H_2 by inserting a new degree-two vertex into each edge $u_i v_i$ are simple and the edges $u_i v_i$ form a matching in H_2 .

Proof. Note that H_1 remains intact in the new graph G , so it is still Z_3 -connected (as a subgraph of G). By Lemma 2.5(a), G admits a nowhere-zero 3-flow if and only if G/H_1 admits a nowhere-zero 3-flow.

Since H_2 is a cubic bipartite graph, by Lemma 2.2, it admits a nowhere-zero 3-flow. Furthermore, $G/H_1 = H_2$ if only Operation C is applied, and if Operation A or B is applied, then G/H_1 can be obtained from H_2 by subdividing some edges once and then identifying the new degree-two vertices as one vertex. Therefore, G/H_1 also admits a nowhere-zero 3-flow. $\square$

LEMMA 4.4. *Let $\pi = (d_1, d_2, 4^{n-4}, 3^2)$ be a sequence with $n - 1 \geq d_1 \geq d_2 \geq 4$ and $n \geq 5$. Then π is graphical, provided that $d_1 + d_2$ is even.*

Proof. Assume the contrary: π is a counterexample with minimum n . Then $n \geq 6$, for otherwise $n = 5$, so $\pi = (4^3, 3^2)$, and thus the graph obtained from K_5 by deleting one edge is a realization of π , a contradiction.

If $d_2 \geq 5$, then the residual sequence $\bar{\pi} = (d_1 - 1, d_2 - 1, 4^{n-5}, 3^2)$. From the assumption on π , we see that $\bar{\pi}$ is graphical and hence so is π , by Lemma 2.1(b); this contradiction yields $d_2 = 4$. Therefore $\pi = (d_1, 4^{n-3}, 3^2)$.

Since $d_1 + d_2 = d_1 + 4$ is even, so is d_1 . It follows that the graph G^2 (resp., G^5) in the proof of Lemma 3.3 is a realization of π if $d_1 = 4$ (resp., $d_1 \geq 6$). $\square$

LEMMA 4.5. *Let $\pi = (d_1, d_2, 5, 4^{n-3-k}, 3^k)$ be a graphical sequence with $n - 2 \geq d_1 \geq d_2 \geq 5$, $n - 3 \geq k \geq 5$, and $n \geq 9$. Then π has a realization G such that*

- (a) G admits a nowhere-zero 3-flow;
- (b) G has a Z_3 -connected subgraph H that contains all vertices of G with degree at least four.

Proof. Let us distinguish between two cases according to the parity of k .

Case 1. k is even. Since $k \geq 5$ and since the degree sum of π is even, in this case we have

- (1) $k \geq 6$, and $d_1 + d_2$ is odd.

We propose to construct a realization G of π with properties (a) and (b) using Lemma 4.3, such that the Z_3 -connected graph H_1 (recall Lemma 4.3) has degree sequence $\pi^* = (d_1^*, d_2^*, 4^{n-2-k}, 3^2)$, where d_1^* and d_2^* are to be determined, and the cubic bipartite graph H_2 has $k - 2$ vertices.

We are to determine d_1^* and d_2^* by letting $A_i = \max\{4, d_i - (k - 2)\}$ and $B_i = \min\{n - k + 1, d_i - 1\}$ for $i = 1, 2$. Then

- (2) $A_i \leq B_i$, and equality holds if and only if $A_i = B_i = 4$. To verify this, note that $4 \leq n - k + 1$ (because $k \leq n - 3$), $d_i - (k - 2) < d_i - 1$ (because $6 \leq k$), $4 \leq d_i - 1$ (because $5 \leq d_i$), and $d_i - (k - 2) < n - k + 1$ (because $d_i \leq n - 2$). Combining these inequalities yields $A_i \leq B_i$ and $d_i - (k - 2) < B_i$. It follows that $A_i = B_i$ if and only if both of them are four. So (2) is true.

Clearly (2) guarantees the existence of d_1^* and d_2^* such that

- $A_i \leq d_i^* \leq B_i$ for $i = 1, 2$;
- $d_1^* + d_2^*$ is even.

By Lemma 4.4, the sequence $\pi^* = (d_1^*, d_2^*, 4^{n-2-k}, 3^2)$ is graphical, and hence by Theorem 3.1 it admits a Z_3 -connected realization H_1 . By the definitions of A_i , B_i , and d_i^* , we have $d_i - (k-2) \leq d_i^* \leq d_i - 1$ for $i = 1, 2$, so

$$(3) \quad 1 \leq d_i - d_i^* \leq k - 2.$$

It follows that $2 \leq (d_1 + d_2) - (d_1^* + d_2^*) \leq 2(k-2)$. Since $d_1 + d_2$ is odd (by (1)) while $d_1^* + d_2^*$ is even, the above inequalities can be strengthened as $3 \leq (d_1 + d_2) - (d_1^* + d_2^*) \leq 2(k-2) - 1$, or equivalently

$$(4) \quad 2 \leq \frac{(d_1 + d_2) - (d_1^* + d_2^*) + 1}{2} \leq k - 2.$$

Using Lemma 4.3, a realization of π can be constructed as follows. Let H_2 be a cubic bipartite graph with $k-2$ (≥ 4) vertices, where H_2 is simple if $k \geq 8$ and contains precisely two disjoint 2-circuits, C_1 and C_2 , if $k = 6$. Then there exist two disjoint perfect matchings M_1, M_2 in H_2 such that $M_i \cap C_i \neq \emptyset$ for $i = 1, 2$ if $k = 6$. Without loss of generality, we assume that $d_1 - d_1^*$ is even. So $d_2 - d_2^*$ is odd. By (3), (4) and the selection of M_i , we can find a subset F_i of M_i such that $|F_1| = \frac{d_1 - d_1^*}{2}$, $|F_2| = \frac{d_2 - d_2^* + 1}{2}$, and $(F_1 \cup F_2) \cap C_i \neq \emptyset$ for $i = 1, 2$ if $k = 6$. Let v_i be the vertex with degree d_i^* in H_1 for $i = 1, 2$, and let ab be a special edge in F_2 . For $i = 1, 2$, let us subdivide each edge in F_i once by a degree-two vertex (let c denote this vertex on the special edge ab), then identifying all these degree-two vertices with v_i . At this stage, the resulting graph has the degree sequence $(d_1, d_2 + 1, 4^{n-2-k}, 3^k)$. Finally, switch the edge ca away from vertex v_2 to a degree-four vertex in $H_1 \setminus \{v_1, v_2\}$. Clearly, the resulting graph G is simple and is a desired realization of the sequence π .

Case 2. k is odd. In this case, we have

- (5) $d_1 + d_2$ is even. According to the value of k , we consider two possibilities.

Subcase 2.1. $k \leq n - 4$. The proof of this subcase goes along the same line as that of Case 1: We propose to construct a realization G of π with properties (a) and (b), using Lemma 4.3, such that the Z_3 -connected graph H_1 (recall Lemma 4.3) has degree sequence $\pi^* = (d_1^*, d_2^*, 4^{n-3-k}, 3^2)$, where d_1^* and d_2^* are to be determined, and the cubic bipartite graph H_2 has $k-1$ vertices.

We are to determine d_1^* and d_2^* by letting $A_i = \max\{4, d_i - (k-1)\}$ and $B_i = \min\{n-k, d_i - 1\}$ for $i = 1, 2$. It is a routine matter to check that

- (6) $A_i \leq B_i$, and equality holds if and only if $A_i = B_i = 4$.

Thus (6) guarantees the existence of d_1^* and d_2^* such that

- $A_i \leq d_i^* \leq B_i$ for $i = 1, 2$;
- $d_1^* + d_2^*$ is even.

By Lemma 4.4, the sequence $\pi^* = (d_1^*, d_2^*, 4^{n-3-k}, 3^2)$ is graphical, and hence by Theorem 3.1 it admits a Z_3 -connected realization H_1 . By the definitions of A_i , B_i , and d_i^* , we have $d_i - (k-1) \leq d_i^* \leq d_i - 1$ for $i = 1, 2$, so

$$(7) \quad 1 \leq d_i - d_i^* \leq k - 1.$$

Let H_2 be a cubic bipartite graph with $k-1$ (≥ 4) vertices, where H_2 is simple if $k \geq 7$, or contains precisely two disjoint 2-circuits, C_1 and C_2 , if $k = 5$. Renaming the subscripts if necessary, we assume that $d_1 - d_1^* \geq d_2 - d_2^*$. Set $t_1 = (d_1 - d_1^*)/2$ and $t_2 = (d_2 - d_2^*)/2$ if $d_1 - d_1^*$ is even, or, set $t_1 = (d_1 - d_1^* + 1)/2$ and $t_2 = (d_2 - d_2^* - 1)/2$ otherwise. Since $d_1 + d_2$ and $d_1^* + d_2^*$ have the same parity, t_1 and t_2 are both integers. Moreover, by (7) we have

- (8) $t_2 \leq t_1 \leq (k-1)/2$, and $t_2 < t_1$ if $d_1 - d_1^*$ is odd.

Let M_1, M_2, M_3 be three disjoint perfect matchings in H_2 . In view of (8), we can find a subset F_i of M_i such that $|F_i| = t_i$ for $i = 1, 2$, $|F_3| = 1$, and $(F_1 \cup F_2 \cup F_3) \cap C_i \neq \emptyset$ for $i = 1, 2$ if $k = 5$. Let v_i be the vertex with degree d_i^* in H_1 for $i = 1, 2$, let ab be a special edge in F_1 such that a is covered by no edge in F_2 if $d_1 - d_1^*$ is odd (such edge is available as $t_2 < t_1$), and let cd be the edge in F_3 . For $i = 1, 2$, let us subdivide each edge in F_i once by a degree-two vertex (let x denote this vertex on the special edge ab) and identify all these degree-two vertices with v_i . At this stage, the resulting graph has the degree sequence $(d_1, d_2, 4^{n-3-k}, 3^{k+1})$ if $d_1 - d_1^*$ is even, or $(d_1 + 1, d_2 - 1, 4^{n-3-k}, 3^{k+1})$ if $d_1 - d_1^*$ is odd. Then switch edge xa away from vertex v_1 to v_2 if $d_1 - d_1^*$ is odd, and finally cut cd ($\in F_3$) into two edges cx and yd and identify x (resp., y) with a degree-four (resp., degree-three) vertex in $H_1 \setminus \{v_1, v_2\}$. Clearly, the resulting graph G is simple and is a desired realization of the sequence π .

Subcase 2.2. $k = n - 3$. In this subcase, by the hypothesis of the theorem we have (9) $\pi = (d_1, d_2, 5, 3^k)$, where $n - 2 \geq d_1 \geq d_2 \geq 5$ and $n \geq 9$.

According to the hypothesis of Case 2, k is odd. From $k = n - 3$ and $n \geq 9$ we deduce that

$$(10) \quad k \geq 7.$$

By Lemma 4.4, the sequence $(4^3, 3^2)$ is graphical and hence, by Theorem 3.1, it admits a Z_3 -connected realization H_1 ; let x_1, x_2, x_3 be the three vertices of degree four in H_1 . Since $k \geq 7$ by (10), we can find a simple cubic bipartite graph H_2 with $k - 1$ vertices. Let u be a vertex of H_2 , let v_1, v_2, v_3 be the neighbors of u in H_2 , and let M_1, M_2, M_3 be three disjoint perfect matchings of H_2 . Renaming the subscripts if necessary, we assume $uv_i \in M_i$ for $i = 1, 2, 3$. Let G^* be the graph obtained from the union of H_1 and H_2 by first splitting u into three vertices $\{u_1, u_2, u_3\}$ (so the three edges incident with u in H_2 become u_1v_1, u_2v_2, u_3v_3) and then identifying u_i with x_i in H_1 for $i = 1, 2, 3$. At this stage, the resulting graph has the degree sequence $(5^3, 3^k)$.

Set $t_1 = (d_1 - 5)/2$ and $t_2 = (d_2 - 5)/2$ if d_1 is odd, and set $t_1 = (d_1 - 4)/2$ and $t_2 = (d_2 - 6)/2$ otherwise. Since $d_1 + d_2$ is even by (5), t_1 and t_2 are both integers. Moreover, since $k = n - 3$ and $n - 2 \geq d_1 \geq d_2 \geq 5$, we have

$$(11) \quad t_2 \leq t_1 \leq (k - 3)/2, \text{ and } t_2 < t_1 \text{ if } d_1 \text{ is even.}$$

Since $M_i - \{uv_i\}$ contains $(k - 3)/2$ edges, we can find a subset F_i of $M_i - \{uv_i\}$ such that $|F_i| = t_i$ for $i = 1, 2$. Let ab be a special edge in F_1 such that a is covered by no edge in F_2 if d_1 is even (such an edge is available as $t_2 < t_1$). We construct a graph from G^* as follows: For $i = 1, 2$, subdivide each edge in F_i once by a degree-two vertex (let c denote this vertex on the special edge ab), then identify all these degree-two vertices with x_i , and finally switch edge ca away from vertex x_1 to x_2 if d_1 is even. Clearly, the resulting graph G is simple and is a desired realization of the sequence π . $\square$

Now we are ready to establish the main results of this section.

Proof of Theorem 4.1. Assume the contrary: π is a counterexample with minimum n . By Theorem 3.1, we have the following:

- (1) $d_n = 3$.
- (2) The residual sequence $\bar{\pi}$ does not satisfy the hypothesis of the theorem.

Otherwise, from the assumption on π we deduce that $\bar{\pi}$ has a realization $\bar{G}$ such that

- $\bar{G}$ admits a nowhere-zero 3-flow;
- $\bar{G}$ has a Z_3 -connected subgraph $\bar{H}$ that contains all vertices of $\bar{G}$ with degree at least four.

Let G be the realization of π obtained from $\bar{G}$ by adding a new vertex v and three edges between v and corresponding vertices in $\bar{G}$ (recall (1)). Let H be the subgraph induced by $V(\bar{H}) \cup \{v\}$ in G . Since $\bar{H}$ contains all vertices of $\bar{G}$ with degree at least four, it also contains all vertices of G with degree at least four as $d_3 \geq 5$ and $d_n = 3$. So the degree of v in H is also three. Hence H is Z_3 -connected by Lemma 2.6. Note that the existence of a nowhere-zero 3-flow is preserved under edge contractions, so $\bar{G}/\bar{H}$ and hence G/H (as $\bar{G}/\bar{H} = G/H$) admits a nowhere-zero 3-flow. It follows from Lemma 2.5(a) that so does G . This contradiction implies (2).

(3) $d_1 \leq n - 2$. Otherwise, $d_1 = n - 1$. Since $d_3 \geq 5$, π has a Z_3 -connected realization by Theorem 3.2, a contradiction. So we get (3).

(4) $d_3 = 5$. Otherwise, $d_3 \geq 6$; combining this with (1), we see that $\bar{\pi}$ satisfies the hypothesis of the theorem, contradicting (2). So (4) holds.

Throughout the proof, let m_k denote the multiplicity of k in π . Then

(5) $m_3 \geq 5$. Otherwise, by Theorem 3.1, we have $m_3 = 4$. Thus precisely three entries of $\bar{\pi}$ are three by (1) and (4). It follows from Theorem 3.1 that $\bar{\pi}$ has a Z_3 -connected realization. Hence so does π by Lemma 2.6. This contradiction yields (5).

(6) $n \geq 9$. Otherwise, $n \leq 8$. From (4) and (5) we deduce that $n = 8$ and $m_3 = 5$. In view of (3), $d_1 \leq 6$. So $\pi = (5^3, 3^5)$ or $(6^2, 5, 3^5)$.

For $\pi = (5^3, 3^5)$, let G be the graph obtained from the disjoint union of a W_4 , with hub v_0 and rim $v_1v_2v_3v_4v_1$, and a path $v_5v_6v_7$ by adding edges $v_5v_1, v_5v_3, v_7v_1, v_7v_3, v_6v_0$. Then G/W_4 is triangularly connected and contains two 2-circuits. Since 2-circuits are Z_3 -connected, so is G/W_4 by Lemma 2.9(a). It follows from Lemmas 2.9(c) and 2.5(b) that G is Z_3 -connected.

For $\pi = (6^2, 5, 3^5)$, we have $\bar{\pi} = (5^2, 4, 3^4)$. Let $\bar{G}$ be the graph obtained from the disjoint union of a W_4 and a W_3 by identifying one rim edge of W_4 with a rim edge of W_3 . Then $\bar{G}$ is a realization of $\bar{\pi}$. Using the same proof employed in the preceding paragraph, we can justify that $\bar{G}$ is Z_3 -connected. Now let G be the graph obtained from $\bar{G}$ by adding a new vertex v and three edges between v and vertices of degree at least four in $\bar{G}$. Clearly, G is a realization of π and is Z_3 -connected by Lemma 2.6. This contradiction proves (6).

From (3), (4), (5), (6), and Lemma 4.5, we deduce the following:

(7) $d_4 = 5$.

(8) $d_2 = 5$. Otherwise, $d_2 \geq 6$. In view of (7), $\bar{\pi}$ satisfies the conditions of the theorem, contradicting (2).

(9) $d_5 \leq 4$. Suppose to the contrary that $d_5 \geq 5$. By (7), we have $d_5 = 5$. It follows that $d_1 = 5$ and $d_6 \leq 4$, for otherwise $\bar{\pi}$ satisfies the conditions of the theorem, contradicting (2). Thus $\pi = (5^5, 4^{m_4}, 3^{m_3})$. Consider the sequence $\pi^* = (4^{m_4+4}, 3^2)$. By Lemma 4.4, π^* is graphical. Thus it has a Z_3 -connected realization H_1 by Theorem 3.1. Note that m_3 is odd since $\pi = (5^5, 4^{m_4}, 3^{m_3})$ (by Lemma 2.1). Let H_2 be a bipartite cubic graph on $m_3 - 1$ (≥ 4) vertices, where H_2 is simple if $m_3 \geq 7$, or contains precisely two 2-circuits, C_1 and C_2 , if $m_3 = 5$. We take three edges e_1, e_2, e_3 in H_2 such that e_1 and e_2 are independent and that $e_i \in C_i$ for $i = 1, 2$ if $m_3 = 5$. Let us construct a graph from the disjoint union of H_1 and H_2 as follows: Cut each of e_1 and e_2 into two edges (let u_1, u_2, u_3, u_4 denote the new vertices), then subdivide e_3 once by a degree-two vertex v , and finally identify u_1, u_2, u_3, u_4 with four degree-four vertices in H_1 , respectively, and v with a degree-three vertex in H_1 . Clearly, G is a realization of π . It follows from Lemma 4.3 that G admits a nowhere-zero 3-flow. This contradiction implies (9).

From the above observations, we conclude

(10) $\pi = (d_1, 5^3, 4^{n-m_3-4}, 3^{m_3})$. So $m_3 \leq n - 4$.

Let us distinguish between two cases according to the parity of d_1 .

Case 1. d_1 is even. In view of (10), we have

(11) m_3 is odd.

We propose to construct a realization G of π with properties (a) and (b), using Lemma 4.3, such that the Z_3 -connected graph H_1 (recall Theorem 3.1) has degree sequence $\pi^* = (d_1^*, 4^{n-m_3-2}, 3^2)$, where d_1^* is to be determined, and the cubic bipartite graph H_2 has $m_3 - 1$ vertices.

In order to determine d_1^* , set $A = \max\{4, d_1 - (m_3 - 1)\}$ and $B = \min\{n - m_3, d_1\}$. By virtue of (3), (5), (10), and (11), it is a routine matter to check that

(12) $A \leq B$ and equality holds if and only if $A = B = 4$.

Thus (12) guarantees the existence of d_1^* such that $A \leq d_1^* \leq B$ and that d_1^* is even. By Lemma 4.4, the sequence $\pi^* = (d_1^*, 4^{n-m_3-2}, 3^2)$ is graphical, and hence by Theorem 3.1 it admits a Z_3 -connected realization H_1 . By the definitions of A , B , and d_1^* , we have $d_1 - (m_3 - 1) \leq d_1^* \leq d_1$, so

(13) $0 \leq d_1 - d_1^* \leq m_3 - 1$, and hence $0 \leq \frac{d_1 - d_1^*}{2} \leq \frac{m_3 - 1}{2}$ (recall Case 1).

Let H_2 be a cubic bipartite graph with $m_3 - 1$ (≥ 4) vertices, where H_2 is simple if $m_3 \geq 7$ and contains precisely two disjoint 2-circuits, C_1 and C_2 , if $m_3 = 5$ (see (11)). By (13), there exist two disjoint matchings M_1, M_2 in H_2 such that $|M_1| = \frac{d_1 - d_1^*}{2}$, $|M_2| = 2$, and $(M_1 \cup M_2) \cap C_i \neq \emptyset$ for $i = 1, 2$ if $m_3 = 5$. Let us first subdivide each edge in M_1 once by a new vertex and identify all these vertices with the vertex of degree d_1^* in H_1 , and then cut each edge in M_2 into two edges (let v_1, v_2, v_3, v_4 be the new vertices) and identify v_1, v_2, v_3 with three degree-four vertices in H_1 , respectively, and v_4 with a degree-three vertex in H_2 . Clearly, the resulting graph G is simple and is a desired realization of the sequence π . By Lemma 4.3, G admits a nowhere-zero 3-flow.

Case 2. d_1 is odd. From (10) we see that

(14) m_3 is even. So $m_3 \geq 6$ by (5).

We propose to construct a realization G of π with properties (a) and (b), using Lemma 4.3, such that the Z_3 -connected graph H_1 (recall Theorem 3.1) has degree sequence $\pi^* = (d_1^*, 4^{n-m_3-1}, 3^2)$, where d_1^* is to be determined, and the cubic bipartite graph H_2 has $m_3 - 2$ vertices.

To this end, set $A = \max\{4, d_1 - (m_3 - 2)\}$ and $B = \min\{n - m_3 + 1, d_1\}$. By virtue of (3), (5), (10), and (14), we get

(15) $A \leq B$ and equality holds if and only if $A = B = 4$.

Thus (15) guarantees the existence of d_1^* such that $A \leq d_1^* \leq B$ and that d_1^* is even. By Lemma 4.4, the sequence $\pi^* = (d_1^*, 4^{n-m_3-1}, 3^2)$ is graphical, and hence by Theorem 3.1 it admits a Z_3 -connected realization H_1 . By the definitions of A , B , and d_1^* , we have $d_1 - (m_3 - 2) \leq d_1^* \leq d_1$, so

(16) $0 \leq d_1 - d_1^* \leq m_3 - 2$, and hence $1 \leq \frac{d_1 - d_1^* + 1}{2} \leq \frac{m_3 - 2}{2}$ (as d_1 and d_1^* have different parities).

Let H_2 be a cubic bipartite graph with $m_3 - 2$ (≥ 4) vertices, where H_2 is simple if $m_3 \geq 8$, or contains precisely two disjoint 2-circuits, C_1 and C_2 , if $m_3 = 6$ (see (14)). By (16), there exist a pair of edge-disjoint matchings M_1, M_2 in H_2 such that $|M_1| = \frac{d_1 - d_1^* + 1}{2}$, $|M_2| = 1$, and $(M_1 \cup M_2) \cap C_i \neq \emptyset$ for $i = 1, 2$ if $m_3 = 6$. Let ab be a special edge in M_1 . We construct a graph from the disjoint union of H_1 and H_2 as follows: First subdivide each edge in M_1 once by a new vertex (let x be the new vertex on the special ab) and identify all these new vertices with the vertex v_1 of degree d_1^* in H_1 , then cut the edge in M_2 into two edges (let y, z be

the new vertices) and identify y, z with two degree-four vertices in H_1 , respectively. At this stage, the resulting graph has the degree sequence $(d_1 + 1, 5^2, 4^{n-m_3-3}, 3^{m_3})$. Finally, switch the edge ax away from v_1 to some degree-four vertex. Clearly, the resulting graph G is simple and is a desired realization of the sequence π (see (10)). By Lemma 4.3, G admits a nowhere-zero 3-flow. This contradiction completes the proof of the theorem. $\square$

Proof of Theorem 4.2. The proof goes along the same line as that of Lemma 4.5, so we give only a sketch here. Let us consider four cases according to the values of k and d_4 : In each case we propose to construct a realization G of π with properties (a) and (b) using Lemma 4.3; the degree sequence π^* of the Z_3 -connected graph H_1 (recall Lemma 4.3) and the number of vertices in the cubic bipartite graph H_2 are given below.

Case 1. k is even and at least six. The number of vertices in H_2 is $k - 2$, and the degree sequence π^* of H_1 is $(d_1^*, d_2^*, 4^{n-2-k}, 3^2)$. In order to determine d_1^* and d_2^* , we

- set $A_i = \max\{4, d_i - (k - 2)\}$ for $i = 1, 2$,
- set $B_1 = \min\{n - k + 1, d_1 - p\}$, where $p = 3$ if $d_2 = 4$ and 2 if $d_2 \geq 5$, and
- set $B_2 = \min\{n - k + 1, d_2 - q\}$, where $q = 0$ if $d_2 = 4$ and 1 if $d_2 \geq 5$.

Case 2. $k = 4$. The number of vertices in H_2 is k , and the degree sequence π^* of H_1 is $(d_1^*, d_2^*, 4^{n-4-k}, 3^2)$. In order to determine d_1^* and d_2^* , we

- set $A_i = \max\{4, d_i - k\}$ for $i = 1, 2$,
- set $B_1 = \min\{n - k - 1, d_1 - p\}$, where $p = 3$ if $d_2 = 4$ and 2 if $d_2 \geq 5$, and
- set $B_2 = \min\{n - k - 1, d_2 - q\}$, where $q = 0$ if $d_2 = 4$ and 1 if $d_2 \geq 5$.

Case 3. k is odd and $d_4 = 4$. The number of vertices in H_2 is $k - 1$, and the degree sequence π^* of H_1 is $(d_1^*, d_2^*, 4^{n-3-k}, 3^2)$. From $d_4 = 4$ it can be seen that $n - 3 - k \geq 1$. In order to determine d_1^* and d_2^* , we

- set $A_i = \max\{4, d_i - (k - 1)\}$ for $i = 1, 2$,
- set $B_1 = \min\{n - k, d_1 - p\}$, where $p = 3$ if $d_2 = 4$ and 2 if $d_2 \geq 5$, and
- set $B_2 = \min\{n - k, d_2 - q\}$, where $q = 0$ if $d_2 = 4$ and 1 if $d_2 \geq 5$.

In each of the above three cases, it is a routine matter to check that $A_i \leq B_i$ and equality holds if and only if $A_i = B_i = 4$. Thus there exist d_1^* and d_2^* such that

- $A_i \leq d_i^* \leq B_i$ for $i = 1, 2$;
- $d_1^* + d_2^*$ is even.

By Lemma 4.4, the sequence π^* is graphical, and hence by Theorem 3.1 it admits a Z_3 -connected realization H_1 . Clearly, we can choose H_2 so that it is simple if it has at least six vertices and contains precisely two disjoint 2-circuits otherwise. Note that H_2 contains three disjoint perfect matchings. By subdividing or cutting a certain number (at least two) of edges in these matchings, we can get a realization G of π , as desired.

Case 4. k is odd and $d_4 = 3$. In this case $n - k - 2 = 1$, so $k = n - 3$. Hence $\pi = (d_1, d_2, 4, 3^k)$, $k \geq 7$, and n is even (recall the hypothesis of the theorem). By Lemma 2.1(a), $d_1 + d_2$ is odd. From $d_1 + d_2 \geq 11$ we further deduce that $n - 3 \geq d_1 \geq 7$ if d_1 is odd.

Let H_1 be a Z_3 -connected realization of $(4^3, 3^2)$ in which vertices x_1, x_2, x_3 are of degree four, and let H_2 be a cubic bipartite graph with $k - 1$ vertices. Let u be a vertex of H_2 , let v_1, v_2, v_3 be the neighbors of u in H_2 , and let M_1, M_2, M_3 be three disjoint perfect matchings of H_2 . Renaming the subscripts if necessary, we assume $uv_i \in M_i$ for $i = 1, 2, 3$. For odd d_1 , let G^* be the graph obtained from the union of H_1 and H_2 by identifying x_1 and u . For even d_1 , let G^* be the graph obtained from the union of H_1 and H_2 by first splitting u into three vertices $\{u_1, u_2, u_3\}$ (so

the three edges incident with u in H_2 become u_1v_1, u_2v_2, u_3v_3) and then identifying u_1, u_2 with x_1 and u_3 with x_2 .

Set $t_1 = (d_1 - 7)/2$ and $t_2 = (d_2 - 4)/2$ if d_1 is odd, and set $t_1 = (d_1 - 6)/2$ and $t_2 = (d_2 - 5)/2$ otherwise. Since $d_1 + d_2$ is odd, t_1 and t_2 are both integers. Moreover, since $k = n - 3$, $n - 2 \geq d_1 \geq d_2 \geq 4$, and $d_1 \leq n - 3$ if d_1 is odd, we have

- $t_1 \leq (k - 1)/2 - 3$ if d_1 is odd and $t_1 \leq (k - 1)/2 - 2$ otherwise, and
- $t_2 \leq (k - 1)/2 - 1$ if d_1 is odd and $t_2 \leq (k - 1)/2 - 1$ otherwise.

For odd d_1 , let F_1 be a subset of M_1 such that F_1 covers none of v_1, v_2, v_3 and $|F_1| = t_1$, and let F_2 be a subset of M_2 such that $uv_2 \notin F_2$ and $|F_2| = t_2$. For even d_1 , let F_1 be a subset of M_1 such that F_1 covers neither v_1 nor v_2 and $|F_1| = t_1$, and let F_2 be a subset of M_3 such that $uv_3 \notin F_2$ and $|F_2| = t_2$. Finally, we construct a graph from G^* as follows: For $i = 1, 2$, subdivide each edge in F_i once by a degree-two vertex, and then identify all these degree-two vertices with x_i . Clearly, the resulting graph G is simple and is a realization of π that admits a nowhere-zero 3-flow, a contradiction. $\square$

5. Realizations with 3-flows. To establish the main theorem of this paper, we shall break the proof into two parts and turn to proving the following two theorems.

THEOREM 5.1. *Let $\pi = (d_1, d_2, \dots, d_n)$ be a graphical sequence with $d_1 \geq d_2 \geq \dots \geq d_n \geq 3$. Then π has a realization G that admits a nowhere-zero 3-flow if and only if $\pi \neq (k, 3^k), (k^2, 3^{k-1})$, where k is odd.*

THEOREM 5.2. *Let $\pi = (d_1, d_2, \dots, d_n)$ be a graphical sequence with $d_1 \geq d_2 \geq \dots \geq d_n = 2$. Then π has a realization G that admits a nowhere-zero 3-flow if and only if $\pi \neq (3^4, 2)$.*

The following lemma will be used repeatedly in our proof.

LEMMA 5.3. *Let $k \geq 4$ be an even integer, and let $\pi = (a_1, a_2, a_3, 3^k)$ be a sequence with $0 \leq a_i \leq k$. Then π has a realization that admits a nowhere-zero 3-flow if one of the following holds:*

- (a) a_i is even for $i = 1, 2, 3$, and $a_1 + a_2 + a_3 \geq 4$ if $k = 4$;
- (b) one of a_1, a_2, a_3 is even, and the remaining two are odd and at least three.

Proof. (a) Let H be a cubic bipartite graph with k vertices, where H is simple if $k \geq 6$, or contains precisely two disjoint 2-circuits, C_1 and C_2 , if $k = 4$. Then the edge set of H can be decomposed into three perfect matchings M_1, M_2, M_3 . Since $a_i \leq k$ and since $a_1 + a_2 + a_3 \geq 4$ if $k = 4$, we can find a subset F_i of M_i such that $|F_i| = \frac{a_i}{2}$ for $i = 1, 2, 3$ and that $(F_1 \cup F_2 \cup F_3) \cap C_i \neq \emptyset$ for $i = 1, 2$ if $k = 4$. For $i = 1, 2, 3$, let us subdivide each edge in F_i once by a degree-two vertex and then identify all these degree-two vertices as a single vertex v_i . Then the resulting graph G is a realization of the sequence π . By Lemma 2.2, H admits a nowhere-zero 3-flow, and so does G .

(b) Renaming the subscripts if necessary, we may assume that a_3 is even. By Lemma 2.1, k is even and so $a_i \leq k - 1$ for $i = 1, 2$. Let H be a simple cubic bipartite graph with $k + 2$ vertices, and let M_1, M_2, M_3 be three disjoint perfect matchings of H . Take an arbitrary edge u_1v_1 in M_1 , and take a subset F_i of M_i such that $|F_i| = \frac{a_i - 3}{2}$ for $i = 1, 2$, $|F_3| = \frac{a_3}{2}$, and F_1 (resp., F_2) covers no vertices of $N(u_1)$, the neighborhood of u_1 (resp., $N(v_1)$). Let us now subdivide each edge of F_1 (resp., F_2) once by a degree-two vertex and then identify all these degree-two vertices with u_1 (resp., v_1), and subdivide each edge of F_3 once by a degree-two vertex and then identify all these degree-two vertices. Then the resulting graph G is a realization of the sequence π . By Lemma 2.2, H admits a nowhere-zero 3-flow, and so does G . $\square$

Proof of Theorem 5.1. Since the “only if” part is already established by Lemma 2.10, let us proceed to the “if” part.

Assume the contrary: π is a counterexample with minimum n . By Theorems 3.1, we have

(1) $m_3 \geq 4$.

From Theorem 4.1 it can be seen that

(2) $d_3 \leq 4$.

Let us further make some simple observations.

(3) $4 \leq d_1 \leq n - 2$. The upper bound follows instantly from Theorem 3.2. To justify lower bound, we assume to the contrary that $d_1 = 3$. So $\pi = (3^n)$. Using the upper bound, we have $n \geq 5$. From Lemma 2.1(a) we deduce that n is even and so $n \geq 6$. Thus π can be realized by a bipartite cubic graph, which, by Lemma 2.2, admits a nowhere-zero 3-flow, a contradiction. So (3) holds.

(4) $d_2 \geq 4$. Otherwise, $d_2 = 3$. Thus $\pi = (d_1, 3^{n-1})$. If $n - 1$ is even, then so is d_1 . Set $a_1 = d_1$, $a_2 = a_3 = 0$, and $k = n - 1$. By (3) and Lemma 5.3(a), π has a realization that admits a nowhere-zero 3-flow, a contradiction.

So we assume that $n - 1$ is odd. In this case d_1 is also odd. Write $\pi = (d_1, 3, 3^{n-2})$. Set $a_1 = 0$, $a_2 = d_1$, $a_3 = 3$, and $k = n - 2$. Then, by (3) and Lemma 5.3(b), π has a realization that admits a nowhere-zero 3-flow, a contradiction. This proves (4).

(5) $d_3 = 4$. By (2), $d_3 \leq 4$. We prove by contradiction and assume $d_3 = 3$. Thus $\pi = (d_1, d_2, 3^{n-2})$. Observe that $d_1 + d_2$ is odd, for otherwise (1), (3), and Lemma 5.3 (with $a_i = d_i$ for $i = 1, 2$ and $a_3 = 0$) would guarantee the existence of a realization of π that admits a nowhere-zero 3-flow, a contradiction. It follows that n and exactly one of d_1 and d_2 are odd.

If $d_1 \leq n - 3$, then, by Lemma 5.3(b) (with $a_i = d_i$ for $i = 1, 2$, $a_3 = 3$, and $n - 3 \geq 4$ because n is odd), π has a realization that admits a nowhere-zero 3-flow; this contradiction implies that $d_1 = n - 2$. Hence d_1 is odd and d_2 is even. Now let us take odd wheel W_{n-2} with hub v_0 and rim $v_1 v_2 \dots v_{n-2} v_1$ and take $M = \{v_{2i} v_{2i+1} : 0 \leq i \leq d_2/2 - 1\}$. Clearly, M is a matching of size $d_2/2$. Let us subdivide each edge in M once and identify all the new vertices as a single vertex. Then the resulting graph is a realization of π and admits a nowhere-zero 3-flow (to find it, direct each of edges $v_0 u$, $v_0 v_2$, $v_0 v_3$ from v_0 to the other end, where u is the vertex subdividing $v_0 v_1$; then directions of the remaining edges can be determined accordingly). This contradiction implies (5).

(6) $n \geq 9$. Otherwise, $n \leq 8$. By (5) and (1), we have $d_3 = 4$ and $m_3 \geq 4$. So $n \geq 7$.

If $n = 7$ then, by (3), (5), and Lemma 2.1(a), we have $\pi = (4^3, 3^4)$ or $(5^2, 4, 3^4)$. For $\pi = (4^3, 3^4)$, clearly $K_{3,4}$ is a realization of π that admits a nowhere-zero 3-flow. For $\pi = (5^2, 4, 3^4)$, let G be the graph obtained from the union of W_4 and W_3 by identifying a rim edge of W_4 with a rim edge of W_3 . Then G is a realization of π that, by Lemma 2.9(c) and (a), admits a nowhere-zero 3-flow. So we have $n = 8$.

Since $d_3 = 4$ by (5), we have $m_3 \leq 5$; combining this with (1), we further have $m_3 = 4$ or 5 . Since $d_1 \leq 6$ by (3), one of the following cases must occur:

- $\pi = (6^2, 4^2, 3^4)$,
- $\pi = (6, 4^3, 3^4)$,
- $\pi = (4^4, 3^4)$,
- $\pi = (5^2, 4^2, 3^4)$,
- $\pi = (5, 4^2, 3^5)$, or
- $\pi = (6, 5, 4, 3^5)$.

For each π , we shall exhibit a realization G that admits a nowhere-zero 3-flow, thereby reaching a contradiction.

For $\pi = (6^2, 4^2, 3^4)$, let $\bar{G}$ be the graph obtained from the union of W_4 and W_3 by identifying a rim edge of W_4 with a rim edge of W_3 . Then $\bar{G}$ is the realization of the residual sequence $\bar{\pi}$ and $\bar{G}$ is Z_3 -connected. So it is easy to obtain a realization G of π from $\bar{G}$ such that G admits a nowhere-zero 3-flow, a contradiction.

For $\pi = (6, 4^3, 3^4)$, let G be the graph obtained from W_6 by adding an edge between two nonadjacent vertices, then subdividing two independent edges once each, and finally identifying these new vertices as one vertex. (Since W_6 is Z_3 -connected, so is the graph obtained from W_6 by adding an edge by Lemma 2.7.)

For $\pi = (4^4, 3^4)$, let G be the graph obtained from the cubic bipartite graph with four vertices by subdividing each of the four multiple edges once and then connecting these four degree-two vertices with a 4-circuit. (Note that G can be decomposed into a subdivision of a cubic bipartite graph and a 4-circuit.)

For $\pi = (5^2, 4^2, 3^4)$, let G be the graph obtained from the union of two W_4 's by identifying a rim edge of one W_4 with a rim edge of the other W_4 . (In fact G is Z_3 -connected.)

For $\pi = (5, 4^2, 3^5)$, let G be the graph obtained from a $K_{3,3}$ (with color classes $\{u_1, u_2, u_3\}$ and $\{v_1, v_2, v_3\}$) by subdividing $u_i v_i$ once with a degree-two vertex w_i for $i = 2, 3$, and then adding a triangle $u_1 w_2 w_3 u_1$.

For $\pi = (6, 5, 4, 3^5)$, let G be the graph obtained from W_4 (in which $u_1 u_2$ is a rim edge) by adding a path $v_1 v_3 v_2$ and then adding edges $v_1 u_1, v_1 u_2, v_2 u_1, v_2 u_2, v_3 u_1$. (Note that G is triangularly connected and contains W_4 . So it is Z_3 -connected, by Lemma 2.9(a).)

It is a routine matter to check that G is a realization of π and admits a nowhere-zero 3-flow in each case. This contradiction implies (6).

From (3), (5), (6), and Theorem 4.2, we deduce that

$$(7) \quad d_1 + d_2 \leq 10.$$

Since $d_1 \geq d_2 \geq 4$, we see that (d_1, d_2) is $(6, 4)$, or $(5, 5)$, or $(5, 4)$, or $(4, 4)$. So the following is the complete list of all possible configurations of π :

- $\pi = (6, 4^{m_4}, 3^{m_3})$,
- $\pi = (5^2, 4^{m_4}, 3^{m_3})$,
- $\pi = (5, 4^{m_4}, 3^{m_3})$, and
- $\pi = (4^{m_4}, 3^{m_3})$,

where m_k is the multiplicity of k in π . Let us process these cases one by one: For each π , we shall construct a realization G that admits a nowhere-zero 3-flow, thereby reaching a contradiction.

Case 1. $\pi = (4^{m_4}, 3^{m_3})$. Note that m_3 is even. Depending on the value of m_3 , we consider two subcases.

Subcase 1.1. $m_3 = 4$. Our proof relies on the following statement.

(8) The sequence $(4^k, 2^4)$, with $k \geq 1$, can be realized by a simple connected graph H that admits a nowhere-zero 3-flow. To justify this, we apply induction on k . For $k = 1$, the graph H_1 obtained from two triangles by gluing them at a common vertex is as desired. Suppose that H_k is the desired realization of $(4^k, 2^4)$. Let e, f be two independent edges in H_k , and let H_{k+1} be the graph obtained from H_k by first subdividing each of e, f once with a degree-two vertex and then identifying these degree-two vertices. Clearly, H_{k+1} is a realization of $(4^{k+1}, 2^4)$ and admits a nowhere-zero 2-flow since it is Eulerian. So (8) holds.

By (6), we have $m_4 \geq 5$. In view of (8), we can find a connected realization H of the sequence $(4^{m_4-4}, 2^4)$. Let G be the graph obtained from H by adding a bipartite cubic graph F with four vertices, then subdividing each of the four multiple edges in

F with a degree-two vertex, and finally identifying these four degree-two vertices with the degree-two vertices of H , respectively. Clearly, G is a realization of π and admits a nowhere-zero 3-flow.

Subcase 1.2. $m_3 \geq 6$. For $m_4 \geq 5$, let H_1 be a bipartite cubic simple graph with m_3 vertices, and let H_2 be a Z_3 -connected realization of $(4^{m_4-2}, 3^2)$ (recall Lemmas 4.4 and 3.3). Then we can get a desired realization G of π by cutting one edge of H_1 and then connecting these two half edges to the degree-three vertices of H_2 , respectively. For $3 \leq m_4 \leq 4$, let G be the graph obtained from the union of H_1 and an m_4 -circuit H_3 by subdividing m_4 edges of H_1 once with degree-two vertices and then identifying these new vertices with m_4 vertices of H_3 , respectively. Clearly, G is a realization of π and admits a nowhere-zero 3-flow.

Case 2. $\pi = (5, 4^{m_4}, 3^{m_3})$. Note that m_3 is odd, so $m_3 \geq 5$ by (1). We distinguish two subcases according to the value of m_3 .

Subcase 2.1. $m_3 = 5$. Recall that $n \geq 9$ by (6). For $n = 9$, let H_1 be the graph obtained from W_3 by adding a new vertex and joining it to two vertices of the W_3 . Clearly, H_1 admits a nowhere-zero 3-flow and has degree sequence $(4^2, 3^2, 2)$. For $n \geq 10$, we have $n - m_3 - 2 \geq 3$. By Lemma 4.4, the sequence $(4^{n-m_3-2}, 3^2)$ is graphical and hence, by Lemma 3.3, admits a Z_3 -connected realization F . Let H_1 be the graph obtained from F by subdividing one edge once. Clearly, H_1 admits a nowhere-zero 3-flow and has degree sequence $(4^{n-m_3-2}, 3^2, 2)$. Let H_2 be the cubic bipartite graph on four vertices in which both u_1v_1 and u_2v_2 are of multiplicity two, and let G be the graph obtained from the union of H_1 and H_2 by subdividing u_iv_i once for $i = 1, 2$, and then identifying one new vertex with the degree-two vertex of H_1 and the other new vertex with a degree-three vertex of H_1 . Clearly, G is a realization of π and admits a nowhere-zero 3-flow.

Subcase 2.2. $m_3 \geq 7$. Since $d_3 = 4$ by (5), we have $m_4 \geq 2$. Let H_1 be a realization of $(4^{n-m_3-2}, 3^2, 2)$ as exhibited in the preceding paragraph, and let H_2 be a cubic bipartite simple graph with $m_3 - 1$ vertices. Using H_1 and H_2 and following the same argument as the preceding paragraph, we can obviously get a realization G of π that admits a nowhere-zero 3-flow.

Case 3. $\pi = (6, 4^{m_4}, 3^{m_3})$. Let H be an arbitrary realization of π , and let u be the vertex of degree six in H . Then the configuration of π implies the existence of two nonadjacent neighbors v, w of u in H . Let H' be the graph obtained from H by replacing path $vu w$ with edge vw . Then the degree sequence of H' is $\pi' = (4^{m_4+1}, 3^{m_3})$. By Case 1, π' has a realization G' that admits a nowhere-zero 3-flow. Moreover, if $\pi' = (4^5, 3^4)$, by Subcase 1.1, G' can be chosen such that there is a degree-four vertex x and an edge e such that x is not incident with e and is not adjacent to the end-vertices of e . Let x be a degree-four vertex in G' . Let $X_1 = N(x) \cup \{x\}$ and $X_2 = V(G') \setminus X_1$. Then there must be an edge e not incident with x such that x is not adjacent to the end-vertices of e if $\pi' \neq (4^5, 3^4)$. Otherwise, G' is connected and X_2 is an independent set. Let $[X_1, X_2]$ denote the set of edges with one end in X_1 and the other in X_2 . Then $(n - 5) \times 3 \leq \sum_{u \in X_2} d(u) = |[X_1, X_2]| \leq \sum_{v \in X_1 \setminus \{x\}} (d(v) - 1) \leq 4 \times 3 = 12$. Hence, $n \leq 9$ with equality if and only if X_1 consists of all degree-four vertices and X_2 consists of all degree-three vertices. By (6), we have $n = 9$. Therefore, the degree sequence of G' is $(4^5, 3^4)$, a contradiction to the assumption that $\pi' \neq (4^5, 3^4)$. Therefore, in any case, we can find a degree-four vertex a and an edge e such that a is not incident with e and is not adjacent to the end-vertices of e . Let G be the graph obtained from G' by subdividing e once and then identifying the new vertex with a . Clearly, G is a realization of π and admits a nowhere-zero 3-flow.

Case 4. $\pi = (5^2, 4^{m_4}, 3^{m_3})$. Since π is graphical, using the same argument

employed in the preceding paragraph we deduce that the sequence $(5, 4^{m_4}, 3^{m_3+1})$ is graphical and hence, by Case 2, has a realization H that admits a nowhere-zero 3-flow. Let x be a degree-three vertex. Let $X_1 = N(x) \cup \{x\}$ and $X_2 = V(H) \setminus X_1$. We first show that there exists an edge e such that x is not incident with e and is not adjacent to the end-vertices of e . Otherwise, H is connected and X_2 is independent. Let $[X_1, X_2]$ denote the set of edges with one end in X_1 and the other in X_2 . Since the degree of each vertex in X_2 is at least three and the degree of each vertex in X_1 is at most five, we have $(n-4) \times 3 \leq \sum_{u \in X_2} d(u) = |[X_1, X_2]| \leq \sum_{v \in X_1 \setminus \{x\}} (d(v) - 1) \leq 4 \times 3 = 12$. Therefore, $n \leq 4$, contradicting (6), i.e., that $n \geq 9$. Let e be an edge and x be a degree vertex not incident with e such that x is not adjacent to any end-vertices of e . Let G be the graph obtained from H by subdividing an edge e once and then identifying the new vertex with a degree-three vertex not incident to e . Clearly, G is a realization of π and admits a nowhere-zero 3-flow. This completes the proof of Theorem 5.1. $\square$

Let us make some preparation before presenting the proof of Theorem 5.2.

LEMMA 5.4. *Let k be an integer with $k = 2$ or $k \geq 4$, and let $\pi = (k, 3^k, 2)$ or $(k^2, 3^{k-1}, 2)$. If π is graphical, then it has a realization that admits a nowhere-zero 3-flow.*

Proof. Note that if $k = 2$, then $\pi = (3^2, 2^2)$. Let G be the graph obtained from W_3 by deleting one edge. Clearly, G is a realization of π and admits a nowhere-zero 3-flow. So we assume

(1) $k \geq 4$. According to the configurations of π , we consider two cases.

Case 1. $\pi = (k, 3^k, 2)$. If k is even, then, by Lemma 5.3(a), π has a realization that admits a nowhere-zero 3-flow. It remains to consider the subcase when k is odd. Thus $k \geq 5$ by (1). Let H be a bipartite cubic simple graph with $k+1$ vertices, let u be a vertex of H , and let $\{v_1, v_2, v_3\}$ be the neighbors of u . Then $H \setminus \{u, v_1, v_2\}$ contains a matching M of size $(k-3)/2$. Let G be the graph obtained from H by subdividing each edge in M once, then identifying all the degree-two vertices with u , and finally subdividing one edge uv_3 once. Clearly, G is a realization of π and admits a nowhere-zero 3-flow.

Case 2. $\pi = (k^2, 3^{k-1}, 2)$. In this case k is odd, so $k \geq 5$ by (1). Write $k = 2t+1$. Let H be the graph obtained from the disjoint union of a 4-circuit and $t-1$ triangles by gluing them at a common vertex x . Then H has $2t+2$ vertices and degree sequence $(2t, 2^{2t+1})$. Let G be the graph obtained from H by adding a new vertex y and making it adjacent to all vertices of H except precisely one degree-two vertex in a triangle. Then the degree sequence of G is $((2t+1)^2, 3^{2t}, 2)$, which is exactly π . Since G is triangularly connected and contains W_4 , it is Z_3 -connected by Lemma 2.9 and hence admits a nowhere-zero 3-flow. $\square$

Proof of Theorem 5.2. The “only if” part is already established by Lemma 2.10, so we proceed to the “if” part.

Assume the contrary: π is a counterexample with minimum n . Observe that

(1) $d_2 \geq 3$. Otherwise, $d_2 = d_3 = \dots = d_n = 2$. So d_1 is even. Thus each realization of π admits a nowhere-zero 2-flow; this contradiction leads to (1).

(2) The sequence $\sigma = (d_1, d_2, \dots, d_{n-1})$ is not graphical. Assume to the contrary that σ is graphical. Then $\sigma \neq (3^4, 2), (k, 3^k), (k^2, 3^{k-1})$, where k is an odd integer, for otherwise $\pi = (3^4, 2^2), (k, 3^k, 2)$, or $(k^2, 3^{k-1}, 2)$, so π has a realization that admits a nowhere-zero 3-flow by Lemma 5.3(a) or Lemma 5.4, a contradiction. By Theorem 5.1 and the assumption on π , the sequence σ has a realization H that admits a nowhere-zero 3-flow. Let G be a graph obtained from H by subdividing an edge once. Clearly,

G is a realization of π and admits a nowhere-zero 3-flow; this contradiction implies (2).

(3) Let $\bar{G}$ be an arbitrary realization of the residual sequence $\bar{\pi} = (d_1 - 1, d_2 - 1, d_3, \dots, d_{n-1})$ and let v_i be a vertex of $\bar{G}$ with degree $d_i - 1$ for $i = 1, 2$. Then v_1v_2 is an edge of $\bar{G}$.

Otherwise, v_1 and v_2 are nonadjacent in $\bar{G}$. Thus the graph obtained from $\bar{G}$ by adding edge v_1v_2 is a realization of the sequence $\sigma = (d_1, d_2, \dots, d_{n-1})$, contradicting (1). So (3) holds.

(4) The residual sequence $\bar{\pi}$ is $(3^4, 2)$, $(k, 3^k)$, or $(k^2, 3^{k-1})$, where k is an odd integer.

Otherwise, $\bar{\pi}$ has a realization $\bar{G}$ that admits a nowhere-zero 3-flow. Let v_i be a vertex of $\bar{G}$ with degree $d_i - 1$ for $i = 1, 2$. By (4), v_1v_2 is an edge of $\bar{G}$. Let G be the graph obtained from $\bar{G}$ by adding a new vertex w and making it adjacent to both v_1 and v_2 . Since G contains the triangle wv_1v_2w and since $\bar{G}$ admits a nowhere-zero 3-flow, it is easy to see that so does G . Hence (4) is justified.

From (4) we deduce that one of the following four cases must occur:

- $\pi = (4, 3^4, 2)$,
- $\pi = (4^2, 3^2, 2^2)$,
- $\pi = (k + 1, 4, 3^{k-1}, 2)$ or
- $\pi = ((k + 1)^2, 3^{k-1}, 2)$,

where k is an odd integer. In each case we shall construct a realization of π that admits a nowhere-zero 3-flow, thereby reaching a contradiction.

For $\pi = (4, 3^4, 2)$, π has a realization G obtained by subdividing one edge once of a W_4 . Since W_4 admits a nowhere-zero 3-flow, so does G .

For $\pi = (4^2, 3^2, 2^2)$, let G be the graph obtained from a W_3 by adding a new vertex, making it adjacent to two vertices of the W_3 , and then subdividing an edge. Clearly G is a realization of π . To see that G admits a nowhere-zero 3-flow, let H be the graph obtained from W_3 by duplicating an edge. Then H is triangularly connected and contains a 2-circuit. By Lemma 2.9, H is Z_3 -connected. So G admits a nowhere-zero 3-flow as it is a subdivision of H .

For $\pi = (k + 1, 4, 3^{k-1}, 2)$, let G be a graph obtained from W_k by adding a new vertex and making it adjacent to the hub and a rim vertex. It is easy to see that G is a realization of π and admits a nowhere-zero 3-flow.

For $\pi = ((k + 1)^2, 3^{k-1}, 2)$, let G be the graph obtained from the disjoint union of $\frac{k-1}{2}$ copies of W_3 by gluing all of them along an edge uv , and then adding a new vertex and making it adjacent to both u and v . Clearly, G is a realization of π and admits a nowhere-zero 3-flow.

This completes the proof of Theorem 5.2 and hence of Theorem 1.2. $\square$

REFERENCES

- [1] D. ARCHDEACON, *Realizing Degree Sequences with Graphs Having 3-Flows*, <http://www.cems.uvm.edu/~archdeac/problems/seqflow.html>.
- [2] P. J. CAMERON, *Problems from the 16th British combinatorial conference*, Discrete Math., 197/198 (1999), pp. 799–812.
- [3] P. A. CATLIN, *Double cycle covers and the Petersen graph*, J. Graph Theory, 13 (1989), pp. 465–483.
- [4] M. DEVOS, R. XU, AND G. YU, *Nowhere-zero Z_3 -flows through Z_3 -connectivity*, Discrete Math., 306 (2006), pp. 26–30.
- [5] M. T. HAJIAGHAEE, E. S. MAHMOODIAN, V. S. MIRROKNI, A. SABERI, AND R. TUSSEKANI, *On the simultaneous edge-coloring conjecture*, Discrete Math., 216 (2000), pp. 267–272.
- [6] S. L. HAKIMI, *On realizability of a set of integers as degrees of the vertices of a linear graph. I*, SIAM J. Appl. Math., 10 (1962), pp. 496–506.

- [7] S. L. HAKIMI, *On realizability of a set of integers as degrees of the vertices of a linear graph. II. Uniqueness*, SIAM J. Appl. Math., 11 (1963), pp. 135–147.
- [8] F. JAEGER, *Nowhere-zero flow problems*, in Selected Topics in Graph Theory III, L. W. Beineke and R. J. Wilson, eds., Academic Press, London, 1988, pp. 71–95.
- [9] F. JAEGER, N. LINIAL, C. PAYAN, AND M. TARSI, *Group connectivity of graphs—A nonhomogeneous analogue of nowhere-zero flow properties*, J. Combin. Theory Ser. B, 56 (1992), pp. 165–182.
- [10] A. D. KEEDWELL, *Critical sets and critical partial Latin squares*, in Combinatorics, Graph Theory, Algorithms and Applications (Beijing, 1993), World Scientific Publishing, River Edge, NJ, 1994, pp. 111–123.
- [11] A. D. KEEDWELL, *Critical sets for Latin squares, graphs and block designs: A survey*, in Festschrift for C. St. J. A. Nash-Williams, Congr. Numer., 113 (1996), pp. 231–245.
- [12] D. J. KLEITMAN AND D. L. WANG, *Algorithm for constructing graphs and digraphs with given valences and factors*, Discrete Math., 6 (1973), pp. 79–88.
- [13] H.-J. LAI, R. XU, AND C.-Q. ZHANG, *3-Flow Reducible Configurations and Triangularly Connected Graphs*, manuscript.
- [14] R. LUO, W. ZANG, AND C.-Q. ZHANG, *Nowhere-zero 4-flows, simultaneous edge-colorings, and critical partial Latin squares*, Combinatorica, 24 (2004), pp. 641–657.
- [15] M. MAHDIAN, E. S. MAHMOODIAN, A. SABERI, M. R. SALAVATIPOUR, AND R. TUSSEKANI, *On a conjecture of Keedwell and the cycle double cover conjecture*, Discrete Math., 216 (2000), pp. 287–292.
- [16] P. D. SEYMOUR, *Nowhere-zero 6-flows*, J. Combin. Theory Ser. B, 30 (1981), pp. 130–135.
- [17] P. D. SEYMOUR, *Nowhere-zero flows*, in Handbook of Combinatorics, R. L. Graham, M. Grötschel, and L. Lovász, eds., Elsevier, Amsterdam, 1995, pp. 289–299.
- [18] W. T. TUTTE, *On the imbedding of linear graphs in surfaces*, Proc. London Math. Soc. (2), 51 (1949), pp. 474–483.
- [19] W. T. TUTTE, *A contribution to the theory of chromatical polynomials*, Canadian J. Math., 6 (1954), pp. 80–91.
- [20] W. T. TUTTE, *Graph Theory*, Addison–Wesley, Reading, MA, 1984.
- [21] C.-Q. ZHANG, *Integer Flows and Cycle Covers of Graphs*, Marcel Dekker, New York, 1997.

HAMILTON CYCLES IN RANDOM LIFTS OF DIRECTED GRAPHS*

PRASAD CHEBOLU[†] AND ALAN FRIEZE[†]

Abstract. An n -lift of a digraph K is a digraph with a vertex set $V(K) \times [n]$, and for each directed edge $(i, j) \in E(K)$ there is a perfect matching between fibers $\{i\} \times [n]$ and $\{j\} \times [n]$, with edges directed from fiber i to fiber j . If these matchings are chosen independently and uniformly at random, then we say that we have a random n -lift. We show that if h is sufficiently large, then a random n -lift of the complete digraph $\vec{K}_h$ is a Hamiltonian **whp**.

Key words. digraph, random, lifts, Hamilton cycle

AMS subject classifications. 05C80, 05C45, 05C20

DOI. 10.1137/060670808

1. Introduction. For a graph K , an n -lift G of K has a vertex set $V(K) \times [n]$, where for each vertex $v \in V(K)$, $\{v\} \times [n]$ is called the *fiber* above v and will be denoted by F_v . The edge set of an n -lift G consists of a perfect matching between fibers F_u and F_w for each edge $(u, w) \in E(K)$. The set of n -lifts will be denoted $\Lambda_n(K)$. In this paper we discuss random n -lifts, chosen uniformly from $\Lambda_n(K)$. In this case, the matchings between fibers are chosen independently and uniformly at random.

Lifts of graphs were introduced by Amit and Linial in [1], where they proved that if K is a connected, simple graph with minimum degree $\delta \geq 3$, and G is chosen randomly from $\Lambda_n(K)$, then G is δ -connected **whp**, where the asymptotics are for $n \rightarrow \infty$. They continued the study of random lifts in [2], where they proved expansion properties of lifts. Together with Matoušek, they gave bounds on the independence number and chromatic number of random lifts in [3]. Linial and Rozenman [13] give a tight analysis for when a random n -lift has a perfect matching.

Burgin et al. [5] showed that a random n -lift of the complete graph K_h is Hamiltonian, provided h is sufficiently large. In this paper we study a directed version of the question. An n -lift of a digraph K is a digraph with a vertex set $V(K) \times [n]$, and for each directed edge $(i, j) \in E(K)$ there is a perfect matching between fibers $\{i\} \times [n]$ and $\{j\} \times [n]$, with edges directed from fiber i to fiber j .

We use the notation $y \overset{r}{\in} Y$ for “ y is chosen uniformly at random from Y .” We let $\vec{K}_h$ denote the complete digraph on vertex set $[h]$. Note that here there are edges in both directions (u, v) and (v, u) for all $u \neq v \in [h]$.

THEOREM 1. *If h is sufficiently large and $D \overset{r}{\in} \Lambda_n(\vec{K}_h)$, then D is Hamiltonian **whp**.* We will use the 3-phase method used in Cooper and Frieze [6, 7], Cooper, Frieze, and Molloy [8], and Frieze, Karp, and Reed [11].

A *permutation digraph* is a set of vertex disjoint directed cycles that cover all n vertices. Its *size* is the number of cycles.

*Received by the editors September 26, 2006; accepted for publication (in revised form) December 13, 2007; published electronically March 20, 2008.

<http://www.siam.org/journals/sidma/22-2/67080.html>

[†]Department of Mathematics, Carnegie Mellon University, Pittsburgh PA 15213 (pchebolu@andrew.cmu.edu, alan@random.math.cmu.edu). The second author’s research was supported in part by NSF grant CCF 0502793.

Phase 1. We show that **whp** the lift D contains a directed permutation digraph of size at most $2 \ln n$.

Phase 2. We increase the minimum cycle length in the permutation digraph to at least

$$n_0 = \left\lceil \frac{100nh^3}{\ln n} \right\rceil.$$

Phase 3. We convert the Phase 2 permutation digraph to a Hamilton cycle.

The main difficulty involved in implementing this strategy comes from Phase 3. This is basically a second moment calculation, but it needs a *trick* to reduce the variance. The idea of the trick is from [6], but implementing the idea has turned out to be quite difficult. This is basically the content of section 4, where we prove a lower bound on the number of Hamilton cycles of a certain type in a digraph of high degree.

We use the following standard inequalities for the tails of the binomial distribution:

$$(1) \quad \Pr(|B(n, p) - np| \geq \epsilon np) \leq 2e^{-\epsilon^2 np/3}, \quad 0 \leq \epsilon \leq 1,$$

$$(2) \quad \Pr(B(n, p) \geq anp) \leq (e/a)^{anp}.$$

2. Phase 1. Making a permutation digraph with at most $2 \log n$ cycles.

LEMMA 1. Suppose that $D \stackrel{r}{\in} \Lambda_n(\vec{K}_h)$. Then **whp** D contains a permutation digraph with at most $2 \ln n$ cycles.

Proof. Let X_0 denote the Hamilton cycle $(1, 2, \dots, h)$ of K_h . Let F_i be the fiber of D corresponding to i . Let $r_{i,j}$ be the permutation defined by the matching $M_{i,j}$ from fiber F_i to fiber F_j in the lift D ; i.e., the edges of $M_{i,j}$ are $\{(i, k), (j, r_{i,j}(k)) : k \in [n]\}$. Let $r = r_{h,1} \circ r_{h-1,h} \circ \dots \circ r_{1,2}$. r defines a permutation of fiber F_1 . The permutation digraph $\{(i, x), (i+1, r_{i,i+1}(x)) : i \in [h], x \in [n]\}$ has as many cycles as the permutation r . The permutation r is a random permutation as it is the composition of random permutations. We know that the number of cycles in a random permutation is at most $2 \ln n$ **whp**; see, for example, Bollobás [4]. $\square$

We partition the cycles of the permutation digraph Σ_0 into sets SMALL and LARGE, containing small cycles C of length $|C| < n_0$ and large cycles $|C| \geq n_0$, respectively.

In a random permutation the expected number of vertices on cycles of length at most n_0 is precisely n_0 [12]. Thus, by the Markov inequality, **whp** Σ_0 contains at most $nh \log \log n / (4 \log n)$ vertices on small cycles.

Thus at the end of Phase 1, we can assume that we have a permutation digraph Σ_0 of size at most $2 \ln n$ which contains at most $nh \log \log n / (4 \log n)$ vertices on cycles of length $\leq n_0$. Let E_0 denote the edges D that are *not* in Σ_0 .

3. Phase 2. Removing small cycles. We now denote the vertices in the lift by $v_{i,k}$, where $i \in [h]$ and $k \in [n]$. We define a near permutation digraph (NPD) to be a digraph obtained from a permutation digraph by removing one edge. Thus an NPD Γ consists of a path $P(\Gamma)$ plus a permutation digraph $PD(\Gamma)$, which covers $([h] \times [n]) \setminus V(P(\Gamma))$.

Each step of the process we are about to describe involves the exposure of an edge $(v_{i,k}, v_{i',k'})$. When an edge is exposed in this way in Phase 2, we say that the two endpoints are *used*. The set of used vertices is denoted by W . Initially, $W = \emptyset$, and we ensure that $|W| \leq n^{3/4}$ throughout. Therefore, at any time the process is conditioned by the knowledge of partial matchings $M'_{i,j}$ between the fibers F_i and F_j

for $i \neq j$. (The matchings $M_{i,i+1}$ have, of course, been completely exposed in Phase 1.) For $j \neq i, i+1$, the unexposed part of $M_{i,j}$ will be a uniform extension of $M'_{i,j}$. So, in particular, when we examine a vertex $v_{i,k} \notin W$, the $M_{i,j}$ edge incident with $v_{i,k}$ is chosen uniformly from a set of size $n - o(n)$.

We now give an informal description of a process which removes a small cycle C from a *current* permutation digraph Σ . We break this process into an *out-phase* and an *in-phase*. We start by choosing an (arbitrary) edge (v_{i,j_0}, v_{i+1,k_0}) of C and delete it to obtain an NPD Γ_0 with $P_0 = P(\Gamma_0) \in \mathcal{P}(v_{i+1,k_0}, v_{i,j_0})$, where $\mathcal{P}(x, y)$ denotes the set of paths from x to y in D . The aim of this process is to produce a *large* set S of NPDs such that for each $\Gamma \in S$, (i) $P(\Gamma)$ has at least n_0 edges and (ii) the small cycles of $PD(\Gamma)$ are a strict subset of the small cycles of Σ . We will show that **whp** the endpoints of one of the $P(\Gamma)$'s can be joined by an edge to create a permutation digraph with (at least) one less small cycle. This completes the informal description.

We now give a fairly formal description, but we leave out some details for later. We produce a sequence $S_0 = \{\Gamma_0\}, S_1, S_2, \dots, S_t, \dots$ of sets of NPDs. Fix $t > 0$ and $\Gamma \in S_{t-1}$, and let v be the terminal endpoint of $P(\Gamma)$. We examine the $h-2$ edges of E_0 leaving v , i.e., the edges going *out* from the end of the path. Let w_j , $1 \leq j \leq h-2$, be the terminal vertices of these edges, and assume that Γ contains edges (x_j, w_j) , $1 \leq j \leq h-2$. Then for $1 \leq j \leq h-2$, $\Gamma_j = \Gamma \cup \{(v, w_j)\} \setminus \{(x_j, w_j)\}$ is added to S_t , assuming that the edge (v, w_j) is *acceptable* w.r.t. Γ . We call this an acceptable *out-step*. An (v, w) is acceptable if the following is true: Suppose that $P(\Gamma) \in \mathcal{P}(\cdot, v)$. Let $\Gamma' = \Gamma \cup \{(v, w)\} \setminus \{(x, w)\}$, where $(x, w) \in E(\Gamma)$. We then say that we *use* w .

- (i) $P(\Gamma')$ contains at least n_0 edges.
- (ii) Any new cycle created (i.e., in Γ' and not in Γ) also has at least n_0 edges.
- (iii) w does not lie on a small cycle of Γ .
- (iv) $P(\Gamma') \in \mathcal{P}(\cdot, x)$, where x has not been used before in Phase 2.

If Γ_j contains no edge (x_j, w_j) , then $w_j = v_{i+1,k_0}$. We accept the edge if $P(\Gamma_j)$ has at least n_0 edges. This would create a permutation digraph (PD) and (prematurely) end an iteration, although it is unlikely to occur.

Each member of S_{t-1} (usually) has $h-2$ descendants in S_t , and in this way we build a tree T_0 of NPDs in a natural breadth-first fashion where each nonleaf Γ at depth t is an element of S_t . By construction, all paths $P(\Gamma)$, $\Gamma \in T_0$, will have the same start vertex, viz. the head v_{i+1,k_0} of the edge deleted from the small cycle C . The construction of T_0 ends when we first have $\nu = \lceil \sqrt{n \log n} \rceil$ leaves. The construction of T_0 constitutes an out-phase of our procedure to eliminate small cycles. Having constructed T_0 we need to do a further in-phase, which is similar to a set of out-phases.

Then **whp** we close at least one of the paths $P(\Gamma)$ to a cycle of length at least n_0 . If this process fails, then we try again with a different independent edge of C in place of (v_{i,j_0}, v_{i+1,k_0}) . If we succeed, then we move on to the next small cycle.

We now fill in the details. We start Phase 2 with a PD Σ_0 , and a general iteration of Phase 2 starts with a PD Σ whose small cycles are a subset of those in Σ_0 . Iterations continue until there are no more small cycles. At the start of an iteration we choose some small cycle C of Σ . There then follows an out-phase in which we construct a tree $T_0 = T_0(\Sigma, C)$ of NPDs as follows: The root of T_0 is Γ_0 which is obtained by deleting an edge (v_{i,j_0}, v_{i+1,k_0}) of C .

We grow T_0 to a depth $O(\log n)$. The set of nodes at depth t will be S_t . $\Gamma \in S_{t-1}$ with $P = P(\Gamma) \in \mathcal{P}(v_{i,j_0}, v)$, $v \in F_i$, has up to $h-2$ descendants in S_t .

LEMMA 2. Let $C \in \text{SMALL}$. Then, where $\nu = \lceil \sqrt{n \log n} \rceil$,

$$\Pr\left(\exists t < \lceil \log_{h/2} \nu + 500 \log \log n \rceil \text{ such that } |S_t| \in [\nu, h\nu]\right) = 1 - O((\log \log n)^3 / \log n).$$

Proof. We assume that we stop construction of T_0 , in midphase if necessary, when $|S_t| \in [\nu, h\nu]$. Let us consider a generic construction in the growth of T_0 .

For an NPD Γ with $P(\Gamma) \in \mathcal{P}(v_{i,j_0}, v)$ we let $Z_j(\Gamma)$ be the 0-1 indicator for the edge (v, w_j) being unacceptable. If $Z_j(v) = 1$, then either (a) w_j lies on $P(\Gamma)$ and is too close to an endpoint—this has probability bounded above by $201h^3 / \log n$ —or (b) the corresponding vertex x_j is in W —this has probability bounded above by $2n^{-1/4}$ —or (c) w_j lies on a small cycle of Γ and hence of Σ_0 ; this has probability bounded above by $\log \log n / 3 \log n$. Then $\Pr(Z_j(\Gamma) = 1) \leq \log \log n / 2 \log n$ regardless of the history of the process to this point.

Let $Z_t = \sum_{\Gamma \in S_{t-1}} \sum_j Z_j(\Gamma)$. Z_t is the sum of possibly dependent random variables, but it is stochastically dominated by the binomial $B((h-2)|S_t|, \log \log n / \log n)$.

We write

$$|S_{t+1}| = (h-2)|S_t| - Z_t.$$

Now let $t_0 = \lceil 1000 \log \log n \rceil$ and $t_1 = \lceil \log_{h/2} \nu + 1000 \log \log n \rceil$.

- (a) $\Pr(\exists t \leq t_0 : |S_t| \leq 500 \log \log n \text{ and } Z_t > 0) = O((\log \log n)^3 / \log n)$.
- (b) $\Pr(\exists t \leq t_1 : |S_t| \geq 500 \log \log n \text{ and } Z_t > h|S_t|/100) \leq (\log n)^{-\Omega(\log \log n)}$.
- (a) $\Pr(Z_t > 0 \mid |S_t| \leq 500 \log \log n) = O((\log \log n)^2 / \log n)$ by the Markov inequality.
- (b) Immediate from (2).

Let $\mathcal{E}_a$ and $\mathcal{E}_b$ be the low probability events described in (a) and (b) above. Assume the occurrence of $\bar{\mathcal{E}}_a \cap \bar{\mathcal{E}}_b$. $\bar{\mathcal{E}}_a$ implies that $|S_t|$ reaches a size of at least $500 \log \log n$ before t reaches t_0 . Once this happens, $\bar{\mathcal{E}}_b$ implies that $|S_t|$ then grows geometrically with t up to time t_1 at a rate of at least $h/2$. The lemma follows. $\square$

The total number of vertices added to W in this way throughout the whole of Phase 2 is $O(\nu |\text{SMALL}|) = o(n^{3/4})$. We try this process once or twice for each $C \in \text{SMALL}$. Let t^* denote the value of t when we stop the growth of T_0 . At this stage we have leaves Γ_k , for $k = 1, \dots, \nu$, each with a path of length at least n_0 (unless we have already successfully made a cycle).

We now execute an in-phase. This involves the construction of trees $T_k, k = 1, 2, \dots, \nu$. The reader may while reading become concerned that in building ν trees of size ν we will need to use too many vertices. This will not be the case. This is a consequence of (3) later. Assume that $P(\Gamma_k) \in \mathcal{P}(v_{i+1,k_0}, v_{i,\ell_k})$. Notice that the start vertex of each of these paths is the same, viz. the head v_{i+1,k_0} of the edge deleted from the small cycle C . We start with Γ_k and build T_k in a similar way to T_0 except that here all paths generated end with v_{i,ℓ_k} . This is done as follows: If a current NPD Γ has $P(\Gamma) \in \mathcal{P}(u, v_{i,\ell_k})$, then we consider adding an edge $(w, u) \in E_0$ and deleting an edge $(w, x) \in \Gamma$. Thus our trees are grown by considering edges directed into the start vertex of each $P(\Gamma)$ rather than directed out of the end vertex. Some technical changes are necessary, however. We consider the construction of our ν trees in two stages. First of all we grow the trees only enforcing (an in-analogue of) condition (iv) of acceptability and thus allow the formation of small cycles and paths. We try to grow them to depth t_1 . The growth of the ν trees can naturally be considered to occur simultaneously. Let $L_{k,\ell}$ denote the set of start vertices of the paths associated

with the nodes at depth ℓ of the k th tree, $k = 1, 2, \dots, \nu$, $\ell = 0, 1, \dots, t_1$. Thus $L_{k,0} = \{v_{i+1,k_0}\}$ for all k . We prove inductively that

$$(3) \quad L_{k,\ell} = L_{1,\ell} \text{ for all } k, \ell.$$

In fact if $L_{k,\ell} = L_{1,\ell}$, then the acceptable E_0 edges have the same set of initial vertices, and since all of the deleted edges are Σ_0 -edges (enforced by (iv)) we have $L_{k,\ell+1} = L_{1,\ell+1}$. Note that the number of nodes in each tree is $O(h^{t_1+1}) = O(n^{3/5})$ if h is sufficiently large. Although we grow many trees, because of (3), the actual number of vertices used in total is $O(n^{3/5})$, and this is why we can claim that **whp** $|W| \leq n^{3/4}$ throughout.

The probability that we succeed in constructing trees $T_1, T_2, \dots, T_\nu$ is, by the analysis of Lemma 3, $1 - O((\log \log n)^3 / \log n)$.

We now consider the fact that in some of the trees some of the leaves may have been constructed in violation of (i)–(iii). We imagine that we prune the trees $T_1, T_2, \dots, T_\nu$ by disallowing any node that was constructed in violation of (i)–(iii). Let a tree be BAD if after pruning it has less than ν leaves and is GOOD otherwise. Now an individually pruned tree has been constructed in the same manner as the tree T_0 obtained in the out-phase. (We have chosen t_1 to obtain ν leaves even at the slowest growth rate of $h/2$ per node as asked for at the end of the proof of Lemma 2.) Thus

$$\Pr(T_1 \text{ is BAD}) = O\left(\frac{(\log \log n)^3}{\log n}\right)$$

and

$$\mathbf{E}(\text{number of BAD trees}) = O\left(\frac{\nu(\log \log n)^3}{\log n}\right)$$

and

$$\Pr(\exists \geq \nu/2 \text{ BAD trees}) = O\left(\frac{(\log \log n)^3}{\log n}\right).$$

Thus

$$\begin{aligned} \Pr(\exists < \nu/2 \text{ GOOD trees after pruning}) \\ \leq \Pr(\text{failure to construct } T_1, T_2, \dots, T_\nu) + \Pr(\exists \geq \nu/2 \text{ BAD trees}) \\ = O\left(\frac{(\log \log n)^3}{\log n}\right). \end{aligned}$$

Thus with probability $1 - O((\log \log n)^3 / \log n)$ we end up with $\nu/2$ sets of ν paths, each of length at least $100nh^3 / \log n$. All paths in one set have the same terminal vertex. Suppose that $v \in F_i$ is the common terminal vertex of one set. If a path in this set begins with a vertex in F_{i+1} , then we attempt to change the fiber of its start vertex to F_k , $k \neq i+1$, by performing an acceptable in-step. We will succeed with probability $1 - o(1)$, given the previous history. Thus assume that we have $\nu/2$ sets of $\nu/2$ paths where if the initial vertex is in fiber F_i , then the terminal vertex is not

in fiber F_{i-1} . Given this,

$$\begin{aligned} \Pr(\text{no } E_0 \text{ edge closes one of these paths}) &\leq \left(1 - \frac{\nu}{2n(1 - o(1))}\right)^{\nu/2} \\ &= O(n^{-1/4}). \end{aligned}$$

Consequently, the probability that we fail to eliminate a particular small cycle C after breaking an edge is $O((\log \log n)^3 / \log n)$. For every $C \in \text{SMALL}$, we have $|C| \geq h$. Therefore, it is possible to try once or twice using independent edges of C , and so the probability that we fail to eliminate a given small cycle C is certainly $O((\log \log n)^6 / (\log n)^2)$ (remember that we calculated all probabilities conditional on previous outcomes and assuming $|W| \leq n^{3/4}$). Hence, since **whp** $|C| = O(\log n)$, we have the following lemma.

LEMMA 3. *The probability that Phase 2 fails to produce a PD with minimal cycle length at least n_0 is $o(1)$.*

At this stage we have shown that if h is sufficiently large, then D almost always contains a PD Σ^* in which the minimum cycle length is at least n_0 .

We shall refer to Σ^* as the *Phase 2 PD*.

Now all the cycles of the PD Σ_0 defined in Phase 1 have lengths divisible by h . Also, as we traverse a cycle the fibers encountered are $F_1, F_2, \dots, F_h, F_1, \dots$. Now Σ^* is obtained from Σ_0 by replacing $O((\log n)^2)$ edges. By considering the unbroken sections of Σ_0 which pass through the fibers in order, we see that for each i , a cycle of length ℓ in Σ^* contains at least $\ell/h - O((\log n)^2)$ edges from fiber F_i to F_{i+1} .

4. Deterministic problem. In this section we give a lower bound to the number of a certain type of Hamilton cycle in a digraph with large minimum in-degree and out-degree.

Let Γ be a digraph with vertex set $[m]$ and minimum in-degree and out-degree at least $.99m$.

For a permutation ρ of $[m]$, let PD_ρ be its associated PD. Let

$$\begin{aligned} T_\Gamma &= \{\rho : PD_\rho \text{ is a subgraph of } \Gamma\}, \\ S_\Gamma &= \{\rho \in T_\Gamma : \rho \text{ is cyclic}\}. \end{aligned}$$

Let ϕ be a fixed *even* permutation of $[m]$. Let

$$R_\Gamma = \{\rho \in S_\Gamma : \phi\rho \text{ is cyclic}\}.$$

The permutations ρ thus correspond to a restricted class of Hamilton cycles in Γ .

In the next few sections we will prove the following theorem.

THEOREM 2. $|R_\Gamma| \geq m!e^{-3m}$.

This will help us resolve a strange technical problem, already met in [6]. It facilitates a second moment calculation.

The proof is quite long and is deferred to a later section, so as not to interrupt the flow of the probabilistic part of the argument.

5. Second moment calculation. Let $C_1, C_2, \dots, C_k$ be the cycles of Σ^* produced by Phase 2, and let $c_i^* = \min_j c_{ij}$, where c_{ij} is the number of *clean* edges in cycle C_i from fiber F_j to F_{j+1} for $j = 1, 2, \dots, h$. An edge is clean if it is not incident with W . Recall that W is the set of vertices v for which Phase 2 exposed an edge incident with v . E_1 will denote the set of clean edges that do not join fibers F_i, F_{i+1} for some i .

The cycles are numbered so that $c_1^* \leq c_2^* \leq \dots \leq c_k^*$ and $c_1^* \geq n_0/h - 2n^{3/4} \geq \frac{99h^2n}{\log n}$. If $k = 1$, there is nothing more to do. Otherwise let $a = \lceil \frac{nh^2}{\log n} \rceil$. We will show that **whp** it is possible to delete a set of edges from each C_i and then replace them so that the resulting structure is a Hamilton cycle. We will use the second moment method to do this.

We select an odd number of edges from each C_i and delete them. The choice of parity is related to the need to keep permutation ϕ (defined next) even. See (4). We will then be able to apply Theorem 2. Continuing, when h is odd, we choose the edges in the following manner. For each C_i , we select a set of $l_i = 2\lfloor \frac{c_i^*}{a} \rfloor + 1$ vertices $v \in C_i \setminus W$ from each fiber F_j , where $j = 1, 2, \dots, h$, and delete the corresponding edges (v, u) in Σ^* . The number of edges deleted from cycle C_i is $m_i = l_i h$, which is odd. Since $v \in C_i \setminus W$, the deleted edge $(v, u) \in \Sigma_0$ is an edge between fibers F_j and F_{j+1} for some j . The above procedure is not acceptable when h is even as we would end up deleting an even number of edges from each cycle. We circumvent this problem by choosing $l_i - 1$ vertices from fiber F_j , where $j \equiv i \pmod{h}$ and l_i from the rest of the fibers for cycle C_i . This ensures that the number of edges deleted from cycle C_i is $m_i = l_i h - 1$, which is odd.

Let $m = \sum_{i=1}^k m_i$. In each cycle C_i choose a vertex x_i which loses a cycle edge directed out of it. Let $v_1 = x_1$ and then go around C_1 defining $v_2, v_3, \dots, v_{m_1}$ in order as the endpoints of the path sections. Then let $v_{m_1+1} = x_2$ and so on. Now relabel the broken edges as $(v_i, u_i), i \in [m]$. We thus have m path sections $P_j \in \mathcal{P}(u_{\phi(j)}, v_j)$ in Σ^* for permutation ϕ , where $\phi(1) = m_1, \phi(2) = 1, \dots, \phi(m_1) = 1, \phi(m_1 + 1) = m_1 + m_2, \phi(m_1 + 2) = m_1 + 1$, etc. (Some path sections could be just a single vertex.)

Note that since ϕ is made up of cycles of odd length,

$$(4) \quad \phi \text{ is an even permutation of } [m].$$

The number of paths starting (ending) at any of the fibers is either $\lfloor \frac{m}{h} \rfloor$ or $\lfloor \frac{m}{h} \rfloor + 1$.

We will attempt to rejoin these paths in a different order so that a Hamilton cycle is constructed. Suppose that path section P starts in fiber $F_\xi(P)$ and ends in fiber $F_\eta(P)$. Because we do not wish to put back edges that we have just deleted: If P immediately precedes Q in some reordering, then we will have $\xi(Q) \neq \eta(P), \eta(P) + 1$. We write $P \rightarrow Q$ if this holds for P, Q .

Fix, for now, a choice of $P_1, P_2, \dots, P_m$, and let Γ be the digraph with vertex set $[m]$ and a directed edge (i, j) whenever $P_i \rightarrow P_j$. We are interested in the number of Hamilton cycles in Γ that satisfy a certain property. Each vertex of Γ has an in-degree and out-degree of at least $m - 2\lfloor \frac{m}{h} \rfloor - 2$. By choosing h to be suitably large, we can assume that $m - 2\lfloor \frac{m}{h} \rfloor - 2 \geq .99m$. Thus, Γ satisfies the conditions of Theorem 2.

Let Ω denote the set of ordered pairs of selections of edges for deletion and cycle rearrangements ρ satisfying the condition that a path ending on fiber F_i is not joined to a path starting on the same fiber or F_{i+1} and such that $\lambda = \phi\rho$ is cyclic. $\omega \in \Omega$ is a *success* if E_1 contains the edges needed for the associated Hamilton cycle.

Let H stand for the union of the PD Σ^* and E_1 . We finish our proof by proving the following lemma.

LEMMA 4. $\mathbf{Pr}(H \text{ does not contain a Hamilton cycle}) = o(1)$.

Proof. Let X be the number of Hamilton cycles in H obtainable by deleting edges as above, rearranging the path sections generated by ϕ , and, if possible, reconnecting all of the sections using edges of E_1 . We can think of this as the sum of indicator

variables indexed by Ω . We will use the well-known inequality

$$(5) \quad \Pr(X > 0) \geq \frac{\mathbf{E}(X)^2}{\mathbf{E}(X^2)}.$$

Probabilities in (5) are thus, with respect to the space of E_1 -choices for edges incident with vertices not in W .

Now the definition of l_i yields that

$$\frac{2nh - |W|h}{a} - (h+1)k \leq m \leq \frac{2nh}{a} + hk$$

and so

$$\frac{1.98}{h} \log n \leq m \leq \frac{2.01}{h} \log n.$$

Also

$$k \leq \log n / 100h^3, \quad l_i \geq 199, \quad \text{and} \quad \frac{c_i^*}{l_i} \geq \frac{a}{2.02}, \quad 1 \leq i \leq k.$$

Now fix a set of m paths and a permutation $\omega \in \Omega$. The probability that the edges exist for a success is at least n^{-m} . For having conditioned on the existence of a set of edges A , the probability that edge (u, v) exists is $1/\mu$, where $\mu \leq n$ is the number of vertices in the fiber containing v which are not incident with an A -edge whose other endpoint is in the same fiber as u . Recall that up to this point all we have done is condition on certain edges being present. The remaining edges form random partial matchings between the fibers. Furthermore, we have only conditioned on the presence of $O(n^{3/4})$ edges in total.

Let $\theta = 1$ if h is even and $\theta_{ij} = \theta \times 1_{i \equiv j \pmod{h}}$.

$$\begin{aligned} \mathbf{E}(X) &= \sum_{\omega \in \Omega} \Pr(\omega \text{ is a success}) \\ &\geq \sum_{\omega \in \Omega} n^{-m} \\ &\geq n^{-m} m! e^{-3m} \prod_{i=1}^k \binom{c_i^*}{l_i - \theta} \prod_{\substack{j=1 \\ j \neq i}}^h \binom{c_i^*}{l_i} \\ &\geq n^{-m} m! e^{-3m} \prod_{i=1}^k \binom{c_i^*}{l_i}^h \left(\frac{l_i}{c_i^*} \right)^\theta, \quad \text{after using Theorem 2,} \\ &\geq e^{-3m} \left(\frac{m}{en} \right)^m \prod_{i=1}^k \left(\left(\frac{c_i^* e^{1-1/12l_i}}{l_i^{1+(1/2l_i)}} \right)^{l_i} \left(\frac{1-2l_i^2/c_i^*}{\sqrt{2\pi}} \right)^h \right) \left(\frac{l_i}{c_i^*} \right)^\theta \\ &\geq e^{-3m} \left(\frac{m}{en} \right)^m \prod_{i=1}^k \left(\frac{e}{1.01} \right)^{l_i h} \left(\frac{c_i^*}{l_i} \right)^{l_i h - \theta} \\ &= e^{-3m} \left(\frac{mea}{(1.01)(2.02)en} \right)^m \\ &\geq e^{-3m} (.9h)^m \\ (6) \quad &\geq n^{\Omega(1)}. \end{aligned}$$

Let M, M' be two sets of selected edges which have been deleted in Σ^* and whose path sections have been rearranged into Hamilton cycles according to ρ, ρ' , respectively. Let N, N' be the corresponding sets of edges which have been added to make the Hamilton cycles. What is the interaction between these two Hamilton cycles?

Let $s = |M \cap M'|$ and $t = |N \cap N'|$. Now $t \leq s$ since if $(v, u) \in N \cap N'$, then there must be a unique $(\tilde{v}, u) \in M \cap M'$ which is the unique Σ^* -edge into u .

We claim that $t = s > 0$ implies $t = s = m$ and $(M, \rho) = (M', \rho')$. (This is why we have restricted our attention to $\rho \in R_\Gamma$.) For the following argument recall that we delete edges (v_i, u_i) , $i \in [m]$, and our path segments go from $u_{\phi(i)}$ to v_i . In our reordering, v_i will be connected to $u_{\phi\rho(i)} = u_{\lambda(p)}$. Suppose then that $t = s > 0$ and $(v_i, u_i) \in M \cap M'$. Now the edge $(v_i, u_{\lambda(i)}) \in N$, and since $t = s$ this edge must also be in N' . But this implies that $(v_{\lambda(i)}, u_{\lambda(i)}) \in M'$ and hence in $M \cap M'$. Repeating the argument we see that $(v_{\lambda^k(i)}, u_{\lambda^k(i)}) \in M \cap M'$ for all $k \geq 0$. But λ is cyclic, and so our claim follows.

We adopt the following notation. Let $\langle s, t \rangle$ denote $|M \cap M'| = s$ and $|N \cap N'| = t$. So

$$\begin{aligned} \mathbf{E}(X^2) &\leq \mathbf{E}(X) + (1 + O(m|W|/n)) \sum_{\omega \in \Omega} n^{-m} \sum_{\omega': N' \cap N = \emptyset} n^{-m} \\ &\quad + (1 + O(m|W|/n)) \sum_{\omega \in \Omega} n^{-m} \sum_{s=2}^m \sum_{t=1}^{s-1} \sum_{\omega': \langle s, t \rangle} n^{t-m} \\ (7) \quad &= \mathbf{E}(X) + (1 + o(1))(A_1 + A_0) \text{ say.} \end{aligned}$$

Clearly

$$(8) \quad A_1 \leq \left(\sum_{\omega \in \Omega} n^{-m} \right)^2 \leq \mathbf{E}(X)^2.$$

For given M, M', ρ , how many ρ' satisfy the condition $\langle s, t \rangle$? We bound it from above by $(m-t)!$ (consider fixing t edges of Λ'). Thus

$$\begin{aligned} A_0 &\leq \sum_{\omega \in \Omega} n^{-m} \sum_{s=2}^m \sum_{t=1}^{s-1} \left[\sum_{\sigma_{11} + \dots + \sigma_{kh} = s} \prod_{i=1}^k \prod_{j=1}^h \binom{l_i - \theta_{ij}}{\sigma_{ij}} \binom{c_i^* - l_i}{l_i - \sigma_{ij}} \right] (m-t)! n^{t-m} \\ &\leq \mathbf{E}(X)^2 \sum_{s=2}^m \sum_{t=1}^{s-1} \binom{s}{t} \left[\sum_{\sigma_{11} + \dots + \sigma_{kh} = s} \prod_{i=1}^k \prod_{j=1}^h \frac{\binom{l_i - \theta_{ij}}{\sigma_{ij}} \binom{c_i^* - l_i}{l_i - \sigma_{ij}}}{\binom{c_i^*}{l_i}} \right] \frac{(m-t)! e^{3m}}{m!} n^t. \end{aligned}$$

Now

$$\frac{\binom{c_i^* - l_i}{l_i - \sigma_{ij}}}{\binom{c_i^*}{l_i}} \leq \frac{\binom{c_i^*}{l_i - \sigma_{ij}}}{\binom{c_i^*}{l_i}} \leq (1 + o(1)) \left(\frac{l_i}{c_i^*} \right)^{\sigma_{ij}} \leq (1 + o(1)) \left(\frac{2.02}{a} \right)^{\sigma_{ij}},$$

where the $o(1)$ term is $O((\log n)^3/n)$. Also

$$\sum_{\sigma_{11} + \dots + \sigma_{kh} = s} \prod_{i=1}^k \prod_{j=1}^h \binom{l_i - \theta_{ij}}{\sigma_{ij}} = \binom{m}{s}.$$

Hence,

$$\begin{aligned}
\frac{A_0}{\mathbf{E}(X)^2} &\leq (1 + o(1)) \sum_{s=2}^m \sum_{t=1}^{s-1} \binom{s}{t} \left(\frac{2.02}{a}\right)^s \binom{m}{s} \frac{(m-t)! e^{3m} n^t}{m!} \\
&\leq n^{.01} \sum_{s=2}^m \sum_{t=1}^{s-1} \binom{s}{t} \left(\frac{2.02}{a}\right)^s \frac{m^{s-t} n^t}{s!}, \quad \text{using } e^{3m} \leq n^{.005} \text{ for } h \text{ large,} \\
&= n^{.01} \sum_{s=2}^m \left(\frac{2.02}{a}\right)^s \frac{m^s}{s!} \sum_{t=1}^{s-1} \binom{s}{t} \left(\frac{n}{m}\right)^t \\
&\leq \frac{2m}{n^{.99}} \sum_{s=2}^m \left(\frac{(2.02)n}{a}\right)^s \frac{1}{s!} \\
&\leq \frac{2m}{n^{.99}} \sum_{s=2}^m \left(\frac{2.02 \log n}{h^2}\right)^s \frac{1}{s!} \\
&\leq \frac{2m}{n^{.99}} n^{2.02/h^2} \\
(9) \quad &= o(1).
\end{aligned}$$

The lemma follows from (5) to (9).

This completes the proof of Theorem 1. $\square$

6. Proof of Theorem 2.

6.1. Lower bound for $|T_\Gamma|$.

LEMMA 5. $|T_\Gamma| \geq m! e^{-5m/2}$.

Proof. Consider a bipartite graph $B = (V_1 \cup V_2, E)$, where V_1, V_2 are disjoint copies of $[m]$. The edge set $E(B) = \{(i, j) : (i, j) \in E(\Gamma)\}$. It follows from the above that the degree of each vertex is at least $\delta_B = .99m$. A perfect matching M in B gives rise to a member ρ of T_Γ , where $j = \rho(i)$ iff $(i, j) \in M$. For a matching M of $K_{m,m}$, let $\nu(M)$ be the number of edges of M which are also in B . Let $\mathcal{M}_k$ denote the set of perfect matchings of $K_{m,m}$ with $\nu(M) = k$, $k = 0, 1, 2, \dots, m$. We first prove that if $\mu_k = |\mathcal{M}_k|$, $k = 0, 1, \dots, m$, and $k \geq .26m + 1$, then

$$(10) \quad \frac{\mu_{k+1}}{\mu_k} \geq \frac{.23m(m-k)}{\binom{m}{2}} \geq \frac{.46(m-k)}{m}.$$

Consider the set $\mathcal{X}$ of ordered pairs (M_1, M_2) , where $M_1 \in \mathcal{M}_k$, $M_2 \in \mathcal{M}_{k+1}$, and the symmetric difference $M_1 \oplus M_2$ is an alternating cycle of length 4. Now each $M \in \mathcal{M}_k$ is in at least

$$(m-k)(2(\delta_B - 1) - (m-1) - (m-k)) \geq .23m(m-k)$$

such cycles.

Explanation. There are $m-k$ choices for an edge $e = (v, w) \in M \setminus E(B)$. Then there are at least $2(\delta_B - 1) - (m-1)$ pairs of vertices (a, b) such that $(a, b) \in M$ and $(v, b) \in E(B)$, $(a, w) \in E(B)$, and at least $2(\delta_B - 1) - (m-1) - (m-k)$ pairs $(a, b) \in M \cap E(B)$, in which case $M' = (M \cup (a, w), (v, b)) \setminus ((v, w), (a, b))$ is a member of $\mathcal{M}_{k+1}$ and $(M, M') \in \mathcal{X}$.

On the other hand, each M is in at most $\binom{m}{2}$ pairs and (10) follows.

Now, if $k \leq .26m$, then

$$(11) \quad \frac{\mu_{k+2}}{\mu_k} \geq \frac{.35m(m-k)}{\binom{m}{2}} \geq \frac{.7(m-k)}{m} \geq \frac{.7(m-k)(m-(k+1))}{m^2}.$$

Explanation. As in the previous case, there are $m-k$ choices for an edge $e = (v, w) \in M \setminus E(B)$. Then there are at least $2(\delta_B - 1) - (m-1)$ pairs of vertices (a, b) such that $(a, b) \in M$ and $(v, b) \in E(B)$, $(a, w) \in E(B)$, and at least $2(\delta_B - 1) - (m-1) - k \geq .72m - 1$ pairs $(a, b) \in M \setminus E(B)$, in which case $M' = (M \cup (a, w), (v, b)) \setminus ((v, w), (a, b))$ is a member of $\mathcal{M}_{k+2}$ and $(M, M') \in \mathcal{X}$. Now, every such pair of edges $\{(v, w), (a, b)\}$ gets counted twice. Hence, the actual number of pairs is at least $.35m$ and (11) follows.

It follows that $\mu_m \geq (.46)^k m^{-k} k! \mu_{m-k}$ for $k \geq 0$ and so

$$m! = \mu_0 + \mu_1 + \cdots + \mu_m \leq \mu_m \sum_{k=0}^m \frac{m^k}{(.46)^k k!} \leq \mu_m e^{m/.46}. \quad \square$$

6.2. Upper bound for $|T_\Gamma|/|S_\Gamma|$. We will now extend an approach of Dyer, Frieze, and Jerrum [9] to directed graphs.

Let $k^* = \lfloor 7 \ln m \rfloor$, and for $1 \leq k \leq m/2$, define $g(k) = m^7 k! (7 \ln m)^{-k}$ and

$$f(k) = \begin{cases} g(k) & \text{if } k \leq k^*, \\ g(k^*) & \text{otherwise.} \end{cases}$$

LEMMA 6. *Let f be the function defined above. Then*

(a) *f is nonincreasing and satisfies*

$$f(k-1) \geq 7k^{-1} f(k) \ln m;$$

(b) *$f(k) \geq 1$ for all k .*

Proof. Observe that g is unimodal and k^* is the value of k minimizing $g(k)$; it follows that f is nonincreasing. When $k \leq k^*$, we have $f(k-1) = g(k-1) = (7 \ln m) k^{-1} g(k) = (7 \ln m) k^{-1} f(k)$. Otherwise, $f(k-1) = g(k^*) = f(k) \geq (7 \ln m) k^{-1} f(k)$. In either case, the inequality in part (a) of the lemma holds.

Part (b) of the lemma follows from the chain of inequalities.

$$\frac{1}{f(k)} \leq \frac{1}{g(k^*)} = \frac{(7 \ln m)^{k^*}}{m^7 k^*!} \leq m^{-7} \sum_{k=0}^{\infty} \frac{(7 \ln m)^k}{k!} = m^{-7} \exp(7 \ln m) = 1.$$

LEMMA 7. $|S_\Gamma| \geq m^{-8} |T_\Gamma|$.

Proof. For $1 \leq k \leq \lfloor m/2 \rfloor$, let Φ_k be the set of all PDs in Γ containing exactly k cycles, and let $\Phi = \cup_k \Phi_k$ be the set of all PDs. Define

$$\Psi = \{(F, F') : F \in \Phi_k, F' \in \Phi_{k-1} \text{ and } F \oplus F' \cong C_0\},$$

where $\oplus$ denotes symmetric difference and C_0 is the directed graph on four vertices with two vertices having in-degree two, out-degree zero, and two vertices having

out-degree two, in-degree zero. Observe that if $(F, F') \in \Psi$, then F' can be obtained from F by deleting two edges and adding two edges; this operation reduces the number of cycles by exactly one.

Our proof strategy is to define a positive weight function on the arc set Ψ such that the total weight of arcs leaving each node (PD) $F \in \Phi \setminus \Phi_1$ is at least one greater than the total weight of arcs entering F . This will imply that the total weight of arcs entering Φ_1 is an upper bound on the number of non-Hamiltonian PDs in Γ and that the maximum total weight of arcs entering a single node in Φ_1 is an upper bound on the ratio $|\Phi \setminus \Phi_1|/|\Phi_1|$.

The weight function $w : \Psi \rightarrow R^+$ is defined as follows: For any arc (F, F') with $F' \in \Phi_k$, if the PD F' is obtained from F by coalescing two cycles of length γ_1 and γ_2 into a single cycle of length $\gamma_1 + \gamma_2$, then $w(F, F') = (\gamma_1^{-1} + \gamma_2^{-1})f(k)$.

Let $F \in \Phi_k$ be a PD with $k > 1$ cycles $C_1, C_2, \dots, C_k$ of lengths $\gamma_1, \gamma_2, \dots, \gamma_k$. Let us try to obtain a lower bound on the weight of the arcs going out of F . Suppose we chose cycle C_i to be one of the two cycles that coalesces with the other to form a new cycle. The number of ways to pick an edge from C_i is γ_i . Having picked an edge (u, v) from C_i , we need to find an edge (s, t) on the remaining cycles such that (u, t) and (s, v) are edges in the digraph Γ . Since the out-degree of u is at least $.99m$, we have at least $.99m - (\gamma_i - 1)$ possible choices for s . Similarly, we have at least $.99m - (\gamma_i - 1)$ possible choices for t . The number of feasible edges (s, t) is

$$\begin{aligned} &\geq .99m - (\gamma_i - 1) + .99m - (\gamma_i - 1) - (m - \gamma_i) \\ &> .98m - \gamma_i \end{aligned}$$

for each edge in C_i . The total number of ways to form a new cycle using C_i is thus at least $\gamma_i(.98m - \gamma_i)$. The weight of any arc leaving F is at least $\gamma_i^{-1}f(k-1)$, which, by Lemma 2, is bounded below by $(7 \ln m)(k\gamma_i)^{-1}f(k)$. We also have to note that every pair $(u, v), (s, t)$ may get counted twice, once from each cycle. Thus, the total weight of arcs leaving F is bounded as follows:

$$\begin{aligned} \sum_{F^+ : (F, F^+) \in \Psi} w(F, F^+) &\geq \frac{1}{2} \sum_{i=1}^k \gamma_i [.98m - \gamma_i] \frac{(7 \ln m)f(k)}{k\gamma_i} \\ &= \frac{1}{2} m [.98k - 1] \frac{(7 \ln m)f(k)}{k} \\ (12) \qquad &\geq 3f(k)m \ln m, \end{aligned}$$

where we have used the fact that $k \geq 2$.

We now give an upper bound to the weight of the arcs $(F^-, F) \in \Psi$ entering F . Suppose once again that F has cycles $C_1, C_2, \dots, C_k$ of lengths $\gamma_1, \gamma_2, \dots, \gamma_k$. A directed arc from F^- to F implies that one of the cycles of F was formed by coalescing two cycles of F^- . Suppose C_i was formed by coalescing two cycles from F^- . By removing the added edges from C_i and putting back the deleted edges, we can generate F^- . Let a and b be the lengths of the corresponding cycles in F^- . Remove one of the added edges from C_i and go around the cycle and remove the edge which is at a distance a . This way we make sure that the deletion of the first of the two added edges uniquely determines the two paths of lengths $a - 1$ and $b - 1$. Since the added edge in F could be any of the γ_i edges, we have at most γ_i choices.

The total weight of arcs entering F can thus be bounded previously as follows:

$$\begin{aligned}
 \sum_{F^-: (F^-, F) \in \Psi} w(F^-, F) &\leq \sum_{i=1}^k \gamma_i f(k) \sum_{\substack{a, b \geq 1 \\ a+b=\gamma_i}} \left(\frac{1}{a} + \frac{1}{b} \right) \\
 &= \sum_{i=1}^k \gamma_i f(k) \sum_{a=1}^{\gamma_i-1} \left(\frac{1}{a} + \frac{1}{\gamma_i-a} \right) \\
 (13) \qquad &\leq 2f(k)mH(m),
 \end{aligned}$$

where $H_m = \sum_{i=1}^m i^{-1} \leq \ln m + 1$ is the m th harmonic number. Combining inequalities (12) and (13), we have

$$\begin{aligned}
 \sum_{F^+: (F, F^+) \in \Psi} w(F, F^+) - \sum_{F^-: (F^-, F) \in \Psi} w(F^-, F) &\geq 3f(k)m \ln m - 2f(k)mH(m) \\
 &\geq f(k)m(\ln m - 2) \\
 &\geq m(\ln m - 2),
 \end{aligned}$$

where the final inequality is by Lemma 6. Thus the total weight of arcs leaving F exceeds the total weight of arcs entering by at least 1, provided $m \geq e^3$. The number of non-Hamiltonian PDs $|\Phi \setminus \Phi_1|$ is bounded above by the total weight of arcs entering Φ_1 , which in turn is bounded, see inequality (13), by $|\Phi_1| \times 2f(1)mH_m \leq m^8 |\Phi_1|/2$. $\square$

6.3. Lower bound on $|R_\Gamma|$. In this section, we will use an argument similar in flavor to the argument of the previous section to give a lower bound on $|R_\Gamma|$. Each $\rho \in S_\Gamma$ yields another permutation $\lambda = \lambda(\rho) = \phi\rho$. Recall that ϕ is considered to be fixed and $\rho \in R_\Gamma$ iff $\lambda = \phi\rho$ is cyclic.

We show next that if $\rho \in S_\Gamma$, then λ has an odd number of cycles. Now ϕ is even, and so λ and ρ have the same sign, $(-1)^{m-1}$. If λ has cycles $C_1, C_2, \dots, C_k$ of sizes $\gamma_1, \gamma_2, \dots, \gamma_k$, then the sign of λ is $(-1)^{\gamma_1-1+\dots+\gamma_k-1} = (-1)^{m-k} = (-1)^{m-1}$.

We let

$$k_0 = 2\lfloor 50001 \ln m \rfloor + 1,$$

and we will consider the following conditions for ρ : (i) λ must have at least k_0 cycles, (ii) the longest cycle in λ has length at most $\leq .8m$, and (iii) the sum of the lengths of the longest two cycles is at most $\leq .92m$.

The λ with less than k_0 cycles will be considered later in this section. Given (i), it is not clear whether or not we can find ρ satisfying conditions (ii) and (iii). Let S_k , k odd be the set of all $\rho \in S_\Gamma$ such that $\lambda(\rho)$ has $k \geq k_0$ cycles. We define a partition $S_k = P_k \cup Q_k \cup R_k$. Let P_k be the set of all ρ which do not satisfy condition (ii), Q_k be the set of those which satisfy condition (ii) but not (iii), and R_k be the set of those which satisfy both conditions (ii) and (iii).

We will first show that a constant fraction satisfying condition (i) also satisfies conditions (ii) and (iii).

6.3.1. $|P_k|/|Q_k \cup R_k| \leq 1/60$. We define $X_k = \{(\rho_1, \rho_2) : \rho_1 \in P_k, \rho_2 \in Q_k \cup R_k \text{ and } \Lambda_1 \oplus \Lambda_2 \cong H_0\}$, where $\Lambda_i = PD_{\lambda(\rho_i)}$, $i = 1, 2$, and $\oplus$ stands for symmetric

difference and H_0 is the directed graph on six vertices as in Figure 1(c). The weight function $w : X_k \rightarrow R^+$ is defined as follows: If Λ_2 is obtained by breaking and patching two cycles of Λ_1 of length γ_1 and γ_2 into two cycles of length γ' and γ'' , then the weight $w(\rho_1, \rho_2) = \gamma_1^{-1} + \gamma_2^{-1}$. We restrict X_k to those (ρ_1, ρ_2) for which $\gamma', \gamma'' \geq m/4$.

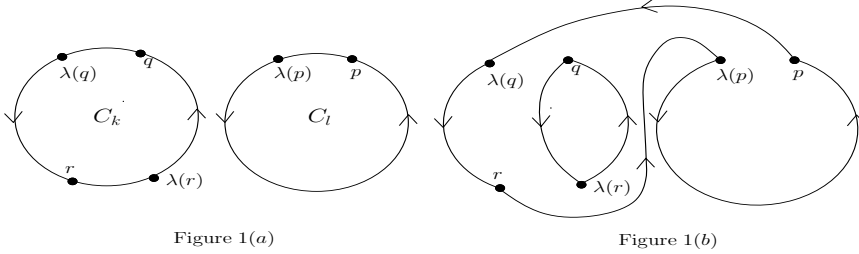


Figure 1(a)

Figure 1(b)

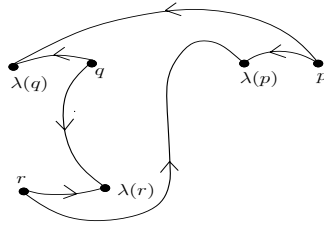


Figure 1(c)

FIG. 1.

Suppose $\rho_1 \in P_k$ and $C_1, C_2, \dots, C_k$ are the cycles of Λ_1 in increasing order of size. Let γ_t denote the length of cycle C_t for $t \in [k]$. We will combine C_k with a smaller cycle C_ℓ , $\ell \neq k$, to obtain two cycles (as in Figure 1(b)) such that the new auxiliary graph constructed, say $\Lambda_2 = \Lambda(\rho_2)$, belongs to $Q_k \cup R_k$. The weight of (ρ_1, ρ_2) is $\gamma_k^{-1} + \gamma_\ell^{-1}$. While combining the two cycles, we will have to ensure that ρ_2 is a cyclic permutation.

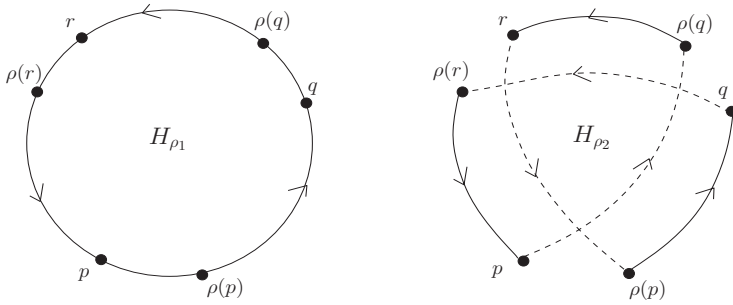


FIG. 2.

To create $\rho_2 \in Q_k \cup R_k$ we delete edges $(p, \lambda(p))$, $(q, \lambda(q))$, and $(r, \lambda(r))$, from Λ_1 , where p, q, r appear in this order on the Hamilton cycle H_{ρ_1} of Γ , corresponding to ρ_1 ,

and replace them with the edges $(p, \lambda(q))$, $(q, \lambda(r))$, and $(r, \lambda(p))$ as shown in Figure 2 to maintain Hamiltonicity. Vertices q and r are chosen on cycle C_k and vertex p is chosen on C_ℓ , $\ell \neq k$. We will of course have to impose the restriction on p, q , and r in this construction that $(p, \rho(q))$, $(q, \rho(r))$, and $(r, \rho(p))$ are all edges of Γ . There will be other necessary restrictions.

The number of choices for p is γ_ℓ . Given p we rule out at most $m - \gamma_k$ choices for q due to not being on C_k . Recall that C_k is the longest cycle of Λ_1 and it corresponds to $\rho_1 \in P_k$ and so $\gamma_k \geq .8m$. This leaves us with at least $.99m + .8m - m = .79m$ choices. Having chosen q we rule out $m/2$ choices for r within distance $\leq m/4$ from q on C_k . For each such choice, both of the two new cycles replacing C_k, C_ℓ are of size $\geq m/4$, placing ρ_2 in $Q_k \cup R_k$.

Some of these choices are inadmissible because we need to ensure that ρ_2 is a cyclic permutation. First, we will restrict our choice of q to one of the first $m/4$ vertices following p on H_{ρ_1} . We have at least $m/4 - .21m \geq .04m$ such choices. Now we must choose r from the remaining $3m/4$ vertices. We have ruled out $m/2$ already and also rule out further $\leq .02m$ choices of r for which r is not an in-neighbor of $\lambda(p)$ or $\lambda(r)$ is not an out-neighbor of q . This yields at least $(.75 + .8 - .5 - 1 - .02)m = .03m$ choices.

Thus

$$\begin{aligned}
 \sum_{\rho_2: (\rho_1, \rho_2) \in X_k} w(\rho_1, \rho_2) &\geq \sum_{\ell=1}^{k-1} \gamma_\ell [(.04m)(.03m)](\gamma_k^{-1} + \gamma_\ell^{-1}) \\
 &\geq .0012 \sum_{\ell=1}^{k-1} \gamma_\ell m^2 (m^{-1} + \gamma_\ell^{-1}) \\
 &= .0012 \sum_{\ell=1}^{k-1} [\gamma_\ell + m] m \\
 &= .0012 [m - \gamma_k + (k-1)m] m \\
 &\geq .0012 (k-1) m^2 \\
 &\geq 120 m^2 \ln m.
 \end{aligned}
 \tag{14}$$

We now obtain an upper bound on the total weight of pairs containing ρ_2 of $Q_k \cup R_k$. Suppose that $C'_1, C'_2, \dots, C'_k$ are the cycles of Λ_2 in increasing order of length $\gamma'_1 \leq \gamma'_2 \leq \dots \leq \gamma'_k$. Our choices for Λ_1 are restricted as follows: We must choose two cycles C'_r, C'_s such that $\gamma'_r + \gamma'_s = \gamma_k + \gamma_\ell \geq \gamma_k \geq .8m$ and $\gamma'_r, \gamma'_s \geq m/4$. This implies $\{r, s\} = \{k-1, k\}$. Given this, we see that the total weight of edges entering Λ_2 can be bounded as follows: The parameter a is the length of the path from q to r on C_k :

$$\begin{aligned}
 \sum_{\rho_1: (\rho_1, \rho_2) \in X_k} w(\rho_1, \rho_2) &\leq m^2 \max_{M \leq m} \left\{ \sum_{\substack{a+b=M \\ a, b \geq 2}} \left(\frac{1}{a} + \frac{1}{b} \right) \right\} \\
 &\leq m^2 \sum_{a=2}^m \frac{2}{a} \\
 &\leq 2m^2 \ln m.
 \end{aligned}
 \tag{15}$$

It follows that

$$(16) \quad 120m^2|P_k|\ln m \leq \sum_{(\rho_1, \rho_2) \in X_k} w(\rho_1, \rho_2) \leq 2m^2|Q_k \cup R_k|\ln m.$$

We have thus shown that

$$(17) \quad |P_k| \leq |Q_k \cup R_k|/60.$$

6.3.2. $|Q_k|/|R_k| \leq 1/200$. The argument that follows is similar to the one used previously to show that $|Q_k \cup R_k|$ is a constant fraction of $|S_k|$. Now let $Y_k = \{(\rho_1, \rho_2) : \rho_1 \in Q_k, \rho_2 \in R_k \text{ and } \Lambda_1 \oplus \Lambda_2 \cong H_0\}$, where H_0 is the same directed graph on six vertices as in Figure 1(c). The weight function $w : Y_k \rightarrow R^+$ is defined as follows: If Λ_2 is obtained by breaking and patching two cycles of Λ_1 of length γ_i and γ_j into two cycles of length γ' and γ'' , then $w(\rho_1, \rho_2) = \gamma_i^{-1} + \gamma_j^{-1}$. We again restrict our attention to (ρ_1, ρ_2) such that $\gamma', \gamma'' \geq m/10$.

Suppose $\rho_1 \in Q_k$ and $C_1, C_2, \dots, C_k$ are the cycles of Λ_1 in the order of increasing size $\gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_k$, where $\gamma_k \leq .8m$ and $\gamma_k + \gamma_{k-1} > .92m$. We will combine C_k with a smaller cycle C_ℓ , $\ell \notin \{k, k-1\}$, to obtain two cycles (as in Figure 1(b)) such that the new auxiliary graph constructed, Λ_2 , belongs to R_k .

While q and r are chosen from cycle C_k , p is chosen from C_ℓ . The number of choices for p is γ_ℓ . Since $\gamma_k + \gamma_{k-1} \geq .92m$, we have $\gamma_k \geq .46m$. The number of *feasible* choices for q and r is therefore at least $.99m + .46m - m = .45m$. Having chosen q we rule out $.2m$ choices for r within distance $\leq .1m$ from q on C_k . For each such choice, the size of the new largest cycle is at most $\max\{\gamma_{k-1}, \gamma_k + .08m - .1m\}$, and both of the two new cycles replacing C_k, C_ℓ are of size $\geq .1m$, placing Γ_2 in R_k .

Some of these choices are inadmissible because we need to ensure that ρ_2 is a cyclic permutation. First, we will restrict our choice of q so that q is one of the first $.1m$ feasible C_k -vertices following p on H_ρ . Now we must choose r from the remaining $\geq .35m$ feasible C_k -vertices. We have ruled out $.2m$ already and also rule out a further $\leq .02m$ choices of r for which r is not an in-neighbor of $\lambda(p)$ or $\lambda(r)$ is not an out-neighbor of q . This yields at least $.13m$ choices.

Thus,

$$(18) \quad \begin{aligned} \sum_{\rho_2: (\rho_1, \rho_2) \in Y_k} w(\rho_1, \rho_2) &\geq \sum_{\ell=1}^{k-2} \gamma_\ell [(.1m)(.13m)](\gamma_k^{-1} + \gamma_\ell^{-1}) \\ &\geq .013 \sum_{\ell=1}^{k-2} \gamma_\ell m^2 (m^{-1} + \gamma_\ell^{-1}) \\ &= .013 \sum_{\ell=1}^{k-2} [\gamma_\ell + m]m \\ &= .013[m - \gamma_k - \gamma_{k-1} + (k-2)m]m \\ &\geq .013(k-2)m^2 \\ &\geq 1300m^2 \ln m. \end{aligned}$$

We now obtain an upper bound on the total weight of pairs containing a member ρ_2 of R_k . Suppose that $C'_1, C'_2, \dots, C'_k$ are the cycles of Λ_2 in increasing order of length $\gamma'_1 \leq \gamma'_2 \leq \dots \leq \gamma'_k$. Our choices for Λ_1 are restricted as follows: We must choose

two cycles C'_r, C'_s such that $\gamma'_r + \gamma'_s = \gamma_k + \gamma_\ell \geq \gamma_k \geq .46m$ and $\gamma'_r, \gamma'_s \geq .1m$. This implies that $\{r, s\} \subseteq \{k-2, k-1, k\}$. (Those Λ_2 that are paired with Λ_1 have at most 3 cycles of size greater than $.08m$.) Given this, we see that

$$\begin{aligned} \sum_{\rho_1: (\rho_1, \rho_2) \in Y_k} w(\rho_1, \rho_2) &\leq 3m^2 \max_{M \leq m} \left\{ \sum_{\substack{a+b=M \\ a, b \geq 2}} \left(\frac{1}{a} + \frac{1}{b} \right) \right\} \\ &\leq 3m^2 \sum_{a=2}^m \frac{2}{a} \\ &\leq 6m^2 \ln m. \end{aligned}$$

It follows that

$$(19) \quad 1300m^2 |Q_k| \ln m \leq \sum_{(\rho_1, \rho_2) \in Y_k} w(\rho_1, \rho_2) \leq 6m^2 |R_k| \ln m.$$

From (17) and (19), we have

$$(20) \quad |R_k| \geq 9|S_k|/10.$$

6.4. Final estimate. We can now complete our proof of Theorem 2. Fix $k_0 \leq k \leq \lfloor m/2 \rfloor$, and let

$$Z_k = \{(\rho_1, \rho_2) : \rho_1 \in R_k, \rho_2 \in S_{k-2}, \text{ and } \rho_1 \oplus \rho_2 \cong H_1\},$$

where H_1 is the graph shown in Figures 3(c) and 4(c). Here we combine two cases. We take 3 (see Figure 3) or 4 (see Figure 4) cycles from Λ_1 and reduce this number by 2 to create Λ_2 .

Let $\rho_1 \in R_k$ give rise to Λ_1 which has cycles $C_1, C_2, \dots, C_k$ of lengths $\gamma_1, \gamma_2, \dots, \gamma_k$. For $(\rho_1, \rho_2) \in Z_k$ we will delete four edges from Λ_1 and add four new edges to obtain Λ_2 . If Λ_2 is obtained from Λ_1 by breaking and patching cycles of length $\gamma_1, \gamma_2, \dots, \gamma_s$, then $w(\Lambda_1, \Lambda_2) = (\gamma_1^{-1} + \gamma_2^{-1} + \dots + \gamma_s^{-1})$. We will first obtain a lower bound on the total weight of pairs containing ρ_1 .

We choose cycles C_i and C_j from Λ_1 and delete one edge from each $(p, \lambda(p))$ and $(q, \lambda(q))$, respectively. The remaining two edges to be deleted should lie outside the cycles, C_i and C_j . Let the edges deleted from outside the cycles, C_i and C_j , be $(r, \lambda(r))$ and $(s, \lambda(s))$.

We will give a lower bound on the number of sets of four edges that can be deleted from the auxiliary graph. Suppose $(p, \lambda(p))$ and $(q, \lambda(q))$ are chosen from C_i and C_j , respectively. The number of choices for $(p, \lambda(p))$ and $(q, \lambda(q))$ are γ_i and γ_j , respectively. Now, we remove two edges, $(r, \lambda(r))$ and $(s, \lambda(s))$, outside of cycles γ_i and γ_j such that edges of the form $(x, \lambda(y))$ and edges of the form $(y, \lambda(x))$ are edges in Γ for $x \in \{p, q\}$ and $y \in \{r, s\}$. The number of choices for $(r, \lambda(r))$ such that $(p, \lambda(r))$ is an edge in Γ is at least $.99m$. The number of choices for $(r, \lambda(r))$ such that both $(p, \lambda(r))$ and $(q, \lambda(r))$ are edges of in Γ is at least $.99m + .99m - m \geq .98m$. By a similar argument, we have that the number of choices for $(r, \lambda(r))$ such that both $(r, \lambda(p))$ and $(r, \lambda(q))$ are edges of Γ is at least $.98m$. Thus the number of choices for $(r, \lambda(r))$ is at least $.98m + .98m - m \geq .96m$. Since $(r, \lambda(r))$ should lie outside of cycles C_i and C_j , the number of choices for $(r, \lambda(r))$ is $.96m - \gamma_i - \gamma_j$. The same applies

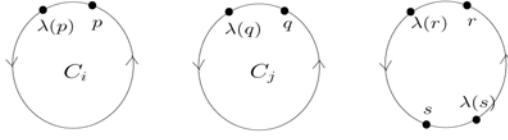


Figure 3(a)

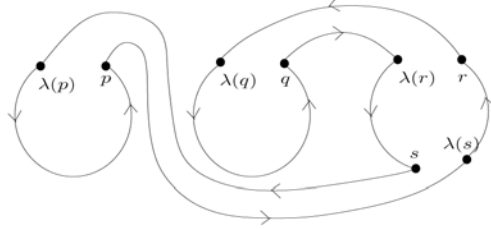


Figure 3(b)

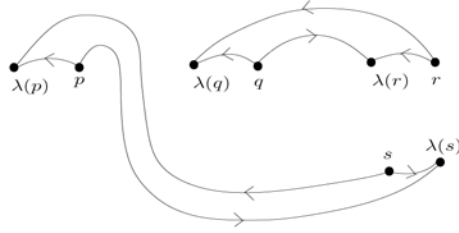


Figure 3(c)

FIG. 3. Figure (a) shows three cycles of Λ_1 and Figure (b) shows the cycle of Λ_2 obtained by coalescing the three cycles of Λ_1 . Figure (c) gives the graph $H_1 = \Lambda_1 \oplus \Lambda_2$.

to $(s, \lambda(s))$. We also add the extra condition that p, q, r, s lie in the order p, q, r, s or p, s, r, q on H_{ρ_1} (see Figure 5), thereby ensuring that ρ_2 is a cyclic permutation. The number of choices for r and s is at least $\binom{(.96m - \gamma_i - \gamma_j)/2}{2}$. The number of choices for p, q, r, s is therefore at least

$$(21) \quad \gamma_i \gamma_j \binom{(.96m - \gamma_i - \gamma_j)/2}{2} \geq .0001 \gamma_i \gamma_j m^2.$$

We can pick the same set of four edges at most 12 times by picking one of the four edges to be the first edge and one of the remaining three to be the second edge from a different cycle. The weight of (Λ_1, Λ_2) is at least $\gamma_i^{-1} + \gamma_j^{-1}$. Consequently, the weight of arcs emanating from $\rho_1 \in R_k$ is

$$(22) \quad \begin{aligned} &\geq \frac{m^2}{12000} \sum_{i=1}^k \sum_{j \neq i} \gamma_i \gamma_j \left(\frac{1}{\gamma_i} + \frac{1}{\gamma_j} \right) \\ &= \frac{m^2}{12000} \sum_{i=1}^k \sum_{j \neq i} (\gamma_i + \gamma_j) \\ &\geq \frac{(k-1)m^3}{12000} \\ &> 8m^3 \ln m. \end{aligned}$$

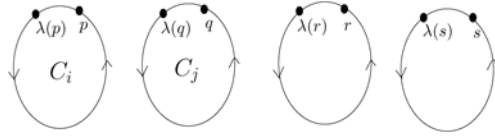


Figure 4(a)

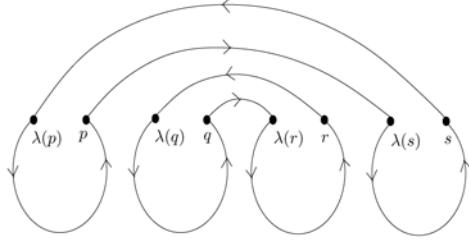


Figure 4(b)

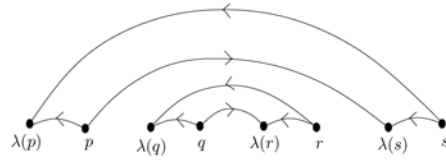


Figure 4(c)

FIG. 4. Figure (a) shows three cycles of Λ_1 and Figure (b) shows the two cycle of Λ_2 obtained by coalescing the four cycles of Λ_1 . Figure (c) gives the graph $H_1 = \Lambda_1 \oplus \Lambda_2$.

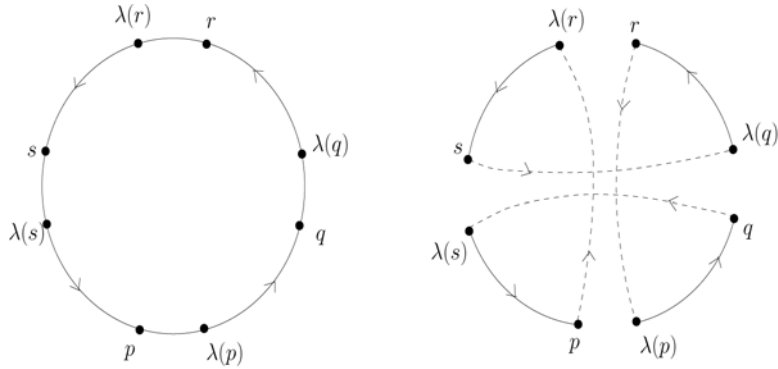


FIG. 5.

Now, we will give an upper bound on the total weight of pairs $(\Lambda_1, \Lambda_2) \in Z_k$ for a fixed $\Lambda_2 \in S_k$. Suppose Λ_2 has cycles $C_1, C_2, \dots, C_k$ of lengths $\gamma_1, \gamma_2, \dots, \gamma_k$. Either one cycle of Λ_2 was formed by coalescing three cycles of Λ_1 or two cycles of Λ_2 were formed by coalescing four cycles of Λ_1 into two.

Case 1. Suppose $C_i \in \Lambda_2$ was formed by coalescing three cycles of lengths a, b , and c from Λ_1 with one edge deleted from the cycles of length a and b and two edges deleted from the cycle of length c , as in Figure 3. There is a choice $1 \leq d \leq c - 1$ for length of the *first* of the paths created from the cycle of length c . Given a, b, c, d ,

there are γ_i choices for p say. Thus the total weight of pairs containing a fixed Λ_2 can be bounded from above in this case by

$$\begin{aligned}
 & \sum_{i=1}^k \gamma_i \sum_{\substack{(a,b,c,d): a,b,c \geq 1 \\ 1 \leq d \leq c-1 \\ a+b+c=\gamma_i}} \left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \\
 & \leq m \sum_{i=1}^k \gamma_i \sum_{\substack{(a,b,c): a,b,c \geq 1 \\ a+b+c=\gamma_i}} \left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \\
 & \leq 3m \sum_{i=1}^k \gamma_i^2 (\ln \gamma_i + 1) \\
 (23) \quad & \leq 4m^3 \ln m.
 \end{aligned}$$

Case 2. Suppose $C_i, C_j \in \Lambda_2$ were formed by combining four cycles of lengths a, b, c , and d from Λ_1 with one edge deleted from each of the four cycles, as in Figure 4. The total weight of pairs containing a fixed Λ_2 can be bounded from above in this case by

$$\begin{aligned}
 & \sum_{i=1}^k \sum_{j=i+1}^k \gamma_i \gamma_j \sum_{\substack{(a,c): a,c \geq 2 \\ a+c=\gamma_i}} \sum_{\substack{(b,d): b,d \geq 2 \\ b+d=\gamma_j}} \left(\frac{1}{a} + \frac{1}{c} + \frac{1}{b} + \frac{1}{d} \right) \\
 & = \frac{1}{2} \sum_{i=1}^k \sum_{j \neq i} \gamma_i \gamma_j \sum_{a=2}^{\gamma_i-2} \sum_{b=2}^{\gamma_j-2} \left(\frac{1}{a} + \frac{1}{\gamma_i-a} + \frac{1}{b} + \frac{1}{\gamma_j-b} \right) \\
 & \leq \frac{1}{2} \sum_{i=1}^k \sum_{j \neq i} \gamma_i \gamma_j (2\gamma_j \ln \gamma_i + 2\gamma_i \ln \gamma_j) \\
 & \leq \frac{1}{2} \sum_{i=1}^k 4\gamma_i m^2 \ln m \\
 & = 2m^3 \ln m. \\
 (24) \quad &
 \end{aligned}$$

Combining inequalities (22), (23), and (24), we have that for $k \geq k_0$,

$$(6m^3 \ln m) |S_{k-2}| \geq w(Z_k) \geq (8m^3 \ln m) |R_k|,$$

and hence

$$|S_{k-2}| \geq 4|R_k|/3 \geq 6|S_k|/5$$

after using (20). It follows that

$$(25) \quad \sum_{\substack{k \geq k_0 \\ k \text{ odd}}} |S_k| \leq 6|S_{k_0-2}|.$$

Recall that all λ have an odd number of cycles.

We will now consider λ with at most $3 \leq k \leq k_0 - 2$ cycles. We will show that

$$(26) \quad \frac{|S_{k-2}|}{|S_k|} \geq \frac{1}{m^4}$$

which implies

$$|S_k| \leq m^{2(k-1)} |S_1|.$$

Hence,

$$\frac{|R_\Gamma|}{|S_\Gamma|} = \frac{|S_1|}{\sum_{\substack{k \geq 1 \\ k \text{ odd}}} |S_k|} = \Omega(m^{-2k_0}).$$

Therefore we have

$$|R_\Gamma| = e^{-o(m)} |S_\Gamma| \geq e^{-o(m)} |T_\Gamma| \geq m! e^{-(5/2+o(1))m}.$$

It remains to prove (26).

Fix $k \leq k_0 - 2$, and let P_k, Q_k , and R_k partition S_k as before. Suppose $\Lambda \in S_k$ has cycles $C_1, C_2, \dots, C_k$ in increasing order of size $\gamma_1, \gamma_2, \dots, \gamma_k$.

Suppose $\rho \in S_k$. Suppose $(p, \lambda(p)) \in C_i$ and $(q, \lambda(q)) \in C_j$ are deleted where $i \neq j$ and $i, j \in \{1, 2, \dots, k-1\}$. We delete two edges $(r, \lambda(r))$ and $(s, \lambda(s))$ from C_k such that the four paths created can be patched into one cycle, reducing the number of cycles by two. We also ensure that the corresponding ρ' is cyclic. If $\rho \in R_k$, then $\gamma_i + \gamma_j \leq .92m$. If $\rho \in P_k \cup Q_k$, then $\gamma_i + \gamma_j \leq m - \gamma_k \leq .54m$. The binomial term in (21) is thus at least $\binom{(.96m-.92m)/2}{2} > 0$, and this ensures that we have at least one pair of edges in γ_k that can be deleted. Thus there is at least one way to transform $\rho \in S_k$ into $\rho' \in S_{k-2}$. Fix one way for each $\rho \in S_k$. Clearly, each ρ' arises in at most m^4 times in this way.

This verifies (26) and completes our proof of Theorem 2. $\square$

REFERENCES

- [1] A. AMIT AND N. LINIAL, *Random graph coverings. I. General theory and graph connectivity*, Combinatorica, 22 (2002), pp. 1–18.
- [2] A. AMIT AND N. LINIAL, *Random lifts of graphs. Edge expansion*, Combin. Probab. Comput., 15 (2006), pp. 317–332.
- [3] A. AMIT, N. LINIAL, AND J. MATOUŠEK, *Random lifts of graphs. Independence and chromatic number*, Random Structures Algorithms, 20 (2002), pp. 1–22.
- [4] B. BOLLOBÁS, *Random Graphs*, 2nd ed., Cambridge University Press, Cambridge, UK, 2001.
- [5] K. BURGIN, P. CHEBOLU, C. COOPER, AND A. M. FRIEZE, *Hamilton cycles in random lifts of graphs*, European J. Combin., 27 (2006), pp. 1282–1293.
- [6] C. COOPER AND A. M. FRIEZE, *Hamilton cycles in a class of random directed graphs*, J. Combin. Theory Ser. B, 62 (1994), pp. 151–163.
- [7] C. COOPER AND A. M. FRIEZE, *Hamilton cycles in random graphs and directed graphs*, Random Structures Algorithms, 16 (2000), pp. 369–401.
- [8] C. COOPER, A. M. FRIEZE, AND M. MOLLOY, *Hamilton cycles in random regular digraphs*, Combin. Probab. Comput., 3 (1994), pp. 39–49.
- [9] M. DYER, A. FRIEZE, AND M. JERRUM, *Approximately counting Hamilton paths and cycles in dense graphs*, SIAM J. Comput., 27 (1998), pp. 1262–1272.
- [10] T. I. FENNER AND A. M. FRIEZE, *On the connectivity of random m -orientable graphs and digraphs*, Combinatorica, 2 (1982), pp. 347–359.
- [11] A. M. FRIEZE, R. M. KARP, AND B. REED, *When is the assignment bound tight for the asymmetric traveling-salesman problem?*, SIAM J. Comput., 24 (1995), pp. 484–493.
- [12] V. F. KOLCHIN, *Random Mappings*, Optimization Software Inc., New York, 1986.
- [13] N. LINIAL AND E. ROZENMAN, *Random lifts of graphs: Perfect matchings*, Combinatorica, 25 (2005), pp. 407–424.

COLORING OF TRIANGLE-FREE GRAPHS ON THE DOUBLE TORUS*

DANIEL KRÁL'[†] AND MATĚJ STEHLÍK[‡]

Abstract. We show that every triangle-free graph on the double torus is 4-colorable. This settles a problem raised by Gimbel and Thomassen [*Trans. Amer. Math. Soc.*, 349 (1997), pp. 4555–4564].

Key words. triangle-free graphs, double torus, graphs on surfaces

AMS subject classifications. 05C15, 05C10

DOI. 10.1137/07067920X

1. Introduction. Colorings of graphs on surfaces permanently attract the attention of researchers in graph theory. The most classical result is the four color theorem [1, 13], which asserts that every planar graph is 4-colorable. Another classical result is Grötzsch's theorem [8], which states that every planar graph with no triangles is 3-colorable; also see [15, 16] for short proofs of this result.

Gimbel and Thomassen [7] generalized Grötzsch's theorem to surfaces of higher genera. While the chromatic number of a graph embedded on a surface of genus g is bounded by $O(g^{1/2})$ [9], the chromatic number of a triangle-free graph that can be embedded on a surface of genus g is bounded by $O((g/\log g)^{1/3})$ [7]. On the other hand, there exist triangle-free graphs embeddable on a surface of genus g that have a chromatic number $\Omega(g^{1/3}/\log g)$ [7].

Let us focus on triangle-free graphs on surfaces of small genera. As we have already mentioned, triangle-free plane graphs are 3-colorable. Every triangle-free graph in the projective plane is 3-degenerate and thus 4-colorable. On the other hand, there are nonbipartite triangle-free projective planar quadrangulations, and each such quadrangulation is 4-chromatic [18]. Kronk and White [11] established that every triangle-free graph on the torus is 4-colorable; as in the case of the projective plane, the bound cannot be improved since there exist triangle-free graphs on the torus that are not 3-colorable (one example is the Cayley graph for the group Z_{13} with generators 1 and 5 [2]). Gimbel and Thomassen [7] asked whether the result of Kronk and White can be extended to the double torus S_2 .

PROBLEM 1 (see [7, Problem 9]). *Is every graph on S_2 of girth four 4-colorable?*

In the present paper, we answer this question in the affirmative way (see Corollary 23). As there are triangle-free graphs on the torus that are not 3-colorable, the bound cannot be improved.

2. Preliminary observations. In this section, we recall standard graph theory notation related to graph colorings and critical graphs. However, we do not provide any detailed introduction to topological graph theory. We refer the reader to a recent

*Received by the editors January 4, 2007; accepted for publication (in revised form) December 17, 2007; published electronically March 20, 2008.

<http://www.siam.org/journals/sidma/22-2/67920.html>

[†]Institute for Theoretical Computer Science (ITI), Faculty of Mathematics and Physics, Charles University, Malostranské náměstí 25, 118 00 Prague, Czech Republic (kral@kam.mff.cuni.cz). This author's research with the Institute for Theoretical Computer Science was supported as project 1M0545 by the Czech Ministry of Education.

[‡]Na Zábřadlí 1, 110 00 Prague, Czech Republic (stehlik@kam.mff.cuni.cz). This author's research was supported by the Department of Applied Mathematics, Charles University.

monograph [12] if interested. The only fact that we will need in our further considerations is the following corollary of Euler's formula: The number of edges of a simple triangle-free graph that can be embedded on the surface S_g is at most $2n - 4 + 4g$. In particular, the following lemma holds.

LEMMA 1. *The number of edges of an n -vertex triangle-free graph that can be embedded on the double torus is at most $2n + 4$.*

We also assume that the reader is familiar with basic graph theory concepts such as k -colorability or the chromatic number of graphs. In our investigations, the structure of "extremal" non- $(k-1)$ -colorable graphs will play a crucial role: A graph G is k -critical if it is not $(k-1)$ -colorable but each proper subgraph of G is $(k-1)$ -colorable. Clearly, the minimum degree of a k -critical graph is at least $k-1$. The vertices of a k -critical graph can be partitioned into two groups: *low-degree* vertices of degree $k-1$ and *high-degree* vertices of degree k or more. The subgraph induced by the low-degree vertices of G is denoted by $L_{k-1}(G)$ and that induced by the high-degree vertices by $H_{k-1}(G)$. One of the first results on critical graphs is the following theorem of Gallai that restricts the possible structure of the components of the low-degree subgraph. Recall that a *Gallai tree* is a graph such that each block (maximal 2-connected subgraph) is an odd cycle or a complete graph.

THEOREM 2 (see Gallai [6]). *If G is a k -critical graph, then each component of $L_{k-1}(G)$ is a Gallai tree.*

Gallai [6] also studied which colorings of the high-degree subgraph cannot be extended to the components of low-degree subgraphs. In particular, he showed that if each component of $L_{k-1}(G)$ contains a vertex adjacent to two vertices of $H_{k-1}(G)$ of the same color, then the coloring of $H_{k-1}(G)$ can be extended to all of the components of $L_{k-1}(G)$. This is in fact a special case of a more general phenomenon studied in the area of list colorings [3, 5, 17]. Results obtained in this area allow us to replace the original condition of Gallai by several others. In particular, the following lemma is true.

LEMMA 3. *Let G be a 5-critical graph. Any precoloring of $H_4(G)$ with four colors can be extended to any component of $L_4(G)$ that has*

- *a vertex adjacent to two vertices of $H_4(G)$ of the same color, or*
- *two adjacent vertices v_1 and v_2 of degree two in $L_4(G)$ such that their neighbors in $H_4(G)$ have at least three distinct colors.*

We will also use the following result of Stiebitz [14] that allows us to bound the number of components of the high-degree subgraph by the number of components of the low-degree subgraph.

THEOREM 4 (see Stiebitz [14]). *If G is a k -critical graph that contains a vertex of degree $k-1$, then the number of components of $H_{k-1}(G)$ does not exceed the number of components of $L_{k-1}(G)$.*

Our goal is to show that every triangle-free graph on the double torus is 4-colorable. In order to do so, we show that there are no 5-critical triangle-free graphs that can be embedded on the double torus (Theorem 22). As we have already observed, the minimum degree of a 5-critical graph is four. On the other hand, Lemma 1 implies that the average degree of each triangle-free graph on the double torus is at most slightly above four. Hence the high-degree subgraph of a triangle-free 5-critical graph on the double torus cannot contain too many vertices.

LEMMA 5. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. Let ℓ be the number of vertices of $H_4(G)$ and $d_1, \dots, d_\ell$ their degrees in G . The sum $\sum_{i=1}^{\ell} (d_i - 4)$ is at most eight. In particular, the number of vertices of $H_4(G)$ is at most eight.*

Lemma 5 easily follows from Lemma 1 and the fact that the minimum degree of a 5-critical graph is four, and we leave its detailed proof to the reader.

We finish this section with two results on the minimum number of vertices of a triangle-free graph that is not 3- or 4-colorable. Let us remark that the bounds given in the following two theorems are the best possible.

THEOREM 6 (see Chvátal [4]). *Every triangle-free graph on at most ten vertices is 3-colorable.*

THEOREM 7 (see Jensen and Royle [10]). *Every triangle-free graph on at most 21 vertices is 4-colorable.*

3. Structure of the low-degree subgraph. In this section, we focus on the possible structure of the low-degree subgraph of triangle-free 5-critical graphs on the double torus. We define the *weight* of a component G_0 of $L_4(G)$ to be the number of edges between G_0 and $H_4(G)$. Note that the weight of each component of $L_4(G)$ is even. The next lemma provides us with a simple bound on the total weight of the components of $L_4(G)$ in terms of the number of high-degree vertices and the number of edges of $H_4(G)$.

LEMMA 8. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. If $H_4(G)$ has ℓ vertices and m edges, then the total weight of the components of $L_4(G)$ is at most $4\ell + 8 - 2m$.*

Proof. Let $d_1, \dots, d_\ell$ be the degrees of the vertices of $H_4(G)$. By Lemma 5, the sum of the degrees $\sum_{i=1}^{\ell} d_i$ of the vertices of $H_4(G)$ is at most $4\ell + 8$. If the number of edges of $H_4(G)$ is m , then the number of edges between $H_4(G)$ and $L_4(G)$ is at most $4\ell + 8 - 2m$. Hence the total weight of the components of $L_4(G)$ is at most $4\ell + 8 - 2m$. $\square$

By Theorem 2, each component of $L_4(G)$ is a Gallai tree. Let us call a component G_0 of $L_4(G)$ that has no vertices of degree one and that is not an odd cycle a *grunter*. Since each end-block of a grunter must be an odd cycle of length at least five (recall that we are considering triangle-free graphs), the next lemma readily follows from Theorem 2.

LEMMA 9. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus, and let G_0 be a component of $L_4(G)$. If G_0 does not contain a vertex of degree one, then either G_0 is an odd cycle or it is a grunter. In particular, if G_0 does not contain a vertex of degree one, then it contains two adjacent vertices of degree two.*

The simplest components of $L_4(G)$ are trees. A weight of such an n -vertex component is $2n + 2$. In the next lemma, we show that if the weight of an n -vertex component is significantly smaller than $2n + 2$, then n must be quite large. This will allow us to efficiently analyze the structure of components of $L_4(G)$ based on their weights in our further considerations.

LEMMA 10. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. If G' is an n -vertex component of $L_4(G)$, then the weight of G' is equal to $2n + 2$ if and only if G' is a tree, and its weight is equal to $2n$ if and only if G' is a cycle or a unicyclic graph. Moreover, if the weight of G' is $2n - 2\ell$, then $n \geq 5 + 4\ell$. In particular, if the weight of G' is less than $2n$, then the weight of G' is at least 16, and if it is less than $2n - 2$, then it is at least 22.*

Specifically, if G' has no vertex of degree one, then its weight is at least $2n \geq 10$, and if G' is a grunter, then its weight is at least 16.

Proof. Observe first that the blocks of G' are odd cycles and edges only since G' is triangle-free. If G' has k blocks that are odd cycles, then the weight of G' is

$2n + 2 - 2k$. Moreover, if G' has k blocks that are odd cycles, the number of vertices of G' is at least $4k + 1$. Hence if the weight of G' is $2n + 2$, then $k = 0$ and G' is a tree. If the weight of G' is $2n$, then $k = 1$ and G' is a cycle or a unicyclic graph.

Assume now that the weight of G' is $2n - 2\ell$. We conclude that G' contains $k = \ell + 1$ blocks that are odd cycles, and thus it contains at least $4k + 1 = 4\ell + 5$ vertices. In particular, if the weight of G' is less than $2n$, i.e., $\ell \geq 1$, then G' has weight of at least

$$2 \cdot (4\ell + 5) - 2\ell = 6\ell + 10 \geq 16.$$

It remains to prove the last part of the lemma. If G' has no vertex of degree one, then $k \geq 1$ and the weight of G' is at least

$$2n + 2 - 2k \geq 2(4k + 1) + 2 - 2k = 6k + 4 \geq 10.$$

If G' is a grunter, then $k \geq 2$ and, consequently, its weight is at least $6k + 4 \geq 16$. The whole lemma has now been established. $\square$

We now aim to utilize our observations on the structure of the low-degree subgraph of a triangle-free 5-critical graph on the double torus. Let us start by showing that the number of components of $L_4(G)$ must be at least two.

LEMMA 11. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. The subgraph $L_4(G)$ contains at least two components.*

Proof. By Lemma 5 and Theorem 6, $L_4(G)$ contains at least one component. Assume for the sake of contradiction that $L_4(G)$ has a single component G' . Since G' is a Gallai tree it contains a vertex v of degree at most two. Let w_1 and w_2 be two neighbors of v in $H_4(G)$. Note that the vertices w_1 and w_2 are not adjacent since G is triangle-free. Consider a coloring of $H_4(G)$ with three colors, which exists by Lemma 5 and Theorem 6, and recolor the vertices w_1 and w_2 with the fourth (unused) color. By Lemma 3, the precoloring of the vertices of $H_4(G)$ can be extended to a coloring of G with four colors—a contradiction. $\square$

We finish this section with three lemmas which bound the number of components of $L_4(G)$ under certain assumptions on the total number of vertices and the total weight of the components of $L_4(G)$.

LEMMA 12. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. If the total weight of the components of $L_4(G)$ is at most 26 and $L_4(G)$ has at least 15 vertices, then the number of components of $L_4(G)$ does not exceed one.*

Proof. Assume for the sake of contradiction that $L_4(G)$ has at least two components. Since $26 < 2 \cdot 15$, $L_4(G)$ must contain an n_1 -vertex component G_1 of weight at most $2n_1 - 2$.

If the weight of G_1 is exactly $2n_1 - 2$, then $L_4(G)$ must contain another n_2 -vertex component G_2 of weight less than $2n_2$. By Lemma 10, the weight of each of G_1 and G_2 is at least 16. Since the total weight of the components of $L_4(G)$ is at most 26, the component G_2 could not exist. We conclude that the weight of G_1 is at most $2n_1 - 4$.

By Lemma 10, the weight of G_1 is at least 22. Since the total weight of the components of $L_4(G)$ is at most 26, the weight of G_1 is 22 and the weight of the other component of $L_4(G)$ is four. We conclude that G_1 has 13 vertices and the other component of $L_4(G)$ is comprised of a single vertex. Hence $L_4(G)$ has $n_1 + 1 = 13 + 1 = 14$ vertices, which contradicts our assumption that the number of vertices of $L_4(G)$ is at least 15. $\square$

In the next lemma, we show that the number of components of $L_4(G)$ does not exceed two under some weaker assumptions.

LEMMA 13. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. If the total weight of the components of $L_4(G)$ is at most 28 and $L_4(G)$ has at least 14 vertices, then the number of components of $L_4(G)$ does not exceed two.*

Proof. Assume for the sake of contradiction that $L_4(G)$ is comprised of three or more components. Since $28/3 < 10$, $L_4(G)$ contains a component that is a tree by Lemma 10. Let G_1 be such a component and n_1 its number of vertices. By Lemma 10, the weight of G_1 is $2n_1 + 2$.

Since the number of vertices of $L_4(G)$ is at least 14 and the total weight of its components is 28, $L_4(G)$ must contain an n_2 -vertex component G_2 of weight at most $2n_2 - 2$. Since the weight of G_2 is at most $28 - 2 \cdot 4 = 20$, the weight of G_2 is $2n_2 - 2$ by Lemma 10. On the other hand, since the sum of the weights of G_1 and G_2 is at least $4 + 16 = 20$, any component of $L_4(G)$ distinct from G_1 and G_2 is a tree by Lemma 10. Hence G_2 is the only component of $L_4(G)$ which is not a tree. We infer from Lemma 10 that the total weight of the components of $L_4(G)$ is at least $2 \cdot 14 + 2 = 30$, which is impossible. This completes the proof. $\square$

Finally, under even weaker assumptions, the number of components of $L_4(G)$ is at most three.

LEMMA 14. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. If the total weight of the components of $L_4(G)$ is at most 30 and $L_4(G)$ has at least 14 vertices, then the number of components of $L_4(G)$ does not exceed three.*

Proof. Assume for the sake of contradiction that $L_4(G)$ contains four or more components. Since the average weight of the components of $L_4(G)$ is also at most $30/4 < 10$, one of the components of $L_4(G)$ is a tree, say G_1 . Since the average weight of the components of $L_4(G)$ distinct from G_1 is at most $(30 - 4)/3 < 10$, $L_4(G)$ contains another component G_2 that is a tree. Let n_1 and n_2 be the number of vertices of G_1 and G_2 , respectively. By Lemma 10, the sum of the weights of G_1 and G_2 is $2(n_1 + n_2) + 4$.

Since $L_4(G)$ has 14 vertices and the total sum of the weights of the components of $L_4(G)$ is 30, G must contain an n_3 -vertex component G_3 of weight at most $2n_3 - 2$. Since the weight of G_3 is at most $30 - 3 \cdot 4 = 18$, the weight of G_3 is $2n_3 - 2$ by Lemma 10. Note that the weight of G_3 must also be at least 16. Because the total weight of G_1 , G_2 , and G_3 is at least 24, $L_4(G)$ contains four components and the component G_4 is a tree. Let n_4 be the number of vertices of G_4 . By Lemma 10, the total weight of the components of $L_4(G)$ is $2(n_1 + n_2 + n_3 + n_4) + 4$. Since $n_1 + n_2 + n_3 + n_4 \geq 14$, the total weight of the components of $L_4(G)$ must be at least 32, which is impossible. $\square$

4. Number of edges in the high-degree subgraph. In this section we focus on estimates on the number of edges that could be contained in the high-degree subgraph of a triangle-free 5-critical graph on the double torus. We start by showing a lower bound on the number of such edges.

LEMMA 15. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. The subgraph $H_4(G)$ contains at least three edges.*

Proof. Assume to the contrary that there exists a graph G in which $H_4(G)$ contains at most two edges. If $H_4(G)$ has at most one edge or two edges sharing a vertex, we can color the vertices of $H_4(G)$ in such a way that all of the vertices of $H_4(G)$ have the same color except for a single vertex with a different color. Let G' be a component of $L_4(G)$. If G' contains a vertex v of degree one, then v is adjacent to two vertices of $H_4(G)$ of the same color. If G' contains no vertices of degree one, then G' contains two adjacent vertices v_1 and v_2 of degree two by Lemma 9. Since G

is triangle-free, the vertices v_1 and v_2 are adjacent to four different vertices of $H_4(G)$. Hence one of them is adjacent to two vertices of the same color. We conclude that each component of $L_4(G)$ has a vertex adjacent to two vertices of $H_4(G)$ of the same color. Lemma 3 implies that G is 4-colorable, which contradicts our assumption that G is 5-critical.

It remains to consider the case when $H_4(G)$ is formed by two disjoint edges, say x_1x_2 and y_1y_2 , and ℓ isolated vertices, $0 \leq \ell \leq 4$. Let c_{ij} , $i, j = 1, 2$, be the coloring of $H_4(G)$ that assigns the vertices x_i and y_j the same color and all the remaining vertices of $H_4(G)$ another color. Let us consider a component G' of $L_4(G)$. Assume that G' has no vertices of degree one and each vertex of G' is adjacent to vertices of $H_4(G)$ with mutually distinct colors. Since $H_4(G)$ is colored with only two colors, G' has no vertices of degree one. Moreover, each vertex of G' of degree two is adjacent to either x_i or y_j . Since two vertices adjacent in G' cannot both be neighbors of x_i or of y_j , G' is not an odd cycle. Hence Lemma 9 yields that G' is a grunter. Moreover, each vertex of degree two contained in an end-block of the grunter is adjacent to either x_i or y_j (note that the vertices of degree two in an end-block must be adjacent alternately to x_i and y_j).

Let us now distinguish two cases based on the number of grunter components of $L_4(G)$. The number of grunter components of $L_4(G)$ is at most two since each grunter has a weight of at least 16 and the total weight of the components of $L_4(G)$ is at most $40 - 4 = 36$ by Lemma 8.

We first consider the case that $L_4(G)$ has a single grunter component G_1 . If G_1 contains a vertex of degree two adjacent to two vertices in $H_4(G)$ with the same color in the coloring c_{11} or in c_{12} , then the coloring of $H_4(G)$ can be extended to G_1 by Lemma 3. Otherwise, each vertex of degree two is adjacent to two vertices of distinct colors both in the coloring c_{11} and the coloring c_{12} . Hence, the vertices of degree two in the end-blocks of G_1 are adjacent to x_1 and y_1 alternately (because of the coloring c_{11}) and to x_1 and y_2 alternately (because of the coloring c_{12}). Consequently, there is a vertex z of G_1 adjacent to both y_1 and y_2 . We conclude that G contains a triangle y_1y_2z .

It remains to analyze the case that $L_4(G)$ has two grunter components. Let G_1 and G_2 be these two components. Assume that for each of the four colorings c_{ij} , all of the vertices of G_1 or all of the vertices G_2 of degree two are adjacent to vertices of $H_4(G)$ with two different colors. In the previous paragraph, we have shown that G_1 contains a vertex of degree two adjacent to two vertices of $H_4(G)$ with the same color in c_{11} or c_{12} . By symmetry, assume that G_1 contains a vertex of degree two adjacent to two vertices of the same color in c_{12} . Similarly, observe that G_2 contains a vertex of degree two adjacent to two vertices of the same color in c_{12} or c_{22} . Since all vertices of G_2 of degree two must be adjacent to two vertices of distinct colors in c_{12} (otherwise, c_{12} could be extended to both G_1 and G_2), G_2 contains a vertex of degree two adjacent to two vertices of the same color in c_{22} . Along these lines, we conclude that G_1 contains vertices of degree two adjacent to two vertices of $H_4(G)$ of the same color both in c_{12} and c_{21} , and G_2 contains vertices of degree two adjacent to two vertices of the same color both in c_{11} and c_{22} .

Since neither of these four colorings c_{ij} can be extended to both G_1 and G_2 , all vertices of degree two of G_1 are adjacent to two vertices of different colors both in c_{11} and c_{22} , and all vertices of degree two of G_2 are adjacent to two vertices of different colors both in c_{12} and c_{21} . In particular, each vertex of degree two lying in an end-block of G_1 is either adjacent to x_1 and y_2 or to x_2 and y_1 (the other cases

are excluded by our assumption that G is triangle-free). Observe that G_1 has at least four vertices of the former kind and at least four vertices of the latter kind. Similarly, each vertex of degree two lying in an end-block of G_2 is either adjacent to x_1 and y_1 or to x_2 and y_2 , and the number of vertices of each of the two kinds is at least four. We conclude that all of the vertices x_1, x_2, y_1 , and y_2 have at least eight neighbors in G_1 and G_2 . Hence the degree of each of them in G is at least nine, which is impossible by Lemma 5.

We infer from our discussion that for at least one of the colorings $c_{11}, c_{12}, c_{21}, c_{22}$, each component of $L_4(G)$ has a vertex adjacent to two vertices of $H_4(G)$ of the same color: Indeed, we have established this for the components G_1 and G_2 , and for other components, which are not grunterns, this follows from the facts that each other component contains a vertex of degree one and that the vertices of $H_4(G)$ are 2-colored. We can now infer from Lemma 3 that G is 4-colorable, which contradicts our assumption that G is 5-critical. $\square$

We have just established a lower bound on the number of edges of $H_4(G)$. Our next aim is to find an upper bound.

LEMMA 16. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. The subgraph $H_4(G)$ contains at most six edges.*

Proof. Assume for the sake of contradiction that $H_4(G)$ has seven or more edges. By Lemma 5 and Theorem 7, $L_4(G)$ has at least 14 vertices, and, by Lemma 11, $L_4(G)$ is comprised of at least two components. By Lemma 8, the total weight of the components of $L_4(G)$ cannot exceed $40 - 14 = 26$. Since $26 < 2 \cdot 14$, $L_4(G)$ has an n_1 -vertex component G_1 of weight less than $2n_1$. On the other hand, the weight of G_1 is at least 16 by Lemma 10. Hence the weight of any other component G_2 of $L_4(G)$ is at most 10. If G_2 were a tree, then $L_4(G)$ would have to contain another component of weight less than twice the number of its vertices, which is impossible. We conclude that G_2 is a cycle of length 5 and its weight is exactly 10. Consequently, the weight of G_1 is 16 and G_1 is a double-5-cycle; i.e., two 5-cycles sharing a vertex. Since the total weight of G_1 and G_2 is 26, the number of vertices of $H_4(G)$ is eight and it contains exactly seven edges. Moreover, the degree of each vertex of $H_4(G)$ in G is five.

Since $H_4(G)$ has eight vertices and seven edges, $H_4(G)$ contains at least four vertices v of degree at most two. Since G is triangle-free, the vertices of G_2 are adjacent to at least five different vertices of $H_4(G)$, and thus there is a vertex v of $H_4(G)$ of degree at most two adjacent to a vertex of G_2 . Since each vertex of $H_4(G)$ can be adjacent to at most two vertices of G_2 (otherwise G would contain a triangle) and the degree of every vertex of $H_4(G)$ in G is five, there exists a vertex v_0 of $H_4(G)$ adjacent to both a vertex of G_1 and a vertex of G_2 . Let v_i be any of the neighbors of v_0 in G_i , $i = 1, 2$. Note that the degree of v_i in G_i is two. Since G_1 is a double-5-cycle, the vertex v_1 has a neighbor v'_1 of degree two in G_1 . Similarly, v_2 has a neighbor v'_2 of degree two in G_2 .

Now color the vertices of $H_4(G)$ with three colors (this is possible by Theorem 6) and recolor the vertex v_0 with the fourth (unused) color. If the vertex v'_1 is not adjacent to two vertices of $H_4(G)$ of the same color, then the neighbors of v_1 and v'_1 in $H_4(G)$ must have at least three distinct colors (two colors appear because of the neighbors of v'_1 and the third color appears because of the vertex v_0 ; note that v_0 is not a neighbor of v'_1 since G is triangle-free). Similarly, if v'_2 is not adjacent to two vertices of $H_4(G)$ of the same color, then the neighbors of v_2 and v'_2 have at least three distinct colors. We infer from Lemma 3 that the precoloring of the vertices of

$H_4(G)$ can be extended to a coloring of all of the vertices of G with four colors—a contradiction. $\square$

Lemmas 15 and 16 now imply the following.

LEMMA 17. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. The number of edges of $H_4(G)$ is between three and six.*

5. Structure of the high-degree subgraph. In this section, we further refine our knowledge about the possible structure of high-degree subgraphs of triangle-free 5-critical graphs G on the double torus. Let us start by showing that the number of vertices of $H_4(G)$ must be eight for every such graph G .

LEMMA 18. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. The number of vertices of $H_4(G)$ is eight. In particular, the degree of each vertex of $H_4(G)$ is five.*

Proof. Assume first that $H_4(G)$ has at most six vertices. By Lemma 17, $H_4(G)$ contains at least three edges. Hence the total weight of the components of $L_4(G)$ is at most $32 - 6 = 26$ by Lemma 8. By Lemma 11, $L_4(G)$ is comprised of at least two components, and by Theorem 7, $L_4(G)$ has at least $22 - 6 = 16$ vertices, which is impossible by Lemma 12.

It remains to exclude the case when the number of vertices of $H_4(G)$ is seven. By Lemma 17, the number of edges of $H_4(G)$ is between three and six. We first exclude the case when $H_4(G)$ has only three edges. Since $H_4(G)$ has seven vertices, $H_4(G)$ is comprised of at least four components. Hence the number of components of $L_4(G)$ is at least four by Theorem 4. By Lemma 8, the total weight of the components of $L_4(G)$ is $36 - 6 = 30$, and by Theorem 7, the number of vertices of $L_4(G)$ is at least 15. However, such a graph G cannot exist by Lemma 14.

We conclude that $H_4(G)$ has at least four edges. Recall that the number of vertices of $L_4(G)$ is at least 15 by Theorem 7. If $H_4(G)$ has four edges, then Lemma 8 yields that the total weight of the components of $L_4(G)$ is at most $36 - 8 = 28$, and Theorem 4 yields that the number of components of $L_4(G)$ is at least three. Such a graph G cannot exist by Lemma 13. If $H_4(G)$ has five or more edges, then by Lemma 8 the total weight of the components of $L_4(G)$ is at most $36 - 10 = 26$, and Theorem 4 yields that the number of components of $L_4(G)$ is at least two. However, Lemma 12 excludes the existence of such a graph G . We conclude that $H_4(G)$ must have eight vertices. The rest follows from Lemma 5. $\square$

Another fact that we establish in this section is that $H_4(G)$ must be bipartite.

LEMMA 19. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. The subgraph $H_4(G)$ is bipartite.*

Proof. Assume for the sake of contradiction that $H_4(G)$ is not bipartite. By Lemma 18, the number of vertices of $H_4(G)$ is eight. Since the number of edges of $H_4(G)$ does not exceed six by Lemma 17 and G is triangle-free, $H_4(G)$ contains five or six edges. In particular, $H_4(G)$ contains a cycle of length five and possibly one more edge. In the rest, we distinguish two cases based on the number of edges of $H_4(G)$ and eventually obtain contradiction in each of them.

Assume first that $H_4(G)$ contains exactly five edges; i.e., $H_4(G)$ is comprised of a cycle of length five and three isolated vertices. In particular, the number of components of $L_4(G)$ is at least four by Theorem 4. By Lemma 8, the total weight of the components of $L_4(G)$ is at most $40 - 2 \cdot 5 = 30$, and by Theorem 7, the number of vertices of $L_4(G)$ is at least 14. Consequently, Lemma 14 excludes the existence of G .

We have shown that $H_4(G)$ must contain six edges. Consequently, the number

of components of $H_4(G)$ is three. By Lemma 8, the total weight of the components of $L_4(G)$ is $40 - 2 \cdot 6 = 28$, and by Theorem 7, the number of vertices of $L_4(G)$ is at least 14. Theorem 4 and Lemma 13 now exclude the existence of such a graph G . $\square$

6. Bipartite high-degree subgraph with eight vertices. In this section, we utilize our observations to exclude the existence of a 5-chromatic triangle-free graph that can be embedded on the double torus. Let us start by observing that if the high-degree subgraph is bipartite, then at least one of the components of the low-degree subgraph is an odd cycle or a grunter.

LEMMA 20. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. If $H_4(G)$ is bipartite, then $L_4(G)$ has a component that is either an odd cycle or a grunter.*

Proof. Assume the contrary, i.e., that each component of $L_4(G)$ has a vertex of degree one by Theorem 4, and consider any 2-coloring of the vertices of $H_4(G)$. Since each component of $L_4(G)$ has a vertex of degree one (which must be adjacent to two vertices of the same color), G is 4-colorable by Lemma 3, contrary to our assumption that G is 5-critical. $\square$

In the proof of the main theorem, we consider 3-colorings of $H_4(G)$ such that one of the colors is assigned to a single vertex. We claim that all such colorings can be extended to all components of $L_4(G)$ of weight at most eight.

LEMMA 21. *Let G be a triangle-free 5-critical graph that can be embedded on the double torus. If the vertices of $H_4(G)$ are precolored with three colors in such a way that one of the three colors is assigned to a single vertex of G , then the precoloring of $H_4(G)$ can be extended to any component of $L_4(G)$ of weight at most 8.*

Proof. Let G' be a component of $L_4(G)$ of weight at most 8. By Lemma 10, G' is a tree. Hence G' is either a single vertex, an edge, or a path comprised of two edges. If G' contains a vertex adjacent to two vertices of $H_4(G)$ of the same color, then the precoloring can be extended to G' by Lemma 3. Hence if the precoloring of $H_4(G)$ cannot be extended to G' , then G' is a path $v_1v_2v_3$ and the vertices v_1 and v_3 are adjacent to vertices of three distinct colors. Since $H_4(G)$ is precolored with three distinct colors, the vertices v_1 and v_3 can be colored with the fourth color. Since the vertex v_2 has degree four, it is adjacent to vertices of at most three distinct colors (its neighbors v_1 and v_3 have the same color), and thus the coloring can also be extended to v_2 as desired. $\square$

We are now ready to prove our main result.

THEOREM 22. *There is no triangle-free 5-critical graph G that can be embedded on the double torus.*

Proof. By Lemmas 18 and 19, $H_4(G)$ is a bipartite graph with eight vertices. By Lemma 17, the number of edges of $H_4(G)$ is between three and six. Hence the total weight of the components of $L_4(G)$ is at most 34 by Lemma 8.

In the rest of the proof, we establish a series of claims that eventually combine to the proof of the theorem.

CLAIM 1. *If $L_4(G)$ contains a component G_1 whose weight is less than twice the number of the vertices of G_1 , then $L_4(G)$ contains another component of weight ten or more.*

Assume the opposite. Let n_1 be the number of vertices of G_1 . Since G_1 is the only component of $L_4(G)$ of weight ten or more, the remaining components of $L_4(G)$ are trees of weight at most eight by Lemma 10. By Lemmas 10 and 20, G_1 is a grunter (G_1 cannot be a cycle since its weight is less than n_1). Consider any 2-coloring c of

$H_4(G)$. If the precoloring c cannot be extended to $L_4(G)$, then each vertex of G_1 of degree two is adjacent to two vertices of $H_4(G)$ of two distinct colors. Let v_1 and v_2 be any two adjacent vertices of degree two in G_1 . Now recolor any vertex of $H_4(G)$ adjacent to v_1 with the third (unused) color. G_1 now contains two adjacent vertices of degree two, namely v_1 and v_2 , such that their neighbors in $H_4(G)$ are colored with three distinct colors. In particular, the precoloring of $H_4(G)$ can be extended to G_1 by Lemma 3. Since the precoloring of $H_4(G)$ can be extended to the remaining component(s) of $L_4(G)$ by Lemma 21, G is 4-colorable, which is impossible.

CLAIM 2. *The subgraph $L_4(G)$ contains no component of weight less than twice the number of its vertices.*

Assume the contrary, i.e., that $L_4(G)$ contains a component G_1 with n_1 vertices of a weight of at most $2n_1 - 2$. Claim 1 yields that $L_4(G)$ contains another component of a weight of ten or more. Let G_2 be this component of $L_4(G)$ and n_2 the number of its vertices. Since the weight of G_1 is at most $2n_1 - 2$, its weight is at least 16 by Lemma 10. Since the total weight of G_1 and G_2 is at least $16 + 10 = 26$, the weight of any component of $L_4(G)$ distinct from G_1 and G_2 is at most eight. Assume that G_1 contains a vertex w of degree one. Hence G_2 is a grunter or a cycle by Lemma 20. In particular, G_2 contains two adjacent vertices of degree two, say v_1 and v_2 . Since G is triangle-free, v_1 and v_2 have four distinct neighbors in H_4 , say $u_1, \dots, u_4$. By symmetry, we can assume that u_1 is not a neighbor of w .

Now consider a coloring of the vertices of $H_4(G)$ with two colors. If G_2 contains a vertex of degree two adjacent to two vertices of $H_4(G)$ of the same color, then the precoloring of the vertices of $H_4(G)$ can be extended to all of the components of $L_4(G)$ by Lemma 3 since each component of $L_4(G)$ has a vertex adjacent to two vertices of $H_4(G)$ of the same color. Otherwise, each vertex of G_2 of degree two is adjacent to two vertices of $H_4(G)$ of distinct colors. Now recolor the vertex u_1 with the third (unused) color. The precoloring of $H_4(G)$ can be extended to G_1 by Lemma 3 since the vertex w is adjacent to two vertices of $H_4(G)$ of the same color, to G_2 since the vertices v_1 and v_2 are adjacent to vertices of $H_4(G)$ of three distinct colors, and to the remaining components of $L_4(G)$ by Lemma 21.

We conclude that G_1 has no vertex of degree one. An analogous argument yields that G_2 also has no vertex of degree one. Since the weight of G_1 is at most $2n_1 - 2$, G_1 is a grunter. If G_2 were also a grunter, then the sum of the weights of G_1 and G_2 would be at least $2 \cdot 16 = 32$ and G_1 and G_2 would be the only components of $L_4(G)$. By Lemma 8, the number of edges of $H_4(G)$ would be at most four, and thus the number of components of $H_4(G)$ would be at least four, which is impossible by Theorem 4. Hence G_2 is an odd cycle.

Let v_1 and v_2 be two adjacent vertices of degree two in G_1 , and let A_1 be the set of their four neighbors in $H_4(G)$. Let A_2 be the set of the neighbors of the vertices of G_2 in $H_4(G)$. Since G is triangle-free, A_2 contains at least five distinct vertices. Hence there is a vertex u contained in both A_1 and A_2 . By symmetry, we can assume that u is a neighbor of v_1 . Let w_1 be a neighbor of u in G_2 and w_2 a vertex of G_2 adjacent to w_1 .

Now consider a coloring of $H_4(G)$ with three distinct colors such that all of the vertices except u are colored with only two colors and u is colored with the third color. By Lemma 21, the coloring can be extended to all of the components of $L_4(G)$ with a possible exception of G_1 and G_2 . If the vertex v_2 is adjacent to two vertices of $H_4(G)$ of the same color, then the precoloring can be extended to G_1 by Lemma 3. Otherwise, the vertices v_1 and v_2 are adjacent to vertices of $H_4(G)$ with three distinct

colors and the precoloring can be extended to G_1 again by Lemma 3. An analogous argument yields that the precoloring can be extended to G_2 . Hence the graph G is 4-colorable, which is impossible. The proof of Claim 2 is now complete.

We have established that the weight of each component of $L_4(G)$ is at least twice the number of its vertices. In particular, all the components of $L_4(G)$ are trees, cycles, and unicyclic graphs. By Lemma 20, at least one of the components is an odd cycle. Let G_1 be this component and n_1 the number of its vertices. In addition, let A_1 be the neighbors of G_1 in $H_4(G)$. Note that $|A_1| \geq 5$.

CLAIM 3. *The subgraph $L_4(G)$ has at least three components of a weight of ten or more.*

First, assume that G_1 is the only component of $L_4(G)$ of a weight of ten or more. Now choose a vertex $w \in A_1$ and color the vertices of $H_4(G)$ with three colors in such a way that all of the vertices of $H_4(G)$ except w are assigned only two colors. Then, the precoloring can be extended to $L_4(G)$ by Lemmas 3 and 21, which is impossible.

Assume next that $L_4(G)$ has two components of a weight of ten or more; let G_2 be such a component distinct from G_1 and let n_2 be the number of its vertices. If G_2 has a vertex v of degree one, then choose $w \in A_1$, that is not a neighbor of v , and proceed as in the previous paragraph. Otherwise, G_2 must be an odd cycle. Since the vertices of G_2 have at least five neighbors in $H_4(G)$, there exists a vertex $w \in A_1$ adjacent to both G_1 and G_2 . Analogously to the preceding cases, the precoloring assigning the vertices of $H_4(G)$ except w two colors and w the third color can be extended to each component of $L_4(G)$.

CLAIM 4. *There is no triangle-free 5-critical graph that can be embedded on the double torus.*

Claim 3 implies that there are at least three components, say G_1 , G_2 , and G_3 , that have a weight of ten or more. If $L_4(G)$ had four components, then their total weight would be at least $3 \cdot 10 + 4 = 34$. Hence $H_4(G)$ would have only three edges by Lemma 8. On the other hand, $H_4(G)$ would have five components, which is impossible by Theorem 4. We conclude that G_1 , G_2 , and G_3 are the only components of $L_4(G)$.

Since $L_4(G)$ is comprised of only three components, $H_4(G)$ has at least five edges by Theorem 4. Hence the total weight of the components of $L_4(G)$ is at most $40 - 5 \cdot 2 = 30$ by Lemma 8. Consequently, the weight of each G_i , $i = 1, 2, 3$, is 10, and $H_4(G)$ is a forest with exactly five edges.

By Lemma 20 and symmetry, we can assume that G_1 is a cycle of length five. If both G_2 and G_3 were trees, then the number of vertices of $L_4(G)$ would be $5 + 4 + 4 = 13$, which is impossible by Theorem 7. We conclude that G_2 is also a cycle of length five. Let A_i , $i = 1, 2$, be the set of vertices of $H_4(G)$ that are adjacent to a vertex of G_i . Note that $|A_i| \geq 5$ for $i = 1, 2$. Moreover, if $|A_i| = 5$, then each vertex of A_i is adjacent to exactly two vertices of G_i .

Now assume that G_3 is a tree, and let v_1 and v_2 be two of its leaves. If $|A_1| > 5$ or $|A_2| > 5$, then there exist three vertices u_1 , u_2 , and u_3 of $H_4(G)$ adjacent to vertices of both G_1 and G_2 . Now consider a 2-coloring of the vertices of $H_4(G)$. If there is a vertex u_i , $i = 1, 2, 3$, not adjacent to v_1 , then recolor u_i with the third (unused) color. The coloring of the vertices of $H_4(G)$ can now be extended to G_1 , G_2 , and G_3 by Lemma 3. Hence the vertices u_1 , u_2 , and u_3 are neighbors of v_1 . By symmetry, we can assume that the colors of u_1 and u_2 are the same. Let us now recolor the vertex u_3 with the third (unused) color. Again, Lemma 3 yields that the precoloring can be extended to the vertices of $L_4(G)$, which is impossible since G is not 4-colorable.

It remains to consider the case when $|A_1| = |A_2| = 5$. Note that each vertex of

A_i is adjacent to exactly two vertices of G_i since G is triangle-free. Since $H_4(G)$ has eight vertices, there exist two vertices u_1 and u_2 adjacent to vertices of both G_1 and G_2 . If u_i is not adjacent to v_1 or v_2 , then we consider a coloring of the vertices of $H_4(G)$ with three colors that assigns the vertex u_i a color different from all of the other vertices of $H_4(G)$ and proceed as in the previous paragraph. Hence both u_1 and u_2 are adjacent to v_1 and v_2 . However, this implies that the degree of u_1 is six, which contradicts our previous deduction that all the vertices of $H_4(G)$ have degree five (see Lemma 18).

We can now conclude that all of the three components G_1 , G_2 , and G_3 are cycles of length five. As in the previous cases, let A_i be the set of neighbors of G_i in $H_4(G)$. If there exists a vertex $u \in A_1 \cap A_2 \cap A_3$, then we color the vertices of $H_4(G)$ with two colors and recolor u with the third (unused) color. We infer from Lemma 3 that the precoloring can be extended to $L_4(G)$, which contradicts our assumption that G is not 4-colorable.

It remains to consider the case when $A_1 \cap A_2 \cap A_3 = \emptyset$. Let u_{12} be a vertex of $A_1 \cap A_2$ and u_{23} a vertex of $A_2 \cap A_3$. Note that such vertices exist since $|A_i| \geq 5$ for every $i = 1, 2, 3$. Now color the vertices of $H_4(G)$ with two colors, recolor u_{12} with the third (unused) color, and u_{23} with the fourth (unused) color. It is straightforward to check that each of the cycles G_1 , G_2 , and G_3 either contains a vertex adjacent to two vertices of the same color or its vertices are adjacent to vertices with at least three distinct colors; thus it has two adjacent vertices of degree two adjacent to vertices of at least three distinct colors in $H_4(G)$. Lemma 3 now yields that the precoloring can be extended to $L_4(G)$, which is impossible since G is not 4-colorable. $\square$

As a corollary of Theorem 22, we can now settle Problem 1.

COROLLARY 23. *Every graph on S_2 of girth four is 4-colorable.*

Acknowledgment. The authors would like to thank both of the anonymous referees for their comments that helped to improved the clarity of the arguments used.

REFERENCES

- [1] K. APPEL AND W. HAKEN, *Every planar map is four colorable*, Bull. Amer. Math. Soc., 82 (1976), pp. 711–712.
- [2] D. ARCHDEACON, J. HUTCHINSON, A. NAKAMOTO, S. NEGAMI, AND K. OTA, *Chromatic numbers of quadrangulations on closed surfaces*, J. Graph Theory, 37 (2001), pp. 100–114.
- [3] O. V. BORODIN, *Criterion of chromaticity of a degree prescription*, in Abstracts of the Fourth All-Union Conference on Theoretical Cybernetics, Novosibirsk, 1977, pp. 127–128 (in Russian).
- [4] V. CHVÁTAL, *The minimality of the Mycielski graph*, in Graphs and Combinatorics, Lecture Notes in Math. 406, Springer-Verlag, Berlin, 1974, pp. 243–246.
- [5] P. ERDŐS, A. L. RUBIN, AND H. TAYLOR, *Choosability in graphs*, in Proceedings of the West Coast Conference on Combinatorics, Graph Theory, and Computing (Arcata, CA, 1979) Congress Numer. 26, Utilitas Math., Winnipeg, MB, Canada, 1980, pp. 125–157.
- [6] T. GALLAI, *Kritische Graphen I*, Publ. Math. Inst. Hungar. Acad. Sci., 8 (1963), pp. 165–192.
- [7] J. GIMBEL AND C. THOMASSEN, *Coloring graphs with fixed genus and girth*, Trans. Amer. Math. Soc., 349 (1997), pp. 4555–4564.
- [8] H. GRÖTZSCH, *Zur theorie der diskreten gebilde*, VII, *Ein Dreifarbensatz für dreikreisfreie Netze auf der Kugel*, Wiss. Z. Martin-Luther-Universität, Halle-Wittenberg, Math.-Nat. Reihe, 8 (1959), pp. 109–120.
- [9] P. J. HEAWOOD, *Map colour theorem*, Quart. J. Pure Appl. Math., 24 (1890), pp. 332–338.
- [10] T. JENSEN AND G. F. ROYLE, *Small graphs with chromatic number 5: A computer search*, J. Graph Theory, 19 (1995), pp. 107–116.
- [11] H. V. KRONK AND A. T. WHITE, *A 4-color theorem for toroidal graphs*, Proc. Amer. Math. Soc., 34 (1972), pp. 83–86.

- [12] B. MOHAR AND C. THOMASSEN, *Graphs on Surfaces*, Johns Hopkins University Press, Baltimore, MD, 2001.
- [13] N. ROBERTSON, D. SANDERS, D. SEYMOUR, AND R. THOMAS, *The four-color theorem*, J. Combin. Theory Ser. B, 70 (1997), pp. 2–44.
- [14] M. STIEBITZ, *Proof of a conjecture of T. Gallai concerning connectivity properties of colour-critical graphs*, Combinatorica, 2 (1982), pp. 315–323.
- [15] C. THOMASSEN, *3-list-coloring planar graphs of girth 5*, J. Combin. Theory Ser. B, 64 (1995), pp. 101–107.
- [16] C. THOMASSEN, *A short list color proof of Grötzsch’s theorem*, J. Combin. Theory Ser. B, 88 (2003), pp. 189–192.
- [17] V. G. VIZING, *Colouring the vertices of a graph with prescribed colours*, Diskret. Anal., 29 (1976), pp. 3–10 (in Russian).
- [18] D. A. YOUNGS, *4-chromatic projective graphs*, J. Graph Theory, 21 (1996), pp. 219–227.

POWER DOMINATION IN PRODUCT GRAPHS*

PAUL DORBEČ†, MICHEL MOLLARD†, SANDI KLAVŽAR‡, AND SIMON ŠPACAPAN§

Abstract. The power system monitoring problem asks for as few as possible measurement devices to be put in an electric power system. The problem has a graph theory model involving power dominating sets in graphs. The power domination number $\gamma_P(G)$ of G is the minimum cardinality of a power dominating set. Dorfling and Henning [*Discrete Appl. Math.*, 154 (2006), pp. 1023–1027] determined the power domination number of the Cartesian product of paths. In this paper the power domination number is determined for all direct products of paths except for the odd component of the direct product of two odd paths. For instance, if n is even and C a connected component of $P_m \times P_n$, where m is odd or $m \geq n$, then $\gamma_P(C) = \lceil n/4 \rceil$. For the strong product we prove that $\gamma_P(P_n \boxtimes P_m) = \max\{\lceil n/3 \rceil, \lceil (n+m-2)/4 \rceil\}$, unless $3m-n-6 \equiv 4 \pmod{8}$. The power domination number is also determined for an arbitrary lexicographic product.

Key words. domination, power domination, electric power monitoring, graph products, paths

AMS subject classification. 05C69

DOI. 10.1137/060661879

1. Introduction. Electric power systems need to be continually monitored. One way to fulfill this task is to place phase measurement units at (carefully) selected locations in the system. The power system monitoring problem, as introduced in [2], asks for as few as possible measurement devices to be put in an electric power system.

The power system monitoring problem has been formulated as a graph-theory domination problem by Haynes et al. in [6]. This problem is of somehow different flavor than standard domination type problems, since putting a phase measurement unit into a vertex of a graph can have global effects. For instance, if an electric power system can be modeled by a path, then a single measurement unit suffices to monitor the system no matter how long the path.

For a vertex v of a graph G , let $N(v)$ denote the open neighborhood of v , and for a subset S of $V(G)$ let $N(S) = \cup_{v \in S} N(v) - S$. In the same way, the closed neighborhood $N[S]$ of a subset S is the set $N[S] = N(S) \cup S$.

Let G be a connected graph and S a subset of its vertices. Then we denote the set monitored by S with $M(S)$ and define it algorithmically as follows:

1. (*domination*)

$$M(S) \leftarrow S \cup N(S)$$

2. (*propagation*)

As long as there exists $v \in M(S)$ such that

$$N(v) \cap (V(G) - M(S)) = \{w\}$$

*Received by the editors June 2, 2006; accepted for publication (in revised form) August 27, 2007; published electronically March 21, 2008. This work was supported in part by Proteus Project BI-FR/04-002, by the Ministry of Science of Slovenia under grant P1-0297, by ENS Lyon, and by CNRS.

<http://www.siam.org/journals/sidma/22-2/66187.html>

†Institut Fourier, BP 74, 100 rue des Maths, 38402 St Martin d'Hères Cedex, France (paul.dorbec@ujf-grenoble.fr, michel.mollard@ujf-grenoble.fr).

‡Department of Mathematics and Computer Science, FNM, University of Maribor, Gosposvet-ska 84, 2000 Maribor, Slovenia. Also with the Institute of Mathematics, Physics and Mechanics, Jadranska 19, 1000 Ljubljana (sandi.klavzar@uni-mb.si).

§University of Maribor, FME, Smetanova 17, 2000 Maribor, Slovenia; Institute of Mathematics, Physics and Mechanics, Jadranska 19, 1000 Ljubljana (simon.spacapan@uni-mb.si).

set $M(S) \leftarrow M(S) \cup \{w\}$.

In other words, the set $M(S)$ is obtained from S as follows. First put into $M(S)$ the vertices from the closed neighborhood of S . Then repeatedly add to $M(S)$ vertices w that have a neighbor v in $M(S)$ such that all the other neighbors of v are already in $M(S)$. After no such vertex w exists, the set monitored by S has been constructed.

The set S is called a *power dominating set* of G if $M(S) = V(G)$ and the *power domination number* $\gamma_P(G)$ is the minimum cardinality of a power dominating set.

In [6] these concepts have been introduced in a slightly more complicated way by treating both vertices and edges of a given graph. However, it is easily shown that both approaches are equivalent in the sense that the power dominating sets correspond. In fact, this is also (implicitly) observed in [3].

The power domination number has received considerable attention from the algorithmic point of view. As it turned out, the problem is quite difficult—it is NP-complete even when restricted to bipartite graphs and chordal graphs [6], to planar graphs and circle graphs [4], and to split graphs [4, 11]. On the other hand, the problem has efficient solutions on trees [6] and more generally on graphs with bounded treewidth [10], as well as on interval graphs [11] and block graphs [13]. We also add that the power domination number of block graphs has been further studied in [1], and that in [14] upper bound for the power domination number of general graphs and claw-free graphs are presented.

Dorfling and Henning [3] obtained closed formulas for the power domination numbers of grid graphs. This result is in striking contrast to the fact that a determination of such formulas for the usual domination number of grid graphs is a notorious open problem; see [7, section 2.6.2]. Now, a natural description of a grid is as the Cartesian product of two paths. Besides the Cartesian one, there are three more standard graph products: the strong, the direct, and the lexicographic product [8]. Hence it is natural to ask whether the power domination number can also be determined for these products of paths. In this paper we indeed show this to be the case.

We proceed as follows. In the rest of this section we give definitions used in this paper. In section 2 we determine the power domination number for direct products of paths with the exception of the odd components of products of odd paths. For this case an upper bound is given for which we believe to be optimal. Then, in section 3, we determine the power domination number for the strong product of paths, except that in one of the eight cases the power domination number is only bounded. But also in this case the bounds are almost tight. While it takes quite an effort to obtain the results for the direct and the strong product of paths, it is rather straightforward to deal with the lexicographic product. In fact, in the last section we show that the power domination number can be determined exactly for the lexicographic product of arbitrary graphs in terms of the (total) domination number of the factors.

All of the four standard graph products constructed from graphs G and H have vertex set $V(G) \times V(H)$. Let (g, h) and (g', h') be two vertices in $V(G) \times V(H)$. They are adjacent in the *Cartesian product* $G \square H$ if they are equal in one coordinate and adjacent in the other. They are adjacent in the *direct product* $G \times H$ if they are adjacent in both coordinates. The edge set of the *strong product* $G \boxtimes H$ is $E(G \square H) \cup E(G \times H)$. Finally, (g, h) and (g', h') are adjacent in the *lexicographic product* $G \circ H$ if $gg' \in E(G)$, or if $g = g'$ and $hh' \in E(H)$. All these products are associative; all but the lexicographic product are also commutative. See [8] for more information on all the products.

Let $G * H$ be any of the four standard graph products. Then the subgraph of $G * H$ induced by $\{g\} \times V(H)$ is called an *H-fiber* and denoted ${}^gH$. Similarly, one

defines the G -fiber G^h for a vertex h of H . If $*$ is the Cartesian, the strong, or the lexicographic product, then fibers are isomorphic to the corresponding factors. For the direct product the fibers are discrete, that is, they induce edgeless graphs.

Observe that if one of the factors is the trivial graph, then the direct product is a stable graph, and every other product is isomorphic to the other factor. Therefore, we will assume in the following that both factors are nontrivial.

Finally, for the path P_m we will always assume that $V(P_m) = \{0, \dots, m-1\}$, where vertices are adjacent in the natural way. Moreover, the vertices of $P_m * P_n$ will be denoted by (i, j) , where $0 \leq i < m$ and $0 \leq j < n$.

2. The direct product. For any integers $m, n \geq 2$, the direct product $P_m \times P_n$ is not connected. If m or n is even, $P_m \times P_n$ consists of two isomorphic connected components; otherwise it has two nonisomorphic connected components (see [12]; cf. also [5, 9]). We call the component of $P_m \times P_n$ containing $(0, 0)$ the *even component* and the other component the *odd component*.

In this section we determine the power domination number of the direct product of two paths, except for the odd component of the product of two odd paths. In the first subsection we treat the case when at least one of the paths is even. In this case $\gamma_P(C) = \lceil n/4 \rceil$, where C is a connected component of $P_m \times P_n$ with n even and m odd or $m \geq n$. In the subsequent subsection we follow with products of odd paths proving that for the even component C , $\gamma_P(C) = \lceil (m+n)/6 \rceil$. We close by constructing a power dominating set of the odd component of such products. We conjecture that the construction is optimal.

Let $p_1 : P_m \times P_n \rightarrow P_m$ and $p_2 : P_m \times P_n \rightarrow P_n$ be the natural projections onto the first and the second factor of $P_m \times P_n$, respectively. Then we have the following lemma.

LEMMA 2.1. *Let S be a power dominating set of a connected component C of $P_m \times P_n$. If $m \geq 3$ is odd, then for every subpath $P \subseteq P_n$ of length 3, $p_2(S) \cap P \neq \emptyset$.*

Proof. Suppose to the contrary that there exists a subpath $P = x_1x_2x_3x_4$ of length 3, such that $p_2(S) \cap P = \emptyset$.

Assume C is even and x_1 is even. During the domination step, no vertices of $P_m^{x_3}$ are monitored. Moreover, any neighbor of a vertex from $P_m^{x_3} \cap C$ has two neighbors in $P_m^{x_3}$, not monitored after the domination step. Therefore, no vertex of $P_m^{x_3} \cap C$ may be monitored after the propagation step. Similarly we infer that if x_1 is odd, then no vertices of $P_m^{x_2} \cap C$ may be monitored, a contradiction. Arguments for C being odd are analogous. $\square$

2.1. At least one factor is even. To solve the case when at least one factor is an even path, we begin with two lemmas.

LEMMA 2.2. *Let m and n be even and let S be a power dominating set of a connected component C of $P_m \times P_n$. If there is a subpath $P \subseteq P_m$ of length 3, such that $P \cap p_1(S) = \emptyset$, then for every subpath $Q \subseteq P_n$ of length 3, $Q \cap p_2(S) \neq \emptyset$.*

Proof. We will prove this lemma by contradiction. Let m and n be even, and let S be a power dominating set of a connected component C of $P_m \times P_n$. Both components are isomorphic, so we assume without loss of generality that C is the even component. Suppose that $P = p_1p_2p_3p_4 \subseteq P_m$ and $Q = q_1q_2q_3q_4 \subseteq P_n$ are paths of length 3 with $p_1(S) \cap P = p_2(S) \cap Q = \emptyset$ (see Figure 1). Let

$$F = \{(i, j) \mid i \notin \{p_2, p_3\}, j \notin \{q_2, q_3\}\} \cap C.$$

After the domination step, all the monitored vertices are necessarily in F . Suppose that all the vertices of F are monitored. Suppose (u, v) is a vertex of F such that all

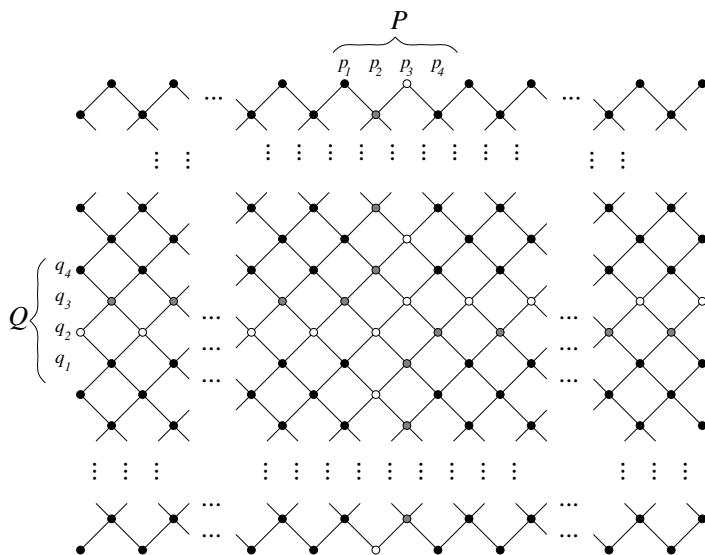


FIG. 1. Proof of Lemma 2.2. Black vertices are in F and gray vertices are in M .

except one neighbor of (u, v) are in F (and thus monitored). Necessarily, (u, v) is on the boundary of $P_m \times P_n$, so $(u, v) \in \{(p_2, 0), (p_2, n-1), (p_3, 0), (p_3, n-1), (0, q_2), (m-1, q_2), (0, q_3), (m-1, q_3)\}$. Since n and m are even, exactly four vertices among these are in the component C . As we can see in Figure 1, when p_1 and q_1 are odd, the set M of the vertices obtained by propagation from F does not contain all the vertices of C . Therefore, S may not power dominate C . The cases for other parities of p_1 and q_1 are similar. $\square$

LEMMA 2.3. *Let m be even and let C be a connected component of $P_m \times P_n$. If all the vertices from $(P_m^u \cup P_m^v) \cap C$ are monitored, where u and v are adjacent vertices of P_n , then C is power dominated.*

Proof. Since m is even, both $P_m^u \cap C$ and $P_m^v \cap C$ contain a vertex that has exactly one neighbor not in $(P_m^u \cup P_m^v) \cap C$. It follows easily that the two neighboring fibers (or the neighboring fiber if there is only one such) are monitored. Induction completes the argument. $\square$

THEOREM 2.4. *Let n be even and C be a connected component of $P_m \times P_n$. If m is odd or $m \geq n$, then $\gamma_P(C) = \lceil n/4 \rceil$.*

Proof. Let n be an even integer, and let C be a connected component of $P_m \times P_n$. Since both components are isomorphic, we assume without loss of generality that C is the even component. We will first prove the lower bound, then give a construction that reaches it and thus prove the upper bound.

We first consider the case when m is odd. All the vertices in $P_m^1 \cap C$ have two neighbors in $P_m^0 \cap C$, so any power dominating set $S \subseteq C$ has nonempty intersection with the fiber P_m^0 or P_m^1 .

By Lemma 2.1 a power dominating set S of C has the property that for any subpath $P \subseteq P_n$ of length 3, $P \cap p_2(S) \neq \emptyset$. Without loss of generality assume that $P_m^1 \cap S$ is nonempty. Since the fiber P_m^0 as well as the fibers P_m^{n-2} and P_m^{n-1} (but not P_m^{n-3}) might have empty intersection with S and at least every fourth fiber has nonempty intersection with S , we find that $|S| \geq |\{1, 1+4, \dots, 1+4 \lceil (n-4)/4 \rceil\}| = \lceil n/4 \rceil$. Hence $\gamma_P(G) \geq \lceil n/4 \rceil$.

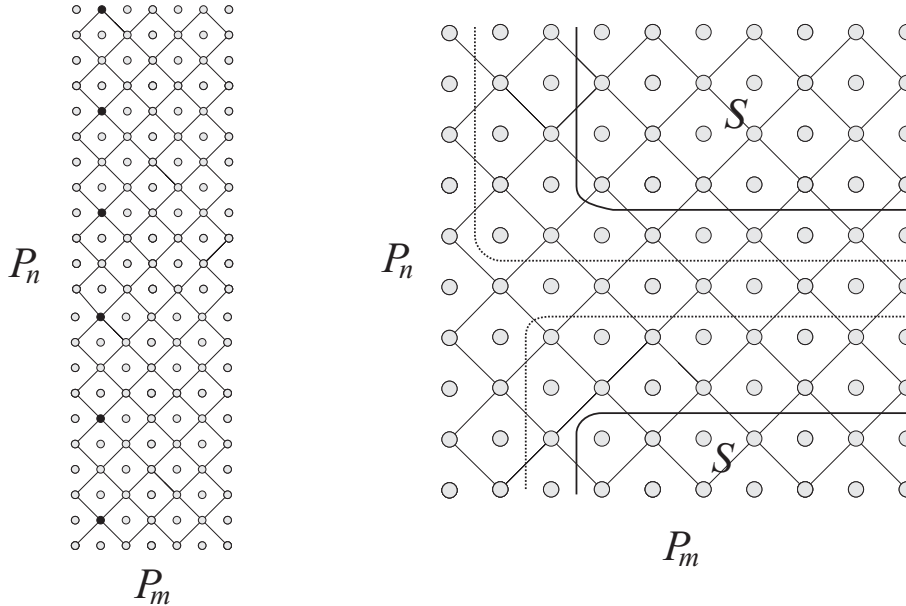


FIG. 2. An optimal power dominating set of $P_m \times P_n$ is shown on the left. The maximal set of monitored vertices is marked with broken line on the right.

Now let m be even, $m \geq n$. Let S be a power dominating set of $P_m \times P_n$. By Lemma 2.2, there is either no subpath of P_m of length 3 disjoint with $p_1(S)$ or there is no subpath of P_n of length 3 disjoint with $p_2(S)$. Hence $|S| \geq \lceil n/4 \rceil$. This is the desired lower bound if $n = 4k$.

If $n = 4k + 2$ and $|S| = k$, then either there is a subpath $P \subseteq P_n$ of length 3 disjoint with $p_2(S)$, or the first three or the last three P_m -fibers (that is, P_m^0, P_m^1 , and P_m^2 or P_m^{n-3}, P_m^{n-2} , and P_m^{n-1}) have empty intersection with $p_2(S)$. In the first case there is no subpath of P_m of length 3 disjoint with $p_1(S)$; therefore in this case $m = n$. But this means that S is disjoint with fibers ${}^0P_n, {}^1P_n$, and ${}^2P_n$ (or ${}^{m-3}P_n, {}^{m-2}P_n$, and ${}^{m-1}P_n$). Therefore, using the right-hand side of Figure 2, we verify that ${}^0P_n$ is not monitored, a contradiction. The second case is when P_m^0, P_m^1 , and P_m^2 have empty intersection with $p_2(S)$. If also ${}^0P_n, {}^1P_n$, and ${}^2P_n$ have empty intersection with $p_1(S)$, then ${}^0P_n$ and P_m^0 are not monitored. Otherwise there is a subpath of P_m of length 3 disjoint with $p_1(S)$. This case leads to a contradiction as shown in Figure 2. Thus if $n = 4k + 2$ and $|S| = k$, then S cannot be a power dominating set. Hence in either of the cases $|S| \geq \lceil n/4 \rceil$.

To prove the reverse inequality we need to construct a power dominating set S of C with $|S| = \lceil n/4 \rceil$. Let

$$(1) \quad S = \{(1, 1 + 4\alpha) \mid \alpha = 0, \dots, \lceil n/4 \rceil - 1\}.$$

Observe that all the vertices from ${}^0P_n$ are monitored since for every vertex from ${}^0P_n$ there is a neighboring vertex from S which dominates it; see the left-hand side of Figure 2. Next observe that also all the vertices from ${}^1P_n$ are monitored, since for every vertex from ${}^1P_n$ there is a vertex from ${}^0P_n$ with all but one vertices in S . Since the consecutive fibers ${}^0P_n$ and ${}^1P_n$ are monitored and n is even, Lemma 2.3 implies that S is a power dominating set of C . This completes the proof. $\square$

Note that Theorem 2.4 holds also for $n = 2$. Indeed, a connected component of $P_m \times P_2$ is isomorphic to P_m whence its power domination number is 1.

2.2. Both factors are odd. We continue with the odd times odd case. For this case we introduce the following concept on a grid $P_m \square P_n$.

Consider an initial set S of vertices. Let D be initialized by $D = S$. As long as there is a vertex v in $V \setminus D$ such that v has at least two neighbors in D , add v to D . If at the end of the process $D = V(P_m \square P_n)$, then we call the set S a *life winning (LWS) set* of $P_m \square P_n$. Let us call this concept the *life-like game*.

THEOREM 2.5. *For any positive integers m, n , there exists a LWS set of $P_m \square P_n$ of cardinality $\lceil \frac{m+n}{2} \rceil$ and it is a minimum LWS of $P_m \square P_n$.*

Proof. Let

$$\begin{aligned} S = & \left\{ (2i, 0) \mid 0 \leq i \leq \frac{m-1}{2} \right\} \\ & \cup \left\{ (m-1, 2j + \mathcal{E}) \mid 0 \leq j \leq \frac{n-2}{2}, \mathcal{E} = 1 \text{ if } m \text{ is even, } 0 \text{ otherwise} \right\} \\ & \cup \{(m-1, n-1)\}. \end{aligned}$$

For example, if m and n are odd, S consists of vertices $(0, 0), (2, 0), \dots, (m-1, 0), (m-1, 2), (m-1, 4), \dots, (m-1, n-1)$. It is straightforward to verify that S is an LWS set of the desired cardinality.

Let us prove that S is minimum. For $X \subseteq V(P_m \square P_n)$ let $\Pi(X)$ be the *perimeter* of X , that is, $\Pi(X) = \sum_{v \in X} (4 - |N(v) \cap X|)$. If D' is a subset of $V(P_m \square P_n)$ obtained by propagation from D , then $\Pi(D') \leq \Pi(D)$. Indeed, any vertex added to D decreases $\Pi(D)$ by at least 2 and adds at most 2 to $\Pi(D)$. At the end of the propagation, $\Pi(D) = 2(m+n)$. At the beginning, $\Pi(S) \leq 4|S|$. Hence, $|S| \geq (m+n)/2$. $\square$

The above theorem might be of independent interest, but for our purposes it is important that during the propagation process the perimeter of the monitored set does not increase.

In the following we will consider the direct product $P_m \times P_n$ and at the same time the Cartesian product $P_m \square P_n$, where the vertex sets of both products are the same.

LEMMA 2.6. *Let m, n be odd integers and P a power dominating set of the even component of $P_m \times P_n$. Let S be the set monitored by P after the domination step of the algorithm. Then S is a LWS of $P_m \square P_n$.*

Proof. We first remark that all the vertices of the even component either have both their coordinates odd or both even. We will call them odd and even vertices, respectively. Since both m and n are odd, the set of even vertices is a LWS of $P_m \square P_n$.

Let M and D , initialized to S , be the sets that are propagating for power domination and life-like game, respectively. We will prove by induction that any even vertex in M is in D . From the choice $M = D = S$, this condition is initially true.

Suppose an even vertex w is added to M . Then, it is monitored by propagation from an odd vertex v . Any odd vertex has four neighbors, all even. So all the three other neighbors of v are monitored and belong to D . In Figure 3, we show how w is thus also put in D by the rule of the life-like game. $\square$

THEOREM 2.7. *Let $m \leq n$ be odd and C the even component of $P_m \times P_n$. Then*

$$\gamma_P(C) = \max \left\{ \left\lceil \frac{n}{4} \right\rceil, \left\lceil \frac{m+n}{6} \right\rceil \right\}.$$

Proof. Since any power dominating set of C contains a vertex from P_m^0 or P_m^1 , and a vertex from P_m^{n-2} or P_m^{n-1} , Lemma 2.1 implies that $\gamma_P(C) \geq n/4$.

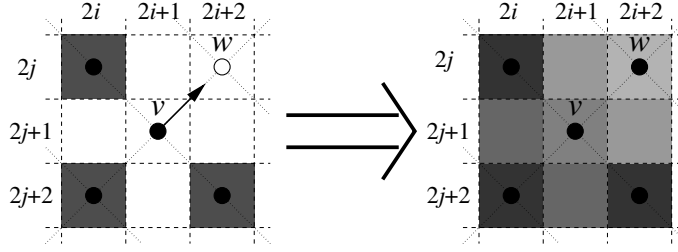


FIG. 3. Propagation from an odd vertex to an even.

Let P be a power dominating set of $P_m \times P_n$. Let S be the set monitored after the domination step. Let us consider the life-like game starting on S . Every vertex from P contributes at most five vertices to S which propagate to a 3×3 box in the grid. Denoting the obtained set S' we infer that every vertex in P contributes at most 12 to $\Pi(S')$. The set S' is a LWS of the grid, so $12|P| \geq 2(m+n)$. (Here we have used the fact that during the propagation process the perimeter of the monitored set does not increase.) Therefore $|P| \geq \lceil (m+n)/6 \rceil$.

To establish the upper bound we construct a corresponding power dominating set P of C as follows. Suppose first that $m \leq n \leq 2m$ and set

$$S_1 = \{(1+2i, 1+4i) \mid 0 \leq i \leq (2n-m-3)/6\},$$

$$S_2 = \{((2n-m)/3+4j, (4n-2m-3)/3+2j) \mid 1 \leq j \leq (2m-n-3)/6\}.$$

Suppose that $n+m \equiv 0 \pmod{6}$ (the perfect case). Then set $P = S_1 \cup S_2$. Otherwise, construct P from a maximal perfect case contained in $P_m \times P_n$ by adding the vertex $(m-2, n-2)$ to it. It is straightforward to verify that in all cases P is a power dominating set. Moreover, $((2n-m-3)/6+1) + (2m-n-3)/6 = (m+n)/6$, hence $|P| = \lceil (m+n)/6 \rceil$.

Let $n > 2m$. Then set

$$S_1 = \{(1+2i, 1+4i) \mid 0 \leq i \leq (m-3)/2\},$$

$$S_2 = \{(1, 1+4j) \mid (m-3)/2 < j \leq (n-3)/4\}.$$

Now the perfect case is when $n \equiv 3 \pmod{4}$, in which case we set $P = S_1 \cup S_2$. Otherwise, proceed as above by adding $(m-2, n-2)$ to the power dominating set of a maximal perfect case contained in $P_m \times P_n$. Since $(m-3)/2 + 1 + ((n-3)/4 - (m-3)/2) = (n+1)/4$, in this case $|P| = \lceil n/4 \rceil$. $\square$

The construction for the first case from the previous proof is illustrated in Figure 4 with black dots. Moreover, another optimal power dominating set is presented with black squares in order to illustrate that such a set can be unusual.

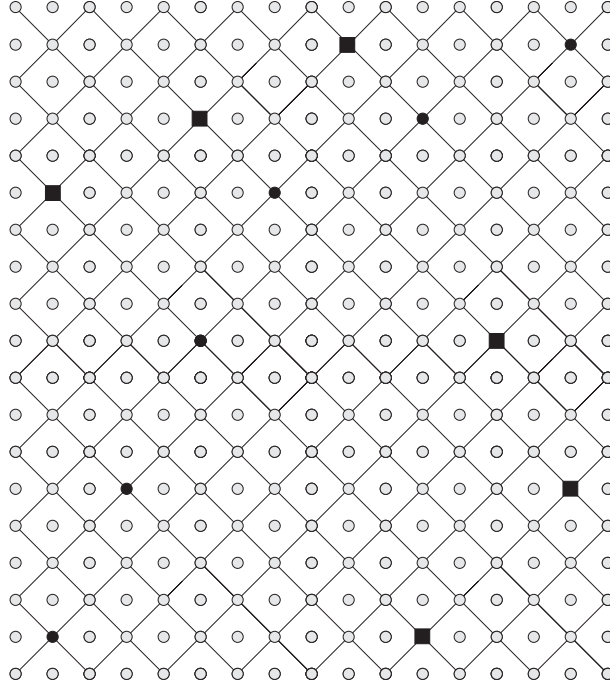
THEOREM 2.8. *Let $m \leq n$ be odd and C the odd component of $P_m \times P_n$. Then*

$$\gamma_P(C) \leq \max \left\{ \left\lceil \frac{n-2}{4} \right\rceil, \left\lceil \frac{m+n-2}{6} \right\rceil \right\}.$$

Proof. We give a construction of a power dominating set of the given cardinality. Suppose $m \leq n \leq 2m+2$ and $m+n-2 \equiv 0 \pmod{6}$ (the perfect case); then the construction is as follows. Let

$$S_1 = \{(3+2i, 4i) \mid 1 \leq i \leq (2n-m-7)/6\} \text{ and}$$

$$S_2 = \{((2n-m-1)/3+4j, (4n-2m-11)/3+2j) \mid 1 \leq j \leq (2m-n-7)/6\}.$$

FIG. 4. Two optimal power dominating sets of the even component of $P_{17} \times P_{19}$.

Then $P = S_1 \cup S_2 \cup \{(2, 1), (m - 2, n - 3)\}$ is a power dominating set of C . Since $(2n - m - 7)/6 + (2m - n - 7)/6 + 2 = (m + n - 2)/6$, this construction has the claimed cardinality.

In the nonperfect case, construct S_1 and S_2 for a minimal perfect case containing $P_m \times P_n$. Add to these sets the vertices $(2, 1)$ and $(m - 2, n - 3)$ (thus, only this last vertex differs). This set clearly has the cardinality $\lceil (m + n - 2)/6 \rceil$.

If $n > 2m + 2$ and $n \equiv 1 \pmod{4}$, then let

$$S_1 = \{(1 + 2i, 2 + 4i) \mid 0 \leq i \leq (m - 3)/2\} \text{ and} \\ S_2 = \{(1, 2 + 4j) \mid (m - 3)/2 \leq j \leq (n - 5)/4\}.$$

In the nonperfect case ($n \equiv 3 \pmod{4}$), construct S_1 and S_2 for $P_m \times P_{n+2}$ and it will form a PDS of $P_m \times P_n$. Since $(m - 3)/2 + 1 + ((n - 5)/4 - (m - 3)/2) = (n - 1)/4$, in this case $|P| = \lceil (n - 2)/4 \rceil$. $\square$

We think that the upper bound of Theorem 2.8 is optimal. However, proving that this is the case seems harder than it was for the even component.

3. The strong product. In this section we prove the following theorem.

THEOREM 3.1. Let $n \geq m \geq 2$. Then

$$\gamma_P(P_n \boxtimes P_m) = \max \left\{ \left\lceil \frac{n}{3} \right\rceil, \left\lceil \frac{n + m - 2}{4} \right\rceil \right\}$$

unless $3m - n - 6 \equiv 4 \pmod{8}$, in which case

$$\max \left\{ \left\lceil \frac{n}{3} \right\rceil, \left\lceil \frac{n + m - 2}{4} \right\rceil \right\} \leq \gamma_P(P_m \boxtimes P_n) \leq \max \left\{ \left\lceil \frac{n}{3} \right\rceil, \left\lceil \frac{n + m - 2}{4} \right\rceil + 1 \right\}.$$

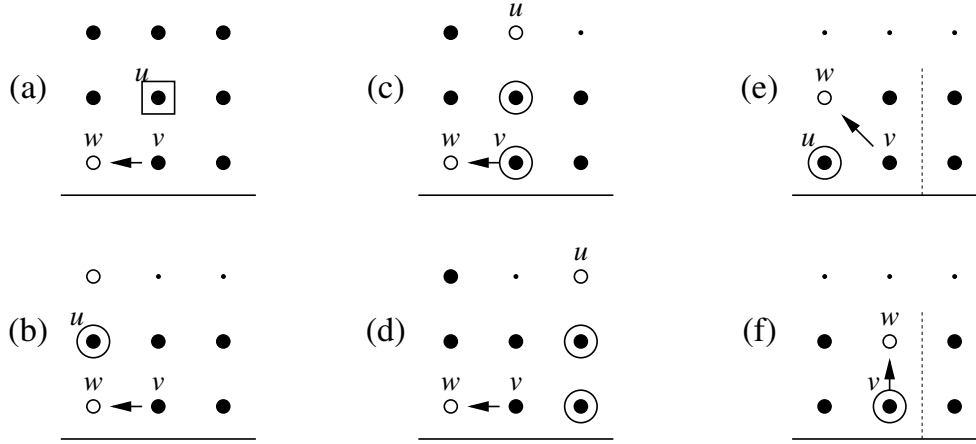


FIG. 5. Propagation from the border.

In the first subsection we prove the lower bounds, while in the subsequent two subsections corresponding power dominating sets are constructed. It is easily verified that the result holds if $n < 4$ and $m < 4$. Hence in the rest of this section we assume that $n \geq m \geq 4$.

3.1. Proof of the lower bound.

LEMMA 3.2. $\gamma_P(P_m \boxtimes P_n) \geq \lceil \frac{n}{3} \rceil$.

Proof. Let S be a power dominating set of $P_m \boxtimes P_n$. Suppose there is a P_m -fiber that does not contain any vertex of $N[S]$. Then, the vertices of this P_m -fiber have to be monitored during the propagation step. Consider the first vertex w of this P_m -fiber that is monitored, and let v be the vertex that monitors w . The vertex v has at least one other neighbor in the same P_m -fiber as w . This neighbor is not yet monitored, so the propagation cannot happen. Thus S must dominate at least one vertex of each P_m -fiber, which implies that $\gamma_P(P_m \boxtimes P_n) \geq \lceil n/3 \rceil$. $\square$

LEMMA 3.3. $\gamma_P(P_m \boxtimes P_n) \geq \frac{n+m-2}{4}$.

Proof. Let S be a minimum power dominating set of $P_m \boxtimes P_n$. We can suppose that S contains no vertices on the boundary (that is, the vertices of $P_m \boxtimes P_n$ of degree 3 or 5), since any closed neighborhood of a vertex on the boundary is contained in the closed neighborhood of a vertex not on the boundary.

Let M be the set of monitored vertices of $P_m \boxtimes P_n$ during a propagation step of the algorithm. Let $B(M)$ be the set of vertices of M that have less than eight neighbors in M . We claim that during the propagation step, $|B(M)|$ cannot increase.

Suppose that during the propagation step we monitor a new vertex w , where w is the only neighbor not in M of some vertex v from M . Then w may now be added to $B(M)$. We distinguish two cases.

Case 1. The degree of v is 8. Then v now has all its eight neighbors in M , so v is removed from $B(M)$. Some other vertices may be removed from $B(M)$, if w was their eighth neighbor to enter M , but none other may be added. So $|B(M)|$ cannot be increasing.

Case 2. The degree of v is less than 8. In this case we check that w could have been monitored with a propagation from a vertex not on the boundary. The cases to be considered are shown in Figure 5, where $\bullet$ are vertices monitored, $\circ$ are vertices not monitored, and $\cdot$ may or may not be monitored.

In case (a) of the figure, the vertex u (in a box) has all its neighbors except w monitored, so we can suppose the propagation comes from this vertex. Cases (b), (e), and (f) cannot occur: the vertex v in a circle cannot have been monitored during the domination step, so it must have been monitored with some propagation. But any neighbor of v also has another neighbor not in $M(S)$, a contradiction. In cases (c) and (d), the two vertices in circles are monitored. By our assumption that no vertices of S are on the boundary, they must have been monitored by propagation (otherwise u would be monitored). But none of them may have been monitored before the other, since all their neighbors are either neighbor of both, or also neighbor of u . The case of a propagation from a corner is induced by cases e and f . This proves the claim.

Hence $|B(M)|$ cannot increase during a propagation step of the algorithm. After the domination step of the algorithm, $|B(M)|$ is not greater than $8|S|$. Moreover, at the end of the algorithm, when all the vertices are monitored, $|B(M)|$ is equal to $2n + 2m - 4$. So we must have $|S| \geq (n + m - 2)/4$. $\square$

3.2. A construction for $n \geq 3m - 6$. In this subsection we construct power dominating sets for the case when one path is much longer than the other, more precisely, when $n \geq 3m - 6$. Then

$$\left\lceil \frac{n}{3} \right\rceil \geq \left\lceil \frac{3n}{12} + \frac{3m-6}{12} \right\rceil = \left\lceil \frac{n+m-2}{4} \right\rceil,$$

hence the power domination number depends only of the size of the longer path. Thus we need to construct a power dominating set S of cardinality $\lceil n/3 \rceil$. We do this as follows. Set $S = S_1 \cup S_2$, where

$$S_1 = \{(1+i, 1+3i) \mid 0 \leq i \leq m-3\}$$

and

$$S_2 = \begin{cases} \{(1, 3m-9+3j) \mid 1 \leq j \leq \lceil (n-3m+6)/3 \rceil\}, & n \equiv 1 \pmod{3}, \\ \{(1, 3m-8+3j) \mid 1 \leq j \leq \lceil (n-3m+6)/3 \rceil\}, & \text{otherwise.} \end{cases}$$

The construction is illustrated in Figure 6.

It is easy to check that the set S will power dominate the whole strong product. Moreover, $|S| = |S_1| + |S_2| = m-2 + \lceil (n-3m+6)/3 \rceil = \lceil n/3 \rceil$.

3.3. A construction for $n \leq 3m - 6$. In this case m and n are comparable. Since $\lceil n/3 \rceil \leq \lceil (n+m-2)/4 \rceil$ we need to construct power dominating sets S of cardinality $\lceil (n+m-2)/4 \rceil$. We start with the case when 8 divides $3m-n-6$, and modify it in all the other cases.

Case 1. $3m-n-6 \equiv 0 \pmod{8}$. Set $S = S_1 \cup S_2$, where

$$\begin{aligned} S_1 &= \{(1+3i, 1+i) \mid 0 \leq i \leq (3m-n-6)/8\} \\ S_2 &= \{(1+3(3m-n-6)/8+j, 1+(3m-n-6)/8+3j) \mid \\ &\quad 1 \leq j \leq (n+m-6)/4 - (3m-n-6)/8\}. \end{aligned}$$

The construction is illustrated in Figure 7.

For coherence of the above construction we should remark that

- $\frac{3m-n-6}{8}$ and $\frac{n+m-6}{4} = m-3 - \frac{3m-n-6}{4}$ are integers.
- $\frac{n+m-6}{4} - \frac{3m-n-6}{8} = \frac{3n-m-6}{8}$ is positive since $n \geq m$ and $n \geq 4$.
- The coordinates of the last vertex which is furthest up and right are $(m-2, n-2)$, so all the vertices are in the graph.

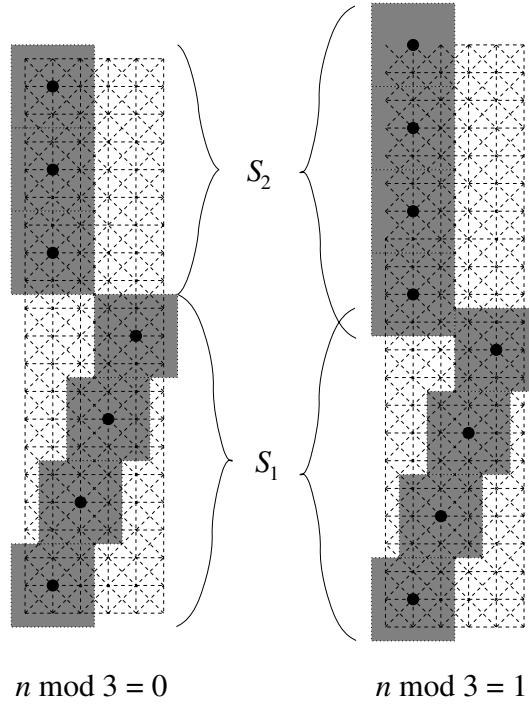


FIG. 6. Optimal power dominating sets when one path is relatively long.

- The set S does not contain any vertex on the border of the grid. (This remark will be very useful in the following.)

The set S contains $(n + m - 2)/4$ vertices and is therefore optimal.

Case 2. $3m - n - 6 \equiv 1 \pmod{8}$. In this case $3m - (n + 1) - 6 \equiv 0 \pmod{8}$. Then we construct the power dominating set S of $P_m \boxtimes P_{n+1}$ from Case 1. Removing the fiber P_m^0 from $P_m \boxtimes P_{n+1}$ we get a graph $G = P_m \boxtimes P_n$. Since S does not contain any border vertex, the set S restricting to G does not change. Consequently, $|S| = (n + 1 + m - 2)/4 = \lceil (n + m - 2)/4 \rceil$.

Case 3. $3m - n - 6 \equiv 2 \pmod{8}$. Then $3m - (n + 2) - 6 \equiv 0 \pmod{8}$, hence construct the power dominating set S of $P_m \boxtimes P_{n+2}$ as in Case 1 and restrict it to the graph $P_m \boxtimes P_{n+2}$ with the fibers P_m^0 and P_m^{n+1} removed. Now $|S| = (n + 2 + m - 2)/4 = \lceil (n + m - 2)/4 \rceil$.

Case 4. $3m - n - 6 \equiv 3 \pmod{8}$. Since $3(m + 2) - (n + 1) - 6 \equiv 0 \pmod{8}$ we take the power dominating set S of $P_{m+2} \boxtimes P_{n+1}$ and remove the fibers ${}^0P_n$, ${}^{m+1}P_n$, and P_m^0 . Then $|S| = (n + 1 + m + 2 - 2)/4 = \lceil (n + m - 2)/4 \rceil$.

Case 5. $3m - n - 6 \equiv 4 \pmod{8}$. As $3(m + 2) - (n + 2) - 6 \equiv 0 \pmod{8}$ we proceed as before, removing the fibers ${}^0P_n$, ${}^{m+1}P_n$, P_m^0 , and P_m^{n+1} . Then $|S| = (n + 2 + m + 2 - 2)/4 = \lceil (n + m - 2)/4 \rceil + 1$. Note that in this special case we miss the lower bound by 1.

Case 6. $3m - n - 6 \equiv 5 \pmod{8}$. Then $3(m + 1) - n - 6 \equiv 0 \pmod{8}$ so we construct the power dominating set S of $P_{m+1} \boxtimes P_n$ and remove the fiber ${}^0P_n$. Then $|S| = (n + m + 1 - 2)/4 = \lceil (n + m - 2)/4 \rceil$.

Case 7. $3m - n - 6 \equiv 6 \pmod{8}$. Then $3(m + 1) - (n + 1) - 6 \equiv 0 \pmod{8}$ and we start from the power dominating set S of $P_{m+1} \boxtimes P_{n+1}$, and remove the fibers ${}^0P_n$ and P_m^0 . Now $|S| = (n + 1 + m + 1 - 2)/4 = \lceil (n + m - 2)/4 \rceil$.

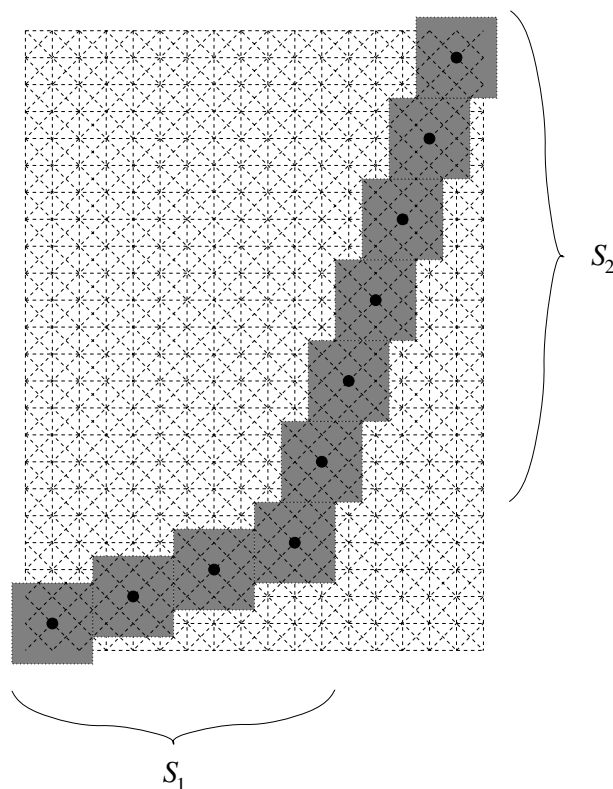


FIG. 7. Power dominating sets of products of path of similar length.

Case 8. $3m - n - 6 \equiv 7 \pmod{8}$. Then $3(m + 1) - (n + 2) - 6 \equiv 0 \pmod{8}$. Take the power dominating set S in $P_{m+1} \boxtimes P_{n+2}$ and remove the fibers ${}^0P_n$, P_m^0 , and P_m^{n+1} . We have $|S| = (n + 2 + m + 1 - 2)/4 = \lceil (n + m - 2)/4 \rceil$.

4. The lexicographic product. The remaining standard graph product to consider is the lexicographic one. For it, the power domination problem is much easier. In fact we will determine the power domination number for any lexicographic product in terms of the domination number and the total domination number of its factors.

Recall that for a graph $G = (V, E)$, a set $S \subseteq V$ is a *dominating set* if each vertex in $V \setminus S$ is adjacent to at least one vertex of S . If in addition each vertex of S has a neighbor in S , then S is called a *total dominating set*. The *domination* (resp., *total domination*) *number* $\gamma(G)$ (resp., $\gamma_t(G)$) of G is the minimum cardinality of a dominating (resp., total dominating) set. Then we have the next theorem.

THEOREM 4.1. *Let G be a nontrivial graph without isolated vertices. Then for any nontrivial graph H ,*

$$\gamma_P(G \circ H) = \begin{cases} \gamma(G); & \gamma_P(H) = 1, \\ \gamma_t(G); & \gamma_P(H) > 1. \end{cases}$$

Proof. Suppose $\gamma_P(H) = 1$. Let $\{v\}$ be a power dominating set of H and D a dominating set of G . Then $D \times \{v\}$ is a power dominating set of $G \circ H$. Indeed, for any vertex u of G , if $u \notin D$, any vertex of ${}^uH$ is in the neighborhood of (u', v) , for

some $u' \in D$ with $uu' \in E(G)$, so it is monitored. If $u \in D$, any neighbor of a vertex of ${}^uH$ not in ${}^uH$ is monitored, and $\{(u, v)\} \cup N(u, v)$ is monitored. So since $\{v\}$ is a power dominating set of H , the fiber is monitored. Therefore $\gamma_P(G \circ H) \leq \gamma(G)$.

Assume that there is a power dominating set S of $G \circ H$ that contains less than $\gamma(G)$ vertices. Then there is an H -fiber ${}^uH$ that contains no vertex of $S \cup N(S)$. So the vertices of ${}^uH$ are monitored by the propagation. Consider the first propagation that occurs to a vertex (u, v) of ${}^uH$. The vertex (u, v) must be the only neighbor not monitored of some vertex not in ${}^uH$. But this other vertex is also adjacent to all the other vertices of ${}^uH$, which are not monitored yet. Since H has more than one vertex, this implies that there can be no propagation on ${}^uH$. Hence $\gamma_P(G \circ H) \geq \gamma(G)$.

Suppose now $\gamma_P(H) > 1$. Let D be a total dominating set of G . Then for any vertex v of H , $D \times \{v\}$ is a dominating set of $G \circ H$ and so a power dominating set of H of cardinality $|D| = \gamma_t(G)$. Thus $\gamma_P(G \circ H) \leq \gamma_t(G)$.

Let S be a minimum power dominating set of $G \circ H$. Suppose there is an H -fiber ${}^uH$ that contains at least two vertices of S . Let S' be obtained by removing from S all vertices of ${}^uH$ but one and adding an arbitrary vertex of a neighboring H -fiber (if there is none yet). Then $N[S] \subseteq N[S']$, so S' is a power dominating set of $G \circ H$, with $|S'| \leq |S|$. Repeating this process if necessary, we may now assume that every H -fiber of $G \circ H$ contains at most one vertex in S .

Suppose that there exists an H -fiber ${}^uH$ such that for any neighbor v of u in G , ${}^vH \cap S$ is empty. So the vertices monitored by P in ${}^uH$ are exactly the vertices monitored by $S \cap {}^uH$. Since there is no power dominating set of H of cardinality 1, $S \cap {}^uH$ contains at least two vertices, a contradiction. Thus, any vertex $u \in V(G)$ has a neighbor v in G such that the H -fiber ${}^vH$ contains a vertex of S . So the set $\{v \in V(G) \mid {}^vH \cap S \neq \emptyset\}$ is a total dominating set of G and we conclude that $\gamma_P(G \circ H) = |S| \geq \gamma_t(G)$. $\square$

In Theorem 4.1 we have assumed that G has no isolated vertices. In the case that G contains r isolated vertices, let G' be the subgraph of G induced by nonisolated vertices. Then $G = G' \cup rK_1$ and $G \circ H$ is the disjoint union of $G' \circ H$ and k copies of H . Consequently, $\gamma_P(G \circ H) = \gamma_P(G' \circ H) + r\gamma_P(H)$.

Acknowledgments. We thank Sylvain Gravier for introducing us to the concept of the power domination and for useful conversations. We are also grateful to one of the referees for a careful reading of the manuscript.

REFERENCES

- [1] D. ATKINS, T. W. HAYNES, AND M. A. HENNING, *Placing monitoring devices in electric power networks modelled by block graphs*, Ars Combin., 79 (2006), pp. 129–143.
- [2] T. L. BALDWIN, L. MILI, M. B. BOISEN, JR., AND A. ADAPA, *Power system observability with minimal phasor measurement placement*, IEEE Trans. Power Syst., 8 (1993), pp. 707–715.
- [3] M. DORFLING AND M. A. HENNING, *A note on power domination in grid graphs*, Discrete Appl. Math., 154 (2006), pp. 1023–1027.
- [4] J. GUO, R. NIEDERMEIER, AND D. RAIBLE, *Improved Algorithms and Complexity Results for Power Domination in Graphs*, in Lecture Notes in Comput. Sci. 3623, 2005, pp. 172–184.
- [5] R. HAMMACK, *Isomorphic components of direct products of bipartite graphs*, Discuss. Math. Graph Theory, 26 (2006), pp. 231–248.
- [6] T. W. HAYNES, S. M. HEDETNIEMI, S. T. HEDETNIEMI, AND M. A. HENNING, *Domination in graphs applied to electric power networks*, SIAM J. Discrete Math., 15 (2002), pp. 519–529.
- [7] T. W. HAYNES, S. T. HEDETNIEMI, AND P. J. SLATER, EDS., *Fundamentals of Domination in Graphs*, Marcel Dekker, Inc. New York, 1998.
- [8] W. IMRICH AND S. KLAVŽAR, *Product Graphs: Structure and Recognition*, John Wiley & Sons, New York, 2000.

- [9] P. K. JHA, S. KLAVŽAR, AND B. ZMAZEK, *Isomorphic components of Kronecker product of bipartite graphs*, Discuss. Math. Graph Theory, 17 (1997), pp. 301–309.
- [10] J. KNEIS, D. MÖLLE, S. RICHTER, AND P. ROSSMANITH, *Parameterized power domination complexity*, Inform. Process. Lett., 98 (2006), pp. 145–149.
- [11] C.-S. LIAO AND D.-T. LEE, *Power Domination Problem in Graphs*, in Lecture Notes in Comput. Sci. 3595, 2005, pp. 818–828.
- [12] P. M. WEICHSEL, *The Kronecker product of graphs*, Proc. Amer. Math. Soc., 13 (1962), pp. 47–52.
- [13] G. XU, L. KANG, E. SHAN, AND M. ZHAO, *Power domination in block graphs*, Theoret. Comput. Sci., 359 (2006), pp. 299–305.
- [14] M. ZHAO, L. KANG, AND G. J. CHANG, *Power domination in graphs*, Discrete Math., 306 (2006), pp. 1812–1816.

PLANAR GRAPHS OF ODD-GIRTH AT LEAST 9 ARE HOMOMORPHIC TO THE PETERSEN GRAPH*

Z. DVOŘÁK†, R. ŠKREKOVSKI‡, AND T. VALLA†

Abstract. Let G be a graph and let $c : V(G) \rightarrow \binom{1, \dots, 5}{2}$ be an assignment of 2-element subsets of the set $1, \dots, 5$ to the vertices of G such that for every edge vw , the sets $c(v)$ and $c(w)$ are disjoint. We call such an assignment a $(5, 2)$ -coloring. A graph is $(5, 2)$ -colorable if and only if it has a homomorphism to the Petersen graph. The *odd-girth* of a graph G is the length of the shortest odd cycle in G (∞ if G is bipartite). We prove that every planar graph of odd-girth at least 9 is $(5, 2)$ -colorable, and thus it is homomorphic to the Petersen graph. Also, this implies that such graphs have a fractional chromatic number at most $\frac{5}{2}$. As a special case, this result holds for planar graphs of girth at least 8.

Key words. planar graphs, coloring, homomorphism, girth

AMS subject classification. 05C15

DOI. 10.1137/060650507

1. Introduction. One of the most natural generalizations of the usual graph coloring is the k -tuple coloring. We assign to each vertex a set of k distinct colors, instead of just one color, and require that the adjacent vertices obtain disjoint sets of colors. Such an assignment is called a k -tuple coloring, or a k -tuple n -coloring if n colors are used (where $k \leq n$). To be brief, we use the term (n, k) -coloring instead. Obviously, an $(n, 1)$ -coloring is just the usual n -coloring.

It is easy to see that these multicolorings are homomorphisms to Kneser graphs. The *Kneser graph* $K(n, k)$ (where $2k \leq n$) is a graph with a vertex set consisting of all k -element subsets of $\{1, 2, \dots, n\}$ and with two vertices adjacent if and only if the subsets are disjoint. An (n, k) -coloring of G also can be viewed as a homomorphism of G to $K(n, k)$. Let us note that $K(5, 2)$ is isomorphic to the Petersen graph.

The *fractional chromatic number* of G , denoted $\chi_f(G)$, is the infimum of the fractions n/k such that G admits an (n, k) -coloring. The notion can also be alternatively defined using the linear relaxation of the integer programming formulation of the graph coloring problem. For a fixed k , the minimum n such that G admits an (n, k) -coloring is called the k -tuple chromatic number of G and it is denoted by $\chi^k(G)$. Therefore, $\chi_f(G) = \inf \chi^k(G)/k$. More details about this coloring can be found in the monograph of Scheinerman and Ullman [6].

Pirnazar and Ullman [5] showed that the fractional chromatic number of a planar graph with girth at least $8n - 4$ is at most $2 + \frac{1}{n}$, in particular, that planar graphs with girth 12 have a fractional chromatic number at most $\frac{5}{2}$. We improve this result by showing that every planar graph of odd-girth at least 9 (and thus also every

*Received by the editors January 20, 2006; accepted for publication (in revised form) September 13, 2007; published electronically March 21, 2008. This work was supported in part by bilateral projects SLO-CZ/04-05-002 and MSMT-07-0405 between Slovenia and the Czech Republic.

<http://www.siam.org/journals/sidma/22-2/65050.html>

†Faculty of Mathematics and Physics, DIMATIA, and Institute for Theoretical Computer Science (ITI), Charles University, Malostranské nám. 2/25, 118 00, Prague, Czech Republic (rakdver@kam.mff.cuni.cz, valla@kam.mff.cuni.cz). These authors were supported in part by project 1M0545 of the Czech Ministry of Education.

‡Department of Mathematics, University of Ljubljana, Jadranska 19, 1111 Ljubljana, Slovenia (bluesky2high@yahoo.com). This author was supported in part by Ministry of Science and Technology of Slovenia Research Program P1-0297.

planar graph of girth 8) is $(5, 2)$ -colorable, and hence it has a fractional chromatic number at most $\frac{5}{2}$. Another interpretation of this result is that such graphs have a homomorphism to the Petersen graph. Regarding the lower bounds, Pirnazar and Ullman [5] show that there exists a planar graph of girth 5 with fractional chromatic number $\frac{13}{5}$. This implies that our result cannot be improved to odd-girth 5. We leave the case of odd-girth 7 as an open problem.

Our result also has connections to several important conjectures regarding flows and circular chromatic number. Let C be a circle of (Euclidean) length r . An r -circular coloring of a graph G is a mapping c which assigns to each vertex v of G an open unit length arc $c(v)$ of C , such that for every edge $xy \in E(G)$, $c(x) \cap c(y) = \emptyset$. We say a graph G is r -circular colorable if there is an r -circular coloring of G . The circular chromatic number of a graph, denoted by $\chi_c(G)$, is defined as

$$\chi_c(G) = \inf\{r \mid G \text{ is } r\text{-circular colorable}\}.$$

The following well-known inequalities hold for each graph G :

$$\chi(G) - 1 < \chi_c(G) \leq \chi(G) \quad \text{and} \quad \chi_f(G) \leq \chi_c(G) \leq \chi(G).$$

It is also well known that the circular chromatic number is always a rational number. For more details on the circular chromatic number see [10].

The concept of (n/k) -circular colorings is closely related to the concept of (n, k) -flows. Let G be a graph and D an orientation of G . For positive integers k and $n \geq 2k$, an (n, k) -flow f of D is a mapping that assigns to each edge e of D an integer $f(e) \in \{k, k+1, \dots, n-k\}$ such that $\sum_{e \in E^+(v)} f(e) = \sum_{e \in E^-(v)} f(e)$ for every vertex v ; Here $E^+(v)$ is the set of edges incident to v that are directed away from v , and $E^-(v)$ is the set of edges incident to v that are directed towards v in D . A graph G is said to admit an (n, k) -flow if G has an orientation D that admits an (n, k) -flow.

Notice that $(n, 1)$ -flows are precisely nowhere-zero integer n -flows introduced by Tutte [7, 8]. One can find more about flows in the monograph of Zhang [9] and the survey of Jaeger [2]. The following well-known conjecture was proposed by Jaeger [3].

CONJECTURE 1 (circular-flow conjecture). *For any integer $k \geq 1$, every $4k$ -edge-connected graph admits a $(2k+1, k)$ -flow.*

Note that the circular-flow conjecture for $k = 1$ and $k = 2$ implies the famous 3-flow conjecture and 5-flow conjecture of Tutte, respectively. For planar graphs, the flow problem can be dualized to the circular coloring problem. More precisely, an (n/k) -circular coloring of a planar graph G corresponds to an (n, k) -flow of the dual graph of G . Therefore, the restriction of Jaeger's conjecture to planar graphs is equivalent to the following.

CONJECTURE 2. *Every planar graph G of girth at least $4k$ has a circular chromatic number at most $2 + \frac{1}{k}$.*

The odd-edge connectivity of a graph G is the size of the smallest odd edge-cut of G . The dual of odd-edge connectivity is an odd-girth. The odd-girth of a graph G is the length of the shortest odd cycle in G (∞ if G is bipartite). Klostermeyer and Zhang [4] proposed a strengthening of Conjecture 2, where the edge-connectivity condition is replaced by an odd edge-connectivity condition as follows.

CONJECTURE 3. *Every planar graph G of odd-girth at least $4k+1$ has a circular chromatic number at most $2 + \frac{1}{k}$.*

Regarding Conjecture 3, Zhu [11] proved that for every $k \geq 2$, planar graphs of odd-girth at least $8k-3$ have a circular chromatic number at most $2 + \frac{1}{k}$. Later Borodin et al. [1] proved that a planar graph with girth at least $\frac{20k-2}{3}$ has a circular

chromatic number at most $2 + \frac{1}{t}$. Both results imply that planar graphs of girth at least 13 are $(5, 2)$ -colorable.

Conjecture 3 for $k = 2$ implies that the fractional chromatic of each planar graph of odd-girth at least 9 is at most $\frac{5}{2}$. Our result shows that this conclusion is indeed true.

2. Main theorem. The following theorem is the main result presented in this paper.

THEOREM 1. *Every planar graph of odd-girth at least 9 is $(5, 2)$ -colorable.*

To prove this theorem, we first show that we may restrict our attention to graphs without any short cycles. To deal with short faces, we use the following lemma proved by Klostermeyer and Zhang [4].

LEMMA 2 (folding lemma). *Let G be a planar graph with odd-girth $g > 3$. If $C = v_0v_1 \dots v_{r-1}$ is a facial circuit of G with $r \neq g$, then there is an integer $i \in \{0, \dots, r-1\}$ such that the graph G' obtained from G by identifying v_{i-1} and v_{i+1} (where indices are taken modulo r) is also of odd-girth g .*

To handle short cycles that are not faces, we show that under some assumptions, a coloring of the exterior of such a cycle can be extended to its interior as follows.

THEOREM 3. *Let H be a 2-connected planar graph such that H does not contain a cycle of length smaller than 9, except possibly for its outer face O_H whose length is at least 4. The graph H is $(5, 2)$ -colorable. Furthermore, if the length of O_H is at most 8, then any proper $(5, 2)$ -coloring of O_H can be extended to a proper coloring of H .*

Assuming that Theorem 3 holds, we can prove Theorem 1 easily. Given a plane graph G and a cycle C , let $\text{Int}_G(C)$ be the subgraph of G consisting of vertices and edges that lie on C or inside C , and let $\text{Out}_G(C)$ be the subgraph of G consisting of vertices and edges that lie on C or outside C . Thus, $\text{Int}_G(C)$ and $\text{Out}_G(C)$ share only the vertices and edges of C .

Proof of Theorem 1. For the sake of contradiction, let us assume that Theorem 1 is false and let G be a counterexample with minimum number of vertices. Obviously, G is connected. We may also assume that G is 2-connected; otherwise we can color each block separately and permute the colors so that the colorings agree on the cut-vertices.

Let us show that each face of G has size at least 9. Let $g \geq 9$ be the odd-girth of G . Assume that G contains a face of length different from g . By Lemma 2, there exist vertices v_1 and v_2 on that face such that the graph G' obtained from G by identifying v_1 and v_2 is a planar graph of odd-girth g . Note that the vertices v_1 and v_2 are not adjacent because of the restriction on the odd-girth of G . By the minimality of G , the graph G' is $(5, 2)$ -colorable. However, since v_1 and v_2 are not adjacent, this also gives a proper $(5, 2)$ -coloring of G , which is a contradiction.

If G does not contain any cycle of length less than 9, then G is $(5, 2)$ -colorable by Theorem 3. Therefore, assume that G contains a short cycle. Let C be a cycle in G of length at most 8 such that the number of vertices of $\text{Int}_G(C)$ is minimal. Since G does not contain any face of length less than 9, the graph $\text{Out}_G(C)$ has less vertices than G . By the minimality of G , there exists a proper $(5, 2)$ -coloring c of $\text{Out}_G(C)$. The graph $\text{Int}_G(C)$ does not contain any cycle shorter than 9, and thus the coloring c can be extended to the vertices of $\text{Int}_G(C)$ by Theorem 3. This shows that G can be $(5, 2)$ -colored, which is a contradiction. $\square$

It remains to prove Theorem 3. We proceed by the discharging method. Let us assume that H is a minimal counterexample to the theorem throughout the rest of the paper, and let O_H be its outer face. We first assign charge to vertices and

faces of H in such a way that the total amount of charge is negative. The bound on the length of faces ensures that the charge of each face is nonnegative, and we never change the charge of faces, and hence we know that at each moment the total charge of vertices is negative. The initial distribution of the charge is described in section 4. We then redistribute the charge, and assuming that H is a minimal counterexample to Theorem 3, we show that the final charge of each vertex is nonnegative. Since the amount of charge does not change during its redistribution, we obtain a contradiction. This will establish Theorem 3.

In section 3, we describe some subgraphs that cannot appear in H ; we call such subgraphs (configurations) *reducible*. The proof of reducibility of a configuration C usually proceeds in the following way. By the minimality of H , we know that $H - C$ can be properly colored. We then show that an arbitrary coloring of $H - C$ can be extended to C , thus showing that H is (5,2)-colorable and obtaining a contradiction.

The discharging proceeds in two phases. Let us call vertices of degree at least 3 *important*. In the first phase (section 5), we consider vertices of degree 2 in H and make their charge nonnegative by moving charge from adjacent important vertices. We then analyze the charge of important vertices and show that the subgraph H' induced by the important vertices with negative or small positive charge has a very special structure—in particular, it has maximum degree at most 3, and only vertices of degree at most 2 may have negative charge.

In section 7, we study the reducible configurations in H' . This enables us to restrict the structure of H' . Finally, we run the second phase of discharging over H' , moving the charge from vertices of degree three to vertices of smaller degree in H' , and show that one of 3-vertices is contained in a reducible configuration (section 8).

Throughout the proof, we need to handle the precolored outer face O_H . We would run into problems if a vertex v of a reducible configuration C were precolored, since we would be unable to prove that the coloring of $H - C$ can be extended in a way that is consistent with the color of v . To avoid this problem, we add an extra charge to the vertices of O_H while still preserving the fact that the total amount of charge is negative. This extra charge ensures that the charge of the vertices of O_H is large enough, and thus they do not appear in any configuration we consider—only configurations whose total charge is negative are important in the discharging method.

3. Reducible configurations. We say that a configuration (a subgraph) is *reducible* if it cannot appear in the minimal counterexample H , except possibly if some of its vertices are precolored. To use the discharging method we first need to prove that some configurations are reducible. We start with the following lemma that enables us to handle vertices of degree 2 in H .

LEMMA 4. *The following claims hold:*

- (1) *Let $P = v_1v_2v_3$ be a path of H such that $d(v_2) = 2$ and v_2 is not a precolored vertex of the outer face. If c is a coloring of $H - v_2$ such that $c(v_1) \cap c(v_3) \neq \emptyset$, then c can be extended to a proper coloring of H .*
- (2) *Let $P = v_1v_2v_3v_4$ be a path of H such that $d(v_2) = d(v_3) = 2$ and neither v_2 nor v_3 is precolored. If c is a coloring of $H - \{v_2, v_3\}$ such that $c(v_1) \neq c(v_4)$, then the coloring c can be extended to a proper coloring of H .*
- (3) *If P is a path $v_1v_2v_3v_4v_5$ in H such that $d(v_2) = d(v_3) = d(v_4) = 2$ and neither of v_2, v_3 , and v_4 is precolored, then P is reducible.*

Proof. Let us prove each claim separately:

- (1) Let $X = \{1, \dots, 5\} \setminus [c(v_1) \cup c(v_2)]$. The size of X is at least two, and thus we can choose $c(v_2) \subseteq X$.

- (2) If $c(v_1) \cap c(v_4) = \emptyset$, then set $c(v_2) = c(v_4)$ and $c(v_3) = c(v_1)$. Otherwise $c(v_1) = \{a, b_1\}$ and $c(v_4) = \{a, b_2\}$ have exactly one common element a . Let $\{b_3, b_4\} = \{1, \dots, 5\} \setminus \{a, b_1, b_2\}$. We set $c(v_2) = \{b_2, b_3\}$ and $c(v_3) = \{b_1, b_4\}$.
- (3) Suppose that this configuration appears in H . There exists a coloring c of $H - \{v_2, v_3, v_4\}$ by the minimality of H . Let $c(v_2)$ be a color-pair disjoint with $c(v_1)$ and distinct from $c(v_5)$ and use the second claim of this lemma to color v_3 and v_4 . In this way, we extend the coloring to H , thus obtaining a contradiction. $\square$

We need to show that many subgraphs are reducible. For each such subgraph C , we consider the coloring of $H - C$ and extend it to C . We usually use counting arguments of the following type: If a vertex v of C is adjacent to exactly one vertex u of $H - C$, the vertex v can be colored by one of three colors that are not forbidden by the color of u . If we determine the numbers of available colors for each vertex of C in a similar way, and show that C can be colored from all possible lists of colors of the given length, then we prove that C is reducible. To make the statements and the proofs of the theorems simpler, we introduce a notion closely related to the list coloring.

Let $\mathcal{P} = \binom{\{1, \dots, 5\}}{2}$ be the set of all unordered pairs of the elements of $\{1, \dots, 5\}$ and let us call the elements of $\mathcal{P}$ *color-pairs*. A *list* is a subset of $\mathcal{P}$, and if its size is k , then we call it a *k-list*. Lists are used to specify which color-pairs can be used at vertices in the considered configurations. For a given $X \subseteq \{1, \dots, 5\}$, let $\text{ss}(X) = \{p \in \mathcal{P}; p \subseteq X\}$ be the list of all 2-element subsets of X . Let $\text{int}(X) = \{p \in \mathcal{P}; p \cap X \neq \emptyset\}$ be the list of all color-pairs that intersect X .

A *type* is a set of lists. Types are used to express properties of the possible colorings of the considered configurations. For example, if we know that a vertex v has one neighbor precolored with a color-pair p , then the color-pairs that can be used to color v must be disjoint with p , and the list of such color-pairs belongs to the type $\mathcal{N}_3$ defined below. Let us point out the following special types:

- $\mathcal{P}_{\text{all}} = \mathcal{P}$ is the set of all lists, and $\mathcal{P}_k$ is the set of all k -lists. Thus, $\mathcal{P}_{10} = \{\mathcal{P}\}$.
- $\mathcal{N}_4 \subset \mathcal{P}_4$ is the set of all 4-lists such that no color belongs to all color-pairs in the list. For example, the list $L = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}\}$ does not belong to $\mathcal{N}_4$.
- $\mathcal{N}_3 \subset \mathcal{P}_3$ is the type of all 3-lists that can be expressed as $\text{ss}(X)$ for some 3-element subset X of $\{1, \dots, 5\}$.
- $\mathcal{N}_7 \subset \mathcal{P}_7$ is the set of all 7-lists, whose complements belong to $\mathcal{N}_3$. In other words, $\mathcal{N}_7$ is the type of all lists that can be expressed as $\text{int}(X)$ for some 2-element subset X of $\{1, \dots, 5\}$.

If $\mathcal{A}$ and $\mathcal{B}$ are two types, we write $\mathcal{B} \geq \mathcal{A}$ if every list in $\mathcal{B}$ has a (not necessarily proper) subset that belongs to $\mathcal{A}$. That is, the lists in $\mathcal{B}$ are “greater” than the lists in $\mathcal{A}$. In particular, if $\mathcal{B} \geq \mathcal{A}$ and a coloring can be extended to a vertex whose list of available colors belongs to $\mathcal{A}$, then the coloring can also be extended to the same vertex with a list from $\mathcal{B}$. Obviously, if $\mathcal{B} \subseteq \mathcal{A}$, then $\mathcal{B} \geq \mathcal{A}$. Also, $\mathcal{P}_m \geq \mathcal{P}_n$ for $m \geq n$.

LEMMA 5. *The following relations hold:*

- (1) $\mathcal{P}_7 \geq \mathcal{N}_3$, and
- (2) $\mathcal{P}_9 \geq \mathcal{N}_7$.

Proof. Consider the first claim. We may assume that there exists a set X of seven color-pairs that does not contain a subset in $\mathcal{N}_3$. Each of the color-pairs has two elements; therefore, there must be a color that belongs to at least three color-pairs

in X . We may assume that these color-pairs are $\{1, 2\}$, $\{1, 3\}$, and $\{1, 4\}$. Hence, the color-pairs $\{2, 3\}$, $\{2, 4\}$, and $\{3, 4\}$ cannot belong to X . By the size of X , all remaining color-pairs must belong to X . In particular, $\{1, 5\}$ and $\{2, 5\}$ belong to X and together with $\{1, 2\}$ they form a list from $\mathcal{N}_3$, which is a contradiction.

Let us now prove the second claim. Let X be an arbitrary list in $\mathcal{P}_9$, i.e., $X = \mathcal{P} \setminus \{p\}$ for a color-pair p . Let Y be any 2-element subset of $\{1, \dots, 5\} \setminus p$. The set $\text{int}(Y) \in \mathcal{N}_7$ is a subset of X , witnessing that $\mathcal{P}_9 \geq \mathcal{N}_7$. $\square$

We say that a color-pair p_1 *forbids* a color-pair p_2 if p_1 and p_2 are not disjoint, and that a list L *forbids* a color-pair p if every color-pair in L forbids p . This means that if vertices u and v are adjacent and the color of u belongs to L , then the color of v cannot be p ; conversely, if the color of v is p , then we cannot color u from the list L . We often use a counting argument—if we determine that less than 10 color-pairs can be forbidden at a vertex v by a coloring of some of its neighbors, such a coloring can be extended to the vertex v . A color-pair that is not forbidden is called *free*.

So far, we were using the terms “configuration” and “reducible” intuitively. Let us now define them formally using the concept of types. A *configuration* C is a graph $G_C = (V_C, E_C)$ together with a function $\text{type}_C : V_C \rightarrow \wp(\mathcal{P}_{\text{all}})$ that assigns a type to each vertex of C . Let us remark that we only consider configurations that do not contain precolored vertices of the face O_H . The configuration represents an induced subgraph of H whose reducibility we study. We want to show that any precoloring of the vertices of $V(H) \setminus V_C$ can be extended to a $(5, 2)$ -coloring of the whole H . The types of the vertices of the configuration reflect the constraints imposed by colors of vertices of $V(H) \setminus V_C$ —each proper $(5, 2)$ -coloring of $V(H) \setminus V_C$ gives us a list of free colors at each vertex v of C , and the type of v in the configuration consists of all such lists.

An *instance* Q of the configuration C is a function from V_C to $\mathcal{P}_{\text{all}}$ such that $Q(v) \in \text{type}_C(v)$ for each vertex v . The instance Q is *colorable* if there exists a proper coloring c of G_C such that $c(v) \in Q(v)$ for each vertex of C . The configuration C is *reducible* if each instance of C is colorable.

Note that if $\mathcal{X}$ and $\mathcal{Y}$ are types with $\mathcal{X} \geq \mathcal{Y}$ and we change the type of a vertex in a reducible configuration from $\mathcal{Y}$ to $\mathcal{X}$, then the new configuration is reducible as well. We use the following notation to simplify the descriptions of configurations. If $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_k$ are types, then $\text{path}(\mathcal{X}_1 \mathcal{X}_2 \dots \mathcal{X}_n)$ is a path $v_1 v_2 \dots v_n$ on n vertices with $\text{type}(v_i) = \mathcal{X}_i$ for each i . Similarly, $\text{cycle}(\mathcal{X}_1 \mathcal{X}_2 \dots \mathcal{X}_n)$ is a cycle $v_1 v_2 \dots v_n$ on n vertices with $\text{type}(v_i) = \mathcal{X}_i$.

LEMMA 6. *The following configurations are reducible:*

- (1) $\text{path}(\mathcal{P}_1 \mathcal{P}_8)$,
- (2) $\text{path}(\mathcal{P}_2 \mathcal{P}_6)$,
- (3) $\text{path}(\mathcal{P}_3 \mathcal{P}_5)$, and
- (4) $\text{path}(\mathcal{N}_4 \mathcal{P}_4)$.

Proof. Observe that a configuration $\text{path}(\mathcal{X} \mathcal{P}_n)$ is reducible if and only if every list in $\mathcal{X}$ forbids at most $n - 1$ color-pairs. In the rest of this proof, let a, b, c, d , and e be mutually distinct elements of $\{1, \dots, 5\}$. Let us consider the configurations separately as follows:

- (1) A single color-pair $\{a, b\}$ forbids exactly the seven color-pairs of $\text{int}(\{a, b\})$; therefore, $\text{path}(\mathcal{P}_1 \mathcal{P}_8)$ is reducible.
- (2) If the two color-pairs are disjoint, say $\{a, b\}$ and $\{c, d\}$, then they forbid four color-pairs $\{a, c\}$, $\{a, d\}$, $\{b, c\}$, and $\{b, d\}$. Otherwise they share a common element, say, the pairs are $\{a, b\}$ and $\{a, c\}$; then they forbid five color-pairs—the four color-pairs that contain a and the color-pair $\{b, c\}$. Therefore, $\text{path}(\mathcal{P}_2 \mathcal{P}_6)$ is reducible.

(3) Without loss of generality, there are the following four possibilities for three color-pairs:

- $\{a, b\}$, $\{a, c\}$, and $\{a, d\}$, forbidding the four color-pairs that contain a .
- $\{a, b\}$, $\{a, c\}$, and $\{b, c\}$, forbidding three color-pairs $\{a, b\}$, $\{a, c\}$, and $\{b, c\}$.
- $\{a, b\}$, $\{a, c\}$, and $\{b, d\}$, forbidding three color-pairs $\{a, b\}$, $\{a, d\}$, and $\{b, c\}$.
- $\{a, b\}$, $\{a, c\}$, and $\{d, e\}$, forbidding two color-pairs $\{a, d\}$ and $\{a, e\}$.

Therefore, $\text{path}(\mathcal{P}_3 \mathcal{P}_5)$ is reducible.

(4) Note that if a list X is a subset of a list Y and X forbids n color-pairs, then Y forbids at most n color-pairs. If L is any list from $\mathcal{N}_4$, then there is a list $L' \subset L$ of size three such that the intersection of elements of L' is empty. But such a list L' forbids at most three color-pairs, as follows from the proof of the previous claim of this lemma. Therefore, L also forbids at most three color-pairs, and hence $\text{path}(\mathcal{N}_4 \mathcal{P}_4)$ is reducible. $\square$

Before we proceed with the next lemma, let us introduce the following convention: We use $\mathcal{X}^*$ inside a path or a cycle description of a configuration for an arbitrarily long (possibly empty) chain of vertices of type $\mathcal{X}$. For example, $\text{path}(\mathcal{P}_2 \mathcal{P}_7^* \mathcal{P}_9)$ may be any of $\text{path}(\mathcal{P}_2 \mathcal{P}_9)$, $\text{path}(\mathcal{P}_2 \mathcal{P}_7 \mathcal{P}_9)$, $\text{path}(\mathcal{P}_2 \mathcal{P}_7 \mathcal{P}_7 \mathcal{P}_9)$, etc.

Note that a chain of vertices of type $\mathcal{P}_7$ “propagates” the type of the vertex adjacent to it. For example, let $uv_1v_2 \dots v_kw$ be a path, let the type of u be $\mathcal{P}_6$ (or any other type $\mathcal{X} \geq \mathcal{P}_6$), and let the type of each v_i ($1 \leq i \leq k$) be $\mathcal{P}_7$. Let Q be an instance of this configuration. The list $Q(u)$ forbids at most one color-pair q , because $\text{path}(\mathcal{P}_6 \mathcal{P}_2)$ is reducible by Lemma 6(2). Let $Q_1 = Q(v_1) \setminus \{q\}$ be the list of v_1 without the forbidden color-pair. The size of Q_1 is at least six, and thus it also forbids at most one color-pair q_1 in the list of v_2 . Similarly the list $Q_2 = Q(v_2) \setminus \{q_1\}$ forbids at most one color-pair in the list of v_3 , etc. Finally, the nonforbidden part of $Q(v_k)$ forbids at most one color-pair in $Q(w)$. This in particular means that if $|Q(w)| > 1$, we may color the vertex w by a remaining color and then choose a free color for $v_k, v_{k-1}, \dots, v_1$, and u . Thus, we can say that $\text{path}(\mathcal{P}_6 \mathcal{P}_7^*)$ behaves as $\text{path}(\mathcal{P}_6)$. Similarly, since a list in $\mathcal{P}_5$ forbids at most two color-pairs by Lemma 6(3), $\text{path}(\mathcal{P}_5 \mathcal{P}_7^*)$ behaves as $\text{path}(\mathcal{P}_5)$, and since a list in $\mathcal{P}_2$ forbids five color-pairs by Lemma 6(2), $\text{path}(\mathcal{P}_2 \mathcal{P}_7^*)$ behaves as $\text{path}(\mathcal{P}_2)$. The lists in types $\mathcal{P}_4$ and $\mathcal{P}_3$ may forbid four color-pairs by Lemma 6(3), so $\text{path}(\mathcal{P}_4 \mathcal{P}_7^*)$ and $\text{path}(\mathcal{P}_3 \mathcal{P}_7^*)$ behave as $\text{path}(\mathcal{P}_3)$.

LEMMA 7. *The following configurations are reducible:*

- (1) $\text{path}(\mathcal{P}_2 \mathcal{P}_8 \mathcal{P}_5)$,
- (2) $\text{path}(\mathcal{P}_2 \mathcal{P}_7^* \mathcal{P}_{10} \mathcal{P}_7^* \mathcal{P}_2)$, and
- (3) $\text{path}(\mathcal{P}_2 \mathcal{P}_7^* \mathcal{P}_9 \mathcal{P}_7^* \mathcal{P}_3)$.

Proof. Let us prove the reducibility of each of the configurations separately.

- (1) The reducibility of the configuration $\text{path}(\mathcal{P}_2 \mathcal{P}_8 \mathcal{P}_5)$ is a simple consequence of Lemma 6(2), (3). A vertex with a list from type $\mathcal{P}_2$ forbids at most five color-pairs. A vertex with a list from type $\mathcal{P}_5$ forbids at most two color-pairs. Therefore, there is at least one free color-pair in the list of the vertex of type $\mathcal{P}_8$.
- (2) Since the type of vertices propagates through the chain of vertices of type $\mathcal{P}_7$, it is sufficient to show that $\text{path}(\mathcal{P}_2 \mathcal{P}_{10} \mathcal{P}_2)$ is reducible. Let $v_1v_2v_3$ be the configuration and let Q be an instance of the configuration. A vertex of type $\mathcal{P}_2$ forbids at most five color-pairs. If either $Q(v_1)$ or $Q(v_3)$ forbids less than five color-pairs at v_2 , then the instance is colorable. By the proof of Lemma 6(2), a list $L \in \mathcal{P}_2$ can forbid five color-pairs only if the color-pairs in

L intersect each other. Suppose that both $Q(v_1)$ and $Q(v_3)$ forbid five color-pairs. Let a be the intersection of color-pairs in $Q(v_1)$ and b the intersection of color-pairs in $Q(v_3)$. If $a \neq b$, then let $p = \{a, b\}$; otherwise let p be any color-pair that contains a . Observe that p is forbidden by both $Q(v_1)$ and $Q(v_3)$. This means that the total number of forbidden color-pairs in $Q(v_2)$ is at most nine; hence the instance is colorable.

- (3) We argue similarly to the previous claim that it suffices to show that $\text{path}(\mathcal{P}_2 \mathcal{P}_9 \mathcal{P}_3)$ is reducible. We use an argument similar to the previous case. Let $v_1 v_2 v_3$ be the configuration and let Q be an instance of the configuration. A vertex of type $\mathcal{P}_2$ forbids at most five color-pairs and a vertex of type $\mathcal{P}_3$ forbids at most four color-pairs. Therefore, if the instance is not colorable, $Q(v_1)$ must forbid precisely five color-pairs and $Q(v_3)$ must forbid precisely four color-pairs. The proofs of Lemma 6(2), (3) show that this may happen only if both $\bigcap_{p \in Q(v_1)} p \neq \emptyset$ and $\bigcap_{p \in Q(v_3)} p \neq \emptyset$. Let a be the element of intersection of $Q(v_1)$ and b the element of intersection of $Q(v_3)$. If $a \neq b$, then let $p = \{a, b\}$; otherwise let p be any color-pair that contains a . Observe that p is forbidden by both $Q(v_1)$ and $Q(v_3)$. This means that the total number of forbidden color-pairs in $Q(v_2)$ is at most eight, which shows that the instance is colorable. $\square$

LEMMA 8. *The configurations $\text{path}(\mathcal{P}_2 \mathcal{P}_7^* \mathcal{P}_9 \mathcal{P}_7^* \mathcal{P}_8 \mathcal{P}_7^* \mathcal{P}_2)$ are reducible.*

Proof. Let $v_1 v_2 \dots v_n$ be one of the configurations described by the statement of this lemma, and let Q be its instance. Since $\mathcal{P}_2$ propagates over a chain of vertices of type $\mathcal{P}_7^*$, it suffices to consider the case when v_2 has type $\mathcal{P}_9$ and v_{n-1} has type $\mathcal{P}_8$.

By Lemma 7(3), the configuration $\text{path}(\mathcal{P}_2 \mathcal{P}_9)$ on $v_1 v_2$ forbids at most two color-pairs at v_3 . If $n = 4$, then v_3 is of type $\mathcal{P}_8$, and hence $Q(v_3)$ contains at least six free color-pairs. This instance of the configuration $\text{path}(\mathcal{P}_6 \mathcal{P}_2)$ on $v_3 v_4$ is colorable by Lemma 6(2). If $n > 4$, then the type of v_3 is $\mathcal{P}_7$ and $Q(v_3)$ contains at least five free color-pairs. Since the type $\mathcal{P}_5$ propagates over chain of vertices of type $\mathcal{P}_7$, it suffices to show that the configuration $\text{path}(\mathcal{P}_5 \mathcal{P}_8 \mathcal{P}_2)$ is reducible. This follows by Lemma 7(1). $\square$

LEMMA 9. *The cyclic configurations $\text{cycle}(\mathcal{N}_3 X \mathcal{N}_3 \mathcal{P}_{10} \mathcal{P}_{10})$, where X is an arbitrary path consisting of one vertex of type $\mathcal{P}_9$ and at least three vertices of type $\mathcal{N}_7$, are reducible.*

Proof. Let $v_1 v_2 \dots v_k$ be one of these configurations, where v_1 and v_{k-2} are the vertices of type $\mathcal{N}_3$ and v_{k-1} and v_k have type $\mathcal{P}_{10}$. By symmetry, we may assume that the types of v_2 and v_3 are $\mathcal{N}_7$. Let Q be an instance of this configuration. We may assume that $Q(v_1) = \text{ss}(\{1, 2, 3\})$. By Lemma 4(2), any coloring of the vertices v_1 and v_{k-2} by distinct color-pairs can be extended to v_{k-1} and v_k .

Suppose first that $Q(v_{k-2}) \neq Q(v_1)$, and thus $|Q(v_{k-2}) \cap Q(v_1)| \leq 1$. Now restrict the configuration to the vertices $v_1, \dots, v_{k-2}$ and create an instance Q' such that $Q'(v_1) = Q(v_1) \setminus Q(v_{k-2})$ (i.e., a list that consists of at least two color-pairs) and $Q'(v_i) = Q(v_i)$ for $i > 1$. Every coloring of Q' can be extended to a coloring of Q by Lemma 4(2), and the instance Q' is colorable by Lemma 7(3). Hence, Q is colorable as well.

In the rest of the lemma, we assume that $Q(v_{k-2}) = Q(v_1)$. We know that $Q(v_2) = \text{int}(p)$, where p is some color-pair. Let us now consider the case that $p \notin Q(v_1)$. Without loss of generality, we may assume that either $p = \{1, 4\}$ or $p = \{4, 5\}$. Restrict the configuration to vertices $v_2, \dots, v_{k-2}$ and create an instance Q' such that $Q'(v_2) = \{\{1, 4\}, \{1, 5\}, \{4, 5\}\} \subset Q(v_2)$, $Q'(v_{k-2}) = \{\{1, 2\}, \{1, 3\}\}$ and $Q'(v_i) =$

$Q(v_i)$ for $2 < i < k - 2$. It is easy to check that any coloring of Q' can be extended to a coloring of Q , since for each color-pair in $Q'(v_2)$ we may color v_1 with $\{2, 3\}$. The instance Q' is colorable by Lemma 7(3). Thus, Q is colorable as well.

Now we consider the case $p \in Q(v_1)$. Without loss of generality, we may assume $p = \{1, 2\}$ and $Q(v_3) = \text{int}(q)$, where q is one of the following color-pairs: $\{1, 2\}$, $\{1, 3\}$, $\{1, 5\}$, $\{3, 5\}$, or $\{4, 5\}$. Let us fix $c(v_1) = \{1, 3\}$; then, the free color-pairs at v_2 are $\{2, 4\}$ and $\{2, 5\}$. Consider the configuration restricted to vertices $v_3, \dots, v_{k-2}$. Let Q' be the instance such that $Q'(v_{k-2}) = \{\{1, 2\}, \{2, 3\}\}$, $Q'(v_i) = Q(v_i)$ for $4 < i < k - 2$ and $Q'(v_3)$ is the list of color-pairs that are not forbidden by the free color-pairs $\{\{2, 4\}, \{2, 5\}\}$ of v_2 . That is,

- if $1 \in q$, then $\{\{1, 3\}, \{1, 4\}, \{1, 5\}\} \subseteq Q'(v_3)$.
- if $q = \{3, 5\}$ or $q = \{4, 5\}$, then $\{\{1, 5\}, \{3, 4\}, \{3, 5\}\} \subseteq Q'(v_3)$.

The instance Q' is colorable by Lemma 7(3). We have chosen the list of v_3 such that the color $c(v_3)$ forbids at most one color-pair at v_2 ; thus we can color v_2 by $\{2, 4\}$ or $\{2, 5\}$. Since $c(v_1) \neq c(v_{k-2})$, this coloring can be extended by Lemma 4(2) to the coloring of the whole configuration; thus finishing the proof of this lemma. $\square$

We define the following operations on types:

- $\text{cut}_1(\mathcal{X}) = \{X'; (\exists X \in \mathcal{X}) X' \subseteq X, |X \setminus X'| \leq 1\}$, i.e., the lists in this set can be obtained from those in $\mathcal{X}$ by removing at most one color-pair.
- $\text{cut}_3(\mathcal{X}) = \{X'; (\exists X \in \mathcal{X}) (\exists p \in \mathcal{P}) X' = X \cap \text{int}(p)\}$, i.e., the lists in this set can be obtained from those in $\mathcal{X}$ by removing color-pairs that are disjoint with a color-pair p .
- $\text{cut}_7(\mathcal{X}) = \{X'; (\exists X \in \mathcal{X}) (\exists p \in \mathcal{P}) X' = X \setminus \text{int}(p)\}$, i.e., the lists in this set can be obtained from those in $\mathcal{X}$ by removing color-pairs that intersect a color-pair p .

Let us write cut_a^k for applying k times the operation cut_a ; in particular, cut_a^0 is the identity. These operations are used to reflect the effect of precoloring a vertex of a configuration to the lists of remaining vertices—for example, if a list of a vertex v belongs to type $\mathcal{X}$ and we color a neighbor of v , the list of color-pairs that are not forbidden by this precoloring belongs to type $\text{cut}_7(\mathcal{X})$. Note that the operations commute with each other; i.e., $\text{cut}_a(\text{cut}_b(\mathcal{X})) = \text{cut}_b(\text{cut}_a(\mathcal{X}))$ for each $a, b \in \{1, 3, 7\}$. The index of the cut operator describes the number of colors removed from the lists; i.e., at most a colors are removed from each list of $\mathcal{X}$ in $\text{cut}_a(\mathcal{X})$ for $a \in \{1, 3, 7\}$.

LEMMA 10. *The following relations hold:*

- (1) $\text{cut}_1(\mathcal{P}_n) \geq \mathcal{P}_{n-1}$,
- (2) $\text{cut}_3(\mathcal{P}_{10}) = \mathcal{N}_7$,
- (3) $\text{cut}_7(\mathcal{P}_{10}) = \mathcal{N}_3$, and
- (4) $\text{cut}_3^2(\mathcal{P}_{10}) \geq \mathcal{N}_4$.

Proof. The first three claims are trivial. Let us consider the fourth one. Let $L \in \text{cut}_3^2(\mathcal{P}_{10})$. Note that $|L| \geq 4$. If $|L| \geq 5$, then L obviously contains four color-pairs whose intersection is empty. Thus, we need only consider the case when $|L| = 4$. This happens only when the color-pairs p_1 and p_2 used to obtain L from $\mathcal{P}_{10}$ by cut_3 operations are disjoint. We may assume that $p_1 = \{1, 2\}$ and $p_2 = \{3, 4\}$. The list $L = \{\{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}\}$ belongs to $\mathcal{N}_4$, thus proving the fourth relation. $\square$

A vertex is called *important* if its degree is at least 3, and an edge is called *important* if both of its vertices are important. Let us call a path with k inner vertices of degree 2 connecting two important vertices a k -link. In particular, an important edge is a 0-link. Links with $k > 0$ are called *proper*. By Lemma 4(3), only 0-, 1-, and 2-links may appear in H .

Let G be a fixed subgraph and v an important vertex belonging to G . We say that G is *closed* with respect to v if all links incident to v are either completely contained in G (including the other important vertex of the link) or share only the vertex v with G . For such a vertex v , let $U(v)$ be the set of important vertices that do not belong to G and that are joined with v by a link, and let $U^*(v)$ be the set consisting of v vertices in $U(v)$ and the vertices of the links connecting v to vertices of $U(v)$. Suppose that v is joined by n_k k -links to vertices in $U(v)$ for $0 \leq k \leq 2$. Let $\mathcal{U}_v = \text{cut}_7^{n_0}(\text{cut}_3^{n_1}(\text{cut}_1^{n_2}(\mathcal{P}_{10})))$. We say that G is closed if it is closed with respect to all important vertices in G .

The following lemma shows how we can determine the type of a vertex v in a configuration C from the links through that v is connected to vertices that do not belong to C . This lemma basically restates Lemma 4 and tells us that links to vertices of C can be eliminated and replaced by setting the types of vertices of the configuration using the appropriate cut operations.

LEMMA 11. *Let v be an important vertex of a subgraph G such that G is closed with respect to v . Let c be a proper coloring of the graph induced by $U(v)$. If L is the list of color-pairs p such that $c(v) = p$ for some extension of c to the subgraph induced by $U^*(v)$, then $L \in \mathcal{U}_v$.*

Proof. If $u \in U(v)$ is adjacent to v , then the color-pair $c(u)$ forbids all color-pairs that intersect $c(u)$ at v . By Lemma 4, if u is joined to v by a 1-link, then the color-pair $c(u)$ forbids all color-pairs that are disjoint with $c(u)$. Finally, if u is joined to v by a 2-link, then the color-pair $c(u)$ forbids the color-pair $c(u)$ at v . The operations cut_7 , cut_3 , and cut_1 modify the lists in the type they are applied to in the same way, and since the lists for all possible choices of the colors assigned to the vertices of $U(v)$ are included in $\mathcal{U}_v$, we infer that $L \in \mathcal{U}_v$. $\square$

If we determine the types of important vertices of an induced subgraph of H using Lemma 11 and the resulting configuration is reducible, then we obtain a contradiction with the fact that H is a counterexample to Theorem 3, as the following lemma shows. This provides us with a powerful tool for determining the subgraphs that cannot appear in H .

LEMMA 12. *Let G be a closed subgraph without precolored vertices. Let C be the configuration on G , where the type of each important vertex v of C is $\mathcal{U}_v$ and the type of each 2-vertex is $\mathcal{P}_{10}$. If the configuration C is reducible, then G cannot appear in H .*

Proof. Suppose for the sake of contradiction that G appears in H . Let $W = V_C \cup \bigcup_{v \in V_C} (U^*(v) \setminus U(v))$ be the set consisting of V_C and vertices of degree 2 contained in proper links that share only one vertex with C . Let $V' = V(H) \setminus W$. By the minimality of H , the subgraph induced by V' has a proper $(5, 2)$ -coloring c_1 .

Let v be an important vertex of C . Let F_v be the list of color-pairs that

- intersect the color-pair $c_1(u)$ for some neighbor u of v belonging to V' , or
- are disjoint with the color-pair $c_1(u)$ for some $u \in V'$ that is connected with v by a 1-link, or
- are equal to the color-pair $c_1(u)$ for some $u \in V'$ that is connected with v by a 2-link.

Let $L_v = \mathcal{P} \setminus F_v$. By Lemma 11, $L_v \in \mathcal{U}_v$. Let Q be the instance of C such that $Q(v) = L_v$ for each important vertex $v \in V(C)$ and $Q(v) = \mathcal{P}$ for each 2-vertex v of C . Since C is reducible, the instance Q is colorable. Let c_2 be a proper coloring c_2 of this instance.

Note that for two adjacent important vertices $u \in V'$ and $v \in C$, the color-pairs $c_1(u)$ and $c_2(v)$ are disjoint. Let $W' = W \setminus V_C$ be the set of 2-vertices that belong

to the links joining C with V' . Note that the vertices of W' do not belong to the precolored outer face of H , since the vertices of G do not belong to it either. Observe that we have chosen the lists L_v in such a way that the union of the colorings c_1 and c_2 can be extended to W' , using Lemma 4(1), (2). This yields a proper $(5, 2)$ -coloring of H , which is a contradiction. $\square$

4. Initial charge. We assign a charge to vertices and faces of H . For every $v \in V(H)$, we define the initial vertex charge $\text{ch}_0(v)$ as

$$\text{ch}_0(v) = \frac{7}{3}d(v) - 6,$$

where $d(v)$ denotes the degree of v in the graph H . Let $F(H)$ be the set of faces of the graph H . For every face $f \in F(H)$, distinct from O_H , we define the initial face charge $\text{ch}_0(f)$ as

$$\text{ch}_0(f) = \frac{2}{3}r(f) - 6,$$

where $r(f)$ denotes the length of f . The charge of the outer face O_H is set to $\frac{2}{3}r(O_H) + \frac{11}{2}$. By Euler's formula, the total amount of charge is

$$\begin{aligned} \sum_{v \in V(H)} \text{ch}_0(v) + \sum_{f \in F(H)} \text{ch}_0(f) &= \left(\frac{14}{3}|E(H)| - 6|V(H)| \right) \\ &\quad + \left(\frac{4}{3}|E(H)| - 6|F(H)| \right) + \frac{23}{2} \\ &= 6(|E(H)| - |V(H)| - |F(H)|) + \frac{23}{2} \\ &= -\frac{1}{2}. \end{aligned}$$

Here we see the charge assigned to the vertices of small degrees:

$d(v)$	2	3	4	5	6
$\text{ch}_0(v)$	$-\frac{4}{3}$	1	$\frac{10}{3}$	$\frac{17}{3}$	8

Note that the charge of each face f distinct from O_H is at least $\frac{2}{3}r(f) - 6 \geq 0$, since $r(f) \geq 9$. We redistribute the charge of the vertices in two phases such that the total amount of charge does not change. However, we show that in the minimal counterexample H , the final charge of each vertex and face is nonnegative, thus contradicting the existence of the minimal counterexample and proving Theorem 3.

5. First phase of discharging. In the first phase, we ensure that all 2-vertices have a nonnegative charge. We use the following redistribution rules to move a charge from the important vertices to links:

- (R1) If xyz is a 1-link in H such that x has degree 3, 4, or 5 and y is not precolored, then x sends $\frac{2}{3}$ to y .
- (R2) If $xyzv$ is a 2-link in H such that x has degree 3, 4, or 5 and y is not precolored, then x sends $\frac{4}{3}$ to y .
- (R3) If x is a vertex of degree at least 6 adjacent to a 2-vertex y , and y is not precolored, then x sends $\frac{4}{3}$ to y .
- (R4) If the length r of O_H is at most 8, the face O_H sends $\frac{4}{3}$ to each vertex incident to it.

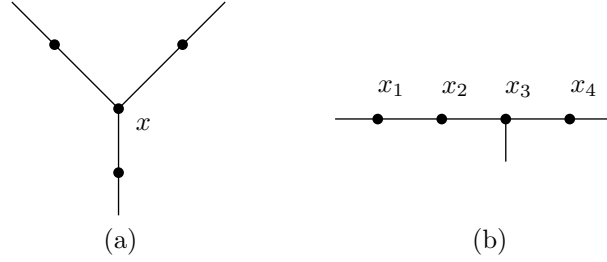


FIG. 1. Reducible 3-vertex subgraphs.

The charge after the first phase is denoted by ch_1 . The charge of 2-vertices becomes nonnegative after the first phase. Furthermore, note that the charge of a vertex w of degree at least 6 is nonnegative:

$$\text{ch}_1(w) \geq \frac{7}{3}d(w) - 6 - \frac{4}{3}d(w) = d(w) - 6 \geq 0.$$

If the face O_H is precolored, then each vertex w of O_H has at least two neighbors to that it does not send any charge, and hence its charge is at least

$$(5.1) \quad \text{ch}_1(w) \geq \frac{7}{3}d(w) - 6 - \frac{4}{3}(d(w) - 2) + \frac{4}{3} = d(w) - 2 \geq 0.$$

The charge of inner faces does not change, and thus it is still nonnegative. The charge of O_H changes only if its length r is at most 8, and then the final charge of O_H is at least $(\frac{2}{3}r + \frac{11}{2}) - \frac{4}{3}r \geq \frac{1}{6}$. Therefore, the charge of each face after the first phase of discharging is nonnegative, and hence the sum of the charges of vertices is negative. In the rest of the paper, we only take into account the charge of vertices.

Let us now analyze the possible neighborhoods of important vertices of H . We first consider neighborhoods of 3-vertices.

LEMMA 13. *Subgraphs (a) and (b) in Figure 1 cannot appear in H unless the 3-vertex is precolored.*

Proof. Consider first the subgraph (a). Using Lemma 12, the type of x after the links are eliminated is $\text{cut}_3^3(\mathcal{P}_{10}) \geq \mathcal{P}_1$. Since a configuration consisting of a single vertex of type $\mathcal{P}_1$ is reducible, this subgraph cannot appear in H .

Now consider the subgraph (b). Let v_1, v_3 , and v_4 be the vertices adjacent to x_1, x_3 , and x_4 outside of the subgraph, respectively. Let c be a coloring of $H - \{x_1, x_2\}$. Note that $c(v_3) \neq c(v_4)$ because of the vertices x_3 and x_4 . The set of colors p such that the coloring c' defined by $c'(v_3) = c(v_3)$, $c'(v_4) = c(v_4)$, and $c'(x_3) = p$ is proper and can be extended to x_4 is $\text{int}(c(v_4)) \setminus \text{int}(c(v_3))$, and since $c(v_3) \neq c(v_4)$, this set has size at least two. Hence, we can change colors of x_3 and x_4 so that $c(x_3) \neq c(v_1)$. We extend the coloring c to vertices x_1 and x_2 using Lemma 4. Since H is not $(5, 2)$ -colorable, the subgraph (b) cannot appear in H . $\square$

The complete list of possible subgraphs in the neighborhood of a nonprecolored 3-vertex that do not contain one of these two reducible subgraphs is given in Figure 2. The final charge and the type of the central vertex when all incident proper links are eliminated (as in Lemma 11) are also shown in the figure. The edges of 1-links and 2-links are drawn thin and the important edges are drawn thick.

Let us now consider neighborhoods of 4-vertices.

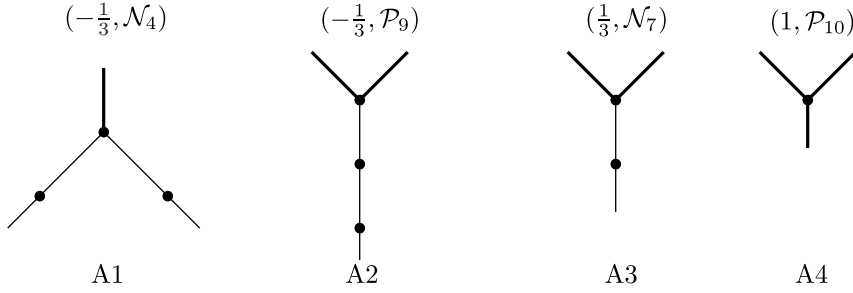


FIG. 2. 3-vertex neighborhoods.

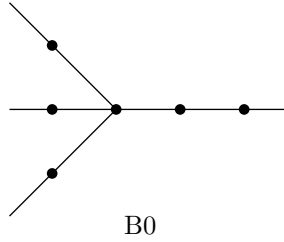


FIG. 3. A reducible 4-vertex subgraph.

LEMMA 14. If A_1 , A_2 , and A_3 are three subsets of $\{1, \dots, 5\}$ of size three, then there exists a color-pair that is a subset of two of the sets A_1 , A_2 , and A_3 . Consequently, $\text{cut}_3^3(\mathcal{P}_{10}) \geq \mathcal{P}_2$.

Proof. We may assume that $A_1 = \{1, 2, 3\}$. No two of the sets A_1 , A_2 , and A_3 may be disjoint, and thus we assume that $|A_1 \cap A_2| = |A_1 \cap A_3| = 1$. But then $\{4, 5\} \subset A_2, A_3$. $\square$

LEMMA 15. The subgraph B0 in Figure 3 cannot appear in H unless the 4-vertex is precolored.

Proof. By Lemma 14, $\mathcal{X} = \text{cut}_3^3(\mathcal{P}_{10}) \geq \mathcal{P}_2$. The type of the important vertex of the configuration is $\text{cut}_1(\mathcal{X}) \geq \mathcal{P}_1$, and the subgraph cannot appear in H by Lemma 12. $\square$

The complete list of possible subgraphs in the neighborhood of a nonprecolored 4-vertex that do not contain B0 is given in Figure 4. Next, we study neighborhoods of nonprecolored 5-vertices.

LEMMA 16. The subgraph C0 in Figure 5 cannot appear in H unless the 5-vertex is precolored.

Proof. The type of the important vertex of the configuration is $\text{cut}_1^3(\text{cut}_3^2(\mathcal{P}_{10})) \geq \mathcal{P}_1$, and the subgraph cannot appear in H by Lemma 12. $\square$

Let $d_{\text{imp}}(v)$ be the number of important neighbors of a vertex v of H . We show that vertices of large degree have a charge of at least $\frac{2d_{\text{imp}}(v)}{3}$, which will provide enough charge to distribute in the second phase of discharging.

LEMMA 17. If v is a 5-vertex in H such that $\text{ch}_1(v) < \frac{2d_{\text{imp}}(v)}{3}$, then v belongs to the subgraph C1 depicted in Figure 6.

Proof. If $d_{\text{imp}}(v) = 0$, then v would be incident to at most two 2-links by Lemma 16, and hence $\text{ch}_1(v) \geq \frac{17}{3} - 2 \cdot \frac{4}{3} - 3 \cdot \frac{2}{3} = 1$. If $d_{\text{imp}}(v) \geq 2$, then $\text{ch}_1(v) \geq \frac{17}{3} - (5 - d_{\text{imp}}(v))\frac{4}{3} = \frac{4d_{\text{imp}}(v)}{3} - 1 \geq \frac{2d_{\text{imp}}(v)}{3} + \frac{1}{3}$.

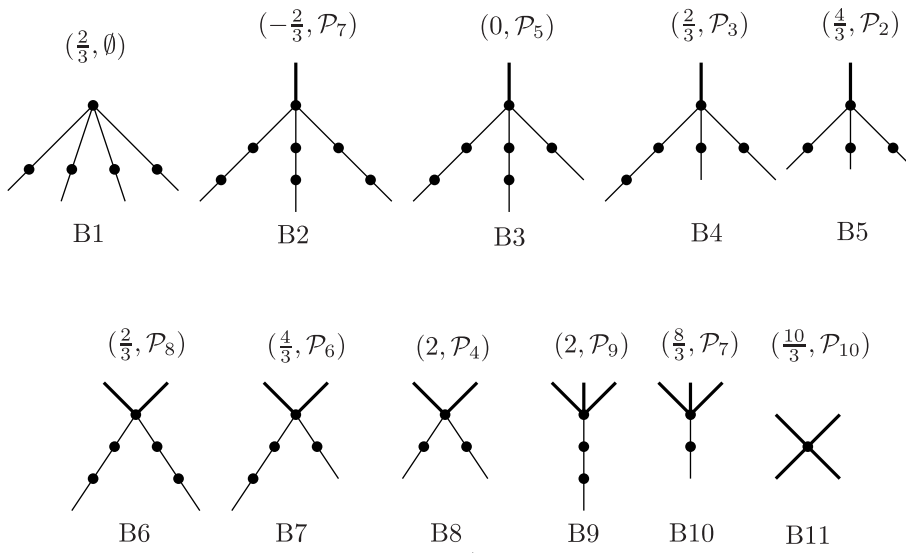


FIG. 4. 4-vertex neighborhoods.

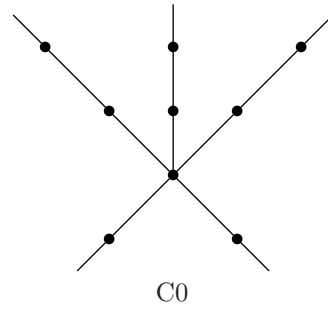


FIG. 5. A reducible 5-vertex subgraph.

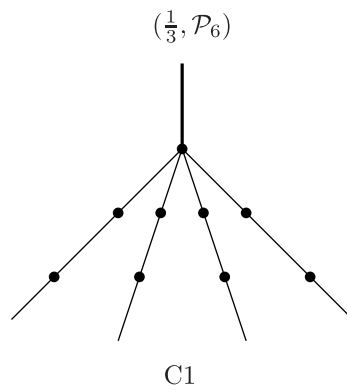


FIG. 6. Another 5-vertex subgraph.

Now we may assume that $d_{\text{imp}}(v) = 1$. If v is incident to a 1-link, then $\text{ch}_1(v) \geq \frac{17}{3} - 3 \cdot \frac{4}{3} - \frac{2}{3} = 1 > \frac{2}{3}$. Therefore, v is incident to four 2-links and belongs to the configuration C1. $\square$

Finally, we consider vertices of degree at least 6.

LEMMA 18. *Every vertex v of H of degree $d \geq 6$ has charge $\text{ch}_1(v) \geq \frac{2d_{\text{imp}}(v)}{3}$.*

Proof. The charge of v is at least

$$\frac{7}{3}d - 6 - (d - d_{\text{imp}}(v))\frac{4}{3} = d - 6 + \frac{4}{3}d_{\text{imp}}(v) \geq \frac{2d_{\text{imp}}(v)}{3}. \quad \square$$

6. Vertex classes and graph H' . After the first phase of discharging, some of the vertices still have a negative charge. In the second phase, we want to make the charge of all vertices nonnegative by moving the charge over the important edges. It turns out that it is always sufficient to move the charge at most $\frac{2}{3}$ over each edge, and thus we need only consider configurations consisting of vertices whose charge is smaller than $\frac{2}{3}d_{\text{imp}}$. We may classify the important vertices by type and charge after the first phase as follows:

- A. A 4-vertex v incident to one important edge and three 2-links (B2 in Figure 4). Its charge after the first stage is $-\frac{2}{3}$. By Lemma 11, when we consider reducibility of a configuration C containing v , and none of the three important vertices to that v is connected through the 2-links belongs to C , we may assume that the type of v is $\mathcal{P}_7$. For brevity we omit the references to Lemma 11 in the rest of the enumeration.
- B. A 3-vertex incident to one important edge and two 1-links (A1 in Figure 2). Its charge is $-\frac{1}{3}$, and when the 1-links are omitted, its type is $\mathcal{N}_4$.
- C. A 4-vertex incident to one important edge, one 1-link, and two 2-links (B3 in Figure 4), or a 5-vertex incident to one important edge and four 2-links (C1 in Figure 6). The charge of such a vertex is at least 0 and when the proper links are omitted, its type is $\geq \mathcal{P}_5$.
- D. A 3-vertex incident to two important edges and one 2-link (A2 in Figure 2). Its charge is $-\frac{1}{3}$ and when the 2-link is omitted, its type is $\mathcal{P}_9$.
- E. A 3-vertex incident to two important edges and one 1-link (A3 in Figure 2). Its charge is $\frac{1}{3}$ and when the 1-link is omitted, its type is $\mathcal{N}_7$.
- F. A 4-vertex incident to two important edges and two 2-links (B6 in Figure 4). Its charge is $\frac{2}{3}$ and when the 2-links are omitted, its type is $\mathcal{P}_8$.
- G. A 3-vertex incident to three important edges (A4 in Figure 2). Its charge is 1.
- H. A vertex incident to d important edges and with charge at least $\frac{2d}{3}$.

To distinguish the items in this classification from “types” introduced earlier, we say that the vertex described in the item $\mathcal{X}$ of this classification has *class* $\mathcal{X}$. If $\mathcal{X}$ and $\mathcal{Y}$ are classes, let us say that a vertex v is of class $(\mathcal{X}\mathcal{Y})$ if v is of class $\mathcal{X}$ or $\mathcal{Y}$.

We now want to restrict our attention to only the subgraphs of H that do not have enough charge. In order to do this, we create a new planar multigraph H in that we replace the parts of the graph H that have enough charge by vertices of class $\mathcal{H}$, which turn out to be easy to handle. More precisely, we obtain H' from H by iterating the following operation as long as possible: If there is an induced closed connected subgraph G with at least two important vertices such that G is joined to the rest of the graph by d important edges and the total charge of vertices in G is $c \geq \frac{2d}{3}$, then we contract this subgraph to a single vertex v . We keep the parallel edges that might arise during this operation, but we remove loops. We let the charge of v be c and its class $\mathcal{H}$. This operation does not change the total charge of the vertices, and thus it

is enough to prove that the total charge of vertices of H' is nonnegative in order to obtain a contradiction. Note that the construction does not necessarily determine H' uniquely, but it ensures that any closed connected subgraph of H' with at least two important vertices has a charge of at most $\frac{2}{3}$ per outgoing edge.

The following crucial observation enables us to work with the graph H' instead of H in the rest of the paper: The configurations whose reducibility we are going to consider do not contain vertices of class $\mathcal{H}$. We also never pose any assumptions on whether the vertices adjacent to the configurations we consider are distinct or not in the following proofs; in particular, decontracting the vertices of class $\mathcal{H}$ does not affect the reducibility of the configurations. Therefore, if we find a reducible configuration in H' , the corresponding configuration in H is reducible as well, thus obtaining a contradiction with the minimality of H .

If the outer face O_H is precolored, we contract it to a vertex of class $\mathcal{H}$ as the first step in the construction of H' : The degree of the created vertex is $\sum_{v \in O_H} (d(v) - 2)$, and each vertex v of the face has a charge of at least $d(v) - 2$ by the analysis we performed in section 5 (inequality (5.1)). Additionally, note that any cycle in H' of length less than 9 contains a vertex of class $\mathcal{H}$, and all parallel edges must be incident with at least one vertex of class $\mathcal{H}$.

7. Further reducible configurations. In this section we focus on the reducibility of the configurations in H' that have a negative charge. A *string* is a path or a cycle (if it matters, we use the terms *path string* and *cycle string* to distinguish between these two possibilities) in H' whose edges are all important and whose vertices do not have class $(\mathcal{GH})$. Vertices of class $(\mathcal{ABC})$ may, of course, be only the end vertices of a path string, while the inner vertices of a string must be of class $(\mathcal{DEF})$. A path string that contains a vertex of class $(\mathcal{ABC})$ is called a *leaf string*. Note that we do not impose any restrictions on where the proper links from the vertices of the string lead to (i.e., they may connect vertices inside the string). We call a string *spanned* in case there exists a proper link such that both of its important vertices belong to the string. The *charge* of a string is the sum of the charges of its vertices.

We use the notation $\text{path}(T_1 \dots T_k)$ and $\text{cycle}(T_1 \dots T_k)$ for the strings whose vertices have classes $T_1, \dots, T_k$, in the order of their occurrence on the path or the cycle. We let $\mathcal{X}^k$ stand for a path of k vertices of class $\mathcal{X}$. We also use the following notation for describing the inner vertices of strings: We let $(\mathcal{D}, \mathcal{EF})^{m,n}$ mean a string that consists of m vertices of class $\mathcal{D}$ and n vertices of class $(\mathcal{EF})$, in any order. We write $(\mathcal{D}, \mathcal{EF})^{\leq m,n}$ for a string that consists of at most m vertices of class $\mathcal{D}$ and exactly n vertices of class $(\mathcal{EF})$, in any order, and similarly for other possible combinations of numbers and $\geq$ or $\leq$. We sometimes use this notation inside $\text{path}(\dots)$ or $\text{cycle}(\dots)$ to denote a substring in this form. For example, $\text{path}((\mathcal{D}, \mathcal{EF})^{2,1})$ is one of the strings $\text{path}(\mathcal{D}^2 \mathcal{E})$, $\text{path}(\mathcal{D}^2 \mathcal{F})$, $\text{path}(\mathcal{D} \mathcal{E} \mathcal{D})$, $\text{path}(\mathcal{D} \mathcal{F} \mathcal{D})$, $\text{path}(\mathcal{E} \mathcal{D}^2)$, and $\text{path}(\mathcal{F} \mathcal{D}^2)$. For consistency, we also sometimes use $(\mathcal{D}, \mathcal{EF})^{0,n}$ for strings that consist entirely of vertices of class $(\mathcal{EF})$, and $(\mathcal{D}, \mathcal{EF})^{n,0}$ for strings that consist of only vertices of class $\mathcal{D}$.

In the following lemmas, we describe the strings in H' that are not reducible. Lemmas 19–21 deal with the leaf strings that contain a vertex of class $\mathcal{A}$, $\mathcal{B}$, or $\mathcal{C}$. Lemma 22 considers the nonleaf path strings. Finally, Lemma 23 shows that H' does not contain any cycle strings with negative charge (and thus no cycle strings at all, by construction of H').

LEMMA 19. *The strings in H' that are not reducible and that contain a vertex of class $\mathcal{A}$ are of the form $\text{path}(\mathcal{A}(\mathcal{D}, \mathcal{EF})^{0,\leq 3})$. Moreover, these strings are not spanned.*

Proof. Let s be a string in H' that contains a vertex of class $\mathcal{A}$. Note that if s is one of the strings from the statement of the lemma, s cannot be spanned, since every cycle in H' of length less than 9 contains a vertex of class $\mathcal{H}$.

Suppose now that s is not one of the strings described in the statement. The string s cannot contain more than three vertices of class $(\mathcal{EF})$ in a row; otherwise the charge of the substring formed by the vertices of class $(\mathcal{EF})$ would be at least $\frac{4}{3}$ and this substring would become a vertex of class $\mathcal{H}$ during the construction of the graph H' . Therefore, s must contain a substring s' of form $\text{path}(\mathcal{A}(\mathcal{D}, \mathcal{EF})^{0, \leq 3} \mathcal{X})$, where the class $\mathcal{X}$ is different from $(\mathcal{EF})$ (i.e., $\mathcal{X}$ is one of $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, or $\mathcal{D}$).

Since H' does not contain a cycle of length less than 9 without a vertex of class $\mathcal{H}$, the string s' cannot be spanned. Therefore, s' is a closed subgraph of H' , and we may apply Lemma 11. Let us show that regardless of the choice of $\mathcal{X}$, the configuration C obtained from s' by the procedure described in statement of Lemma 11 is reducible. Let v be the vertex of s' with class $\mathcal{X}$.

According to the analysis we performed in section 6, the vertex of class $\mathcal{A}$ has type $\mathcal{P}_7$, and the vertices of class $(\mathcal{EF})$ have type at least $\mathcal{P}_7$. Similarly, if the class of v is $(\mathcal{ABC})$, its type is at least $\mathcal{P}_4$. If the class of v is $\mathcal{D}$, then v is incident to one edge and one 2-link outside of C , and thus its type is $\text{cut}_1(\text{cut}_7(\mathcal{P}_{10})) \geq \mathcal{P}_2$.

Therefore, it suffices to show that the configuration $\text{path}(\mathcal{P}_7^* \mathcal{P}_2)$ is reducible. By Lemma 6(2) and the fact that the type $\mathcal{P}_2$ propagates over vertices of type $\mathcal{P}_7$, this is indeed the case. It follows that s' cannot appear in H' . $\square$

LEMMA 20. *The strings in H' that are not reducible and that contain a vertex of class $\mathcal{B}$ are of the form $\text{path}(\mathcal{B}(\mathcal{D}, \mathcal{EF})^{0, \leq 2})$ or $\text{path}(\mathcal{B}(\mathcal{D}, \mathcal{EF})^{1, \leq 3})$. Moreover, these strings are not spanned.*

Proof. The string cannot contain a vertex of class $\mathcal{A}$ by Lemma 19. Let us consider several cases regarding the string of vertices of class $(\mathcal{DEF})$ adjacent to the vertex of class $\mathcal{B}$. Note that the string cannot contain more vertices of class $(\mathcal{EF})$ than specified in each of the cases; otherwise it would contain a subgraph with a charge of at least $\frac{2}{3}$ per outgoing important edge. None of the strings may be spanned; otherwise H' would contain a cycle of length at most 8 not containing a vertex of class $\mathcal{H}$.

$\text{path}(\mathcal{B}(\mathcal{D}, \mathcal{EF})^{0, \leq 1} \mathcal{B})$: By Lemma 11, it suffices to show that configurations $\text{path}(\mathcal{N}_4 \mathcal{N}_4)$ and $\text{path}(\mathcal{N}_4 \mathcal{P}_7 \mathcal{N}_4)$ are reducible. For the former one, this follows from Lemma 6(4). For the latter one, by Lemma 6(4), a list from $\mathcal{N}_4$ forbids at most three color-pairs, and hence we may always choose a color-pair that is not forbidden for the middle vertex.

$\text{path}(\mathcal{B}(\mathcal{D}, \mathcal{EF})^{\leq 1, \leq 1} \mathcal{C})$: The corresponding configuration $\text{path}(\mathcal{N}_4 \mathcal{P}_7^* \mathcal{P}_5)$ is reducible, as $\mathcal{P}_5$ is propagated over a chain of $\mathcal{P}_7$'s and the final configuration $\text{path}(\mathcal{N}_4 \mathcal{P}_5)$ is reducible by Lemma 6(4).

$\text{path}(\mathcal{B}(\mathcal{D}, \mathcal{EF})^{1, \leq 2} \mathcal{B})$: The corresponding configurations $\text{path}(\mathcal{N}_4 \mathcal{P}_7^* \mathcal{P}_9 \mathcal{P}_7^* \mathcal{N}_4)$ are reducible by Lemma 7(3), since $\mathcal{N}_4 \geq \mathcal{P}_3$.

$\text{path}(\mathcal{B}(\mathcal{D}, \mathcal{EF})^{1, \leq 3})$: This is one of the configurations described in the statement of the lemma.

A string with prefix $\text{path}(\mathcal{B}(\mathcal{D}, \mathcal{EF})^{1, \leq 3} \mathcal{D})$: Let s be this prefix. Note that s cannot be spanned: In order to avoid a cycle of length at most 8, the link would have to be a 2-link between the end-vertices of the path. However, a vertex of class $\mathcal{B}$ is incident only to 1-links. Consequently, the string $\text{path}(\mathcal{B}(\mathcal{D}, \mathcal{EF})^{1, \leq 3} \mathcal{D})$ corresponds to a configuration $\text{path}(\mathcal{N}_4 \mathcal{P}_7^* \mathcal{P}_9 \mathcal{P}_7^* \mathcal{P}_2)$ —the type $\mathcal{P}_2$ of the last vertex of the path is obtained from $\mathcal{P}_9$ (the type of vertex of class $\mathcal{D}$) by applying operation cut_7 , to take into account the important vertex adjacent to it outside of the configuration. This configuration is reducible by Lemma 7(3).

This shows that indeed the only such strings in H' that are not reducible are the nonspanned strings described in the claim of this lemma. $\square$

LEMMA 21. *The strings in H' that are not reducible and that contain a vertex of class $\mathcal{C}$ are of the form $\text{path}(\mathcal{C}(\mathcal{D}, \mathcal{EF})^{0, \leq 1})$ or $\text{path}(\mathcal{C}(\mathcal{D}, \mathcal{EF})^{1, \leq 2})$. Moreover, these strings are not spanned.*

Proof. Let s be a string in H' that contains a vertex of class $\mathcal{C}$. By Lemmas 19 and 20, the string s cannot contain a vertex of class $(\mathcal{AB})$. We distinguish several cases. Note that in each of the cases the considered string cannot contain more vertices of class $(\mathcal{EF})$ than specified; otherwise s would contain a subgraph with charge at least $\frac{2}{3}$ per outgoing important edge and this subgraph would be eliminated during the construction of H' . The considered strings also cannot be spanned, since the length of the cycle would be less than 9. If s contains at least two vertices of class $\mathcal{D}$, then s has a prefix $\text{path}(\mathcal{C}(\mathcal{D}, \mathcal{EF})^{1, \leq 2} \mathcal{D})$. However, the corresponding configuration $\text{path}(\mathcal{P}_5 \mathcal{P}_7^* \mathcal{P}_9 \mathcal{P}_7^* \mathcal{P}_2)$ is reducible by Lemma 7(3). Otherwise, if s contains two vertices of class $\mathcal{C}$, then s is $\text{path}(\mathcal{C}(\mathcal{D}, \mathcal{EF})^{\leq 1, 0} \mathcal{C})$, but the corresponding configurations $\text{path}(\mathcal{P}_5 \mathcal{P}_5)$ and $\text{path}(\mathcal{P}_5 \mathcal{P}_9 \mathcal{P}_5)$ are reducible by Lemmas 6(3) and 7(3). Therefore, the only possible strings are indeed those described in the statement of this lemma. $\square$

Finally, there is a classification of path strings that are composed from vertices of class $(\mathcal{DEF})$ only. To simplify the arguments we characterize only the strings whose charge is negative.

LEMMA 22. *The path strings in H' that are not reducible, contain only vertices of class $(\mathcal{DEF})$, and have negative charge are nonspanned strings of one of the following forms: $\text{path}(\mathcal{D})$, $\text{path}((\mathcal{D}, \mathcal{EF})^{2, \leq 1})$, or $\text{path}((\mathcal{D}, \mathcal{EF})^{3, \leq 2})$.*

Proof. A string with m vertices of class $\mathcal{D}$ may contain at most $m - 1$ vertices of class $(\mathcal{EF})$ in order for it to have a negative charge. The strings described in the statement of this lemma are nonspanned, because the length of the cycle that a spanning link would create is at most 7.

Let us assume that $s = v_1 \dots v_k$ is a string in H' with a negative charge that consists of only vertices of class $(\mathcal{DEF})$ and that contains at least four vertices of class $\mathcal{D}$. We want to show that s is reducible. We may assume that s is minimal; i.e., it does not contain a substring with more than three vertices of class $\mathcal{D}$ and negative charge.

By the minimality of s , its substring $s' = v_2 \dots v_k$ either has nonnegative charge or contains at most three vertices of class $\mathcal{D}$. Since the charge of vertices of classes $(\mathcal{EF})$ is positive, this means that v_1 is of class $\mathcal{D}$. Similarly, the class of v_k is $\mathcal{D}$. Furthermore, if s contains more than four vertices of class $\mathcal{D}$, then v_2 and v_{k-1} must also be of class $\mathcal{D}$; otherwise we can remove the first (or the last) two vertices from s .

Since the charge of a vertex of class $\mathcal{D}$ is $-\frac{1}{3}$, if the charge of s is less than $-\frac{1}{3}$, the string s must contain exactly four vertices of class $\mathcal{D}$; otherwise we can remove the first vertex from s . In case the charge of s is exactly $-\frac{1}{3}$, the string s contains at most six vertices of class $\mathcal{D}$ —if there are at least 7 vertices of class $\mathcal{D}$, let s_1 be the shortest prefix of s that contains four vertices of class $\mathcal{D}$, and let s_2 be the suffix of s that starts at the last vertex of class $\mathcal{D}$ that belongs to s_1 . Obviously s_1 and s_2 both contain at least four vertices of class $\mathcal{D}$. Let c be the charge of s and c_i be the charge of s_i for $i = 1, 2$. The strings s_1 and s_2 share exactly one vertex of class $\mathcal{D}$, and hence $c = c_1 + c_2 + \frac{1}{3}$. Therefore, at least one of c_1 and c_2 must be negative, contradicting the minimality of s .

The string s must be spanned; otherwise by Lemma 8, the configuration $\text{path}(\mathcal{P}_2 \mathcal{P}_7^* \mathcal{P}_9 \mathcal{P}_7^* \mathcal{P}_9 \mathcal{P}_7^* \mathcal{P}_2)$ corresponding to s is reducible. It follows that there

are no minimal strings with a charge less than $-\frac{1}{3}$, since such strings consist of four vertices of class $\mathcal{D}$ and at most two vertices of class $(\mathcal{EF})$, and such strings cannot be spanned, as H' does not contain a cycle of length at most 8 without a vertex of class $\mathcal{H}$. Thus, we may assume that the charge of s is $-\frac{1}{3}$.

Let us first consider the case that s contains exactly four vertices of class $\mathcal{D}$. Such a string must contain three vertices of class $\mathcal{E}$ in order to be spanned, and there must be a 2-link between v_1 and v_7 . But the resulting configuration is reducible by Lemma 9.

If s contains five vertices of class $\mathcal{D}$, we consider several possibilities:

- *The string s contains four vertices of class $\mathcal{E}$.* The path $v_1 \dots v_8$ is a substring of s with four vertices of class $\mathcal{D}$ and four vertices of class $\mathcal{E}$. Let us recall that v_1, v_2 , and v_8 are vertices of class $\mathcal{D}$. The vertices of the configuration cannot be joined by a 1-link, because the vertices of class $\mathcal{D}$ are incident only to 2-links and the vertices of class $\mathcal{E}$ are too close to each other. As the vertices of class $\mathcal{E}$ are not incident with 2-links, there must be a 2-link joining two vertices of class $\mathcal{D}$. Let v_i and v_j ($i < j$) be two vertices joined by a 2-link such that $j - i$ is the smallest possible. The difference $j - i$ must be at least 6, and at least one of $v_{i+1}, \dots, v_{j-1}$ must be of class $\mathcal{D}$, since there are only four vertices of class $\mathcal{E}$. Therefore, the configuration $v_i \dots v_j$ together with the 2-link is reducible by Lemma 9.
- *The string s contains two vertices of class $\mathcal{E}$ and one vertex of class $\mathcal{F}$.* Let v_a be the vertex of s of class $\mathcal{D}$ different from v_1, v_2, v_7 , and v_8 , and let v_b be the vertex of class $\mathcal{F}$. By symmetry, we may assume that $a < b$. The vertices of class $(\mathcal{EF})$ cannot be all consecutive in s , since they would form a substring with charge $\frac{4}{3}$ that would be replaced by a vertex of class $\mathcal{H}$ during the construction of H' , and hence $a > 3$. Consider the substring $v_a \dots v_8$. The vertices v_a and v_8 are of class $\mathcal{D}$, and there is one more vertex of class $\mathcal{D}$ and a vertex of class $\mathcal{F}$ inside the substring. This substring is too short to be spanned and it is reducible by Lemma 8.
- *The string s contains two vertices of class $\mathcal{F}$.* We consider the substring on vertices $v_1 \dots v_6$. This substring cannot be spanned and it contains four vertices of class $\mathcal{D}$, and thus it is reducible by Lemma 8.

Finally, consider the case that s contains six vertices of class $\mathcal{D}$ and at most five vertices of class $(\mathcal{EF})$. We know that classes of v_1 and v_2 are $\mathcal{D}$. If classes of v_3 and v_4 are both $(\mathcal{EF})$, the substring on vertices $v_4 \dots v_k$ has negative charge, contradicting the minimality of s . Similarly, it is not possible for v_3 to have class $\mathcal{F}$. Since the string $\text{path}(\mathcal{D}^4)$ is reducible by Lemma 8, s cannot contain four consecutive vertices of class $\mathcal{D}$, and thus exactly one of v_3 and v_4 has class $\mathcal{D}$. Similarly, exactly one of v_{k-3} and v_{k-2} is of class $\mathcal{D}$. Since s has charge $-\frac{1}{3}$, it contains either five vertices of class $\mathcal{E}$, three vertices of class $\mathcal{E}$, and one of class $\mathcal{F}$, or one vertex of class $\mathcal{E}$ and two of class $\mathcal{F}$.

If s contains five vertices of class $\mathcal{E}$, considering that s cannot contain four consecutive vertices of class $\mathcal{E}$ by the construction of H' , it follows that s is equal to $s_1 = \text{path}(\mathcal{D}^2 \mathcal{E} \mathcal{D} \mathcal{E}^3 \mathcal{D} \mathcal{E} \mathcal{D}^2)$. If s contains three vertices of class $\mathcal{E}$ and one of class $\mathcal{F}$, it similarly follows that s is equal to $s_2 = \text{path}(\mathcal{D}^2 \mathcal{E} \mathcal{D} \mathcal{E} \mathcal{F} \mathcal{D} \mathcal{E} \mathcal{D}^2)$. In case s contains one vertex of class $\mathcal{E}$ and two of class $\mathcal{F}$, the classes of v_3 and v_7 cannot be $\mathcal{F}$ by the minimality of s . However, the two vertices of class $\mathcal{F}$ would then have to be adjacent in s . This is not possible, as their total charge is $\frac{4}{3}$ and they would be eliminated during the construction of H' .

The string s_1 contains a substring $s'_1 = \text{path}(\mathcal{D}^2 \mathcal{E} \mathcal{D} \mathcal{E}^3 \mathcal{D})$. If s'_1 is nonspanned, then it is reducible by Lemma 8. There may be a 2-link between two vertices of

class $\mathcal{D}$, but in that case s_1 together with the link is reducible by Lemma 9. The string s_2 contains a substring $\text{path}(\mathcal{D}\mathcal{E}\mathcal{D}\mathcal{E}\mathcal{F}\mathcal{D})$ that cannot be spanned, and hence is reducible by Lemma 8.

Therefore, the only nonleaf path strings with negative charge in H' are those listed in the statement of this lemma. $\square$

Let us now show that no cycle strings appear in H' .

LEMMA 23. *H' does not contain a cycle string.*

Proof. A cycle string consists of a vertices of class $(\mathcal{D}\mathcal{E}\mathcal{F})$, and thus its length must be at least 9. If its charge is negative, then it contains a substring of length 6 with a negative charge. Such a string cannot be spanned and it is reducible by Lemma 22. A cycle string has no outgoing important edges and it forms a closed subgraph, and hence if its charge is nonnegative, then it would be replaced by a vertex of class $\mathcal{H}$ during the construction of H' . Therefore, H' does not contain any cycle strings. $\square$

Let us note that strings in H' do not have large negative charge.

LEMMA 24. *Each leaf string in H' has a charge of at least $-\frac{2}{3}$, and each nonleaf path string has a charge of at least -1 .*

Proof. Let us inspect all possible strings. By Lemma 19, the leaf strings that contain a vertex of class $\mathcal{A}$ are of form $\text{path}(\mathcal{A}(\mathcal{D}, \mathcal{E}\mathcal{F})^{0, \leq 3})$, and thus their charge is at least $-\frac{2}{3}$.

By Lemma 20, the leaf strings that contain a vertex of class $\mathcal{B}$ contain only one such vertex and at most one vertex of class $\mathcal{D}$, and hence their charge is at least $-\frac{2}{3}$.

By Lemma 21, the leaf strings that contain a vertex of class $\mathcal{C}$ contain at most one vertex of class $\mathcal{D}$, and hence their charge is at least $-\frac{1}{3}$.

By Lemma 22, a nonleaf path string with a negative charge contains at most three vertices of class $\mathcal{D}$, and thus its charge is at least -1 . $\square$

8. Second phase of discharging. Now we are ready to perform the second phase of discharging in the graph H' . Note that all vertices whose charge after the first phase is negative are inside strings. A path string $s = v_1v_2 \dots v_k$ is *adjacent* to a vertex v of class $(\mathcal{GH})$ if v and v_1 are adjacent. A string is *maximal* if it is not a proper substring of another string in H' . In this phase, the vertices of classes $(\mathcal{GH})$ send a charge to the adjacent maximal strings, thus making the charge of the vertices in the strings nonnegative. We then argue that the vertices of class $\mathcal{G}$ with negative charge must belong to reducible configurations. The rules for discharging are the following:

- (R5) A vertex of class $\mathcal{H}$ sends a charge of $\frac{2}{3}$ to each adjacent maximal string.
- (R6) A vertex of class $\mathcal{G}$ adjacent to a maximal leaf string of charge $c < 0$ sends a charge of $-c$ to the string.
- (R7) A vertex of class $\mathcal{G}$ adjacent to a maximal nonleaf string of charge $c < 0$ sends $-\frac{c}{2}$ to the string.

Each vertex of class $\mathcal{H}$ sends a charge of at most $\frac{2}{3}$ per important edge by (R5), and thus its final charge is nonnegative. By Lemma 24, each leaf string in H' has a charge of at least $-\frac{2}{3}$ and each nonleaf path string has a charge of at least -1 . Therefore, each maximal leaf or path string with charge $c < 0$ receives a charge of at least $-c$ from the adjacent vertices of class $(\mathcal{GH})$ and its final charge is nonnegative. By Lemma 23, the cycle strings do not appear in H' . It remains to show that also the charge of vertices of class $\mathcal{G}$ is nonnegative. Let us start with some definitions and lemmas.

By the classification of the strings presented in the previous section, the maximal path strings with negative charge in H' are nonspanned, and we can divide them into the following sets according to the charge they receive from the adjacent vertices of $\mathcal{G}$:

- $C_{\frac{2}{3}}$ is the set of strings to which vertices of class $\mathcal{G}$ send $\frac{2}{3}$, i.e., $\text{path}(\mathcal{A})$ and $\text{path}(\mathcal{D}\mathcal{B})$.
- $C_{\frac{1}{2}}$ is the set containing a single string $\text{path}(\mathcal{D}^3)$ to which vertices of class $\mathcal{G}$ send $\frac{1}{2}$.
- $C_{\frac{1}{3}}$ is the set of strings to which vertices of class $\mathcal{G}$ send $\frac{1}{3}$, consisting of $\text{path}(\mathcal{E}\mathcal{A})$, $\text{path}(\mathcal{B})$, $\text{path}(\mathcal{E}\mathcal{D}\mathcal{B})$, $\text{path}(\mathcal{D}\mathcal{E}\mathcal{B})$, $\text{path}(\mathcal{D}\mathcal{C})$, $\text{path}(\mathcal{D}^2)$, and all path strings that consist of three vertices of class $\mathcal{D}$ and one of class $\mathcal{E}$.
- $C_{\frac{1}{6}}$ is the set of strings to which vertices of class $\mathcal{G}$ send $\frac{1}{6}$, consisting of all path strings with n_d vertices of class $\mathcal{D}$, n_e vertices of class $\mathcal{E}$, and n_f vertices of class $\mathcal{F}$, where the triple (n_d, n_e, n_f) is one of $(1, 0, 0)$, $(2, 1, 0)$, $(3, 0, 1)$, or $(3, 2, 0)$.

Let $C_{\text{all}} = C_{\frac{2}{3}} \cup C_{\frac{1}{2}} \cup C_{\frac{1}{3}} \cup C_{\frac{1}{6}}$. We are going to consider configurations formed by strings in C_{all} adjacent to a vertex of $\mathcal{G}$. The main problem is dealing with the spanned configurations. Since some of these strings are quite long, it is impossible to directly use the girth restriction to show that the configurations cannot be spanned. To solve this problem, we restrict our attention to prefixes of these strings that are short enough so that the girth argument is applicable. The following lemma claims that we can replace each string in C_{all} by a single vertex of type $\mathcal{P}_2$ when considering the reducibility of a subgraph of H' . We recommend that the reader recalls Lemma 12 and the related definitions at this point.

LEMMA 25. *Let G be a closed subgraph of H' and let s be a path string $v_1v_2 \dots v_k$ in G belonging to C_{all} such that the important vertex adjacent to v_1 distinct from v_2 also belongs to G , but (in case s is a nonleaf string) the important vertex adjacent to v_k distinct from v_{k-1} does not belong to G . Let K' be a configuration obtained from G by removing the vertices $v_2, \dots, v_k$ and the 2-vertices of the proper links incident with them and letting the type of each important vertex v of K' be $\mathcal{U}_v$. We create a configuration K from K' by changing the type of v_1 to $\mathcal{P}_2$. If K is reducible, then G is reducible as well.*

Proof. We proceed similarly as in the proof of Lemma 12. We consider G as a subgraph of H . Depending on s , we select a suitable number m ($2 \leq m \leq k+1$) as described below. Let W be the set consisting of $V(K)$, the vertices $v_2, v_3, \dots, v_{m-1}$, and the 2-vertices of the proper links incident to them, and $V' = V(H) \setminus W$. By the minimality of H , there exists a coloring c'_1 of the subgraph of H induced by V' . As we demonstrate below, it is possible to extend the coloring c'_1 to vertices $v_2, \dots, v_{m-1}$ in such a way that there are at least two free color-pairs a and b at v_1 , thus obtaining a partial coloring c_1 . We define the lists $L_v \in \mathcal{U}_v$ for each important vertex v in the same way as we did in the proof of Lemma 12. We now create an instance Q of K by setting $Q(v) = L_v$ for each important vertex v of K except for v_1 , and by setting $Q(v_1) = \{a, b\}$ and $Q(v) = \mathcal{P}$ for each 2-vertex of K . Since K is reducible, its instance Q has a coloring c_2 . As in the proof of Lemma 12, we observe that the union of the colorings c_1 and c_2 can be extended to a proper $(5, 2)$ -coloring of H . This will show the reducibility of G . It remains to show how to determine m and how to extend c'_1 to c_1 . We discuss several cases that cover all strings in C_{all} .

If $k = 1$ or the class of v_1 is $\mathcal{D}$, we set $m = 2$ and $c_1 = c'_1$. In this case, the class of v_1 is $(\mathcal{A}\mathcal{B}\mathcal{D})$, and $\mathcal{U}_v \geq \mathcal{P}_2$. This implies that for any coloring c_1 , the list L_v contains at least two color-pairs a and b as required.

If the class of v_1 is $\mathcal{E}$ and the class of v_2 is $(\mathcal{A}\mathcal{D})$, then we set $m = 3$. Similarly to the previous case, with respect to the coloring c'_1 , L_{v_2} contains at least two color-pairs a_1 and a_2 , and $L_{v_1} \in \mathcal{N}_7$. Suppose that $L_{v_1} = \text{int}(p)$. At least one of a_1 or a_2 (say a_1)

is distinct from p . Thus, a_1 forbids at most five color-pairs in L_{v_1} . We set $c_1(v_2) = a_1$ and $\{a, b\} \subseteq \text{int}(p) \setminus \text{int}(a_1)$.

If the class of v_1 and v_2 is $\mathcal{E}$ and the class of v_3 is $\mathcal{D}$, then we set $m = 4$. Similarly as in the previous case, there are at least two free colors at v_3 and we may choose a color for v_3 such that it keeps two free colors at v_2 . Repeating the same argument, we color v_2 in such a way that there are at least two free color-pairs at v_1 .

The remaining case is that the class of v_1 is $\mathcal{F}$ and the class of v_2 is $\mathcal{D}$. In this case we set $m = 3$. There are at least two free colors a_1 and a_2 at v_2 . Let L be a subset of L_{v_1} (where L_{v_1} is defined with respect to c'_1) that contains exactly eight color-pairs. It cannot be the case that both $\text{int}(a_1)$ and $\text{int}(a_2)$ are subsets of L , since $|\text{int}(a_1) \cup \text{int}(a_2)| \geq 9$. We may assume that $\text{int}(a_1) \not\subseteq L$, and hence $|L \setminus \text{int}(a_1)| \geq 2$. We color v_2 with the color-pair a_1 , which leaves at least two free color-pairs at v_1 . $\square$

A string in $C_{\frac{1}{3}} \cup C_{\frac{1}{2}}$ contains one of the following substrings adjacent to v : $\text{path}(\mathcal{E}\mathcal{D}^2)$, $\text{path}(\mathcal{E}\mathcal{D}\mathcal{B})$, $\text{path}(\mathcal{D}\mathcal{E}\mathcal{B})$, $\text{path}(\mathcal{D}\mathcal{E}\mathcal{D})$, $\text{path}(\mathcal{D}^2)$, $\text{path}(\mathcal{D}\mathcal{C})$, $\text{path}(\mathcal{E}\mathcal{A})$, or $\text{path}(\mathcal{B})$. Let $C_{\#}$ be the set consisting of these strings. We show that each of these strings forbids only a few color-pairs. Note that unlike in Lemma 25, here we must assume that there are no links from such a string to the rest of the configuration.

LEMMA 26. *Suppose that $s = v_1 \dots v_k \in C_{\#}$, and let us consider a closed configuration that includes s and a vertex v of class $\mathcal{G}$ adjacent to s , but contains neither any of the proper links incident to s nor (in case s is not a leaf string) the important vertex adjacent to v_k distinct from v_{k-1} . The string s forbids at most two color-pairs at v , with the exception of $\text{path}(\mathcal{B})$ which forbids at most three color-pairs.*

Proof. Let us consider the elements of $C_{\#}$ separately.

If $s = \text{path}(\mathcal{E}\mathcal{D}^2)$ or $s = \text{path}(\mathcal{E}\mathcal{D}\mathcal{B})$, then the type of v_1 is $\mathcal{N}_7$, the type of v_2 is $\mathcal{P}_9$, and the type of v_3 is $\geq \mathcal{P}_2$, by Lemma 12. By Lemma 7(3), the configuration $\text{path}(\mathcal{P}_3\mathcal{P}_7\mathcal{P}_9\mathcal{P}_2)$ is reducible, and thus the string s forbids at most two color-pairs at v .

Similarly, if $s = \text{path}(\mathcal{D}\mathcal{E}\mathcal{D})$ or $s = \text{path}(\mathcal{D}\mathcal{E}\mathcal{B})$, the type of v_1 is $\mathcal{P}_9$, the type of v_2 is $\mathcal{P}_7$, and the type of v_3 is $\geq \mathcal{P}_2$. Again by Lemma 7(3), the configuration $\text{path}(\mathcal{P}_3\mathcal{P}_9\mathcal{P}_7\mathcal{P}_2)$ is reducible and s forbids at most two color-pairs at v .

If $s = \text{path}(\mathcal{D}^2)$ or $s = \text{path}(\mathcal{D}\mathcal{C})$, then the type of v_1 is $\mathcal{P}_9$ and the type of v_2 is $\geq \mathcal{P}_2$. Since the configuration $\text{path}(\mathcal{P}_3\mathcal{P}_9\mathcal{P}_2)$ is reducible by Lemma 7(3), s forbids at most two color-pairs at v .

If $s = \text{path}(\mathcal{E}\mathcal{A})$, then the type of v_1 is $\mathcal{N}_7$ and the type of v_2 is $\mathcal{P}_7$. The vertex v_2 forbids at most one color-pair at v_1 , and v_1 then forbids at most one color-pair at v , by Lemma 6(2).

Finally, if $s = \text{path}(\mathcal{B})$, then the type of v_1 is $\mathcal{N}_4$, and it forbids at most three color-pairs at v by Lemma 6(4). $\square$

The strings in $C_{\frac{2}{3}} \cup C_{\frac{1}{2}}$ forbid even fewer color-pairs.

LEMMA 27. *Suppose that $s = v_1 \dots v_k \in C_{\frac{2}{3}} \cup C_{\frac{1}{2}}$, and let K be a configuration that includes s and a vertex v of class $\mathcal{G}$ adjacent to s , but contains neither any of the proper links incident to s nor (in case s is not a leaf string) the important vertex adjacent to v_k distinct from v_{k-1} . The string s forbids at most one color-pair at v .*

Proof. Let us consider each possible string separately.

If $s = \text{path}(\mathcal{A})$, then the type of v_1 is $\mathcal{P}_7$. By Lemma 6(2), the configuration $\text{path}(\mathcal{P}_2\mathcal{P}_7)$ is reducible, and thus s forbids at most one color-pair at v .

If $s = \text{path}(\mathcal{D}\mathcal{B})$, then the type of v_1 is $\mathcal{P}_9$ and the type of v_2 is $\mathcal{N}_4$. By Lemma 7(3), the configuration $\text{path}(\mathcal{P}_2\mathcal{P}_9\mathcal{N}_4)$ is reducible, thus showing that s forbids at most one color-pair at v .

Finally, if $s = \text{path}(\mathcal{D}^3)$, then the type of v_1 and v_2 is $\mathcal{P}_9$ and the type of v_3 is $\mathcal{P}_2$. By Lemma 8, the configuration $\text{path}(\mathcal{P}_2 \mathcal{P}_9 \mathcal{P}_9 \mathcal{P}_2)$ is reducible, thus also in this case, s forbids at most one color-pair at v . $\square$

We are now ready to show that the final charge of vertices of class $\mathcal{G}$ is nonnegative.

THEOREM 28. *The charge of each vertex of classes $(\mathcal{G})$ after the second phase of discharging is nonnegative.*

Proof. Let v be a vertex of class $\mathcal{G}$. The charge of v before the second phase is 1. We prove that if the charge of the vertex v would become negative after applying the rules (R5)–(R7), then the configuration in the neighborhood of v is reducible. If v is adjacent to at most one maximal string with a negative charge, then v sends at most one unit of charge by Lemma 24, and hence the final charge of v is nonnegative. The vertex v cannot be joined to a single string with negative charge through more than one edge; otherwise the string together with v forms a cycle of length less than 9. Let s_1 , s_2 , and s_3 be the strings adjacent to v . There are the following cases in that the charge of v after the second phase is negative (up to symmetry):

- s_1 belongs to $C_{\frac{2}{3}}$ and s_2 belong to $C_{\frac{2}{3}} \cup C_{\frac{1}{2}}$. The configuration formed by s_1 , v , and s_2 cannot be spanned by the girth constraints. By Lemma 27, each of s_1 and s_2 forbids at most one color-pair at v . The precoloring of s_3 forbids at most seven color-pairs at v . Since there is at least one free color-pair at v , the configuration is reducible.
- $s_1 \in C_{\frac{2}{3}}$, $s_2 \in C_{\frac{1}{3}}$, and $s_3 \in C_{\frac{1}{3}} \cup C_{\frac{1}{6}}$. Consider the configuration formed by v , s_1 , the prefix substring of s_2 contained in $C_{\#}$, and a single vertex of type $\mathcal{P}_2$ that replaces s_3 using Lemma 25. This configuration cannot be spanned. By Lemmas 27, 26, and 6(2), the strings forbid at most 1, 3, and 5 color-pairs at v , respectively. Therefore, there is a free color-pair at v and the configuration is reducible.
- $s_1, s_2 \in C_{\frac{1}{2}}$, and $s_3 \in C_{\frac{1}{6}} \cup C_{\frac{1}{3}} \cup C_{\frac{1}{2}}$. We consider the configuration formed by v , the substrings $\text{path}(\mathcal{D}^2) \in C_{\#}$ of s_1 and s_2 , and a vertex of type $\mathcal{P}_2$ that replaces the string s_3 using Lemma 25. This configuration cannot be spanned by the girth constraints. By Lemmas 26 and 6(2), the strings forbid at most two, two, and five color-pairs at v , respectively, and hence the configuration is reducible.
- $s_1 \in C_{\frac{1}{2}}$ and $s_2, s_3 \in C_{\frac{1}{3}}$. We consider the configuration formed by v , $s_1 = \text{path}(\mathcal{D}^3)$, the prefix substring s'_2 of s_2 that belongs to $C_{\#}$, and a vertex of type $\mathcal{P}_2$ that replaces s_3 . Consider first the case that the configuration is not spanned. By Lemmas 27, 26, and 6(2), the numbers of color-pairs forbidden by the strings at v are at most 1, 3, and 5, respectively, and hence the configuration is reducible.

Let us now consider the case that the configuration is spanned. By the girth restrictions, this can happen only if there is a 2-link from the last vertex of s_1 to s'_2 . The girth restrictions also exclude the case $s'_2 = \text{path}(\mathcal{B})$, and thus s'_2 forbids at most two color-pairs, by Lemma 26. We shorten s_1 to its substring $\text{path}(\mathcal{D}^2) \in C_{\#}$ in the configuration. The new configuration cannot be spanned, and by Lemmas 26 and 6(2), the numbers of forbidden color-pairs at v are at most two, two, and five, respectively. Therefore, this smaller configuration is reducible.

This shows that the charge of v is nonnegative after the second phase of discharging. $\square$

This is the last piece needed in order to finish the proof of our main result.

Proof of Theorem 3. We proceed by contradiction. We assume that H is a minimal counterexample to this theorem. We assign a charge to the vertices and faces of H as described in section 4 such that the total amount of charge is negative. We run the first phase of discharging over H as described in section 5, arguing that the charge of all faces and 2-vertices is nonnegative after this phase. We construct the graph H' described in section 6 such that the sum of the charges of vertices of H' is negative. Finally, we run the second phase of discharging as described in section 8. The final charge of each vertex of class $(\mathcal{GH})$ and of each string is nonnegative, which is a contradiction. $\square$

9. Conclusion. An immediate consequence of Theorem 1 is the following.

COROLLARY 29. *Every planar graph of girth at least 8 is $(5, 2)$ -colorable.*

From R. Naseraser we learned that no planar graph of odd-girth 7 is known that does not map to the Petersen graph. This motivates the following question.

PROBLEM 1. *Is there a planar graph of odd-girth 7 with fractional chromatic number greater than $\frac{5}{2}$? Or is there at least one that does not map to the Petersen graph?*

REFERENCES

- [1] O. BORODIN, S. KIM, A. KOSTOCHKA, AND D. WEST, *Homomorphisms from sparse graphs with large girth*, J. Combin. Theory Ser. B, 90 (2004), pp. 147–159.
- [2] F. JAEGER, *Nowhere-zero flow problems*, in Selected Topics in Graph Theory 3, L. W. Beineke and R. J. Wilson, eds., Academic Press, London, 1988, pp. 71–95.
- [3] F. JAEGER, *On circular flows in graphs*, in Finite and Infinite Sets, Colloq. Math. Soc. János Bolyai 37, North-Holland, Amsterdam, 1984, pp. 391–402.
- [4] W. KLOSTERMEYER AND C. Q. ZHANG, $(2 + \epsilon)$ -coloring of planar graphs with large odd-girth, J. Graph Theory, 33 (2000), pp. 109–119.
- [5] A. PIRNAZAR AND D. H. ULLMAN, *Girth and fractional chromatic number of planar graphs*, J. Graph Theory, 39 (2002), pp. 201–217.
- [6] E. SCHEINERMAN AND D. ULLMAN, *Fractional Graph Theory*, Wiley-Interscience, New York, 1997, pp. 200–211.
- [7] W. T. TUTTE, *A contribution to the theory of chromatic polynomials*, Canad. J. Math., 6 (1954), pp. 80–91.
- [8] W. T. TUTTE, *On the imbeddings of linear graphs in surfaces*, Proc. London Math. Soc., 51 (1954), pp. 474–483.
- [9] C. Q. ZHANG, *Integer Flows and Cycle Covers of Graphs*, Marcel Dekker, New York, 1997.
- [10] X. ZHU, *Circular chromatic number: A survey*, Discrete Math., 229 (2001), pp. 371–410.
- [11] X. ZHU, *Circular chromatic number of planar graphs of large odd girth*, Electronic J. Combin., 8 (2001), research paper 25 (electronic).

HAT GUESSING GAMES*

STEVE BUTLER[†], MOHAMMAD T. HAJIAGHAYI[‡], ROBERT D. KLEINBERG[§], AND
TOM LEIGHTON[¶]

Abstract. Hat problems have become a popular topic in recreational mathematics. In a typical hat problem, each of n players tries to guess the color of the hat he or she is wearing by looking at the colors of the hats worn by some of the other players. In this paper we consider several variants of the problem, united by the common theme that the guessing strategies are required to be deterministic and the objective is to maximize the number of correct answers in the worst case. We also summarize what is currently known about the worst-case analysis of deterministic hat guessing problems with a finite number of players.

Key words. hat game, deterministic strategies, sight graph, Tutte–Berge formula, hypercube

AMS subject classifications. 91A12, 05C20

DOI. 10.1137/060652774

1. Introduction. Consider the following game. There are n distinguishable players and one adversary. The adversary will place on the heads of the players hats of k different colors, at which point players are allowed to see all hats but their own. No communication is allowed. Each player then makes a private guess as to what color of hat he or she is wearing. The goal of the players is to maximize the number of correct guesses.

To help players maximize their correct guesses, the players are allowed to meet before the hats are placed on their heads and to determine a public deterministic strategy (public in that everyone, including the adversary, knows the strategy, and deterministic in that the guesses are determined completely by the hat placement). What is the maximum number of correct guesses that can be guaranteed, and what strategy should be implemented to achieve the maximum?

These questions have been answered in [3, 9]. Here we will present the answers and consider some variations on the game, showing that they lead to some surprisingly subtle combinatorial and algorithmic problems and theorems. We proceed as follows. In the remainder of the introduction we answer these questions for the game as stated. In section 2 we consider what happens when players are not able to see everyone, which might be applicable as an abstraction of the problem of guessing information globally in a market or a decentralized computing system, in which each person (or software

*Received by the editors February 22, 2006; accepted for publication (in revised form) September 23, 2007; published electronically March 21, 2008.

<http://www.siam.org/journals/sidma/22-2/65277.html>

[†]Department of Mathematics, University of California, San Diego, La Jolla, CA 92093 (sbutler@math.ucsd.edu).

[‡]Department of Mathematics and Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar Street, Cambridge, MA 02139 (hajiagha@theory.csail.mit.edu).

[§]Computer Science Division, University of California, Berkeley, Berkeley, CA 94720, Department of Computer Science, Cornell University, Ithaca, NY 14853, and Department of Mathematics and Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar Street, Cambridge, MA 02139 (rdk@theory.csail.mit.edu). Part of this author's work was supported by a National Science Foundation Mathematical Sciences Postdoctoral Research Fellowship.

[¶]Akamai Technologies, Eight Cambridge Center, Cambridge, MA 02139 and Department of Mathematics and Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar Street, Cambridge, MA 02139 (ftl@theory.csail.mit.edu).

agent) has only partial/local knowledge. In section 3 we give a hypercube interpretation of the game which gives insight into the nature of optimal strategies, and we explore a version of the hats game where the adversary has a restricted hat supply.

Hat guessing games have long been a popular source of problems in recreational mathematics, and variations of hat guessing games have recently attracted increasing attention [6], partly because of their connections with coding theory (particularly Hamming codes). As another example, Aggarwal et al. [1] have used hat problems in the design of deterministic auction mechanisms. When constructing truthful auction mechanisms, a mechanism designer must devise a procedure for assigning a price to each bidder based only on the bids of other players (or else she may have an incentive to lie about her bid), with the goal of charging many bidders a price which is close to their own bidders. Note the formal similarity to hat problems, in which the goal is to devise a procedure for assigning a guess to each player based only on the hat colors of other players, with the goal of assigning to many players a guess which matches their own hat colors. By exploiting this similarity between the two problems, Aggarwal et al. used hat guessing strategies for a variant of the balanced hat problem (discussed in section 3.1) to provide a generic procedure for converting a randomized auction into a deterministic auction with approximately the same revenue, in markets with single-parameter bidders and no supply constraints.

Finally, it is worth mentioning that the fact that we are considering deterministic strategies instead of randomized ones is very important in this paper. The main reason is that these deterministic strategies focus on the worst-case scenarios instead of average or even almost-all scenarios. As a result, for several problems in this paper obtaining a randomized algorithm which guesses a constant fraction of the desired hat colors on average is easy, though we cannot even guess one hat color or a constant number of hat colors deterministically (see sections 2.3 and 2.4).

1.1. A winning approach to the hat guessing game.

Example 1. Consider the case where there are 2 players and 2 colors of hats. Then a winning strategy for these players is to have the first player guess what the second player is wearing and the second player to guess the color opposite of what the first player is wearing. If they are wearing hats of the same color, then the first player guesses correctly. If they are wearing different colors, then the second player guesses correctly. In any case there is one correct guess (and one incorrect guess).

It is interesting to note that in this example the expected number of correct guesses is one, which is the same as if they had guessed randomly. What their strategy has done is eliminate the variance involved in the guessing. The other thing to note is that neither player has any idea who guessed correctly, but they do know collectively that one of them did. These two properties will hold in general.

We have the following general result, first proved for 2 colors by Winkler [9] and later generalized to k colors by Feige [3].

THEOREM 2. *If there are n players and hats of k different colors, then there exists a strategy guaranteeing at least $\lfloor n/k \rfloor$ correct guesses. No strategy can improve on this.*

Proof. We first demonstrate a strategy. Number the players 1 to n and the colors of the hats 1 to k . The i th player will guess as if the sum of all the hats (including his/her own) is congruent to $i \bmod k$. At least $\lfloor n/k \rfloor$ of the players will be acting correctly and will therefore guess correctly.

To see that this cannot be improved upon we use an averaging argument. If a player sees a particular placement of hats, then this player is in one of k situations

and will guess correctly in exactly one of these situations. Since there are k^{n-1} ways to place the hats on the remaining players, we see that each player will make k^{n-1} correct guesses over all possible placements of hats. Since there are n players and k^n ways to place the hats, then on average we have $nk^{n-1}/k^n = n/k$ correct guesses. It follows that the adversary can find some placement of hats with at most $\lfloor n/k \rfloor$ correct guesses. $\square$

2. Restricting our vision in the game. In the original version of the game every player can see every other player. In an actual implementation of the game with a large group of people this might be difficult to achieve. So we now consider a variation where each player sees some subset of the other players.

To do this we introduce another layer to the game. We consider the “sight graph” where the vertices are the players and we have a directed edge from $a \rightarrow b$ if player a can see player b . As an example, in the original version of the game the graph was the complete graph on n vertices. For a given sight graph G we will let $H(G)$ denote the maximum number of correct guesses that the players can guarantee using an optimal strategy when there are 2 colors of hats.

In this section we will first consider the undirected case, i.e., the case in which every directed edge (u, v) is accompanied by the reverse edge (v, u) . For this case, an exact answer to $H(G)$ is known. In the directed case no exact answer for $H(G)$ is known but simple lower and upper bounds do exist. Finally, we consider the case when there are more than 2 colors of hats, for which little is known.

2.1. The undirected case. When we have an undirected graph the obvious strategy is to have players pair up as best as possible. Then in each pair we can implement the strategy in Example 1. This shows that we can guarantee at least $|M|$ correct guesses, where M is a maximum matching of G . The next result shows that this cannot be improved upon.

THEOREM 3. *Let G be an undirected graph with M a maximum matching of G . Then $H(G) = |M|$.*

Proof. It remains to show that $H(G) \leq |M|$. To do this we use the Tutte–Berge formula [2, 8], which says that there is a subset U of the vertices such that

$$|M| = \frac{|V| + |U| - o(G - U)}{2},$$

where $o(G - U)$ is the number of connected components of the induced subgraph $G - U$ which have an odd number of vertices. For $j = o(G - U)$ let $W_1, \dots, W_j$ be the connected components of $G - U$, which have an odd number of vertices, and let Y be the union of all the connected components of $G - U$ which have an even number of vertices.

Given any strategy we place hats as follows. First, place hats on U arbitrarily. Having fixed the hat placement on U , the guess of each player in W_i is now completely determined by the hat placement on W_i (since the only other players that can be seen are U which have already been placed). Applying the arguments from Theorem 2 we see that there is some placement of hats on each W_i with at most $(|W_i| - 1)/2$ correct guesses. Similarly we can place hats on Y so that there are at most $|Y|/2$ correct guesses. Therefore the total number of correct guesses is bounded above by

$$|U| + \frac{|Y|}{2} + \frac{|W_1| - 1}{2} + \dots + \frac{|W_j| - 1}{2} = \frac{|V| + |U| - j}{2} = |M|.$$

This concludes the proof. $\square$

2.2. The directed case. For the directed case there is no obvious strategy to adopt, and no sharp bound for $H(G)$ is known. However, there exist simple upper and lower bounds, as shown in the following.

LEMMA 4. *Given a directed graph G let $c(G)$ denote the maximal number of vertex disjoint cycles in G , and let $F(G)$ denote the minimum number of vertices whose removal from G makes the graph acyclic. Then $c(G) \leq H(G) \leq F(G)$.*

Proof. The lower bound follows by noting that for every cycle we can guarantee one correct guess. For example, if we have a cycle $a_1 \rightarrow a_2 \rightarrow \cdots \rightarrow a_k \rightarrow a_1$, then by having players $a_1, \dots, a_{k-1}$ guess the opposite color of the next player and a_k guess the color of the hat that a_1 is wearing, we guarantee at least one correct guess.

For the upper bound we note that we can arrange the vertices in order so that the removal of $v_1, \dots, v_{F(G)}$ leaves the graph acyclic and the remaining vertices $v_{F(G)+1}, \dots, v_n$ are such that if $i > F(G)$ and there is an edge from v_i to v_j , then $j < i$. In other words for the last $n - F(G)$ vertices, all outgoing edges point to the left. We place hats on the first $F(G)$ players arbitrarily and then we can place hats on players $F(G) + 1$ to n in turn, choosing each of the last $n - F(G)$ hat colors so as to force the corresponding player to guess incorrectly, given the colors of the preceding players. $\square$

By a theorem of Reed et al. [4], formerly known as Younger's conjecture (namely, for every integer $k \geq 0$ there exists an integer $t \geq 0$ such that every digraph G has k vertex disjoint directed cycles, or G can be made acyclic by deleting at most t vertices), this implies a criterion for determining whether a family of directed graphs has an unbounded "hat number."

COROLLARY 5. *Let $\mathcal{G}$ be a set of finite directed graphs. The set $\{H(G) : G \in \mathcal{G}\}$ is unbounded if and only if the set $\{F(G) : G \in \mathcal{G}\}$ is unbounded.*

Neither bound in Lemma 4 is sharp. For the upper bound, the undirected triangle has $F(G) = 2$, but we know from Theorem 2 that $H(G) = 1$. An example showing that the lower bound is not sharp is a little more involved and is given below.

Example 6. Let G consist of a directed four cycle $a \rightarrow b \rightarrow c \rightarrow d \rightarrow a$ with a fifth node e joined to the other four by bidirected edges. This graph, shown in Figure 1, has $c(G) = 1$ and $H(G) = 2$.

To describe a strategy we will let A, B, C, D, E denote the actual colors of hats placed on players a, b, c, d, e respectively, while g_a, g_b, g_c, g_d, g_e denote their guesses. We can describe their strategy in mod 2 arithmetic as follows:

$$g_a = B + E; \quad g_b = C + E; \quad g_c = D + E; \quad g_d = A + E + 1;$$

$$g_e = \begin{cases} 1 & \text{if } (A + B, B + C, C + D, A + D + 1) \text{ has Hamming weight 1;} \\ 0 & \text{if } (A + B, B + C, C + D, A + D + 1) \text{ has Hamming weight 3.} \end{cases}$$

What happens is that the players a, b, c, d will make either one or three correct guesses, depending on what e is wearing. So e guesses as though his/her hat would force one correct guess among the other four players. Thus, either e guesses incorrectly and

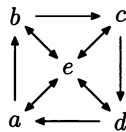


FIG. 1.

there are three correct guesses among a, b, c, d ; or e guesses correctly and there is one correct guess among a, b, c, d for a total of two correct guesses.

Question. For an undirected graph G how do we calculate $H(G)$? The obvious algorithm for deciding if $H(G) \geq h$ requires nondeterministic exponential time nondeterministically comes up with a guessing strategy for the players, and then spends exponential time verifying that this strategy produces at least h correct answers on every input. We do not know if there is a more efficient algorithm for deciding if $H(G) \geq h$.

We note that in answering this question for directed graphs it suffices to consider graphs G which are strongly connected. In particular, $H(G) = \sum_k H(G_k)$ where G_k are the strongly connected components of G . The proof of this is similar to the argument of the upper bound in Lemma 4. Namely, we can order the strongly connected components so that there is an edge from G_i to G_j only if $i > j$. Then the adversary acts optimally on each of the connected components in turn.

A related question concerns the complexity of optimal guessing strategies. Define a guessing strategy with sight graph G to be *optimal* if it achieves at least $H(G)$ correct answers on every input. We saw in the proof of Theorem 3 that when G is undirected, there is always an optimal guessing strategy in which each player's guess is computed by evaluating a linear function (over the field with two elements) whose inputs are the other players' hat colors together with the constant 1. For directed graphs this is not the case. The sight graph in Example 6 has $H(G) = 2$, but the reader may verify by a simple case analysis that for every linear guessing strategy, there exists an input on which fewer than 2 players answer correctly.

Question. Is it true that for every directed graph G , there is an optimal guessing strategy in which every player's guess is computed by inserting the other players' hat colors into a Boolean circuit of size polynomial in $|G|$?

Note that an affirmative answer to this question would imply that the problem of deciding if $H(G) \geq h$ belongs to the complexity class Σ_2^P , providing a partial answer to the preceding question.

2.3. More than 2 colors of hats. Considerably less is known when there are more than 2 colors involved in the game. Let $H_k(G)$ denote the maximum number of correct guesses that the players can guarantee using an optimal strategy when there are k colors of hats. From the proof of Theorem 2, we know that $H_k(G) = 1$ when G is an undirected k -clique, and therefore $H_k(G) \geq \ell$ whenever G contains ℓ disjoint undirected k -cliques. However, it is possible to avoid k -cliques altogether and still guarantee at least one correct guess, as shown below.

THEOREM 7. *For every number k , there exists a bipartite graph G with $H_k(G) > 0$.*

Proof. Let G be a complete bipartite graph with $n = k - 1$ vertices on the left side and $m = k^{k^n}$ vertices on the right side. Let C denote the set of all k -colorings of the left side of G . Note that $|C| = k^n$ and $m = k^{|C|}$; hence m is equal to the number of mappings from C to $\{1, 2, \dots, k\}$. Pick a one-to-one correspondence between the vertices on the right side of G and the mappings from C to $\{1, 2, \dots, k\}$, and let each vertex on the right side of G guess its color using the corresponding mapping.

We will need the following claim.

Claim. Let c_R denote a fixed coloring of the right side of G , and let C' denote the set of all colorings c_L of the left side of G such that the combined coloring (c_L, c_R) causes every vertex on the right side to guess its color incorrectly. Then $|C'| < k$.

Now it's time to define the guessing strategies used by the vertices on the left side of G . Given the coloring of the right side, the set C' defined in the lemma above has

at most $n = k - 1$ elements. So let $c_1, c_2, \dots, c_n$ be a list of colorings which contains every element of C' . For $i = 1, 2, \dots, n$, vertex i on the left guesses that its color is $c_i(i)$. This guessing strategy (combined with the guessing strategy for the vertices on the right side as defined above) guarantees at least one correct answer. This is because the above claim guarantees that at least one vertex on the right side guesses correctly unless the coloring of the left side belongs to C' . But if the coloring of the left side belongs to C' , then it is equal to c_i for some i in $1, 2, \dots, n$, in which case vertex i on the left guesses its color correctly.

It remains to prove the claim. The proof follows from noting that if C' contains k distinct elements $c_1, c_2, \dots, c_k$, then there exists a function f from C to $\{1, 2, \dots, k\}$ which assumes k distinct values on the set $\{c_1, \dots, c_k\}$. Let v denote the vertex on the right side of G corresponding to f . Since the set $\{f(c_1), f(c_2), \dots, f(c_k)\}$ contains all k colors, we must have $f(c_i) = c_R(v)$ for some i in $1, 2, \dots, k$. Thus, the combined coloring (c_i, c_R) causes vertex v to guess its color correctly, contradicting our assumption that c_i belongs to C' , ending the proof. $\square$

Question. Is there a bipartite graph G satisfying $H_k(G) > 0$ whose size is polynomial in k ? What if, instead of bipartite, we consider k -clique-free graphs?

Question. Call an undirected sight graph G “edge-critical for the hats game with k colors” if G has the property that there exists a guessing strategy which guarantees at least one correct answer for the hats game with k colors, but no proper subgraph of G has this property. For $k = 2$, the only edge-critical graph is a 2-clique. For $k > 2$, there are at least two (undirected) edge-critical graphs, namely a k -clique and a subgraph of the complete bipartite $(k - 1)$ -by- (k^{k-1}) graph. For $k > 2$, are there infinitely many graphs which are edge-critical for the hats game with k colors?

We close this section by proving that $H_k(G) = 0$ whenever $k > 2$ and G is an undirected tree. This fact is a consequence of the following more general lemma.

LEMMA 8. *Suppose we are given an undirected tree T ; a guessing strategy Γ for the hat k -coloring problem on T ; a node v in T ; and a pair of colors c_1, c_2 . Then there exists a k -coloring ($k \geq 3$) of T such that every node guesses its color incorrectly, and the color of node v is either c_1 or c_2 .*

Proof. The proof is by induction on the size of T . When $|V(T)| = 1$ the result is trivial. Otherwise deleting v from T partitions the remaining vertices into a collection of disjoint subtrees $T_1, T_2, \dots, T_j$. For $i = 1, 2, \dots, j$, let $r(T_i)$ denote the unique neighbor of v in T_i . Let $\Gamma_1(T_i)$ (respectively, $\Gamma_2(T_i)$) denote the guessing strategy applied in T_i when the color of v is c_1 (respectively, c_2). Note that $\Gamma_1(T_i)$ and $\Gamma_2(T_i)$ differ only in the function which $r(T_i)$ uses to guess its color based on the colors of its neighbors in T_i . Let $B_1(T_i)$ denote the set of “bad colorings” for guessing strategy $\Gamma_1(T_i)$, i.e., the colorings which cause every node of T_i to guess its color incorrectly. Let $C_1(T_i)$ denote the set of colors assigned to $r(T_i)$ by colorings in $B_1(T_i)$. Define sets $B_2(T_i), C_2(T_i)$ similarly, but use the guessing strategy $\Gamma_2(T_i)$ in place of $\Gamma_1(T_i)$. The induction hypothesis implies that $C_1(T_i)$ and $C_2(T_i)$ each have at least $k - 1$ elements. (If not, then the complement of one of these sets, say $C_1(T_i)$, contains at least 2 colors, say c_3, c_4 . Applying the induction hypothesis with tree T_i , guessing strategy $\Gamma_1(T_i)$, node $r(T_i)$, and color pair c_3, c_4 would lead to an element of $B_1(T_i)$ in which the color of $r(T_i)$ is either c_3 or c_4 , contradicting the assumption that c_3 and c_4 are both in the complement of $C_1(T_i)$.) Having established that $C_1(T_i)$ and $C_2(T_i)$ each have at least $k - 1$ elements, it follows (from the fact that $k > 2$) that the intersection of $C_1(T_i)$ and $C_2(T_i)$ is nonempty. Choose a color c_i from the intersection of these two sets and assign it to $r(T_i)$. Do this for each i in $\{1, 2, \dots, j\}$. Having

assigned a color to each neighbor of v , the guess of node v is now determined. At least one of the colors c_1 and c_2 differs from this guess, so we may assign this color to node v and thereby ensure that it guesses incorrectly. Assume without loss of generality that color c_1 is assigned to v . For each subtree T_i , the set $B_1(T_i)$ contains a coloring which satisfies that

- the color of $r(T_i)$ is c_i ;
- every node guesses its color incorrectly using guessing strategy $\Gamma_1(T_i)$.

We choose one such coloring and use it to assign colors to the nodes of T_i . Doing this for every i in $\{1, 2, \dots, j\}$ yields a coloring of T which satisfies the two properties in the statement of the lemma. $\square$

COROLLARY 9. *If G is an undirected tree and $k > 2$, then $H_k(G) = 0$.*

2.4. Generalized guessing graphs. In this section we consider a variation in which players are not necessarily trying to guess their own hat colors. Instead there is a set P (“players”), a set H (“hats”), and two directed graphs, G_v (“visibility graph”) and G_g (“guessing graph”). Both graphs have a vertex set which is the union of P and H . Every edge of G_v has its tail in P and its head in H ; we think of edge (u, v) as indicating that person u can see the color of hat v . Every edge of G_g has its tail in H and its head in P ; we think of edge (v, u) as indicating that person u must guess the color of hat v . (Note that the orientation of these edges is from hats to people, the reverse of the orientation convention used in G_v . This orientation convention is adopted because it will be convenient later on.) The hat problem considered in earlier sections corresponds to the case when there is a bijection $\phi : H \rightarrow P$ and the edge set of G_g is $\{(v, \phi(v)) : v \in H\}$.

A “guessing strategy” is a set of functions, one for each edge in G_g . Each such function maps the set of k -colorings of H to the set of colors and has the property that the value of the function corresponding to edge $e = (v, u)$ depends only on the colors of the elements of H which are adjacent to u in G_v . Given a k -coloring of H and a guessing strategy, we say that edge $e = (v, u)$ of G_g gives a correct answer if its function evaluates to the color which was assigned to v . We define $H_k(G_v, G_g)$ to be the maximum number of correct guesses that the players can guarantee using an optimal strategy when there are k colors of hats.

THEOREM 10. *When $k = 2$, $H_k(G_v, G_g) > 0$ if and only if at least one of the following properties holds:*

- There is a vertex of G_g whose out-degree is greater than 1.*
- There is a directed cycle in the union of G_v and G_g .*

Proof. Identify the set of colors with the set $\{0, 1\}$. If property (a) is satisfied and G_g contains edges (v, u) and (v, u') for some v in H and u, u' in P , then assign the constant function 0 to edge (v, u) and the constant function 1 to edge (v, u') . Clearly, on every input, at least one of these edges gives a correct answer.

If property (b) is satisfied, let the vertices of the cycle be

$$v_1 \rightarrow u_1 \rightarrow v_2 \rightarrow u_2 \rightarrow \dots \rightarrow v_n \rightarrow u_n$$

and adopt the following guessing strategy. For $i = 1, 2, \dots, n - 1$, player u_i guesses that the color of v_i is different from the color of v_{i+1} . Player u_n guesses that the color of v_n is the same as the color of v_1 . Observe that this is a legal guessing strategy since each of the edges $(u_1, v_2), (u_2, v_3), \dots, (u_n, v_1)$ belongs to G_v . Also, for any input on which none of $u_1, u_2, \dots, u_{n-1}$ guess correctly, it must be the case that $v_1, v_2, \dots, v_n$ are all assigned the same color. But then (u_n, v_1) guesses correctly.

Finally, suppose neither (a) nor (b) is satisfied; we must prove that for every guessing strategy there exists an input on which every edge guesses incorrectly. We will only sketch this part of the proof. Let G be the union of G_v and G_g , and let G' be the directed graph obtained from G by contracting the edges of G_g . If G' contains a directed cycle, then G must also contain a directed cycle. (In fact, our assumption that property (a) is not satisfied implies that every edge of G' corresponds to a 2-hop path between two elements of P in G .) Since we are assuming G contains no directed cycles, it follows that G' is acyclic. An elementary induction argument, using a topological sort of G' , produces a coloring of H which causes every edge to guess incorrectly. $\square$

Question. Above we characterize visibility and guessing graphs for which we can guarantee at least one correct answer. It would be nice if we could determine exactly how much more information in the guessing graph we can obtain by adding a particular edge to the visibility graph. More generally, given m , G_v , and G_g determine the smallest value j such that there exists a graph G'_v consisting of G_v with j additional edges such that the hat number $H_2(G'_v, G_g)$ is at least m .

The above question can be loosely considered along the same lines as the Aanderaa–Rosenberg conjecture [7] which asks for the minimum number of edges of a graph that should be revealed in order to determine whether the graph has a given monotone property P or not (see also [5]).

3. Using hypercubes to approach the game. One of the interesting connections between hat guessing games and applications lies in interpreting the game in terms of various structure restrictions on hypercubes (the restriction depending on the rule of the particular game). In this section we will consider several questions for which we can use hypercubes to give an answer. One drawback to using hypercubes is that the number of vertices is exponential in the number of players and so many of the constructions are not polynomial in n . Nevertheless, insight into the game can be achieved by considering hypercubes (see, for example, Proposition 13).

Recall that the n -cube has as vertices all 2^n binary words of length n . This has a natural connection with the possible placements of hats. The edges of the n -cube join two vertices which differ by one letter. As an example, in the 4-cube there is an edge between 1011 and 1010; we can represent this in shorthand as 101* where * indicates an indeterminate bit which is either 0 or 1. The edges of the n -cube represent the decisions which must be made in forming the strategy. So 101* indicates that the fourth player (the *) sees 1, 0, and 1 on the first, second, and third players, respectively. In this situation they must guess either 0 or 1. To indicate his/her guess we will “orient” the edge in the direction of the guess. So for instance, if the player guesses 0 in this case, then we will have 1011→1010.

The original version of the game reduces to finding an orientation on the edges of the n -cube to maximize the minimum in-degree.

3.1. Balanced strategies. We now return to the original game. In Theorem 2 we gave one example of how to construct a strategy to guarantee $\lfloor n/k \rfloor$ guesses. This strategy is far from unique and may not have some desired property. For example, while that strategy is easy to implement, the correct guesses are not reflective of the actual hats that are placed on the players. For 2 colors of hats we will show how to construct different strategies, and in particular how to construct a balanced strategy.

LEMMA 11. *If there are n players and 2 different hat colors, then there exists a strategy which is balanced. That is, if there are b blue hats and r red hats placed on*

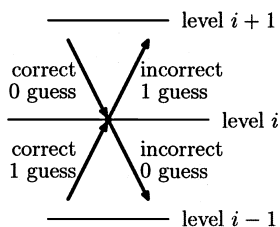


FIG. 2.

the players ($r + b = n$), then at least $\lfloor b/2 \rfloor$ of the people wearing blue guess correctly and $\lfloor r/2 \rfloor$ of the people wearing red guess correctly.

We first approach the problem by using hypercubes. To construct a balanced strategy we will group the vertices of the hypercube in i levels, where a vertex is at level i if the word indexing the vertex has Hamming weight i . The up-degree (respectively, down-degree) at a vertex in level i will be the number of edges between that vertex and vertices in level $i + 1$ (respectively, $i - 1$). If we consider how directing the edges corresponds to guesses, we have the picture shown in Figure 2.

We now see that the balanced strategy in Lemma 11 would correspond to an orientation on the edges of the hypercube so that at each node the number of directed edges from level $i + 1$ to that node is $\lfloor \text{up-degree}/2 \rfloor$ while the number of directed edges from level $i - 1$ to that node is $\lfloor \text{down-degree}/2 \rfloor$.

Proof. Construction of Lemma 11. For n even we start with any edge and orient it arbitrarily and then continue to lengthen the directed path to be as long as possible by continually directing an undirected edge, which is incident with the current terminal vertex. The only restriction is that if an edge is between level i and level $i + 1$, then if possible the next edge will also be between level i and level $i + 1$. When the path can no longer be extended, if we have not oriented all the edges, then we pick an unoriented edge and repeat the process.

When n is odd a similar construction works, with the only caveat being that we must be careful in selecting our initial edges. Now an initial edge cannot be chosen arbitrarily but must be directed up from a vertex of odd up-degree or down from a vertex of odd down-degree.

It is easy to see that the strategy we construct is balanced, since at each vertex we pair directed in- and out-edges in such a way that every edge coming in from below (respectively, above) is paired, if possible, with an edge going down (respectively, up). $\square$

Using a different technique based on network flow, it is possible to construct balanced strategies for every k ; see [1] for details.

THEOREM 12 (see [1]). *If there are n players and k different hat colors, then there exists a strategy which is balanced. That is, if a_i of the players are wearing hats of color i ($1 \leq i \leq k$), then at least $\lfloor a_i/k \rfloor$ of the people wearing color i guess correctly for each value of i .*

3.2. Optimal strategies are unbiased. If one constructs many optimal strategies for the 2-color game when n is even, one starts to see a pattern emerge. Namely, each player is as likely to guess one hat color as to guess the other. We give a short proof that uses hypercubes.

PROPOSITION 13. *Suppose the set of hat colors is $\{0, 1\}$ and there are n players playing an optimal strategy, where n is even. For any fixed player i , looking over all*

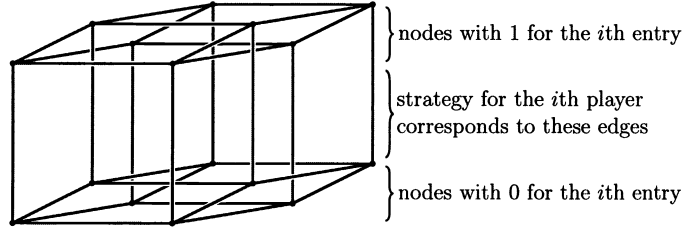


FIG. 3.

possible hat placements, we have that the number of times that player i guesses 0 is the same as the number of times that player i guesses 1.

Proof. When n is even an optimal strategy corresponds to an orientation on the edges of the n -cube with the in-degree equal to the out-degree at each vertex. In particular, the strategy corresponds to some Eulerian walk on the n -cube. We now redraw the n -cube as in Figure 3.

Then note that the number of edges in the center that point up is the number of times the i th player guesses 1, and the number of edges that point down is the number of times that the i th player guesses 0. Since we have an Eulerian walk the number of up-edges equals the number of down-edges. $\square$

Proposition 13 can be generalized to games with more than 2 colors.

PROPOSITION 14. *Suppose that n players are playing an optimal strategy of the k -color game, where k is a divisor of n . If the players' hat colors are drawn independently from the uniform distribution on $\{1, 2, \dots, k\}$, then for each player i and each hat color c ,*

$$\Pr(i \text{ guesses } c) = 1/k.$$

Equivalently, in an optimal strategy each player guesses each hat color an equal number of times.

Proof. Let X denote the random variable which counts the number of correct answers provided by the players. Our assumption that the players use an optimal guessing strategy means that $X \geq n/k$ at every point of the sample space. Now let $\mathcal{E}(i, c)$ denote the event that player i is assigned hat color c . We have

$$\begin{aligned} (1) \quad \mathbf{E}[X \mid \mathcal{E}(i, c)] &= \sum_{j=1}^n \Pr(j \text{ guesses correctly} \mid \mathcal{E}(i, c)) \\ (2) \quad &= \frac{n-1}{k} + \Pr(i \text{ guesses } c \mid \mathcal{E}(i, c)) \\ (3) \quad &= \frac{n-1}{k} + \Pr(i \text{ guesses } c). \end{aligned}$$

Here (1) follows from linearity of expectation, and (2) follows from the fact that, conditional on $\mathcal{E}(i, c)$, every player except i has a hat color which is uniformly distributed and is independent of its own guess. Finally, (3) follows from the fact that player i 's guess is independent of its own hat color.

Recalling now that $X \geq n/k$ at every point of the sample space, we see that

$$\mathbf{E}[X \mid \mathcal{E}(i, s)] \geq n/k,$$

and according to (3) this implies $\Pr(i \text{ guesses } c) \geq 1/k$. Since this inequality applies to every hat color c , it must be the case that $\Pr(i \text{ guesses } c) = 1/k$ for every color c . $\square$

3.3. The limited hats game. We now consider another variation on the original hats game. The setup is as before, but now the adversary has a limited supply of hats to choose from. We will let $H(n; a_1, a_2, \dots, a_k)$ denote the maximum number of correct guesses that we can guarantee when there are n players and a_1 hats of the first color, a_2 hats of the second color, and so on up through a_k hats of the k th color. We need $a_1 + a_2 + \dots + a_k \geq n$ (to ensure that we have enough hats for the players) and without loss of generality we can assume that $0 < a_i \leq n$ for all i .

Example 15. Suppose that there are three players and the adversary has two blue hats and two red hats. The players can choose to ignore this information and use the same strategy as in Theorem 2, guaranteeing one correct guess. However, if they modify their strategy, then they can guarantee 2 correct guesses. If the players are a, b, c , then such a strategy would be for a to guess the opposite of what b is wearing, b to guess the opposite of what c is wearing, and c to guess the opposite of what a is wearing. So $H(3; 2, 2) = 2$. (More generally, it was shown in [3] that $H(4k - 1; 2k, 2k) = 3k - 1$.)

THEOREM 16. *We have the following properties:*

(i)

$$H(n; \underbrace{n, n, \dots, n}_{k \text{ times}}) = \lfloor n/k \rfloor.$$

(ii) *If $a_1 + a_2 + \dots + a_k = n$, then $H(n; a_1, a_2, \dots, a_k) = n$.*

(iii) *If m is even or k is odd (or both), then*

$$H(mk - 1; \underbrace{m, \dots, m}_{k \text{ times}}) = \frac{mk + m - 2}{2}.$$

(iv) *$H(n; a_1, a_2, \dots, a_k) = H(n; a_{\sigma(1)}, a_{\sigma(2)}, \dots, a_{\sigma(k)})$ for any permutation σ .*

(v) *If $b_i \leq a_i$ for all $i = 1, 2, \dots, k$, then $H(n; a_1, a_2, \dots, a_k) \leq H(n; b_1, b_2, \dots, b_k)$.*

(vi)

$$H(n; a_1, a_2, \dots, a_k) \leq \left\lfloor \sum_{\substack{b_i \leq a_i, 1 \leq i \leq k \\ b_1 + \dots + b_k = n-1}} n \binom{n-1}{b_1, \dots, b_k} \right\rfloor / \sum_{\substack{b_i \leq a_i, 1 \leq i \leq k \\ b_1 + \dots + b_k = n}} \binom{n}{b_1, \dots, b_k}.$$

Proof. Item (i) is Theorem 2. Item (ii) is obvious because the strategy is to have each player guess the hat not seen. Item (iv) says we can permute the hat colors. Item (v) follows by noting that the optimal strategy for the $H(n; a_1, a_2, \dots, a_k)$ game is also a (not necessarily optimal) strategy for the $H(n; b_1, b_2, \dots, b_k)$ game.

To prove item (vi) we first give a hyper-hypercube interpretation of the game. The k^n vertices are words of length n from the alphabet $\{0, \dots, k-1\}$ and correspond to the k^n possible placements of hats. The tuples (i.e., edges) represent the decisions which must be made in deciding a strategy. So, for example, if $n = 5$ and $k = 3$, then one tuple would be $210*1 = \{21001, 21011, 21021\}$, representing the situation when the fourth player (the $*$) sees 2, 1, 0, and 1 on the first, second, third, and fifth players, respectively. A strategy corresponds to marking one vertex on each tuple,

the marking indicating the guess that the player will make. Note each marking is one correct guess.

We now use an averaging argument similar to that given in Theorem 2.

$$\begin{aligned} H(n; a_1, a_2, \dots, a_k) &\leq \left\lfloor \frac{\# \text{ of correct guesses}}{\# \text{ of possible hat placements}} \right\rfloor \\ &= \left\lfloor \frac{\# \text{ of tuples available for marking}}{\# \text{ of vertices to be marked}} \right\rfloor \\ &= \left\lfloor \frac{\sum_{\substack{b_i \leq a_i, 1 \leq i \leq k \\ b_1 + \dots + b_k = n-1}} n \binom{n-1}{b_1, \dots, b_k}}{\sum_{\substack{b_i \leq a_i, 1 \leq i \leq k \\ b_1 + \dots + b_k = n}} \binom{n}{b_1, \dots, b_k}} \right\rfloor \end{aligned}$$

The numerator is the n possible positions of the $*$ along with the allowable combinations of the remaining $n-1$ entries. The denominator is the number of ways to place the n hats in allowable combinations.

For item (iii) we have that $H(mk-1; m, \dots, m) \leq (mk+m-2)/2$ from (vi). So it suffices to show that we can construct a strategy guaranteeing at least $(mk+m-2)/2$ correct guesses. There are two types of tuples; those which involve only one markable vertex and those that involve two. The first kind is for players who see a full set of all but one type of hat and so automatically know their hats. The second kind is for players who see a full set of all but two types of hats and so have to make one of two choices.

Every vertex will be associated with $m-1$ tuples of the first type (one for each hat of the deficient color in the placement); we mark these tuples and put them aside. We now construct a bipartite graph with the remaining edges and tuples as follows: The vertex set is $S \cup T$, where S is the set of tuples we have not yet marked and T is the set of vertices to be marked; there is an edge between an element in S and an element in T if the corresponding vertex can be marked by the corresponding tuple.

Every element in S has degree 2 and every element in T has degree $mk-m$, which by assumption is even. We now split the elements in T by duplicating each element $(mk-m)/2$ times and then distributing the edges of the original element so that each resulting piece has degree 2. It is easy to now construct a perfect matching between S and T (for example, start with any edge and going through a cycle alternatively include/not include the edges). This perfect matching gives a marking on the tuples so that each of the vertices is marked $(mk-m)/2$ times.

In total each vertex was marked $(m-1) + (mk-m)/2 = (mk+m-2)/2$ times, giving our desired strategy. $\square$

Questions. What is $H(n; a_1, a_2, \dots, a_k)$? Is the upper bound given in Theorem 16 tight? (Note that items (i), (ii), and (iii) in Theorem 16 are examples where the bound is tight.)

As a warmup to the above question, the interested reader might enjoy showing that $H(5; 4, 3) = 3$ (an upper bound of 3 immediately follows from Theorem 16 so it suffices to find a strategy guaranteeing at least three correct guesses).

4. Summary and open questions. In this paper we considered hat guessing games in which players wearing hats of various colors use deterministic strategies to guess the color of some hats (usually their own) with the goal of maximizing the number of correct answers in the worst case. In this section we summarize our main results, and we list the open questions which are scattered throughout this paper.

Focusing on hat games with a *sight graph* that specifies the set of hats visible to each player, we defined the *hat number* $H_k(G)$ to be the number of correct answers

provided in the worst case by an optimal strategy for players guessing their own hat color, when the sight graph is G and there are k colors of hats. We proved that for two colors and undirected sight graphs, the hat number equals the cardinality of a maximum matching in the sight graph. For a directed graph G and two hat colors, we provided lower and upper bounds on the hat number; these bounds suffice to distinguish graph families with a bounded hat number from those with an unbounded hat number. For three or more colors, we proved that the hat number of a tree is zero and that there are bipartite graphs with a nonzero hat number. When there are two hat colors and there is a *guessing graph* which specifies the set of hat colors which each player must guess, we provided necessary and sufficient conditions for the existence of a strategy which guarantees at least one correct answer.

Turning to questions about the distribution of guesses and of correct guesses, we introduced a hypercube interpretation of the game which permitted us to prove that when there are two colors of hats and players can see all hats except their own, there is a guessing strategy which guarantees that roughly half of the players wearing each type of hat guess correctly. (This result was also proved in [1], which contains a generalization to more than two colors.) We also proved that optimal strategies are unbiased, i.e., when the number of hat colors is a divisor of the number of players, each player guesses each hat color with equal probability when the colors are assigned independently and uniformly at random. Finally, turning to a version in which the players are given an a priori upper bound on the number of hats of each color, we exhibited some bounds on the number of correct answers provided by the optimal strategy in the worst case.

In presenting these results, we also introduced many open questions inspired by them. Here we recapitulate the open questions presented earlier.

Complexity of computing hat numbers and optimal strategies. What is the computational complexity of determining $H_k(G)$, given k and G ? There is a polynomial-time algorithm when $k = 2$ and G is undirected (because a maximum matching can be computed in polynomial time), but for all other cases the best known algorithm requires nondeterministic exponential time. On the other hand, we do not know if the problem is NP-hard.

How much computational power is required to implement an optimal guessing strategy? For $k = 2$, can the optimal guessing strategy for a directed sight graph G always be implemented by players using Boolean circuits of size polynomial in the size of G ?

Graphs with positive hat numbers. Is there a bipartite graph G satisfying $H_k(G) > 0$ whose size is polynomial in k ? What if, instead of bipartite, we consider k -clique-free graphs? Are there infinitely many graphs G such that $H_k(G) > 0$ but $H_k(G') = 0$ for every proper subgraph $G' \subset G$?

Augmenting sight graphs to improve hat numbers. Given a positive number m and two graphs G_v, G_g (the sight graph and guessing graph of a hat guessing game), determine the smallest value j such that one can add j additional edges to G_v to obtain a graph G'_v satisfying $H_2(G'_v, G_g) \geq m$.

The limited hats game with unrestricted vision. Suppose players can see every hat color except their own and must guess their own hat color. When there are k hat colors and at most a_i hats in color class i ($1 \leq i \leq k$), compute the maximum number of correct answers that can be guaranteed by a guessing strategy given this limitation on the placement of hats. Is the upper bound in part (vi) of Theorem 16 tight?

Acknowledgments. We thank Joe Buhler, Kevin Costello, Harvey Friedman, and Ron Graham for interesting and helpful discussions on this subject.

REFERENCES

- [1] G. AGGARWAL, A. FIAT, A. V. GOLDBERG, J. D. HARTLINE, N. IMMORLICA, AND M. SUDAN, *Derandomization of auctions*, in Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC), 2005, pp. 619–625.
- [2] C. BERGE, *Sur le couplage maximum d'un graphe*, C. R. Acad. Sci. Paris, 247 (1958), pp. 258–259.
- [3] U. FEIGE, *You Can Leave Your Hat On (If You Guess Its Color)*, Technical report MCS04-03, Computer Science and Applied Mathematics, The Weizmann Institute of Science, Rehovot, Israel, 2004.
- [4] B. REED, N. ROBERTSON, P. SEYMOUR, AND R. THOMAS, *Packing directed circuits*, Combinatorica, 16 (1996), pp. 535–554.
- [5] R. L. RIVEST AND J. VUILLEMIN, *A generalization and proof of the Aanderaa–Rosenberg conjecture*, in Proceedings of the Seventh Annual ACM Symposium on Theory of Computing (STOC), 1975, pp. 6–11.
- [6] S. ROBINSON, *Why mathematicians now care about their hat color*, New York Times, April 10, 2001, D5.
- [7] A. L. ROSENBERG, *On the time required to recognize properties of graphs: A problem*, ACM SIGACT News, 5 (1973), pp. 15–16.
- [8] W. T. TUTTE, *The factorization of linear graphs*, J. London Math. Soc., 22 (1947), pp. 107–111.
- [9] P. WINKLER, *Games people don't play*, in Puzzler's Tribute: A Feast for the Mind, A. K. Peters, Baltimore, 2002, pp. 301–313.

THE WINDY GENERAL ROUTING POLYHEDRON: A GLOBAL VIEW OF MANY KNOWN ARC ROUTING POLYHEDRA*

ANGEL CORBERÁN[†], ISAAC PLANA[‡], AND JOSÉ M. SANCHIS[§]

Abstract. The windy postman problem consists of finding a minimum cost traversal of all of the edges of an undirected graph with two costs associated with each edge, representing the costs of traversing it in each direction. In this paper we deal with the windy general routing problem (WGRP), in which only a subset of edges must be traversed and a subset of vertices must be visited. This is also an *NP*-hard problem that generalizes many important arc routing problems (ARPs) and has some interesting real-life applications. Here we study the description of the WGRP polyhedron, for which some general properties and some large families of facet-inducing inequalities are presented. Moreover, since the WGRP contains many well-known routing problems as special cases, this paper also provides a global view of their associated polyhedra. Finally, for the first time, some polyhedral results for several ARPs defined on mixed graphs formulated by using two variables per edge are presented.

Key words. arc routing problems, windy general routing problem, polyhedra, facets

AMS subject classifications. 90C27, 90C57, 90B06

DOI. 10.1137/050640886

1. Introduction. Arc routing problems (ARPs) have their origin in the celebrated Königsberg bridge problem solved by Euler. They basically consist of finding a traversal on a graph satisfying some conditions related to the links of the graph. ARPs have been studied in depth during the past 40 years due to the large number of real situations that can be modeled in this way: collection or delivery of goods, mail distribution, network maintenance (electrical lines or gas main inspection), snow removal, garbage collection, etc. The book edited by Dror [15] summarizes the state-of-the-art and real-life applications of the ARPs. Other interesting papers are those by Assad and Golden [1] and Eiselt, Gendreau, and Laporte [17, 18].

The first ARP presented in the literature is the well-known Chinese postman problem (CPP) [22]. It consists of finding a shortest tour (closed walk) traversing all of the *links* of a given graph G . Since then, this problem has been generalized in several aspects. Orloff [29] proposed the rural postman problem (RPP), in which not all of the links of the graph have to be traversed by the tour but only those in a given subset of “required” links. Orloff also proposed the general routing problem (GRP), in which the graph has a given set of required links to be traversed and a given set of required vertices to be visited. The GRP can also be considered as a generalization of the graphical traveling salesman problem (GTSP), studied by Cornuéjols, Fonlupt, and Naddef [14], Fleischmann [19], and Naddef and Rinaldi [27]. The GTSP consists of finding a shortest tour visiting all of the vertices of a given graph G at least once.

*Received by the editors September 22, 2005; accepted for publication December 3, 2007; published electronically March 21, 2008. This work was supported by the Ministerio de Educación y Ciencia of Spain (project MTM2006-14961-C05-02).

<http://www.siam.org/journals/sidma/22-2/64088.html>

[†]Corresponding author. Departamento d’Estadística i Investigació Operativa, Universitat de València, València, Spain (angel.corberan@uv.es).

[‡]Departamento d’Estadística i Investigació Operativa, Universitat de València, València, Spain (isaac.plana@uv.es).

[§]Departamento de Matemática Aplicada, Universidad Politécnica de València, València, Spain (jmsanchis@upv.es).

All of these problems were defined on *undirected* graphs, where all of the links are *edges* that can be traversed in both directions with the same cost. These four problems are shown in Figure 1 (box labeled “undirected”), where an arc between two problems means that the first one is generalized by the second one.

Alternatively, the CPP can be defined on a *directed* graph, where all of the links are *arcs* that must be traversed in a given direction. This problem is known as the directed Chinese postman problem (DCPP) [16]. Obviously the RPP and the GRP can also be defined on a *directed* graph to obtain their “directed” versions, the DRPP [5, 33, 24] and the DGRP. The GTSP on a directed graph is known as the graphical asymmetric TSP (GATSP) [3]. These four problems are also represented in Figure 1 (box labeled “directed”).

A graph G is called a *mixed* graph if it has edges and arcs simultaneously. The CPP on a mixed graph, the mixed Chinese postman problem (MCP) [16, 4, 21, 28] is then a generalization of both the CPP and the DCPP. Note that any routing problem defined on a mixed graph generalizes the corresponding problems defined on undirected and directed graphs. This is the case for the mixed RPP (MRPP) [32, 11] and for the mixed GRP (MGRP) [6]. Note that since the GTSP is a pure node routing problem, the GTSP defined on a mixed graph is equivalent to the GATSP. These four “mixed” problems are also shown in Figure 1 (box labeled “mixed”). An arc between two “boxes” means that each problem in the first box is generalized by the corresponding problem in the second box.

Finally, we have the windy postman problem (WPP), proposed by Minieka [26] and also studied by Guan [23], Win [34], and Grötschel and Win [20, 21]. This problem consists of finding a minimum cost traversal of all of the edges of an undirected graph in which the cost of traversing an edge (i, j) in a given direction, c_{ij} , can be different from the cost of traversing it in the opposite direction, c_{ji} . Obviously, the WPP

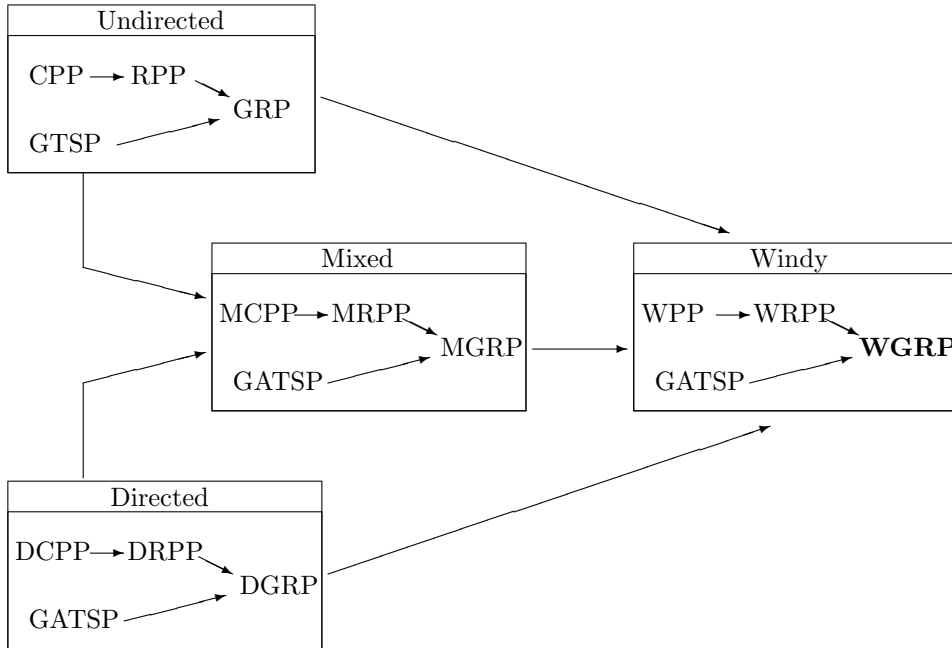


FIG. 1. Relationship among routing problems.

generalizes the undirected CPP (when $c_{ij} = c_{ji}$), but also the DCP (each arc (i, j) with cost c can be modeled as an edge with costs $c_{ij} = c$ and $c_{ji} = \infty$) and therefore the MCP. As before, windy routing problems are also represented in Figure 1 (box labeled “windy”). As for the mixed case, the windy version of the GTSP is equivalent to the GATSP.

In this paper we deal with the windy general routing problem (WGRP). This is the most general routing problem among those represented in Figure 1 and contains all of the other problems as special cases. Consequently, most of the theoretical results obtained for the WGRP and the algorithms designed for its approximate or exact resolution can be applied to many other ARPs.

Except for the CPP and the DCP, which can be solved in polynomial time, all of the other problems in Figure 1 are *NP*-hard problems and therefore the WGRP is also *NP*-hard. The most successful approach to optimally solving these difficult routing problems is polyhedral combinatorics. It is based on the description of the polyhedron defined by the convex hull of all of the solutions associated with a formulation of the problem. This approach has recently been used by Corberán, Plana, and Sanchis [9] to optimally solve WGRP instances of large size, and its success is mostly due to the partial description of the WGRP polyhedron we present here. This description is the subject of this paper.

On the other hand, all of the formulations proposed in the literature for ARPs defined on undirected or directed graphs use only one variable per link. This variable represents, in the directed case, the number of times its associated arc is traversed (in the fixed direction), while, in the undirected case, it represents only the number of times the edge is traversed in either of the two possible directions. All windy problems are formulated by using two variables associated with each edge, representing the number of times the edge is traversed in each direction. However, ARPs defined on mixed graphs can be formulated by using either one variable per edge [28, 11, 6] or two variables [4, 31]. Although the polyhedra associated with ARPs defined on mixed graphs and formulated by using only one variable per edge have already been studied [11, 6], this is not the case when these problems are formulated by using two variables per edge. Then, as a by-product of the polyhedral study of the WGRP presented here, we also obtain some results on the polyhedra associated with ARPs defined on mixed graphs whose formulation uses two variables per edge.

Although the facet-inducing inequalities presented in this paper are associated with a formulation with two variables per edge, they can be easily transformed into facets of the polyhedra of the ARPs shown in Figure 1. Hence, a final contribution of this work is that it also gives a global view of the polyhedra associated with many well-known ARPs.

More specifically, the WGRP is formally defined and some notation is presented in section 2. Section 3 introduces the WGRP polyhedron and presents some results characterizing its facet-defining inequalities. In section 4 several families of valid inequalities and their facet-defining property are described. Section 5 discusses the application of these results to mixed ARPs formulated with two variables per edge. Finally, section 6 presents the conclusions.

2. Problem definition and notation. The WGRP can be defined as follows. Let $G = (V, E)$ be an undirected and connected graph with two nonnegative costs c_{ij} and c_{ji} associated with each edge $(i, j) \in E$, corresponding to the costs of traversing it from i to j and from j to i , respectively. Given a subset $E_R \subseteq E$ of required edges and a subset $V_R \subseteq V$ of required vertices, the problem is to find a minimum cost tour

(closed walk) on G traversing each required edge and visiting each required vertex at least once.

Note that the vertices incident with required edges can also be considered as required vertices and then included in V_R . Without loss of generality, we will assume that the original graph has been transformed to satisfy $V = V_R$. This is not a serious restriction as it is easy to transform WGRP instances which do not satisfy the assumption into equivalent instances which do. Hence, we will assume, in what follows, that we are working with a connected graph $G = (V, E) := (V_R, E_R \cup E_{NR})$.

If we remove all of the nonrequired edges from graph $G = (V, E)$, the resulting graph $G_R = (V, E_R)$ is, in general, not connected. Let p be the number of connected components of graph G_R , and let $V_1, V_2, \dots, V_p$ denote the vertex sets of these connected components, which will be called *R-sets*. We will call the subgraphs $G(V_i)$ of G *R-connected components*. Note that each isolated required vertex itself defines an *R-set*.

If S_1 and S_2 are two vertex sets, (S_1, S_2) represents the set of edges with an end point in S_1 and another end point in S_2 , while $(S_1, S_2)_R$ and $(S_1, S_2)_{NR}$ represent, respectively, the sets of required and nonrequired edges in (S_1, S_2) . Given a vertex subset S , $\delta(S)$ denotes the cutset $(S, V \setminus S)$, i.e., the set of edges with an end point in S and another end point not in S , while $E(S)$ denotes the set of edges with both end points in S . A vertex is called *R-even* (*R-odd*) if it is incident with an even (odd) number of required edges, and a subset of vertices $S \subseteq V$ is *R-even* (*R-odd*) if it contains an even (odd) number of *R-odd* vertices.

A tour for the WGRP is a closed walk on the edges of G traversing each required edge and visiting each required vertex at least once. Each tour is represented by an incidence vector $x \in \mathbb{R}^{2|E|}$ with two variables x_{ij}, x_{ji} associated with each edge $e = (i, j) \in E$ representing the number of times e is traversed in each direction. Given an edge set $F \subseteq E$, $x(F)$ denotes the sum of the variables in x corresponding to the edges in F , i.e., $x(F) := \sum_{(i,j) \in F} (x_{ij} + x_{ji})$. Moreover, a tour for the WGRP can be represented by a directed graph (V, A) , where A contains as many copies of arc (i, j) as the number of times the original edge (i, j) has been traversed from i to j . In a similar way to that given in [2] for the WRPP, the WGRP can be formulated as follows:

$$\begin{aligned}
 & \text{Minimize} && \sum_{(i,j) \in E} (c_{ij}x_{ij} + c_{ji}x_{ji}) \\
 & \text{subject to} \\
 (1) & && x_{ij} + x_{ji} \geq 1 \quad \forall (i, j) \in E_R, \\
 (2) & && \sum_{(i,j) \in \delta(i)} (x_{ij} - x_{ji}) = 0 \quad \forall i \in V, \\
 (3) & && \sum_{i \in S, j \notin S} x_{ij} \geq 1 \quad \forall S = \bigcup_{k \in Q} V_k, \quad Q \subset \{1, \dots, p\}, \\
 (4) & && x_{ij}, x_{ji} \geq 0 \quad \forall (i, j) \in E, \\
 (5) & && x_{ij}, x_{ji} \text{ integer} \quad \forall (i, j) \in E.
 \end{aligned}$$

Traversing inequalities (1) force the tour x to traverse each required edge at least once (in any direction). *Symmetry* equations (2) force the tour x to depart from each vertex as many times as it arrives at it. Given any edge cutset of G , $(S, V \setminus S)$, by

combining some of the equations (2),

$$\sum_{i \in S, j \notin S} x_{ij} = \sum_{i \in S, j \notin S} x_{ji}$$

is obtained. *Connectivity* inequalities (3) prevent the formation of subtours. Since inequalities (1) guarantee that x traverses each edge cutset containing required edges, inequalities (3) are needed only for edge cutsets joining different R -sets. Finally, (4) and (5) assure that all of the variables are nonnegative and integer.

3. The WGRP polyhedron. Let $\text{WGRP}(G)$ denote the convex hull of all of the tours for the WGRP on graph G , i.e., all of the integer vectors $x \in \mathbb{R}^{2|E|}$ satisfying (1)–(4). $\text{WGRP}(G)$ is an unbounded polyhedron which has the following properties.

THEOREM 1. *If G is a connected graph, the dimension of $\text{WGRP}(G)$ is $2|E| - |V| + 1$.*

Proof. The matrix defined by the coefficients of (2) has rank $|V| - 1$, and therefore it follows that $\dim(\text{WGRP}(G)) \leq 2|E| - (|V| - 1)$. On the other hand, it is not difficult to see that $\text{aff}(\text{WGRP}(G)) = \{x \in \mathbb{R}^{2|E|} : x \text{ satisfies (2)}\}$. Therefore, $\dim(\text{WGRP}(G)) \geq 2|E| - |V| + 1$, and the result follows. $\square$

Note that $\text{WGRP}(G)$ is not of full dimension, and different inequalities can induce the same facet of it. Such inequalities are called equivalent. In this section we show that inequalities (4), (1), and (3) from the formulation of the problem are facet-inducing for $\text{WGRP}(G)$.

THEOREM 2. *If edge $e = (i, j) \in E$ is not a bridge of G , then inequalities $x_{ij} \geq 0$ and $x_{ji} \geq 0$ are facet-inducing for $\text{WGRP}(G)$. If in addition $e = (i, j) \in E_R$, then inequality $x_{ij} + x_{ji} \geq 1$ is facet-inducing for $\text{WGRP}(G)$.*

Proof. We will study only the case when $e \in E_R$. If e is not a bridge of G , graph $G \setminus \{e\}$ is connected, and there are $K - 1$ affinely independent tours for the WGRP on $G \setminus \{e\}$, $x^1, x^2, \dots, x^{K-1}$, where $K = \dim(\text{WGRP}(G))$. All of these tours are not tours for the WGRP on G because they do not traverse the required edge e . Let C^1 be a cycle on G traversing e once from j to i . Then $x^1 + C^1, x^2 + C^1, \dots, x^{K-1} + C^1, x^1 + 2C^1$ are K affinely independent tours for the WGRP on G satisfying $x_{ij} = 0$. Hence, inequality $x_{ij} \geq 0$ is facet-inducing for $\text{WGRP}(G)$.

Consider now the inequality $x_{ij} + x_{ji} \geq 1$, and let C^2 be a cycle on G traversing e once from i to j . Then $x^1 + C^1, x^2 + C^1, \dots, x^{K-1} + C^1, x^1 + C^2$ are K affinely independent tours for the WGRP on G satisfying $x_{ij} + x_{ji} = 1$. $\square$

Before discussing the facet-defining property of connectivity and other families of inequalities, we will present some properties that are common to all of them in what follows.

3.1. Weak configuration inequalities. The polyhedra based on standard formulations using only one variable per edge corresponding to the routing problems in Figure 1 satisfy the following property. All of the facet-inducing inequalities (except the trivial ones) are *configuration inequalities* [27]. An inequality $F(x) \geq b_0$ is a configuration inequality if there is a partition $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_r\}$ of V such that the subgraphs $G(\mathcal{B}_k)$ are connected, the variables associated with links in the sets $E(\mathcal{B}_k)$ have coefficient zero in the inequality, and the variables associated with links from a given set $\mathcal{B}_p$ to another given set $\mathcal{B}_q$ have the same coefficient in the inequality.

It is easy to see that $\text{WGRP}(G)$ does not satisfy the above property. Connectivity inequalities (3), for example, are configuration inequalities with $\mathcal{B} = \{S, V \setminus S\}$, while traversing inequalities (1), and others that will be presented later, are not configuration inequalities. In what follows we will show that all of the facet-inducing

inequalities for $\text{WGRP}(G)$, except the trivial and the traversing ones, share some of the properties described above for the configuration inequalities. This will allow us to define a new wider class, the *weak configuration inequalities*.

LEMMA 3. *Let $F(x) \geq b_0$ be a facet-inducing inequality for $\text{WGRP}(G)$ different from the trivial and the traversing ones.*

(a) *For each variable x_{ij} there is a tour x satisfying $F(x) = b_0$ and $x_{ij} \geq 1$.*

(b) *For each required edge $e = (i, j) \in E_R$ there is a tour x satisfying $F(x) = b_0$ and $x_{ij} + x_{ji} \geq 2$.*

Proof. If (a) is not satisfied, every tour x such that $F(x) = b_0$ satisfies $x_{ij} = 0$. Then

$$\{x \text{ WGRP tour on } G : F(x) = b_0\} \subseteq \{x \text{ WGRP tour on } G : x_{ij} = 0\},$$

and, hence, either $F(x) \geq b_0$ is equivalent to a trivial inequality or it is not facet-inducing. In a similar way, if (b) is not satisfied, every tour x such that $F(x) = b_0$ satisfies $x_{ij} + x_{ji} = 1$, and either $F(x) \geq b_0$ is equivalent to a traversing inequality or it is not facet-inducing. $\square$

Part (a) in Lemma 3 implies that, for each variable x_{ij} associated with a *nonrequired* edge (i, j) , there is a tour x such that $F(x) = b_0$ and uses x_{ij} *an extra time*. This means that if a copy of the arc (i, j) is removed from the directed graph associated with x and is replaced by any other path from i to j , another WGRP tour is obtained. Part (a) also implies that, for each variable x_{ij} associated with a *required* edge (i, j) , there is a tour x such that $F(x) = b_0$ and uses x_{ij} at least once, although not necessarily *an extra time*. Finally, part (b) implies that, at least for one variable of each pair x_{ij}, x_{ji} associated with a *required* edge (i, j) , there is a tour x such that $F(x) = b_0$ and uses it *an extra time*.

An inequality $F(x) \geq b_0$ will be called a *weak configuration inequality* if there is a partition $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_r\}$ of V such that the subgraphs $G(\mathcal{B}_k)$ are connected, the variables associated with edges in the sets $E(\mathcal{B}_k)$ have coefficient zero in the inequality, and the variables associated with *nonrequired* edges from a given set $\mathcal{B}_p$ to another given set $\mathcal{B}_q$ have the same coefficient in the inequality. Note that if, in addition, the variables associated with *required* edges from $\mathcal{B}_p$ to $\mathcal{B}_q$ also have the same coefficient in the inequality, we have a (strong) configuration inequality.

THEOREM 4. *All of the facet-inducing inequalities for the $\text{MGRP}(G)$, except those equivalent to the trivial or the traversing ones, are weak configuration inequalities.*

Proof. Let $F(x) := \sum_{(i,j) \in E} (a_{ij}x_{ij} + a_{ji}x_{ji}) \geq b_0$ be such an inequality, and let $E^0 = \{e = (i, j) \in E : a_{ij} = a_{ji} = 0\}$. Let S be the node set of a connected component of $G^0 = (V, E^0)$.

Assume that there is an edge $e = (i, j) \in E(S)$ with $a_{ij} \neq 0$ (necessarily $a_{ij} > 0$). Let x^* be a tour for the WGRP on G such that $F(x^*) = b_0$ and $x_{ij}^* \geq 1$ (it exists from Lemma 3). Let us replace in x^* arc (i, j) by a path from i to j in $E^0(S)$. If x^* uses x_{ij} an extra time, we obtain a tour x^{**} for the WGRP on G satisfying $F(x^{**}) < b_0$, which is a contradiction. From Lemma 3, this is true for all of the variables associated with nonrequired edges and for one variable of each pair associated with a required edge. Then, if x^* does not use x_{ij} an extra time, we add to x^* the arcs in a cycle formed by the arc from j to i (whose associated variable x_{ji} , which is used an extra time, has coefficient zero) and a path in E^0 from i to j , obtaining a tour for the WGRP on G violating the inequality. Therefore, if $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_r\}$ are the vertex sets of the connected components of G^0 , the induced graphs $G(\mathcal{B}_k)$ are connected and $a_{ij} = a_{ji} = 0$, for all $e = (i, j) \in E(\mathcal{B}_k)$, $k = 1, 2, \dots, r$.

Let us suppose now that there are two nonrequired edges $e=(i, j)$, $f=(u, v)$ with $i, u \in \mathcal{B}_p$ and $j, v \in \mathcal{B}_q$ such that $a_{ij} > a_{uv}$. Let x^* be a tour for the WGRP on G using x_{ij} an extra time and $F(x^*) = b_0$. Replace in x^* arc (i, j) by a path in $G(\mathcal{B}_p)$ from i to u , the edge f traversed from u to v , and a path in $G(\mathcal{B}_q)$ from v to j . This way, a tour violating the inequality is obtained, which is a contradiction. $\square$

Then, associated with each facet-inducing inequality for the WGRP(G), we have a *configuration graph* $G_C = (V^C, E^C)$. This is the graph resulting from shrinking vertex sets $\mathcal{B}_k$, $k = 1, \dots, r$, into a single vertex each and, after that, shrinking each set of nonrequired parallel edges into one single nonrequired edge but keeping all of the required edges. If two costs equal to the coefficients of the corresponding variables in the inequality are associated with each edge in G_C , this graph keeps all of the information about the inequality, since b_0 is the length of the shortest tour for the WGRP on graph G_C . Hence, each facet-inducing inequality for the WGRP(G) can be represented by means of its corresponding configuration graph G_C . The following theorem implies that, in order to show that a given inequality is facet-inducing for WGRP(G), it will suffice to prove it for the polyhedron associated with the configuration graph, which is simpler in general.

THEOREM 5. *If a weak configuration inequality is facet-inducing for WGRP(G_C), it is also facet-inducing for WGRP(G).*

Proof. It suffices to prove the result for a graph G obtained from G_C by replacing vertex $\mathcal{B}_1$ (for example) by a connected graph $G(\mathcal{B}_1) = (V^{\mathcal{B}_1}, E^{\mathcal{B}_1})$. For simplicity, $F(x) \geq b_0$ will denote both the inequality on G and the inequality on G_C . Note that, given a tour x for the WGRP on G , by shrinking $G(\mathcal{B}_1)$ into a single vertex we obtain a tour x^* for the WGRP on G_C satisfying $F(x) = F(x^*) \geq b_0$.

Let $K = \dim(\text{WGRP}(G_C))$, and let us suppose that $b_0 \neq 0$. Then there are K linearly independent tours for the WGRP on G_C satisfying $F(x) = b_0$, say, $x^1, x^2, \dots, x^K$. After replacing $\mathcal{B}_1$ by the graph $G(\mathcal{B}_1)$, some nonrequired edges may appear between the vertices of $G(\mathcal{B}_1)$ and nodes $\mathcal{B}_i$, $i \neq 1$. Let us call G'_C the graph obtained from G_C by adding Q nonrequired edges $e'_1, e'_2, \dots, e'_Q$, where each e'_i is parallel to a nonrequired edge (not necessarily distinct) e_j of G_C .

For each e'_i , $i = 1, 2, \dots, Q$, we will construct two tours in G'_C . Let $e_j = (\mathcal{B}_1, \mathcal{B}_q)$ be the original edge of G_C to which e'_i is parallel. From Lemma 3, there is a tour in G_C satisfying $F(x) = b_0$ and that traverses e_j from $\mathcal{B}_1$ to $\mathcal{B}_q$ and another one that traverses it in the opposite direction. By replacing the edge e_j by the edge e'_i in these two tours, we obtain two new tours on G'_C satisfying $F(x) = b_0$. Then we have $K+2Q$ linearly independent tours on G'_C satisfying $F(x) = b_0$. These tours can be completed with edges of $G(\mathcal{B}_1)$ to obtain $K+2Q$ tours on G , $x^1, x^2, \dots, x^{K+2Q}$, also satisfying $F(x) = b_0$ because the variables associated with the edges of $G(\mathcal{B}_1)$ do not appear in the inequality.

Consider now the WGRP defined on graph $G(\mathcal{B}_1)$. Let $M = \dim(\text{WGRP}(G(\mathcal{B}_1)))$, and let $y^1, y^2, \dots, y^M$ be M linearly independent tours for the WGRP on $G(\mathcal{B}_1)$. Then $x^1 + y^1, x^1 + y^2, \dots, x^1 + y^M$ are also tours for the WGRP on G satisfying $F(x) = b_0$. By expressing the $K+2Q+M$ tours as rows of a matrix and subtracting the first row x^1 from the last M rows, we obtain a full rank matrix. Then, since $K = 2|E^C| - |V^C| + 1$ and $M = 2|E^{\mathcal{B}_1}| - |V^{\mathcal{B}_1}| + 1$, we have $K+2Q+M = 2(|E^C| + |E^{\mathcal{B}_1}| + Q) - (|V^C| - 1 + |V^{\mathcal{B}_1}|) + 1 = \dim(\text{WGRP}(G))$ linearly independent tours satisfying $F(x) = b_0$. $\square$

As mentioned before, strong and weak configuration inequalities differ only in the fact that, in the first one, “parallel” variables (variables x_{ij} , x_{km} associated with edges in $(\mathcal{B}_p, \mathcal{B}_q)$, with $i, k \in \mathcal{B}_p$ and $j, m \in \mathcal{B}_q$) have equal coefficients, while the

last one can have parallel variables with different coefficients. By arguing as in the proof of Theorem 4 above, it can be shown that, if for each variable x_{ij} there is a tour satisfying $F(x) = b_0$ and using x_{ij} an extra time, then $F(x) \geq b_0$ is a strong configuration inequality, with all of the variables in $(\mathcal{B}_p, \mathcal{B}_q)$ having a coefficient equal to a_{pq} or a_{qp} . A variable with a different coefficient in a weak configuration inequality is then a variable x_{ij} associated with a required edge $(i, j) \in (\mathcal{B}_p, \mathcal{B}_q)$ such that there is no tour satisfying $F(x) = b_0$ which uses x_{ij} an extra time. It can be seen that, in this case, the coefficients of variables x_{ij} and x_{ji} are, respectively, $a_{pq} + k$ and a_{qp} , for some k satisfying $0 < k \leq a_{pq} + a_{qp}$.

Therefore, if a given facet-inducing inequality $F(x) \geq b_0$ for $\text{WGRP}(G)$ is not a strong configuration inequality, then there is at least one variable x_{ij} (associated with a required edge) such that every tour x satisfying $F(x) = b_0$ also satisfies either $x_{ij} = 0$ or $x_{ij} + x_{ji} = 1$. In other words, if a given facet $\mathcal{F}$ of $\text{WGRP}(G)$ is not induced by a strong configuration inequality, then every tour for the WGRP on G which lies in $\mathcal{F}$ also lies either in a trivial facet or in a traversing facet of $\text{WGRP}(G)$. Hence, we think that there must be relatively few such facets in $\text{WGRP}(G)$. In fact, the only inequalities of this type that we have found are a subset of the class of zigzag inequalities. All of the other inequalities presented in the following sections are strong configuration inequalities.

3.2. Connectivity inequalities. To finish this section, we will study the conditions under which connectivity inequalities (3) are facet-inducing for $\text{WGRP}(G)$. Furthermore, the proof of Theorem 6 shows the usefulness of the configuration issue and illustrates the use of Theorem 5 above.

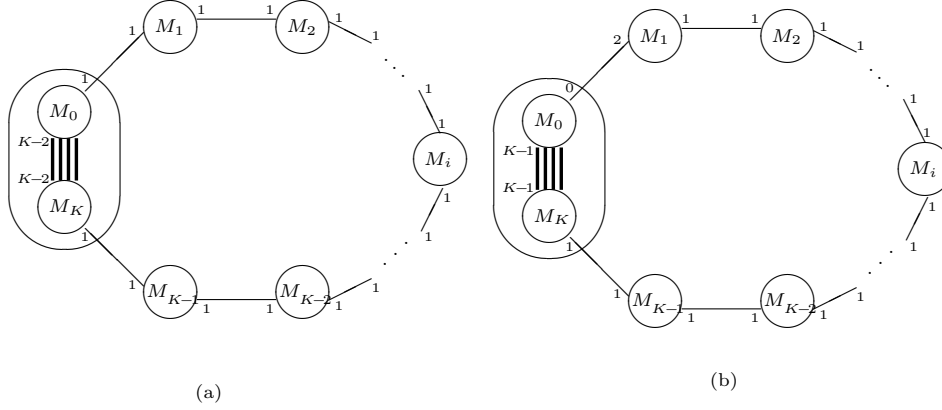
THEOREM 6. *Connectivity inequalities (3) are facet-inducing for $\text{WGRP}(G)$ if subgraphs $G(S)$ and $G(V \setminus S)$ are connected.*

Proof. The configuration graph G_C associated with connectivity inequalities has only two nodes, say, 1 and 2, corresponding to S and to $V \setminus S$, respectively, joined by a single nonrequired edge. Therefore, $\dim(\text{WGRP}(G_C)) = 1$, and, since the tour $x_{12} = x_{21} = 1$ satisfies $\sum_{i \in S, j \notin S} x_{ij} = x_{12} = 1$, the inequality is facet-inducing for $\text{WGRP}(G_C)$. By applying Theorem 5, connectivity inequalities (3) are facet-inducing for $\text{WGRP}(G)$. $\square$

4. Other facets of the WGRP polyhedron. In the presence of the integrality conditions (5), equations (2) and inequalities (4), (1), and (3) in the formulation are enough to describe all of the tours for the WGRP on G . Nevertheless, this is not the case when the integrality conditions are relaxed. In this section we study some other families of inequalities which are facet-inducing for $\text{WGRP}(G)$.

We first obtain valid inequalities for the WGRP from known valid inequalities for the undirected GRP. Let $G = (V, E)$ be an undirected graph, and let $\text{GRP}(G)$ be the polyhedron associated with the formulation of the undirected GRP using one variable y_e per edge $e = (i, j) \in E$, representing the number of times the edge is traversed in any direction. Consider now the same graph $G = (V, E)$ but with two nonnegative costs associated with each edge, and let $\text{WGRP}(G)$ be the polyhedron studied in this paper. In [2] it is shown that, if $\sum_{e \in E} \beta_e y_e \geq b_0$ is a valid inequality for $\text{GRP}(G)$, the inequality $\sum_{e=(i,j) \in E} \beta_e (x_{ij} + x_{ji}) \geq b_0$ is valid for $\text{WGRP}(G)$. Therefore, the R -odd cut, K-C, path-bridge, and honeycomb inequalities presented in what follows are valid inequalities for the $\text{WGRP}(G)$.

4.1. R -odd cut inequalities. An edge cutset $\delta(S)$ is called R -odd if it contains an odd number of required edges. Given that any tour must cross any given edge cutset

FIG. 2. K -C and K -C₀₂ configuration graphs.

an even number of times, the R -odd cut inequalities

$$(6) \quad x(\delta(S)) \geq |\delta_R(S)| + 1 \quad \forall S \subset V \quad \text{such that } \delta(S) \text{ is } R\text{-odd}$$

are valid for $\text{WGRP}(G)$. It can be proved [30] that they are facet-inducing for $\text{WGRP}(G)$ if subgraphs $G(S)$ and $G(V \setminus S)$ are connected.

4.2. K-C inequalities. K-C inequalities were introduced in [12] for the undirected RPP, and in [11] they were generalized to the mixed GRP in two families, the K-C and the K-C₀₂ inequalities. In this section we will study their generalization to the WGRP.

Consider a partition of V into $K+1$ subsets $\{M_0, M_1, M_2, \dots, M_{K-1}, M_K\}$, $K \geq 3$, such that each R -set V_i is contained by one of the node sets $M_0 \cup M_K, M_1, M_2, \dots, M_{K-1}$, the induced subgraphs $G(M_i)$, $i = 0, 1, 2, \dots, K$, are connected, and (M_0, M_K) contains a positive and even number of required edges. Furthermore, we will assume that sets (M_i, M_{i+1}) are nonempty, $i = 0, 1, \dots, K-1$. Such a partition defines the configuration graph shown in Figure 2(a), where the costs of its edges are $c(M_0, M_K) = c(M_K, M_0) = K-2$ and $c(M_i, M_j) = |i-j|$ for all i, j such that $\{i, j\} \neq \{0, K\}$. These costs, shown in Figure 2(a), correspond to the coefficients of the variables in the inequality. Each number represents the coefficient of the variable associated with the traversal of the edge from the nearest node to the farthest one. The K -C inequality is then

$$(7) \quad (K-2) x(M_0, M_K) + \sum_{\substack{0 \leq i < j \leq K \\ \{i, j\} \neq \{0, K\}}} |i-j| x(M_i, M_j) \geq 2(K-1) + (K-2)|(M_0, M_K)_R|.$$

Since K-C inequalities are a particular case of path-bridge inequalities, which will be presented later, the proof that K-C inequalities are facet-inducing is omitted here.

4.3. K-C₀₂ inequalities. With the same underlying configuration graph, K-C₀₂ inequalities differ from the standard ones in that integer K can now take value 2 and in that the coefficients of the variables and the right-hand side (RHS) change as follows:

- $c(M_0, M_K) = c(M_K, M_0) = K-1$;
- $c(M_0, M_1) = 0$, $c(M_1, M_0) = 2$;
- $c(M_i, M_{i+1}) = c(M_{i+1}, M_i) = 1$, $i = 1, 2, \dots, K-1$;

- otherwise, $c(M_i, M_j)$ is the shortest path cost from M_i to M_j in G_C , i.e.,
 - $c(M_0, M_j) = j - 1$, $1 \leq j \leq K - 1$,
 - $c(M_j, M_0) = j + 1$, $1 \leq j \leq K - 1$,
 - $c(M_i, M_j) = |i - j|$, $1 \leq i, j \leq K$.

The corresponding K -C₀₂ inequality is then

$$(8) \quad \sum_{(i,j) \in E} (a_{ij}x_{ij} + a_{ji}x_{ji}) \geq 2(K-1) + (K-1)|(M_0, M_K)_R|,$$

where $a_{ij} = c(M_p, M_q)$ if $i \in M_p$ and $j \in M_q$. These inequalities are also a special case of the path-bridge₀₂ inequalities presented next.

4.4. Path-bridge inequalities. Path-bridge (PB) inequalities were introduced by Letchford [25] for the undirected GRP and are based on the path inequalities for the GTSP proposed by Cornuéjols, Fonlupt, and Naddef [14]. In this section we study their extension to the WGRP. Let $P \geq 1$ and $B \geq 0$ be integers such that $P + B \geq 3$ is an odd number, and let $n_i \geq 2$, $i = 1, \dots, p$, be integer numbers. Let us define a partition $\{A, Z, \{V_j^i\}_{j=1, \dots, n_i}^{i=1, \dots, P}\}$ of V satisfying that each R -set is contained either in $A \cup Z$ or in a set V_j^i , the induced subgraphs $G(V_j^i)$ are connected, $i = 1, 2, \dots, P$, $j = 0, 1, 2, \dots, n_i + 1$ (where for the sake of simplicity we identify A with V_0^i and Z with $V_{n_i+1}^i$ for all i), and (A, Z) contains a number B of required edges. Furthermore, we will suppose that sets (V_j^i, V_{j+1}^i) are nonempty. Then the configuration graph has P paths from A to Z , each of them with $n_i + 2$ vertices and $n_i + 1$ edges. The edges joining the sets A and Z form the *bridge*. This configuration graph is shown in Figure 3, where only one number near an edge means that the coefficients of its two corresponding variables are equal. These coefficients are

- $c(A, Z) = c(Z, A) = 1$,
- $c(V_j^i, V_k^i) = \frac{|j-k|}{n_i-1}$ for all $j, k \in \{0, 1, \dots, n_i+1\}$, $i \in \{1, \dots, P\}$,
- $c(V_j^i, V_k^r) = \frac{1}{n_i-1} + \frac{1}{n_r-1} + \left| \frac{j-1}{n_i-1} - \frac{k-1}{n_r-1} \right|$ for all $i \neq r \in \{1, \dots, P\}$, $j \in \{1, \dots, n_i\}$, $k \in \{1, \dots, n_r\}$.

The PB inequality is then

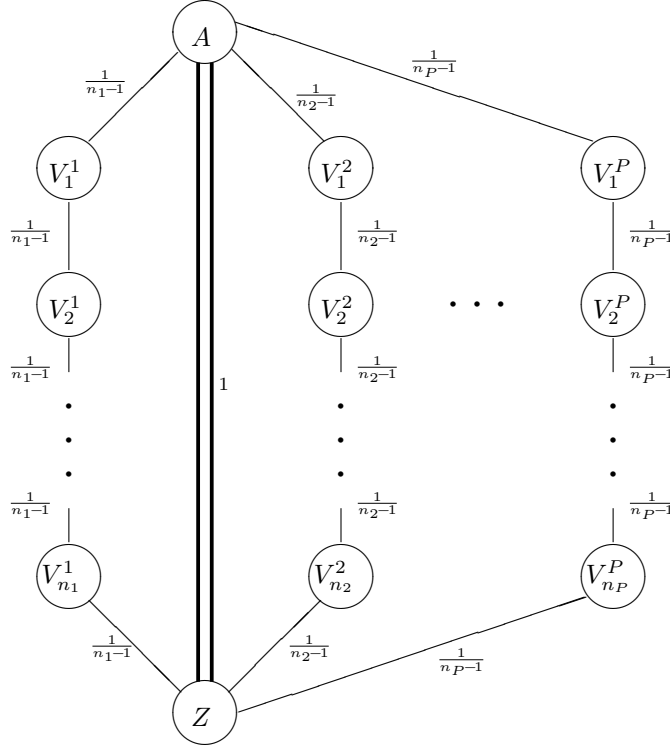
$$(9) \quad \sum_{(i,j) \in E} (a_{ij}x_{ij} + a_{ji}x_{ji}) \geq 1 + \sum_{i=1}^P \frac{n_i + 1}{n_i - 1} + |(A, Z)_R|,$$

where, again, $a_{ij} = c(V_p^k, V_q^l)$ if $i \in V_p^k$ and $j \in V_q^l$.

A PB inequality is called n -regular if all of the paths have the same length n . In this case, by multiplying it by $n-1$, the inequality (9) can be rewritten in an easier way. Note that a K-C inequality is a particular case of a PB inequality with $P = 1$ and that PB inequalities reduce to the s -path inequalities for the GATSP, introduced by Chopra and Rinaldi [3], when there are no required links crossing from A to Z .

THEOREM 7. *PB inequalities (9) are facet-inducing for WGRP(G).*

Proof. Let us first suppose that the number B of required edges joining A and Z is even. Then the number P of paths from A to Z is odd. If we consider in G_C all of the required edges as nonrequired and we replace each edge by a pair of opposite arcs, we obtain an instance G'_C of the GATSP. By subtracting $|(A, Z)_R|$ from the RHS of the PB inequality, an s -path inequality is obtained, which is facet-inducing for GATSP(G'_C) (see [3]). The dimension of this polyhedron is $2|E_C| - |V_C| + 1$, and there is this number of affinely independent tours for the GATSP on G'_C satisfying the inequality as an equality. Each one of these tours can be easily transformed into

FIG. 3. *PB configuration graph.*

tours for the WGRP on G_C (by adding to them one arc per each required edge in (A, Z) , half of them directed from A to Z and the other half from Z to A) satisfying the PB inequality (9) as an equality. Hence, PB inequalities are facet-inducing for $\text{WGRP}(G_C)$ and, from Theorem 5, for $\text{WGRP}(G)$.

Let us now suppose that B is odd and P is even. Let $\bar{e}$ be a given required edge in (A, Z) . We construct the tour x^1 as the one traversing edge $\bar{e}$ once in each direction, half of the edges in $(A, Z)_R \setminus \{\bar{e}\}$ from A to Z and the other half from Z to A . Tour x^1 traverses half of the paths from A to Z and the other half from Z to A . From x^1 we construct the following tours:

- (a) For each edge $e \in (A, Z) \setminus \{\bar{e}\}$ we define a tour traversing e from A to Z once more than tour x^1 does and traversing $\bar{e}$ only from Z to A . All of the other edges are traversed as in x^1 . Another similar tour traversing e from Z to A and $\bar{e}$ from A to Z is considered. Then we have $2|(A, Z)| - 2$ tours.
- (b) For each edge $e = (V_i^k, V_{i+1}^k)$, we construct a tour x that traverses all of the edges in the path k , except e , in both directions. Edge $\bar{e}$ is traversed by x in the same direction as path k is traversed by x^1 . All of the other edges are traversed by x as they are traversed by x^1 . Then we have $\sum_{i=1}^P (n_i + 1)$ tours.
- (c) For each path k , it is possible to build $n_i + 1$ different tours by using the opposite orientations to those used in (b) for the other $P - 1$ paths. From among them we select, for each path k , only the tour that does not traverse the edge (V_0^k, V_1^k) , and we then have P tours.
- (d) For each of the remaining edges (if any) there are always two tours for the

	$\bar{e}$	$(A, Z) \setminus \{\bar{e}\}$	Paths	E'
(a)	*	I	0	0
(b)+(c)	*	0	B	0
(d)	*	*	*	I

FIG. 4. Matrix appearing in the proof of Theorem 7.

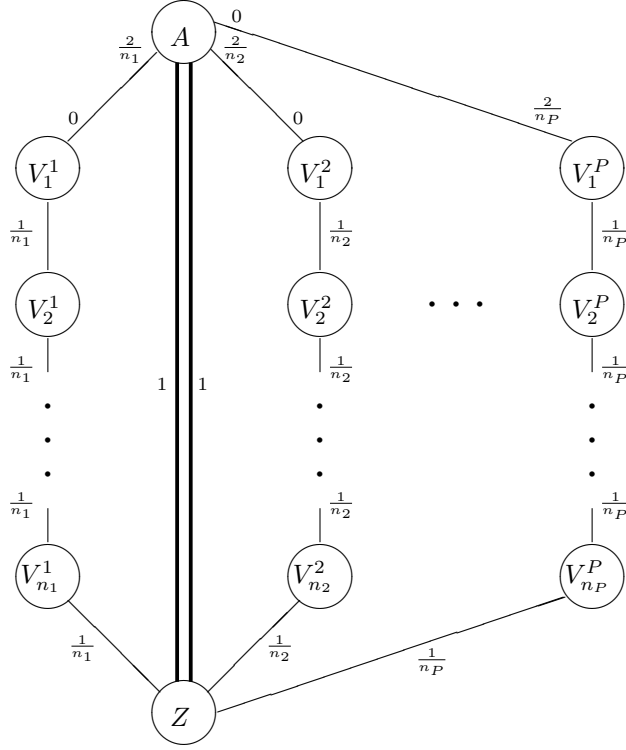
WGRP in G_C satisfying (9) as an equality, traversing this edge exactly once in one direction each, edges in (A, Z) and edges in the P paths. Then, if E' is the set of such remaining edges, we have obtained $2|E'|$ tours.

By including x^1 we have $1+2|(A, Z)|-2+\sum_{i=1}^P(n_i+1)+P+2|E'|$ tours, all of them satisfying the PB inequality as an equality. Hence we have $2|E_C| - \sum_{i=1}^P n_i - 1 = \dim(\text{WGRP}(G_C))$ tours. If we subtract x^1 from all of the other tours and express them as the rows of a matrix (sorted as (a), (b)+(c), and (d) above) and the variables associated with the edges as columns (sorted as $\bar{e}$, $(A, Z) \setminus \{\bar{e}\}$, (V_i^k, V_{i+1}^k) for all i, k , and E'), we obtain the matrix shown in Figure 4. Submatrix B is a full rank matrix (see [30] for details) and, hence, also the complete matrix in Figure 4. Therefore, the PB inequality is facet-inducing for $\text{WGRP}(G_C)$, and then it is also facet-inducing for $\text{WGRP}(G)$. $\square$

4.5. Path-bridge₀₂ inequalities. Path-bridge₀₂ inequalities were presented in [6] for the MGRP. The configuration graph corresponding to a path-bridge₀₂ inequality is similar to the previous one except that integers n_i can now take value 1 and the paths are classified into two types: AZ paths and ZA paths. The associated configuration graph is shown in Figure 5, where the two coefficients associated with an edge are represented only if they are different.

The coefficients are

- $c(A, Z) = c(Z, A) = 1$,
- $c(V_0^i, V_1^i) = 0$ and $c(V_1^i, V_0^i) = \frac{2}{n_i}$ for each AZ path i ,
- $c(V_0^i, V_1^i) = \frac{2}{n_i}$ and $c(V_1^i, V_0^i) = 0$ for each ZA path i ,
- $c(V_j^i, V_{j+1}^i) = c(V_{j+1}^i, V_j^i) = \frac{1}{n_i}$,
- $c(V_j^i, V_k^i)$, with $|j-k| > 1$, is the shortest path cost from V_j^i to V_k^i using edges of the path i , i.e.,
 - $c(V_j^i, V_k^i) = \frac{|j-k|}{n_i}$ for all $j \neq k \in \{1, 2, \dots, n_i+1\}$,
 - $c(V_0^i, V_k^i) = \frac{k-1}{n_i}$ and $c(V_k^i, V_0^i) = \frac{k+1}{n_i}$ if path i is of type AZ , $1 \leq k \leq n_i$,
 - $c(V_0^i, V_k^i) = \frac{k+1}{n_i}$ and $c(V_k^i, V_0^i) = \frac{k-1}{n_i}$ if path i is of type ZA , $1 \leq k \leq n_i$.
- Order the remaining edges (edges in (V_j^i, V_l^k) , $i \neq k$, if any) in an arbitrary way $e_1, e_2, \dots, e_h$. For $i=1$ to h , if $e_i = (u, v)$, we assign to c_{uv} the maximum value such that there is a WGRP tour on G_C with cost $P+1+|(A, Z)_R|$ traversing e_i just from u to v but not traversing the edges $\{e_{i+1}, \dots, e_h\}$. Similarly, we assign to c_{vu} the maximum value such that there is a tour of c -length $P+1+|(A, Z)_R|$ traversing e_i from v to u (perhaps also from u to v) but not traversing the edges $\{e_{i+1}, \dots, e_h\}$ (sequential lifting).

FIG. 5. PB_{02} configuration graph.

The path-bridge $_{02}$ inequality PB_{02} is then

$$(10) \quad \sum_{(i,j) \in E} (a_{ij}x_{ij} + a_{ji}x_{ji}) \geq P + 1 + |(A, Z)_R|.$$

Note that the main difference between PB_{02} and PB inequalities is that variables in sets (A, V_1^i) have different coefficients for its two associated variables. One of these coefficients is $\frac{2}{n_i}$, twice the cost of the other edges in path i , while the other coefficient is 0 (see Figure 5). The direction represented by the variable with coefficient 0 in each path determines whether the path is of type AZ or ZA . In a path-bridge $_{02}$ configuration graph, the pair of asymmetric coefficients $(\frac{2}{n_i}, 0)$ for each path i can be associated with any edge (V_j^i, V_{j+1}^i) in the path, thus obtaining different but equivalent inequalities (given the same ordering in the sequential lifting process). Finally, when $P=1$, the path-bridge $_{02}$ inequality (10) reduces to a $K-C_{02}$ inequality.

THEOREM 8. *Path-bridge $_{02}$ inequalities (10) are valid for $WGRP(G)$.*

Proof. The proof is similar to that in [6] for the MGRP and is omitted here for the sake of brevity. $\square$

THEOREM 9. *Let $F(x) \geq b_0$ be a path-bridge $_{02}$ inequality (10), and let P_{AZ} and P_{ZA} be the number of paths of types AZ and ZA , respectively. Then $F(x) \geq b_0$ is facet-inducing of $WGRP(G)$ if $|(A, Z)_R| \geq |P_{AZ} - P_{ZA}| + 1$.*

Proof. The proof is similar to that of Theorem 7 and is omitted here. Note only that condition $|(A, Z)_R| \geq |P_{AZ} - P_{ZA}| + 1$ is needed in order to construct $WGRP$ tours traversing the paths of type AZ (type ZA) from A to Z (from Z to A) and

satisfying $F(x) = b_0$. $\square$

As for the (standard) PB inequalities, when all of the paths are of the same length n , we obtain the n -regular PB₀₂ inequalities, which can be written in an easier way.

4.6. Honeycomb inequalities. Honeycomb inequalities were first proposed by Corberán and Sanchis [13] for the undirected GRP and by Corberán, Mejía, and Sanchis [6] for the MGRP. These inequalities also generalize the K-C inequalities. Consider a partition $\{A_1, A_2, \dots, A_L, A_{L+1}, \dots, A_K\}$ of V , $3 \leq K \leq p$, $1 \leq L \leq K$, such that each R -set is contained by one A_i and the induced subgraphs $G(A_i)$ are connected. Each set A_i , $i = 1, 2, \dots, L$, is divided into $\gamma_i \geq 2$ subsets, $A_i = B_i^1 \cup B_i^2 \cup \dots \cup B_i^{\gamma_i}$, satisfying that each B_i^j contains an even number of R -odd nodes and the induced subgraphs $G(B_i^j)$ are connected, $j = 1, 2, \dots, \gamma_i$, and the graph defined by nodes $B_i^1, B_i^2, \dots, B_i^{\gamma_i}$ and the required edges (B_i^j, B_i^k) is connected (and even).

Note that, when A_i consists of several R -sets, the last condition implies that the partition of A_i into the B_i^j is made by cutting the R -sets. For notational convenience, we denote $B_i^0 = A_i$, $i = L+1, \dots, K$. Therefore we have the following partition of V :

$$\mathcal{B} = \{B_1^1, B_1^2, \dots, B_1^{\gamma_1}, B_2^1, B_2^2, \dots, B_2^{\gamma_2}, \dots, B_L^1, B_L^2, \dots, B_L^{\gamma_L}, B_{L+1}^0, \dots, B_K^0\}.$$

This partition $\mathcal{B}$ defines the configuration graph $G_C = (\mathcal{B}, \mathcal{E})$. Let T be a set of nonrequired edges in G_C joining nodes corresponding to different A_j such that $(\mathcal{B}, T)$ is a spanning tree (see Figure 6, where the arcs in T are represented by thin lines and the required links by bold lines). Then, for each pair of nodes B_i^j, B_p^q in $\mathcal{B}$, let $d(B_i^j, B_p^q)$ denote the number of edges in the unique path in $(\mathcal{B}, T)$ joining B_i^j to B_p^q . We will also assume that $d(B_i^j, B_i^q) \geq 3$ for all $i = 1, \dots, L$ and for all $j \neq q$.

We divide the set $\mathcal{E}$ into 3 subsets: the set $\mathcal{C}$ formed by the edges joining nodes B_i^p, B_i^q , with $p, q \neq 0$ (the nodes obtained by “cutting” the sets A_i , $i = 1, \dots, L$), the set of edges in T , and the set formed by the remaining edges, which will be called $\mathcal{I}n$. The costs of the edges of graph $G_C = (\mathcal{B}, \mathcal{E})$ are (see Figure 6) the following:

- (I) For the edges $(B_q^i, B_q^j) \in \mathcal{C}$ such that the path in $(\mathcal{B}, T)$ joining B_q^i and B_q^j does not contain more than one node related to the same A_s (except the nodes B_q^i and B_q^j), $c(B_q^i, B_q^j) = c(B_q^j, B_q^i) = d(B_q^i, B_q^j) - 2$.

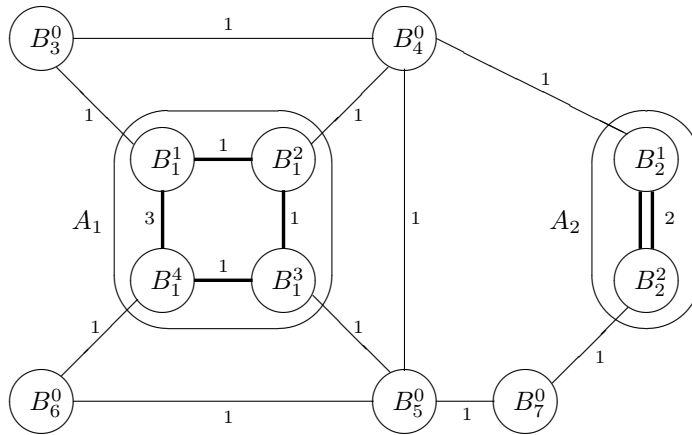


FIG. 6. Honeycomb configuration graph.

- (II) For the edges $(B_r^i, B_q^j) \in T \cup \mathcal{I}n$, $r \neq q$, such that the path in $(\mathcal{B}, T)$ joining B_r^i and B_q^j does not contain more than one node related to the same A_s , $c(B_r^i, B_q^j) = c(B_q^j, B_r^i) = d(B_r^i, B_q^j)$.
- (III) For the remaining edges (if any), we compute their two costs with a process of sequential lifting such as the one described for PB_{02} inequalities.

The corresponding *honeycomb inequality* is then

$$(11) \quad \sum_{(i,j) \in E} (a_{ij}x_{ij} + a_{ji}x_{ji}) \geq 2(K-1) + \sum_{(B_q^i, B_q^j) \in \mathcal{E}_R} c(B_q^i, B_q^j),$$

where $a_{ij} = c(B_r^i, B_q^j)$ if $i \in B_r^i$ and $j \in B_q^j$.

Note that, when the configuration graph has no edges of type III, all of the coefficients in the honeycomb inequality can be computed in terms of the shortest distances in the graph $(\mathcal{B}, T)$. This occurs, for example, when every node B_q^i , $i \neq 0$, has degree 1 in $(\mathcal{B}, T)$. Given that the sequential lifting process for a set of edges guarantees the validity of an inequality if it is valid without this set of edges, and that this is also true for the facet-inducing property, in what follows we will assume that the honeycomb configuration has no edges of type III.

Before proving that the honeycomb inequalities are facet-inducing for $WGRP(G)$, let us show how to build tours for the WGRP on $G_C = (\mathcal{B}, \mathcal{E})$ satisfying the honeycomb inequality $F(x) \geq b_0$ as an equality. Let $(\bar{\mathcal{A}}, T_{\bar{\mathcal{A}}})$ denote the graph with node set $\bar{\mathcal{A}} = \{A_1, A_2, \dots, A_L, A_{L+1}, \dots, A_K\}$ and having an edge (A_i, A_j) for each edge (B_i^p, B_j^q) in T (Figure 7(a)). Then, by traversing all of the $K-1$ edges of any spanning tree of $(\bar{\mathcal{A}}, T_{\bar{\mathcal{A}}})$ once in each direction, and each required edge in G_C once, we obtain a tour of cost b_0 (Figure 7(b)).

THEOREM 10. *Honeycomb inequalities (11) are facet-inducing of $WGRP(G)$ if the shrunk graph $(\bar{\mathcal{A}}, T_{\bar{\mathcal{A}}})$ is 2-connected.*

Proof. The dimension of $WGRP(G_C)$ is $2|\mathcal{E}| - |\mathcal{B}| + 1 = 2|\mathcal{E}| - |T| = 2|\mathcal{C}| + 2|\mathcal{I}n| + |T|$, and this is the number of linear independent WGRP tours x satisfying the inequality as an equality that we build in what follows. Let x^d be the incidence vector of a set of L cycles traversing each required edge exactly once (such cycles exist since the required edges in G_C form even graphs).

For each variable x_{uv} associated with an edge in $\mathcal{C} \cup \mathcal{I}n$, consider the tour that uses x_{uv} exactly once, the path in T from v to u , and the required edges as in x^d . The nodes still not visited are connected with edges in T used once in each direction. Then we have $2|\mathcal{C}| + 2|\mathcal{I}n|$ tours.

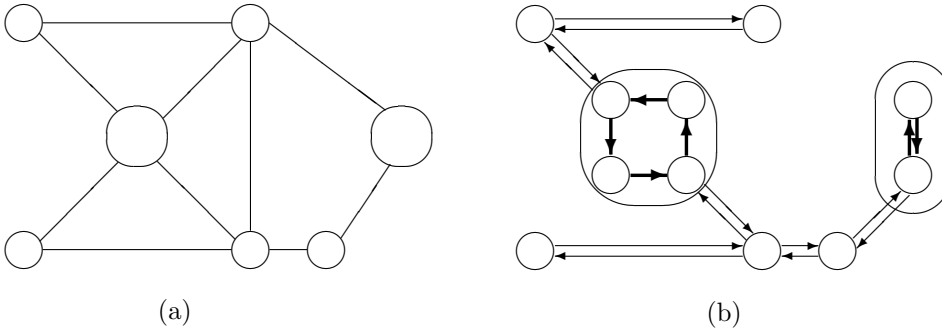


FIG. 7. Graph $(\bar{\mathcal{A}}, T_{\bar{\mathcal{A}}})$ and a WGRP tour on G_C satisfying (11) as an equality.

Since the graph $(\bar{\mathcal{A}}, T_{\bar{\mathcal{A}}})$ is 2-connected, it can be proved that the only equation satisfied by the incidence vectors of all its spanning trees is $\sum_{e \in T_{\bar{\mathcal{A}}}} x_e = |\bar{\mathcal{A}}| - 1$. Hence, $|T_{\bar{\mathcal{A}}}| = |T|$ different spanning trees can be selected in such a way that their incidence vectors are linearly independent. For each spanning tree of graph $(\bar{\mathcal{A}}, T_{\bar{\mathcal{A}}})$ we can define a WGRP tour, such as the one shown in Figure 7(b), which uses each edge in the tree once in each direction and the required edges as in x^d .

If we subtract x^d from the $2|\mathcal{C}| + 2|\mathcal{I}n| + |T|$ tours above and we express them as rows of a matrix, then we obtain a full rank matrix. Hence, the honeycomb inequality is facet-inducing for $\text{WGRP}(G_{\mathcal{C}})$, and, by applying Theorem 5, it is facet-inducing for $\text{WGRP}(G)$. $\square$

4.7. Honeycomb₀₂ inequalities. As for K-C and PB inequalities, there is another version of the honeycomb inequalities in which some edges have coefficients 0 and 2. Honeycomb₀₂ inequalities were presented in [6] for the MGRP. As in that paper, we consider here only the case $L = 1$, i.e., when only one R -set (or cluster of R -sets) is divided into γ_1 parts. We have not studied the general case because it is quite a bit more complicated. The reason is that, after replacing some pairs of coefficients (1, 1) in a honeycomb configuration with $L > 1$ by coefficients (0, 2), a nonvalid inequality can be obtained.

With a similar configuration graph $G_{\mathcal{C}} = (\mathcal{B}, \mathcal{E})$ and tree T (see Figure 8), the nodes in set A_1 are classified into two types, nodes of type $\mathcal{O}$ (those that will be incident with edges in T with coefficients 0 and 2) and nodes of type $\mathcal{I}$, in such a way that there is at least one node of each type. We will assume that $d(B_1^i, B_1^j) \geq 3$ if B_1^i, B_1^j are nodes of the same type, while $d(B_1^i, B_1^j) \geq 2$ if B_1^i, B_1^j are of different types. The coefficients are defined as

- $c(B_1^i, B_1^j) = d(B_1^i, B_1^j) - 2$ if B_1^i, B_1^j are nodes of the same type;
- $c(B_1^i, B_1^j) = d(B_1^i, B_1^j) - 1$ if B_1^i, B_1^j are nodes of different types;
- $c(B_1^i, B_q^0) = 0$ and $c(B_q^0, B_1^i) = 2$ if B_1^i is of type $\mathcal{O}$ and $(B_1^i, B_q^0) \in T$;
- for the remaining edges in T , $c(B_q^i, B_r^j) = c(B_r^j, B_q^i) = 1$;
- otherwise, $c(B_q^i, B_r^j)$ is the shortest path cost in T from B_q^i to B_r^j .

With these coefficients, the honeycomb₀₂ inequality is expressed as in (11).

THEOREM 11. *Honeycomb₀₂ inequalities are valid for $\text{WGRP}(G)$.*

Proof. The proof is similar to that in [6] for the MGRP and is omitted here for the sake of brevity. $\square$

THEOREM 12. *Honeycomb₀₂ inequalities are facet-inducing for $\text{WGRP}(G)$ if the shrunk graph $(\bar{\mathcal{A}}, T_{\bar{\mathcal{A}}})$ is 2-connected.*

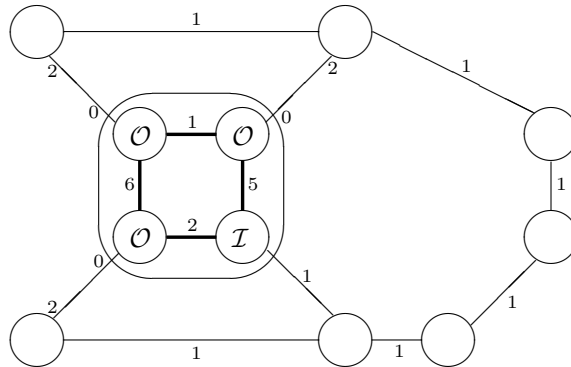


FIG. 8. Honeycomb₀₂ configuration graph with $L = 1$.

Proof. The proof is similar to that of Theorem 10, and the details can be found in [30]. The main difference is the way the tours associated with variables that correspond to traversing an edge $e = (u, v)$ in $\mathcal{C}$ from a node u of type $\mathcal{O}$ to a node v of type $\mathcal{I}$ are defined. If (u, v) is required, we construct a tour traversing each edge in the path in T from u to v twice and some edges in T once in each direction to connect those nodes B_i^0 that have still not been visited. In this case not all of the required edges can be traversed as in x^d , and some of them are redirected to obtain a symmetric graph. If e is not required, the tour is constructed in a similar way. $\square$

4.8. Zigzag inequalities. Zigzag inequalities were introduced in [8]. Here we first describe two families of zigzag inequalities: the even and odd zigzag inequalities. Consider a partition of set V into 4 subsets: M^1 , M^2 , M^3 , and M^4 . Let α_{ij} denote the number of required edges in (M^i, M^j) , and suppose that one of the following conditions is satisfied:

- *Even case:* M^1 , M^2 , M^3 , and M^4 are R -even and $\alpha_{12} = \alpha_{34} = \alpha_{14} = \alpha_{23} = 0$.
- *Odd (simple) case:* M^1 , M^2 , M^3 , and M^4 are R -odd and $\alpha_{12} + \alpha_{34} = \alpha_{14} + \alpha_{23}$.

The configuration graphs $G_{\mathcal{C}}$ associated with even and odd (simple) zigzag inequalities are shown in Figures 9(a) and 9(b) (where the required edges are represented in bold lines) and are defined by the partition of V above and by the following coefficients:

$$c(M^1, M^2) = c(M^3, M^4) = 0, \quad c(M^2, M^1) = c(M^4, M^3) = 2, \quad c(M^i, M^j) = 1 \text{ otherwise.}$$

Then the even and odd (simple) zigzag inequalities can be expressed as

$$(12) \quad x(\delta(M^1 \cup M^2)) + 2x(M^2 : M^1) + 2x(M^4 : M^3) \geq \alpha_{13} + \alpha_{24} + \alpha_{14} + \alpha_{23} + 2,$$

where, given any set of edges (S_1, S_2) , $x(S_1 : S_2) = \sum_{(i,j): i \in S_1, j \in S_2} x_{ij}$.

In the above-mentioned paper it is shown that even and odd (simple) zigzag inequalities (12) are valid and facet-inducing for $\text{WGRP}(G)$ if the following conditions hold: If $G_{\mathcal{C}}$ is a complete graph and $\alpha_{13}, \alpha_{24} \geq 2$ (even case) or $\alpha_{13} \geq |\alpha_{12} - \alpha_{14}| + 1$ and $\alpha_{24} \geq |\alpha_{12} - \alpha_{23}| + 1$ (odd case). All of these inequalities are strong configuration inequalities. Furthermore, it is also proved that odd (simple) zigzag inequalities are equivalent to 3-wheel inequalities for the WPP [34].

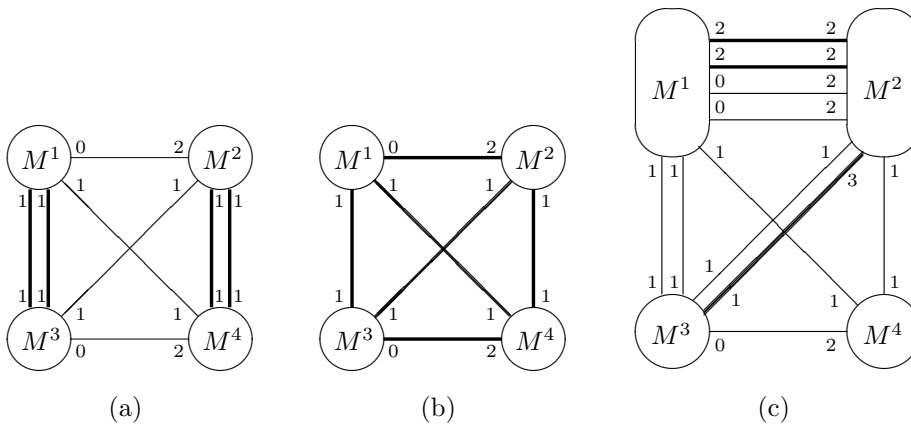


FIG. 9. Zigzag configuration graphs.

Condition $\alpha_{12} + \alpha_{34} = \alpha_{14} + \alpha_{23}$ for the odd (simple) case above is very strong and can be relaxed as follows to obtain a more general version of odd zigzag inequalities. Let M^1, M^2, M^3 , and M^4 be a partition of V into R -odd sets. Let us represent the horizontal edges by $\mathcal{H} = (M^1, M^2) \cup (M^3, M^4)$ and the diagonal ones by $\mathcal{D} = (M^2, M^3) \cup (M^1, M^4)$, and note that $\mathcal{H} \cup \mathcal{D} = \delta(M^1 \cup M^3)$. Let $\mathcal{F} \subset (\mathcal{H} \cup \mathcal{D})_R$ be a subset of required edges (shown as bold lines in Figure 9(c), where all of the edges are required) satisfying

$$(13) \quad |\mathcal{H}_R \setminus \mathcal{F}| + |\mathcal{D}_R \cap \mathcal{F}| = |\mathcal{D}_R \setminus \mathcal{F}| + |\mathcal{H}_R \cap \mathcal{F}|.$$

The more general odd zigzag inequality is

$$(14) \quad \begin{aligned} & x(\delta(M^1 \cup M^2)) + 2x(M^2 : M^1) + 2x(M^4 : M^3) + 2x(F_{zz}) \\ & \geq \alpha_{13} + \alpha_{24} + \alpha_{14} + \alpha_{23} + 2|\mathcal{H} \cap \mathcal{F}| + 2, \end{aligned}$$

where $x(F_{zz})$ denotes the variables associated with the edges in $\mathcal{F}$ in the direction given by the zigzag, i.e., $M^1 \rightarrow M^2 \rightarrow M^3 \rightarrow M^4 \rightarrow M^1$. The pairs of coefficients associated with each edge are shown in Figure 9(c). Odd zigzag inequalities (14) are valid and facet-inducing for $\text{WGRP}(G)$ if $G_C \setminus \mathcal{F}$ is a complete graph, there are two required edges $e_{13} \in (M^1, M^3)$ and $e_{24} \in (M^2, M^4)$, and the remaining required edges in G_C can be oriented to induce a symmetric graph satisfying that all of the edges in $(\mathcal{H} \cup \mathcal{D})_R \setminus \mathcal{F}$ are oriented in the direction of the zigzag and all of the edges in $\mathcal{F}$ are oriented in the opposite direction.

Odd zigzag inequalities (14) are weak configuration inequalities when $\mathcal{F} \neq \emptyset$. If $\mathcal{F} = \emptyset$, they reduce to inequalities (12), and condition (13) becomes $\alpha_{12} + \alpha_{34} = \alpha_{14} + \alpha_{23}$.

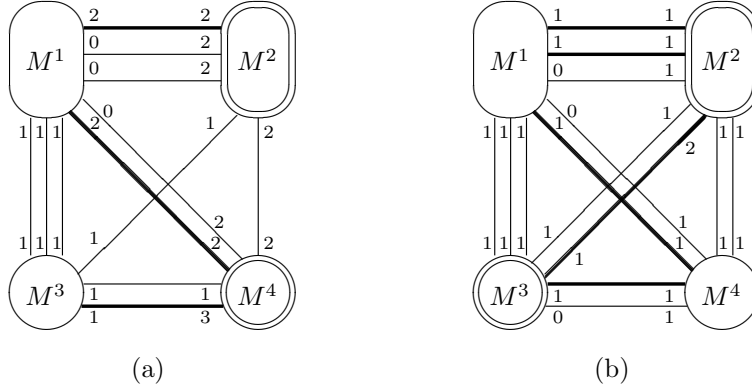
In the odd zigzag inequalities, the direction of the zigzag is $M^1 \rightarrow M^2 \rightarrow M^3 \rightarrow M^4 \rightarrow M^1$ and is determined by the orientation of some edges whose associated variables take fractional values in the linear program (LP) solution. More recently, in [10], two more versions of zigzag inequalities have been described for the WPP, the even-even, and the odd-odd zigzag inequalities. These two kinds of inequalities are designed to cut fractional LP solutions in which some edges in the zigzag have a different orientation, $M^1 \rightarrow M^2 \rightarrow M^3$ and $M^1 \rightarrow M^4 \rightarrow M^3$ for the even-even zigzag inequalities and $M^1 \rightarrow M^2 \rightarrow M^3 \rightarrow M^4$ and $M^1 \rightarrow M^4$ for the odd-odd zigzag inequalities. The name of the inequalities also refers to the degree of the two shores of the edge cutset $\delta(M^1 \cup M^3)$ (see Figures 10(a) and 10(b), where all of the edges are required and a set M^i represented with a double circle is R -odd). These two last types of inequalities are proved to be facet-inducing for the WPP polyhedron under mild conditions. In what follows we describe the generalization of these inequalities to the WGRP.

For the even-even zigzag inequalities, sets M^2 and M^4 are R -odd, sets M^1 and M^3 are R -even, and $\mathcal{F}$ is a subset of required edges $\mathcal{F} \subset \delta_R(M^1 \cup M^3)$ (shown as bold lines in Figure 10(a)), satisfying

$$(15) \quad \begin{aligned} & |(M^1, M^2 \cup M^4)_R \setminus \mathcal{F}| + |(M^2 \cup M^4, M^3)_R \cap \mathcal{F}| \\ & = |(M^1, M^2 \cup M^4)_R \cap \mathcal{F}| + |(M^2 \cup M^4, M^3)_R \setminus \mathcal{F}|. \end{aligned}$$

The corresponding even-even zigzag inequality is

$$(16) \quad \begin{aligned} & x(\delta(M^3)) + 2x(M^2, M^4) + 2x(M^2 : M^1) + 2x(M^4 : M^1) + 2x(F_{zz}) \\ & \geq \alpha_{13} + 2\alpha_{24} + \alpha_{23} + \alpha_{34} + 2|\delta(M^1) \cap \mathcal{F}| + 2, \end{aligned}$$

FIG. 10. Even-even and odd-odd zigzag configurations, with sets $\mathcal{F}$ in bold lines.

where $x(F_{zz})$ denotes the variables associated with the edges in $\mathcal{F}$ in the direction of the zigzag, i.e., $M^1 \rightarrow M^2 \rightarrow M^3$ and $M^1 \rightarrow M^4 \rightarrow M^3$. The coefficients of the variables are also shown in Figure 10(a).

For the odd-odd zigzag inequalities (see Figure 10(b)), sets M^2 and M^3 are R -odd, sets M^1 and M^4 are R -even, set $\mathcal{F} \subset \delta_R(M^1 \cup M^3)$ satisfies the condition

$$(17) \quad \begin{aligned} & |(M^1, M^2 \cup M^4)_R \setminus \mathcal{F}| + |(M^3, M^4)_R \setminus \mathcal{F}| + |(M^2, M^3)_R \cap \mathcal{F}| + 1 \\ & = |(M^1, M^2 \cup M^4)_R \cap \mathcal{F}| + |(M^3, M^4)_R \cap \mathcal{F}| + |(M^2, M^3)_R \setminus \mathcal{F}|, \end{aligned}$$

and the corresponding odd-odd zigzag inequality is

$$(18) \quad \begin{aligned} & x(M^2 \cup M^4 : M^1 \cup M^3) + x(M^1, M^3) + x(M^2, M^4) + x(M^3 : M^2) + x(F_{zz}) \\ & \geq \alpha_{13} + \alpha_{24} + |(M^2, M^3) \setminus \mathcal{F}| + |\mathcal{F}| + 1, \end{aligned}$$

where, again, $x(F_{zz})$ denotes the variables associated with the edges in $\mathcal{F}$ in the direction of the zigzag, i.e., $M^1 \rightarrow M^2 \rightarrow M^3 \rightarrow M^4$ and $M^1 \rightarrow M^4$.

The proofs of the following two theorems are similar to those for the WPP in [10] and are omitted here.

THEOREM 13. *Even-even zigzag inequalities (16) are valid for $WGRP(G)$. They are facet-inducing for $WGRP(G)$ if $G_C \setminus \mathcal{F}$ is a complete graph, there are three edges $e_{13}, f_{13} \in (M^1, M^3)_R$ and $e_{24} \in (M^2, M^4)_R$, and the remaining required edges in G_C can be oriented to induce a (directed) symmetric graph where all of the edges in $\delta_R(M^1 \cup M^3) \setminus \mathcal{F}$ are oriented in the direction of the zigzag and all of the edges in $\mathcal{F}$ are oriented in the opposite direction.*

THEOREM 14. *Odd-odd zigzag inequalities (18) are valid for $WGRP(G)$. They are facet-inducing for $WGRP(G)$ if $G_C \setminus \mathcal{F}$ is a complete graph, there are three edges $e_{13} \in (M^1, M^3)_R$, $e_{24} \in (M^2, M^4)_R$, and $e_{14} \in (M^1, M^4)_R \cap \mathcal{F}$, and the remaining required edges in G_C can be oriented to induce a (directed) symmetric graph where all of the edges in $\delta_R(M^1 \cup M^3) \setminus \mathcal{F}$ are oriented in the direction of the zigzag and all of the edges in $\mathcal{F}$ are oriented in the opposite direction.*

5. The mixed GRP polyhedron. As mentioned in the introduction, the MGRP is the general routing problem defined on a mixed graph $G = (V, E, A)$, where V is the set of vertices, E is the set of edges, and A is the set of arcs, and the MCP is a special case of it.

Nobert and Picard [28] proposed a formulation for the MCPP using only one variable per edge expressing the number of times a given edge is traversed in any direction. Such an approach has been also used for the MGRP in [11] and [6], where a wide polyhedral study is carried out. We call this kind of formulation F1. On the other hand, different formulations using two variables per edge were proposed for the MCPP by Christofides et al. [4] and Ralphs [31]. They will be called F2 formulations. A theoretical and computational comparison of both formulations for ARPs on mixed graphs can be found in [7]. Although formulation F2 is perhaps a more intuitive approach to a problem defined on a mixed graph, polyhedral investigations on it have not been presented up to now. That is the subject of this section, where we will take advantage of the previous WGRP study.

As for the WGRP, we will assume again that the original graph has been transformed to satisfy $V = V_R$. In addition, it can also be assumed that $E \setminus E_R = \emptyset$, because each nonrequired edge can be replaced by a pair of opposite nonrequired arcs [11]. Although this last transformation is not as important for formulation F2 as it is for F1, in what follows we suppose that we are working on a strongly connected graph $G = (V, E, A) := (V_R, E_R, A_R \cup A_{NR})$.

An MGRP instance can be transformed into a WGRP instance in which the cost of traversing the edges in the forbidden directions are set to infinity. Then the variables associated with the forbidden directions can be removed from the WGRP formulation, thus obtaining an MGRP formulation with just one variable associated with each arc and two variables associated with each edge. Hence, formulation F2 for the MGRP is exactly the one presented for the WGRP in section 2 except that some variables do not exist (all of the variables x_{ji} associated with arcs (i, j)). Therefore, constraints (1) and (4) are replaced by

$$(19) \quad x_{ij} + x_{ji} \geq 1 \quad \forall (i, j) \in E (= E_R),$$

$$(20) \quad x_{ij} \geq 1 \quad \forall (i, j) \in A_R,$$

$$(21) \quad x_{ij} \geq 0 \quad \text{otherwise.}$$

Let $\text{MGRP}(G)$ represent the convex hull of all of the feasible solutions of the F2 formulation for the MGRP on $G = (V, E, A)$ (MGRP tours). From the results obtained for the WGRP polyhedron, it can be shown that, if G is a strongly connected graph, the dimension of $\text{MGRP}(G)$ is $2|E| + |A| - |V| + 1$. Furthermore, *traversing* and *trivial* inequalities are facet-inducing for $\text{MGRP}(G)$ if the corresponding link (edge or arc) $e = (i, j)$ satisfies that $G \setminus \{e\}$ is a strongly connected graph. All other facet-inducing inequalities for the $\text{MGRP}(G)$ are *weak configuration inequalities*, and their associated configuration graphs can have parallel variables with different coefficients associated with their edges.

Obviously any valid inequality for the WGRP is also a valid inequality for the MGRP. Furthermore, the facet-inducing inequalities for $\text{WGRP}(G)$ described in this paper are also facet-inducing for $\text{MGRP}(G)$ if the arcs in G have the “appropriate direction,” i.e., if the removed variables do not avoid the existence of all of the tours needed to define a facet. Before presenting the following inequalities, we need some more notation. If $G = (V_R, E_R, A_R \cup A_{NR})$ is a mixed graph and S_1, S_2 are two vertex sets, we denote

$$E(S_1, S_2) = \{e = (i, j) \in E : i \in S_1, j \in S_2 \text{ or } j \in S_1, i \in S_2\},$$

$$A(S_1, S_2) = \{a = (i, j) \in A : i \in S_1, j \in S_2\},$$

$$A_R(S_1, S_2) = \{a = (i, j) \in A_R : i \in S_1, j \in S_2\}.$$

A mixed graph is called *balanced* if, for every $S \subset V$, the difference between the number of arcs leaving S and the number of arcs entering S is less than or equal to the number of edges in $E(S, V \setminus S)$.

It can be shown that the following inequalities are facet-inducing for $\text{MGRP}(G)$:

- Connectivity inequalities (3), if graphs $G(S)$ and $G(V \setminus S)$ are strongly connected.
- R -odd cut inequalities (6), if $G(S)$ and $G(V \setminus S)$ are strongly connected and $\epsilon > |\alpha - \beta|$, where $\epsilon = |E(S, V \setminus S)|$, $\alpha = |A_R(S, V \setminus S)|$, and $\beta = |A_R(V \setminus S, S)|$.
- K-C inequalities (7), if $\epsilon \geq |\alpha - \beta|$, and K-C₀₂ inequalities (8), if $\epsilon \geq |\alpha + 1 - \beta| + 1$, where $\epsilon = |E(M_0, M_K)|$, $\alpha = |A_R(M_0, M_K)|$, and $\beta = |A_R(M_K, M_0)|$. Here we assume that all of the sets $A(M_i, M_{i+1})$, $A(M_{i+1}, M_i)$ are nonempty.
- PB inequalities (9), if $\epsilon \geq |\alpha - \beta|$ for P odd and $\epsilon \geq |\alpha - \beta| + 1$ for P even, and path-bridge₀₂ inequalities (10), if $\epsilon \geq |\alpha + P_{AZ} - \beta - P_{ZA}| + 1$, where P_{AZ} and P_{ZA} are the number of paths of types AZ and ZA , respectively, $\epsilon = |E(A, Z)|$, $\alpha = |A_R(A, Z)|$, and $\beta = |A_R(Z, A)|$. The existence of all of the pairs of opposite (nonrequired) arcs forming the P paths is assumed here.
- Honeycomb inequalities (11), if the subgraph of G_C induced by the required links $E \cup A_R$ is balanced and the graph $(\bar{A}, T_{\bar{A}})$ is 2-connected. Honeycomb₀₂ inequalities if, in addition, for each edge or arc (u, v) , with node u of type $\mathcal{O}$ and node v of type $\mathcal{I}$, the subgraph of G_C induced by the required links plus two extra arcs from u to v is a balanced graph. We assume the existence of all of the pairs of opposite (nonrequired) arcs forming the tree T .
- Zigzag inequalities (12) and (14), if there are 4 links defining a zigzag $M^1 \rightarrow M^2 \rightarrow M^3 \rightarrow M^4 \rightarrow M^1$ and,
 - (a) for the even case, the subgraph of G_C induced by the required links is balanced;
 - (b) for the odd case, there are two required edges $e_{13} \in (M^1, M^3)$ and $e_{24} \in (M^2, M^4)$, and the remaining required edges in G_C can be oriented in such a way that, together with the required arcs in G_C , they induce a symmetric graph in which all of the links in $(\mathcal{H} \cup \mathcal{D})_R \setminus \mathcal{F}$ are oriented in the direction of the zigzag and all of the links in $\mathcal{F}$ are oriented in the opposite direction.
- Even-even zigzag inequalities (16) and odd-odd zigzag inequalities (18), under conditions similar to those above for the odd zigzag inequalities.

Given that the MCP and MRPP are special cases of the MGRP, the above inequalities are also facet-inducing for the polyhedra associated with these problems when they are formulated with two variables.

6. Conclusions. Given a “windy” graph (an undirected graph with two costs associated with each edge, representing the costs of traversing it in each direction) and given a subset of “required” edges and a subset of required vertices, the windy general routing problem consists of finding a minimum cost tour traversing each required edge and visiting each required vertex at least once. Since undirected, directed, and mixed graphs can be modeled with a windy graph, and the GRP generalizes the CPP, RPP, and GTSP, all of the routing problems represented in Figure 1 are special cases of the WGRP. Therefore, most of the theoretical and practical results obtained for the WGRP can be applied to these important problems.

In this paper we have presented a polyhedral study of the WGRP. In addition to describing some basic properties of its associated polyhedron $\text{WGRP}(G)$, we have shown that all of its facet-inducing inequalities, except those equivalent to the trivial or

the traversing ones, are *weak* configuration inequalities. This is a generalization of the (strong) configuration property related to the polyhedra associated with the problems in Figure 1. Then each facet-inducing inequality can be represented by means of a configuration graph, and we have proved that, if a given inequality is facet-inducing for the polyhedron defined on this simple configuration graph, it is also facet-inducing for the whole polyhedron $\text{WGRP}(G)$.

We have also described wide families of facet-inducing inequalities. This partial description of $\text{WGRP}(G)$ has been recently used in [9] to develop a branch and cut algorithm capable of solving WGRP instances of large size and other difficult instances of some of the problems shown in Figure 1. Although slightly different versions of most of these inequalities were already known for other routing problems polyhedra, it is in this paper that they have been generalized for the WGRP and proved to be facet-inducing of $\text{WGRP}(G)$.

On the other hand, a partial description of the polyhedra associated with ARPs defined on mixed graphs whose formulation uses two variables per edge has also been obtained. This polyhedral study had not been done before.

A final contribution of this paper is that it also provides a global view of most of the known facet-inducing inequalities for the polyhedra associated with the problems in Figure 1, since versions of facet-inducing inequalities for other problems with formulations using only one variable per link are very similar.

REFERENCES

- [1] A. A. ASSAD AND B. L. GOLDEN, *Arc routing methods and applications*, in Network Routing, Handbooks Oper. Res. Management Sci. 8, M. O. Ball, T. L. Magnanti, C. L. Monma, and G. L. Nemhauser, eds., North-Holland, Amsterdam, 1995, pp. 375–483.
- [2] E. BENAVENT, A. CARROTTA, A. CORBERÁN, J. M. SANCHIS, AND D. VIGO, *Lower bounds and heuristics for the windy rural postman problem*, European J. Oper. Res., 176 (2007), pp. 855–869.
- [3] S. CHOPRA AND G. RINALDI, *The graphical asymmetric traveling salesman polyhedron: Symmetric inequalities*, SIAM J. Discrete Math., 9 (1996), pp. 602–624.
- [4] N. CHRISTOFIDES, E. BENAVENT, V. CAMPOS, A. CORBERÁN, AND E. MOTA, *An optimal method for the mixed postman problem*, in System Modelling and Optimization, Lecture Notes in Control and Inform. Sci. 59, P. Thoft-Christensen, ed., Springer, Berlin, 1984, pp. 641–649.
- [5] N. CHRISTOFIDES, V. CAMPOS, A. CORBERÁN, AND E. MOTA, *An algorithm for the rural postman problem on a directed graph*, Math. Program. Stud., 26 (1986), pp. 155–166.
- [6] A. CORBERÁN, G. MEJÍA, AND J. M. SANCHIS, *New results on the mixed general routing problem*, Oper. Res., 53 (2005), pp. 363–376.
- [7] A. CORBERÁN, E. MOTA, AND J. M. SANCHIS, *A comparison of two different formulations for arc routing problems on mixed graphs*, Comput. Oper. Res., 33 (2006), pp. 3384–3402.
- [8] A. CORBERÁN, I. PLANA, AND J. M. SANCHIS, *Zigzag inequalities: A new class of facet-inducing inequalities for arc routing problems*, Math. Program., 108 (2006), pp. 79–96.
- [9] A. CORBERÁN, I. PLANA, AND J. M. SANCHIS, *A Branch & Cut algorithm for the Windy General Routing problem*, Networks, 49 (2007), pp. 245–257.
- [10] A. CORBERÁN, I. PLANA, G. REINELT, J. M. SANCHIS, AND D. THEIS, *New Results on the Windy Postman Problem*, Technical report TR05-2007, Department of Statistics and OR, University of Valencia, Valencia, Spain, 2007.
- [11] A. CORBERÁN, A. ROMERO, AND J. M. SANCHIS, *The mixed general routing problem polyhedron*, Math. Program., 96 (2003), pp. 103–137.
- [12] A. CORBERÁN AND J. M. SANCHIS, *A polyhedral approach to the rural postman problem*, European J. Oper. Res., 79 (1994), pp. 95–114.
- [13] A. CORBERÁN AND J. M. SANCHIS, *The general routing problem: Facets from the RPP and GTSP polyhedra*, European J. Oper. Res., 108 (1998), pp. 538–550.
- [14] G. CORNUÉJOLS, J. FONLUP, AND D. NADDEF, *The traveling salesman problem on a graph and some related integer polyhedra*, Math. Programming, 33 (1985), pp. 1–27.
- [15] M. DROR, ED., *ARC Routing: Theory, Solutions, and Applications*, Kluwer Academic Publishers, Norwell, MA, 2000.

- [16] J. EDMONDS AND E. JOHNSON, *Matching, Euler tours and the Chinese postman*, Math. Programming, 5 (1973), pp. 88–124.
- [17] H. A. EISELT, M. GENDREAU, AND G. LAPORTE, *Arc routing problems. 1: The Chinese postman problem*, Oper. Res., 43 (1995), pp. 231–242.
- [18] H. A. EISELT, M. GENDREAU, AND G. LAPORTE, *Arc routing problems. 2: The rural postman problem*, Oper. Res., 43 (1995), pp. 399–414.
- [19] B. FLEISCHMANN, *A cutting-plane procedure for the traveling salesman problem on road networks*, European J. Oper. Res., 21 (1985), pp. 307–317.
- [20] M. GRÖTSCHEL AND Z. WIN, *On the Windy Postman Polyhedron*, Report 75, Schwerpunkt-program der Deutschen Forschungsgemeinschaft, Universität Augsburg, Augsburg, Germany, 1988.
- [21] M. GRÖTSCHEL AND Z. WIN, *A cutting plane algorithm for the windy postman problem*, Math. Programming, 55 (1992), pp. 339–358.
- [22] M. GUAN, *Graphic programming using odd and even points*, Chinese Math., 1 (1962), pp. 273–277.
- [23] M. GUAN, *On the windy postman problem*, Discrete Appl. Math., 9 (1984), pp. 41–46.
- [24] H. GUN, *Polyhedral Structure and Efficient Algorithms for Certain Classes of Directed Rural Postman Problem*, Ph.D. dissertation, University of Maryland at College Park, College Park, MD, 1993.
- [25] A. LETCHFORD, *New inequalities for the general routing problem*, European J. Oper. Res., 96 (1997), pp. 317–322.
- [26] E. MINIEKA, *The Chinese postman problem for mixed networks*, Management Sci., 25 (1979), pp. 643–648.
- [27] D. NADDEF AND G. RINALDI, *The symmetric traveling salesman polytope and its graphical relaxation: Composition of valid inequalities*, Math. Programming, 51 (1991), pp. 359–400.
- [28] Y. NOBERT AND J. C. PICARD, *An optimal algorithm for the mixed Chinese postman problem*, Networks, 27 (1996), pp. 95–108.
- [29] C. ORLOFF, *A fundamental problem in vehicle routing*, Networks, 27 (1974), pp. 95–108.
- [30] I. PLANA, *The Windy General Routing Problem*, Ph.D. dissertation, University of Valencia, Valencia, Spain, 2005.
- [31] T. RALPHS, *On the mixed Chinese postman problem*, Oper. Res. Lett., 14 (1993), pp. 123–127.
- [32] A. ROMERO, *On the Mixed Rural Postman Problem*, Ph.D. dissertation, University of Valencia, Valencia, Spain, 1997.
- [33] J. V. SAVALL, *Polyhedral Results and Approximate Algorithms for the Directed Rural Postman Problem*, Ph.D. dissertation, University of Valencia, Valencia, Spain, 1990.
- [34] Z. WIN, *Contributions to Routing Problems*, Ph.D. dissertation, University of Augsburg, Augsburg, Germany, 1987.

ON COMPLEXITY OF THE SUBPATTERN PROBLEM*

SHLOMO AHAL[†] AND YURI RABINOVICH[‡]

Abstract. We study various computational aspects of the problem of determining whether, given a (fixed) permutation π on k elements and an input permutation σ on $n > k$ elements, π can be *embedded* in σ in an order-preserving manner. Formally, the goal is to determine whether there exists a strictly increasing function f from $[1, k]$ to $[1, n]$ which is order preserving, i.e., f satisfies $\sigma(f(i)) > \sigma(f(j))$ whenever $\pi(i) > \pi(j)$. We call this decision problem the *subpattern problem*. The study falls into two parts. In the first part we develop and analyze an algorithmic paradigm for this problem. We introduce two naturally defined (related) permutation-complexity measures $C(\pi)$ and a somewhat finer $C^T(\pi)$, and, we show that our algorithms run in time $O(n^{1+C(\pi)})$ and $O(n^{2 \cdot C^T(\pi)})$, respectively; i.e., the hardness of the problem crucially depends on the structure of π , as measured by $C(\pi)$ or by $C^T(\pi)$. In the second part of the paper we study the above complexity measures. In particular, we show that in the general case, $C(\pi) \leq 0.47k + o(k)$. Thus, the time complexity of the subpattern problem is at most $O(n^{0.47k+o(k)})$, improving over the trivial $O(n^k)$. Unfortunately, it turns out that for most permutations $C^T(\pi) = \Omega(k)$, and thus, in general, the upper bound on the running time cannot be significantly improved using this approach. Yet, for many natural classes of permutations the complexity of $C(\pi)$ is sublinear in k . To demonstrate this, we study two interesting classes of “linear” permutations and show that their complexity is $C(\pi) = O(\sqrt{k})$. In addition, we study some structural properties of the complexity measures, show that $C^T(\pi) \leq C(\pi) \leq O(\log k) \cdot C^T(\pi)$, and relate $C(\pi)$ and $C^T(\pi)$ to the pathwidth and the treewidth of a certain graph G_π defined by the permutation π .

Key words. permutation, forbidden pattern, recognition, treewidth

AMS subject classifications. 68R15, 68Q25, 05C78

DOI. 10.1137/S0895480104444776

1. Introduction. The question studied in this paper belongs to a vast family of questions dealing with finding a specified substructure within a given (large) structure. Our structures here are permutations, viewed as order types of finite sequences of distinct numbers. The key relation between two such sequences is *embeddability*. We say that π is embeddable in σ if there exists a subsequence σ' of σ (i.e., a sequence obtained by erasing some of σ 's entries) that has the same order type as π . The embeddability of π in σ will be denoted by $\pi \hookrightarrow \sigma$. Formally, we have the following definition.

DEFINITION 1.1. Let σ, π be permutations of fully ordered sets of sizes n and k , respectively ($n \geq k$). A function $f : [1, k] \rightarrow [1, n]$ (mapping indices to indices, not values to values!) will be called a (π, σ) -legal mapping, or alternatively, an embedding of π into σ , if it is

1. strictly increasing, i.e., $f(i) > f(j)$ whenever $i > j$;
2. order preserving, i.e., $\sigma(f(i)) > \sigma(f(j))$ whenever $\pi(i) > \pi(j)$.

For a subset $A \subset [1, k]$, a function $f : A \rightarrow [1, n]$ will be called a partial (π, σ) -legal mapping if the above two conditions hold for all $i, j \in A$.

For example, the permutation $\pi = (213)$ can be embedded in $\sigma = (124536)$, with

*Received by the editors June 29, 2004; accepted for publication (in revised form) August 9, 2007; published electronically March 28, 2008.

<http://www.siam.org/journals/sidma/22-2/44477.html>

[†]Mathematics Department, Ben-Gurion University of the Negev, Beer Sheva, Israel.

[‡]Department of Computer Science, University of Haifa, Haifa 31905, Israel (yuri@cs.haifa.ac.il). This author was supported in part by a grant ISF-247-02-10.5. Part of this work was done while on a sabbatical at the University of Toronto Computer Science Department.

f being, e.g., $f(1) = 3$, $f(2) = 5$, $f(3) = 6$. The corresponding subsequence σ' is (436).

Families of permutations of the type $S_n(\pi) = \{\sigma \in S_n \mid \pi \not\prec \sigma\}$ occur naturally, e.g., in the study of stack-sortable permutations, or, more notably, in the well-known, recently confirmed [9] Stanley–Wilf conjecture, claiming that for a fixed π the size of $S_n(\pi)$ grows at most exponentially in n . The main obstacle for proving this conjecture was a poor understanding of the structure of $S_n(\pi)$. This structure, however, is clearly present and easily detectable for special π 's. This suggests that some (and, in particular, random) permutations are more complex in this respect than the others. The present paper can be viewed as an attempt to quantify this complexity.

We study the following algorithmic problem, to be called the *subpattern (SP) problem*:

Let $\pi \in S_k$ be fixed. Given an input $\sigma \in S_n$, determine whether the relation $\pi \hookrightarrow \sigma$ holds.

To distinguish between π and σ , we shall call the former a *pattern* and the latter an *input* permutation. The SP problem is, of course, polynomially tractable, and the brute-force approach (checking all the subpatterns of σ of size k) gives an $O(n^k)$ upper bound. A closer look soon reveals that the complexity of the SP problem crucially depends on the structure of the pattern π , and while, for instance, for $\pi = (12 \dots k)$ the problem can be solved in time $\approx n \cdot \log k$ (using, e.g., the patience sorting), there is no obvious way to substantially improve upon the brute-force upper bound for a random $\pi \in S_k$. Possible evidence for the hardness of this problem is the following fact from [5]. When π is *not* fixed, but comes as a part of the input, the problem of deciding whether $\pi \hookrightarrow \sigma$ holds is NP-complete. In our SP problem, however, only the exponent of n (asymptotically) matters, while the dependence on k can be exponential and beyond, as k is constant. This is a typical situation in the parameterized complexity theory (see, e.g., [6]).

How then does the complexity of the SP problem depend on π ? The question has been addressed before only for a special class of *separable* permutations (see [5] and the bibliography therein), but, to our best knowledge, the general case has not been studied so far. We introduce two suitable related notions of *complexity* of a pattern π , a $C(\pi)$, and a somewhat finer (and smaller) $C^{\mathbf{T}}(\pi)$, and show that the SP problem for π can be solved in time $O(n^{1+C(\pi)})$, and also in time $O(n^{\alpha_\pi \cdot C^{\mathbf{T}}(\pi)})$, where $1 \leq \alpha_\pi \leq 2$ is a constant depending on π . This is the main algorithmic contribution of the paper.

We go on to study the two complexity measures. The good news is that $C(\pi)$, and hence $C^{\mathbf{T}}(\pi)$, never exceeds $0.47k + o(k)$. Unfortunately, for most $\pi \in S_k$, $C^{\mathbf{T}}(\pi) = \Omega(k)$. However, this is not *always* the case, and many nonrandom classes of patterns have much smaller complexity; e.g., the separable patterns studied in [5] have complexity $C^{\mathbf{T}}(\pi) \leq 4$ (see the remark at the end of section 3.3). To stress this point, we present two other naturally defined classes of patterns π for which $C(\pi) = O(\sqrt{k})$.

We establish some general structural properties of the two permutation-complexity measures. In particular, we show that $C^{\mathbf{T}}(\pi) \leq C(\pi) \leq O(\log k) \cdot C^{\mathbf{T}}(\pi)$. It also turns out that $C(\pi)$ and $C^{\mathbf{T}}(\pi)$ are closely related to the pathwidth and the treewidth of a graph G_π defined by the permutation π . We show also that there exist $\pi, \rho \in S_k$ with $C(\pi), C(\rho) \leq 6$, while $C(\pi \circ \rho) = \Omega(\sqrt{k})$.

While the exact computation of $C(\pi)$ appears to be computationally difficult (it probably requires time exponential in k , and thus is not practical for, say, $k = 30$), an $O(\log k)$ -approximation of $C(\pi)$ can be obtained in time polynomial in k , yielding

a reasonably good practical algorithm for the SP problem with low-complexity π 's. The same holds for $C^{\mathbf{T}}(\pi)$. In fact, due to relation to the treewidth of G_π , it can be approximated even better, within a factor of $O(\sqrt{\log C^{\mathbf{T}}(\pi)})$ (by [7]).

To sum up, in this paper we take the first steps in the study of the asymptotic complexity of the SP problem and its dependence on the properties of the (fixed) pattern $\pi \in S_k$. A natural question arising from the present work is whether the parameterization of π 's by $C^{\mathbf{T}}(\pi)$ is the correct one. Or, phrasing it more concretely, *Can the SP problem be solved in time $O(n^{o(C^{\mathbf{T}}(\pi))})$ for any infinite family of patterns $\{\pi\}$?*

2. The algorithms. We shall be mainly interested here in either finding a (single) embedding $\pi \hookrightarrow \sigma$ or concluding that no such embedding exists. The approach is basically that of dynamic programming. The pattern π will be gradually *exposed*, and at each stage the corresponding tables will be suitably refined.

Consider a chain of subsets of $[1, k]$ (i.e., the indices), $\emptyset = A_0 \subset A_1 \subset \dots \subset A_k = [1, k]$, such that each $A_i \setminus A_{i-1}$ consists of a single element. Setting $\tau(i) = A_i \setminus A_{i-1}$, we get a 1-1 function $\tau : [1, k] \rightarrow [1, k]$, to be called the *order of exposure*, defined by (and defining) the chain of subsets. The restrictions $\pi_i = \pi|_{A_i}$ form a chain of subsequences $\emptyset = \pi_0 \hookrightarrow \pi_1 \hookrightarrow \dots \hookrightarrow \pi_k = \pi$. For example, let $\pi = (35412)$, and consider the chain of subsets

$$\emptyset \subset \{4\} \subset \{2, 4\} \subset \{2, 4, 5\} \subset \{2, 3, 4, 5\} \subset \{1, 2, 3, 4, 5\}.$$

The corresponding exposure order is (42531). The corresponding chain of subsequences is

$$\emptyset \hookrightarrow (1) \hookrightarrow (51) \hookrightarrow (512) \hookrightarrow (5412) \hookrightarrow (35412) = \pi.$$

Roughly speaking, our strategy is to create tables $T_0, T_1, \dots, T_k$, where T_i will store partial (π, σ) -legal mappings $f_i : A_i \rightarrow [1, n]$. Equivalently, each $f_i \in T_i$ describes an embedding $\pi_i \hookrightarrow \sigma$. The last table T_k will contain (π, σ) -legal mappings, if they exist.

In order to obtain T_i , we take the partial (π, σ) -legal mappings $\{f_{i-1}\}$ stored in T_{i-1} , and extend, if possible, each $f_{i-1} \in T_{i-1}$ to a set of (π, σ) -legal mappings $\{f_i\}$ to be stored in T_i . Extending an f_{i-1} means assigning a value to $\tau(i) = A_i \setminus A_{i-1}$.

Storing in T_i all possible partial (π, σ) -legal mappings is infeasible: the size of T_i can get as large as $\Omega(n^i)$ (e.g., for a random σ), and one gains little compared to the brute-force algorithm. The goal then is to condense the content of T_i without incurring an information loss.

Let us first examine the situation in a particularly clear and simple case when $\pi = (123 \dots k)$, and the exposure order τ is also $(123 \dots k)$. Equivalently, $A_i = \{1, 2, \dots, i\}$ for $i = 1, \dots, k$. Clearly, the *only* information about any particular f_i to be used in the subsequent extensions is solely the value of $f_i(i)$. Thus, storing in T_i just the values $\{f_i(i)\}$, a row per value, we get tables of size at most $O(n)$ that succinctly represent all possible embeddings of π_i into σ ; different partial (π, σ) -legal f_i 's may of course have the same representative in T_i . This leads to an algorithm of complexity $O(n^2)$, improving on $O(n^k)$.

For a better understanding of the nature of the gain in the above example, let us return to the general case and have a closer look on the structure of legal extensions $\{f_i\}$ of some f_{i-1} . Since any such f_i must agree with f_{i-1} on A_{i-1} , the only question is where the new index $\tau(i)$ can be mapped. In view of Definition 1.1, the obvious restrictions on the value of $j = f_i(\tau(i))$ are the following.

Monotonicity: Let $p_i^- \in A_{i-1}$ and $p_i^+ \in A_{i-1}$ be, respectively, the largest element in A_i smaller than $\tau(i)$, and the smallest element in A_i larger than $\tau(i)$. Then, in order to maintain the monotonicity property, it must hold that

$$f_{i-1}(p_i^-) < j < f_{i-1}(p_i^+) .$$

Order preservation: Let $x_i \in \pi_{i-1}$ be the largest element smaller than $\pi(\tau(i))$ appearing in the subpattern π_{i-1} , and let $y_i \in \pi_{i-1}$ be the smallest element larger than $\pi(\tau(i))$ appearing in π_{i-1} . Let $q_i^- \in A_{i-1}$ and $q_i^+ \in A_{i-1}$ be the indices of x_i and y_i , respectively, i.e., $\pi(q_i^-) = x_i$ and $\pi(q_i^+) = y_i$. Then, in order to maintain the order preservation property, it must hold that

$$\sigma(f_{i-1}(q_i^-)) < \sigma(j) < \sigma(f_{i-1}(q_i^+)) .$$

Call $\{p_i^-, p_i^+, q_i^-, q_i^+\} \subseteq A_{i-1}$ the *significant* elements for the i th stage. (Note that this set may contain less than four elements; e.g., for $i = 1$, it is empty.)

EXAMPLE. Consider, e.g., the pattern $\pi = (318257946)$ and the exposure order $\tau = (742936518)$. Then $A_5 = \{7, 4, 2, 9, \underline{3}\} = \{2, \underline{3}, 4, 7, 9\}$ and $\pi_5 = (18296)$ (the underlined index/element is the one last exposed). Respectively, $p_5^- = 2$, $p_5^+ = 4$, $x_5 = 6$, $y_5 = 9$, $q_5^- = 9$, $q_5^+ = 7$.

In the simple case when $\pi = (123456789)$ and $\tau = (123456789)$, $p_5^- = q_5^- = 4$, while p_5^+ and q_5^+ do not exist.

PROPOSITION 2.1. Let π, σ, τ be as above. Let f_{i-1} be a partial (π, σ) -legal mapping. Then, the extension f_i of f_{i-1} obtained by setting $f_i(\tau(i)) = j$ is a partial (π, σ) -legal mapping if and only if the above monotonicity and the order preservation restrictions are met. Consequently, the set of all possible values $\{f_i(\tau(i))\}$ for any legal extension of f_{i-1} is fully determined by the values of f_{i-1} on the i th stage significant elements.

Proof. The necessity is obvious. The sufficiency is also rather obvious. Since f_{i-1} is legal, if the extended f_i is not strictly increasing as required by Definition 1.1, then either $f_{i-1}(p^-) \geq j$ or $f_{i-1}(p^+) \leq j$ must hold. Also if f_i is not order preserving, then either $\sigma(f_{i-1}(q^-)) \geq \sigma(j)$ or $\sigma(f_{i-1}(p^+)) \leq \sigma(j)$. $\square$

Next, we define for each $i = 0, 1, \dots, k-1$ the set $U_i \subseteq [1, k]$ of *unforgettable* elements:

$$U_i = \cup_{j=i+1}^k \{p_j^-, p_j^+, q_j^-, q_j^+\} .$$

A corollary to Proposition 2.1 is the following.

COROLLARY 2.2. For any partial (π, σ) -legal mapping $f_i : A_i \rightarrow [1, n]$, the only elements of A_i whose value will ever be significant for the subsequent extensions are $U_i \cap A_i$. In other words, in order to determine whether or not a partial (π, σ) -legal f_i can be extended to an entire (π, σ) -legal mapping, it suffices to know $f_i|_{U_i \cap A_i}$.

Before going on with the formal description of the emerging algorithm, let us have a different, clearer look at the system of sets $\{U_i \cap A_i\}_{i=1, \dots, k}$.

DEFINITION 2.3. Let $\pi \in S_k$ be a pattern. The incidence graph G_π is an undirected multigraph on vertices $[1, k]$ with two types of edges: blue ones and red ones. The blue edges are $E^{\text{blue}}(G_\pi) = \{(i, j) \mid |i - j| = 1\}$. The red edges are $E^{\text{red}}(G_\pi) = \{(i, j) \mid |\pi(i) - \pi(j)| = 1\}$; i.e., G_π is a union of two Hamiltonian paths, blue: $(1 - 2 - 3 - \dots - n)$, and red: $(\pi^{-1}(1) - \pi^{-1}(2) - \pi^{-1}(3) - \dots - \pi^{-1}(n))$. Curiously, G_π is almost identical to the breakpoint graph of π as defined by Bafna and Pevzner [2] for completely different purposes.

DEFINITION 2.4. For $A \subseteq [1, n]$ define $\partial_\pi A \subseteq A$, the boundary of A in G_π , as the set of all elements in A that have a neighbor (in G_π) outside of A . For future use,

we refine $\partial_\pi A$ to $\partial_\pi^{\text{blue}} A$ and $\partial_\pi^{\text{red}} A$ (which may or may not be disjoint) according to the color of the edge between the element and its external neighbor.

PROPOSITION 2.5. $U_i \cap A_i = \partial_\pi A_i$.

Proof. Let $v \in \partial_\pi A_i$; this means that v possesses a neighbor $u \notin A_i$. Let $\tau^{-1}(u) > i$ be the stage when u is first exposed by τ . It is readily checked that v is a significant element for this stage, and therefore $v \in U_i$.

Conversely, assume that $v \in A_i$ does not belong to $\partial_\pi A_i$. This means that $v \pm 1$ (the blue neighbors) and $\pi^{-1}(\pi(v) \pm 1)$ (the red neighbors) have already been exposed or do not exist. But then v cannot possibly become significant at any stage in the future because its exposed neighbors will prevent it. Thus, $v \notin U_i$. $\square$

From now on, we shall use $\partial_\pi A_i$ for $U_i \cap A_i$. Returning to our algorithmic strategy, we follow the minimalistic logic of dynamic programming (preserve only what is essential for the future) and observe that for every $f_i \in T_i$ it suffices to store only the values of f_i on $\partial_\pi A_i$. This leads to the following algorithm for the SP problem. Consider (for now, fixed) exposure order τ . Compute offline all $\{\partial_\pi A_i\}_{i=1, \dots, k-1}$. This takes constant (more precisely, $O(k^2)$) time. The data structure of the main procedure is a sequence of tables $\{T_i\}_{i=0}^k$. The rows of each T_i will contain a numerical field per each element of $\partial_\pi A_i$ and an additional single field for a pointer to a row of T_{i-1} . A row represents a partial (π, σ) -legal f_i on A_i by its values on $\partial_\pi A_i$, and it also keeps a pointer to some $f_{i-1} \in T_{i-1}$ whose legal extension resulted in the present f_i . It is required that no two rows of T_i will be identical with respect to the numerical fields. Observe that T_k has a single row containing only the pointer, and T_0 is empty.

ALG1-SP $_{\pi, \tau}(\sigma)$:

INPUT: $\sigma \in S_n$ - an input permutation;

$\pi \in S_k$ - a pattern; (* fixed *)

$\tau \in S_k$ - an exposure order; (* for now fixed *)

OUTPUT: an embedding of π into σ , or a message “NONE”;

```

set  $T_0 = \text{EMPTY ROW}$ ;
for  $i = 1$  to  $k$  do {
  set  $T_i = \emptyset$ ;
  for each row of  $T_{i-1}$  (* corresponding to some  $f_{i-1}$  *) do
    (* create all legal extensions *) {
      for  $j = 1$  to  $n$  do {
        (* check whether  $j$  is a legitimate value for  $f_i(\tau(i))$  *)
        check whether  $f_{i-1}(p_i^-) < j < f_{i-1}(p_i^+)$ ; (* Monotonicity *)
        check whether  $\sigma(f_{i-1}(q_i^-)) < \sigma(j) < \sigma(f_{i-1}(q_i^+))$ ; (* Order Preservation *)
        if both checks succeeded, do {
          create a new row of  $T_i$ ;
          extend  $f_{i-1}$  by setting  $f(\tau(i)) = j$ ;
          fill the numerical fields with the values of the extended  $f_{i-1}$  on  $\partial_\pi A_i$ ;
          if the new row does not already occur in  $T_i$ ,
            add it to  $T_i$ , and direct its pointer to the current row in  $T_{i-1}$ ;
        }
      }
    }
}
}
if  $T_k$  is empty return “NONE”;
else, using the back pointers, retrieve an actual embedding  $f_k$ , and print it.
```

The correctness of the algorithm follows from Corollary 2.2. The time complexity is governed by $\max_i |T_i|$, the size of the largest table. Indeed, at every stage $i > 2$, $O(n)$ operations are performed per each row of T_{i-1} . In addition, one must take care not to create identical rows in T_i ; this can be achieved by, e.g., an online bucket sort. In total, we have $O(|T_i| + n \cdot |T_{i-1}|)$ operations for stage i . In the end, the reconstruction of the resulting f_k may require an additional time $O(kn)$. Combining all this, and ignoring the additive and multiplicative factors depending solely on k , we get $O(\sum_{i=0}^k n \cdot |T_i|) = O(n \cdot \max_i |T_i|)$. Now, by its definition, $|T_i| \leq \binom{n}{|\partial_\pi A_i|} = O(n^{|\partial_\pi A_i|})$. Setting $C_\tau(\pi) = \max_i |\partial_\pi A_i|$, we conclude that the time complexity of **ALG1-SP** $_{\pi,\tau}(\sigma)$ is $O(n^{1+C_\tau(\pi)})$.

Observe that the exposure order does not have to be fixed and that it makes good sense to choose the best possible τ for the given π . Thus, we have naturally arrived at the suitable notion of permutation complexity as follows.

DEFINITION 2.6. Define $C(\pi)$, the complexity of a permutation $\pi \in S_k$, as

$$C(\pi) = \min_{\tau} C_\tau(\pi).$$

Our discussion so far can be summarized by the following theorem.

THEOREM 2.7. The SP problem for a pattern $\pi \in S_k$ can be solved in time $O(n^{1+C(\pi)})$.

Next, we present a generalization of **ALG1-SP** $_{\pi,\tau}$. The basic idea is simple: instead of extending the partial embeddings by a single element, one can *merge* partial embeddings of (almost) disjoint sets. The following lemma, a generalization of Proposition 2.1, provides the basis for this approach.

LEMMA 2.8. Let A, B be subsets of $[1, k]$ satisfying $A \cap B \subseteq \partial_\pi A \cap \partial_\pi B$. Let also f_A and f_B be partial (π, σ) -legal mappings of A and B , respectively, into $[1, k]$. Denote by $f_{A \cup B} : A \cup B \rightarrow [1, k]$ their common extension, possibly ill-defined (this happens when f_A and f_B disagree on $A \cap B$). Then, $f_{A \cup B}$ is well-defined partial (π, σ) -legal if and only if its restriction to $\partial_\pi A \cup \partial_\pi B$ is such.

Proof. The well-definedness is obvious, as $A \cap B \subseteq \partial_\pi A \cap \partial_\pi B$. Assuming $f_{A \cup B}$ is well defined, we address (π, σ) -legality. The necessity is obvious: any restriction of a partial (π, σ) -legal mapping is legal as well. For the sufficiency part, assume, e.g., that $f_{A \cup B}$ violates the monotonicity requirement of Definition 1.1. It is, however, strictly increasing on A and B separately. Assume that the violation $f_A(i) \geq f_B(j)$ occurs on a pair $i < j$ where, w.l.o.g., $i \in A$ and $j \in B$. Let i_+ be the largest element in A such that $[i, i_+] \subseteq A$, and let j_- be the smallest element in B such that $[j_-, j] \subseteq B$. Clearly, $i_+ \in \partial_\pi^{\text{blue}} A$ and $j_- \in \partial_\pi^{\text{blue}} B$. (Recall that the blue path represents $[1, k]$ in the natural order.) By the assumptions on A, B there are two possibilities: either $[i, i_+] = [j_-, j]$, or $i_+ \leq j_-$, implying $f_A(i_+) > f_B(j_-)$. In both cases we conclude that a violation must occur on $\partial_\pi^{\text{blue}} A \cup \partial_\pi^{\text{blue}} B$.

By the same argument, if $f_{A \cup B}$ violates the order-preservation requirement, the violation must occur on $\partial_\pi^{\text{red}} A \cup \partial_\pi^{\text{red}} B$, as the red path represents $[1, k]$ in the order of increasing value of π . $\square$

A rooted binary tree T directed towards the root, with leaves labeled by elements of $[1, k]$ (a unique element per leaf, all the elements appearing, and possibly more than once), will be called an *exposure tree* of π provided it satisfies the following additional requirement. A vertex $v \in T$ *dominates* another vertex u if there is a directed path from u to v in T . Let $S_v^T \subseteq [1, k]$ denote the set of the labels on the leaves of T dominated by v . In particular, if v is the root of T , then $S_v^T = [1, k]$. The additional requirement on T is that if u_l, u_r are the parents of some $u \in T$, then the sets $S_{u_l}^T, S_{u_r}^T$ must satisfy the requirements of Lemma 2.8. As we shall see in section 3.3, this

requirement turns out to be quite natural. Observe, e.g., that when each label from $[1, k]$ appears exactly once, the requirement is automatically satisfied.

We now present the second algorithm. To avoid confusion about the meaning of the T_v 's, let us denote here (and only here) the exposure tree by Z rather than by T . Given an exposure tree Z , compute first, as a part of preprocessing, all S_v^Z 's and $\partial_\pi S_v^Z$'s. The data structure for the main procedure is a set of tables $\{T_v\}_{v \in Z}$, logically organized in some postorder of Z ; the root is the last. The rows of T_v will contain a numerical field per each element of $\partial_\pi S_v^Z$, and two additional fields for pointers to rows of the tables corresponding to v 's two parents. We require that no two rows of T_v are identical with respect to the numerical fields. As before, the rows of T_v represent partial (π, σ) -legal mappings $f_{S_v^Z} : S_v^Z \rightarrow [1, n]$ by storing their values on $\partial_\pi S_v^Z$. The pointers are used to keep track of the two mappings whose merge (i.e., the common extension) yielded the current $f_{S_v^Z}$.

ALG2-SP $_{\pi, Z}(\sigma)$:

INPUT: $\sigma \in S_n$ - an input permutation;

$\pi \in S_k$ - a pattern; (* fixed *)

Z - a tree as above; (* for now fixed *)

OUTPUT: an embedding of π into σ , or a message "NONE";

process the nodes v of Z in some order consistent with the partial order induced by Z {

if v is a leaf, set $T_v = [1, n]$, and set all pointers to NULL;

else (* $T_v = \text{Merge}(T_u, T_w)$, where u and w are the parents of v *)

for each $r_u \in T_u$, $r_w \in T_w$ (* merge the two rows *) {

unite the domains of the corresponding partial mappings

to get a partial mapping f^* of $\partial_\pi S_u^Z \cup \partial_\pi S_w^Z$;

verify that f^* is well-defined partial (π, σ) -legal;

if yes, create a new row by assigning the numerical fields

the values of f^* on $\partial_\pi(S_u^Z \cup S_w^Z)$;

check whether there is no such row in T_v already;

if not, add it to T_v , and direct its pointers to r_u and r_w ;

}

}

if T_{root} is empty return "NONE";

else, using the back pointers, retrieve an actual embedding f , and print it.

(* this is done by running a DFS from the root according to the back pointers,

and retrieving $f(S_v^T)$ upon blackening v . *)

The correctness of **ALG2-SP $_{\pi, Z}$** follows from Lemma 2.8, as representing each legal $f_{S_v^Z}$ by its values on $\partial_\pi S_v^Z$ preserves all the information about $f_{S_v^Z}$ that will be needed in the future.

The time complexity of the algorithm depends on the implementation of the procedure **Merge**(T_u, T_w). The trivial implementation compares each row in T_u to each row in T_w . Ignoring the additive and multiplicative factors depending solely on k , we see that with this implementation, for inner vertex v of Z with parents u and w , one performs $O(|T_u| \cdot |T_w|)$, i.e., $O(n^{|\partial_\pi S_u^Z| + |\partial_\pi S_w^Z|})$ operations. In addition, the online bucket sorting (needed to ensure the uniqueness of a row) costs $O(n^{|\partial_\pi S_v^Z|})$, which is dominated by the previous term. Thus, the time complexity in this case can be as large as

$$O\left(n^{\max_{v \in Z} |\partial_\pi S_u^Z| + |\partial_\pi S_w^Z|}\right).$$

Let us outline a slightly improved implementation, starting with an explanation of its logic. An uninterested reader can safely skip this part.

For brevity, set $\Delta_z = \partial_\pi S_z^Z$ for $z \in Z$. Consider the legality comparisons performed while creating r from r_u and r_w . There are at most $\min\{|\Delta_u|, |\Delta_w|\}$ comparisons of the type “ $r_u[i] ? r_w[j]$ ” (for monotonicity) and at most the same number of comparisons of the type “ $\sigma(r_u[i]) ? \sigma(r_w[j])$ ” (for order preservation). Since all the relevant indices $i \in \Delta_u$ and $j \in \Delta_w$ are fixed and known in advance, there are two simple (linear) transformations $r_u \mapsto R_u$ and $r_w \mapsto R_w$, such that the corresponding r is legal if and only if R_u strictly dominates R_w on all coordinates. These transformations can be specified in terms of π and σ alone; the specification can be done in advance, prior to starting with the main algorithm. The vectors R_u and R_w have the same length, being the number of legality comparisons as above. It is at most $2 \cdot \min\{|\Delta_u|, |\Delta_w|\} < 2k$. The entries of R_u are $\pm r_u[i]$ or $\pm \sigma(r_u[i])$; the entries of R_w are $\pm r_w[j]$ or $\pm \sigma(r_w[j])$.

Given two sets of vectors $X, Y \subset \mathbb{R}^d$, finding a single pair $x \in X, y \in Y$ such that y strictly dominates x on all coordinates, or deciding that no such pair exists, can be done in time $\tilde{O}(|X| + |Y|) = O((|X| + |Y|) \cdot \log^{d-1}(|X| + |Y|))$, which is much better than the trivial $O(|X| \cdot |Y|)$.¹

Since the table T_v should represent all possible partial (π, σ) -legal mappings of S_v^Z , producing a single such mapping will not help. Observe, however, that in creating a row r of T_v from $r_u \in T_u$ and $r_w \in T_w$, only the numerical values on D_v will be stored in r , while the rest of the values are needed only for the current step legality comparisons. Thus, it makes sense to bucket the rows of T_u according to their value on $\Delta_u \cap \Delta_v$, and the rows T_w according to their value on $\Delta_w \cap \Delta_v$. Instead of comparing the rows of T_u and T_w in a pairwise manner, let us compare them bucketwise. Now, in comparing two such buckets, it is enough to get a *single* legal pair, or decide that there is none. Thus, the above idea applies to bucket-comparison, improving the overall complexity of the stage v from $O(n^{|\Delta_u|+|\Delta_w|})$ to

$$\tilde{O} \left(n^{\max_{v \in T} \{ |\Delta_u| + |\Delta_w \cap \Delta_v|, |\Delta_w| + |\Delta_u \cap \Delta_v| \}} \right).$$

We omit the straightforward calculation.

To simplify the presentation, we define the second natural notion of permutation complexity as follows.

DEFINITION 2.9. *For a permutation $\pi \in S_k$ and a fixed exposure tree T , define $C_T = \max_{v \in T} |\partial_\pi S_v^T|$. The tree-complexity $C^{\mathbf{T}}(\pi)$ of π is the minimum of $C_T(\pi)$ over all T 's.*

The above discussion can be summarized by the following theorem.

THEOREM 2.10. *The SP problem for a pattern $\pi \in S_k$ can be solved in time $O(n^{\alpha_\pi \cdot C^{\mathbf{T}}(\pi)})$, where $1 \leq \alpha_\pi \leq 2$.*

Observe that $C(\pi) \geq C^{\mathbf{T}}(\pi)$. Indeed, let τ be the optimal exposure order, and let T_τ be the following rooted caterpillar tree: the root's left son is a leaf labeled $\tau(k)$, right-left grandson is a leaf labeled $\tau(k-1)$, right-right-left grandson is a leaf labeled $\tau(k-2)$, and so on, with $\tau(1)$ and $\tau(2)$ being brothers at the lowest level. Direct the edges towards the root to obtain an exposure tree T_τ . Clearly, it holds that $C_{T_\tau}(\pi) = C_\tau(\pi)$. Moreover, the running time of **ALG2-SP** $_{\pi, T_\tau}$ with the trivial implementation is precisely $O(n^{1+C(\pi)})$, the optimal running time of the first algorithm.

The complexity measures $C(\pi)$ and $C^{\mathbf{T}}(\pi)$ are natural graph-theoretic parameters

¹Bernard Chazelle, private communication.

of the graph G_π . Their general properties and interrelations, and their relations to standard graph-theoretic parameters, are studied in the next section.

3. Complexity of permutations.

3.1. A general upper bound. The first natural question to ask is how well a fixed exposure order τ can perform for $\pi \in S_k$. The answer is given by the following proposition.

PROPOSITION 3.1. *For the identity exposure order Id , $C_{Id}(\pi) \leq \frac{2}{3} \cdot k + 1$. This is essentially best possible for a fixed exposure order, as for every τ there exists π' such that $C_\tau(\pi') \geq \lfloor \frac{2}{3} \cdot k \rfloor$.*

Proof. Clearly, $|\partial_\pi^{\text{blue}} A_i| = 1$ for all i . For the red boundary, consider any $A \subseteq [1, k]$. Since the red degree of a vertex $v \in V(G_\pi)$ is at most 2, $|\partial_\pi^{\text{red}} A| \leq \min\{2|A|, |A|\} \leq \frac{2}{3} \cdot k$. This establishes the first part of the proposition.

For the second part, consider first $\tau = Id$. Assume for convenience that $k = 3q$. It suffices to find a pattern $\pi_k \in S_{3q}$ such that $\partial_\pi^{\text{red}} A_{2q} = A_{2q}$. A generic construction of such an A is exemplified for $k = 9$: $\pi_9 = (1\ 3\ 4\ 6\ 7\ 9\ 2\ 5\ 8)$; i.e., 2, 5, 8, ... are removed and placed at the end in this order. For an arbitrary τ , observe that exposing the permutation $\rho = \pi \circ \tau^{-1}$ according to exposure order τ yields exactly the same values, and hence the same red boundaries, as in exposing π according to Id . $\square$

REMARKS.

1. *So far we have treated τ as the exposure order, but in fact τ is also a permutation in S_k , and from now it often will be treated as such.*
2. *The complexity $C(\pi_k)$ of the above generic “hard” permutation π_k is just 4 for all k ’s.*

The next natural question is how well a random exposure order behaves. The answer is as follows.

THEOREM 3.2. *For any pattern $\pi \in S_k$, almost all exposure orders τ satisfy $C_\tau(\pi) \leq 0.54 \cdot k + o(k)$.*

Proof. The proving mechanism to be used here and in the following theorem is as follows.

1. Given a pattern $\pi \in S_k$, define a random process $\mathcal{P}$ which produces the exposure order $\tau \in S_k$ by exposing $[1, k]$ an index at a time. The process $\mathcal{P}$ will be described by specifying the conditional distribution according to which the next step is performed.
2. For each $i \leq k$, consider a random variable $X_i = |\partial_\pi A_i|$ defined by τ . Compute its expectation $E[X_i]$, and show that X_i is well concentrated, i.e., that for some function $g(k) = o(k)$ it holds that

$$\Pr[X_i - E[X_i] \geq g(k)] = o(1/k).$$

To show this, a standard martingale technique is used (see, e.g., [1] and [10] for a detailed discussion of it). For each i and $t = 0, 1, \dots, k$ define a random variable Y_i^t as the expected value of X_i conditional on the values of $\tau([1, t])$. Clearly, $E[X_i] = Y_i^0$ and $\{Y_i^t\}_{t=0}^k$ form a martingale, i.e., $E[Y_i^{t+1} | Y_i^t] = Y_i^t$. The next step is to establish the bounded differences property, i.e., show that there exists a constant L such that $|Y_i^{t+1} - Y_i^t| \leq L$ for all i and t . Finally, employing the Azuma inequality, conclude that

$$\Pr[X_i - E[X_i] \geq 2L\sqrt{k \ln k}] \leq 1/k^2 = o(1/k).$$

3. Having achieved this, we conclude that a random permutation τ drawn according to $\mathcal{P}$ almost surely satisfies $C_\tau(\pi) = \max_i X_i \leq \max_i \mathbb{E}[X_i] + o(k)$.

This is all we need.

Let us see how the above strategy works for analyzing a random uniform $\tau \in S_k$. It will be convenient to use G_π and speak of vertices instead of indices. Thus, $\mathcal{P}$ at each stage will expose a random uniformly chosen previously unexposed vertex.

First we estimate $\mathbb{E}[X_i]$. What is the probability that vertex $v \in V(G)$ belongs to $\partial_\pi A_i$? A simple combinatorial argument shows that if v has degree 4, then

$$\Pr[v \in \partial_\pi A_i] = \frac{i}{k} \left[1 - \frac{i-1}{k-1} \cdot \frac{i-2}{k-2} \cdot \frac{i-3}{k-3} \cdot \frac{i-4}{k-4} \right] \leq h\left(\frac{i}{k}\right) + o(1),$$

where h is a real valued function $h(x) = x(1-x^4)$. If the degree of v is less than 4, the above probability only decreases. Thus, by linearity of expectation, $\mathbb{E}[X_i] = k \cdot h(\frac{i}{k}) + o(k)$. To show that X_i are concentrated as required, it suffices to verify that the exposure martingale $\{Y_i^t\}_{i=0}^k$ has the bounded differences property. But this is easy: since the exposure of the vertex $\tau(i)$ can influence no more than four vertices besides itself, $L = |Y_i^{t+1} - Y_i^t| \leq 1 + 4 = 5$.

We conclude that almost surely,

$$\max_i X_i \leq \left(\max_{x \in [0,1]} h(x) \right) \cdot k + o(k) + O(\sqrt{k \ln k}) = 0.54 \cdot k + o(k). \quad \square$$

How can the constant 0.54 of Theorem 3.2 be improved? Clearly, the process $\mathcal{P}$ should be more “sensitive” to the current boundary and not attempt to increase it needlessly. A natural improvement on $\mathcal{P}$ would be to fill up the “holes” (i.e., the unexposed vertices, all of whose neighbors are exposed) upon their creation. Such an operation can only be beneficial for the size of the boundaries, and there is no gain in delaying it. Therefore, the improved $\mathcal{P}$ would at each “long step” randomly expose a vertex and then deterministically fill the holes it created. Call the new process $\mathcal{P}_1$.

THEOREM 3.3. *Almost all exposure orders $\tau \in S_k$ produced by $\mathcal{P}_1$ satisfy $C_\tau(\pi) \leq 0.47 \cdot k + o(k)$.*

Proof (sketch). Define the *closure* of a set $A \subset V(G_\pi)$ as the set obtained from A by filling all its holes, i.e., $A \cup \{u \in \bar{A} \mid \Gamma(u) \subseteq A\}$, where $\Gamma(u)$ is the set of u ’s neighbors in G_π . Notice that filling a hole cannot create a new hole. Let $\{A_i\}$ be the sets produced by $\mathcal{P}$ as before, and let $\{B_i\}$ be their respective closures (some of the B_i ’s may be identical). It is easy to be convinced that the system of sets produced by the “long steps” of process $\mathcal{P}_1$ under the same random choices as in $\mathcal{P}$ is precisely $\{B_i\}$. In addition, there will be intermediate sets created during the “short” steps of $\mathcal{P}_1$, i.e., the random exposure and the subsequent deterministic hole filling. Since the size of the boundary of any intermediate set created during the “long step” $B_i \rightarrow B_{i+1}$ is at most $|\partial_\pi B_i| + 1$, it suffices to analyze $\{\partial_\pi B_i\}$. From here, we follow closely the proof scheme of Theorem 3.2. Let $X_i = |\partial_\pi B_i|$. We start with estimating $\mathbb{E}[X_i]$. What is the probability that vertex $v \in V(G)$ belongs to $\partial_\pi B_i$? Clearly, this happens when $v \in A_i$, while at least one of v ’s neighbors is not in A_i and has at least one neighbor which is not in A_i as well. The probability of this event is determined by the local structure of G_π in the vicinity of v . There are numerous cases to consider, each quite simple in itself. We shall discuss in detail here only the typical (and the worst) case, which is when the 2-neighborhood of v is a 4-regular tree. Then, $\Pr[v \in \partial_\pi B_i] \leq h_1(|A_i|/k) + o(1)$, where $h_1(x)$ is a sum of terms corresponding to

the number of neighbors of v in the compliment of A_i :

$$h_1(x) = x \cdot \sum_{j=1}^4 \binom{4}{j} (1-x)^j x^{4-j} (1-x^{3j}).$$

Since the maximum value of $h_1(x)$ on $[0, 1]$ is less than 0.47, by linearity of expectation we get

$$\max_i \mathbb{E}[X_i] \leq k \left(o(1) + \max_{x \in [0,1]} h_1(x) \right) \leq 0.47 \cdot k + o(k).$$

As before, in order to complete the proof it remains to verify that the exposure martingale $\{Y_i^t\}$ has the bounded differences property. It is easy to verify that an exposure of an extra vertex v at any particular moment may influence at most the 2-ball of v , i.e., at most $1 + 4 + 3 \cdot 4 = 17$ vertices, and thus $|Y_i^{t+1} - Y_i^t| \leq 17$. $\square$

The process $\mathcal{P}_1$ can be further improved. Besides holes filling, there is another beneficial operation: if the exposure of a vertex does not increase the size of the current boundary, there is definitely no damage in exposing it. This defines a new process $\mathcal{P}_2$. The analysis of the latter process presents some difficulties, since now an exposure of a single vertex may cause a cascading effect, destabilizing the process. We shall not attempt to carry out an analysis of $\mathcal{P}_2$ here (although it is not out of hand). Let us just mention that for all (large) π 's we have checked, the numerical simulations show that for the random τ created by $\mathcal{P}_2$, $C_\tau(\pi)$ is bounded by $0.35 \cdot k$.

Before concluding this section, we would like to mention an entirely different approach for finding a good exposure order τ for the given π . Take the Laplacian matrix of G_π , find its eigenvector corresponding to the first positive eigenvalue, and expose the vertices in increasing order of their value in the eigenvector. Numerical simulations seem to indicate that this method is superior to the other methods described so far, yielding a constant of about 0.22. While there are good intuitive reasons why this approach should work (see, e.g., [12] for a related discussion), we have at the moment no rigorous results in this direction.

3.2. General lower bound. Next, we show that there exist patterns with $C^T(\pi) = \Omega(k)$. In fact, almost all permutations $\pi \in S_k$ have this property.

THEOREM 3.4. *Let $\pi \in S_k$ be a random, uniformly chosen permutation. Then there exists a universal constant $c > 0.036$ such that*

$$\Pr [C^T(\pi) \leq c \cdot k] = o(1).$$

Proof. The reason the theorem holds is that for a random π , the graph G_π is almost surely an edge-expander with some universal constant $c_1 > 0$. From here the road is short. Consider any exposure tree T . Since it is binary, there exists a vertex $v \in D$ such that $\frac{2}{3}k \geq |S_v^T| \geq \frac{1}{3}k$. The edge-boundary of S_v^T in G_π must be at least $\frac{c_1}{3} \cdot k$, and hence $|\partial_\pi S_v^T| \geq \frac{c_1}{12} \cdot k$.

Establishing the edge-expansion of a random G_π can be done with “bare hands” or using the available master-tools. Since the latter path is both shorter and more instructive, we shall follow it here. Observe that in the proofs of Theorems 3.2 and 3.3 the only structural property of G_π exploited was that it has degree at most 4. Did we lose much by making this simplification? It turns out that, at least for a random π , nothing was lost! More precisely, the model of a uniformly random 4-regular graph on $[1, n]$ is *contiguous* to the model of a graph obtained by taking the union of two

random independent Hamiltonian cycles on $[1, n]$ (see [13] and Theorem 4.15 there for a clear and detailed discussion of this and related issues). In particular, the expansion ratio of a random graph in the first model is (asymptotically) the same as that in the second model. Clearly, the expansion ratio of a graph obtained by taking one fixed and one uniformly random Hamiltonian path on $[1, n]$, i.e., our G_π for a random π , has this property as well.

Now, by a well-known result of Bollobas [4], the edge-expansion ratio c_1 for a uniformly random 4-regular graph is at least 0.44. Hence, $c \geq 0.036$. $\square$

3.3. The two complexity measures: $C(\pi)$ and $C^T(\pi)$. Since $C(\pi)$ and $C^T(\pi)$ are purely graph-theoretic parameters of a special graph G_π , it is natural to extend their definitions (in an obvious manner) to arbitrary graphs G . Thus we arrive (with a slight abuse of notation) at graph-theoretic parameters $C(G)$ and $C^T(G)$. As we shall see, these parameters are closely related to *pathwidth* and *treewidth* of G , respectively.

The pathwidth of G is defined as follows. An *interval representation* of G is a mapping $v \mapsto I_v$ of vertices v of G to closed intervals $I_v \subseteq \mathbb{R}$ such that whenever $(u, v) \in E(G)$, the corresponding intervals I_v and I_u have a nonempty intersection. The *width* of such a representation is the maximum over $x \in \mathbb{R}$ of $|\{v \in V(G) \mid x \in I_v\}|$. The pathwidth of G , $\text{pw}(G)$, is the minimum possible width over all interval representations of G minus 1. The treewidth $\text{tw}(G)$ is defined similarly, only the vertices of G are mapped to subtrees of an (arbitrary) host tree T . The width is optimized over all subtree-representations and all host trees T .

PROPOSITION 3.5. $C(G) = \text{pw}(G)$.

Proof (sketch). Consider an optimal interval representation of G . Sort the intervals $\{I_v\}$ by their starting points $\{s_v\}$ and expose the vertices according to this order. Clearly, at the time of exposure of a vertex v , the current boundary, i.e., the set of already exposed vertices that have not-yet-exposed neighbors, is a subset of $\{u \mid s_v \in I_u\}$. In particular, the size of the current boundary is at most the width of the representation at s_v . Hence, $C(G) \leq \text{pw}(G) + 1$. To refine the argument, assume that our optimal representation of G is also minimal, i.e., no interval in the set can be shortened without harming the representation. Then, every s_v is also the endpoint t_u of some interval I_u of the representation. Let the maximum width be attained at a point t_u . Observe that the vertex u will not contribute to the boundary at the moment when the last of the v 's with $s_v = t_u$ is exposed, and hence $C(G) \leq \text{pw}(G)$.

For the other direction, consider the optimal exposure order τ . For each $v \in V(G)$, let $s_v = \tau^{-1}(v)$, i.e., the moment when v is exposed by τ , and let t_v be the smallest $j \geq s_v$ such that $v \notin \partial_G A_j$, i.e., the first moment when both v and all its neighbors are exposed. Map $v \mapsto [s_v, t_v] \subset \mathbb{R}$. It is readily checked that this is an interval representation of G . Indeed, if $(v, u) \in E(G)$, and, w.l.o.g., $s_u > s_v$, then $t_v \geq s_u$, since up to the moment s_u the exposed vertex v has a nonexposed neighbor u . Hence $s_u \in [s_v, t_v]$.

The maximum width of this representation is attained, as one can easily check, at a moment t_v for some $v \in G$. At this moment, v does not contribute to the present boundary. Hence, $\text{pw}(G) \leq C(G)$. $\square$

Returning for a moment to permutations, we observe that the complexity measure $C(\pi)$ is equal, within a multiplicative factor in $[\frac{1}{4}, 1]$, to the *minimum cut linear arrangement*, also called the *cutwidth*, of G_π . This is due to the fact that G_π has degree bounded by 4. While finding the best τ apparently requires time exponential in k , a suboptimal $\tilde{\tau}$ can be found in time polynomial in k by using Rao and Richa's [11]

$O(\log(k))$ -factor approximation algorithm for the minimum cut linear arrangement.

Next, we turn to $C^T(G)$ and relate it to $\text{tw}(G)$. We assume for simplicity that G is connected.

PROPOSITION 3.6. $C^T(G) \leq \text{tw}(G) + 1 \leq 2 \cdot C^T(G)$.

Proof (sketch). Consider an optimal subtree-representation of G , and let T be its host tree. Choose an arbitrary vertex $r \in T$, and make it the root of T . For each of the subtrees $\{T_v\}_{v \in V(G)}$ of the representation, let r_v be the vertex in T_v closest to r . Performing a series of simple local modifications without ever increasing the width, one can arrive at a representation of G with the following special properties. The rooted tree T is a binary tree. Each leaf of T belongs to a unique subtree T_v of the representation, and, moreover, the leaves of any subtree T_v of the representation (with a possible exception of r_v) are also leaves of T ; besides, no T_v is a single vertex. Now, assigning to each leaf of T the label v of the unique subtree T_v that contains it, we obtain a candidate for an exposure tree for G . It remains to verify that the additional requirement on exposure trees (as in Lemma 2.8) holds for T . Clearly, it suffices to show that for each $v \in V(G)$, $\{x \in T \mid v \in \partial_G S_x^T\} = V(T_v) - \{r_v\}$.

We claim that unless the representation is redundant, this is indeed the case. Observe that $\{x \in T \mid v \in S_x^T\}$ is precisely T_v plus the path from r_v to r , the root of T . Since the T_v 's constitute a subtree-representation of G in T , it follows that $v \notin \partial_G S_{r_v}^T$. The same applies, of course, to the vertices on the path from r_v to r . Consider now any vertex $x \in T_v$ different from r_v , and assume that $v \notin \partial_G S_x^T$. This means that the part of T_v dominated by x in T has already met all T_u 's such that $(v, u) \in E(G)$. Thus, T_v can be reduced to this subtree without causing any harm. Therefore, taking the most "economical" optimal special subtree representation of G , we conclude that T is indeed an exposure tree for G , and, moreover, it holds that $C^T(G) \leq \text{tw}(G) + 1$.

For the other direction, consider an optimal exposure tree T for G . For $v \in V(G)$, let r_v be a vertex in T such that $v \in S_{r_v}^T$ but $v \notin \partial_G S_{r_v}^T$, while this fails for any ancestor x of r_v (i.e., a vertex dominated by r_v in T). The additional requirement on exposure trees implies that r_v is uniquely defined, and that it dominates all the leaves of T labeled v . Let T_v be a subtree of T rooted at r_v , whose vertices are all $x \in T$ lying on a path from a leaf labeled v to r_v . We claim that $\{T_v\}_{v \in G}$ is a tree representation of G in T . Indeed, let $(v, u) \in E(G)$, and assume that r_u is at least as low (i.e., far from the root) as r_v in T . Since $u \notin \partial_G S_{r_u}^T$, r_u must dominate some leaf l labeled v . Thus, r_u lies on a path from l to r_v , i.e., $r_u \in T_v$.

What is the width of this representation? For $x \in T$ with parents z, y , the width at x is

$$|\partial_G S_x^T| + |\{v \mid x = r_v\}| \leq |\partial_G S_y^T| + |\partial_G S_z^T|.$$

Hence, $\text{tw}(G) + 1 \leq 2 \cdot C_T(G)$. $\square$

Our next goal is to show that $C^T(G)$ cannot be much smaller than $C(G)$, and that $C(G) \leq O(\log k) \cdot C^T(G)$ holds, where $k = |V(G)|$. In fact, this relation between $\text{tw}(G)$ and $\text{pw}(G)$ is well known (see, e.g., Bodlaender et al. [3]). We give here a different, direct and simple proof of this fact, phrased in the terms of this paper. Then, we show that the factor $O(\log k)$ is tight even for G_π 's, which are the incidence graphs of permutations and the primary concern of this paper, by presenting and analyzing an explicit family of permutations $\{\pi^\tau\}$.

THEOREM 3.7. $\log_2 k \cdot C^T(\pi) \geq C(\pi)$.

Proof. To draw a tree T means to specify for each inner vertex v of T the left and the right parents. Each drawing of T defines the exposure order τ : traverse the

leaves of the drawing from left to right, and expose at each step the corresponding label (i.e., vertex of G), unless it has already been exposed. Given a drawing of T , for each inner vertex $v \in T$, in addition to S_v^T define also Q_v^T , the set of all labels exposed by τ prior to getting to the first leaf of T_v , the (complete) subtree of T rooted at v .

Let T be the optimal exposure tree of G . Draw it so that for *each* inner vertex v it holds that $|S_l^T \setminus Q_v^T| \geq |S_r^T \setminus Q_v^T|$, where l and r are, respectively, the left and the right parents of v . We claim that for the resulting τ , $C_\tau(G) \leq \log_2 k \cdot C_T(G)$.

Indeed, consider the first moment when $C_\tau(G)$ is attained, i.e., when the current exposed set A has the maximum boundary, $|\partial_G A| = C_\tau(\pi)$. Let K denote the set of traversed leaves of T . Call a vertex $v \in T$ *maximal* if all the leaves of T_v belong to K , but v 's son (i.e., the neighbor closer to the root) in T does not have this property. Let $\{v_1, v_2, \dots, v_q\}$ be the set of maximal vertices. Clearly, $A = \cup_{i=1}^q S_{v_i}^T$. The exposure order τ on A defines a total order on maximal vertices: for (maximal) v_i and v_j , we say that v_i precedes v_j if the leaves of T_{v_i} are traversed prior to the leaves of T_{v_j} . Assume that the maximal vertices are numbered according to this order. Now, on one hand, it holds that

$$|\partial_G A| \leq |\partial_G S_{v_1}^T| + |\partial_G S_{v_2}^T| + \dots + |\partial_G S_{v_q}^T|,$$

and hence $C_\tau(G) \leq q \cdot C^T(G)$.

On the other hand, q can be upper-bounded as follows. For each v_i , let u_i be its (right) brother. A simple check shows that u_i is well defined, even for $i = q$ (as $A \neq [1, k]$), and that for each $j > i$, v_j must be an ancestor of u_i . Let $n_i = |S_{v_i}^T \setminus Q_{v_i}^T|$ and $m_i = |S_{u_i}^T \setminus Q_{u_i}^T|$. By our construction, $n_i \geq m_i$ for $i = 1, \dots, q$, and, moreover, by the above observation, $m_i \geq n_{i+1} + m_{i+1}$ for $i = 1, \dots, q-1$. Consequently, $m_i \geq 2m_{i+1}$ for $i = 1, \dots, q-1$, and $m_1 \geq 2^{q-1} \cdot m_q$. Since $m_1 \leq k/2$ and $m_q \geq 1$, we conclude that $q \leq \log_2 k$. $\square$

To argue that the bound of Theorem 3.7 is tight for incidence graphs G_π , let us define the following family of patterns $\{\pi^r \in S_{3^r}\}$. For $r = 0$, $\pi^1 = (1)$. For $r > 1$, define $\pi^r = (\pi^{r-1}, \pi^{r-1} + 2 \cdot 3^{r-1}, \pi^{r-1} + 3^{r-1})$, where by $\pi + x$ we mean the pattern obtained by taking π and adding x to each of its entries. Thus, $\pi^1 = (132)$, $\pi^2 = (132798465)$, etc. We claim that for this family the gap of Theorem 3.7 is indeed attained up to a multiplicative factor.

PROPOSITION 3.8. $C^T(\pi^r) \leq 4$, while $C(\pi^r) > r$.

Proof. For the first statement, we recursively construct an exposure tree T^r for π^r with the property that each label appears at a unique leaf of T^r . We need to specify the tree structure and the labels on the leaves. In what follows, we describe a drawing of an unlabeled T^r , i.e., for each node we specify the left and the right sons. The drawing defines a total (left-to-right) order on the leaves; the labels of the leaves, 1 to 3^r , left to right, are assigned accordingly.

For $r = 0$, T^r is just a single node. For $r > 0$, the right son of the root is (isomorphic to) T^{r-1} ; the left son has two sons, each (isomorphic to) T^{r-1} .

An easy check shows that for each vertex $v \in T^r$ as above, the set $S_v^{T^r}$ of labels constitutes an interval in both the blue and red paths, and therefore has boundary of size at most $2 + 2 = 4$. Thus, $C_{T^r}(\pi^r)$ is at most 4.

The second statement is proved by induction. For $r = 0$ it is obviously true. For $r > 0$, we claim that $C_\tau(\pi^r) \geq 1 + C(\pi^{r-1})$ for any τ . Let K_1 , K_2 , and K_3 denote, respectively, the first, second, and third groups of indices in $[1, 3^r]$. By definition of π^r , its restriction to each K_i is equivalent to π^{r-1} and constitutes an interval in both the blue and the red paths. Therefore, the exposure order τ must proceed as follows.

First, at some moment t_1 the relative boundary of the exposed set with respect to some K_x (i.e., the neighbors lying outside this third are ignored) must become as large as $C(\pi^{r-1})$. At some later moment t_2 this happens for another K_y , and, finally, this happens for the remaining K_z . Let us have a closer look at the actual boundary at the moment t_2 . Clearly, $|\partial_{\pi^r} A_{t_2}| = \sum_{j=1}^3 |K_j \cap \partial_{\pi^r} A_{t_2}|$, and each $|K_j \cap \partial_{\pi^r} A_{t_2}|$ is at least as big as the current relative boundary with respect to K_j . Thus, the only possibility for $|\partial_{\pi^r} A_{t_2}|$ not to exceed $C(\pi^{r-1})$ may occur when the contributions of K_x and K_z are null, while the contribution of K_y is $C(\pi^{r-1})$. That is, K_x must be entirely exposed, while K_z must be entirely unexposed. However, a trivial case analysis shows that in this situation, no matter how K_x, K_y , and K_z are situated, K_x must contain an element with a neighbor (either blue or red) in K_z . Thus, even if the favorable possibility does occur, still the total size of the boundary at this moment must be at least $1 + C(\pi^{r-1})$. $\square$

REMARK. The patterns $\{\pi^r\}$ used in the proof of Proposition 3.8 belong to a larger family of separable permutations, characterized (using our terminology) by the existence of a suitable read-once (i.e., each label appears on a unique leaf) exposure tree T , such that all S_v^T constitute a contiguous interval in both the red and the blue paths. Such patterns were treated in [5], and, in particular, it is shown there that for this family of patterns the SP problem can be solved in time $O(n^6)$. The method, specific to this family, uses dynamic programming in a manner somewhat different than, yet reminiscent of **ALG2-SP**. The bound was further improved to $O(n^4)$ in [8]. For comparison, our results imply that the separable π 's have complexity $C^T(\pi) \leq 4$, and the running time of **ALG2-SP** $_{\pi,T}$ with the improved implementation is $O(n^6 \log n)$.

3.4. Special families: Complexity of the linear permutations. In this section we study in detail two special families of patterns and demonstrate that for such patterns $C(\pi_k) = O(\sqrt{k})$. Throughout this section, for technical reasons we treat the patterns of length k and their exposure orders as permutations of $[0, k-1]$ rather than the usual $[1, k]$; e.g., the pattern $\pi = (201)$ denotes the function $\pi(0) = 2$, $\pi(1) = 0$, $\pi(2) = 1$.

The first family is the set of patterns corresponding to invertible linear transformations of the ring $\mathbb{Z}_k$ to itself, i.e., transformations of the form $\pi(x) = ax + b \bmod k$, where $\text{GCD}(a, k) = 1$.

DEFINITION 3.9. For an invertible $a \in \mathbb{Z}_k$, define $R(a, k) \in \mathbb{N}$ as the smallest natural number R such that there exist $x, y \in [-R, R] \setminus \{0\}$ satisfying $ax + y = 0 \bmod k$.

THEOREM 3.10. Let π be of the form $\pi(x) = ax + b \bmod k$, where $\text{GCD}(a, k) = 1$. Then,

$$R(a, k) - 4 \leq C(\pi) \leq 4 \cdot R(a, k) + 4.$$

Proof. It will be convenient to work in $\mathbb{Z}_k$ under a natural identification between $\mathbb{Z}_k$ and $[0, k-1]$. In what follows, all operations (except square root) are mod k , and so are the intervals $[i, j] = [i, i+1, \dots, j]$.

The lower bound follows from the fact that for a linear π the graph G_π is locally a grid. The upper bound is implied by the manner in which these grid patches overlap. We start with the easier lower bound. Observe that $(i, i + a^{-1})$ is always a red edge except when $i = \pi^{-1}(k-1)$. Let $I = [0, r-1]$, where $r = R(a, k)$. Let $I_x = x \cdot a^{-1} + I$. Consider the restriction of G_π to the vertices in $I_0 \cup I_1 \cup I_2 \cup \dots \cup I_{(r-1)}$. The definition of $R(a, k)$ implies that these intervals are disjoint. It is easy to verify that the induced subgraph contains the $r \times r$ grid with blue horizontal and red vertical edges. One vertical and one horizontal edge may be missing. Exposing G_π is at least as hard as

exposing its subgraph. It is a simple (folklore) exercise to show that any set of size about $\lfloor r^2/2 \rfloor$ in the $r \times r$ grid ($r > 1$) has boundary of size at least r . Subtracting 4 in order to compensate the 2 possibly missing edges, we arrive at the lower bound.

For the upper bound, it will be convenient to consider G_π augmented by a blue edge $(k-1, 0)$ and by a red edge $(\pi^{-1}(k-1), \pi^{-1}(0))$. Let $J = [0, r]$ and $J_x = x \cdot a^{-1} + J$. The exposure order is defined as follows. First, expose J_0 (according to the natural order on it), then J_1 , J_2 , and so on. If the current index/vertex has already been exposed previously, skip to the next one. It is clear that at any moment the set of the exposed indices constitutes a union of at most $r+1$ intervals in the red path, and thus the size of the red boundary never exceeds $2(r+1)$. In order to analyze the blue boundary, observe that by definition of $R(a, k)$, the first $r+1$ (blue) intervals J_x cannot be disjoint. Since $J_s \cap J_t \neq \emptyset$ implies $J_{s+1} \cap J_{t+1} \neq \emptyset$, we conclude that for every $t \geq r$, the interval J_t intersects one or more of the previously exposed intervals. This implies that at any moment the exposed set has at most $r+1$ connected components in the blue path, and thus its blue boundary is of size at most $2(r+1)$. $\square$

COROLLARY 3.11. *For any linear permutation π as above, $C(\pi) \leq 4\sqrt{k} + O(1)$. The upper bound is tight; for instance, when $a = \lfloor \sqrt{k} \rfloor$, it holds that $C(\pi) \geq \sqrt{k} + O(1)$.*

Proof. For the upper bound, divide the domain $[0, k-1]$ into $\lfloor \sqrt{k} \rfloor$ disjoint intervals of length at most $\lceil \sqrt{k} \rceil + 2$. By the pigeonhole principle, some two elements of $\{0, a, 2a, \dots, \lceil \sqrt{k} \rceil \cdot a\}$ must fall into the same interval; subtracting one such element from another yields a pair x, y in the interval $[-\lceil \sqrt{k} \rceil - 1, \lceil \sqrt{k} \rceil + 1] \setminus \{0\}$ satisfying $ax + b = 0$. Thus, $R(a, k) \leq \sqrt{k} + 2$.

The lower bound for the special a is obvious. $\square$

The second family of linear permutations that we would like to discuss corresponds to the family of invertible linear transformations $A : \mathbb{Z}_2^m \rightarrow \mathbb{Z}_2^m$. Let $k = 2^m$. The value $\pi(i)$ of a number $i \in [0, k-1]$ is obtained by taking the binary representation of i (with leading zeroes!), interpreting it as a vector $v \in \mathbb{Z}_2^m$, applying A_π to v to get $u = A_\pi(v)$, and interpreting u as the binary presentation of another number in $[0, k-1]$. Formally, let $e_1, \dots, e_m$ be the vectors of the standard basis of $\mathbb{Z}_2^m$. The correspondence between $\mathbb{Z}_2^m$ and $[0, 2^m - 1]$ is given by $\sum_{i=1}^m b_i e_i \mapsto (b_m \dots b_2 b_1)_2$, a number in the binary representation.

Let $V_i = \text{span}\{e_1, e_2, \dots, e_i\}$. Call a linear transformation $A : \mathbb{Z}_2^m \rightarrow \mathbb{Z}_2^m$ *semi-triangular* if for every $i = 1, 2, \dots, m$ it holds that $A(V_i) \supseteq V_{\lfloor i/2 \rfloor}$.

LEMMA 3.12. *Let A be invertible and semitriangular, and let $\pi \in S_k$ be the corresponding permutation. Then, $C_{Id}(\pi) = O(2^{\frac{m}{2}})$.*

Proof. To start with, observe that there is a natural 1-1 correspondence between the affine subspaces of $\mathbb{Z}_2^m$ of the form $(u + V_i)$ and the subintervals of $[0, 2^m - 1]$ of the form $x \cdot 2^i + [0, 2^i - 1]$, where $x \in [0, 2^{m-i} - 1]$. By definition of A , the image $A(p + V_i)$ of any affine subspace of the form $(p + V_i)$ is a disjoint union of affine subspaces of the form $(u + V_{\lfloor i/2 \rfloor})$. Moreover, by a size argument, it is a union of $2^{\lceil \frac{i}{2} \rceil}$ such subspaces. Keeping in mind the correspondence between subspaces and subintervals, we conclude that the image of an interval $J_i \subseteq [0, 2^m - 1]$ of the above form under π consists of at most $2^{\lceil \frac{i}{2} \rceil}$ intervals.

Consider now the image of $[0, t]$ under π . Observe that for every $0 \leq t \leq 2^m - 1$, the interval $[0, t]$ can be (uniquely) represented as $\cup_{i=0}^{m-1} J_i$, where the J_i 's are of the above form. For example,

$$[0, 54] = [0, 31] \cup 32 + [0, 15] \cup 48 + [0, 3] \cup 52 + [0, 1] \cup 54 + [0].$$

Since $[0, t]$ can be represented as $\cup_{i=0}^{m-1} J_i$, we conclude that $\pi([0, t])$ is a union of at most $2 \sum_{i=0}^{m-1} 2^{\frac{i}{2}} = O(2^{\frac{m}{2}})$ intervals. Hence, the red boundary of $[0, t]$ in G_π is at most $O(2^{\frac{m}{2}})$. The blue boundary is 1, and the lemma follows. $\square$

REMARK. A similar argument implies that for a triangular A , i.e., when $A(V_i) = V_i$ for all i , $C_{Id}(\pi) = O(\log n)$. It is also not hard to see that in this case $C^T(\pi) = 3$. The corresponding exposure tree is a complete binary tree of height m , and the leaves are labeled 0 to $2^m - 1$ in this order, starting from the leftmost leaf and ending at the rightmost one.

In order to upper-bound $C(\pi)$ for π originating in a general linear transformation, we need the following two propositions.

PROPOSITION 3.13. For any pair of permutations $\tau, \pi \in S_k$ it holds that

$$C_\tau(\pi) \leq C_{Id}(\tau) + C_{Id}(\pi \circ \tau).$$

Proof. Let us first establish two simple inequalities. The first is

$$(3.1) \quad C_{Id}(\gamma) \geq C_\gamma(Id).$$

Indeed, it is easy to verify that at every step i , a vertex j belongs to the current red boundary in G_γ if and only if $\gamma(j)$ belongs to the current blue boundary in G_{Id} . Since in G_{Id} the blue and the red boundaries are identical, the inequality holds.

Second, we claim that for any $A \subseteq [0, k-1]$, it holds that

$$(3.2) \quad |\partial_\pi A| \leq |\partial_{Id} A| + |\partial_{Id}(\pi(A))|.$$

Indeed, $\partial_{Id} A = \partial_\pi^{\text{blue}} A$, while the second term $\partial_{Id}(\pi(A))$ is, by the previous argument, in a 1-1 correspondence with $\partial_\pi^{\text{red}} A$.

Now we are ready to prove the proposition. Let $A_i = [0, i-1]$. Using (3.2), we conclude that

$$\begin{aligned} C_\tau(\pi) &= \max_i |\partial_\pi(\tau(A_i))| \leq \max_i |\partial_{Id}(\tau(A_i))| + \max_i |\partial_{Id}(\pi \circ \tau(A_i))| \\ &= C_\tau(Id) + C_{\pi \circ \tau}(Id). \end{aligned}$$

Applying (3.1), we conclude the proof. $\square$

In view of Proposition 3.13, our next goal is to show the following fact from linear algebra.

PROPOSITION 3.14. For any invertible $A : \mathbb{Z}_2^m \rightarrow \mathbb{Z}_2^m$ there exists an invertible semitriangular $D : \mathbb{Z}_2^m \rightarrow \mathbb{Z}_2^m$ such that $A \circ D$ is also invertible semitriangular.

Proof. Restating the problem, we need to construct an invertible D such that for every $k \leq \frac{m}{2}$, $V_k \subset D(V_{2k})$, and also $V_k \subset A \circ D(V_{2k})$. Or, combining the two conditions,

$$\begin{aligned} \text{for } k = 1, 2, \dots, \left\lfloor \frac{m}{2} \right\rfloor, \\ \text{span}\{e_1, A^{-1}(e_1), e_2, A^{-1}(e_2), \dots, e_k, A^{-1}(e_k)\} \subseteq D(\text{span}\{e_1, e_2, \dots, e_{2k}\}). \end{aligned}$$

For k as above, define $u_{2k-1} = e_k$ and $u_{2k} = A^{-1}(e_k)$. D is defined as follows. For $i = 1, \dots, m$, set $D(e_i) = u_i$ unless $u_i \in D(\text{span}\{e_1, \dots, e_{i-1}\})$; in the latter case assign to $D(e_i)$ an arbitrary vector not in $D(\text{span}\{e_1, \dots, e_{i-1}\})$. It is a simple matter to verify that D is well defined, and that it indeed satisfies the requirements. $\square$

Combining Propositions 3.13 and 3.14 we conclude the following.

THEOREM 3.15. *For any permutation $\pi \in S_k$ ($k = 2^m$) associated with an invertible linear transformation $A : \mathbb{Z}_2^m \rightarrow \mathbb{Z}_2^m$, it holds that $C(\pi) = O(\sqrt{k})$.*

The upper bound is tight. Consider, e.g., the following simple transformation inverting the binary representation: $\pi^m(b_m b_{m-1} \dots b_1)_2 = (b_1 b_2 \dots b_m)_2$.

PROPOSITION 3.16. $C^T(\pi^m) \geq 2^{\lfloor \frac{m-1}{2} \rfloor}$.

Proof. It is a simple matter to verify that π^m has the following recursive structure. For any natural $r \in [1, m-1]$,

$$\pi^m = (\pi^{m-r}(0) + 2^r \cdot \pi^r, \pi^{m-r}(1) + 2^r \cdot \pi^r, \dots, \pi^{m-r}(2^{m-r} - 1) + 2^r \cdot \pi^r),$$

where the operations of addition and multiplication of a pattern by a scalar are performed entrywise. For example,

$$\begin{aligned} \pi^4 &= (0, 8, 4, 12, 2, 10, 6, 14, 1, 9, 5, 13, 3, 11, 7, 15) \\ &= (0 + 4 \cdot (0213), 2 + 4 \cdot (0213), 1 + 4 \cdot (0213), 3 + 4 \cdot (0213)). \end{aligned}$$

Assume for simplicity $m = 2r$, and let $J_i = i \cdot 2^r + [0, 2^r - 1]$, $i = 0, 1, \dots, 2^r - 1$, denote subintervals of $[0, 2^m - 1]$. For $q \in [0, 2^r - 1]$, let $J_i(q) = i \cdot 2^r + q$, i.e., the q th position in J_i . Observe that $J_a(q)$ and $J_b(q)$ are red neighbors in G_{π^m} if and only if a and b are red neighbors in G_{π^r} . This means that G_{π^m} contains a $2^r \times 2^r$ grid with blue horizontal and red vertical edges. The rows are the J_a 's, and their order of appearance (top to bottom) is as in the red path of G_{π^r} . We had already encountered a similar situation in the course of the proof of Theorem 3.10. Essentially, we are done. A standard technical argument follows.

Let T be any exposure tree, and let $v \in T$ be a vertex such that $S = S_v^T$ intersects at least half of the intervals $\{J_i\}$, but not all of them. Such a vertex must exist. We claim that $|\partial_{\pi^m} S| \geq 2^{r-1}$. Call an interval J_i *white* if $J_i \cap S = \emptyset$, *black* if $J_i \subseteq S$, and *gray* otherwise. By our assumption, there is at least one white interval. Since each gray interval contributes at least 1 to the blue boundary, either there exists a black interval, or the blue boundary is as large as 2^{r-1} , and we are done. Assume we are in the former case. Then, there exists a chain of intervals $\{J_{a_i}\}_{i=0}^t$ such that J_{a_0} is black, J_{a_t} is white, and $\{a_i, a_{i+1}\}$ are red neighbors in G_{π^r} for $i = 0, 1, \dots, t-1$. Let $K_i = J_i \cap S - i \cdot 2^r$. By the previous discussion,

$$|\partial_{\pi^m}^{\text{red}}(S)| \geq |K_0 \setminus K_1| + |K_1 \setminus K_2| + \dots + |K_{t-1} \setminus K_t| \geq |K_0 \setminus K_t| = 2^r. \quad \square$$

3.5. Permutation complexity and operations on strings. In this last section we would like to make a number of observations about the effect of some standard edit and group theoretic operations on the permutation complexity. All arguments below equally apply to $C(\pi)$ and $C^T(\pi)$. To avoid repetition, we shall restrict the discussion to $C(\pi)$.

CLAIM 3.17. *For a subpattern $\pi' = \pi|_S$ of π , it holds that $C(\pi') \leq 2C(\pi)$.*

Proof. The key point is that for any $A \subseteq [1, k]$ it holds that $|\partial_{\pi'}^{\text{blue}}(A \cap S)| \leq |\partial_{\pi}^{\text{blue}}(A)|$ and also that $|\partial_{\pi'}^{\text{red}}(A \cap S)| \leq |\partial_{\pi}^{\text{red}}(A)|$. Let τ' be the exposure order induced on S by the optimal exposure order τ for π . Comparing the boundaries of $\{A_i\}$ exposed by τ with the S -boundaries of the corresponding $\{S \cap A_i\}$ exposed by τ' , we arrive at the desired conclusion. $\square$

Next, let us consider the edit operations *block reversal* and *block move*. Instead of formal definitions, we provide generic examples of these operations:

$$\text{block reversal: } (156 \underline{2497} 83) \rightarrow (156 \underline{7942} 83);$$

$$\text{block move: } (156 \underline{249} \overline{78} 3) \rightarrow (156 \overline{78} \underline{249} 3).$$

Another operation of interest is *split*, $\pi \rightarrow (\pi|_S, \pi|_{\bar{S}})$ for some $S \subseteq [1, k]$. For example, $(\overline{1} \ 5 \ 6 \ 2 \ 4 \ 9 \ 7 \ 8 \ 3) \rightarrow (12473 \ 5698)$. All these operations correlate well with permutation complexity.

CLAIM 3.18. *Let π' be a pattern obtained from π by one of the above operations. Then, $C(\pi')$ is at most $C(\pi) + 2$ for block reversals, at most $C(\pi) + 4$ for block moves, and at most $2C(\pi) + 1$ for splits.*

Proof (sketch). Let τ denote the optimal exposure order for π . The exposure order τ' (in all three cases) is defined by following requirement: $\pi \circ \tau(i) = \pi' \circ \tau'(i)$ for $i = 1, 2, \dots, n$; i.e., at every given moment i , the same *value* is exposed in π and π' . Let A_i and A'_i denote the corresponding stage- i exposed sets. Then, $\pi'(A'_i) = \pi(A_i)$ and $\pi'(\partial_{\pi'}^{\text{red}}(A'_i)) = \pi(\partial_{\pi}^{\text{red}}(A_i))$. It remains to estimate the blue boundaries.

It is a simple matter to verify that the size of $\partial_{\pi'}^{\text{blue}}(A'_i) \setminus \partial_{\pi}^{\text{blue}}(A_i)$ is at most 2 for a block reversal, and at most 4 for a block move, implying the statement for these operations.

For a split $\pi' = (\pi|_S, \pi|_{\bar{S}})$, observe that the exposure orders induced by τ' on S and $\bar{S}$ coincide with those induced on S and $\bar{S}$, respectively, by τ . Arguing as in the proof of Claim 3.17, we conclude that $|\partial_{\pi'}^{\text{blue}}(A'_i)| \leq 2 \cdot |\partial_{\pi}^{\text{blue}}(A_i)| + 1$, and the statement follows. The factor of 2 is tight, since any pattern can be obtained from the identity by performing at most $\lceil \log_2 n \rceil$ splits. $\square$

Next, let us discuss the effect of group-theoretic operations. It clearly holds that $C(\pi) = C(\pi^{-1})$ since the graphs G_{π} and $G_{\pi^{-1}}$ are isomorphic. Viewing Claim 3.18 from a group-theoretic perspective, observe that π' is of the form $\pi \circ \gamma$, where γ is obtained from the identity permutation by a single block reversal, block move, or split, respectively. Thus, in all cases $C(\gamma)$ is low. This makes one wonder whether $C(\pi \circ \gamma)$ can always be upper-bounded in terms of $C(\pi)$ and $C(\gamma)$ alone. Our next (and last) goal is to provide a negative answer to this question.

Let $k = 2r$, and let $\rho_r, \zeta_r \in S_r$. We will be working with permutations of the form $\pi_k = (\rho_r, r + \zeta_r)$. The permutation π'_k is obtained from π_k by an operation of *shuffle* (an inverse operation of split) of a very special kind:

$$(3.3) \quad \pi_k = (\rho_r, r + \zeta_r) \mapsto (\rho_r(1), r + \zeta_r(1), \rho_r(2), r + \zeta_r(2), \dots, \rho_r(r), r + \zeta_r(r)) = \pi'_k.$$

For example, $(\overline{2431} \ 8756) \mapsto (\overline{2} \ 8 \ 4 \ 7 \ 3 \ 5 \ 1 \ 6)$. Thus, $\pi'_k = \pi_k \circ \gamma_k$, where γ_k is of (generic) form $(1 \ 5 \ 2 \ 6 \ 3 \ 7 \ 4 \ 8)$. Clearly, $C_{Id}(\gamma_k) \leq 4$ for all k .

In what follows, we use the notation $C^{\mathbf{T}}(G)$ introduced in section 3.3. Observe that if H is a minor of G , then $\text{tw}(G) \geq \text{tw}(H)$, and hence, by Proposition 3.6, $C^{\mathbf{T}}(G) \geq 0.5 \cdot C^{\mathbf{T}}(H)$.

CLAIM 3.19. *There exists an infinite family of permutations $\{\pi_k \in S_k\}$ as in (3.3), so that $C(\pi_k) \leq 5$, and yet for their shuffles $\{\pi'_k = \pi_k \circ \gamma_k\}$ as in (3.3) it holds that $C^{\mathbf{T}}(\pi'_k) = \Omega(\sqrt{k})$.*

Proof. In what follows, $k = 2r$, and r is a square of a natural number. For π_k as in (3.3) it holds that $C(\pi_k) \leq \max\{C(\rho_r), C(\zeta_r)\} + 2$, as one can first (optimally) expose ρ_r , and then (optimally) expose ζ_r . We shall construct ρ_r and ζ_r of complexity 3; the resulting π_{2r} will thus have complexity ≤ 5 .

To establish a lower bound on $C^{\mathbf{T}}(\pi'_k)$, we use the following line of reasoning. Observe that contracting the blue edges $\{(2i-1, 2i)\}_{i=1}^r$ of $G_{\pi'_k}$ results in a graph isomorphic to $G_{\{\rho_r, \zeta_r\}}$, the graph obtained from G_{ρ_r} and G_{ζ_r} by gluing them together along their blue Hamiltonian paths; i.e., $G_{\{\rho_r, \zeta_r\}}$ is a union of *three* Hamiltonian paths:

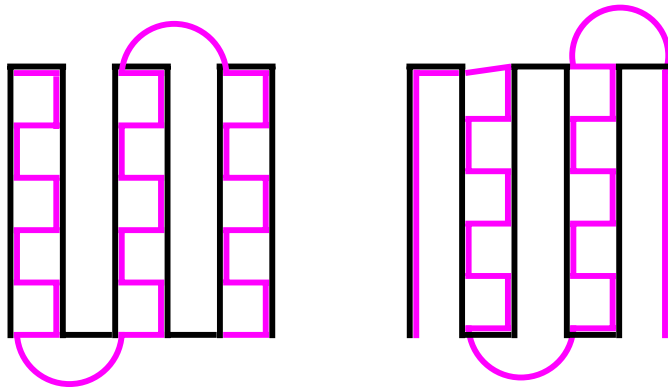


FIG. 3.1. The graphs $G_{\rho_{36}}$ and $G_{\zeta_{36}}$. Black is for blue, and gray is for red. Their union contains a 6×6 grid.

the joint blue one, ρ -red, and ζ -red. This graph, as we shall see, can be considerably more complex than either G_{ρ_r} or G_{ζ_r} . Also, by (rough) monotonicity with respect to taking minors, $C^{\mathbf{T}}(G_{\pi'_k}) \geq 0.5 \cdot C^{\mathbf{T}}(G_{\{\rho_r, \zeta_r\}})$.

It remains to construct the permutations $\{\rho_r, \zeta_r\}$ with the required properties. It is easier to describe such permutations by drawing the corresponding incidence graphs rather than by a formal description. Figure 3.1 provides the generic drawings of G_{ρ_r} and G_{ζ_r} for $r = 6^2$. Clearly, each of the patterns has complexity 3 (e.g., expose according to the red order), and hence, by the previous discussion, $C(\pi_{2r}) \leq 5$. However, $G_{\{\rho_r, \zeta_r\}}$ contains an $\sqrt{r} \times \sqrt{r}$ grid, and thus, as we have seen before (e.g., at the end of the proof of Proposition 3.16), $C^{\mathbf{T}}(G_{\{\rho_r, \zeta_r\}}) = \Omega(\sqrt{r})$. Therefore, $C^{\mathbf{T}}(\pi'_{2r}) = \Omega(\sqrt{r})$. $\square$

Acknowledgments. We are grateful to the anonymous referees for their helpful remarks and criticism.

REFERENCES

- [1] N. ALON, J. H. SPENCER, AND P. ERDOS, *The Probabilistic Method*, Wiley-Interscience, New York, 1991, pp. 83–93.
- [2] V. BAFNA AND P. A. PEVZNER, *Genome rearrangements and sorting by reversals*, SIAM J. Comput., 25 (1996), pp. 272–289.
- [3] H. L. BODLAENDER, J. R. GILBERT, H. HAFSTEINSSON, AND T. KLOKS, *Approximating treewidth, pathwidth, frontsize, and shortest elimination tree*, J. Algorithms, 18 (1995), pp. 238–255.
- [4] B. BOLLOBAS, *The isoperimetric number of random graphs*, European J. Combin., 9 (1988), pp. 241–244.
- [5] P. BOSE, J. F. BUSS, AND A. LUBIW, *Pattern matching for permutations*, in Proceedings of the 3rd Workshop on Algorithms and Data Structures, Lecture Notes in Comput. Sci. 709, Springer, Berlin, 1993, pp. 200–209.
- [6] R. G. DOWNEY AND M. R. FELLOWS, *Parameterized Complexity*, Springer-Verlag, New York, 1999.
- [7] U. FEIGE, M. HAJIAGHAYI, AND J. LEE, *Improved approximation algorithms for minimum-weight vertex separators*, in Proceedings of the 37th Annual ACM Symposium on Theory of Computing, ACM, New York, 2005, pp. 563–572.
- [8] L. IBARRA, *Finding pattern matchings for permutations*, Inform. Process. Lett., 61 (1997), pp. 293–295.
- [9] A. MARCUS AND G. TARDOS, *Excluded permutation matrices and the Stanley-Wilf conjecture*, J. Combin. Theory Ser. A, 107 (2004), pp. 153–160.

- [10] C. McDIARMID, *Concentration*, in Probabilistic Methods for Algorithmic Discrete Mathematics, M. Habib et al., eds., Springer, 1998, pp. 195–248.
- [11] S. RAO AND A. W. RICHA, *New approximation techniques for some ordering problems*, in Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms, ACM, New York, 1998, pp. 211–218.
- [12] D. A. SPIELMAN AND S. TENG, *Spectral partitioning works: Planar graphs and finite element meshes*, in Proceedings of the 37th Annual IEEE Conference on Foundations of Computer Science, IEEE, Piscataway, NJ, 1996, pp. 96–105.
- [13] N. C. WORMALD, *Models of random regular graphs*, in Surveys in Combinatorica, London Math. Soc. Lecture Notes Ser. 267, Cambridge University Press, Cambridge, UK, 1999, pp. 239–298.

CUBIC MONOMIAL BENT FUNCTIONS: A SUBCLASS OF $\mathcal{M}^*$

PASCALE CHARPIN[†] AND GOHAR M. KYUREGHYAN[‡]

Abstract. Based on a computer search, Anne Canteaut conjectured that the exponent $2^{2r} + 2^r + 1$ in $\mathbf{F}_{2^{6r}}$ and the exponent $(2^r + 1)^2$ in $\mathbf{F}_{2^{4r}}$ yield bent monomial functions. These conjectures are proved in [A. Canteaut, P. Charpin, and G. Kyureghyan, *A new class of monomial bent functions*, in Proceedings of the 2006 IEEE International Symposium on Information Theory, (ISIT 06 Seattle), IEEE Press, Piscataway, NJ, 2006, pp. 903–906] and [N. G. Leander, *IEEE Trans. Inform. Theory*, 52 (2006), pp. 738–743]. Both exponents are of binary weight 3 and define functions from the Maiorana–McFarland class $\mathcal{M}$ of bent functions to the subfield. In this paper we show that these are the only such exponents. Our proof is based on the classification of the permutation binomials $X^{2^k+2} + \nu X$ of finite fields of even characteristics. We also extend the result of Leander, determining all bent monomial functions with the exponent $(2^r + 1)^2$.

Key words. cubic bent function, monomial Boolean function, permutation polynomial, Maiorana–McFarland family of bent functions

AMS subject classifications. 11T71, 11T06, 68R01

DOI. 10.1137/060677768

1. Introduction. A bent function is a Boolean function with an even number of variables which has the maximal possible Hamming distance from the set of affine Boolean functions. More precisely, a function $f : \mathbf{F}_2^n \rightarrow \mathbf{F}_2$ (n even) is called bent if

$$\sum_{x \in \mathbf{F}_2^{2t}} (-1)^{f(x) + \langle a, x \rangle} = \pm 2^{\frac{n}{2}}$$

for any $a \in \mathbf{F}_2^n$. Observe that the bent functions are defined in a vector space over $\mathbf{F}_2$ with an inner product. Sometimes it is more convenient to consider bent functions in the finite field $\mathbf{F}_{2^n}$ with the inner product taken to be the absolute trace function. Working in the finite field can be advantageous because of its additional multiplicative structure. The classification of the bent functions in finite fields and vector spaces is completely equivalent. Let us restate the definition of a bent function in $\mathbf{F}_{2^n}$, where n is even. A function $f : \mathbf{F}_{2^n} \rightarrow \mathbf{F}_2$ is called bent if

$$\sum_{x \in \mathbf{F}_{2^n}} (-1)^{f(x) + T_1^n(\alpha x)} = \pm 2^{\frac{n}{2}}$$

for any $\alpha \in \mathbf{F}_{2^n}$, where T_1^n is the absolute trace over $\mathbf{F}_{2^n}$.

Several recent papers are devoted to the study of bent functions which are *monomial*, i.e., which are of the form

$$x \in \mathbf{F}_{2^n} \mapsto f(x) = T_1^n(\lambda x^d) \text{ for some } \lambda \in \mathbf{F}_{2^n},$$

*Received by the editors December 15, 2006; accepted for publication (in revised form) September 4, 2007; published electronically March 28, 2008. An extended abstract of this paper was published in the proceedings of the workshop “Algebraic and Combinatorial Coding Theory-10,” Zvenigorod, Russia, Moscow, IITP RAS, 2006, pp. 52–55.

<http://www.siam.org/journals/sidma/22-2/67776.html>

[†]INRIA projet CODES, B.P. 105, 78153 Le Chesnay Cedex, France (Pascale.Charpin@inria.fr).

[‡]Department of Mathematics, Otto-von-Guericke University of Magdeburg, Universitätplatz 2, 39106 Magdeburg, Germany (Gohar.Kyureghyan@Mathematik.Uni-Magdeburg.de).

where d is said to be the *exponent of f* . One of the goals in this research is to find exponents d and the field elements λ defining a bent function. Also, some properties of bent functions are observed by studying monomial ones. For instance, the only known so-called nonnormal bent functions are constructed with monomial bent functions [3]. A further example is

$$\{ f_\lambda : \mathbf{F}_{2^{6r}} \rightarrow \mathbf{F}_2 \mid f_\lambda(x) = T_1^{6r}(\lambda x^{2^{2r}+2^r+1}), \lambda \in \mathbf{F}_{2^{3r}}^* \text{ and } T_r^{3r}(\lambda) = 0 \},$$

which is closed under the addition of different elements and consists of bent functions of degree 3 (see [1, 2]). To our knowledge, this is the first example of such a family of nonquadratic bent functions.

Based on a computer search carried out for $n \leq 24$, Anne Canteaut conjectured that the exponent $2^{2r} + 2^r + 1$ in $\mathbf{F}_{2^{6r}}$ and the exponent $2^{2r} + 2^{r+1} + 1$ in $\mathbf{F}_{2^{4r}}$ yield bent monomial functions. These conjectures are proved in [1, 10]. Both exponents are of binary weight 3 and define bent functions from the Maiorana–McFarland class $\mathcal{M}$ of bent functions. More precisely, the obtained bent functions are concatenations of affine Boolean functions of the subfield $\mathbf{F}_{2^{\frac{n}{2}}}$ or shifts of such functions. Using the techniques introduced in [1], this paper continues the study of cubic monomial functions. After introducing some preliminary definitions and results in section 2, in section 3 the cubic monomial functions, which can be represented as a concatenation of affine Boolean functions of the subfield $\mathbf{F}_{2^{\frac{n}{2}}}$, are classified. Furthermore, it is shown that among these functions only the ones with the exponents $2^{\frac{n}{2}+2^j+1}$ may define new bent functions. In section 4 the Walsh spectrum of the corresponding monomial functions are studied. It is shown that only exponent $2^{2r} + 2^{r+1} + 1$ for $n = 4r$ of that type defines a bent function. This uses the classification of all permutation polynomials $X^{2^k+2} + \nu X$ over a finite field of even characteristics obtained in section 5. Finally, the result of [10] is extended by determining all bent monomial functions with the exponent $2^{2r} + 2^{r+1} + 1$. In particular, all presently known (found by computer search) monomial bent functions are proved to belong to an infinite family of bent monomial functions. These families are as follows:

- $d = 2^i + 1$ and $\lambda \notin \{y^d \mid y \in \mathbb{F}_{2^n}\}$ (folklore).
- $d = (2^{\frac{n}{2}} - 1)\ell$, $\gcd(\ell, 2^{\frac{n}{2}} + 1) = 1$, and λ corresponds to a zero of the Kloosterman sum (Dillon [6], Lachaud and Wolfmann [9]).
- If n is not a multiple of 3 and $(i, n) = 1$, then $d = 2^{2i} - 2^i + 1$ and $\lambda \notin \{y^3 \mid y \in \mathbb{F}_{2^n}\}$ (Dillon and Dobbertin [7]).
- If $n = 4r$ and r is odd, then $d = 2^{2r} + 2^{r+1} + 1$ (Leander [10]) and $\lambda = \lambda' a^{2^{2r}+2^{r+1}+1}$, where $\lambda' \in \omega \mathbf{F}_{2^r}$, $\omega \in \mathbf{F}_4 \setminus \mathbf{F}_2$, and $a \in \mathbf{F}_{2^{4r}}$ (this paper).
- If $n = 6r$, then $d = 2^{2r} + 2^r + 1$ and $\lambda = \lambda' a^{2^{2r}+2^r+1}$, where $a \in \mathbf{F}_{2^{6r}}$ and $\lambda' \in \mathbf{F}_{2^{3r}}$ such that $T_r^{3r}(\lambda') = 0$ (Canteaut, Charpin, and Kyureghyan [1]).

2. Preliminaries. Let $\mathbf{F}_{2^n}$ be the finite field with 2^n elements. If $n = ms$, then T_m^n is the trace function from $\mathbf{F}_{2^n}$ onto $\mathbf{F}_{2^m}$ given by

$$T_m^n(x) = x + x^{2^m} + \cdots + x^{2^{m(s-1)}} \quad \text{for any } x \in \mathbf{F}_{2^n}.$$

Note that T_m^n is an $\mathbf{F}_{2^m}$ -linear function.

2.1. Boolean functions. Boolean functions on $\mathbf{F}_{2^n}$ are given by $T_1^n(p(x))$, where $p(x)$ is a polynomial over $\mathbf{F}_{2^n}$. Further, we will use the notation

$$\mathcal{F}(f) = \sum_{x \in \mathbf{F}_{2^n}} (-1)^{f(x)}$$

for a given Boolean function f . The weight of f , denoted $wt(f)$, is the number of x such that $f(x) = 1$ and it holds that

$$\mathcal{F}(f) = 2^n - 2wt(f).$$

Recall that f is said to be *balanced* when $wt(f) = 2^{n-1}$ or, equivalently, $\mathcal{F}(f) = 0$. For a given $u \in \mathbf{F}_{2^n}$, a linear Boolean function φ_u is defined by

$$\varphi_u(x) = T_1^n(ux) \text{ for any } x \in \mathbf{F}_{2^n},$$

and every linear Boolean function is of such form. The *Walsh spectrum* of f is the multiset

$$\{ \mathcal{F}(f + \varphi_u) \mid u \in \mathbf{F}_{2^n} \}.$$

Thus a Boolean function is bent if and only if its Walsh spectrum contains only values $\pm 2^{n/2}$. In particular, the weight of a bent function is $2^{n-1} \pm 2^{\frac{n-2}{2}}$.

2.2. The Maiorana–McFarland bent functions. The Maiorana–McFarland class of bent functions was introduced in [12] and extensively studied by Dillon [6, pp. 90–95]. It is usually called the *class* $\mathcal{M}$ of bent functions. The Maiorana–McFarland bent functions on $\mathbf{F}_{2^n}$, $n = 2t$, are the concatenation of 2^t different affine functions, defined on some fixed subspace of dimension t of $\mathbf{F}_{2^{2t}}$. We are interested in the subclass of $\mathcal{M}$, where the fixed subspace is the subfield $\mathbf{F}_{2^t}$. This subclass we denote by $\mathcal{M}(\mathbf{F}_{2^t})$ and define more accurately in Corollary 2.2.

PROPOSITION 2.1. *Let W be a subspace of $\mathbf{F}_{2^{2t}}$ such that $\mathbf{F}_{2^{2t}} = \mathbf{F}_{2^t} \oplus W$. Assume that a Boolean function $f : \mathbf{F}_{2^{2t}} \rightarrow \mathbf{F}_2$ can be expressed in the form*

$$(2.1) \quad f(x) = f(y + a) = T_1^t(y\pi(a) + h(a)),$$

where $x = y + a$, with $y \in \mathbf{F}_{2^t}$ and $a \in W$, and $\pi, h : W \rightarrow \mathbf{F}_{2^t}$. Let us define, for any $u \in \mathbf{F}_{2^{2t}}$,

$$W_u = \{ a \in W \mid \pi(a) = u + u^{2^t} \}.$$

Then

$$\mathcal{F}(f + \varphi_u) = \begin{cases} 0 & \text{if } W_u = \emptyset \\ 2^t \sum_{a \in W_u} (-1)^{T_1^t(h(a) + T_t^{2t}(ua))} & \text{otherwise.} \end{cases}$$

Proof. For any $u \in \mathbf{F}_{2^{2t}}$ it holds that

$$\begin{aligned} \mathcal{F}(f + \varphi_u) &= \sum_{x \in \mathbf{F}_{2^{2t}}} (-1)^{f(x) + T_1^{2t}(ux)} \\ &= \sum_{a \in W} \sum_{y \in \mathbf{F}_{2^t}} (-1)^{f(y+a) + T_1^{2t}(u(y+a))} \\ &= \sum_{a \in W} \sum_{y \in \mathbf{F}_{2^t}} (-1)^{T_1^t(y\pi(a) + h(a) + T_t^{2t}(u(y+a)))} \\ &= \sum_{a \in W} ((-1)^{T_1^t(h(a) + T_t^{2t}(ua))} \sum_{y \in \mathbf{F}_{2^t}} (-1)^{T_1^t(y(\pi(a) + T_t^{2t}(u)))}). \end{aligned}$$

The second sum over $\mathbf{F}_{2^t}$ is 0 unless $\pi(a) = u + u^{2^t}$; in this case it is equal to 2^t . Thus, this sum is 0 for all a if and only if W_u is empty. Otherwise we get

$$\mathcal{F}(f + \varphi_u) = 2^t \sum_{a \in W_u} (-1)^{T_1^t(h(a) + ua + (ua)^{2^t})},$$

completing the proof. $\square$

COROLLARY 2.2. *Let $W \subset \mathbf{F}_{2^{2t}}$ be such that $\mathbf{F}_{2^{2t}} = \mathbf{F}_{2^t} \oplus W$. Assume that a Boolean function $f : \mathbf{F}_{2^{2t}} \rightarrow \mathbf{F}_2$ can be expressed in the form*

$$f(x) = f(y + a) = T_1^t(y\pi(a) + h(a)),$$

where $x = y + a$, with $y \in \mathbf{F}_{2^t}$ and $a \in W$, and $\pi, h : W \rightarrow \mathbf{F}_{2^t}$. Then f is bent if and only if π is bijective. Such functions are said to be the elements of the subclass $\mathcal{M}(\mathbf{F}_{2^t})$ of $\mathcal{M}$.

Proof. Suppose that π is not a permutation. Then there is an element $b \in \mathbf{F}_{2^t}$ such that $\pi(a) \neq b$ for any $a \in W$. For every $u \in \mathbf{F}_{2^{2t}}$ with $u + u^{2^t} = b$ it holds that $W_u = \emptyset$, and thus Proposition 2.1 implies $\mathcal{F}(f + \varphi_u) = 0$, proving the necessity of the condition. To derive the sufficiency, note that in the case of bijective π the cardinality of W_u equals 1 for any $u \in \mathbf{F}_{2^{2t}}$. $\square$

3. Decomposable cubic monomial Boolean functions. In [1], it is shown that the monomial Boolean functions on $\mathbf{F}_{2^n}$ defined by the new bent exponents $2^{2r} + 2^r + 1$ (for $n = 6r$) and $(2^r + 1)^2$ (for $n = 4r$), can be decomposed in form (2.1). In this section we describe all cubic monomial Boolean functions possessing such a decomposition.

Let $0 \leq k \leq 2^n - 2$. We denote by C_k the cyclotomic coset modulo $2^n - 1$ containing k , i.e.,

$$C_k = \{k, 2k, \dots, 2^{n-1}k\} \pmod{2^n - 1}.$$

Recall that if $|C_k| = l$, then $\{x^k : x \in \mathbf{F}_{2^n}\} \subset \mathbf{F}_{2^l}$ and $\mathbf{F}_{2^l}$ is the smallest subfield of $\mathbf{F}_{2^n}$ with this property.

Let $d = 2^i + 2^j + 1$ and $\lambda \in \mathbf{F}_{2^{2t}}$. We consider the function $f : x \mapsto T_1^{2t}(\lambda x^d)$ on $\mathbf{F}_{2^{2t}}$. Taking the smallest representative of C_d , we may assume that

$$(3.1) \quad d = 2^i + 2^j + 1 \quad \text{with} \quad 0 < j < t, \quad i - j < t \quad \text{and} \quad j < i.$$

If $\mathbf{F}_{2^{2t}} = \mathbf{F}_{2^t} \oplus W$, then setting $x = y + a$ ($y \in \mathbf{F}_{2^t}$ and $a \in W$), we have

$$\begin{aligned} T_1^{2t}(\lambda x^d) &= T_1^{2t}(\lambda(y + a)^{2^i + 2^j + 1}) \\ &= T_1^{2t}(\lambda y^{2^i + 2^j + 1}) \\ &\quad + T_1^{2t}(\lambda a^{2^i} y^{2^j + 1} + \lambda a^{2^j} y^{2^i + 1} + \lambda a y^{2^i + 2^j}) \\ &\quad + T_1^{2t}(\lambda y^{2^i} a^{2^j + 1} + \lambda y^{2^j} a^{2^i + 1} + \lambda y a^{2^i + 2^j}) \\ &\quad + T_1^{2t}(\lambda a^{2^i + 2^j + 1}). \end{aligned}$$

Using the transitivity of the trace function we get that $T_1^{2t}(\lambda x^d)$ is equal to the

following sum:

$$\begin{aligned}
 (3.2) \quad & T_1^t \left(T_t^{2t}(\lambda) y^{2^i+2^j+1} \right) \\
 (3.3) \quad & + T_1^t \left(T_t^{2t}(\lambda a^{2^i}) y^{2^j+1} + T_t^{2t}(\lambda a^{2^j}) y^{2^i+1} + T_t^{2t}(\lambda a) y^{2^i+2^j} \right) \\
 & + T_1^t \left(T_t^{2t}(\lambda a^{2^j+1}) y^{2^i} + T_t^{2t}(\lambda a^{2^i+1}) y^{2^j} + T_t^{2t}(\lambda a^{2^i+2^j}) y \right) \\
 & + T_1^{2t} \left(\lambda a^{2^i+2^j+1} \right).
 \end{aligned}$$

Our goal is to describe the set of pairs (i, j) defining a function $T_1^{2t}(\lambda x^{2^i+2^j+1})$ of form (2.1). Equivalently, we are going to find i, j , and λ such that (3.2) and (3.3) are constant or linear on y for any $a \in W$. The following easy result will be used several times in what follows. We present a brief proof of it for the convenience of the reader.

LEMMA 3.1. *Consider the function on $\mathbf{F}_{2^n}$ given by $f(x) = T_1^n(\delta x^d)$, $\delta \in \mathbf{F}_{2^n}^*$. Assume that d is the smallest element of its cyclotomic coset.*

- (a) *Let $d = 2^i + 2^j + 1$. Then the function f is constantly c , $c \in \mathbf{F}_2$, if and only if $c = 0$ and*

$$n = 3\ell, \quad d = 2^{2\ell} + 2^\ell + 1, \quad \text{and} \quad \delta + \delta^{2^\ell} + \delta^{2^{2\ell}} = 0.$$

- (b) *Let $d = 2^i + 1$. Then the function f is constantly c , $c \in \mathbf{F}_2$, if and only if $c = 0$ and*

$$n = 2\ell, \quad i = \ell \quad \text{and} \quad \delta \in \mathbf{F}_{2^\ell}.$$

Proof. Assume that $f(x) = c$ for all x . Then $c = 0$ since $f(0) = 0$. So the polynomial

$$T_1^n(\delta x^d) = \delta x^d + (\delta x^d)^2 + \cdots + (\delta x^d)^{2^{n-1}}$$

is the null polynomial since it has degree less than 2^n . Hence the cyclotomic coset C_d has cardinality ℓ with $\ell < n$ and $n = k\ell$. Then $\{x^k : x \in \mathbf{F}_{2^n}\} \subset \mathbf{F}_{2^\ell}$ and we have

$$T_1^n(\delta x^d) = T_1^\ell(x^d T_1^n(\delta)).$$

To complete the proof we find the exponents d of binary weights 3 and 2 satisfying $2^\ell d \equiv d \pmod{2^n - 1}$.

- (a) $2^\ell(2^i + 2^j + 1) \equiv 2^i + 2^j + 1 \pmod{2^n - 1}$ if and only if $n - i = i - j = j = \ell$. Thus $n = 3\ell$, $i = 2\ell$, and $j = \ell$. In this case, to have $T_1^n(\delta x^d)$ constantly zero it must also hold that $\delta + \delta^{2^\ell} + \delta^{2^{2\ell}} = 0$.
- (b) $2^\ell(2^i + 1) \equiv 2^i + 1 \pmod{2^n - 1}$ if and only if $n - i = i = \ell$. Then $n = 2\ell$ and, moreover, $\delta + \delta^{2^\ell} = 0$. $\square$

We will also need the following observation.

LEMMA 3.2. *Let $n = 6j$ and $d = 2^{4j} + 2^{2j} + 1$. Let f_λ be the function on $\mathbf{F}_{2^n}$ defined by $f_\lambda(x) = T_1^n(\lambda x^d)$, $\lambda \in \mathbf{F}_{2^n}$. Then, f_λ is the null function when $T_{2^j}^{6j}(\lambda) = 0$. Otherwise the weight of f_λ is equal to $2^{2j-1}(2^{4j} + 2^{2j} + 1)$. In particular, f_λ cannot be bent.*

Proof. Recall that the weight of f_λ , say $wt(f_\lambda)$, is the number of x such that $f_\lambda(x) = 1$. Since $(2^{2j} - 1)d = 2^{6j} - 1$, the image set of $x \mapsto x^d$ is equal to $\mathbf{F}_{2^{2j}}$ and every nonzero element from $\mathbf{F}_{2^{2j}}$ has $2^{4j} + 2^{2j} + 1$ many preimages. Using

$$T_1^{6j}(\lambda x^{2^{4j}+2^{2j}+1}) = T_1^{2j} \left(T_{2^j}^{6j}(\lambda) x^{2^{4j}+2^{2j}+1} \right)$$

it is clear that $wt(f_\lambda) = 0$ when λ satisfies $T_{2^j}^{6j}(\lambda) = 0$ (see also Lemma 3.1). Otherwise

$$wt(f_\lambda) = (2^{4j} + 2^{2j} + 1) \# \{ y \in \mathbf{F}_{2^{2j}} \mid T_1^{2j}(T_{2^j}^{6j}(\lambda)y) = 1 \},$$

where $\#B$ denotes the cardinality of a set B . So we get $wt(f_\lambda) = 2^{2j-1}d$. The weights 0 and $2^{2j-1}d$ both do not correspond to a weight of a bent function, completing the proof. $\square$

THEOREM 3.3. *Let $n = 2t$ and $d = 2^i + 2^j + 1$, where d satisfies (3.1). Given $\lambda \in \mathbf{F}_{2^n}^*$, let us define*

$$f_\lambda : x \in \mathbf{F}_{2^n} \mapsto T_1^n(\lambda x^d).$$

Then f_λ can be represented in the form (2.1) exactly in one of the following cases:

- (i) The trivial case: $n = 6j$ and $d = 2^{4j} + 2^{2j} + 1$ with $\lambda + \lambda^{2^{2j}} + \lambda^{2^{4j}} = 0$ —in this case f_λ is the null function;
- (ii) $n = 6j$ and $d = 2^{2j} + 2^j + 1$ with $\lambda \in \mathbf{F}_{2^{3j}}$ satisfying $\lambda + \lambda^{2^j} + \lambda^{2^{2j}} = 0$;
- (iii) $d = 2^t + 2^j + 1$ with $\lambda \in \mathbf{F}_{2^t}$.

Proof. The function f_λ can be represented in the form (2.1) if and only if the sum of the summands in (3.2) and (3.3) is either constant or linear on y . Note that since $y \in \mathbf{F}_{2^t}$ we are interested in the behavior of $d \pmod{2^t - 1}$.

Let us consider (3.2). We want to determine d such that $c(y) = T_1^t(T_t^{2t}(\lambda)y^d)$ is of degree strictly less than 3. Because of the conditions on i, j , if $c(y)$ is not constantly zero, then i must be equal to t . We will consider the cases $i = t$ and $i \neq t$ separately.

Let $d = 2^t + 2^j + 1$. Then the sum of terms in (3.2) and (3.3) becomes

$$g(y) = T_1^t \left(T_t^{2t}(\lambda)y^{2^j+2} \right) + T_1^t \left(T_t^{2t}(\lambda)(a^{2^t} + a)y^{2^j+1} + T_t^{2t}(\lambda a^{2^j})y^2 \right).$$

Suppose that $\lambda \notin \mathbf{F}_{2^t}$. So the function g is of degree 2. Indeed, the integer $2^j + 1$ cannot be a power of 2 and it is not in the cyclotomic coset (modulo $2^t - 1$) of $2^j + 2$. We must have

$$(3.4) \quad T_1^t \left(T_t^{2t}(\lambda)(a^{2^t} + a)y^{2^j+1} \right) = 0,$$

for any $a \in W$, where $j < t$. By Lemma 3.1(b), (3.4) holds if and only if

$$t = 2\ell, \quad j = \ell \quad \text{and} \quad T_t^{2t}(\lambda)T_t^{2t}(a) \in \mathbf{F}_{2^\ell}.$$

But this last condition cannot hold for any $a \in W$, since the map $a \mapsto T_t^{2t}(a)$, from W to $\mathbf{F}_{2^t}$, is bijective. We conclude that we must have $T_t^{2t}(\lambda) = 0$. In this case, we get $g(y) = T_1^t(T_t^{2t}(\lambda a^{2^j})y^2)$, which proves case (iii).

Let $d = 2^i + 2^j + 1$ with $i \neq t$. Set $d' = d \pmod{2^t - 1}$, $0 < d' < 2^t - 1$. In this case $c(y)$ must be constantly zero, and thus applying Lemma 3.1(a), we get

$$t = 3k, \quad d' = 2^{2k} + 2^k + 1 \quad \text{and} \quad T_{3k}^{6k}(\lambda + \lambda^{2^k} + \lambda^{2^{2k}}) = T_k^{6k}(\lambda) = 0.$$

Hence there are two possibilities for d : either $d = d'$ or $d = 2^{4k} + 2^{2k} + 1$. Suppose $d = d'$. Then the sum in (3.3) is as follows:

$$\begin{aligned} & T_1^{3k} \left(T_{3k}^{6k}(\lambda a^{2^{2k}})y^{2^k+1} + T_{3k}^{6k}(\lambda a^{2^k})y^{2^{2k}(2^k+1)} + T_{3k}^{6k}(\lambda a)y^{2^{2k}(2^k+1)} \right) \\ &= T_1^{3k} \left(T_{3k}^{6k}((\lambda + \lambda^{2^k} + \lambda^{2^{2k}})a^{2^{2k}})y^{2^k+1} \right) \\ &= T_1^{3k} \left((\lambda + \lambda^{2^k} + \lambda^{2^{2k}})T_{3k}^{6k}(a^{2^{2k}})y^{2^k+1} \right), \end{aligned}$$

where the last equality follows from previously obtained condition $T_k^{6k}(\lambda) = 0$. Lemma 3.1(b), implies that $\lambda + \lambda^{2^k} + \lambda^{2^{2k}} = 0$, which is possible only for $\lambda \in \mathbf{F}_{2^{3k}}$, because the polynomial $X + X^{2^k} + X^{2^{2k}}$ divides $X^{2^{3k}} + X$. Hence we have proved case (ii).

The last possibility is $d = 2^{4k} + 2^{2k} + 1$. We apply Lemma 3.2. If f_λ is not the null function, then $\mathcal{F}(f_\lambda) = 2^n - 2^{2j}d$. Since $\mathcal{F}(f_\lambda)$ is not divisible by 2^{3j} , we deduce from Proposition 2.1 that f_λ cannot be represented in the form (2.1). Hence we obtain the trivial case (i). $\square$

Further, we want to see when the decomposable cubic monomial functions are bent. We have seen that the exponent $d = 2^{4j} + 2^{2j} + 1$ does not lead to a bent monomial function. The bent monomial functions with the exponent $2^{2j} + 2^j + 1$ are studied in [1], where the following result is proved.

THEOREM 3.4 (see [1]). *The monomial function $T_1^{6r}(\lambda'x^{2^{2r}+2^r+1})$ in $\mathbf{F}_{2^{6r}}$ is bent if and only if there are $\lambda \in \mathbf{F}_{2^{3r}}$ and $a \in \mathbf{F}_{2^{6r}}$ such that $T_r^{3r}(\lambda) = 0$ and $\lambda' = \lambda a^{2^{2r}+2^r+1}$. Moreover, all these bent functions are from the Maiorana–McFarland class.*

By Theorem 3.3 only the exponents $2^t + 2^j + 1$ may provide further examples of cubic monomial bent functions from $\mathcal{M}(\mathbf{F}_{2^t})$. The rest of this paper is devoted to the study of this exponent.

4. Exponent $2^t + 2^j + 1$. In the study of the Walsh spectrum of the monomial functions with exponents $2^t + 2^j + 1$, we may restrict ourselves to the elements λ from the subfield $\mathbf{F}_{2^t}$. This is a consequence of the following observation.

LEMMA 4.1. *Let b be such that $\gcd(b, 2^t + 1) = 1$. Then, for any $\mu \in \mathbf{F}_{2^{2t}}$, there are $\lambda \in \mathbf{F}_{2^t}$ and $\delta \in \mathbf{F}_{2^{2t}}$ such that $\mu = \lambda\delta^b$. Consequently, the Walsh spectrum of the monomial function $T_1^{2t}(\mu x^b)$ is the same as the one of $T_1^{2t}(\lambda x^b)$.*

Proof. Let α be a primitive element of $\mathbf{F}_{2^{2t}}$. Since $\gcd(2^t - 1, 2^t + 1) = 1$, any element $\mu \in \mathbf{F}_{2^{2t}}^*$ can be expressed as follows:

$$\mu = \alpha^{\ell(2^t-1)} \alpha^{k(2^t+1)}, \quad 0 \leq \ell \leq 2^t, \quad 0 \leq k \leq 2^t - 2.$$

Since b and $2^t + 1$ are coprime, it holds that $\mu = \lambda\delta^b$ with $\lambda = \alpha^{k(2^t+1)}$ and $\delta^b = \alpha^{\ell(2^t-1)}$. $\square$

The next proposition gives the explicit decomposition of the function

$$(4.1) \quad g_\lambda(x) = T_1^{2t}(\lambda x^{2^t+2^j+1}), \quad 0 < j < t, \quad \lambda \in \mathbf{F}_{2^t}^*.$$

PROPOSITION 4.2. *Let $W \subset \mathbf{F}_{2^{2t}}$ be such that $\mathbf{F}_{2^{2t}} = \mathbf{F}_{2^t} \oplus W$. Then the function g_λ defined in (4.1) can be written as follows:*

$$g_\lambda(y, a) = T_1^t(y\pi(a) + h(a)), \quad y \in \mathbf{F}_{2^t}, \quad a \in W,$$

where $\pi, h : W \rightarrow \mathbf{F}_{2^t}$ are given by

$$\pi(a) = \lambda^{2^{t-1}}(a + a^{2^t})^{2^{j-1}} + \lambda(a^{2^t} + a)^{2^j+1}$$

and

$$h(a) = \lambda a^{2^t+1}(a + a^{2^t})^{2^j}.$$

Proof. For any $y \in \mathbf{F}_{2^t}$ and $a \in W$, we compute

$$g_\lambda(y, a) = g_\lambda(y + a) = T_1^{2t}(\lambda(y + a)^d).$$

We have $T_1^{2t}(\lambda y^d) = 0$ since λ and y are in $\mathbf{F}_{2^t}$. The part which is (a priori) quadratic relative to y is

$$\begin{aligned} B &= T_1^{2t} \left(\lambda(y^{1+2^j}a + y^{2^j+1}a^{2^t} + y^2a^{2^j}) \right) \\ &= T_1^{2t} \left(\lambda y^{2^j+1}(a^{2^t} + a) + \lambda y^2a^{2^j} \right) \\ &= T_1^{2t}(\lambda y^2a^{2^j}) = T_1^t(\lambda y^2(a + a^{2^t})^{2^j}), \end{aligned}$$

since $\lambda y^{2^j+1}(a^{2^t} + a) \in \mathbf{F}_{2^t}$. Further, we compute the part which is linear relative to y :

$$\begin{aligned} C &= T_1^{2t} \left(y(\lambda(a^{2^t+2^j} + a^{2^j+1}) + \lambda^{2^{t-j}}a^{2^{t-j}(2^t+1)}) \right) \\ &= T_1^{2t} \left(y\lambda(a^{2^t+2^j} + a^{2^j+1}) \right) \\ &= T_1^t \left(y\lambda(a^{2^t} + a)^{2^j+1} \right), \end{aligned}$$

since $\lambda^{2^{t-j}}a^{2^{t-j}(2^t+1)} \in \mathbf{F}_{2^t}$. Finally,

$$g_\lambda(y, a) = T_1^t \left(y(\lambda^{2^{t-1}}(a + a^{2^t})^{2^j+1} + \lambda(a^{2^t} + a)^{2^j+1}) \right) + T_1^{2t}(\lambda a^d).$$

So, we have $g_\lambda(y, a) = T_1^t(y\pi(a) + h(a))$, where $\pi(a)$ is the coefficient of y above and

$$h(a) = \lambda T_1^{2t}(a^d) = \lambda a^{2^t+1}(a + a^{2^t})^{2^j},$$

since $d = 2^t + 2^j + 1$ and $2^t d \equiv 1 + 2^t + 2^{t+j}$. $\square$

The obtained decomposition allows us to get some information on the Walsh spectrum of g_λ in general. By Proposition 2.1 we obtain

$$(4.2) \quad \mathcal{F}(g_\lambda + \varphi_u) = 2^t \sum_{a \in W_u} (-1)^{T_1^t(h(a) + T_1^{2t}(ua))},$$

where $W_u = \{ a \in W \mid \pi(a) = u + u^{2^t} \}$ and, by convention, the sum above is null if $W_u = \emptyset$. Note that

$$W_u = W_{u+\beta} \quad \text{for all } \beta \in \mathbf{F}_{2^t}.$$

Recall that g_λ is bent if and only if π is a permutation (see Corollary 2.2).

As a direct consequence of (4.2) we get a lower bound on the multiplicity of the value 0 in the Walsh spectrum of g_λ . This is because $W_u = \emptyset$ when $u + u^{2^t}$ is not in the image of π .

PROPOSITION 4.3. *Let I be the cardinality of the image of π in $\mathbf{F}_{2^t}$. Then*

$$\#\{ v \in \mathbf{F}_{2^{2t}} \mid \mathcal{F}(g_\lambda + \varphi_v) = 0 \} \geq 2^t(2^t - I).$$

Equation (4.2) gives also an upper bound on the magnitude of the value of Walsh transform at u in terms of the cardinality of W_u .

PROPOSITION 4.4. *For any u , $\mathcal{F}(g_\lambda + \varphi_u) \equiv 0 \pmod{2^t}$. Moreover,*

$$|\mathcal{F}(g_\lambda + \varphi_u)| \leq 2^t \times \#W_u.$$

As we have seen, the Walsh spectrum of g_λ could be determined as soon as we know the set W_u , which describes the preimages of the mapping $\pi : W \rightarrow \mathbf{F}_{2^t}$. Recall that

$$\pi(a) = \lambda^{2^{t-1}}(a + a^{2^t})^{2^{j-1}} + \lambda(a^{2^t} + a)^{2^j+1}.$$

Since the mapping $a \mapsto a + a^{2^t}$ is a permutation from W to $\mathbf{F}_{2^t}$, we have $\pi(a) = \pi'(a + a^{2^t})$, where $\pi' : \mathbf{F}_{2^t} \rightarrow \mathbf{F}_{2^t}$ is given by

$$(4.3) \quad \pi'(y) = \lambda^{2^{t-1}}y^{2^{j-1}} + \lambda y^{2^j+1}, \quad y \in \mathbf{F}_{2^t}.$$

Hence our problem is linked with the one of determining the image of π' or, equivalently, of $(\pi')^2$.

PROBLEM 1. *Let $\lambda \in \mathbf{F}_{2^t}$. For any $\tau \in \mathbf{F}_{2^t}$, compute the number of $y \in \mathbf{F}_{2^t}$ such that*

$$\lambda^2 y^{2(2^j+1)} + \lambda y^{2^j} + \tau = 0.$$

This problem can be easily solved for $\tau = 0$, which yields information on $\mathcal{F}(g_\lambda + \varphi_u)$ for $u \in \mathbf{F}_{2^t}$ as in subsection 4.1 shown. Subsection 4.2 is devoted to the characterization of λ and j such that the equation of Problem 1 has exactly one solution.

4.1. The case $u \in \mathbf{F}_{2^t}$. Let $u \in \mathbf{F}_{2^t}$. Then $W_u = \{ a \in W \mid \pi(a) = 0 \}$ and $W_u = W_{u'}$ for all $u' \in \mathbf{F}_{2^t}$. Clearly, $\pi'(0) = 0$, where π' is given by (4.3). For nonzero y , it holds that

$$\pi'(y) = 0 \Leftrightarrow y^{2^{j-1}+1} = \lambda^{2^{t-1}-1} \Leftrightarrow y^{2^j+2} = \frac{1}{\lambda}.$$

Thus, for any $u \in \mathbf{F}_{2^t}$,

$$(4.4) \quad W_u = \left\{ a \in W \mid a = 0 \text{ or } (a + a^{2^t})^{2^j+2} = \frac{1}{\lambda} \right\}.$$

LEMMA 4.5. *Let $u \in \mathbf{F}_{2^t}$. Then the constant term of g_λ , given by*

$$h(a) = \lambda a^{2^t+1}(a + a^{2^t})^{2^j},$$

satisfies $T_1^t(h(a)) = 1$ for any nonzero $a \in W_u$.

Proof. Note that $h(0) = 0$ and assume now that $a \neq 0$. Using (4.4), we have $(a + a^{2^t})^{2^j+2} = \lambda^{-1}$. By replacing $(a + a^{2^t})^{2^j}$ in $h(a)$, we get

$$h(a) = \lambda a^{2^t+1} \lambda^{-1} (a + a^{2^t})^{-2} = \frac{a^{2^t+1}}{(a + a^{2^t})^2}.$$

The elements a and a^{2^t} are the solutions in $\mathbf{F}_{2^{2t}}$ of the equation

$$x^2 + (a + a^{2^t})x + a^{2^t+1} = 0,$$

while they do not belong to $\mathbf{F}_{2^t}$. This implies that the polynomial $X^2 + (a + a^{2^t})X + a^{2^t+1} \in \mathbf{F}_{2^t}[X]$ is irreducible over $\mathbf{F}_{2^t}$. Hence

$$T_1^t \left(\frac{a^{2^t+1}}{(a + a^{2^t})^2} \right) = T_1^t(h(a)) = 1. \quad \square$$

Now we can describe the values $\mathcal{F}(g_\lambda + \varphi_u)$ for $u \in \mathbf{F}_{2^t}$ more precisely. Notably, we compute the weight of each function g_λ .

PROPOSITION 4.6. *Let $u \in \mathbf{F}_{2^t}$ and W_u be given by (4.4). Set $s = \gcd(2^{j-1} + 1, 2^t - 1)$. Then we have the following:*

(i) *If $s = 1$ (i.e., $t/\gcd(j-1, t)$ odd) then $\#W_u = 2$ and*

$$\mathcal{F}(g_\lambda + \varphi_u) = \begin{cases} 0 & \text{if } T_1^t(u(a + a^{2^t})) = 0, \\ 2^{t+1} & \text{if } T_1^t(u(a + a^{2^t})) = 1, \end{cases}$$

where a is the unique element of W_u different from 0. Note that each value above occurs 2^{t-1} times. In particular, $\mathcal{F}(g_\lambda) = 0$.

(ii) *Set $S_s = \{u^s | u \in \mathbf{F}_{2^t}\}$. For $s > 1$ we have that*

- *if $\lambda \notin S_s$, then $\mathcal{F}(g_\lambda + \varphi_u) = 2^t$ for any u .*
- *if $\lambda \in S_s$, then $|\mathcal{F}(g_\lambda + \varphi_u)| \leq (s+1)2^t$. In particular, $\mathcal{F}(g_\lambda) = (1-s)2^t$.*

Consequently, if g_λ is bent then $s > 1$ and $\lambda \notin S_s$.

Proof. According to (4.4), we get from (4.2) and Lemma 4.5 that

$$\begin{aligned} \mathcal{F}(g_\lambda + \varphi_u) &= 2^t \sum_{a \in W_u} (-1)^{T_1^t(h(a) + u(a + a^{2^t}))} \\ &= 2^t - 2^t \sum_{a \in W_u, a \neq 0} (-1)^{T_1^t(u(a + a^{2^t}))}. \end{aligned}$$

(i) When $s = 1$ then W_u contains only 0 and a unique a such that $(a + a^{2^t})^{2^j+2} = \lambda^{-1}$. Consequently, $\mathcal{F}(g_\lambda + \varphi_u) = 0$ or 2^{t+1} depending on whether $T_1^t(u(a + a^{2^t})) = 0$ or 1. Since u runs through $\mathbf{F}_{2^t}$, each such case occurs 2^{t-1} times. In particular, for $u = 0$, we get $\mathcal{F}(g_\lambda) = 0$.

(ii) Now assume that $s > 1$. If $\lambda \notin S_s$, then $W_u = \{0\}$, implying $\mathcal{F}(g_\lambda + \varphi_u) = 2^t$ for any u .

Denote by G_s the subgroup of $\mathbf{F}_{2^t}^*$ of order s . If $\lambda \in S_s$, then

$$W_u = \{0\} \cup \{a \in W \mid a + a^{2^t} \in yG_s \text{ with } y^{2^j+2} = \lambda^{-1}\}$$

(where $y \in \mathbf{F}_{2^t}$). Using Proposition 4.4, we conclude that $|\mathcal{F}(g_\lambda + \varphi_u)| \leq (s+1)2^t$ when $\lambda \in S_s$. Finally, observe that for $u = 0$, we get $\mathcal{F}(g_\lambda) = (1-s)2^t$. $\square$

4.2. The bent functions. In this section we find all j and $\mu \in \mathbf{F}_{2^{2t}}$ defining bent monomial functions $x \mapsto T_1^{2t}(\mu x^{2^t+2^j+1})$. Remember that this is equivalent to describing those $\lambda \in \mathbf{F}_{2^t}$ for which π' , defined in (4.3), is a permutation. Consider

$$\begin{aligned} (\pi'(y))^2 &= \lambda y^{2^j} + \lambda^2 y^{2^j+1+2} \\ &= \lambda(y + \lambda^{2^t-j} y^{2^t-j(2^j+1)+2})^{2^j} \\ &= \lambda(y + \lambda^{2^t-j} y^{2^t-j+1+2})^{2^j} \\ &= \lambda^2(\lambda^{-2^t-j} y + y^{2^t-j+1+2})^{2^j}. \end{aligned}$$

Thus, the function g_λ is bent if and only if the mapping $\rho: \mathbf{F}_{2^t} \rightarrow \mathbf{F}_{2^t}$, given by

$$\rho(y) = y^{2^t-j+1+2} + \nu y,$$

where $\nu = \lambda^{-2^t-j}$, is a permutation. All such permutation polynomials are characterized in Theorem 5.1 of the next section. In particular, Theorem 5.1 implies the following proposition.

PROPOSITION 4.7. Let $g_\lambda = T_1^{2t}(\lambda x^{2^t+2^j+1})$, $0 < j < t$, $\lambda \in \mathbf{F}_{2^t}^*$. Then, g_λ can be bent only if $t = 2r$, r odd, and $j = r + 1$.

In [10] it is shown that the exponent $2^{2r} + 2^{r+1} + 1$ in $\mathbf{F}_{2^{4r}}$, r odd, yields a bent function if λ is chosen to be β^5 , where β is the root of the polynomial $X^4 + X + 1$. Theorem 5.1 allows us to extend this result.

THEOREM 4.8. Let $\omega \in \mathbf{F}_4 \setminus \mathbf{F}_2$. The monomial function $T_1^{4r}(\mu x^{2^{2r}+2^{r+1}+1})$ on $\mathbf{F}_{2^{4r}}$ is bent for some $\mu \in \mathbf{F}_{2^{4r}}$ if and only if r is odd and there are $\lambda \in \omega \mathbf{F}_{2^r}$ and $a \in \mathbf{F}_{2^{4r}}$ such that $\mu = \lambda a^{2^{2r}+2^{r+1}+1}$.

5. On some binomial quadratic permutations in $\mathbf{F}_{2^t}$. In this section we describe all permutation polynomials of type

$$X^{2^k+2} + \nu X$$

in a finite field of characteristics 2. More precisely, we will prove the next theorem.

THEOREM 5.1. Let $0 \leq k \leq t - 1$ and $\nu \neq 0$. Then the polynomial

$$\rho(X) = X^{2^k+2} + \nu X \quad \text{in } \mathbf{F}_{2^t}[X]$$

is a permutation polynomial of $\mathbf{F}_{2^t}$ if and only if t is even and

- either $k = 1$ and ν is not a third power in $\mathbf{F}_{2^t}$, or
- $t = 2r$, $r \geq 3$ with r odd, $k = r$, and $\nu \in \omega \mathbf{F}_{2^r}$, where $\omega \in \mathbf{F}_{2^2} \setminus \mathbf{F}_2$.

We will use some well-known facts about the quadratic monomial Boolean functions. Our main reference is [11, pp. 175–189]; we give only the idea of a proof of Proposition 5.2. The interested reader may see [8, 10] for these proofs and [5, 8] for some other properties of the quadratic monomial functions.

PROPOSITION 5.2. Let $\gcd(t, i) = s$ and $F: \mathbf{F}_{2^t} \rightarrow \mathbf{F}_{2^t}$ be given by $F(x) = x^{2^i+1}$. Let f_μ , $\mu \in \mathbf{F}_{2^t}^*$, be the Boolean function defined by $f_\mu(x) = T_1^t(\mu F(x))$. Then the following properties hold:

- (a) The function F is bijective if and only if t/s is odd.
- (b) If t/s is odd, then the Walsh transform of the Boolean function f_μ takes the values $\{0, \pm 2^{\frac{t+s}{2}}\}$. In particular, f_μ is never bent.
- (c) Assume that t/s is even. The function f_μ is bent if and only if μ is not a $(2^i + 1)$ th power in $\mathbf{F}_{2^t}$. Otherwise, its Walsh transform takes the values $\{0, \pm 2^{\frac{t+2s}{2}}\}$.

Proof. The function F is bijective if and only if $\gcd(2^i + 1, 2^t - 1) = 1$ which is equivalent to $s = \gcd(2i, t)$, i.e., t/s odd.

Set $S = \{y^{2^i+1} | y \in \mathbf{F}_{2^t}\}$. The function f_μ is bent if and only if t/s is even and $\mu \notin S$ (see [10, Thm. 2]). Otherwise the Walsh spectrum of f_μ is the same as that of the function $x \mapsto T_1^t(x^{2^i+1})$. The Walsh transform of this quadratic function takes the values

$$\begin{aligned} &\{0, \pm 2^{\frac{t+s}{2}}\} && \text{if } t/s \text{ is odd,} \\ &\{0, \pm 2^{\frac{t+2s}{2}}\} && \text{if } t/s \text{ is even and } \mu \in S. \end{aligned}$$

(see [11, Thm. 11.13]). $\square$

Further, we need also the following property of the Walsh transform of quadratic Boolean functions.

PROPOSITION 5.3. Let $\gcd(t, i) = s$, t/s be odd. Let f be the Boolean function on $\mathbf{F}_{2^t}$ defined by $f(x) = T_1^t(x^{2^i+1} + \beta x)$. Then f is balanced if and only if $T_s^t(\beta) \neq 1$.

Proof. Note that $T_s^t(1) = 1$ since $t = su$ with odd u . Such a function f is balanced if and only if there is $a \in \mathbf{F}_{2^t}^*$ such that the derivative of f with respect to a constantly equals 1, that is,

$$h_a(x) = f(x) + f(x+a) = 1 \quad \text{for all } x$$

(see a proof in [4, Prop. II-5]). We compute the derivative of f with respect to a :

$$h_a(x) = T_1^t(ax^{2^i} + a^{2^i}x + a^{2^i+1} + \beta a) = T_1^t\left(x(a^{2^{t-i}} + a^{2^i}) + a(a^{2^i} + \beta)\right).$$

Then h_a is constant if and only if $a^{2^{2i}} = a$, i.e., $a \in \mathbf{F}_{2^s}^*$. In this case,

$$h_a(x) = T_1^t(a + \beta a) = T_1^s(aT_s^t(1 + \beta)) = T_1^s(a(1 + T_s^t(\beta))).$$

So, there is $a \in \mathbf{F}_{2^s}^*$ such that $h_a = 1$ if and only if $T_s^t(\beta) \neq 1$. $\square$

Now we are going to prove Theorem 5.1 by means of the two following lemmas.

LEMMA 5.4. *Let $\nu \in \mathbf{F}_{2^t}$ and $\rho : \mathbf{F}_{2^t} \rightarrow \mathbf{F}_{2^t}$ be given by the polynomial*

$$(5.1) \quad \rho(X) = X^{2^k+2} + \nu X, \quad 2 \leq k \leq t-1.$$

Take $c = \gcd(t, k-1)$. Then ρ is a permutation polynomial of $\mathbf{F}_{2^t}$ if and only if

$$\frac{t}{c} \text{ is odd and } T_c^t(\gamma^{2^k+1}\nu) \neq 1 \quad \text{for every } \gamma \in \mathbf{F}_{2^t}.$$

In this case, t is even, $\frac{t}{\gcd(t,k)}$ is even, and k is odd.

Proof. It is well known that a mapping $g : \mathbf{F}_{2^t} \rightarrow \mathbf{F}_{2^t}$ is bijective if and only if

$$\sum_{x \in \mathbf{F}_{2^t}} (-1)^{T_1^t(\beta g(x))} = 0$$

for any $\beta \in \mathbf{F}_{2^t}^*$. Consider the sum

$$S = \sum_{x \in \mathbf{F}_{2^t}} (-1)^{T_1^t(\beta x^{2^k+2} + \beta \nu x)} = \sum_{x \in \mathbf{F}_{2^t}} (-1)^{T_1^t(\beta^{2^{n-1}} x^{2^{k-1}+1} + \beta \nu x)}.$$

Suppose t/c is even; then by Proposition 5.2(c), there is a β such that the function $x \mapsto T_1^t(\beta^{2^{n-1}} x^{2^{k-1}+1})$ is bent. Clearly, $S \neq 0$ for such a β . Hence ρ defines a permutation only if t/c is odd.

We assume that t/c is odd, i.e., by Proposition 5.2(a), the mapping $x \mapsto x^{2^{k-1}+1}$ is a permutation on $\mathbf{F}_{2^t}$. Then $\beta^{2^{n-1}} = \gamma^{2^{k-1}+1}$ for some $\gamma \in \mathbf{F}_{2^t}$ and

$$\begin{aligned} S &= \sum_{x \in \mathbf{F}_{2^t}} (-1)^{T_1^t((\gamma x)^{2^{k-1}+1} + \gamma^{2^k+2} \nu x)} \\ &= \sum_{z \in \mathbf{F}_{2^t}} (-1)^{T_1^t(z^{2^{k-1}+1} + \gamma^{2^k+1} \nu z)}, \end{aligned}$$

where $z = \gamma x$. Hence, we are interested in $\nu \in \mathbf{F}_{2^t}$ satisfying

$$(5.2) \quad \sum_{z \in \mathbf{F}_{2^t}} (-1)^{T_1^t(z^{2^{k-1}+1} + \gamma^{2^k+1} \nu z)} = 0 \quad \text{for every } \gamma \in \mathbf{F}_{2^t}.$$

By Proposition 5.3, condition (5.2) is satisfied if and only if

$$(5.3) \quad T_c^t(\gamma^{2^k+1}\nu) \neq 1 \quad \text{for every } \gamma \in \mathbf{F}_{2^t}.$$

Let $s = \gcd(t, k)$. If (5.3) holds, then t/s must be even because otherwise $x \mapsto x^{2^k+1}$ would be a permutation. Hence t must be even. Further, $k-1$ must be even when t/c is odd. $\square$

Now we are going to prove that condition (5.3) is generally impossible.

LEMMA 5.5. *Let $t = 2r$, $1 \leq k \leq r$, and $c = \gcd(t, k-1)$, where k is odd. Assume that there exists an element $\lambda \in \mathbf{F}_{2^t}^*$ satisfying*

$$T_c^t(\gamma^{2^k+1}\lambda) \neq 1 \quad \text{for every } \gamma \in \mathbf{F}_{2^t}.$$

Then $k = r$ and $c = 2$. In particular, r is odd.

Proof. Note that c is even, by hypothesis, and thus $c \geq 2$. Denote by N the number of $x \in \mathbf{F}_{2^t}$ satisfying

$$T_c^t(x^{2^k+1}\lambda) = 1.$$

We have

$$\begin{aligned} & \sum_{x \in \mathbf{F}_{2^t}} \sum_{a \in \mathbf{F}_{2^t}} (-1)^{T_1^t(a(T_c^t(x^{2^k+1}\lambda)+1))} \\ &= 2^t \times \#\{x \in \mathbf{F}_{2^t} \mid T_c^t(x^{2^k+1}\lambda) = 1\} \\ &= 2^t N. \end{aligned}$$

Hence we can use the results on the Walsh transform of quadratic monomial functions to bound N . Our goal is to show that $N \geq 1$, except for the case $k = r$. Setting $b = T_c^t(a)$, the above sum can be modified to

$$\begin{aligned} N &= \frac{1}{2^t} \sum_{b \in \mathbf{F}_{2^c}} 2^{t-c} (-1)^{T_1^c(b)} \sum_{x \in \mathbf{F}_{2^t}} (-1)^{T_1^c(b(T_c^t(x^{2^k+1}\lambda)))} \\ &= \frac{1}{2^c} \sum_{b \in \mathbf{F}_{2^c}} (-1)^{T_1^c(b)} \sum_{x \in \mathbf{F}_{2^t}} (-1)^{T_1^t(b\lambda x^{2^k+1})} \\ &= \frac{1}{2^c} \left(2^t + \sum_{b \in \mathbf{F}_{2^c}^*} (-1)^{T_1^c(b)} \sum_{x \in \mathbf{F}_{2^t}} (-1)^{T_1^t(b\lambda x^{2^k+1})} \right). \end{aligned}$$

Let $s = \gcd(t, k)$. Since t/s is even, we have, by Proposition 5.2,

$$\mathcal{F}(h) = \sum_{x \in \mathbf{F}_{2^t}} (-1)^{T_1^t(b\lambda x^{2^k+1})} \in \{0, \pm 2^{r+s}, \pm 2^r\}.$$

Thus,

$$N = \frac{1}{2^c} \left(2^t + \sum_{b \in \mathbf{F}_{2^c}^*} (-1)^{T_1^c(b)} \epsilon_b \right),$$

where $\epsilon_b \in \{0, \pm 2^r, \pm 2^{r+s}\}$. In particular,

$$(5.4) \quad N > \frac{1}{2^c} (2^t - 2^c 2^{r+s}) = 2^{t-c} - 2^{r+s} = 2^r (2^{r-c} - 2^s).$$

Let us prove that $r - c \geq s$ unless $k = r$. Recall that $c = \gcd(2r, k - 1)$ is even. Moreover, s is odd, $c < k \leq r$, and, clearly, $\gcd(s, c) = 1$.

First, note that $r - c \geq s$ for $s = 1$, since $r - c \geq 1$. So we assume that $s \geq 3$. Set $c = 2c'$ so that $r = sc'v$ for some c' and some $v \geq 1$. So we have

$$r - c = sc'v - 2c' = c'(sv - 2).$$

If $v > 1$, then for any c' ,

$$r - c = c'(sv - 2) = c'(s + (s(v - 1) - 2)) \geq s.$$

If $v = 1$, then $r = c's$ and we want to check when $c'(s - 2) \geq s$. Setting $s = 2 + \tau$ with τ odd, it holds if and only if $c'\tau \geq \tau + 2$. This last inequality is satisfied unless $c' = 1$ or $c' = 2$ and $\tau = 1$.

In the case where $c' = 2$ and $\tau = 1$, we have $r = 6$ and $s = 3$ so that $k = 3$, since $s = \gcd(r, k)$. Further, $c = 4$ and we get $k > c$, a contradiction.

Finally, it appears that $r - c \geq s$ unless $c' = 1$ and $v = 1$, the case where $r = s = k$. According to (5.4), we conclude that $N \geq 1$ unless $k = r$. In this case, we have $k = r = s$, which implies that r is odd and $c = 2$, completing the proof. $\square$

Now we are able to prove our main theorem.

Proof of Theorem 5.1. We are now able to determine which polynomials of type

$$\rho(X) = X^{2^k+2} + \nu X, \quad \rho \in \mathbf{F}_{2^t}[X],$$

with $0 \leq k \leq t - 1$, are permutations of $\mathbf{F}_{2^t}$. If $k = 0$, then $\rho(X) = X^3 + \nu X$. So ρ cannot be a permutation, since it has two roots in $\mathbf{F}_{2^t}$. For $k = 1$, we get

$$\rho(X) = X^4 + \nu X = X(X^3 + \nu),$$

which is a linearized polynomial. Thus, it is a permutation polynomial if and only if it has only one root, i.e., ν is never equal to x^3 when x runs through $\mathbf{F}_{2^t}^*$.

We now assume that $k > 1$. According to Lemmas 5.4 and 5.5, we know that ρ is a permutation only when $t = 2r$, with r odd, $k = r$, and $c = \gcd(2r, r - 1) = 2$. Moreover, those ν such that ρ is a permutation are the elements of $\mathbf{F}_{2^t}^*$ which satisfy

$$(5.5) \quad T_2^{2r}(\gamma^{2^r+1}\nu) \neq 1 \quad \text{for every } \gamma \in \mathbf{F}_{2^{2r}}.$$

Note that $\{\gamma^{2^r+1} | \gamma \in \mathbf{F}_{2^{2r}}\} = \mathbf{F}_{2^r}$, and therefore (5.5) is equivalent to

$$(5.6) \quad T_2^{2r}(\delta\nu) \neq 1 \quad \text{for every } \delta \in \mathbf{F}_{2^r}.$$

It remains to describe the set of suitable ν . Consider the map $\ell_\nu : \mathbf{F}_{2^r} \rightarrow \mathbf{F}_{2^2}$ defined by $\ell_\nu(z) = T_2^{2r}(\nu z)$. Note that ℓ_ν is $\mathbf{F}_2$ -linear and

$$\ell_\nu(z) = \nu z + \cdots + (\nu z)^{2^{r-1}} + \nu^{2^{r+1}} z^2 + \cdots + \nu^{2^{2(r-1)}} z^{2^{r-2}}.$$

So ℓ_ν cannot be constant.

Moreover, if also (5.6) holds, then

$$(5.7) \quad \{\ell_\nu(z) | z \in \mathbf{F}_{2^r}\} = \{0, \omega\}, \quad \text{where } \omega \in \mathbf{F}_{2^2} \setminus \mathbf{F}_2.$$

Assume that ℓ_ν satisfies (5.7) for such ω . We consider now the map $\ell_{\nu\omega^2}$. Since $\ell_{\nu\omega^2}(z) = \omega^2\ell_\nu(z)$, we have

$$(5.8) \quad \{\ell_{\nu\omega^2}(z) | z \in \mathbf{F}_{2^r}\} = \{0, 1\}.$$

From (5.8) it follows that $T_1^2(\ell_{\nu\omega^2}(z)) = 0$ for every $z \in \mathbf{F}_{2^r}$. Observe, that

$$T_1^2(\ell_{\nu\omega^2}(z)) = T_1^2(T_2^{2r}(\nu\omega^2 z)) = T_1^{2r}(\nu\omega^2 z) = T_1^r(T_r^{2r}(\nu\omega^2 z)).$$

But the function $z \mapsto T_1^r(T_r^{2r}(\nu\omega^2 z))$ is constantly zero if and only if $T_r^{2r}(\nu\omega^2)$ is zero, which is equivalent to $\nu\omega^2 \in \mathbf{F}_{2^r}$. To complete the proof, note that for any $\nu = \omega u$, with $u \in \mathbf{F}_{2^r}$, we have

$$\ell_{\nu\omega}(z) = \omega\ell_1(uz) = \omega T_1^r(uz),$$

implying (5.7). $\square$

Remark 1. Note that we have proved the following: There are $2(2^r - 1)$ permutations of $\mathbf{F}_{2^{2r}}$ with polynomial form $X^{2^r+2} + \nu X$. These correspond to

$$\nu \in (\omega\mathbf{F}_{2^r}^*) \cup (\omega^2\mathbf{F}_{2^r}^*)$$

with notation $\mathbf{F}_4 = \{0, 1, \omega, \omega^2\}$.

6. Conclusion. As one may see from the list of known monomial bent functions in the introduction, the quadratic monomial bent functions are easily characterized. The complete classification of the cubic monomial bent functions $T_1^{2^t}(\lambda x^{2^t+2^j+1})$ seems to be a difficult problem. In this paper all bent exponents $2^t + 2^j + 1$ are found. Further, it is shown that all cubic monomial bent functions from the Maiorana–McFarland family to the subfield are known. All known bent exponents (computer search for $n \leq 24$, exhaustive search for $n \leq 20$) are covered by the examples in the introduction. However, the following question remains open.

OPEN PROBLEM 1. *Are there cubic monomial bent functions besides those listed in the introduction of this paper?*

Acknowledgments. The authors thank Anne Canteaut for many helpful discussions along this work. They wish also to thank Gregor Leander for his valuable comments.

REFERENCES

- [1] A. CANTEAUT, P. CHARPIN, AND G. KYUREGHYAN, *A new class of monomial bent functions*, in Proceedings of the 2006 IEEE International Symposium on Information Theory (ISIT 06, Seattle), IEEE Press, Piscataway, NJ, 2006, pp. 903–906.
- [2] A. CANTEAUT, P. CHARPIN, AND G. KYUREGHYAN, *A new class of monomial bent functions*, Finite Fields Appl., 14 (2008), pp. 221–241.
- [3] A. CANTEAUT, M. DAUM, H. DOBBERTIN, AND G. LEANDER, *Finding nonnormal bent functions*, Discrete Appl. Math., 154 (2006), pp. 202–218.
- [4] A. CANTEAUT, C. CARLET, P. CHARPIN, AND C. FONTAINE, *On cryptographic properties of the cosets of $R(1, m)$* , IEEE Trans. Inform. Theory, 47 (2001), pp. 1494–1513.
- [5] P. CHARPIN, C. TAVERNIER, AND E. PASALIC, *On bent and semi-bent quadratic Boolean functions*, IEEE Trans. Inform. Theory, 51 (2005), pp. 4287–4298.

- [6] J. F. DILLON, *Elementary Hadamard Difference Sets*, Ph.D. dissertation, University of Maryland, College Park, MD, 1974.
- [7] J. F. DILLON AND H. DOBBERTIN, *New cyclic difference sets with Singer parameters*, *Finite Fields Appl.*, 10 (2004), pp. 342–389.
- [8] G. M. KYUREGHYAN, *Crooked maps in $\mathbf{F}_{2^n}$* , *Finite Fields Appl.*, 13 (2007), pp. 713–726.
- [9] G. LACHAUD AND J. WOLFMANN, *The weights of the orthogonals of the extended quadratic binary Goppa codes*, *IEEE Trans. Inform. Theory*, 36 (1990), pp. 686–692.
- [10] N. G. LEANDER, *Monomial bent functions*, *IEEE Trans. Inform. Theory*, 52 (2006), pp. 738–743.
- [11] R. J. McELIECE, *Finite Fields for Computer Scientists and Engineers*, Kluwer, Boston, 1987.
- [12] R. L. McFARLAND, *A family of noncyclic difference sets*, *J. Combin. Theory Ser. A*, 15 (1973), pp. 1–10.

RANK-WIDTH AND WELL-QUASI-ORDERING*

SANG-IL OUM†

Abstract. Robertson and Seymour [*J. Combin. Theory Ser. B*, 48 (1990), pp. 227–254] proved that graphs of bounded tree-width are well-quasi-ordered by the graph minor relation. By extending their arguments, Geelen, Gerards, and Whittle [*J. Combin. Theory Ser. B*, 84 (2002), pp. 270–290] proved that binary matroids of bounded branch-width are well-quasi-ordered by the matroid minor relation. We prove another theorem of this kind in terms of *rank-width* and *vertex-minors*. For a graph $G = (V, E)$ and a vertex v of G , a *local complementation* at v is an operation that replaces the subgraph induced by the neighbors of v with its complement graph. A graph H is called a *vertex-minor* of G if H can be obtained from G by applying a sequence of vertex deletions and local complementations. Rank-width was defined by Oum and Seymour [*J. Combin. Theory Ser. B*, 96 (2006), pp. 514–528] to investigate clique-width; they showed that graphs have bounded rank-width if and only if they have bounded clique-width. We prove that graphs of bounded rank-width are well-quasi-ordered by the vertex-minor relation; in other words, for every infinite sequence $G_1, G_2, \dots$ of graphs of rank-width (or clique-width) at most k , there exist $i < j$ such that G_i is isomorphic to a vertex-minor of G_j . This implies that there is a finite list of graphs such that a graph has rank-width at most k if and only if it contains no one in the list as a vertex-minor. The proof uses the notion of *isotropic systems* defined by Bouchet.

Key words. rank-width, clique-width, well-quasi-ordering, isotropic system, local complementation, pivoting, binary matroid, pivot-minor, vertex-minor

AMS subject classifications. 05C83, 05C40

DOI. 10.1137/050629616

1. Introduction. Oum and Seymour [12] defined *rank-width* in order to investigate *clique-width*, which was defined by Courcelle and Olariu [5]. Rank-width is a complexity measure of graph in a kind of tree structure, called a *rank-decomposition*. Oum and Seymour showed that graphs have bounded rank-width if and only if they have bounded clique-width and they used it to find an approximation algorithm for clique-width.

Later the author [11] showed that the notion of rank-width of graphs has an interesting intersection with the notion of *branch-width* of matroids. More specifically, the branch-width of a binary matroid is exactly one more than the rank-width of its fundamental graph.

The branch-width of matroids generalizes the branch-width of graphs, both defined by Robertson and Seymour [14]. Branch-width of graphs is approximately equal to *tree-width*, also defined by Robertson and Seymour. Informally speaking, tree-width is a measure describing how close a graph is to being a forest.

Both tree-width and branch-width are interesting with respect to the *graph minor* relation. A *contraction* of an edge e is an operation, that deletes e and identifies the ends of e . A graph H is a *minor* of a graph G if H can be obtained from G by a sequence of contractions, vertex deletions, and edge deletions. If H is a minor of G , then the tree-width of H is at most that of G and the same is true for branch-width. This implies that a set of graphs of bounded tree-width is closed under the graph minor relation.

*Received by the editors April 19, 2005; accepted for publication (in revised form) September 26, 2007; published electronically March 28, 2008.

<http://www.siam.org/journals/sidma/22-2/62961.html>

†Department of Mathematical Sciences, KAIST, Daejeon, 305-701 Republic of Korea (sangil@kaist.edu).

In this paper, we consider vertex-minors of graphs, previously called *l-reductions* by Bouchet [4]. For a graph G and a vertex v of G , let $G * v$ be the graph, obtained by *local complementation* at v , that is, by replacing the subgraph induced by the neighbors of v with its complement graph. We say that G is *locally equivalent* to H if H can be obtained from G by applying a sequence of local complementations. A graph H is a *vertex-minor* of G if H can be obtained from G by applying a sequence of vertex deletions and local complementations. A simple fact is that if H is a vertex-minor of G , then the rank-width of H is at most that of G . For an edge uv of G , *pivoting* on uv , denoted by $G \wedge uv$, is the composition of three local complementations, $G \wedge uv = G * u * v * u$. It is an easy exercise to show that $G * u * v * u = G * v * u * v$ if uv is an edge of G . We say that H is a *pivot-minor* of G if H can be obtained from G by applying a sequence of vertex deletions and pivots. Every pivot-minor of G is a vertex-minor of G , but not vice versa.

In this paper, we prove the following theorem.

THEOREM 4.1. *Let k be a constant. If $\{G_1, G_2, G_3, \dots\}$ is an infinite sequence of graphs of rank-width at most k , then there exist $i < j$ such that G_i is isomorphic to a pivot-minor of G_j , and therefore isomorphic to a vertex-minor of G_j .*

This theorem is motivated by the following two theorems. The first one is for graphs of bounded tree-width, proved by Robertson and Seymour [13].

THEOREM 1.1. *Let k be a constant. If $\{G_1, G_2, G_3, \dots\}$ is an infinite sequence of graphs of tree-width at most k , then there exist $i < j$ such that G_i is isomorphic to a minor of G_j .*

The next one, generalizing the previous one, is due to Geelen, Gerards, and Whittle [8].

THEOREM 1.2. *Let k be a constant. Let $\mathbb{F}$ be a finite field. If $\{M_1, M_2, M_3, \dots\}$ is an infinite sequence of $\mathbb{F}$ -representable matroids of branch-width at most k , then there exist $i < j$ such that M_i is isomorphic to a minor of M_j .*

It is straightforward to prove (section 11) that Theorem 1.2 restricted to binary matroids is equivalent to Theorem 4.1 restricted to bipartite graphs. In fact, the main idea of proving Theorem 1.2 remains in our paper, although we have to go through a different technical notion. In the original proof of Theorem 1.2, the authors use “configurations” to represent $\mathbb{F}$ -representable matroids, and then convert the matroid problem into a vector space problem. In our proof, we use the similar approach but a different notion. Bouchet defined *isotropic systems* and developed their minors and their relation to graphs in a series of papers [1, 2, 3]. Informally speaking, an isotropic system can be considered as an equivalence class of graphs by *local equivalence*. Isotropic systems also generalize pairs of a binary matroid and its dual. A detailed definition will be reviewed in section 3.

In sections 2 and 3, we review some notions and results. The overview of the proof is given in section 4. Then in sections 5 till 8, we prove several lemmas, including an extension of Tutte’s linking theorem to isotropic systems, in section 7. Then in section 9 we prove that isotropic systems of bounded branch-width are well-quasi-ordered under the “isotropic system minor” order and in section 10 we prove that the theorem implies the well-quasi-ordering of graphs under the pivot-minor order. In section 11 we explain why the result restricted to bipartite graphs is the same as Theorem 1.2 restricted to binary matroids.

2. Review on rank-width. In this paper, we assume that graphs are simple, undirected, and finite. Let us review the definition of rank-width, introduced by Oum and Seymour [12]. We will first describe the branch-width of connectivity functions and then use this to define the rank-width.

For a finite set V , an integer-valued function c on subsets of V is called a *connectivity function* if $c(X) + c(Y) \geq c(X \cap Y) + c(X \cup Y)$ for all $X, Y \subseteq V$, $c(X) = c(V \setminus X)$ for all $X \subseteq V$, and $c(\emptyset) = 0$. A *subcubic tree* is a tree such that every vertex has exactly one or three incident edges. We call $(T, \mathcal{L})$ a *branch-decomposition* of a connectivity function c on subsets of V if T is a subcubic tree and $\mathcal{L}$ is a bijection from V to the set of all leaves of T . For an edge e of T , connected components of $T \setminus e$ induce a partition (X, Y) of the set of leaves of T . The *width* of an edge e of a branch-decomposition $(T, \mathcal{L})$ is $c(\mathcal{L}^{-1}(X))$. The *width* of $(T, \mathcal{L})$ is the maximum width of all edges of T . The *branch-width* $\text{bw}(c)$ of c is the minimum width of a branch-decomposition of c . (If $|V| \leq 1$, we define that $\text{bw}(c) = c(\emptyset)$.)

For a branch-decomposition $(T, \mathcal{L})$ of c , let e and f be two edges of T . Let E be the set of leaves of T in the component of $T \setminus e$ not containing f , and let F be the set of leaves of T in the component of $T \setminus f$ not containing e . Let P be the shortest path in T containing e and f . We call e and f *linked* if

$$\min_{h \in E(P)} (\text{width of } h \text{ of } (T, \mathcal{L})) = \min_{\mathcal{L}^{-1}(E) \subseteq Z \subseteq V \setminus \mathcal{L}^{-1}(F)} c(Z).$$

We call a branch-decomposition $(T, \mathcal{L})$ *linked* if each pair of edges of T is linked.

For a matrix $M = (m_{ij} : i \in R, j \in C)$ over a field $\mathbb{F}$, if $X \subseteq R$ and $Y \subseteq C$, let $M[X, Y]$ denote the submatrix $(m_{ij} : i \in X, j \in Y)$. For a graph G , let A_G be its adjacency matrix over the binary field $\text{GF}(2)$.

For a graph $G = (V, E)$ and two disjoint subsets $X, Y \subseteq V$, we define $\rho_G^*(X, Y) = \text{rank}(A_G[X, Y])$ where rank is the matrix rank function; and we define the *cut-rank* function ρ_G of G by $\rho_G(X) = \rho_G^*(X, V(G) \setminus X)$ for $X \subseteq V$.

It is well known that the cut-rank function is an instance of a connectivity function (see [12]). A *rank-decomposition* of a graph is a branch-decomposition of its cut-rank function and the *rank-width* of a graph is the branch-width of its cut-rank function.

3. Review on the isotropic system. In this section, the notion of the isotropic system and a few useful theorems will be reviewed. All material is from Bouchet's papers [1, 2, 3]. We change some notations. The author's thesis [10] also contains the proofs of all theorems in this section.

3.1. Definition of the isotropic system. For a vector space W with a bilinear form $\langle \cdot, \cdot \rangle$, a subspace L of W is called *totally isotropic* if and only if $\langle x, y \rangle = 0$ for all $x, y \in L$.

Let $K = \{0, \alpha, \beta, \gamma\}$ be the 2-dimensional vector space over $\text{GF}(2)$. So this means that $\alpha + \beta + \gamma = \alpha + \alpha = \beta + \beta = \gamma + \gamma = 0$. Moreover, let $\langle \cdot, \cdot \rangle$ be the bilinear form on K defined by

$$\langle x, y \rangle = \begin{cases} 1 & \text{if } x \neq y \text{ and } x, y \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Let V be a finite set. Let K^V be the set of functions from V to K , and so K^V is a vector space over $\text{GF}(2)$. We attach the following bilinear form to K^V :

$$\text{for } x, y \in K^V, \quad \langle x, y \rangle = \sum_{v \in V} \langle x(v), y(v) \rangle.$$

DEFINITION 3.1 (see Bouchet [1]). *An isotropic system is a pair $S = (V, L)$ of a finite set V and a totally isotropic subspace L of K^V with $\dim(L) = |V|$. V is called the element set of S .*

For $X \subseteq V$, let $p_X : K^V \rightarrow K^X$ be the *canonical projection* such that

$$(p_X(a))(v) = a(v) \quad \text{for all } v \in X.$$

For $a \in K^V$ and $X \subseteq V$, $a[X]$ is the vector in K^V such that

$$a[X](v) = \begin{cases} a(v) & \text{if } v \in X, \\ 0 & \text{if } v \in V \setminus X. \end{cases}$$

Let L be a subspace of K^V and $v \in V$. Let $x \in K \setminus \{0\} = \{\alpha, \beta, \gamma\}$.

- Let $L^\perp$ be the subspace of K^V such that $L^\perp = \{z \in K^V : \langle z, y \rangle = 0 \text{ for all } y \in L\}$.
- Let $L|_x^v$ be the subspace of $K^{V \setminus \{v\}}$ such that $L|_x^v = \{p_{V \setminus \{v\}}(a) : a \in L, a(v) = 0 \text{ or } x\}$.
- Let $L|_{\subseteq X}$, $L|_X$ be the subspaces of K^X such that

$$\begin{aligned} L|_{\subseteq X} &= \{p_X(a) : a \in L, a(v) = 0 \text{ for all } v \notin X\}, \\ L|_X &= \{p_X(a) : a \in L\}. \end{aligned}$$

Two vectors $a, b \in K^V$ are called *supplementary* if $\langle a(v), b(v) \rangle = 1$ for all $v \in V$. We call $a \in K^V$ *complete* if $a(v) \neq 0$ for all $v \in V$. For $X \subseteq V$ and a complete vector a of K^X , $L|_a^X$ is the subspace of $K^{V \setminus X}$ such that

$$L|_a^X = \{p_{V \setminus X}(b) : b \in L, b(v) \in \{a(v), 0\} \text{ for all } v \in X\}.$$

Note that $L|_{x_1}^{v_1} |_{x_2}^{v_2} |_{x_3}^{v_3} \dots |_{x_k}^{v_k} = L|_x^{\{v_1, v_2, \dots, v_k\}}$ where $x \in K^{\{v_1, v_2, \dots, v_k\}}$ such that $x(v_i) = x_i$.

DEFINITION 3.2. Let $S = (V, L)$ be an isotropic system and $v \in V$. For $x \in K \setminus \{0\}$, $S|_x^v = (V \setminus \{v\}, L|_x^v)$ is called an *elementary minor* of S . An isotropic system S' is called a *minor* of S if S' can be obtained from S by applying a sequence of elementary minor operations; in other words, $S' = S|_{x_1}^{v_1} |_{x_2}^{v_2} |_{x_3}^{v_3} \dots |_{x_k}^{v_k}$ for $v_i \in V$ and $x_i \in K \setminus \{0\}$.

Bouchet [1, (8.1)] proved that an elementary minor of an isotropic system is again an isotropic system and therefore a minor of an isotropic system is an isotropic system.

3.2. Fundamental basis and fundamental graphs. The connection between isotropic systems and graphs was also studied by Bouchet [2].

DEFINITION 3.3. We call $x \in K^V$ an *Eulerian vector* of an isotropic system $S = (V, L)$ if $x[P] \notin L$ for every nonempty subset P of V .

PROPOSITION 3.4 (see Bouchet [2, (4.1)]). For every complete vector c of K^V , there is an Eulerian vector a of S , supplementary to c .

PROPOSITION 3.5 (see Bouchet [2, (4.3)]). Let a be an Eulerian vector of an isotropic system $S = (V, L)$. For every $v \in V$, there exists a unique vector $b_v \in L$ such that

- (i) $b_v(v) \neq 0$,
- (ii) $b_v(w) \in \{0, a(w)\}$ for $w \neq v$.

Moreover, the set $\{b_v : v \in V\}$ is a basis of L .

We call $\{b_v : v \in V\}$ the *fundamental basis* of L with respect to a .

It is straightforward to construct an isotropic system from every graph. Let $G = (V, E)$ be a graph and let a, b be a pair of supplementary vectors of K^V . Let

$n_G(v)$ be the set of neighbors of v . Then, we may construct an isotropic system $S = (V, L)$ [2, (3.1)] by letting L be the subspace of K^V spanned by

$$\{a[n_G(v)] + b[\{v\}] : v \in V\}.$$

We call (G, a, b) a *graphic presentation* of S .

It is interesting that the reverse direction also works. Suppose an isotropic system $S = (V, L)$ is given with an Eulerian vector a . Let $\{b_v : v \in V\}$ be the fundamental basis of $S = (V, L)$ with respect to a . Let $G = (V, E)$ be the graph such that $vw \in E$ if and only if $v \neq w$ and $b_v(w) \neq 0$. Since $\langle b_v, b_w \rangle = 0$ implies $b_v(w) \neq 0 \Leftrightarrow b_w(v) \neq 0$, G is undirected. We call G the *fundamental graph* of S with respect to a . In fact, if S has a graphic presentation (G, a, b) , then G is the fundamental graph of S with respect to a .

Bouchet [2, (7.6)] showed that if (G, a, b) is a graphic presentation of an isotropic system $S = (V, L)$ and $v \in V$, then

$$(G * v, a + b[\{v\}], a[n_G(v)] + b)$$

is also a graphic presentation of S . Thus, local complementations preserve the associated isotropic system. If G and H are locally equivalent, associated isotropic systems can be chosen to be same by an appropriate choice of supplementary vectors. We note that if $uv \in E(G)$, then

$$(G \wedge uv, a[V \setminus \{u, v\}] + b[\{u, v\}], b[V \setminus \{u, v\}] + a[\{u, v\}])$$

is a graphic presentation of S . This fact will be used in section 10.

PROPOSITION 3.6 (see [2, (4.5) and (7.1)]). *If G_1 and G_2 are fundamental graphs of an isotropic system S , then G_1 and G_2 are locally equivalent.*

A minor of an isotropic system is closely related to a vertex-minor of its fundamental graph as follows.

PROPOSITION 3.7 (see [2, (9.1)]). *Let $G = (V, E)$ be a graph. Let $v \in v(G)$ and $x \in K \setminus \{0\}$. If (G, a, b) is a graphic presentation of an isotropic system $S = (V, L)$, then the elementary minor $S|_x^v$ has a graphic presentation, that is,*

- (i) $(G \setminus v, p_{V \setminus \{v\}}(a), p_{V \setminus \{v\}}(b))$ if either $x = a(v)$, or $x = b(v)$ and v is an isolated vertex of G ,
- (ii) $(G \wedge vv \setminus v, p_{V \setminus \{v\}}(a'), p_{V \setminus \{v\}}(b'))$ where $a' = a[V \setminus \{v, w\}] + b[\{v, w\}]$ and $b' = b[V \setminus \{v, w\}] + a[\{v, w\}]$ if $x = b(v)$ and v has a neighbor w in G ,
- (iii) $(G * v \setminus v, p_{V \setminus \{v\}}(a), p_{V \setminus \{v\}}(b + a[n_G(v)]))$ otherwise.

COROLLARY 3.8. *Suppose G_1 and G_2 are fundamental graphs of isotropic systems S_1 and S_2 , respectively. If S_1 is a minor of S_2 , then G_1 is a vertex-minor of G_2 .*

Note that the choice of w in Proposition 3.7 does not affect the isotropic system because of the following proposition, proved in [11].

PROPOSITION 3.9. *If $vv_1, vv_2 \in E(G)$ are two distinct edges incident with v , then $G \wedge vv_1 \wedge v_1 v_2 = G \wedge vv_2$, and therefore $G \wedge vv_1 \setminus v$ is locally equivalent to $G \wedge vv_2 \setminus v$.*

3.3. Connectivity. For a subspace L of K^V , let $\lambda(L) = |V| - \dim(L)$.

DEFINITION 3.10. *For an isotropic system $S = (V, L)$, we call $c : V \rightarrow \mathbb{Z}$ the connectivity function of S if $c(X) = \lambda(L|_{\subseteq X}) = |X| - \dim(L|_{\subseteq X})$.*

If L is a totally isotropic subspace of K^V , then $L|_{\subseteq X}$ is also a totally isotropic subspace of K^X . Thus, $\dim(L|_{\subseteq X}) \leq |X|$, and therefore $c(X) \geq 0$.

Bouchet [3] observed that the connectivity function of an isotropic system is equal to the cut-rank function of its fundamental graph.

PROPOSITION 3.11 (see [3, Theorem 6]). *Let a be an Eulerian vector of an isotropic system $S = (V, L)$ and let c be the connectivity function of S . Let G be the fundamental graph of S with respect to a . Then, $c(X) = \rho_G(X)$ for all $X \subseteq V$.*

By Proposition 3.11, the connectivity function of an isotropic system $S = (V, L)$ is indeed a connectivity function, as defined in section 2. A *branch-decomposition* and the *branch-width* of an isotropic system are defined as a branch-decomposition and the branch-width of its connectivity function, respectively. By Proposition 3.11, the branch-width of an isotropic system is equal to the rank-width of its fundamental graph.

4. Overview of the main proof. Our main objective is to prove the following.

THEOREM 4.1. *Let k be a constant. If $\{G_1, G_2, G_3, \dots\}$ is an infinite sequence of graphs of rank-width at most k , then there exist $i < j$ such that G_i is isomorphic to a pivot-minor of G_j , and therefore isomorphic to a vertex-minor of G_j .*

In general, we say that a binary relation $\leq$ on a set X is a *quasi-order* if it is reflexive and transitive. For a quasi-order $\leq$, we say “ $\leq$ is a *well-quasi-ordering*” or “ X is *well-quasi-ordered* by $\leq$ ” if for every infinite sequence $a_1, a_2, \dots$ of elements of X , there exist $i < j$ such that $a_i \leq a_j$. We may reiterate Theorem 4.1 as follows: a set of graphs of bounded rank-width is well-quasi-ordered up to isomorphism by the vertex-minor relation and also by the pivot-minor relation.

Here is a corollary of Theorem 4.1. Note that this corollary has an elementary proof in [11] and is used in [6] to show the existence of the polynomial-time algorithm to decide whether the rank-width is at most k for a fixed k .

COROLLARY 4.2. *For a fixed k , there is a finite list of graphs $G_1, G_2, \dots, G_m$ such that for every graph H , the rank-width of H is at most k if and only if G_i is not isomorphic to a vertex-minor of H for all i .*

Proof. Let $X = \{G_1, G_2, \dots\}$ be a set of graphs satisfying that for every graph H , the rank-width of a graph H is at most k if and only if G_i is not isomorphic to a vertex-minor of H for all i . We choose X minimal by set inclusion. There are no $G_i, G_j \in X$, such that G_i is isomorphic to a vertex-minor of G_j , because if so, then we may remove G_j from X . By assumption, the rank-width of $G \setminus v$ for $v \in V(G)$ is at most k , and therefore the rank-width of G_i is at most $k + 1$. By Theorem 4.1, X is finite. $\square$

We say that an isotropic system $S_1 = (V_1, L_1)$ is *simply isomorphic* to another isotropic system $S_2 = (V_2, L_2)$ if there exists a bijection $\mu : V_1 \rightarrow V_2$ such that $L_1 = \{a \circ \mu : a \in L_2\}$. A bijection μ is called a *simple isomorphism*. It is clear that if S_1 is simply isomorphic to S_2 , then every fundamental graph of S_1 is isomorphic to a graph locally equivalent to a fundamental graph of S_2 .

We say that an isotropic system S_1 is an $\alpha\beta$ -minor of an isotropic system S if there is $a \in K^X$ with $S_1 = S|_a^X$ and $a(v) \in \{\alpha, \beta\}$ for all $v \in X$. Every $\alpha\beta$ -minor of an isotropic system S is a minor of S , but not vice versa. Pivot-minors of graphs are closely related to $\alpha\beta$ -minors of isotropic systems as follows.

LEMMA 10.2. *For $i \in \{1, 2\}$, let S_i be an isotropic system with a graphic presentation (G_i, a_i, b_i) such that*

$$a_i(v), b_i(v) \in \{\alpha, \beta\}$$

for all $v \in V(G_i)$. If S_1 is an $\alpha\beta$ -minor of S_2 , then G_1 is a pivot-minor of G_2 .

Instead of dealing with graphs, we show the following theorem on isotropic systems.

THEOREM 9.1. *Let k be a constant. If $\{S_1, S_2, S_3, \dots\}$ is an infinite sequence of isotropic systems of branch-width at most k , then there exist $i < j$ such that S_i is simply isomorphic to an $\alpha\beta$ -minor of S_j .*

We deduce Theorem 4.1 from Theorem 9.1.

Proof of Theorem 4.1. Let S_i be the isotropic system with the graphic presentation (G_i, a_i, b_i) where $a_i(v) = \alpha$, $b_i(v) = \beta$ for all $v \in V(G_i)$. Each S_i has branch-width at most k , since its branch-width is equal to the rank-width of G_i . By Theorem 9.1, there exist $i < j$ such that S_i is simply isomorphic to an $\alpha\beta$ -minor of S_j , and therefore by Lemma 10.2, G_i is isomorphic to a pivot-minor of G_j . $\square$

The following theorem was shown by Geelen, Gerards, and Whittle [8]. It was the first step to prove well-quasi-ordering of matroids representable over a fixed finite field having bounded branch-width. Its analogous result by Thomas [15] was used to prove well-quasi-ordering of graphs of bounded tree-width in [13].

THEOREM 4.3 (see [8, Theorem 2.1] and [9]). *A connectivity function with branch-width n has a linked branch-decomposition of width n .*

COROLLARY 4.4. *An isotropic system of branch-width n has a linked branch-decomposition of width n . Equivalently, a graph of rank-width n has a linked rank-decomposition of width n .*

We also use Robertson and Seymour's "lemma on trees," proved in [13]. It enabled them to prove that a set of graphs of bounded tree-width is well-quasi-ordered by the graph minor relation. Geelen, Gerards, and Whittle [8] used it to prove that a set of matroids representable over a fixed finite field and having bounded branch-width is well-quasi-ordered by the matroid minor relation. We need a special case of the "lemma on trees," namely for subcubic forests, that was also used in [8].

The following definitions are as in [8]. A *rooted tree* is a finite directed tree where all but one of the vertices have indegree 1. A *rooted forest* is a collection of countably many vertex-disjoint rooted trees. Its vertices with indegree 0 are called *roots* and those with outdegree 0 are called *leaves*. Edges leaving a root are *root edges* and those entering a leaf are *leaf edges*.

An *n -edge labeling* of a graph F is a map from the set of edges of F to the set $\{0, 1, \dots, n\}$. Let λ be an n -edge labeling of a rooted forest F and let e and f be edges in F . We say that e is λ -linked to f if F contains a directed path P starting with e and ending with f such that $\lambda(g) \geq \lambda(e) = \lambda(f)$ for every edge g on P .

A *binary forest* is a rooted orientation of a subcubic forest with a distinction between left and right outgoing edges. More precisely, we call a triple (F, l, r) a *binary forest* if F is a rooted forest where roots have outdegree 1 and l and r are functions defined on nonleaf edges of F such that the head of each nonleaf edge e of F has exactly two outgoing edges, namely $l(e)$ and $r(e)$.

LEMMA 4.5 (see [8, (3.2) Lemma on Subcubic Trees]). *Let (F, l, r) be an infinite binary forest with an n -edge labeling λ . Moreover, let $\leq$ be a quasi-order on the set of edges of F with no infinite strictly descending sequences such that $e \leq f$ whenever f is λ -linked to e . If the set of leaf edges of F is well-quasi-ordered by $\leq$, but the set of root edges of F is not, then F contains an infinite sequence $(e_0, e_1, \dots)$ of nonleaf edges such that*

- (i) $\{e_0, e_1, \dots\}$ is an antichain with respect to $\leq$,
- (ii) $l(e_0) \leq l(e_1) \leq l(e_2) \leq \dots$,
- (iii) $r(e_0) \leq r(e_1) \leq r(e_2) \leq \dots$.

Informally speaking, at the last stage of proving Theorem 9.1, we need an object describing a piece of isotropic systems such that the number of ways to merge two of such objects into one isotropic system is finite up to simple isomorphisms. More precisely, we call a triple $P = (V, L, B)$ a *scrap* if V is a finite set, L is a totally isotropic subspace of K^V , and B is an ordered basis of $L^\perp/L$. An *ordered basis* is a basis with a linear ordering, and therefore B is of the form $\{b_1 + L, b_2 + L, \dots, b_k + L\}$ with $b_i \in L^\perp$. We denote $V(P) = V$. Note that $L^\perp/L$ is a vector space containing vectors of the form $a + L$ with $a \in L^\perp$ and $a + L = b + L$ if and only if $a - b \in L$. Also note that $|B| = \dim(L^\perp/L) = \dim(L^\perp) - \dim(L) = 2(|V| - \dim(L)) = 2\lambda(L)$.

Two scraps $P_1 = (V, L, B)$ and $P_2 = (V', L', B')$ are called *isomorphic* if there exists a bijection $\mu : V \rightarrow V'$ such that $L = \{a \circ \mu : a \in L'\}$ and $b_i + L = (b'_i \circ \mu) + L$ where $B = \{b_1 + L, b_2 + L, \dots, b_k + L\}$ and $B' = \{b'_1 + L', b'_2 + L', \dots, b'_k + L'\}$.

For $x \in K \setminus \{0\}$ and $v \in V$, let $\delta_x^v \in K^V$ such that $\delta_x^v(v) = x$ and $\delta_x^v(w) = 0$ for all $w \neq v$. We often write δ_x^v without referring to V if that is not ambiguous. If $P = (V, L, B)$ is a scrap and $\delta_x^v \notin L^\perp \setminus L$, we define

$$P|_x^v = (V \setminus \{x\}, L|_x^v, \{p_{V \setminus \{v\}}(b_i) + L|_x^v\}_i),$$

where each $b_i \in L^\perp$ is chosen to satisfy that $B = \{b_i + L\}_i$ and $b_i(v) \in \{0, x\}$. We will prove that $P|_x^v$ is a well-defined scrap in Proposition 6.2. Note that $P|_x^v$ is only defined when $\delta_x^v \notin L^\perp \setminus L$.

A scrap P' is called a *minor* of a scrap P if $P' = P|_{x_1}^{v_1} |_{x_2}^{v_2} \dots |_{x_l}^{v_l}$ for some v_i and x_i . Similarly, a scrap P' is called an $\alpha\beta$ -*minor* of a scrap P if $P' = P|_{x_1}^{v_1} |_{x_2}^{v_2} \dots |_{x_l}^{v_l}$ for some v_i and $x_i \in \{\alpha, \beta\}$.

Two scraps $P_1 = (V, L, B)$ and $P_2 = (V', L', B')$ are called *disjoint* if $V \cap V' = \emptyset$. A scrap $P = (V, L, B)$ is called a *sum* of two disjoint scraps $P_1 = (V_1, L_1, B_1)$ and $P_2 = (V_2, L_2, B_2)$ if

$$V = V_1 \cup V_2, \quad L_1 = L|_{\subseteq V_1}, \quad \text{and} \quad L_2 = L|_{\subseteq V_2}.$$

A sum of two disjoint scraps is not uniquely determined; we, however, will define the *connection types* that will determine a sum of two disjoint scraps such that there are only finitely many connection types. Moreover, we will prove the following.

LEMMA 8.5. *Let P_1, P_2, Q_1, Q_2 be scraps. Let P be the sum of P_1 and P_2 and let Q be the sum of Q_1 and Q_2 . If P_i is a minor of Q_i for $i = 1, 2$ and the connection type of P_1 and P_2 is equal to the connection type of Q_1 and Q_2 , then P is a minor of Q .*

Moreover, if P_i is an $\alpha\beta$ -minor of Q_i for $i \in \{1, 2\}$ and the connection type of P_1 and P_2 is equal to the connection type of Q_1 and Q_2 , then P is an $\alpha\beta$ -minor of Q .

Another requirement to apply Lemma 4.5 is that $e \leq f$ whenever f is λ -linked to e . This condition will be satisfied by the following lemma, which is an extension of Tutte's linking theorem. Tutte's linking theorem for matroids was used by Geelen, Gerards, and Whittle [8] and is an extension of Menger's theorem. Robertson and Seymour also used Menger's theorem in [13].

THEOREM 7.2. *Let V be a finite set and let X be a subset of V . Let L be a totally isotropic subspace of K^V . Let k be a constant. Let b be a complete vector of $K^{V \setminus X}$.*

Then $\lambda(L|_{\subseteq Z}) \geq k$ for all $Z \supseteq X$ if and only if there is a complete vector $a \in K^{V \setminus X}$ such that $\lambda(L|_a^{V \setminus X}) \geq k$ and $a(v) \neq b(v)$ for all $v \in V \setminus X$.

The actual proof of Theorem 9.1 is based on a construction of a forest with a certain k -labeling from branch-decompositions of isotropic systems, and applying the lemmas described above. In the subsequent sections, we will prove these lemmas.

5. Lemmas on totally isotropic subspaces. In this section, L is a totally isotropic subspace of K^V , not necessarily $\dim(L) = |V|$. We prove some general results on totally isotropic subspaces.

LEMMA 5.1. *Let L be a totally isotropic subspace of K^V and $v \in V$, $x \in K \setminus \{0\}$. Then,*

$$(L|_x^v)^\perp = L^\perp|_x^v.$$

Proof. Suppose that $y \in L|_x^v$. There exists $\bar{y} \in L^\perp$ such that $\bar{y}(v) \in \{0, x\}$ and $y = p_{V \setminus \{v\}}(\bar{y})$. For every $z \in L|_x^v$, there exists $\bar{z} \in L$ such that $\bar{z}(v) \in \{0, x\}$ and $p_{V \setminus \{v\}}(\bar{z}) = z$. Since $\langle y, z \rangle = \langle \bar{y}, \bar{z} \rangle - \langle \bar{y}(v), \bar{z}(v) \rangle = 0$, we have $y \in (L|_x^v)^\perp$.

Conversely, suppose that $y \notin L|_x^v$. Let $y \oplus x \in K^V$ be such that $p_{V \setminus \{v\}}(y \oplus x) = y$ and $(y \oplus x)(v) = x$. By assumption, $y \oplus x \notin L^\perp$. Therefore, there exists $z \in L$ such that

$$\langle y, p_{V \setminus \{v\}}(z) \rangle + \langle x, z(v) \rangle = \langle y \oplus x, z \rangle = 1.$$

If $\langle x, z(v) \rangle = 0$, then $p_{V \setminus \{v\}}(z) \in L|_x^v$ and $\langle y, p_{V \setminus \{v\}}(z) \rangle = 1$, and therefore $y \notin (L|_x^v)^\perp$. So, we may assume that $\langle x, z(v) \rangle = 1$.

Let $y \oplus 0 \in K^V$ such that $p_{V \setminus \{v\}}(y \oplus 0) = y$ and $(y \oplus 0)(v) = 0$. By assumption, $y \oplus 0 \notin L^\perp$. Therefore, there exists $w \in L$ such that $\langle y, p_{V \setminus \{v\}}(w) \rangle = \langle y \oplus 0, w \rangle = 1$. If $w(v) \in \{0, x\}$, then $p_{V \setminus \{v\}}(w) \in L|_x^v$ and $y \notin (L|_x^v)^\perp$. Hence we may assume that $\langle x, w(v) \rangle = 1$.

Now, we obtain that $\langle x, w(v) + z(v) \rangle = 0$, and so $w(v) + z(v) \in \{0, x\}$. Therefore $p_{V \setminus \{v\}}(w + z) \in L|_x^v$. Furthermore $\langle p_{V \setminus \{v\}}(w + z), y \rangle = 1$. So, $y \notin (L|_x^v)^\perp$. $\square$

LEMMA 5.2. *If L is a totally isotropic subspace of K^V and $X \subseteq V$, then*

$$(L|_{\subseteq X})^\perp = L^\perp|_X.$$

Proof. We use induction on $|V \setminus X|$. If $|X| < |V| - 1$, then we pick $v \notin X$ and deduce that $(L|_{\subseteq V \setminus \{v\}}|_{\subseteq X})^\perp = (L|_{\subseteq V \setminus \{v\}})^\perp|_X = L^\perp|_{V \setminus \{v\}}|_X = L^\perp|_X$. Therefore we may assume that $V \setminus X = \{v\}$.

For $x \in K^X$ and $y \in K$, we let $x \oplus y$ denote the vector in K^V such that $p_X(x \oplus y) = x$ and $(x \oplus y)(v) = y$.

(1) We claim that $L^\perp|_X \subseteq (L|_{\subseteq X})^\perp$.

Suppose that $a \in L^\perp|_X$. Then there exists $b \in K$ such that $a \oplus b \in L^\perp$. For any $c \in L|_{\subseteq X}$, $\langle a \oplus b, c \oplus 0 \rangle = 0$, and therefore $\langle a, c \rangle = 0$. Thus, $a \in (L|_{\subseteq X})^\perp$.

(2) We claim that $(L|_{\subseteq X})^\perp \subseteq L^\perp|_X$.

Suppose that $a \in (L|_{\subseteq X})^\perp$ and $a \notin L^\perp|_X$. We have $a \oplus x \notin L^\perp$ and therefore for every $x \in K$, there exists $a_x \oplus c_x \in L$ such that $\langle a_x, a \rangle + \langle c_x, x \rangle = \langle a_x \oplus c_x, a \oplus x \rangle = 1$. In particular, $\langle a_0, a \rangle = 1$.

If $c_x = 0$, then $a_x \in L|_{\subseteq X}$ and so $\langle a_x, a \rangle = 0$ and $\langle c_x, x \rangle = 0$, contrary to the fact that $\langle a_x, a \rangle + \langle c_x, x \rangle = 1$. Therefore $c_x \neq 0$ for all $x \in K$.

If $c_x = c_y$ for $x \neq y$, then $a_x + a_y \in L|_{\subseteq X}$. Thus, $\langle c_x, x + y \rangle = 1 + \langle c_x, x \rangle + 1 + \langle c_y, y \rangle = \langle a_x + a_y, a \rangle = 0$. Since $c_x \neq 0$ and $x + y \neq 0$, we have $c_x = c_y = x + y$ and $\langle a_x, a \rangle = 1 + \langle x + y, x \rangle = 1 + \langle x, y \rangle$.

If $c_x = c_y = c_z$ for distinct x, y, z , then $x + y = y + z = z + x$. So, $x = y = z$, which is a contradiction.

If $c_x = c_y$, $c_z = c_w$ for distinct x, y, z, w , then $c_x = x + y = z + w = c_z$. So, $x = y = z = w$. This is a contradiction.

Therefore, there is exactly one pair $x, y \in K$ such that $c_x = c_y$. Let $\{z, w\} = K \setminus \{x, y\}$.

Since $c_z \neq c_w$ and $c_z, c_w \in K \setminus \{0, x + y\}$, we have $c_z + c_w = x + y = c_x = c_y$. Therefore, $a_z + a_w + a_x \in L|_{\subseteq X}$ and $\langle a_z + a_w + a_x, a \rangle = 0$. Since $\langle a_z, a \rangle + \langle a_w, a \rangle = \langle c_z, z \rangle + \langle c_w, w \rangle$, we have

$$\langle c_z, z \rangle + \langle c_w, w \rangle = \langle a_z + a_w + a_x, a \rangle + \langle a_x, a \rangle = 1 + \langle x, y \rangle.$$

If $x = 0$, then $c_z + c_w = y$, hence $c_z, c_w \in \{z, w\}$. So, $\langle c_z, z \rangle + \langle c_w, w \rangle = 0$. A contradiction.

So we may assume that $x \neq 0, y \neq 0, z = 0$, and then $x + y = w$ and $\langle c_w, w \rangle = 0$. But, this implies that $c_w = w = x + y = c_x$. A contradiction. $\square$

PROPOSITION 5.3. *Let V be a finite set and let L be a totally isotropic subspace of K^V and $v \in V, x \in K \setminus \{0\}$. Then,*

$$\dim(L|_x^v) = \begin{cases} \dim(L) & \text{if } \delta_x^v \in L^\perp \setminus L, \\ \dim(L) - 1 & \text{otherwise.} \end{cases}$$

In other words,

$$\lambda(L|_x^v) = \begin{cases} \lambda(L) & \text{if } \delta_x^v \notin L^\perp \setminus L, \\ \lambda(L) - 1 & \text{otherwise.} \end{cases}$$

Proof. Let $L' = \{a \in L : a(v) \in \{0, x\}\}$. Then clearly $L' = L$ if $a(v) \in \{0, x\}$ for each $a \in L$, and $\dim(L') = \dim(L) - 1$ otherwise. Also, $\dim(L|_x^v) = \dim(L')$ when $\delta_x^v \notin L'$, and $\dim(L|_x^v) = \dim(L') - 1$ otherwise. Moreover, obviously, “ $a(v) \in \{0, x\}$ for each $a \in L$ ” is equivalent to “ $\delta_x^v \in L^\perp$ ”; and “ $\delta_x^v \notin L'$ ” is equivalent to “ $\delta_x^v \notin L$.” As $L \subseteq L^\perp$, all this implies this proposition. $\square$

COROLLARY 5.4. *Let V be a finite set and let L be a totally isotropic subspace of K^V and $v \in V$. Let $C \subseteq K \setminus \{0\}$, $|C| = 2$. Then, either there is $x \in C$ such that $\lambda(L|_x^v) = \lambda(L)$ or for all $y \in K \setminus \{0\}$,*

$$L|_y^v = L|_{\subseteq V \setminus \{v\}} \quad \text{and} \quad \lambda(L|_y^v) = \lambda(L) - 1.$$

Proof. Let $C = \{a, b\}$. Suppose there is no such $x \in C$. Then δ_a^v, δ_b^v are in $L^\perp \setminus L$. Therefore, $z(v) = 0$ for all $z \in L$. Thus, $L|_y^v = L|_{\subseteq V \setminus \{v\}}$ and $\lambda(L|_y^v) = \lambda(L) - 1$ for all $y \in K \setminus \{0\}$. $\square$

6. Scraps. In this section, we prove that a minor of a scrap is well defined. From section 4, we recall that a *scrap* is a triple $P = (V, L, B)$ of a finite set V , a totally isotropic subspace L of K^V , and an ordered basis B of $L^\perp/L$.

LEMMA 6.1. *Let $P = (V, L, B)$ be a scrap and $v \in V, x \in K \setminus \{0\}$. If $\delta_x^v \notin L^\perp \setminus L$, then there is a sequence $b_1, b_2, \dots, b_m \in L^\perp$ such that $b_i(v) \in \{0, x\}$ and $B = \{b_1 + L, b_2 + L, \dots, b_m + L\}$.*

Proof. Let $B = \{a_1 + L, a_2 + L, \dots, a_m + L\}$ with $a_i \in L^\perp$. If $\delta_x^v \in L$, then $a_i(v) \in \{0, x\}$ for all i . Hence we may assume that $\delta_x^v \notin L$ and so $\delta_x^v \notin L^\perp$. Therefore there is $y \in L$ such that $\langle y, \delta_x^v \rangle = 1$. Thus, $y(v) \notin \{0, x\}$. Let

$$b_i = \begin{cases} a_i & \text{if } a_i(v) \in \{0, x\}, \\ a_i + y & \text{otherwise.} \end{cases}$$

Then, $b_i + L = a_i + L$ and $b_i(v) \in \{0, x\}$. $\square$

PROPOSITION 6.2. *Let $P = (V, L, B)$ be a scrap. If $\delta_x^v \notin L^\perp \setminus L$, then $P|_x^v$ is well defined and is a scrap.*

Proof. Let us first show that it is well defined. Let $b_1, b_2, \dots, b_k \in L^\perp$ be such that $b_i(v) \in \{0, x\}$ and $B = \{b_i + L : i = 1, 2, \dots, k\}$. We claim that the choice of b_i does not change $P|_x^v$. Suppose $b_i - b'_i \in L$ and $b_i(v), b'_i(v) \in \{0, x\}$. Since $b_i - b'_i \in L$ and $(b_i - b'_i)(v) \in \{0, x\}$, we have $p_{V \setminus \{v\}}(b_i - b'_i) \in L|_x^v$. Therefore, $p_{V \setminus \{v\}}(b_i) + L|_x^v = p_{V \setminus \{v\}}(b'_i) + L|_x^v$.

Now, we claim that $P|_x^v$ is a scrap.

First, we show that $L|_x^v$ is a totally isotropic subspace of $K^{V \setminus \{v\}}$. For all $a, b \in L|_x^v$, there are $\bar{a}, \bar{b} \in L$ such that $\bar{a}(v), \bar{b}(v) \in \{0, x\}$, $p_{V \setminus \{v\}}(\bar{a}) = a$, $p_{V \setminus \{v\}}(\bar{b}) = b$, and $\bar{a}, \bar{b} \in L$. Hence $\langle a, b \rangle = \langle \bar{a}, \bar{b} \rangle = 0$.

Next, we show that $\{p_{V \setminus \{v\}}(b_i) + L|_x^v : i = 1, 2, \dots, k\}$ is a basis of $(L|_x^v)^\perp / (L|_x^v)$. Since $b_i(v) \in \{0, x\}$, we have $p_{V \setminus \{v\}}(b_i) \in (L|_x^v)^\perp = (L^\perp)|_x^v$. Suppose that there exists $C \neq \emptyset$ such that

$$\sum_{i \in C} (p_{V \setminus \{v\}}(b_i) + L|_x^v) = 0 + L|_x^v.$$

Since $\sum_{i \in C} p_{V \setminus \{v\}}(b_i) \in L|_x^v$, there exists $z \in L \subseteq L^\perp$ such that $z(v) \in \{0, x\}$ and $p_{V \setminus \{v\}}(z) = \sum_{i \in C} p_{V \setminus \{v\}}(b_i)$. By assumption, $\sum_{i \in C} b_i \notin L$. Since $p_{V \setminus \{v\}}(\sum_{i \in C} b_i - z) = 0$, we have $\sum_{i \in C} b_i - z = \delta_x^v \in L^\perp \setminus L$, a contradiction. Therefore, $\{p_{V \setminus \{v\}}(b_i) + L|_x^v : i = 1, 2, \dots, k\}$ is linearly independent. Moreover, $\dim((L|_x^v)^\perp / (L|_x^v)) = 2(|V| - 1 - \dim(L|_x^v)) = 2(|V| - \dim(L)) = \dim(L^\perp / L)$ because $\delta_x^v \notin L^\perp \setminus L$. $\square$

7. Generalization of Tutte's linking theorem. In this section, we show a generalization of Tutte's linking theorem [16].

LEMMA 7.1. *Let V be a finite set and $v \in V$. Let L be a totally isotropic subspace of K^V . Let $X_1, Y_1 \subseteq V \setminus \{v\}$. Let $x, y \in K \setminus \{0\}$, $x \neq y$. Then,*

$$\dim(L|_{\subseteq X_1 \cap Y_1}) + \dim(L|_{\subseteq X_1 \cup Y_1 \cup \{v\}}) \geq \dim(L|_x^v|_{\subseteq X_1}) + \dim(L|_y^v|_{\subseteq Y_1}).$$

In other words,

$$\lambda(L|_x^v|_{\subseteq X_1}) + \lambda(L|_y^v|_{\subseteq Y_1}) \geq \lambda(L|_{\subseteq X_1 \cap Y_1}) + \lambda(L|_{\subseteq X_1 \cup Y_1 \cup \{v\}}) - 1.$$

Proof. We may assume that $V = X_1 \cup Y_1 \cup \{v\}$ by taking $L' = L|_{\subseteq X \cup Y \cup \{v\}}$.

Let B be a minimum set of vectors in L such that $p_{X_1 \cap Y_1}(B)$ is a basis of $L|_{\subseteq X_1 \cap Y_1}$ and for every $z \in B$, $z(w) = 0$ for all $w \notin X_1 \cap Y_1$.

Let C be a minimum set of vectors in L such that $p_{X_1}(B \cup C)$ is a basis of $L|_x^v|_{\subseteq X_1}$ and for every $z \in C$, $z(w) = 0$ for all $w \notin X_1 \cup \{v\}$ and $z(v) \in \{0, x\}$. We may assume that at most one vector in C has x on v .

Let D be a minimum set of vectors in L such that $p_{Y_1}(B \cup D)$ is a basis of $L|_y^v|_{\subseteq Y_1}$ and for every $z \in D$, $z(w) = 0$ for all $w \notin Y_1 \cup \{v\}$ and $z(v) \in \{0, y\}$. We may assume that at most one vector in D has y on v .

We claim that $B \cup C \cup D$ is linearly independent. Suppose there are $B' \subseteq B$, $C' \subseteq C$, and $D' \subseteq D$ such that

$$\sum_{b \in B'} b + \sum_{c \in C'} c + \sum_{d \in D'} d = 0.$$

No element of C' has x on v , because all vectors in $B' \cup D'$ have 0 or y on v and at most one element of C' has x on v . Since $\sum_{c \in C'} c(w) = 0$ for all $w \in V \setminus (X_1 \cap Y_1)$,

we have $p_{X_1 \cap Y_1}(\sum_{c \in C'} c) \in L|_{\subseteq X_1 \cap Y_1}$. Since $p_{X_1 \cap Y_1}(B)$ is a basis, there is $B'' \subseteq B$ such that $p_{X_1 \cap Y_1}(\sum_{c \in C'} c) = p_{X_1 \cap Y_1}(\sum_{b \in B''} b)$. So,

$$\sum_{c \in C'} c + \sum_{b \in B''} b = 0.$$

This means that $C' = \emptyset$ because $C \cup B$ is a basis. Similarly $D' = \emptyset$ and so $B' = \emptyset$. Hence $B \cup C \cup D$ is linearly independent indeed. So $\dim(L) \geq |B| + |C| + |D| = \dim(L|_{\subseteq X_1}^v) + \dim(L|_{\subseteq Y_1}^v) - \dim(L|_{\subseteq X_1 \cap Y_1})$. $\square$

Now, we translate Tutte's linking theorem into isotropic subspaces.

THEOREM 7.2. *Let V be a finite set and let X be a subset of V . Let L be a totally isotropic subspace of K^V . Let k be a constant. Let b be a complete vector of $K^{V \setminus X}$.*

Then $\lambda(L|_{\subseteq Z}) \geq k$ for all $Z \supseteq X$ if and only if there is a complete vector $a \in K^{V \setminus X}$ such that $\lambda(L|_a^{V \setminus X}) \geq k$ and $a(v) \neq b(v)$ for all $v \in V \setminus X$.

Proof. ($\Leftarrow$) Let Z be a subset of V such that $X \subseteq Z$. Let $a_1 = p_{V \setminus Z}(a)$, $a_2 = p_{Z \setminus X}(a)$. Since $L|_{\subseteq Z} \subseteq L|_{a_1}^{V \setminus Z}$, we have $\lambda(L|_{\subseteq Z}) \geq \lambda(L|_{a_1}^{V \setminus Z})$. So

$$\lambda(L|_{\subseteq Z}) \geq \lambda(L|_{a_1}^{V \setminus Z}) \geq \lambda(L|_{a_1}^{V \setminus Z}|_{a_2}^{Z \setminus X}) = \lambda(L|_a^{V \setminus X}) \geq k.$$

($\Rightarrow$) We use induction on $|V \setminus X|$. Suppose that there is no such complete vector $a \in K^{V \setminus X}$. We may assume that $|V \setminus X| \geq 1$.

Pick $v \in V \setminus X$. Let $K \setminus \{0, b(v)\} = \{x, y\}$. Since there is no complete vector $a' \in K^{V \setminus \{v\} \setminus X}$ such that $\lambda(L|_{a'}^{V \setminus \{v\} \setminus X}) \geq k$, there exists X_1 such that $X \subseteq X_1 \subseteq V \setminus \{v\}$ and $\lambda(L|_{\subseteq X_1}^v) < k$. Similarly, there exists Y_1 such that $X \subseteq Y_1 \subseteq V \setminus \{v\}$ and $\lambda(L|_{\subseteq Y_1}^v) < k$. By Lemma 7.1, either $\lambda(L|_{\subseteq X_1 \cap Y_1}) < k$ or $\lambda(L|_{\subseteq X_1 \cup Y_1 \cup \{v\}}) < k$, a contradiction. $\square$

COROLLARY 7.3. *Let V be a finite set and let X be a subset of V . Let L be a totally isotropic subspace of K^V . Let b be a complete vector of $K^{V \setminus X}$.*

If $\lambda(L|_{\subseteq Z}) \geq \lambda(L|_{\subseteq X})$ for all $Z \supseteq X$, then there is a complete vector $a \in K^{V \setminus X}$ such that $L|_a^{V \setminus X} = L|_{\subseteq X}$ and $a(v) \neq b(v)$ for all $v \in V \setminus X$.

Proof. By Theorem 7.2, there exists a complete vector $a \in K^{V \setminus X}$ such that

$$\lambda(L|_a^{V \setminus X}) = \lambda(L|_{\subseteq X}) \text{ and } a(v) \neq b(v) \text{ for all } v \in V \setminus X.$$

Since $L|_{\subseteq X} \subseteq L|_a^{V \setminus X}$ and $\dim(L|_{\subseteq X}) = \dim(L|_a^{V \setminus X})$, we have $L|_{\subseteq X} = L|_a^{V \setminus X}$. $\square$

COROLLARY 7.4. *Let $P = (V, L, B)$ be a scrap and $X \subseteq V$. If*

$$\lambda(P) = \lambda(L|_{\subseteq X}) = \min_{X \subseteq Z \subseteq V} \lambda(L|_{\subseteq Z}),$$

then there is an ordered set B' such that $Q = (X, L|_{\subseteq X}, B')$ is a scrap and an $\alpha\beta$ -minor of P .

Proof. By applying Corollary 7.3 with $b(v) = \gamma$ for all $v \in V \setminus X$, there is a complete vector $a \in K^{V \setminus X}$ such that $L|_a^{V \setminus X} = L|_{\subseteq X}$ and $a(v) \in \{\alpha, \beta\}$ for all $v \in V \setminus X$. Let $V \setminus X = \{y_1, y_2, \dots, y_m\}$ and $a_i = a(y_i)$. Then, $L|_{\subseteq X} = L|_{a_1}^{y_1}|_{a_2}^{y_2} \dots |_{a_m}^{y_m}$. Let $L_0 = L$ and $L_i = L_{i-1}|_{a_i}^{y_i}$. By Proposition 5.3, $\lambda(L|_{\subseteq X}) = \lambda(L) = \lambda(L_i)$ implies $\delta_{a_{i+1}}^{y_{i+1}} \notin L_i^\perp \setminus L_i$. So, $P|_{a_1}^{y_1}|_{a_2}^{y_2} \dots |_{a_m}^{y_m} = (X, L|_{\subseteq X}, B')$ is well defined and is an $\alpha\beta$ -minor of P . $\square$

8. Sum. A scrap $P = (V, L, B)$ is called a *sum* of two disjoint scraps $P_1 = (V_1, L_1, B_1)$ and $P_2 = (V_2, L_2, B_2)$ if

$$V = V_1 \cup V_2, \quad L_1 = L|_{\subseteq V_1}, \text{ and } L_2 = L|_{\subseteq V_2}.$$

For given two disjoint scraps, there could be many scraps that are sums of those. In this section, we define the *connection type*, which determines a sum uniquely.

Let $[n]$ denote the set $\{1, 2, 3, \dots, n\}$.

DEFINITION 8.1. Let $P = (V, L, B)$ be a sum of two disjoint scraps $P_1 = (V_1, L_1, B_1)$ and $P_2 = (V_2, L_2, B_2)$ where $B = \{b_1 + L, b_2 + L, \dots, b_n + L\}$, $B_1 = \{b_1^1 + L_1, b_2^1 + L_1, \dots, b_m^1 + L_1\}$, and $B_2 = \{b_1^2 + L_2, b_2^2 + L_2, \dots, b_l^2 + L_2\}$. For $x_1 \in K^{V_1}$ and $x_2 \in K^{V_2}$, let $x_1 \oplus x_2$ denote the vector in K^V such that $p_{V_i}(x_1 \oplus x_2) = x_i$ for $i = 1, 2$. Let

$$C_0 = \left\{ (X, Y) : X \subseteq [m], Y \subseteq [l], \left(\sum_{i \in X} b_i^1 \right) \oplus \left(\sum_{j \in Y} b_j^2 \right) \in L \right\},$$

$$C_s = \left\{ (X, Y) : X \subseteq [m], Y \subseteq [l], \left(\sum_{i \in X} b_i^1 \right) \oplus \left(\sum_{j \in Y} b_j^2 \right) - b_s \in L \right\}, \quad s = 1, \dots, n.$$

The sequence $C(P, P_1, P_2) = (C_0, C_1, C_2, \dots, C_n)$ is called the *connection type* of this sum.

It is easy to see that if $\lambda(P), \lambda(P_1), \lambda(P_2) \leq k$, then the number of distinct connection types is bounded by a function of k , because $|B| = 2\lambda(P) \leq 2k$ and $|B_i| = 2\lambda(P_i) \leq 2k$ for $i = 1$ and 2 .

PROPOSITION 8.2. The connection type is well defined.

Proof. It is enough to show that the choice of b_i , b_i^1 , and b_i^2 does not affect C_i . Suppose $b_i + L = d_i + L$, $b_i^1 + L_1 = d_i^1 + L_1$, and $b_i^2 + L_2 = d_i^2 + L_2$. For any (X, Y) such that $X \subseteq [m]$, $Y \subseteq [l]$, we have $\sum_{i \in X} (b_i^1 - d_i^1) \oplus \sum_{j \in Y} (b_j^2 - d_j^2) \in L$ and $b_s - d_s \in L$, and therefore C_0 and C_s are well defined. $\square$

PROPOSITION 8.3. The connection type uniquely determines the sum of two disjoint scraps P_1 and P_2 .

Proof. Suppose not. Let $P = (V, L, B)$, $Q = (V, L', B')$ be two distinct sums of $P_1 = (V_1, L_1, B_1)$ and $P_2 = (V_2, L_2, B_2)$ by the same connection type $(C_0, C_1, \dots, C_n)$. Let $B_1 = \{b_1^1 + L_1, b_2^1 + L_1, \dots, b_m^1 + L_1\}$ and $B_2 = \{b_1^2 + L_2, b_2^2 + L_2, \dots, b_k^2 + L_2\}$.

We claim that $L = L'$. To show this, it is enough to show that $L \subseteq L'$. For any $a \in L$, $p_{V_1}(a) \in (L|_{\subseteq V_1})^\perp$ and $p_{V_2}(a) \in (L|_{\subseteq V_2})^\perp$. Therefore there is (X, Y) such that

$$x_1 = \sum_{i \in X} b_i^1 - p_{V_1}(a) \in L_1 \quad \text{and} \quad x_2 = \sum_{i \in Y} b_i^2 - p_{V_2}(a) \in L_2.$$

Since $x_1 \oplus 0, 0 \oplus x_2 \in L$, we have $x_1 \oplus x_2 \in L$. We deduce that $\sum_{i \in X} b_i^1 \oplus \sum_{i \in Y} b_i^2 = a + (x_1 \oplus x_2) \in L$. Therefore, $(X, Y) \in C_0$ and $a + (x_1 \oplus x_2) \in L'$. Since $x_1 \oplus 0, 0 \oplus x_2 \in L'$, we have $x_1 \oplus x_2 \in L'$, and so $a \in L'$.

Now, we show that $B = B'$. Let $b_j + L$ be the j th element of B with $b_j \in L^\perp$. Let $b'_j + L$ be the j th element of B' with $b'_j \in L^\perp$. Since $p_{V_i}(b_j) \in (L|_{\subseteq V_i})^\perp$, there is (X, Y) such that

$$x_1 = \sum_{i \in X} b_i^1 - p_{V_1}(b_j) \in L_1 \quad \text{and} \quad x_2 = \sum_{i \in Y} b_i^2 - p_{V_2}(b_j) \in L_2.$$

Since $x_1 \oplus 0, 0 \oplus x_2 \in L$, we have $x_1 \oplus x_2 \in L$, and therefore $\sum_{i \in X} b_i^1 \oplus \sum_{i \in Y} b_i^2 - b_j \in L$. Thus, $(X, Y) \in C_s$, and so

$$\sum_{i \in X} b_i^1 \oplus \sum_{i \in Y} b_i^2 - b'_j \in L' = L.$$

Thus, $b_j + L = b'_j + L = b'_j + L'$. $\square$

PROPOSITION 8.4. *Let $P_1 = (V_1, L_1, B_1)$, $P_2 = (V_2, L_2, B_2)$ be two disjoint scraps. Let P be the sum of P_1 and P_2 by connection type $C(P, P_1, P_2)$. If $v \in V_1$ and $\delta_x^v \notin L_1^\perp \setminus L_1$, then $\delta_x^v \notin L^\perp \setminus L$ and $P|_x^v$ is the sum of $P_1|_x^v$ and P_2 by connection type $C(P, P_1, P_2)$.*

Proof. If $\delta_x^v \in L^\perp \setminus L$, then $\delta_x^v \in (L^\perp)|_{V_1} = (L|_{\subseteq V_1})^\perp = L_1^\perp$ and $\delta_x^v \notin L|_{\subseteq V_1}$. This contradicts $\delta_x^v \notin L_1^\perp \setminus L_1$. So, $\delta_x^v \notin L^\perp \setminus L$.

First, we claim that $P|_x^v$ is a sum of $P_1|_x^v$ and P_2 . This is equivalent to

$$L|_x^v|_{\subseteq V_1 \setminus \{v\}} = L|_{\subseteq V_1}|_x^v \quad \text{and} \quad L|_x^v|_{\subseteq V_2} = L|_{\subseteq V_2}.$$

So, as clearly $L|_x^v|_{\subseteq V_1 \setminus \{v\}} = L|_{\subseteq V_1}|_x^v$ and $L|_{\subseteq V_2} \subseteq L|_x^v|_{\subseteq V_2}$, it suffices to show that

$$L|_x^v|_{\subseteq V_2} \subseteq L|_{\subseteq V_2}.$$

Suppose $z \in L|_x^v|_{\subseteq V_2}$. Let $\bar{z} \in K^V$ such that $p_{V_2}(\bar{z}) = z$, $\bar{z}(v) \in \{0, x\}$, and $p_{V_1 \setminus \{v\}}(\bar{z}) = 0$. If $\bar{z}(v) = 0$, then $z \in L|_{\subseteq V_2}$. If $\bar{z}(v) = x$, then $p_{V_1}(z) = \delta_x^v \in L^\perp|_{V_1} = L_1^\perp$, and therefore $\delta_x^v \in L_1$. So, $\delta_x^v \in L$ and $z + \delta_x^v \in L$. Since $(z + \delta_x^v)(v) = 0$, we have $z \in L|_{\subseteq V_2}$. This proves that $P|_x^v$ is a sum of $P_1|_x^v$ and P_2 .

Now, let us show that $C(P, P_1, P_2) = C(P|_x^v, P_1|_x^v, P_2)$. Let $B_1 = \{b_1^1 + L_1, b_2^1 + L_1, \dots, b_m^1 + L_1\}$ and $B_2 = \{b_1^2 + L_2, b_2^2 + L_2, \dots, b_k^2 + L_2\}$. For $s \in K^{V_1}$ and $t \in K^{V_2}$, let $s \oplus t$ denote the vector in K^V such that $p_{V_1}(s \oplus t) = s$ and $p_{V_2}(s \oplus t) = t$. We may assume that $b_i^1(v) \in \{0, x\}$ for all i by Lemma 6.1. Let $b \in L^\perp$ be such that $b(v) \in \{0, x\}$. Let $a(X, Y) = (\sum_{i \in X} b_i^1) \oplus (\sum_{j \in Y} b_j^2) - b$. Suppose we have (X, Y) such that $X \subseteq [m]$, $Y \subseteq [k]$, and $a(X, Y) \in L$. Since $(\sum_{i \in X} b_i^1(v)) - b(v) \in \{0, x\}$, we have

$$p_{V \setminus \{v\}}(a(X, Y)) = \left(\sum_{i \in X} p_{V_1 \setminus \{v\}}(b_i^1) \right) \oplus \left(\sum_{j \in Y} b_j^2 \right) - p_{V \setminus \{v\}}(b) \in L|_x^v.$$

Conversely, let us suppose that there is (X, Y) such that $X \subseteq [m]$, $Y \subseteq [k]$, and $(\sum_{i \in X} p_{V \setminus \{v\}}(b_i^1)) \oplus (\sum_{j \in Y} b_j^2) - p_{V \setminus \{v\}}(b) \in L|_x^v$. Then, either $a(X, Y) \in L$ or $a(X, Y) + \delta_x^v \in L$. If $\delta_x^v \in L$, then $a(X, Y) \in L$. If $\delta_x^v \notin L^\perp$, then $a(X, Y) + \delta_x^v \notin L^\perp$ because $a(X, Y) \in L^\perp$, and therefore $a(X, Y) \in L$. This proves that $C(P, P_1, P_2) = C(P|_x^v, P_1|_x^v, P_2)$. $\square$

LEMMA 8.5. *Let P_1, P_2, Q_1, Q_2 be scraps. Let P be a sum of P_1 and P_2 and let Q be a sum of Q_1 and Q_2 . If P_i is a minor of Q_i for each $i = 1, 2$ and the connection type of P_1 and P_2 is equal to the connection type of Q_1 and Q_2 , then P is a minor of Q .*

Moreover, if P_i is an $\alpha\beta$ -minor of Q_i for each $i \in \{1, 2\}$ and the connection type of P_1 and P_2 is equal to the connection type of Q_1 and Q_2 , then P is an $\alpha\beta$ -minor of Q .

Proof. We use induction on $|V(Q_1) \setminus V(P_1)| + |V(Q_2) \setminus V(P_2)|$. We may assume $|V(Q_1) \setminus V(P_1)| + |V(Q_2) \setminus V(P_2)| > 0$ and $V(Q_1) \neq V(P_1)$ by symmetry. There are $v \in V(Q_1) \setminus V(P_1)$, $x \in K \setminus \{0\}$, $X = V(Q_1) \setminus V(P_1) \setminus \{v\}$, and a complete vector $a \in K^X$ such that $P_1 = Q_1|_x^v|_a^X$. If P_1 is an $\alpha\beta$ -minor of Q_1 , then we may assume $x \in \{\alpha, \beta\}$ and $a(w) \in \{\alpha, \beta\}$ for all $w \in X$.

By Proposition 8.4, $Q|_x^v$ is the sum of $Q_1|_x^v$ and Q_2 with the connection type $C(Q|_x^v, Q_1|_x^v, Q_2) = C(Q, Q_1, Q_2) = C(P, P_1, P_2)$. So, P is a minor of $Q|_x^v$ by induction and therefore P is a minor of Q .

Similarly, if P_1 is an $\alpha\beta$ -minor of Q_1 and P_2 is an $\alpha\beta$ -minor of Q_2 , then by induction P is an $\alpha\beta$ -minor of Q . $\square$

9. Well-quasi-ordering. We are finally ready to prove our main theorem on isotropic systems.

THEOREM 9.1. *Let k be a constant. If $\{S_1, S_2, S_3, \dots\}$ is an infinite sequence of isotropic systems of branch-width at most k , then there exist $i < j$ such that S_i is simply isomorphic to an $\alpha\beta$ -minor of S_j .*

Proof. By Corollary 4.4, there is a linked branch-decomposition $(T_i, \mathcal{L}_i)$ of $S_i = (V_i, L_i)$ of width at most k for each i . In T_i , we pick an edge and attach a root and direct every edge so that each leaf has a directed path from the root. Let F be the forest such that the i th component is T_i .

For each edge e of T_i , let X_e be the set of leaves of T_i having a directed path from e . Let $A_e = \mathcal{L}_i^{-1}(X_e)$. We associate e with a scrap $P_e = (A_e, L_i|_{\subseteq A_e}, B_e)$ and $\lambda(e) = \lambda(L_i|_{\subseteq A_e}) \leq k$ where B_e is chosen to satisfy the following: If f is λ -linked to e , then P_e is an $\alpha\beta$ -minor of P_f . We claim that we can choose B_e satisfying the above property. We prove it by induction on the length of the directed path from a root edge to e . If no other edge is λ -linked to e , let B_e be a basis of $(L_i|_{\subseteq A_e})^\perp / (L_i|_{\subseteq A_e})$ in an arbitrary order. If f , other than e , is λ -linked to e , choose f such that the distance between e and f is minimal. By induction, there exists B_f so that whenever g is λ -linked to f , P_f is an $\alpha\beta$ -minor of P_g . By Corollary 7.4, there is an ordered basis B_e such that P_e is an $\alpha\beta$ -minor of P_f . Suppose that g is λ -linked to e . Then f is on the path from g to e and therefore g is λ -linked to f . Thus, P_f is an $\alpha\beta$ -minor of P_g and therefore P_e is an $\alpha\beta$ -minor of P_g . We conclude that there is B_e satisfying the property.

For $e, f \in E(F)$, let $e \leq f$ denote that a scrap P_e is isomorphic to an $\alpha\beta$ -minor of a scrap P_f . Clearly, $\leq$ has no infinite strictly descending sequences, because there are finitely many scraps of bounded number of elements up to isomorphisms. By construction if f is λ -linked to e , then $e \leq f$.

The leaf edges of F are well-quasi-ordered, because there are only finitely many distinct scraps of one element up to isomorphisms.

Suppose the root edges are not well-quasi-ordered. By Lemma 4.5, F contains an infinite sequence $(e_0, e_1, \dots)$ of nonleaf edges such that

- (i) $\{e_0, e_1, \dots\}$ is an antichain with respect to $\leq$,
- (ii) $l(e_0) \leq l(e_1) \leq \dots$,
- (iii) $r(e_0) \leq r(e_1) \leq \dots$.

Since $\lambda(e_i) \leq k$ for all i , we may assume that $\lambda(e_i)$ is a constant for all i , by taking a subsequence.

Since the number of distinct connection types $C(P_{e_i}, P_{l(e_i)}, P_{r(e_i)})$ is finite, we may assume that the connection types are the same for all i , also by taking a subsequence.

Then, by Lemma 8.5, P_{e_0} is isomorphic to an $\alpha\beta$ -minor of P_{e_1} , that is, $e_0 \leq e_1$. This contradicts that $\{e_0, e_1, \dots\}$ is an antichain with respect to $\leq$.

Therefore, root edges are well-quasi-ordered, and there exist $i < j$ such that a scrap $(V_i, L_i, \emptyset)$ is isomorphic to an $\alpha\beta$ -minor of a scrap $(V_j, L_j, \emptyset)$. Thus, S_i is simply isomorphic to an $\alpha\beta$ -minor of S_j . $\square$

10. Pivot-minor and $\alpha\beta$ -minor. In this section, we shall show a relation between pivot-minors of graphs and $\alpha\beta$ -minors of isotropic systems.

PROPOSITION 10.1. *For $i \in \{1, 2\}$, let S_i be an isotropic system with a graphic presentation (G_i, a_i, b_i) such that $a_i(v), b_i(v) \in \{\alpha, \beta\}$ for all $v \in V(G_i)$.*

If $S_1 = S_2$, then G_1 can be obtained from G_2 by applying a sequence of pivots.

Proof. Let $V = V(G_1) = V(G_2)$ and let $S = S_1 = S_2 = (V, L)$ be an isotropic system. We show Proposition 10.1 by induction on $N(a_1, a_2) = |\{v \in V : a_1(v) \neq a_2(v)\}|$.

If $N(a_1, a_2) = 0$, then $b_1 = b_2$, so $G_1 = G_2$. Hence we may assume that $N(a_1, a_2) \geq 1$. Let $u \in V$ with $a_1(u) \neq a_2(u)$. This implies that $a_1(u) = b_2(u)$ because $a_i(u), b_i(u) \in \{\alpha, \beta\}$. By Proposition 3.5, there exists a vector $\hat{b}_u \in L$ such that $\hat{b}_u(u) = b_2(u) = a_1(u)$ and $\hat{b}_u(w) \in \{0, a_2(w)\}$ for all $w \neq u$. As a_1 is Eulerian and as $\hat{b}_u \in L$, there exists $v \in V$ such that $\hat{b}_u(v) \notin \{0, a_1(v)\}$. Hence $\hat{b}_u(v) = a_2(v)$ and $a_1(v) \neq a_2(v)$.

Since $\hat{b}_u(u) \neq a_2(u)$, this means that $v \neq u$. Moreover, since $\hat{b}_u(v) \neq 0$, the vertices u and v are adjacent in G_2 . Pivoting (G_2, a_2, b_2) on uv yields $(G_2 \wedge uv, a_2[V \setminus \{u, v\}] + b_2[\{u, v\}], b_2[V \setminus \{u, v\}] + a_2[\{u, v\}])$. As $N(a_1, a_2[V \setminus \{u, v\}] + b_2[\{u, v\}]) = N(a_1, a_2) - 2$, it follows by induction that $G_2 \wedge uv$ can be obtained from G_1 by a sequence of pivots, hence so can G_2 . $\square$

LEMMA 10.2. *For $i \in \{1, 2\}$, let S_i be the isotropic system with a graphic presentation (G_i, a_i, b_i) such that $a_i(v), b_i(v) \in \{\alpha, \beta\}$ for all $v \in V(G_i)$. If S_1 is an $\alpha\beta$ -minor of S_2 , then G_1 is a pivot-minor of G_2 .*

Proof. The lemma is a straightforward consequence of Propositions 10.1 and 3.7. $\square$

11. Binary matroids. We would like to show that Theorem 4.1 implies the well-quasi-ordering theorem of Geelen, Gerards, and Whittle [8] for binary matroids. The proof uses the following lemma.

LEMMA 11.1 (Higman's lemma). *Let $\leq$ be a well-quasi-order on X . For finite subsets $A, B \subseteq X$, we write $A \leq B$ if there is an injective mapping $f : A \rightarrow B$ such that $a \leq f(a)$ for all $a \in A$. Then $\leq$ is a well-quasi-ordering on the set of all finite subsets of X .*

(For proof, see Diestel's book [7, Lemma 12.1.3].)

For a binary matroid M with a fixed base B , the *fundamental graph* of M with respect to B is a bipartite graph $\text{Bip}(M, B)$ such that $V(\text{Bip}(M, B)) = E(M)$ and $v \in E(M) \setminus B$ is adjacent to $w \in B$ if and only if w is in the fundamental circuit of v with respect to B . For a bipartite graph $G = (V, E)$ with a bipartition $V = A \cup B$, $\text{Bin}(G, A, B)$ is a binary matroid on V , represented by an $A \times V$ matrix $(I_A \quad A_G[A, B])$, where I_A is an $A \times A$ identity matrix.

LEMMA 11.2. *Let M_1, M_2 be binary matroids and let B_i be a fixed base of M_i . If M_1 is connected and $\text{Bip}(M_1, B_1)$ is a pivot-minor of $\text{Bip}(M_2, B_2)$, then M_1 is a minor of either M_2 or M_2^* .*

Proof. Let $H = \text{Bip}(M_1, B_1)$ and $G = \text{Bip}(M_2, B_2)$. If H is a pivot-minor of a bipartite graph G , then there is a bipartition (A', B') of H such that a binary matroid $M_3 = \text{Bin}(H, A', B')$ is a minor of $M_2 = \text{Bin}(G, B_2, V(G) \setminus B_2)$. Since M_1 is connected, H is connected, and therefore H has a unique bipartition. So, $M_1 = M_3$ if $A = A'$ or $M_1 = M_3^*$ if $A = B'$. $\square$

COROLLARY 11.3. *Let k be a constant. If $\{M_1, M_2, M_3, \dots\}$ is an infinite sequence of binary matroids of branch-width at most k , then there exist $i < j$ such that M_i is isomorphic to a minor of M_j .*

Proof. First, we claim that if M_i is connected for all i , then the statement is true. Let B_i be a fixed base of M_i and $G_i = \text{Bip}(M_i, B_i)$ for all i . The rank-width of G_i is at most $k - 1$, since rank-width of G_i is equal to (branch-width of M_i) - 1, shown in [11]. By Theorem 4.1, there is an infinite subsequence $G_{a_1}, G_{a_2}, G_{a_3}, \dots$ such that G_{a_i} is isomorphic to a pivot-minor of $G_{a_{i+1}}$ for all i . By Lemma 11.2, M_{a_1} is isomorphic to a minor of either M_{a_2} , or $M_{a_2}^*$, and M_{a_2} is isomorphic to a minor of either M_{a_3} or $M_{a_3}^*$. It follows that M_{a_1} is isomorphic to a minor of M_{a_2} , M_{a_2} is

isomorphic to a minor of M_{a_3} , or M_{a_1} is isomorphic to a minor of M_{a_3} . This proves the above claim.

Now, we prove the main statement. We may consider each M_i as a set of disjoint connected matroids and then M_i is isomorphic to a minor of M_j if and only if there is an injective function f from components of M_i to components of M_j such that a is isomorphic to a minor of $f(a)$ for every component a of M_i . By Higman's lemma, there exist $i < j$ such that M_i is isomorphic to a minor of M_j . $\square$

Acknowledgments. This work was performed while the author was at the Program in Applied and Computational Mathematics, Princeton University, Princeton, New Jersey and partially supported by the SRC program of Korea Science and Engineering Foundation grant funded by Korea government (MOST) (No. R11-2007-035-01002-0). I would like to thank Jim Geelen for encouraging me to obtain Corollary 11.3 from Theorem 4.1. I also thank the anonymous referee for careful reading and the detailed report.

REFERENCES

- [1] A. BOUCHET, *Isotropic systems*, European J. Combin., 8 (1987), pp. 231–244.
- [2] A. BOUCHET, *Graphic presentations of isotropic systems*, J. Combin. Theory Ser. B, 45 (1988), pp. 58–76.
- [3] A. BOUCHET, *Connectivity of isotropic systems*, in Combinatorial Mathematics: Proceedings of the Third International Conference, New York, 1985, Ann. New York Acad. Sci. 555, New York Acad. Sci., pp. 81–93.
- [4] A. BOUCHET, *Circle graph obstructions*, J. Combin. Theory Ser. B, 60 (1994), pp. 107–144.
- [5] B. COURCELLE AND S. OLARIU, *Upper bounds to the clique width of graphs*, Discrete Appl. Math., 101 (2000), pp. 77–114.
- [6] B. COURCELLE AND S.-IL OUM, *Vertex-minors, monadic second-order logic, and a conjecture by Seese*, J. Combin. Theory Ser. B, 97 (2007), pp. 91–126.
- [7] R. DIESTEL, *Graph Theory*, second ed., Grad. Texts in Math. 173, Springer-Verlag, New York, 2000.
- [8] J. F. GEELLEN, A. M. H. GERARDS, AND G. WHITTLE, *Branch-width and well-quasi-ordering in matroids and graphs*, J. Combin. Theory Ser. B, 84 (2002), pp. 270–290.
- [9] J. F. GEELLEN, A. M. H. GERARDS, AND G. WHITTLE, *A Correction to Our Paper “Branch-Width and Well-Quasi-Ordering in Matroids and Graphs,”* manuscript, 2006.
- [10] S. OUM, *Graphs of Bounded Rank-Width*, Ph.D. thesis, Princeton University, 2005.
- [11] S. OUM, *Rank-width and vertex-minors*, J. Combin. Theory Ser. B, 95 (2005), pp. 79–100.
- [12] S. OUM AND P. SEYMOUR, *Approximating clique-width and branch-width*, J. Combin. Theory Ser. B, 96 (2006), pp. 514–528.
- [13] N. ROBERTSON AND P. SEYMOUR, *Graph minors. IV. Tree-width and well-quasi-ordering*, J. Combin. Theory Ser. B, 48 (1990), pp. 227–254.
- [14] N. ROBERTSON AND P. SEYMOUR, *Graph minors. X. Obstructions to tree-decomposition*, J. Combin. Theory Ser. B, 52 (1991), pp. 153–190.
- [15] R. THOMAS, *A Menger-like property of tree-width: The finite case*, J. Combin. Theory Ser. B, 48 (1990), pp. 67–76.
- [16] W. T. TUTTE, *Menger's theorem for matroids*, J. Res. Nat. Bur. Standards Sect. B, 69B (1965), pp. 49–53.

COLORING AN ORTHOGONALITY GRAPH*

C. D. GODSIL[†] AND M. W. NEWMAN[†]

Abstract. We deal with a graph coloring problem that arises in quantum information theory. Alice and Bob are each given a ± 1 -vector of length 2^k and are to respond with k bits. Their responses must be equal if they are given equal inputs, and distinct if they are given orthogonal inputs; however, they are not allowed to communicate any information about their inputs. They can always succeed using quantum entanglement, but their ability to succeed using only classical physics is equivalent to a graph coloring problem. We resolve the graph coloring problem, thus determining that they can succeed without entanglement exactly when $k \leq 3$.

Key words. orthogonality graph, graph coloring, maximum independent sets, ratio bound, quantum telepathy

AMS subject classifications. 05C15, 05C69

DOI. 10.1137/050639715

1. Introduction and background. We are concerned here with a graph coloring problem that arises in quantum information theory. Following this we describe the motivating problem and how it reduces to a question about the chromatic number of a particular type of graph. We close the first section with a review of some results in the literature.

For the reader unfamiliar with graph theory, we give a brief outline of concepts and notation in an appendix.

1.1. Motivating problem. The graph Ω_n has vertex set the set of ± 1 -vectors of length n ; two vertices are adjacent if they are orthogonal. Our main result is that the chromatic number of this graph is equal to n if and only if $n = 2^k$ with $k \leq 3$.

This problem arises in the following scenario, introduced in [3, 2]. Alice and Bob are each given a 01-vector of length 2^k , and they are each to respond with a 01-vector of length k . If their input vectors are equal, then their output vectors must also be equal; if their input vectors differ in exactly 2^{k-1} positions, then their output vectors must be distinct. Also, they are not allowed to communicate any information about their input vectors to each other.

Their cause is hopeless without some shared resource. Of course this resource must not allow them to share information about their inputs. We consider two possibilities, which correspond roughly to classical physics and quantum physics.

If they are allowed to share prior information, then they could agree beforehand on a proper coloring of Ω_{2^k} . They would then each interpret their input as a vertex of Ω_{2^k} and respond with the color of that vertex. Since they are allowed to output only k bits, this works only if $\chi(\Omega_{2^k}) \leq 2^k$.

Now consider that Alice and Bob have some strategy involving some prior shared information, and that their strategy is guaranteed to succeed. Alice and Bob are then given their respective inputs. Now before they actually answer, each writes down a list of the response they would have given to all possible inputs. If they have a

*Received by the editors September 7, 2005; accepted for publication (in revised form) December 13, 2007; published electronically March 28, 2008. This research was supported by NSERC Canada.
<http://www.siam.org/journals/sidma/22-2/63971.html>

[†]Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON, N2L 3G1, Canada (cgodsil@uwaterloo.ca, mikenewman@alumni.uwaterloo.ca).

winning strategy, then they are able to do this. Let $a(x)$ be the entry in Alice's list corresponding to x , and likewise $b(x)$ for Bob. Since Alice and Bob's strategy is guaranteed to succeed for any pair of inputs, we have $a(x) = b(x)$ for all x , and $a(x) \neq b(y)$ whenever x and y correspond to adjacent vertices. Their answers are restricted to k bits, so their lists contain at most 2^k distinct entries. Thus they have a proper coloring with at most 2^k colors. (Note that we have not shown that if Alice is given the same input on different occasions, she must respond in the same way. Rather, at each round, her copy of the shared information amounts to a proper coloring.)

In other words, any strategy based on prior shared information is equivalent to coloring Ω_{2^k} with at most 2^k colors.

If, instead of information, they are allowed a shared resource of *quantum entanglement*, then they can always succeed when $n = 2^k$ for all k . This was first observed by Buhrman, Cleve, and Wigderson [3] (see also Brassard, Cleve, and Tapp [2]).

We do not assume any familiarity with quantum entanglement, qubits, and quantum information theory; however, the interested reader will find a good introduction to these areas in [13]. We summarize briefly the quantum algorithm given in [3, 2] (see [12] for a version that does not assume any background in quantum information theory). Alice and Bob share between them an entangled quantum state consisting of k EPR pairs of qubits. The ± 1 -vectors of length 2^k can be interpreted as indexing a particular family of quantum operations. So given their inputs, they each apply the corresponding operation to their qubits and then measure their qubits. They answer the result of their measurements. This turns out to be a winning strategy.

The point of this scenario is then that Alice and Bob can always succeed using quantum physics (i.e., by sharing quantum entanglement), whereas they cannot always succeed using classical physics (i.e., by sharing prior information). So our result quantifies the difference between what can be accomplished using quantum or classical physics for this particular scenario.

1.2. Prior work. In [3] it is also shown that for sufficiently large k , Alice and Bob cannot succeed by sharing only prior information. This follows directly from a deep result of Frankl and Rödl [7], who show that for large enough $n = 4m$ the size of an independent set in Ω_n is at most $(2 - \epsilon)^n$ for some $\epsilon > 0$. It follows that the chromatic number must eventually be greater than n . We note that the motivation of [7] has nothing to do with any quantum scenario; furthermore, their result is stronger than what we state here.

The reader may easily verify that $\chi(\Omega_n) = n$ for $n = 1, 2$ and with a little more effort for $n = 4$ as well (this case follows trivially from the recursive construction of section 6). Unpublished computations by Gordon Royle determined that $\chi(\Omega_8) = 8$ and characterized all of the proper 8-colorings. Galliard, Tapp, and Wolf [9] show that the size of an independent set in Ω_{16} is at most 3912, which implies that its chromatic number is at least 17.

2. Some simple cases. It is not hard to see that Ω_n is edgeless if n is odd.

If n is an odd multiple of two, then the vertices can be divided into the *even* vertices (those with an even number of -1 's) and the odd vertices; every edge joins an even vertex to an odd vertex, and so the graph is bipartite.

If $n = 4m$, then every edge joins an even vertex to an even vertex or an odd vertex to an odd vertex, and it follows that there are two components: one induced by the even vertices and one induced by the odd vertices. In fact, it is not hard to

see that these two components are isomorphic. Furthermore, if a vertex x is adjacent to y , then it is also adjacent to $-y$, and x is not adjacent to $-x$.

Given graphs G and H , the *lexicographic product* $G[H]$ is obtained by first replacing each vertex of G with a copy of H and then adding edges between every pair of vertices that correspond to adjacent vertices in G . (Thus choosing exactly one vertex from each new copy of H gives a set of vertices that induces a copy of G .) We see then that each component of Ω_n can be written as a lexicographic product $Y_n[\overline{K_2}]$ for some graph Y_n .

It follows from the definition that $\chi(G) = \chi(G[\overline{K_t}])$; in particular $\chi(\Omega_n) = \chi(Y_n)$. Furthermore, maximum independent sets in Ω_n are exactly four copies of maximum independent sets in Y_n . It follows that $\alpha(\Omega_n)$ is a multiple of four. More importantly, it will simplify some of our computational work.

3. Bounding independent sets. One of the main tools we use in analyzing Ω_n is the Delsarte–Hoffman bound on independent sets (see [6, section 3.3] or [4, page 115]; alternatively [12] for more recent work).

THEOREM 3.1. *Let X be a d -regular connected graph on v vertices, and τ the least eigenvalue of its adjacency matrix A . Let S be an independent set of size s and let z be the characteristic vector of S . Then*

$$s \leq v \frac{-\tau}{d - \tau}.$$

Furthermore, equality holds if and only if

$$A \left(z - \frac{s}{v} \mathbf{1} \right) = \tau \left(z - \frac{s}{v} \mathbf{1} \right).$$

The graph Ω_n is a graph in the Hamming scheme. We will not go into details here; the interested reader is directed to [10, Chapter 12] for background material on the Hamming scheme and association schemes in general. For our present purposes, the consequence of this is that we know a complete set of eigenvectors for Ω_n . In our case, Ω_n is a Cayley graph for an Abelian group, so we can use character theory to determine a complete set of eigenvectors (see, e.g., [10, section 12.9]). This gives the following result from [11, section 4.3].

LEMMA 3.2. *Let n be a multiple of four. Then the least eigenvalue of Ω_n is*

$$\tau = -\frac{1}{n-1} \binom{n}{\frac{n}{2}}.$$

Furthermore, let W be the matrix with rows indexed by subsets of $[n]$ and columns indexed by 2-subsets and $(n-2)$ -subsets, with entry in row x and column y equal to $(-1)^{|x \cap y|}$. Then the columns of W form a basis for the τ -eigenspace.

Note that here we are thinking of the vertices of Ω_n as being subsets of $[n]$ instead of ± 1 -vectors ($[n]$ denotes $\{1, 2, \dots, n\}$); two subsets are adjacent when they are at Hamming distance $\frac{n}{2}$. In fact, given any subset $y \subseteq [n]$, the vector w_y defined by $(w_y)_x = (-1)^{|x \cap y|}$ is an eigenvector of Ω_n . As there are 2^n choices for y , we obtain 2^n eigenvectors. As these are characters, they are orthogonal, so we have a complete set of eigenvectors.

Let $\widehat{W} = (W \quad \mathbf{1})$. Putting Theorem 3.1 and Lemma 3.2 together, we obtain the following result for our graphs.

COROLLARY 3.3. *The size of an independent set in Ω_n is bounded by*

$$\alpha(\Omega_n) \leq \frac{2^n}{n}.$$

Furthermore, if equality holds, then the characteristic vector of a maximum independent set lies in the column space of $\widehat{W}$.

Note that if Ω_n is n -colorable, then this bound must hold with equality (in fact, the coloring must be a partition into maximum independent sets). In other words, we have shown that $\chi(\Omega_n) > n$ whenever $n = 4m$ is not a power of two. It is noteworthy that the question “When is $\chi(\Omega_n) \leq n$?” becomes trivial when n is not a power of two: the mathematical analysis is simplest for the cases that are physically less interesting. Of more immediate use is the fact that if the bound does not hold with equality, then Ω_n is not n -colorable. One way to show that the bound is not tight is to use the equality condition of Theorem 3.1: it suffices to show that there are no suitable vectors in the τ -eigenspace.

4. Finding maximum independent sets. Since we know all of the eigenspaces of Ω_n , it is not hard to see that Ω_n , the even component of Ω_n , and Y_n all have the same least eigenvalue τ . If we take only those columns of W that correspond to the 2-subsets, and only those rows that correspond to the even vertices, the resulting column space gives the τ -eigenspace of the even component (this amounts to taking only one of the two eigenspaces of the Hamming scheme that give the τ -eigenspace on Ω_n). If we further reduce by taking only one vertex (i.e., row) from each pair $\{x, -x\}$, then we obtain a matrix whose columns form a basis for the τ -eigenspace for Y_n . We will denote this matrix by H . Furthermore, let $\widehat{H} = \begin{pmatrix} H & \mathbf{1} \end{pmatrix}$.

COROLLARY 4.1. *The size of an independent set in Y_n is bounded by*

$$\alpha(Y_n) \leq \frac{2^n}{4n}.$$

Furthermore, if equality holds then the characteristic vector of a maximum independent set lies in the column space of $\widehat{H}$.

Let z be the characteristic vector of an independent set S that meets the bound of Corollary 4.1. Then $z = \widehat{H}y$ for some vector y . Since Y_n is vertex-transitive, we are free to assume that S contains any particular vertex. As S is maximum, this is equivalent to assuming that S is disjoint from the neighborhood of a vertex. Let $\widehat{N}$ be the submatrix of $\widehat{H}$ with rows corresponding to a neighborhood; then we may assume that z takes the value 0 at the corresponding positions, meaning that y is in the kernel of $\widehat{N}$. Thus we are led to the following result.

LEMMA 4.2. *The kernel of $\widehat{N}$ is given by the row space of $\widehat{B} = \begin{pmatrix} B & \mathbf{1} \end{pmatrix}$, where B is the incidence matrix of K_n .*

Proof. A direct computation shows that

$$NB^T = -\mathbf{1}.$$

This means that

$$\widehat{N}\widehat{B}^T = 0.$$

We will show that this is the whole kernel by a rank argument.

We note that

$$BB^T = (n-1)I + J.$$

Thus the eigenvalues of BB^T are $2n - 1$ and $n - 1$, so B has full rank, and $\text{rk}(B) = \text{rk}(BB^T) = n$. As $B\mathbf{1} = (n - 1)\mathbf{1}$, we see that $\text{rk}(\widehat{B}) = n$.

Essentially the same argument determines the rank of $\widehat{N}$ as well. The rows and columns of $N^T N$ are indexed by 2-subsets, and the (a, b) -entry depends only on $|a \cap b|$. Let L and $\bar{L}$ be the incidence matrices of the line graph of K_n and its complement, respectively. It follows that

$$N^T N = c_0 I + c_1 L + c_2 \bar{L},$$

where

$$\begin{aligned} c_0 &= \binom{n}{\frac{n}{2}}, \\ c_1 &= \binom{n}{\frac{n}{2}} - 8 \binom{n-3}{\frac{n}{2}-1}, \\ c_2 &= \binom{n}{\frac{n}{2}} - 16 \binom{n-4}{\frac{n}{2}-1}. \end{aligned}$$

The matrices $I, L, \bar{L}$ are simultaneously diagonalizable with known eigenvectors (more precisely: the line graph of K_n is strongly regular). It follows that the eigenvalues of $N^T N$ are

$$\frac{n}{2(n-1)} \binom{n}{\frac{n}{2}}, \quad \frac{n(n-2)}{(n-1)(n-3)} \binom{n}{\frac{n}{2}}, \quad 0,$$

with respective multiplicities 1, $\binom{n}{2} - n$, and $n - 1$. As $N\mathbf{1} = \frac{n}{2}\mathbf{1}$, we see that $\text{rk}(\widehat{N}) = \text{rk}(N)$, and so $\dim(\ker(\widehat{N})) = \text{rk}(\widehat{B})$. The result follows. $\square$

We note that the rank arguments in the above proof amount to observing that BB^T and $N^T N$ both lie in the Bose–Mesner algebras of known association schemes.

Let C be the reduced column echelon form of the matrix $\widehat{H}\widehat{N}$ (C is in reduced column echelon form means that C^T is in reduced row echelon form). Then it follows that there are vectors x, x' such that

$$z = \widehat{H}y = \widehat{H}\widehat{B}^T x = Cx'.$$

Furthermore, since z is a 01-vector so is x' . Since H and B^T have full column rank, it follows that the rank of C is n . Thus it suffices to check all 2^n possibilities for x' in order to determine if there exist any independent sets that meet the bound. We have carried out this computation for $n = 8, 16$: for $n = 8$, we find that there are eight independent sets of the required size containing a given vertex; for $n = 16$, there are none.

On its own this computation is not particularly satisfying: we have not contradicted Royle's result mentioned above, and we have established a weaker bound on $\alpha(\Omega_{16})$ than the one given in [9]. However, it will turn out that our computations for $n = 8, 16$ suffice to determine all values of n for which Ω_n is n -colorable. For this purpose, we will rederive the bound of Theorem 3.1 for Ω_n twice more.

5. Coloring Ω_n . It is well-known that for any vertex-transitive graph X on v vertices, $\alpha(X)\omega(X) \leq v$. This is the *clique-coclique bound*. It also holds for any graph that is a union of classes in an association scheme. Ω_n falls into both of these categories, but it is also a normal Cayley graph, for which we can extend this result.

In particular, we will show that for a normal Cayley graph X , if $\alpha(X)\omega(X) = |V(X)|$, then $\chi(X) = \omega(X)$. We will need some preliminary results first.

Recall that the *connection set* of a Cayley graph for a group G is the subset D of G such that $a \sim b$ whenever $ba^{-1} \in D$. If S is a subset of G , then we write

$$S^{-1} = \{g^{-1} : g \in S\}, \\ Sa = \{ga : g \in S\}.$$

LEMMA 5.1. *Let X be a Cayley graph for a group G with connection set $D \subseteq G$. Let S be an independent set of X . If a and b are adjacent, then $S^{-1}a \cap S^{-1}b = \emptyset$.*

Proof. Assume $g^{-1}a = h^{-1}b$ for some $g, h \in S$. Then $hg^{-1} = ba^{-1}$. But $h \sim g$ so $hg^{-1} \notin D$, while $b \sim a$ so $ba^{-1} \in D$. $\square$

COROLLARY 5.2. *If X is a Cayley graph, then $\alpha(X)\omega(X) \leq |V(X)|$.*

Proof. Let S be an independent set and C a clique. Then by the previous result the sets

$$S^{-1}c, c \in C$$

are all disjoint. $\square$

We note parenthetically that this can be extended to a proof for all vertex-transitive graphs. It is an old result of Sabidussi [14] that if X is a vertex-transitive graph, there is an integer m such that the lexicographic product $X[\overline{K_m}]$ is a Cayley graph. Since

$$\alpha(X[\overline{K_m}]) = m\alpha(X), \quad \omega(X[\overline{K_m}]) = \omega(X), \quad |V(X[\overline{K_m}])| = m|V(X)|,$$

the result follows for all vertex transitive graphs.

Recall that a Cayley graph is *normal* if its connection set is closed under conjugation. Our purpose in approaching the clique-coclique bound through Cayley graphs is the following extension from [11].

COROLLARY 5.3. *If X is a normal Cayley graph and $\alpha(X)\omega(X) = v$, then $\chi(X) = \omega(X)$.*

Proof. Let S be a maximum independent set and C be a maximum clique. Again, the sets

$$S^{-1}c, c \in C$$

are disjoint, and so they partition the vertex set. Since X is normal, they are also independent sets and hence form a coloring. $\square$

For our purposes, Corollary 5.3 rederives the bound of Theorem 3.1 for Ω_n , but with a different equality condition. Notice that any clique in Ω_n is a set of pairwise orthogonal vectors; hence it is linearly independent and has size at most n . Furthermore, an n -clique would correspond to a Hadamard matrix, which certainly exists if $n = 2^k$. It follows that we can use Corollary 5.3 to rederive the bound of Theorem 3.1, but with a different equality condition.

COROLLARY 5.4. *Let n be a power of two.*

The size of an independent set in Ω_n is bounded as

$$\alpha(\Omega_n) \leq \frac{2^n}{n}.$$

Furthermore, if equality holds, then $\chi(\Omega_n) = n$.

So $\chi(\Omega_n) = n$ if and only if $\alpha(\Omega_n) = \frac{2^n}{n}$. Our determination of $\alpha(\Omega_8)$ above is now upgraded to a proof that $\chi(\Omega_8) = 8$.

We see that $\chi(\Omega_n) = n$ only if n is a power of two; this is because $\alpha(\Omega_n)$ must be an integer. Furthermore, if $\chi(\Omega_n) = n$, then each color class is an independent set that meets the bound of Theorem 3.1.

6. A recursive construction. The graph Ω_n is an induced subgraph of Ω_{2n} : take exactly those vertices of Ω_{2n} whose last n entries are the same as the first n entries. In fact, we can say much more than this.

For vertices x, r of Ω_n , let $x^{(r)}$ be the vertex of Ω_{2n} obtained by concatenating x with the entrywise product of x and r (recall that vertices are ± 1 -vectors). Let $\Omega_n^{(r)}$ be the subgraph induced by the vertices

$$\{x^{(r)} : x \in V(\Omega_n)\}.$$

Then we see that $\Omega_n^{(r)}$ is isomorphic to Ω_n , for any r . The vertex set of Ω_{2n} can be partitioned as

$$\{V(\Omega_n^r) : r \in V(\Omega_n)\}.$$

Furthermore, every vertex of $\Omega_n^{(r)}$ is adjacent to every vertex of $\Omega_n^{(-r)}$.

The *join* of two graphs X_1 and X_2 is $X_1 + X_2 = \overline{X_1} \cup \overline{X_2}$. (Informally, we first take the disjoint union of the two graphs, and then add edges between each vertex in X_1 and each vertex in X_2 .) We have established the following result.

LEMMA 6.1. *The vertex set of Ω_{2n} can be partitioned into 2^{n-1} copies of $\Omega_n + \Omega_n$.*

Note that any independent set in the join of two graphs must lie entirely within one or the other. This gives us a bound on the size of an independent set in Ω_{2n} : it is at most 2^{n-1} times the size of a maximum independent set in Ω_n . We can use this to again rederive the bound of Theorem 3.1, with yet another equality condition.

COROLLARY 6.2. *Let $n = 2^k$ where $k > 1$.*

The size of an independent set in Ω_n is bounded as

$$\alpha(\Omega_n) \leq \frac{2^n}{n}.$$

Furthermore, if equality holds then it also holds for $n = 2^{k-1}$.

We can in fact apply the recursive construction when n is not a power of two. If $n = m2^k$, where m is odd, then we find that

$$(6.1) \quad \alpha(\Omega_n) \leq \frac{2^n}{2^k}.$$

When $k > 1$ the bound of Corollary 3.3 is better than (6.1) by a factor of m . When $k = 1$, the bound of (6.1) is half the number of vertices, and is tight since Ω_n is then bipartite. But for $k = 1$ the least eigenvalue of Ω_n is no longer given by Lemma 3.2, and applying Theorem 3.1 in this case again gives half the number of vertices.

In other words, this recursion does not ever give a tighter bound than Theorem 3.1; rather, it is useful because of the equality condition of Corollary 6.2.

We mention another point of view on this recursion.

For $n = 2^k$, let the graphs Ψ_n be defined by setting $\Psi_1 := \Omega_1 = \overline{K_2}$, and recursively defining Ψ_{2n} to be the disjoint union of 2^{n-1} copies of $\Psi_n + \Psi_n$. Then Ψ_n

is a spanning subgraph of Ω_n , and so the method of Corollary 6.2 gives a bound on independent sets in Ψ_n as well. It follows that

$$\alpha(\Psi_n) = \frac{2^n}{n},$$

$$\chi(\Psi_n) = n.$$

Furthermore we see that $\Psi_n = \Omega_n$ for $n = 1, 2, 4$, by simply observing that their degrees are the same. This is an easy way to see that $\chi(\Omega_4) = 4$.

There are of course other edges in Ω_n for general n . In fact, Ψ_n is not only spanning, it is asymptotically sparse, in the following sense.

LEMMA 6.3.

$$\lim_{k \rightarrow \infty} \frac{|E(\Psi_{2^k})|}{|E(\Omega_{2^k})|} = 0.$$

In other words, for $n = 2^k$, Theorem 3.1 effectively only “sees” the edges of Ψ_n . This is expected, since one consequence of the Frankl–Rödl result is that for large enough n , the bound of Corollary 3.3 is exponentially too big.

7. Main result. Recall that although we can easily define Ω_n for any positive integer n , it is the cases where $n = 2^k$ that are most of interest. In exactly these cases, we have three different ways of proving the same bound on independent sets: using the Delsarte–Hoffman bound, using the maximum cliques, and a recursive construction. Furthermore, each approach gives different information in the case where the bound is tight.

We now find that our main result follows directly.

THEOREM 7.1. $\chi(\Omega_n) = n$ if and only if $n = 2^k$ with $k \leq 3$.

Proof. It follows from Corollary 3.3 and the comments after it that $\chi(\Omega_n) \neq n$ if n is not a power of two. Furthermore, if $\chi(\Omega_n) = n$, then the bound of Corollary 3.3 holds with equality, and Corollary 5.4 tells us that it is sufficient that this bound holds with equality. Our computations of section 4 deal with the cases $n = 2^k$ for $k = 3, 4$. We then invoke Corollary 6.2 to conclude that $\chi(\Omega_{2^k}) > 2^k$ for all $k > 4$. $\square$

8. Further bounds. Galliard [8] has a construction of an independent set inspired by the methods of Ahlswede and Khachatrian [1]. We state it in terms of the graph Y_n . It is convenient to regard the vertices as being subsets of $[n]$.

Let $n = 2^k$ and $c = \frac{n}{4} - 1$. Then the following collection is an independent set in Y_n :

$$\mathcal{F}_n = \{F \subseteq [n] : |F| = 2i, i \leq 2c; |F \cap [c]| \geq |F \setminus [c]|\}.$$

It is not hard to see that $\mathcal{F}$ is not properly contained in any larger independent set.

It turns out that for $k \leq 3$ this set meets the bound of Corollary 3.3, and hence this construction is maximum. Up to automorphisms of Ω_n , this is unique (this follows both from Gordon Royle’s computations and from our work in section 4). Galliard conjectured that $\mathcal{F}_n$ is maximum for all $n = 2^k$.

A recent computation of de Klerk and Pasechnik [5], using a technique of Schrijver [15] gives that $\alpha(Y_{16}) \leq 576$, which is exactly the size of $\mathcal{F}_{16}$. The reader is referred to [5] for details and further results. We do not know the value of $\alpha(\Omega_{4m})$ for $m > 4$.

There is another way to look at the collection $\mathcal{F}$. If we replace each element F of $\mathcal{F}_n$ with its symmetric difference with $[c]$ we obtain the collection of all odd subsets of $[n]$ of size at most c . More generally, if we assume only that $n = 4m$, we have the following construction:

$$\mathcal{S}_n = \{F \subseteq [n] : |F| \not\equiv m \pmod{2}; |F| < m\}.$$

We conjecture that these are in fact maximum in general. If this is true, it would imply our statement of the Frankl–Rödl result, namely that

$$\alpha(\Omega_{4m}) \leq (2 - c)^{4m}$$

for some $c > 0$, for large enough m .

Appendix. Graph theory notation.

A *graph* is a pair (V, E) where the elements of V are called *vertices* and E is a set of pairs of distinct vertices, called *edges*. If $\{x, y\} \in E$, we say that x and y are adjacent, and we write $x \sim y$. We often write the edge $\{x, y\}$ as xy . A *path* is a sequence of vertices with consecutive vertices adjacent; a graph is *connected* if there is a path between any two distinct vertices.

A *subgraph* of (V, E) is a pair (V', E') with $V' \subseteq V$, $E' \subseteq E$, where $xy \in E'$ implies $x, y \in V'$. It is *induced* if whenever $x, y \in V'$ and $xy \in E$, then $xy \in E'$ as well. The complement of a graph $X = (V, E)$ is the graph $\bar{X}$ which has the same vertex set as X , and where vertices are adjacent in $\bar{X}$ if and only if they are not adjacent in X .

An *independent set* is a set of vertices no two of which are adjacent. By a *maximum independent set*, we will always mean one of maximum size in the graph. A *clique* is a set of vertices where every two (distinct) vertices are adjacent; a clique with m vertices is denoted by K_m . (An independent set is sometimes called a *coclique*.) The size of a largest independent set in a graph X is denoted by $\alpha(X)$, and the size of the largest clique is denoted by $\omega(X)$.

A proper (vertex) coloring of a graph is a function c from the vertex set to a finite set of “colors” (the positive integers will do) such that whenever $x \sim y$, $c(x) \neq c(y)$. The preimage of a color (sometimes called a *color class*) is an independent set; thus a coloring is a partition of the vertex set into independent sets. A *t-coloring* is a coloring using t colors. The *chromatic number* of a graph X is the least positive integer t such that the graph has a t -coloring; it is denoted by $\chi(X)$.

The adjacency matrix of a graph is the (square) matrix with rows and columns indexed by the vertices, whose (i, j) -entry is equal to one or zero according to whether i and j are adjacent or not. The *eigenvalues of a graph* are the eigenvalues of its adjacency matrix.

REFERENCES

- [1] R. AHLWEDE AND L. H. KHACHATRIAN, *The complete intersection theorem for systems of finite sets*, European J. Combin., 18 (1997), pp. 125–136.
- [2] G. BRASSARD, R. CLEVE, AND A. TAPP, *Cost of exactly simulating quantum entanglement with classical communication*, Phys. Rev. Lett., 83 (1999), pp. 1874–1877.
- [3] H. BUHRMAN, R. CLEVE, AND A. WIDGERSON, *Quantum vs. classical communication and computation*, in Proceedings of the 30th Annual ACM Symposium on the Theory of Computing, 1998, pp. 63–68.
- [4] D. M. CVETKOVIĆ, M. DOOB, AND H. SACHS, *Spectra of Graphs*, Academic Press, New York, 1980.

- [5] E. DE KLERK AND D. V. PASECHNIK, *A note on the stability number of an orthogonality graph*, (2004), available online as arXiv:math.CO/0505038.
- [6] P. DELSARTE, *An algebraic approach to the association schemes of coding theory*, Philips Res. Rep. Suppl., (1973), pp. vi+97.
- [7] P. FRANKL AND V. RÖDL, *Forbidden intersections*, Trans. Amer. Math. Soc., 300 (1987), pp. 259–286.
- [8] V. GALLIARD, *Classical Pseudo-Telepathy and Colouring Graphs*, Master's thesis, ETH Zurich, 2001.
- [9] V. GALLIARD, A. TAPP, AND S. WOLF, *The impossibility of pseudo-telepathy without quantum entanglement*, (2002), available online as arXiv:quant-ph/0211011.
- [10] C. D. GODSIL, *Algebraic Combinatorics*, Chapman & Hall, New York, 1993.
- [11] C. D. GODSIL, *Interesting Graphs and Their Colourings*, 2003, unpublished notes, <http://quoll.uwaterloo.ca/pstuff/colours.pdf>.
- [12] M. W. NEWMAN, *Independent Sets and Eigenspaces*, Ph.D. thesis, University of Waterloo, Waterloo, ON, Canada, 2004.
- [13] M. A. NIELSEN AND I. L. CHUANG, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, UK, 2000.
- [14] G. SABIDUSSI, *Vertex-transitive graphs*, Monatsh. Math., 68 (1964), pp. 426–438.
- [15] A. SCHRIJVER, *New code upper bounds from the Terwilliger algebra and semidefinite programming*, IEEE Trans. Inform. Theory, 51 (2005), pp. 2859–2866.

FINDING A MAXIMUM INDEPENDENT SET IN A SPARSE RANDOM GRAPH*

URIEL FEIGE[†] AND ERAN OFEK[‡]

Abstract. We consider the problem of finding a maximum independent set in a random graph. The random graph G , which contains n vertices, is modeled as follows. Every edge is included independently with probability $\frac{d}{n}$, where d is some sufficiently large constant. Thereafter, for some constant α , a subset I of αn vertices is chosen at random, and all edges within this subset are removed. In this model, the planted independent set I is a good approximation for the maximum independent set I_{max} , but both $I \setminus I_{max}$ and $I_{max} \setminus I$ are likely to be nonempty. We present a polynomial time algorithm that with high probability (over the random choice of random graph G and without being given the planted independent set I) finds the maximum independent set in G when $\alpha \geq \sqrt{c_0/d}$, where c_0 is some sufficiently large constant independent of d .

Key words. average case, maximum independent set

AMS subject classifications. 05C80, 05C85, 68R10

DOI. 10.1137/060661090

1. Introduction. Let $G = (V, E)$ be a graph. An independent set I is a subset of vertices which contains no edges. The problem of finding a maximum size independent set in a graph is a fundamental problem in computer science, and it was among the first problems shown to be NP-hard [17]. Moreover, Håstad shows [14] that for any $\epsilon > 0$ there is no $n^{1-\epsilon}$ approximation algorithm for the maximum independent set problem unless $\text{NP} = \text{ZPP}$. The best approximation ratio currently known for a maximum independent set [6] is $O(n(\log \log n)^2/(\log n)^3)$.

In light of the above-mentioned negative results, we may try to design a heuristic that performs well on typical instances. Karp [16] proposed trying to find a maximum independent set in a random graph. However, even this problem appears to be beyond the capabilities of current algorithms. For example, let $G_{n,1/2}$ denote the random graph on n vertices obtained by choosing randomly and independently each possible edge with probability $1/2$. The size of the maximum independent set in a random $G_{n,1/2}$ graph is almost surely $2(1 + o(1)) \log_2 n$. A simple greedy algorithm almost surely finds an independent set of size $\log_2 n$ [13]. However, there is no known polynomial time algorithm that almost surely finds an independent set of size $(1 + \epsilon) \log_2 n$ (for any $\epsilon > 0$).

To further simplify the problem, Jerrum [15] and Kučera [18] proposed a *planted model* $G_{n,1/2,k}$ in which a random graph $G_{n,1/2}$ is chosen and then a clique of size k is randomly placed in the graph. (A clique in a graph G is an independent set in the edge complement of G , and hence all algorithmic results that apply to one of the problems apply to the other.) Alon, Krivelevich, and Sudakov [2] gave an algorithm based on spectral techniques that almost surely finds the planted clique for $k = \Omega(\sqrt{n})$. More

*Received by the editors May 26, 2006; accepted for publication (in revised form) December 17, 2007; published electronically March 28, 2008. This work was supported in part by a grant from the G.I.F., the German–Israeli Foundation for Scientific Research and Development. Part of this work was done while the authors were visiting Microsoft Research in Redmond, WA.

<http://www.siam.org/journals/sidma/22-2/66109.html>

[†]Department of Computer Science and Applied Mathematics, The Weizmann Institute, Rehovot 76100, Israel (uriel.feige@weizmann.ac.il).

[‡]Schema, Herzliya 46905, Israel (eran.ofek@gmail.com).

generally, one may extend the range of parameters of the above model by planting an independent set in $G_{n,p}$, where p need not be equal to $1/2$ and may also depend on n . The $G_{n,p,\alpha}$ model is as follows: n vertices are partitioned at random into two sets of vertices, I of size αn and C of size $(1 - \alpha)n$. No edges are placed within the set I , thus making it an independent set. Every other possible edge (with at least one end point not in I) is added independently at random with probability p . The goal of the algorithm given the input G (but without being given the partition into I and C) is to find a maximum independent set. Intuitively, as α becomes smaller, the size of the planted independent is closer to the probable size of the maximum independent set in $G_{n,p}$, and the problem becomes harder.

We consider values of p as small as d/n , where d is a large enough constant. A difficulty that arises in this sparse regime (i.e., when d is constant) is that the planted independent set I is not likely to be a maximum independent set. Moreover, with high probability I is not contained in a maximum independent set of G . For example, there are expected to be $e^{-d}n$ vertices in C of degree one. It is very likely that two (or more) such vertices $v, w \in C$ will have the same neighbor and that it will be some vertex $u \in I$. This implies that every maximum independent set will contain v and w and not u , and thus I contains vertices that are not contained in the maximum independent set. See Figure 1.

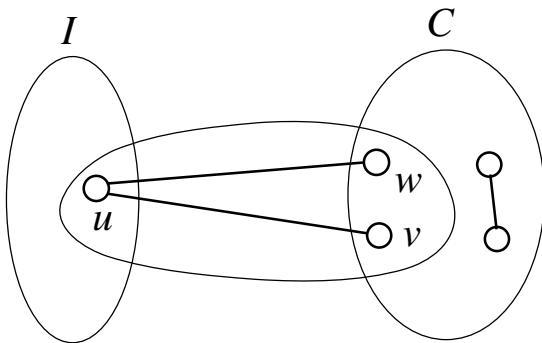


FIG. 1. The vertex $u \in I$ is not contained in any maximum independent set because no other edges touch v, w .

A similar argument shows that there are expected to be $e^{-\Omega(d)}n$ isolated edges. This implies that there will be an exponential number of maximum independent sets.

1.1. Our result. Our main result is summarized in the following theorem.

THEOREM 1.1. *There is a polynomial time algorithm that almost surely finds the maximum independent set of a graph G selected at random from the distribution $G_{n,\frac{d}{n},\alpha}$, when $d > d_0$ and $\alpha \geq \sqrt{c_0/d}$ (d_0, c_0 are some universal constants). The parameter d can also be an increasing function of n .*

The bulk of the paper is devoted to proving the above theorem. To simplify the presentation of the proof, it will be convenient to assume upper bounds on d and on α , namely, $d < n^{1/40}$ and $\alpha < 1/3$. These upper bounds do not limit the generality of our results, because when d or α exceeds these upper bounds the proof becomes simpler (see later discussions in the paper).

1.2. Related work. For $p = 1/2$, $\alpha = \Omega(1/\sqrt{n})$, Alon, Krivelevich, and Sudakov [2] gave an efficient spectral algorithm which almost surely finds the planted

independent set. For the above-mentioned parameters the planted independent set is likely to be the unique maximum independent set.

A few papers deal with *semirandom* models which extend the planted model by enabling a mixture of random and adversarial decisions. Feige and Kilian [7] considered the following model: a random $G_{n,p,\alpha}$ graph is chosen, and then an adversary may add arbitrarily many edges between I and C and make arbitrary changes (adding or removing edges) inside C . For any constant $\alpha > 0$ they give a heuristic that almost surely outputs a list of independent sets containing the planted independent set, whenever $p > (1 + \epsilon) \ln n / \alpha n$ (for any $\epsilon > 0$). The planted independent set may not be the only independent set of size αn since the adversary has full control over the edges inside C . Possibly, this makes the task of finding the planted independent set harder. In [8] Feige and Krauthgamer considered a less adversarial semirandom model in which an adversary is allowed to add edges to a random $G_{n, \frac{1}{2}, \frac{1}{\sqrt{n}}}$ graph. Their algorithm almost surely extracts the planted independent set and certifies its optimality. Coja-Oghlan [4] considered similar semirandom variants of random $G_{n,p,\sqrt{c/pn}}$ graphs, with $\ln(n)^2/n < p < 0.99$ (and c a sufficiently large constant). He shows among other things that in most such graphs a maximum independent set can be found in polynomial time.

By extracting from the above discussion the known previous results on planted independent set models (and ignoring issues such as handling semirandom instances), we see that the same trade-off between d and α as claimed in the current paper (namely, $\alpha \geq \sqrt{c_0/d}$) was previously established for the case $d = n/2$ (first in [2] and then again in [8]) and then extended all the way down to $d = \ln^2 n$ in [4]. The condition $\alpha \geq \sqrt{c_0/d}$ is needed mainly to ensure that the effect of planting the independent set I shows up in the spectrum of the adjacency matrix of the graph. For the previous range of parameters, the planted independent set is likely to be the maximum independent set in the graph. In this paper we address the case that d is a (sufficiently large) constant, which introduces difficulties not present in earlier work, mainly due to the fact that the planted independent set is not likely to be the maximum independent set.

Heuristics for optimization problems different from the max independent set will be discussed in the following section.

1.2.1. Technique and outline of the algorithm. Our algorithm builds on ideas from the algorithm of Alon and Kahale [1], which was used for recovering a planted 3-coloring in a random graph. The algorithm that we propose has four phases and is sketched below. This overall structure is similar to the structure of the algorithm of [1], but phases 3 and 4 include ingredients that are different from those of [1]. These differences are discussed towards the end of this section.

1. Get a coarse approximation I', C' of I, C with $|C \Delta C'| + |I \Delta I'| < \frac{1}{60}|I|$. This phase is based on spectral techniques.
2. Reduce the error in the approximation by using an iterative procedure based on the number of neighbors that each vertex has in I' . The error term $|C \Delta C'| + |I \Delta I'|$ is reduced to at most n/d^{18} .
3. Commit to placing certain vertices of I' in the final independent set. This is done by moving vertices from I' and C' to *OUT* (a new set) and committing to placing the remaining part of I' in the final independent set. The choice of vertices to move to *OUT* depends on their degree into I' . When this process ends, I' is an independent set, every vertex of C' has at least 4 edges

to I' , and no vertex of I' has edges to OUT . By using the fact that sparse random graphs (almost surely) have no small dense sets, it will be shown that $I' \subseteq I_{max}$ and also that OUT is rather small.

4. Extend the independent set I' optimally by using the vertices of OUT . This is done by finding a maximum independent set among the vertices of OUT and adding it to I' . The structure of OUT will be simple enough so that a maximum independent set can be efficiently found. The proof is based to a large extent on the fact that OUT is small. If vertices of OUT were chosen independently at random from the input graph G , then it would have been easy to show that all connected components in OUT are very small (say, of size $O(\log n)$), and then a maximum independent set in OUT can be found in time polynomial in n (even by an exhaustive search). However, OUT is a result of a deterministic process applied to G , which makes the analysis of its structure considerably more difficult.

The technique of [1] was implemented successfully on various problems in the planted model: planted hypergraph coloring, planted 3-SAT, planted 4-NAE, and min-bisection (by Chen and Frieze [3], Flaxman [11], Goerdts and Lanka [12], and Coja-Oghlan [5], respectively).

Perhaps the work closest in nature to the work in the current paper is that of Coja-Oghlan [5] on finding a bisection in a sparse random graph. Both in our work and in that of [5], one is dealing with an optimization problem, and the density of the input graph is such that the planted solution is not an optimal solution. The algorithm for bisection in [5] is based on spectral techniques and has the advantage that it provides a certificate showing that the solution that it finds is indeed optimal. We do not address the issue of certification in this paper. In [5] the random instance is generated as follows. The vertices of the graph are partitioned into two classes of equal size randomly. Then the edges are inserted: edges inside the two classes with probability p' and edges crossing the partition with probability p independently. Intuitively, as $p' - p$ becomes smaller, the problem becomes harder.

Denote by $d_1 = np'/2$ and $d_2 = np/2$ the expected degree of a vertex into its own class and into the other class, respectively. The algorithm in [5] is proven to succeed (almost surely) whenever $d_1 - d_2 \geq \sqrt{c_0 d_1 \log d_1}$. In our independent set model the problem becomes harder as αd becomes smaller. If we denote by $\tilde{d}_1 = d$ and $\tilde{d}_2 = (1 - \alpha)d$ the expected degrees of a vertex in C and I , respectively, then our algorithm (almost surely) succeeds whenever $\tilde{d}_1 - \tilde{d}_2 = \alpha \tilde{d}_1 \geq \sqrt{c_0 \tilde{d}_1}$. We remark that a significant source for complications in our algorithm and its analysis comes from a tightening of the parameters, saving a $\sqrt{\log \tilde{d}}$ factor in the size of α . A preliminary version of this work (see [10]) presents a simpler algorithm but requires $\alpha \tilde{d}_1 \geq \sqrt{c_0 \tilde{d}_1 \log \tilde{d}_1}$.

An important difference between planted models for an independent set and those for other problems such as 3-coloring and min-bisection is that in our case the planted classes I, C are not symmetric. The lack of symmetry between I and C makes some of the ideas used for the more symmetric problems insufficient. In the approach of [1], a vertex is removed from its current color class and placed in OUT if its degree into some other current color class is very different from what one would typically expect to see between the two color classes. This procedure is shown to “clean” every color class C from all vertices that should have been from a different color class but were wrongly assigned to class C in phase 2 of the algorithm. (The argument proving

this goes as follows. Every vertex remaining in the wrong color class by the end of phase 3 must have many neighbors that are wrongly assigned themselves. Thus the set of wrongly assigned vertices induces a small subgraph with a large edge density. But G itself does not have any such subgraphs, and hence by the end of phase 3 it must be the case that all wrongly assigned vertices were moved into OUT .) It turns out that this approach works well when classes are of similar nature (such as color classes or two sides of a bisection) but does not seem to suffice in our case, where I' is supposed to be an independent set whereas C' is not. Specifically, the set I' might still contain wrongly assigned vertices and might not be a subset of a maximum independent set in the graph. Under these circumstances, phase 4 will not result in a maximum independent set. Our solution to this problem involves the following aspects, not present in previous work. In phase 3 we remove from I' every vertex that has even one edge connecting it to OUT . This adds more vertices to OUT and makes the connected components in OUT larger. We do not attempt to analyze the maximum size of connected components in OUT , which is a key ingredient in previous approaches. Instead, we analyze the 2-core of OUT and show that the 2-core has no large components. Then, in phase 4, we use dynamic programming to find a maximum independent set in OUT and use the special structure of OUT to show that the algorithm runs in polynomial time.

1.3. Notation. Let $G = (V, E)$, and let $U \subset V$. The subgraph of G induced by the vertices of U is denoted by $G[U]$. When the set of edges used is clear from the context, we will use $\deg(v)_U$ to denote the degree of a vertex v into a set $U \subset V$. To specify exactly the set of edges used, we use $\deg^E(v)_U$, which is the degree of a vertex v into a set U induced by the set of edges E . We use $\Gamma(U)$ to denote the vertex neighborhood of $U \subset V$ (excluding U). The parameter d (specifying the expected degree in the random graph G) is assumed to be sufficiently large, and some of the inequalities that we shall derive implicitly use this assumption, without stating it explicitly. The term *with high probability* (w.h.p.) is used to denote a sequence of probabilities that tends to 1 as n tends to infinity.

2. The algorithm. The first step of the algorithm is tailored to deal with the case that α is small. When α is large (in fact, $\alpha > \sqrt{100 \log d/d}$ suffices), the combination of steps 1 and 2 can be replaced by a simpler step based on partitioning the vertices according to their degrees (see section 3.2 for details). Hence we shall assume here that $\alpha \leq 1/3$, and this will somewhat simplify the presentation of the algorithm and its analysis. Furthermore, we assume for simplicity that the values of the parameters α, d are given as inputs to the algorithm. This assumption can be avoided by enumerating all possible values of α, d (actually we need only a good approximation of α, d).

Figure 2 depicts the situation after step 3 of the algorithm is done. At that point, I_3 is an independent set, there are no edges between I_3 and OUT_3 , and every vertex $v \in C_3$ has at least four neighbors in I_3 .

3. Correctness. Let I_{max} be a maximum independent set of G . We establish two theorems. Theorem 3.1 guarantees the correctness of the algorithm, and Theorem 3.2 guarantees its efficient running time. Here we present these two theorems, and their proofs are deferred to later sections.

THEOREM 3.1. *W.h.p. there exists some maximum independent set I_{max} such that $I_3 \subseteq I_{max}$, $C_3 \cap I_{max} = \emptyset$.*

ALGORITHM *FindIS*(V, E).

1. Let A' be the adjacency matrix of the graph induced by removing from G all vertices of degree $> 5d$. Compute the eigenvector of the most negative eigenvalue of A' , denoted by $v_{n'}$. Sort the vertices of V' in order of the value of the corresponding entry in $v_{n'}$, breaking ties arbitrarily. Let I_1 be either the first αn vertices in this sorted order or the last αn vertices. (To choose between the above two options, find a maximal matching in each of the two possible induced graphs $G(I_1)$, and pick the option for which the maximal matching found is smaller.) Set $C_1 = V \setminus I_1$.
2. Set $C_2^0 = C_1, I_2^0 = I_1$. Iterate $j = 1, 2, \dots, \log n$:
for every vertex v : if $\deg(v)_{I_2^{j-1}} < \alpha d/2$, then $v \in I_2^j$; otherwise, $v \in C_2^j$.
3. (a) Set $I_3 = I_2^{\log n}$, $C_3 = C_2^{\log n}$, $OUT_3 = \emptyset$.
(b) For every edge (u, v) such that both u, v are in I_3 , move u, v to OUT_3 .
(c) A vertex $v \in C_3$ is *removable* if $\deg(v)_{I_3} < 4$.
Iteratively: find a removable vertex v , and move it from C_3 to OUT_3 . If v has neighbors in I_3 , move these neighbors from I_3 to OUT_3 .
4. Find a maximum independent set in $G[OUT_3]$ (this will be shown to be doable in polynomial time; see Corollary 3.2). Output the union of this independent set and I_3 .

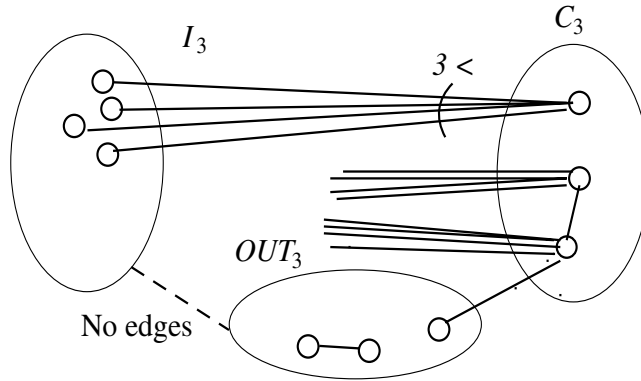


FIG. 2. *FindIS* step 3 outcome.

DEFINITION 3.1. The 2-core of a graph G is the maximal subgraph in which the minimal degree is 2.

It is easy to see that the 2-core is unique and can be found by iteratively removing vertices whose degree < 2 .

THEOREM 3.2. *W.h.p. the largest connected component in the 2-core of $G[OUT_3]$ has cardinality of at most $2 \log n$.*

Let G be any graph. Those vertices of G that do not belong to the 2-core form trees. Each such tree is either disconnected from the 2-core or it is connected by exactly one edge to the 2-core. To find a maximum independent set of $G[OUT_3]$ we need to find a maximum independent set in each connected component of $G[OUT_3]$ separately. For each connected component D_i of $G[OUT_3]$ we find the maximum independent set as follows: let C_i be the intersection of D_i with the 2-core of $G[OUT_3]$. We enumerate all possible independent sets in C_i (there are at most $2^{|C_i|} \leq 2^{2 \log n}$

possibilities), and each one of them can be optimally extended to an independent set of D_i by solving (separately) a maximum independent set problem on each of the trees connected to C_i . For some trees we may have to exclude the tree vertex which is connected to D_i if it is connected to a vertex of the independent set that we try to extend. On each tree the problem can be solved by dynamic programming.

COROLLARY 3.2. *A maximum independent set in $G[OUT_3]$ can be found in polynomial time.*

3.1. Dense sets, degree deviations, and other properties. In proving the correctness of the algorithm, we will use structural properties of the random graph G . In particular, such a random graph most likely has no small dense sets (small sets of vertices that induce many edges). This fact will be used on several occasions to derive a proof by contradiction. Namely, certain undesirable outcomes of the algorithm cannot occur, as otherwise they will lead to a discovery of a small dense set in G . Most of the following lemmas are standard, and those proofs not given in this section will be given in section 4.

LEMMA 3.3. *Let G be a random $G_{n,p,\alpha}$ ($p = \frac{d}{n}$, $d < n^{1/40}$) random graph. W.h.p. the following hold:*

1. *There is no set $U \subset V$ of cardinality smaller than $2n/d^5$ such that the number of edges induced by U is more than $\frac{4}{3}|U|$.*
2. *There is no set $U \subset V$ of size $< \alpha n/40$ containing $\alpha d|U|/14$ edges (and this happens with probability of at least $> 1 - o(n^{-\sqrt{d}})$).*
3. *There is no $C' \subseteq C$ such that $n/2d^5 \leq |C'| \leq \frac{2n \log d}{d}$ and $|\Gamma(C') \cap I| \leq |C'|$.*

COROLLARY 3.4. *Let G be a graph which has the property from Lemma 3.3, part 1. Let A, B be any two disjoint sets of vertices each of size smaller than n/d^5 . If every vertex of B has at least 2 edges going into A , then $|A| \geq |B|/2$.*

Proof. Assume for the sake of contradiction that $|A| = \delta|B|$ for some $0 < \delta < 1/2$. The number of internal edges of $A \cup B$ is at least $2|B| > \frac{4}{3}(1 + \delta)|B| = \frac{4}{3}|A \cup B|$. The last inequality contradicts part 1 of Lemma 3.3. $\square$

The following lemma bounds the number of vertices whose degree largely deviates from its expectation.

LEMMA 3.5. *Let $d < n^{1/40}$. With probability $> 1 - e^{-n^{0.1}}$ the following hold:*

1. *There are at most n/d^{21} vertices in C whose degree into I is $< 0.9\alpha d$.*
2. *There are at most $3e^{-d}dn$ edges that contain a vertex with degree at least $3d$.*

LEMMA 3.6. *W.h.p. the size of a maximum independent set in $G[C]$ is $< \frac{2n \log d}{d}$.*

Proof. The probability for an independent set of size k is at most

$$(1) \quad \binom{n}{k} \left(1 - \frac{d}{n}\right)^{k(k-1)/2} \leq \left(\frac{ne}{k}\right)^k e^{-\frac{dk(k-1)}{2n}} = e^{k(\log(\frac{n}{k}) + 1 - \frac{d(k-1)}{2n})}.$$

By setting $k = \frac{2n \log d}{d}$ the above term becomes

$$(2) \quad e^{k(\log d - \log \log d - \log 2 + 1)} e^{-k(\log d - o(1))} = e^{-k(\log \log d + \log 2 - 1 - o(1))} = o(1). \quad \square$$

The following lemma differs from previous lemmas in this section in the sense that it does not refer to a “static” property that the input graph G is likely to have but rather to the likely outcome of step 1 of the algorithm $FindIS(G)$. It also differs from other lemmas in this section in the sense that its proof is considerably more complicated.

LEMMA 3.7. *Let $\sqrt{\frac{c_0}{d}} \leq \alpha \leq 1/3$, $d < n^{1/40}$. W.h.p. $|I_1 \triangle I| < |I|/60$.*

DEFINITION 3.8. *An instance G (taken from $G_{n,d/n,\alpha}$) is good if it has all of the properties listed in Lemmas 3.3, 3.5, 3.6, and 3.7.*

3.2. Proof of Theorem 3.1. In the proof of Theorem 3.1 we will assume that $d < n^{1/40}$. When d is very large (e.g., $d = n/2$), some of the lemmas that we use, such as Lemma 3.3, are not true. However, when d is very large, these lemmas are also not needed. When $d > n^{1/40}$, the planted independent set is almost surely the unique maximum independent set and can be found by using algorithms described in earlier work [2, 4].

When $\alpha \geq 1/3$, step 1 of the algorithm can be replaced by the following simpler step: put in I_1 all of the vertices of degrees $< (1 - \alpha/2)d$ and the rest of the vertices in C_1 . A simple argument (similar to the proof of Lemma 3.5) shows that in this case w.h.p. $|I \triangle I_1| < |I|/60$ (see [10] for more details). Hence when analyzing step 1 of the algorithm (proving that w.h.p. $|I \triangle I_1| < |I|/60$), we will without loss of generality limit the range of α to $\alpha < 1/3$.

We will show that if G is good, then steps 1–3 of the algorithm give a very good approximation of I, C .

3.2.1. An aid in the analysis. For the sake of the analysis, we introduce Definition 3.9. This definition identifies certain key sets of vertices in the graph (V_2 and V_3). Definition 3.9 involves an algorithm that has complete knowledge of the planted independent set I . To avoid the possibility of confusion, let us stress here that this algorithm is not part of the algorithm *FindIS*, and, moreover, the sets V_2 and V_3 are neither given to nor computed by the algorithm *FindIS*.

DEFINITION 3.9. *Given a graph $G = (V, E)$ with an independent set I (and a corresponding vertex cover $C = V \setminus I$), the P-core of G is a subset of V which is extracted by using the following steps:*

- CR0:** *Initialization:* $V_2 = I \cup \{v \in C \mid \deg(v)_I \geq 0.9\alpha d\}$.
 Iteratively: (i) If there is $v \in I \cap V_2$ with $\deg(v)_{V \setminus V_2} > \alpha d/4$,
 remove v from V_2 .
 (ii) If there is $v \in C \cap V_2$ with $\deg(v)_{I \cap V_2} < 0.8\alpha d$,
 remove v from V_2 .
- CR1:** *Initialization:* $V_3 = V_2$,
 remove from V_3 all of the vertices of $V_3 \cap I$ that have
 edges to $V \setminus V_3$.
 Iteratively: Find a vertex $v \in V_3 \cap C$ with $\deg(v)_{V_3 \cap I} < 4$, and
 remove v and its neighbors in I from V_3 .

The P-core of G is defined to be V_3 . We use $\overline{P\text{-core}}$ to denote $V \setminus P\text{-core}$.

The sets V_2 and V_3 have certain structural properties, as defined in Definition 3.9. We shall also need the fact that if G is a good graph (in the sense of Definition 3.8), then almost all vertices of G are in V_2 and in V_3 .

LEMMA 3.10. *For a good graph G , the following hold:*

1. $|V \setminus V_2| < n/d^{20}$ (where $V_2 = \text{CR0}(G)$).
2. $|\overline{P\text{-core}}| < n/d^{18}$.

Proof. Since G is good, after setting $V_2 = I \cup \{v \in C \mid \deg(v)_I \geq 0.9\alpha d\}$ (and before the iterations of CR0) it holds that $|V \setminus V_2| < n/d^{21}$ (see Lemma 3.5, part 1). In the iteration process, every vertex that we remove from V_2 contributes at least $0.1\alpha d$ edges to $V \setminus V_2$. If the iteration steps are repeated too many times, $V \setminus V_2$ will become dense. Assume by contradiction that at some point the set $V \setminus V_2$ doubled its size (when we compare it to the size before the first iteration). At this point it contains at least $\frac{1}{2}|V \setminus V_2|0.1\alpha d$ edges, and its size is at most $2n/d^{21}$, which contradicts the definition of good G (see Lemma 3.3, part 2). This proves item 1 of the lemma.

We now prove item 2 of the lemma. Initially, $V_3 = V_2$; at this point, $|V \setminus V_3| \leq$

n/d^{20} (using part 1 of this lemma). We then remove from V_3 all of the vertices of I that have edges to $V \setminus V_3$. By doing so, we lose at most $3dn/d^{20} + 3de^{-d}n$ vertices (see the definition of good G in Lemma 3.5, part 2; for a random G this happens with probability $> 1 - e^{-n^{0.1}}$). At this point (just before the iteration steps) $|V \setminus V_3| \leq 4n/d^{19}$. We now begin the iteration process. In every iteration we remove a vertex u of $V_3 \cap C$ whose degree into $V_3 \cap I$ became < 4 (we also remove its neighbors from $V_3 \cap I$). Note that after the initialization of CR1 the degree of u into $V_3 \cap I$ was $\geq 0.8\alpha d$. It follows that after removing u and its neighbors from $V_3 \cap I$ there are at least $0.8\alpha d$ additional edges in $V \setminus V_3$. If the iteration step is repeated too many times, the set $V \setminus V_3$ will become too dense. Assume by contradiction that at some point (for the first time during the iterations) $|V \setminus V_3|$ doubled its size when compared to the size of $V_3 \setminus V$ before the first iteration. At this point the number of edges inside $V \setminus V_3$ is at least $\frac{1}{2}|V \setminus V_3| \cdot \frac{1}{4} \cdot 0.8\alpha d$. Moreover, the size of $V \setminus V_3$ at this point is $\leq 8n/d^{19} + 3 < n/d^5$. This cannot happen in a good G (see Lemma 3.3, part 2; for a random G this happens with probability $o(n^{-\sqrt{d}})$). We conclude that $|V \setminus V_3| < n/d^{18}$. $\square$

Remark. The reader can verify by inspection that those properties of good graphs that are used in the proof of Lemma 3.10 hold with probability at least $1 - o(n^{-\sqrt{d}})$ (for a random graph taken from $G_{n,d/n,\alpha}$). Hence $\Pr[\overline{\text{P-core}} < n/d^{18}] > 1 - o(n^{-\sqrt{d}})$.

3.2.2. Analysis of step 1. Lemma 3.7 states that the spectral phase of the algorithm (step 1) is likely to give a good approximation of the planted independent set. The proof of Lemma 3.7 follows from known principles, but it is rather involved. It is given in section 4.2.

3.2.3. Analysis of step 2. In this section, we assume that G is good (as in Definition 3.8). The approximation I_1, C_1 serves as a bootstrap for the “error reduction” done at step 2. In Lemma 3.11 we show that step 2 significantly reduces the error term, i.e., $|I_2 \triangle I| < n/d^{20} < |I|/d^{19}$.

The idea of the proof is as follows. Recall that in Lemma 3.10 we showed that the set $V_2 = \text{CR0}(G)$ from Definition 3.9 is of size $> (1 - 1/d^{20})n$. Here we will use the assumption that G is good to show that every iteration of step 2 of *FindIS* reduces the number of errors (with respect to I, C) in V_2 by a factor of 2. It then follows that, after step 2 is done, all of the vertices of V_2 are assigned correctly.

LEMMA 3.11 (error reduction). *Let G be a good graph, $V_2 = \text{CR0}(G)$, and $I_2 = I_2^{\log n}$ from step 2 of *FindIS*(G). Then $V_2 \cap (I \triangle I_2) = \emptyset$ and $|I_2 \triangle I| \leq n/d^{20}$.*

Proof. When step CR0 ends, every vertex of $C \cap V_2$ has at least $0.8\alpha d$ edges to $I \cap V_2$, and every vertex of $I \cap V_2$ has at most $\alpha d/4$ edges to $V \setminus V_2$.

For each iteration i of step 2 of *FindIS*(G), define the respective “error” E_i to be the set of those vertices of V_2 that are assigned incorrectly, namely, $E_i = V_2 \cap (I_2^i \triangle I)$. We will show that each iteration of step 2 of *FindIS*(G) reduces the error by a factor of at least 2. This is based on the fact (to be shown shortly) that every vertex of E_i has at least $\alpha d/4$ edges to E_{i-1} . Based on this fact, we show that an event $|E_i| \geq |E_{i-1}|/2$ would lead to a contradiction. Fix E'_i to be any subset of E_i whose size is $|E_{i-1}|/2$ (if $|E_{i-1}|$ is odd, we take a subset of size $\lceil |E_{i-1}|/2 \rceil$, and essentially the same argument goes through). The set $E'_i \cup E_{i-1}$ is of cardinality $\leq \frac{3}{2}|E_{i-1}|$, and it contains at least $\frac{\alpha d}{4}|E_{i-1}|/2$ edges. This contradicts the definition of good G as $|E_{i-1}| \leq |E_0| \leq \alpha n/60$ (see Lemma 3.3, part 2, and note that $|E'_i \cup E_{i-1}| \leq \frac{3}{2}|E_{i-1}| \leq \alpha n/40$).

We now show that each vertex of E_i has at least $\alpha d/4$ edges to E_{i-1} .

Case 1. $v \in E_{i-1} \cap E_i$ (either $v \in I \cap C_2^{i-1}$ or $v \in C \cap I_2^{i-1}$): If $v \in I \cap C_2^{i-1}$ (and $v \in E_i$), then in round $i-1$ it has at least $\alpha d/2$ neighbors in I_2^{i-1} , and these

neighbors are in $C \cap I_2^{i-1}$ since $v \in I$. At least $\alpha d/4$ of these neighbors are in V_2 since v has at most $\alpha d/4$ edges to $V \setminus V_2$ (because $v \in I \cap V_2$). Thus v has $> \alpha d/4$ neighbors in E_{i-1} . If $v \in C \cap I_2^{i-1}$ (and $v \in E_i$), then in round $i-1$ it has at most $\alpha d/2$ neighbors in I_2^{i-1} . Since $v \in V_2 \cap C$, it has $0.8\alpha d$ neighbors in $I \cap V_2$, and thus at least $0.3\alpha d$ of them are in $I \cap C_2^{i-1} \subseteq E_{i-1}$.

Case 2. $v \in E_i \setminus E_{i-1}$: If $v \in E_i \setminus E_{i-1} \cap I$, then v was moved from I_2^{i-1} to C_2^i , and therefore it has at least $\alpha d/2$ neighbors in $I_2^{i-1} \cap C$. Among them, at least $\alpha d/4$ belong to V_2 because v has at most $\alpha d/4$ edges to $V \setminus V_2$. If $v \in E_i \setminus E_{i-1} \cap C$, then v was moved from C_2^{i-1} to I_2^i , and therefore it has at most $\alpha d/2$ neighbors in I_2^{i-1} . Since $v \in V_2 \cap C$, it has at least $0.8\alpha d$ neighbors in $I \cap V_2$, among which at least $0.3\alpha d$ are in $I \cap C_2^{i-1} \subseteq E_{i-1}$.

It follows that when step 2 is done all of the vertices of V_2 are correctly assigned by the algorithm. It then follows that $|I_2 \triangle I| \leq |V \setminus V_2| \leq n/d^{20}$, where the last inequality follows from Lemma 3.10, part 1. $\square$

3.2.4. Analysis of step 3. So far we have shown that at most n/d^{20} vertices of I_2, C_2 are wrongly assigned (with respect to I, C). The goal of step 3 is to “clean” I_2, C_2 yielding I_3, C_3 that can be extended into an optimal solution. Before showing that $I_3 \subseteq I_{max}$ (for some maximum independent set I_{max}), we show that the process of “cleaning” in step 3 does not move too many vertices to OUT_3 . This will be used later for proving that $I_3 \subseteq I_{max}$. For the following lemma, recall the notion of $\bar{P}$ -core from Definition 3.9.

LEMMA 3.12. *Let G be a good graph, and let OUT_3 be the outcome of step 3 of $FindIS(G)$. Then $OUT_3 \subseteq \bar{P}$ -core, and $OUT_3 < n/d^{18}$.*

Proof. As $\overline{OUT}_3 = I_3 \cup C_3$ (where I_3, C_3 are from step 3 of $FindIS(G)$) it is enough to show that $V_3 \subseteq I_3 \cup C_3$. Immediately after step 3(a) it holds that $I_3 \cup C_3 = V$ and thus $V_3 \subseteq I_3 \cup C_3$. We will show that this is kept during steps 3(b) and 3(c). Initially (at CR1) $V_3 = V_2$. Removing from V_3 all of the vertices of $V_3 \cap I$ that have edges to $V \setminus V_3$ ensures that there are no edges between vertices of $V_3 \cap I$ and vertices which were assigned incorrectly (all of the vertices of $V_3 \subseteq V_2$ are correctly assigned after step 2 of the algorithm; see Lemma 3.11). Thus, step 3(b) of the algorithm does not touch any vertex of V_3 because it removes only edges that contain at least one wrongly assigned end point. Finally, the iteration process of CR1 ensures that every vertex of $V_3 \cap C$ has at least 4 edges to vertices in $V_3 \cap I$. Since V_3 is a subset of $I_3 \cup C_3$ at the beginning of step 3(c) and there are no wrongly assigned vertices in V_3 , during step 3(c) there will never be a vertex of $V_3 \cap C$ that has fewer than 4 edges to vertices of $V_3 \cap I$. We conclude that $V_3 \subseteq I_3 \cup C_3$ at the end of step 3.

By combining with Lemma 3.10, part 2, we deduce that w.h.p. $OUT_3 < n/d^{18}$. $\square$

As $|I_3 \triangle I| < |I_2 \triangle I| + |OUT_3|$, by using Lemmas 3.11 and 3.12 we deduce the following.

COROLLARY 3.13. *For a good G , it holds that $|I_3 \triangle I| < 2n/d^{18}$.*

It turns out that I_3, C_3 is also a good approximation of I_{max}, C_{max} (I_{max} is any maximum independent set, and $C_{max} = V \setminus I_{max}$).

LEMMA 3.14. *For a good G it holds that $|I_3 \triangle I_{max}| < n/d^5$.*

Proof.

$$|I_{max} \triangle I_3| \leq |I_{max} \triangle I| + |I \triangle I_3|.$$

By Corollary 3.13, $|I_3 \triangle I| < n/d^{18}$. It remains to bound $|I_{max} \triangle I|$:

$$|I_{max} \triangle I| = |I_{max} \setminus I| + |I \setminus I_{max}| \leq 2|I_{max} \setminus I| = 2|I_{max} \cap C|.$$

$I_{max} = (I_{max} \cap I) \cup (I_{max} \cap C)$. One can always replace $I_{max} \cap C$ with $\Gamma(I_{max} \cap C) \cap I$ to get an independent set $(I_{max} \cap I) \uplus (\Gamma(I_{max} \cap C) \cap I)$. The size of a maximum independent set of C is $< \frac{2n \log d}{d}$ (as G is good; see Lemma 3.6). This upper bounds $|I_{max} \cap C|$. From Lemma 3.3, part 3, if $|I_{max} \cap C| > n/(2d^5)$, then $|\Gamma(I_{max} \cap C) \cap I| > |I_{max} \cap C|$, which contradicts the maximality of I_{max} . $\square$

At this point (for a good G) we know that I_3, C_3 have the following two properties:

- (i) the error term $|(I_3 \cap C_{max}) \cup (I_{max} \cap C_3)| \leq |I_{max} \Delta I_3| < n/d^5$;
- (ii) I_3 is an independent set, and every vertex of C_3 has at least 4 neighbors in I_3 .

The above two properties and the fact that G is good imply that $I_3 \subseteq I_{max}$. This is proven in the following lemma.

LEMMA 3.15 (extension lemma). *Let I_m be any independent set of G (the reader may think of I_m as I_{max}), and let $C_m \triangleq V \setminus I_m$. Let I', C', OUT' be an arbitrary partition of V for which I' is an independent set. If the following hold:*

1. $|(I' \cap C_m) \cup (I_m \cap C')| < n/d^5$;
2. every vertex of C' has at least 4 neighbors in I' (none of the vertices of I' have edges to OUT');
3. the graph G has no small dense subsets as described in Lemma 3.3, part 1,

then there exists an independent set I_{new} (and $C_{new} \triangleq V \setminus I_{new}$) such that $I' \subseteq I_{new}$, $C' \subseteq C_{new}$, and $|I_{new}| \geq |I_m|$.

Proof. If we could show that on average a vertex of $U = (I' \cap C_m) \cup (I_m \cap C')$ contributes at least $4/3$ internal edges to U , then U would form a small dense set that contradicts Lemma 3.3. This would imply that $U = (I' \cap C_m) \cup (I_m \cap C')$ is the empty set, and we could take $I_{new} = I_m$ in the proof of Lemma 3.15. The proof below extends this approach to cases where we cannot take $I_{new} = I_m$.

Every vertex $v \in C'$ has at least 4 edges into vertices of I' . Since I_m is an independent set, it follows that every vertex of $I_m \cap C'$ has at least 4 edges into $I' \cap C_m$. To complete the argument we would like to show that every vertex of $I' \cap C_m$ has at least 2 edges into $I_m \cap C'$. However, some vertices $v \in I' \cap C_m$ might have less than two neighbors in $I_m \cap C'$ (see Figure 3). In this case, we will modify I_m to get an independent set I_{new} (and $C_{new} \triangleq V \setminus I_{new}$) at least as large as I_m , for which every vertex of $I' \cap C_{new}$ has 2 neighbors in $I_{new} \cap C'$. This is done iteratively; after each iteration we set $I_m = I_{new}$, $C_m = C_{new}$. Consider a vertex $v \in (I' \cap C_m)$ with $\deg(v)_{I_m \cap C'} < 2$:

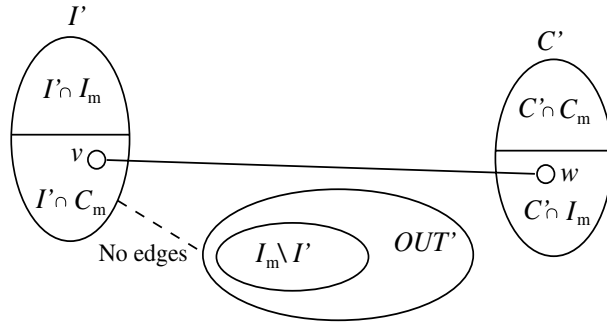


FIG. 3. A vertex $v \in (I' \cap C_m)$ which has strictly less than 2 edges into $I_m \cap C'$.

- If v has no neighbors in $I_m \cap C'$, then define $I_{new} = I_m \cup \{v\}$. I_{new} is an independent set because v (being in I') has no neighbors in I' or in OUT' .

- If v has only one edge into $w \in (I_m \cap C')$, then define $I_{new} = (I_m \setminus \{w\}) \cup \{v\}$. I_{new} is an independent set because v (being in I') has no neighbors in I' or in OUT' . The only neighbor of v in $I_m \cap C'$ is w .

The three properties are maintained also with respect to I_{new}, C_{new} (replacing I_m, C_m): properties 2, 3 are independent of the sets I_m, C_m , and property 1 is maintained since after each iteration it holds that $|(I' \cap C_{new}) \cup (I_{new} \cap C')| < |(I' \cap C_m) \cup (I_m \cap C')|$.

When the process ends, let U denote $(I' \cap C_m) \cup (I_m \cap C')$. Each vertex of $I' \cap C_m$ has at least 2 edges into $I_m \cap C'$, and thus $|I_m \cap C'| \geq \frac{1}{2}|I' \cap C_m|$ (see Corollary 3.4). Each vertex of $I_m \cap C'$ has 4 edges into $I' \cap C_m$, so the number of edges in U is at least $4|I_m \cap C'| \geq 4|U|/3$ and also $|U| < n/d^5$, which implies that U is empty (by Lemma 3.3, part 1). $\square$

Proof of Theorem 3.1. Assume that G is good (w.h.p. this is the case; see section 3.1). By Lemma 3.14, $|I_3 \triangle I_{max}| < n/d^5$. By using Lemma 3.15 (instantiating I_m and I' from Lemma 3.15 to be I_{max} and I_3 , respectively), we derive that I_3 is contained in some maximum independent set. $\square$

It remains to show that the 2-core of OUT_3 has no large connected components.

3.3. Proof of Theorem 3.2. Having established that, for a good G , OUT_3 is small (Lemma 3.12), we would now like to establish that its structure is simple enough to allow one to find a maximum independent set of $G[OUT_3]$ in polynomial time. Establishing such a structure would have been easy if the vertices of OUT_3 were chosen independently at random, because a small random subgraph of a random graph G is likely to decompose into connected components no larger than $O(\log n)$. However, OUT_3 is extracted from G by using some deterministic algorithm and hence might have a more complicated structure. For this reason, we shall now consider the 2-core of $G[OUT_3]$ and bound the size of its connected components. Note that for good G it holds that $OUT_3 \subseteq \overline{\text{P-core}}$ (Lemma 3.12), and thus for such G it is enough to handle $\overline{\text{P-core}}$. This simplifies the proof because the definition of $\overline{\text{P-core}}$ (namely, the procedure that generates it in Definition 3.9) is simpler than that of OUT_3 .

To prove that the 2-core of $G[\overline{\text{P-core}}]$ has no large connected component, we shall introduce a notion of a *balanced* connected component. A set of vertices $U \subset V$ will be called *balanced* if it contains at least $|U|/3$ vertices from C . The following lemma shows the relevance of the notion of being balanced to our context.

PROPOSITION 3.3. *For good G , every connected component of the 2-core of $G[\overline{\text{P-core}}]$ is balanced.*

Proof. Let A_i be such a connected component. Every vertex of A_i has degree of at least 2 in the 2-core (of $G[\overline{\text{P-core}}]$). Since A_i is a connected component of the 2-core, it follows that every vertex of A_i has degree of at least 2 inside A_i . $|A_i| \leq |\overline{\text{P-core}}| < \frac{n}{d^5}$. If $\frac{|A_i \cap I|}{|A_i|}$ is more than $\frac{2}{3}$, then the number of internal edges in A_i is $> 2 \cdot \frac{2}{3}|A_i| > \frac{4}{3}|A_i|$, which contradicts Lemma 3.3, part 1. $\square$

The following lemma shows that a connected component is balanced only if it contains a sufficiently large balanced tree.

LEMMA 3.16. *Let G be a connected graph whose vertices are partitioned into two sets: C and I . Let $\frac{1}{k}$ be a lower bound on the fraction of C vertices, where k is an integer. For any $1 \leq t \leq |V(G)|/2$ there exists a tree whose size is in $[t, 2t-1]$ and at least $\frac{1}{k}$ fraction of its vertices belong to C .*

Proof. We use the following well-known fact: any tree T contains a *center* vertex v such that each subtree hanging on v contains strictly less than half of the vertices of T .

Let T be an arbitrary spanning tree of G . Observe that at least a fraction of $\frac{1}{k}$

of its vertices belongs to C . We show that any such tree T has a subtree T' of size $|T|/2 < |T'| < |T|$ with at least a fraction of $\frac{1}{k}$ of its vertices in C . Repeating this argument until the first time that $|T'| < 2t$ proves the lemma.

Let v be the center of T , and let $T_1, \dots, T_k$ be the subtrees hanging on v . Let T_j be the subtree with the smallest fraction of C vertices, and take T' to be the tree that remains by removing T_j from T . By properties of the center vertex, $|T|/2 < |T'| < |T|$, as desired. Moreover, if the fraction of C vertices in T_j is $\frac{1}{k}$ or less, then clearly the fraction of C vertices in T' is at least $\frac{1}{k}$. If the fraction of C vertices in T_j is strictly more than $\frac{1}{k}$, then this holds also for all other subtrees of T . By the integrality of k , this implies that the fraction of C vertices in T' is at least $\frac{1}{k}$ (because then $k|C \cap T'| > |T'| - 1$ implies $k|C \cap T'| \geq |T'|$). $\square$

To proceed with our discussion, we introduce the following definition.

DEFINITION 3.17. *We say that G is extension friendly if $G[\overline{P\text{-core}}]$ has no balanced trees of size in $[\log n, 2 \log n]$.*

In Lemma 3.18 to follow, we will see that G is likely to be extension friendly. This is a key component in the proof of Theorem 3.2.

Proof of Theorem 3.2. W.h.p. it holds that G is both good and extension friendly (see section 3.1 and Lemma 3.18). Since G is good, Proposition 3.3 implies that every connected component of the 2-core of $G[\text{P-core}]$ is balanced. Furthermore, any balanced connected component of size at least $2 \log n$ (in vertices) must contain a balanced tree of size in $[\log n, 2 \log n - 1]$ (see Lemma 3.16). Since G is extension friendly, the 2-core of $G[\text{P-core}]$ does not contain a balanced tree with size $[\log n, 2 \log n]$. By combining these three facts we derive that the 2-core of $G[\text{P-core}]$ has no connected components of size $> 2 \log n$. Note that since G is good it holds that $\text{OUT}_3 \subseteq \overline{\text{P-core}}$, and thus also the 2-core of $G[\text{OUT}_3]$ has no connected components of size more than $2 \log n$. $\square$

We will now show that G is likely to be extension friendly.

LEMMA 3.18. *Assume that G is taken from $G_{n,d/n,\alpha}$. W.h.p. $G[\overline{P\text{-core}}]$ has no balanced tree of size in $[\log n, 2 \log n]$.*

Proof. As we have seen in Lemma 3.12, $|\overline{\text{P-core}}|$ is likely to be very small, smaller than n/d^{18} . If we were to pick at random such a small subgraph of G , then it would have been very sparse (average degree at most d^{-17}), making it highly unlikely to contain any connected component of size $\log n$ or more and hence also no tree (let alone a balanced tree) of size $\log n$ or more. The difficulty in proving Lemma 3.18 is that $\overline{\text{P-core}}$ is not a random set of vertices in G . To overcome this difficulty, we look more closely at the random processes by which G is generated and decouple the random decisions that govern which set of vertices comprise $\overline{\text{P-core}}$ from the random decisions that govern which edges are present in $G[\overline{\text{P-core}}]$. To achieve this decoupling, we shall modify the notion of a P -core and slightly enlarge $\overline{\text{P-core}}$.

We now proceed with the detailed proof. We start with the empty graph (a set V of n vertices and no edges) and build the random graph G as we go.

Step 1. Choose an arbitrary partition of the vertex set V into I and C , with $|I| = \alpha n$.

Step 2. Choose a value t for the size of the balanced tree T . As $\log n \leq t \leq 2 \log n$, there are $\log n + 1$ possible ways of choosing t .

For a tree T , let $I(T)$ denote those vertices of T that are in I , and let $C(T)$ denote those vertices of T that are in C .

Step 3. Choose $t_1 = |I(T)|$, the number of vertices of T in I . Observe that $t_1 \leq 2t/3$ because the tree is balanced. Let $t_2 = |C(T)| = t - t_1 \geq t/3$. There are at

most $2t/3 + 1 \leq 2 \log n$ ways of choosing t_1 .

Step 4. Choose $I(T)$, the vertices of I that will be in the tree T . There are $\binom{\alpha n}{t_1}$ ways of doing this.

Step 5. In this step we start constructing the edge set E of the random graph G . For every pair of vertices $u \in C$ and $v \in I \setminus I(T)$ we include the edge (u, v) in E with probability $p = d/n$. For other pairs of vertices, we do not yet decide whether to include an edge or not. We call the graph obtained after this stage G' .

Step 6. At this step we construct a P' -core which is a modified version of the P -core. Its construction follows the construction of a P -core, except for two differences: it is constructed based on G' rather than on G , and the vertices of $I(T)$ are forced not to be part of the P' -core. A detailed description of the procedure for constructing the P' -core will be presented in Definition 3.20. The key properties of the P' -core that we shall need are summarized in the following lemma.

LEMMA 3.19. *For the construction of P' -core as described in Definition 3.20, the following hold:*

1. *Regardless of the choice of $I(T)$ and regardless of which other edges are added to G' so as to create the final graph G , P' -core is always contained in P -core (where P -core is constructed with respect to G in Definition 3.9).*
2. *The bad event of $|\overline{P'-core}| > n/d^{18}$ occurs with probability at most $n^{-\sqrt{d}}$ (here probability is taken over choice of edges in G'). Hence the complementary good event $|\overline{P-core}| \leq n/d^{18}$ happens almost surely.*

We now assume Lemma 3.19, deferring its proof to later, and continue with the proof of Lemma 3.18.

Step 7. Choose the vertices in $C(T)$. If the good event in Lemma 3.19 occurs, then there are at most $\binom{n/d^{18}}{t_2}$ ways of doing this. If the bad event occurs, then there are $\binom{n}{t_2}$ ways of doing this.

Step 8. Having chosen all vertices of T , we need to choose those pairs of vertices that form the tree edges. We call it the interconnection pattern $E(T)$. There are at most t^{t-2} ways of doing this (which is the number of labeled trees on t vertices). The number may actually be much smaller than t^{t-2} , because there cannot be any edges between vertices in the set $I(T)$.

Step 9. Only now do we complete the construction of the random graph G . For every pair of vertices $u, v \in C$, the edge (u, v) is included in E with probability p . Likewise, for every pair of vertices $u \in C$ and $v \in I(T)$, the edge (u, v) is included in E with probability p . Now we compute the probability that all edges of $E(T)$ are in. This probability depends only on Step 9 and is exactly p^{t-1} , regardless of the outcome of Steps 1–8 above.

We can now upper bound the expected number of balanced trees T in $G[\overline{P-core}]$ with t vertices and $\log n \leq t \leq 2 \log n$. Let us fix a value for t and a value for $t_1 = |I(T)|$. By Steps 2 and 3 there are at most $2(\log n)^2$ ways of doing so. By Step 4, there are $\binom{\alpha n}{t_1}$ ways of choosing $I(T)$. By item 1 in Lemma 3.19 it suffices to upper bound the number of balanced trees in $G[\overline{P'-core}]$, as this will also be an upper bound on the number of balanced trees in $G[\overline{P-core}]$. By Step 7, the expected number of ways of choosing the vertices of $C(T)$ in $\overline{P'-core}$ is at most $\binom{n/d^{18}}{t_2} + n^{-\sqrt{d}} \binom{n}{t_2}$. By Step 8, the number of interconnection patterns to consider is at most t^{t-2} . Finally, by Step 9, the probability that all edges of the interconnection pattern are in G is $(d/n)^{t-1}$. Multiplying out all of the terms, the expected number of balanced trees is

at most

$$2(\log n)^2 \binom{\alpha n}{t_1} \left(\binom{n/d^{18}}{t_2} + n^{-\sqrt{d}} \binom{n}{t_2} \right) t^{t-2} \left(\frac{d}{n} \right)^{t-1}.$$

The above expression is much smaller than 1. To see this, observe that for sufficiently large d and $t_2 \leq 2 \log n$ the term $n^{-\sqrt{d}} \binom{n}{t_2}$ is negligible compared to $\binom{n/d^{18}}{t_2}$. Observe also that $\frac{t^t}{t_1^{t_1} t_2^{t_2}} \leq 2^t$ because $t_1 + t_2 = t$. By using these observations, the above expression can be upper bounded by essentially

$$\frac{(2e)^t (\log n)^2 \alpha^{t_1} d^{t-1} n}{d^{18t_2}}.$$

By observing further that $t_1 \leq t$, $t_2 \geq t/3$, $\alpha < 1$, and using the fact the d is sufficiently large and hence larger than $2e$, this last expression is upper bounded by $n(\log n)^2/d^{4t}$. Using the fact that $t \geq \log n$ now implies that the expected number of balanced trees is at most $1/n^2$, and hence with probability at least $1 - 1/n^2$ there is no balanced tree of size between $\log n$ and $2 \log n$ in $G[\text{P-core}]$. $\square$

We present here the detailed definition for the P' -core.

DEFINITION 3.20. *Given a graph G' with the vertex set partitioned into I and C as in Step 1 of the proof of Lemma 3.18, a set $I(T) \subset I$ as in Step 4 of the proof of Lemma 3.18, and an edge set as in Step 5 of the proof of Lemma 3.18, the P' -core of G' is a subset of V which is extracted by using the following steps:*

- CR0':** *Initialization:* $V'_2 = (I \setminus I(T)) \cup \{v \in C \mid \deg(v)_{I \setminus I(T)} \geq 0.9\alpha d\}$.
 Iteratively: (i) If there is $v \in I \cap V'_2$ with $\deg(v)_{V \setminus V'_2} > \alpha d/4$,
 remove v from V'_2 .
 (ii) If there is $v \in C \cap V'_2$ with $\deg(v)_{I \cap V'_2} < 0.8\alpha d$,
 remove v from V'_2 .
- CR1':** *Initialization:* $V'_3 = V'_2$,
 remove from V'_3 all of the vertices of $V'_3 \cap I$ that have
 edges to $V \setminus V'_3$.
 Iteratively: Find a vertex $v \in V'_3 \cap C$ with $\deg(v)_{V'_3 \cap I} < 4$, and
 remove v and its neighbors in I from V'_3 .

The P' -core of G is V'_3 .

We now prove Lemma 3.19.

Proof of Lemma 3.19, part 1. The notion of a P-core is defined in Definition 3.9. The notion of a P' -core is defined in Definition 3.20. Both definitions are similar in many respects and differ only in the following details:

1. $V'_2 \cap I$ is initialized to be smaller than $V_2 \cap I$, because all vertices of $I(T)$ are excluded from it. This is in agreement with our goal of showing that P' -core is contained in P-core.
2. $V'_2 \cap C$ is initialized to be smaller than $V_2 \cap C$, because vertices in C have fewer edges into $I \setminus I(T)$ than they have into I and hence less opportunity of achieving a degree of $0.9\alpha d$. Again, this is in agreement with our goal of showing that P' -core is contained in P-core.
3. The notion of P-core refers to the graph G , whereas the notion of P' -core refers to the graph G' . This difference is irrelevant. The two graphs completely agree on all edges between C and $I \setminus I(T)$. The procedure for constructing P' -core looks only at those edges and no other edges. It would produce exactly the same P' -core even if it is run on the graph G instead of G' .

Having understood the differences between the constructions of P-core and P'-core, it is easy to verify by inspection that every vertex that is removed from P-core will be removed from P'-core as well, and hence P'-core is contained in P-core, as desired. $\square$

Proof of Lemma 3.19, part 2. The proof is similar to the proof of Lemma 3.10 (see also the remark after Lemma 3.10). The difference is that in the initialization of CR0' we remove the vertices of $I(T)$ and ignore the edges of C connected to these vertices. The fact that $|I(T)| < 2 \log n$ together with the fact that the number of vertices of C affected by the removal of $I(T)$ is bounded by the sum of degrees in $I(T)$ (this number is small because the probability of G having a vertex of degree above $d^2 \log n$ is $O(n^{-d})$) imply that the total number of vertices removed by the initializations of CR0' and CR0 differ by $O(d^2 \log n)^2 = o(n/d^{20})$. The rest of the proof of Lemma 3.10 applies without change. $\square$

4. Proofs of the technical lemmas.

4.1. Expansion properties and degrees deviation.

Proof of Lemma 3.3. Let $c > 1$ and $k \geq 3$ (k is an integer) satisfy the following three inequalities:

- (i) $c < d$,
- (ii) $k < n/e$,
- (iii) $\left(\frac{d}{c}\right)^c \left(\frac{ke}{n}\right)^{c-1} e^2 < 1/2$.

The probability that there is a set U of cardinality at most k with $c|U|$ internal edges is at most

$$\begin{aligned}
 \sum_{i=3}^k \binom{n}{i} \binom{\binom{i}{2}}{\lceil ic \rceil} \left(\frac{d}{n}\right)^{\lceil ic \rceil} &\leq \sum_{i=3}^k \left(\frac{ne}{i}\right)^i \left(\frac{i^2 e}{2ic}\right)^{\lceil ic \rceil} \left(\frac{d}{n}\right)^{\lceil ic \rceil} \\
 (3) \quad &\leq \sum_{i=3}^k \left(\frac{d}{c}\right)^{\lceil ic \rceil} \left(\frac{ie}{n}\right)^{\lceil ic \rceil - i} e^{2i} \stackrel{(i), (ii)}{\leq} \sum_{i=3}^k \left(\frac{d}{c}\right)^{ic+1} \left(\frac{ie}{n}\right)^{ic-i} e^{2i} \\
 &= \frac{d}{c} \sum_{i=3}^k \left[\left(\frac{d}{c}\right)^c \left(\frac{ie}{n}\right)^{c-1} e^2 \right]^i \stackrel{(iii)}{\leq} \frac{2d}{c} \left(\left(\frac{d}{c}\right)^c \left(\frac{3e}{n}\right)^{c-1} e^2 \right)^3.
 \end{aligned}$$

Proof of part 1. Set $c := 4/3$, $k := 2n/d^5$. For $d > d_0$ conditions (i)–(iii) hold. The last term in (3) is at most $\frac{6e^7 d^5}{n}$.

Proof of part 2. Set $c := \alpha d/14$, $k := \alpha n/40$. For $d > d_0$ conditions (i)–(iii) hold. The last term in (3) is

$$\frac{2d}{c} \left(\left(\frac{d}{c}\right)^c \left(\frac{3e}{n}\right)^{c-1} e^2 \right)^3 \leq \frac{(3ed)^{3c+3}}{n^{3c-3}} = o(n^{-\sqrt{d}}),$$

where in the last inequality we used $d < n^{1/40}$ and $c = \alpha d/20 \geq c_0 \sqrt{d}/20 \gg \sqrt{d}$ (for large enough c_0).

Proof of part 3. We first show that w.h.p. there is no set $U \subset V$ of size $< \frac{10n \log d}{d}$ containing $50(\log d)|U|$ edges. Set $c := 50 \log d$ and $k := \frac{10n \log d}{d}$. For $d > d_0$ conditions (i)–(iii) hold, and the last term in (3) is $o(1)$. It thus follows that w.h.p. there is no set U of cardinality $\leq \frac{10n \log d}{d}$ that contains at least $50 \log d |U|$ edges.

By contradiction, assume that there is a bad set C' for which $|\Gamma(C') \cap I| \leq |C'|$ and $n/2d^5 \leq |C'| \leq \frac{2n \log d}{d}$. By Lemma 3.5, part 1, at least $|C'| - n/d^{20} > \frac{9}{10}|C'|$ vertices of C' have at least $\frac{\alpha d}{2}$ edges to I . It follows that $C' \cup (\Gamma(C') \cap I)$ has at least

$\frac{\alpha d}{5}|C' \cup (\Gamma(C') \cap I)| > \sqrt{d}|C' \cup (\Gamma(C') \cap I)|$ internal edges and its cardinality is at most $2|C'| < \frac{4n \log d}{d} < \frac{10n \log d}{d}$. By the first part of the proof w.h.p. such a dense set does not exist. $\square$

We now prove Lemma 3.5.

Proof of part 1. The degrees into I are independent random variables. Set $\delta = 1/d^{21}$. For a fixed set of size δn the expected sum of degrees is $\mu = \delta n \alpha d$. A bad set has only $0.9\delta n \alpha d$ edges to I . The probability for a bad set of size δn is bounded by

$$\begin{aligned} \binom{n}{\delta n} e^{-\frac{1}{2}(0.1)^2 \mu} &\leq e^{-\delta n (\alpha d / 200 - \log(e/\delta))} \leq e^{-\delta n (\sqrt{c_0 d} / 200 - 21 \log d - 1)} \\ &\leq e^{-\delta n} < e^{-n/d^{21}} < e^{-n^{0.4}}. \end{aligned}$$

In the last inequality we used $d < n^{1/40}$. $\square$

Proof of part 2. The proof of this lemma is very similar to the proof of (26), and details are omitted. $\square$

4.2. Spectral approximation (proof of Lemma 3.7). Recall that eigenvectors of the adjacency matrix of the graph G are of dimension n , equal to the number of vertices of G . Every coordinate of an eigenvector can be naturally associated with a vertex in the graph.

Let V' be the set of vertices of G with degree $< 5d$, and set $n' \triangleq |V'|$. Note that n' is also the dimension of A' —the adjacency matrix of $G[V']$. Let $I' = V' \cap I$ and $C' = V' \cap C$, and set $\alpha' \triangleq \frac{|I'|}{|V'|}$.

Denote by $\bar{A}'$ the $n' \times n'$ matrix such that $\bar{A}'_{i,j} = 0$ for any $\{i, j\} \subset I'$ and $\bar{A}'_{i,j} = p = d/n$ for the other entries. We will use the fact that $\bar{A}'$ (which is the “expectation” of A' if we ignore the diagonal) is almost surely a good spectral approximation of A' (i.e., the spectral norm of $A' - \bar{A}'$ is small). The rank of $\bar{A}'$ is 2, and it has two nonzero eigenvalues. Each of the two nonzero eigenvectors which we denote by $\bar{v}_1, \bar{v}_{n'}$ is constant on I' and constant on C' (this follows from symmetry). Given that each one of $\bar{v}_1, \bar{v}_{n'}$ has only two values, we need to find $\bar{\beta}, \bar{\lambda}$ which satisfy

$$\begin{array}{c} \overbrace{\begin{array}{|ccc|ccc|} \hline 0 & . & . & 0 & p & . & p & p \\ . & . & & & & & & \\ . & & . & & & & & \\ 0 & . & . & 0 & & & & \\ \hline p & & & & p & & & \\ . & & & & & . & & \\ p & & & & & & p & \\ p & & & & & & & p \\ \hline \end{array}}^{I'} \end{array} \begin{bmatrix} 1 \\ . \\ . \\ . \\ 1 \\ \bar{\beta} \\ . \\ . \\ \bar{\beta} \end{bmatrix} = \begin{bmatrix} (1 - \alpha')n'p\bar{\beta} \\ . \\ . \\ . \\ \alpha'n'p + \bar{\beta}(1 - \alpha')n'p \\ . \\ . \\ . \end{bmatrix} = \bar{\lambda} \begin{bmatrix} 1 \\ . \\ . \\ . \\ 1 \\ \bar{\beta} \\ . \\ . \\ \bar{\beta} \end{bmatrix},$$

or equivalently,

$$(4) \quad (1 - \alpha')n'\bar{\beta}p = \bar{\lambda},$$

$$(5) \quad \alpha'n'p + \bar{\beta}(1 - \alpha')n'p = \bar{\lambda}\bar{\beta},$$

and a simple calculation gives a quadratic equation in $\bar{\beta}$ whose solutions are

$$(6) \quad \bar{\beta}_{1,2} = \frac{1}{2} \left(1 \pm \sqrt{1 + \frac{4\alpha'}{1-\alpha'}} \right).$$

By substituting $\bar{\beta}$ at (4) we derive

$$(7) \quad \begin{aligned} \bar{\lambda}_1 &= (1 - \alpha')n' \bar{\beta}_1 p, \\ \bar{\lambda}_{n'} &= (1 - \alpha')n' \bar{\beta}_2 p, \\ \bar{v}_1 &= (\underbrace{1, 1, \dots, 1}_{\alpha' n'}, \underbrace{\bar{\beta}_1, \bar{\beta}_1, \dots, \bar{\beta}_1}_{(1-\alpha')n'}), \\ \bar{v}_{n'} &= (\underbrace{1, 1, \dots, 1}_{\alpha' n'}, \underbrace{\bar{\beta}_2, \bar{\beta}_2, \bar{\beta}_2, \dots, \bar{\beta}_2}_{(1-\alpha')n'}). \end{aligned}$$

Define $\beta_{1,2} := \frac{1}{2}(1 \pm \sqrt{1 + \frac{4\alpha}{1-\alpha}})$.

LEMMA 4.1. (i) For every $x \geq 0$ it holds that $|1 - \sqrt{1+x}| \leq \frac{x}{2}$.

(ii) For $0 \leq x \leq 3$ it holds that $1 - \sqrt{1+x} \leq -\frac{x}{3}$.

Proof of part (i). Note that $1 - \sqrt{1+x} \leq 0$ for $x \geq 0$, and thus it is enough to show that $1 - \sqrt{1+x} \geq -\frac{x}{2}$:

$$-\frac{x}{2} \leq 1 - \sqrt{1+x} \iff \sqrt{1+x} \leq 1 + \frac{x}{2} \iff 1+x \leq 1+x + \frac{x^2}{4}. \quad \square$$

Proof of part (ii).

$$1 - \sqrt{1+x} \leq -\frac{x}{3} \iff 1 + \frac{x}{3} \leq \sqrt{1+x}$$

$$\iff 1 + \frac{2}{3}x + \frac{x^2}{9} \leq 1+x \iff \frac{x^2}{9} \leq \frac{x}{3} \iff \frac{x}{3} \leq 1. \quad \square$$

In the reminder of this section we will use the following facts:

$$(8) \quad \forall x \perp \bar{v}_1, \bar{v}_{n'} \text{ it holds that } \sum_{i \in C'} x_i = 0, \quad \sum_{i \in I'} x_i = 0,$$

$$(9) \quad \alpha' = (1 \pm e^{-\Omega(d)})\alpha, \quad n' \geq (1 - e^{-\Omega(d)})n,$$

$$(10) \quad \sqrt{\frac{c_0}{d}} \leq \alpha \leq 1/3, \quad d < n^{1/40},$$

$$(11) \quad \|\bar{v}_{n'}\|_2^2 \geq \alpha' n' \geq (1 - e^{-\Omega(d)})\alpha n, \quad \|\bar{v}_1\|_2^2 \geq n' \geq (1 - e^{-\Omega(d)})n,$$

$$(12) \quad \beta_2 = (1 \pm e^{-\Omega(d)})\bar{\beta}_2, \quad \beta_1 = (1 \pm e^{-\Omega(d)})\bar{\beta}_1,$$

$$(13) \quad |\bar{\beta}_2| \leq \frac{\alpha'}{1-\alpha'} \text{ (see the definition of } \bar{\beta}_2 \text{ at (6) and Lemma 4.1, part (i))},$$

$$(14) \quad |\bar{\beta}_2 \sqrt{1-\alpha'}| \leq 1, \quad |\beta_2 \sqrt{1-\alpha}| \leq 1.$$

Proof: $\beta_2 \leq 0$, and thus it is enough to show that $\beta_2 \sqrt{1-\alpha} \geq -1$:

$$\beta_2 \sqrt{1-\alpha} = \frac{1}{2}(\sqrt{1-\alpha} - \sqrt{1-\alpha+4\alpha}) \geq -1$$

$$\iff 2 + \sqrt{1-\alpha} \geq \sqrt{1+3\alpha} \iff 4 + 4\sqrt{1-\alpha} + 1 - \alpha \geq 1 + 3\alpha$$

$$\iff 4 + 4\sqrt{1-\alpha} \geq 4\alpha,$$

$$(15) \quad (1 - \alpha')\bar{\beta}_2 \leq -0.66\sqrt{\frac{c_0}{d}}.$$

$$\begin{aligned} \text{Proof: } (1 - \alpha)\beta_2 &= (1 - \alpha)\frac{1}{2}\left(1 - \sqrt{1 + \frac{4\alpha}{1 - \alpha}}\right) \\ &\stackrel{(10)}{\leq} \frac{1}{3}\left(1 - \sqrt{1 + \frac{4\alpha}{1 - \alpha}}\right) \stackrel{(4.1) \text{ part (ii), (10)}}{\leq} -\frac{1}{3}\frac{4\alpha}{3(1 - \alpha)} \stackrel{(10)}{\leq} -\frac{2}{3}\sqrt{\frac{c_0}{d}}. \end{aligned}$$

(15) follows from the last inequality and (9) and (12).

Additional notation. We will use $v_1, v_2, \dots, v_{n'}$ to denote the unit eigenvectors of A' corresponding to the eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n'}$. The vector of all ones is denoted by $\bar{1}$.

LEMMA 4.2. *Let $\bar{v}_{n'}$ be the eigenvector corresponding to the most negative eigenvalue of A' (as defined at (7)). Let $v_{n'}$ be the eigenvector corresponding to the most negative eigenvalue of A' normalized such that $\|v_{n'}\| = \|\bar{v}_{n'}\|$ and also $\langle v_{n'}, \bar{v}_{n'} \rangle \geq 0$. W.h.p. it holds that $\|\bar{v}_{n'} - v_{n'}\|^2 < \frac{1}{800}\|\bar{v}_{n'}\|^2$.*

We first prove Lemma 3.7 and only then give the proof of Lemma 4.2.

Proof of Lemma 3.7. The vector $\bar{v}_{n'}$ equals

$$\underbrace{(1, 1, \dots, 1)}_{\alpha'n'} \underbrace{(\bar{\beta}_2, \bar{\beta}_2, \dots, \bar{\beta}_2)}_{(1-\alpha')n'}$$

($\bar{\beta}_2$ is defined at (6)). In this proof we will assume that the eigenvector $v_{n'}$ used in step 1 of *FindIS* is normalized so that $\|v_{n'}\| = \|\bar{v}_{n'}\|$ and also that $\langle v_{n'}, \bar{v}_{n'} \rangle \geq 0$. We may use this assumption since the outcome of step 1 does not change if we multiply $v_{n'}$ (used by step 1 of the algorithm) by any nonzero constant.

Denote by L the $\alpha'n'$ indices of $v_{n'}$ which correspond to the largest values in $v_{n'}$. Similarly, denote by $\bar{L}$ the $\alpha'n'$ indices of $\bar{v}_{n'}$ which correspond to the largest values in $\bar{v}_{n'}$ (hence $\bar{L} = I'$). For every vertex $i' \in L \cap C$ we match a unique vertex $i \in \bar{L} \setminus L$. Note that $v_{n'}(i') \geq v_{n'}(i)$, $\bar{v}_{n'}(i) - \bar{v}_{n'}(i') = 1 - \bar{\beta}_2$. Summing the last two inequalities gives

$$(16) \quad (\bar{v}_{n'}(i) - v_{n'}(i)) + (v_{n'}(i') - \bar{v}_{n'}(i')) \geq 1 - \bar{\beta}_2 \geq 1$$

and thus

$$(17) \quad (\bar{v}_{n'}(i) - v_{n'}(i))^2 + (v_{n'}(i') - \bar{v}_{n'}(i'))^2 \geq \frac{1}{2}.$$

By Lemma 4.2 w.h.p. it holds that $\|\bar{v}_{n'} - v_{n'}\|^2 \leq \frac{1}{800}\|\bar{v}_{n'}\|^2$. It thus follows that

$$(18) \quad \frac{1}{2}|L \setminus \bar{L}| \leq \frac{1}{800}\|\bar{v}_{n'}\|^2.$$

Option 1. By setting I_1 to contain the indices of the largest αn entries in $v_{n'}$ we get

$$\begin{aligned} |I_1 \setminus I| &\leq |L \setminus \bar{L}| + |\alpha n - \alpha'n'| \leq \frac{1}{400}\|\bar{v}_{n'}\|^2 + |\alpha n - \alpha'n'| \\ (19) \quad &\stackrel{(9)}{\leq} \frac{\alpha'n' + \bar{\beta}_2^2(1 - \alpha')n'}{400} + e^{-\Omega(d)}n \stackrel{(13)}{\leq} \frac{\alpha'n\left(1 + \frac{\alpha'}{1 - \alpha'}\right)}{400} + e^{-\Omega(d)}n \\ &\stackrel{(10)}{\leq} \frac{1.5\alpha'n'}{400} + e^{-\Omega(d)}n \stackrel{(9)}{\leq} \frac{\alpha n}{200}. \end{aligned}$$

In this case the maximum matching in $G[I_1]$ contains at most $2|C' \cap I_1| \leq \frac{\alpha n}{100}$ vertices.

Remark. Note that since $|I| = |I_1|$, (19) implies that $|I \setminus I_1| \leq \frac{\alpha n}{200}$, or in other words: the intersection of I with the indices of the $(1-\alpha)n$ smallest vertices is bounded by $\frac{\alpha n}{200}$.

Option 2. By setting I_1 to contain the indices of the smallest αn entries in $v_{n'}$, we get $|I_1 \cap I| \leq \frac{\alpha n}{200}$. (Observe that $\alpha < 1/3$ and $n' \simeq n$ imply that the largest αn entries are disjoint from the smallest αn entries, and then use the last remark.) The maximum independent set in $G[C]$ is bounded by $\frac{2 \log dn}{d}$ (Lemma 3.6). It follows that the maximum independent set in $G[I_1]$ contains at most $\frac{\alpha n}{200} + \frac{2 \log dn}{d}$ vertices. Since the complement of a maximal matching is an independent set, it follows that every maximal matching in $G[I_1]$ has at least $\frac{199\alpha n}{200} - \frac{2 \log dn}{d}$ vertices.

Since step 1 of the algorithm takes the option with the smallest maximum matching (between 1 and 2), it chooses Option 1. It then follows that

$$(20) \quad |I_1 \triangle I| = 2|I_1 \setminus I| \leq \frac{\alpha n}{100}. \quad \square$$

Proof of Lemma 4.2. For simplicity, we normalize the vectors so that $\|v_{n'}\| = \|\bar{v}_{n'}\| = 1$. We need to prove that $\|\bar{v}_{n'} - v_{n'}\|^2 < \frac{1}{800}$.

Since the vectors $v_1, v_2, \dots, v_{n'}$ are orthogonal, the vector $\bar{v}_{n'}$ can be written as $\sum_{i=1}^{n'} c_i v_i$ such that $\sum_{i=1}^{n'} c_i^2 = 1$. We will use the following properties:

- (i) $\|(A' - \bar{\lambda}_{n'} I) \bar{v}_{n'}\| \leq 2\sqrt{d}$ (see Lemma 4.3, part (ii));
- (ii) all of the eigenvalues of A' except $\lambda_1, \lambda_{n'}$ are bounded by $2c\sqrt{d}$ in absolute value (Lemma 4.4), where c is some absolute constant independent of α and d ;
- (iii) $\bar{\lambda}_{n'} \leq -0.65\sqrt{c_0 d}$:

$$\text{by (7) second line } \bar{\lambda}_{n'} = (1 - \alpha')n' \bar{\beta}_2 p$$

$$\stackrel{(15)}{\leq} -0.66\sqrt{\frac{c_0}{d}} n' p \stackrel{(9)}{\leq} -0.65\sqrt{c_0 d};$$

(iv) $\lambda_1 \geq 0$ (because the trace of A' is 0);

(v) $c_{n'} = \langle v_{n'}, \bar{v}_{n'} \rangle \geq 0$:

$$(21) \quad \begin{aligned} 4d &\stackrel{(i)}{\geq} \|(A' - \bar{\lambda}_{n'} I) \bar{v}_{n'}\|^2 = \left\| (A' - \bar{\lambda}_{n'} I) \left(\sum_{i=1}^{n'} c_i v_i \right) \right\|^2 \\ &= \sum_{i=1}^{n'} (c_i)^2 (\lambda_i - \bar{\lambda}_{n'})^2 \geq \sum_{i=1}^{n'-1} (c_i)^2 (\lambda_i - \bar{\lambda}_{n'})^2. \end{aligned}$$

Facts (iii) and (iv) imply that $\lambda_1 - \bar{\lambda}_{n'} \geq 0.65\sqrt{c_0 d}$. Fact (ii) states that $|\lambda_i| \leq 2c\sqrt{d}$ ($1 < i < n'$) for some constant c independent of α, d . Thus by setting $c_0 > (4c)^2$ we have $\lambda_i - \bar{\lambda}_{n'} > 0.65\sqrt{c_0 d} - 2c\sqrt{d} > \sqrt{c_0 d}/5$. By combining the last inequality with (21) we derive

$$(22) \quad \sum_{i=1}^{n'-1} c_i^2 \leq \frac{4}{c_0/25} < \frac{100}{c_0}.$$

By using

$$(23) \quad \|\bar{v}_{n'} - v_{n'}\|^2 = \sum_{i=1}^{n'-1} c_i^2 + (1 - c_{n'})^2, \quad \sum_{i=1}^{n'} c_i^2 = 1, \quad c_{n'} \geq 0,$$

we derive that for large enough c_0 it holds that $\|\bar{v}_{n'} - v_{n'}\|^2 < \frac{1}{800}$ as needed. $\square$

LEMMA 4.3. *Let $\bar{v}_1$ and $\bar{v}_{n'}$ be the first and last eigenvectors, respectively, of the matrix $\bar{A}'$ with corresponding eigenvalues $\bar{\lambda}_1$ and $\bar{\lambda}_{n'}$. The following inequalities hold w.h.p.:*

- (i) $\|(A' - \bar{\lambda}_1 I)\bar{v}_1\| \leq 2\sqrt{d}\|\bar{v}_1\|$,
- (ii) $\|(A' - \bar{\lambda}_{n'} I)\bar{v}_{n'}\| \leq 2\sqrt{d}\|\bar{v}_{n'}\|$,
- (iii) for all $x \perp \bar{v}_1, \bar{v}_{n'}$ $\|A'x\| \leq c\sqrt{d}\|x\|$ (where $c \geq 2$ is a universal constant independent of d, α).

Proof of Lemma 4.3, part (ii). By (11), $\|\bar{v}_{n'}\| \geq \sqrt{\alpha n}(1 - e^{-\Omega(d)})$. We will prove that $\|(A' - \bar{\lambda}_{n'} I)\bar{v}_{n'}\| < 2\sqrt{d}\|\bar{v}_{n'}\|$. We use (see (7))

$$\bar{\lambda}_{n'} = (1 - \alpha')\beta_2 n' p = (1 - \alpha)\beta d \pm e^{-\Omega(d)}, \quad \bar{v}_{n'} = \underbrace{(1, 1, \dots, 1)}_{\alpha' n'} \underbrace{(\bar{\beta}_2, \bar{\beta}_2, \dots)}_{(1 - \alpha') n'},$$

$$\begin{aligned} \|(A' - \bar{\lambda}_{n'} I)\bar{v}_{n'}\| &= \left\| (A' - \bar{\lambda}_{n'} I) \begin{bmatrix} 1 \\ \cdot \\ \cdot \\ \bar{\beta}_2 \\ \cdot \\ \cdot \end{bmatrix} \right\| = \left\| \begin{bmatrix} \bar{\beta}_2 \deg(v)_{C'} - \bar{\lambda}_{n'} \\ \cdot \\ \cdot \\ \deg(v)_{I'} + \bar{\beta}_2 \deg(v)_{C'} - \bar{\lambda}_{n'} \bar{\beta}_2 \\ \cdot \\ \cdot \end{bmatrix} \right\| \\ &\stackrel{\beta_2 = \bar{\beta}_2(1 \pm e^{-\Omega(d)})}{\leq} \left\| \begin{bmatrix} \beta_2(\deg(v)_{C'} - (1 - \alpha)d) \\ \cdot \\ \cdot \\ \deg(v)_{I'} + \beta_2 \deg(v)_{C'} - (1 - \alpha)\beta_2^2 d \\ \cdot \\ \cdot \end{bmatrix} \right\| + \underbrace{\left\| \begin{bmatrix} e^{-\Omega(d)} \\ \cdot \\ \cdot \\ e^{-\Omega(d)} \\ \cdot \\ \cdot \end{bmatrix} \right\|}_{\leq e^{-\Omega(d)}\sqrt{n}}. \end{aligned}$$

It remains to upper bound the norm of the left vector, and we do that by upper bounding its squared norm with $2.3d\|\bar{v}_{n'}\|^2$. We will estimate the random variable:

$$(24) \quad \beta_2^2 \sum_{v \in I'} (\deg(v)_{C'} - (1 - \alpha)d)^2 + \sum_{v \in C'} (\deg(v)_{I'} + \beta_2 \deg(v)_{C'} - (1 - \alpha)\beta_2^2 d)^2.$$

Note that the difference between the sum in (24) and

$$(25) \quad \underbrace{\beta_2^2 \sum_{v \in I} (\deg(v)_C - (1 - \alpha)d)^2}_{S_1} + \underbrace{\sum_{v \in C} (\deg(v)_I + \beta_2 \deg(v)_C - (1 - \alpha)\beta_2^2 d)^2}_{S_2}$$

can be classified into two types:

- (i) vertices of $V \setminus V'$ that add summands to (25) which do not exist in (24);
- (ii) vertices of $V \cap V'$ for which the degrees $\deg(v)_C, \deg(v)_{C'}$ or $\deg(v)_I, \deg(v)_{I'}$ are different.

It is more convenient to upper bound (25) rather than (24). The differences of type (i) can only make (25) bigger than (24). The difference induced by type (ii) is $O(e^{-\Omega(d)}nd^3)$, and the explanation is as follows. For any fixed vertex v of degree $< 5d$ in G , its summand in S_1 (or in S_2) is bounded by $O(d^3)$ (because $\beta_2 < 5\sqrt{d}$). The number of vertices in V' which have an edge to a vertex in $V \setminus V'$ is bounded by $e^{-\Omega(d)}n$ (since w.h.p. the total number of edges that touch $|V \setminus V'|$ is at most $e^{-\Omega(d)}n$).

We will now prove the following inequalities:

$$(26) \quad S_1 \leq 1.1\alpha n(1-\alpha)d, \quad S_2 \leq 1.1(1-\alpha)n(\alpha d + \beta_2^2(1-\alpha)).$$

We start with S_2 . The graph G is taken from $G_{n,p,\alpha}$. We denote by $e(G)$ the number of edges in G . Let L be the set of all edges which do not have both end points in I . Denote that $l := |L| = \binom{n}{2} - \binom{\alpha n}{2}$. The expectation of $e(G)$ is $\mu := pl$ (where $\mu \geq 0.4dn$ since $\alpha \leq \frac{1}{3}$). By Chernoff's inequality it holds that

$$(27) \quad \Pr_{G_{n,p,\alpha}} [e(G) \in (1 \pm n^{-0.1})\mu] \geq 1 - o(1)$$

(we used here $\mu \geq 0.4dn$). Denote by $G_{n,m,\alpha}$ the uniform distribution over all graphs with m edges such that none of the edges is inside I . For a graph G taken from $G_{n,p,\alpha}$, given that $e(G) = m$, the graph G has the distribution of $G_{n,m,\alpha}$. Thus, it is sufficient to show that for any $m \in (1 \pm n^{-0.1})\mu$ it holds that

$$(28) \quad \Pr_{G_{n,m,\alpha}} [S_2 > 1.1(1-\alpha)n(\alpha d + \beta_2^2(1-\alpha))] = o(1).$$

Fix some $m \in (1 \pm n^{-0.1})\mu$. Technically, it is more convenient to prove a concentration result in a product measure. Denote by $C_{n,m,\alpha}$ the distribution induced by taking a uniformly random m -tuple from L^m . Let G be taken from $C_{n,m,\alpha}$. Note that G is actually a multigraph. Given that G is simple (i.e., no parallel edges), G is distributed as $G_{n,m,\alpha}$. Note that

$$(29) \quad \Pr_{C_{n,m,\alpha}} [G \text{ is simple}] \geq \left(1 - \frac{m}{l}\right)^m \stackrel{(i)}{\geq} e^{-\frac{2m^2}{l}} \geq e^{-(1+o(1))\frac{1}{2}p^2l} \\ \stackrel{l \leq 0.5n^2}{\geq} e^{-0.2d^2} \stackrel{d < n^{1/40}}{\geq} e^{-0.2n^{0.05}}$$

(in inequality (i) we used $1 - x \geq e^{-2x}$, which holds for $x \in [0, 0.5]$).

We will now prove that

$$(30) \quad \Pr_{C_{n,m,\alpha}} [S_2 > 1.1(1-\alpha)n(\alpha d + \beta_2^2(1-\alpha))] \leq 2e^{-n^{0.1}}.$$

We first calculate the expectation of S_2 . Fix a vertex $v \in C$. Denote that $X := \deg(v)_I + \beta_2 \deg(v)_C - (1-\alpha)\beta_2^2 d$.

$$\begin{aligned} \mathbb{E}[X] &= m \frac{\alpha n + \beta_2((1-\alpha)n-1)}{l} - (1-\alpha)\beta_2^2 d \\ &\stackrel{(9)}{=} (1 \pm n^{-0.1}) p \left[(1 \pm e^{-\Omega(d)})(\alpha' n' + \bar{\beta}_2(1-\alpha')n') \right] - (1 \pm e^{-\Omega(d)})(1-\alpha')d\bar{\beta}_2^2 \\ (31) \quad &\stackrel{(14)}{=} (\alpha' n' p + \bar{\beta}_2(1-\alpha')n' p - \underbrace{(1-\alpha')n'\bar{\beta}_2 p \bar{\beta}_2}_{\bar{\lambda}_{n'}}) \pm 4(n^{-0.1} + e^{-\Omega(d)})d \pm e^{-\Omega(d)}d \\ (32) \quad &\stackrel{(5)}{=} \pm 5d(n^{-0.1} + e^{-\Omega(d)}) \stackrel{d < n^{1/40}}{=} \pm(o(1) + e^{-\Omega(d)}). \end{aligned}$$

By using (10) a straightforward calculation (see Lemma 4.5 for details) shows that for any $v \in C$ it holds that

$$(33) \quad \text{VAR}[X] = (1 + o(1))(\alpha d + \beta_2^2(1 - \alpha)d).$$

Thus for large enough d

$$(34) \quad \mathbb{E}[X^2] = \text{VAR}[X] + (\mathbb{E}[X])^2 = (1 \pm 0.01)(\alpha d + \beta_2^2(1 - \alpha)d).$$

By summing over all C we derive

$$(35) \quad \mathbb{E}[S_2] = (1 \pm 0.01)(1 - \alpha)n(\alpha d + \beta_2^2(1 - \alpha)d).$$

We will now show that w.h.p. S_2 is tightly concentrated around its expectation. Recall that G is taken from $C_{n,m,\alpha}$ for some fixed $m \in (1 \pm n^{-0.1})\mu$. By Chernoff's inequality it holds that

$$(36) \quad \Pr[\text{all vertices degrees in } G \text{ are bounded by } dn^{0.1}] \geq 1 - e^{-n^{0.1}}$$

(the expected degree of a vertex in C is $\frac{n}{l}m = \frac{n}{l}(1 \pm n^{-0.1})pl = d(1 \pm n^{-0.1})$). Thus, with probability of at least $1 - e^{-n^{0.1}}$ every summand of S_2 is bounded by $3d^2n^{0.2}$ (we use here $|\beta_2| \leq 1$, (14)). Consider the function

$$(37) \quad f(v) := \min \{ (\deg(v)_I + \beta_2 \deg(v)_C - (1 - \alpha)\beta_2^2 d)^2, 3d^2n^{0.2} \}.$$

Let $f(G) := \sum_{v \in C} f(v)$. Adding/removing a single edge from G changes $f(G)$ by at most $6d^2n^{0.2}$ since all of the summands of $f(G)$ have values in $[0, 3d^2n^{0.2}]$. Since G is taken from $C_{n,m,\alpha}$, which is a product measure (L^m), we can use Azuma's inequality

$$(38) \quad \Pr[f(G) - \mathbb{E}[f(G)] > \lambda] < 2e^{-\lambda^2/2m(6d^2n^{0.2})^2}.$$

Since for any G it holds that $f(G) \leq S_2$ we derive

$$(39) \quad \Pr[f(G) - \mathbb{E}[S_2] > \lambda] < 2e^{-\lambda^2/2m(2d^4n^{0.2})^2},$$

and by setting $\lambda = 0.01\mathbb{E}[S_2]$ and using $m \in (1 \pm n^{-0.1})\mu$, $\mu = pl \leq dn$, and $\mathbb{E}[S_2] \geq \frac{n}{4}$ we derive

$$(40) \quad \Pr[f(G) \leq 1.01\mathbb{E}[S_2]] \geq 1 - 2e^{-\Omega(n^2/dnd^4n^{0.4})} \stackrel{d < n^{1/40}}{\geq} 1 - 2e^{-\Omega(n^{0.475})}.$$

Note that if all of the degrees in G are bounded by $dn^{0.1}$, then $f(G) = S_2$. Inequalities (36) and (40) imply that

$$(41) \quad \Pr_{C_{n,m,\alpha}} [S_2 > 1.01\mathbb{E}[S_2]] \leq 2e^{-n^{0.1}}.$$

By combining (29) with (41) we derive

$$(42) \quad \Pr_{C_{n,m,\alpha}} [S_2 \geq 1.01\mathbb{E}[S_2] \mid G \text{ is simple}] \leq \frac{2e^{-n^{0.1}}}{\Pr[G \text{ is simple}]} \leq \frac{2e^{-n^{0.1}}}{e^{-0.2n^{0.05}}} = o(1),$$

which together with (35) proves (28) and thus also the second part of (26). The proof of the first part of (26) is similar, so we omit the details. By having (26) proved we derive

$$(43) \quad \beta_2^2 S_1 + S_2 \stackrel{(26)}{\leq} \beta_2^2 1.1\alpha n(1-\alpha)d + 1.1(1-\alpha)n(\alpha d + \beta_2^2(1-\alpha))$$

$$(44) \quad \stackrel{(14)}{\leq} 1.1\alpha dn + 1.1(1-\alpha)n(\alpha d + \beta_2^2(1-\alpha)d)$$

$$(45) \quad \leq 2.2\alpha dn + 1.1(1-\alpha)n\beta_2^2 d \leq 2.2d(\alpha n + (1-\alpha)n\beta_2^2)$$

$$(46) \quad \leq 2.2d(1 + e^{-\Omega(d)})(\alpha' n' + (1-\alpha')n'\beta_2^2) \leq 2.3d\|\bar{v}_{n'}\|_2^2. \quad \square$$

Proof of Lemma 4.3, part (i). By (11), $\|\bar{v}_1\| \geq \sqrt{n}(1 - e^{-\Omega(d)})$. We will prove that $\|(A' - \bar{\lambda}_1 I)\bar{v}_1\| < 2\sqrt{d}\|\bar{v}_1\|$. We use (from (7))

$$\bar{\lambda}_1 = (1 - \alpha')\bar{\beta}_1 n' p = (1 - \alpha)\beta_1 d \pm e^{-\Omega(d)}, \quad \bar{v}_1 = (\underbrace{1, 1, \dots, 1}_{\alpha' n'}, \underbrace{\bar{\beta}_1, \bar{\beta}_1, \dots}_{(1-\alpha')n'}),$$

$$\begin{aligned} \|(A' - \bar{\lambda}_1 I)\bar{v}_1\| &= \left\| (A' - \bar{\lambda}_1 I) \begin{bmatrix} 1 \\ \vdots \\ \bar{\beta}_1 \\ \vdots \end{bmatrix} \right\| = \left\| \begin{bmatrix} \bar{\beta}_1 \deg(v)_{C'} - \bar{\lambda}_1 \\ \vdots \\ \deg(v)_{I'} + \bar{\beta}_1 \deg(v)_{C'} - \bar{\lambda}_1 \bar{\beta}_1 \\ \vdots \end{bmatrix} \right\| \\ &\stackrel{\beta_1 = \bar{\beta}_1(1 \pm e^{-\Omega(d)})}{\leq} \left\| \begin{bmatrix} \beta_1 (\deg(v)_{C'} - (1-\alpha)d) \\ \vdots \\ \deg(v)_{I'} + \beta_1 \deg(v)_{C'} - (1-\alpha)\beta_1^2 d \\ \vdots \end{bmatrix} \right\| + \underbrace{\left\| \begin{bmatrix} e^{-\Omega(d)} \\ \vdots \\ e^{-\Omega(d)} \\ \vdots \end{bmatrix} \right\|}_{\leq e^{-\Omega(d)}\sqrt{n}}. \end{aligned}$$

It remains to upper bound the norm of the left vector, and we do that by upper bounding its squared norm with $2.3d\|\bar{v}_1\|_2^2$. The squared norm of the left vector is

$$(47) \quad \beta_1^2 \sum_{v \in I'} (\deg(v)_{C'} - (1-\alpha)d)^2 + \sum_{v \in C'} (\deg(v)_{I'} + \beta_1 \deg(v)_{C'} - (1-\alpha)\beta_1^2 d)^2.$$

An argument similar to the one used in the proof of part (ii) shows that w.h.p. the last term is bounded by $2.3\|\bar{v}_1\|^2$. $\square$

Proof of Lemma 4.3, part (iii). If $x \perp \bar{v}_1$, $x \perp \bar{v}_{n'}$, then $\sum_{i \in I'} x_i = 0$, $\sum_{i \in C'} x_i = 0$ (by (8)). Thus, for such x the following holds:

$$x^t A' x = x^t \left(A' + \overbrace{\begin{bmatrix} p & \cdot & \cdot & p & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & & & & & & \\ \cdot & & \cdot & & & & & \\ p & \cdot & \cdot & p & & & & \\ 0 & & & & 0 & & & \\ \cdot & & & & & \cdot & & \\ 0 & & & & & & 0 & \\ 0 & & & & & & & 0 \end{bmatrix}}^{I'} \right) x = x^t B' x,$$

where B' is derived from A by removing vertices of degree $> 5d$ and adding the value $p = \frac{d}{n}$ in entries of the submatrix which corresponds to $I' \subset I$. This matrix B' is very similar to the matrices analyzed in [9]. The difference is that in [9] the whole matrix is random, whereas, in our case, a small (about $\alpha n \times \alpha n$) portion of the matrix is deterministically fixed to be p (the expectation). One would expect that having this fixed portion in the matrix would make the eigenvalue structure more similar to that of the all p matrix (a similarity that we wish to establish here). Indeed, a simple modification of the arguments in [9] (in sections 2.2.3, 2.2.4, 3.2, and 3.3) yields that w.h.p. for any $x \perp \bar{1}$ it holds that $x^t B' x \leq c\sqrt{d}$, where c is a universal constant independent of d and α (without loss of generality, we may further assume that $c \geq 2$). We omit the proof from the current manuscript and refer the reader to [9]. $\square$

LEMMA 4.4. *Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n'}$ be the eigenvalues of A' . If conditions (i), (ii), and (iii) from Lemma 4.3 hold, then for $i = 2, 3, \dots, n' - 1$ it holds that $|\lambda_i| \leq 2c\sqrt{d}$ ($c \geq 2$ is the universal constant from Lemma 4.3).*

Proof. It is enough to prove that $\lambda_2 \leq 2c\sqrt{d}$ and $\lambda_{n'-1} \geq -2c\sqrt{d}$. We first show $\lambda_{n'-1} \geq -2c\sqrt{d}$. It is well known that

$$\lambda_{n'-1} = \max_{\substack{H \text{ subspace of} \\ \text{dimension } n-1}} \min_{\substack{x \neq 0, \\ x \in H}} \frac{x^t A' x}{x^t x}.$$

Let us fix H to be the subspace perpendicular to $\bar{v}_{n'}$. Consider any vector $x \perp \bar{v}_{n'}$. The vector x can be written as $x = f + s$, where f is a multiple of $\bar{v}_1$ and $s \perp \bar{v}_1, \bar{v}_{n'}$.

$$\begin{aligned} x^t A' x &= (f + s)^t A' (f + s) = f^t A' f + s^t A' s + 2s^t A' f \\ &\geq \bar{\lambda}_1 \|f\|^2 - c\sqrt{d} \|s\|^2 + 2s^t (A' - \bar{\lambda}_1 I) f \\ &\geq -c\sqrt{d} \|s\|^2 - 2(1.2\sqrt{d} \|s\| \|f\|) \geq -2c\sqrt{d} \|x\|^2. \end{aligned}$$

In the first equality we used the symmetry of A' , in the first inequality we used Lemma 4.3(iii) for the second term and $s \perp f$ for the third term, in the following inequality we used $\bar{\lambda}_1 \geq 0$ and Lemma 4.3(i), and in the last inequality we used $c > 1.2$, $2\|s\| \|f\| \leq \|s\|^2 + \|f\|^2 = \|x\|^2$. A similar argument using

$$\lambda_2 = \min_{\substack{H \text{ subspace of} \\ \text{dimension } n-1}} \max_{\substack{x \in H \\ x \neq 0}} \frac{x^t A' x}{x^t x}$$

gives $\lambda_2 \leq 2c\sqrt{d}$. $\square$

LEMMA 4.5. *Fix $l = \binom{n}{2} - \binom{\alpha n}{2}$, and let G be taken from $C_{n,m,\alpha}$, where $m \in (1 \pm n^{-0.1})pl$. Fix any $v \in C$, and define $X := \deg(v)_I + \beta_2 \deg(v)_C - (1 - \alpha)\beta_2^2 d$. It holds that $\text{VAR}[X] = (1 + o(1))(\alpha d + \beta_2^2(1 - \alpha)d)$.*

Proof. The random variable X can be written as the sum of m independent random variables $X_1, \dots, X_m$, where $X_i := \deg(v)_I + \beta \deg(v)_C - (1 - \alpha)\beta^2 d$ when G is composed of only one random edge from L . It holds that $\text{VAR}[X] = m\text{VAR}[X_1]$. We will now calculate

$$(48) \quad \text{VAR}[X_1] = \text{VAR}[\deg(v)_I] + \beta_2^2 \text{VAR}[\deg(v)_C] + 2\beta_2 \text{COV}(\deg(v)_I, \deg(v)_C).$$

Since $\deg(v)_I$ and $\deg(v)_C$ are in fact indicators, we will denote them, respectively,

by z_1 and z_2 , where $\Pr[z_1] = p_1$ and $\Pr[z_2] = p_2$.

$$(49) \quad \text{VAR}[\deg(v)_I] = p_1(1 - p_1) = \frac{\alpha n}{l} \left(1 - \frac{\alpha n}{l}\right) \stackrel{(10)}{=} (1 \pm o(1)) \frac{\alpha n}{l},$$

$$(50) \quad \begin{aligned} \text{VAR}[\deg(v)_C] &= p_2(1 - p_2) \\ &= \frac{(1 - \alpha)n - 1}{l} \left(1 - \frac{(1 - \alpha)n - 1}{l}\right) \stackrel{(10)}{=} (1 \pm o(1)) \frac{(1 - \alpha)n}{l}, \end{aligned}$$

$$\beta_2 \text{COV}[\deg(v)_I, \deg(v)_C] = \beta_2(\Pr[z_1 \wedge z_2] - p_1 p_2) = \beta_2(0 - p_1 p_2)$$

$$(51) \quad = -\beta_2(1 \pm o(1)) \frac{\alpha(1 - \alpha)n^2}{l^2} \stackrel{(10)}{=} o(1).$$

It thus follows that

$$(52) \quad \text{VAR}[X] = (1 \pm o(1)) \frac{mn}{l} (\alpha + \beta_2^2(1 - \alpha)) = (1 \pm o(1)) d (\alpha + \beta_2^2(1 - \alpha)). \quad \square$$

REFERENCES

- [1] N. ALON AND N. KAHALE, *A spectral technique for coloring random 3-colorable graphs*, SIAM J. Comput., 26 (1997), pp. 1733–1748.
- [2] N. ALON, M. KRIVELEVICH, AND B. SUDAKOV, *Finding a large hidden clique in a random graph*, Random Structures Algorithms, 13 (1988), pp. 457–466.
- [3] H. CHEN AND A. FRIEZE, *Coloring bipartite hypergraphs*, in Proceedings of the 5th International Conference on Integer Programming and Combinatorial Optimization, Springer, New York, 1996, pp. 345–358.
- [4] A. COJA-OGHLAN, *Finding large independent sets in polynomial expected time*, Combin. Probab. Comput., 15 (2006), pp. 731–751.
- [5] A. COJA-OGHLAN, *A spectral heuristic for bisecting random graphs*, Random Structures Algorithms, 29 (2006), pp. 351–389.
- [6] U. FEIGE, *Approximating maximum clique by removing subgraphs*, SIAM J. Discrete Math., 18 (2004), pp. 219–225.
- [7] U. FEIGE AND J. KILIAN, *Heuristics for semirandom graph problems*, J. Comput. System Sci., 63 (2001), pp. 639–671.
- [8] U. FEIGE AND R. KRAUTHGAMER, *Finding and certifying a large hidden clique in a semirandom graph*, Random Structures Algorithms, 16 (2000), pp. 195–208.
- [9] U. FEIGE AND E. OFEK, *Spectral techniques applied to sparse random graphs*, Random Structures Algorithms, 27 (2005), pp. 251–275.
- [10] U. FEIGE AND E. OFEK, *Finding a maximum independent set in a sparse random graph*, in Proceedings of the 9th International Workshop on Randomization and Computation, RANDOM 2005, Lecture Notes in Comput. Sci. 3624, Springer, New York, 2005, pp. 282–293.
- [11] A. FLAXMAN, *A spectral technique for random satisfiable 3CNF formulas*, in Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 2003, pp. 357–363.
- [12] A. GOERDT AND A. LANKA, *On the hardness and easiness of random 4-sat formulas*, in Proceedings of the 15th International Symposium on Algorithms and Computation (ISAAC), Lecture Notes in Comput. Sci. 3341, Springer, New York, 2004, pp. 470–483.
- [13] G. GRIMMET AND C. MCDIARMID, *On colouring random graphs*, Math. Proc. Cambridge Philos. Soc., 77 (1975), pp. 313–324.
- [14] J. HÅSTAD, *Clique is hard to approximate within $n^{1-\epsilon}$* , Acta Math., 182 (1999), pp. 105–142.
- [15] M. JERRUM, *Large cliques elude the Metropolis process*, Random Structures Algorithms, 3 (1992), pp. 347–359.
- [16] R. M. KARP, *The probabilistic analysis of some combinatorial search algorithms*, in Algorithms and Complexity: New Directions and Recent Results, J. F. Traub, ed., Academic Press, New York, 1976, pp. 1–19.
- [17] R. M. KARP, *Reducibility among combinatorial problems*, in Complexity of Computer Computations, R. E. Miller and J. W. Thatcher, eds., Plenum Press, New York, 1972, pp. 85–104.
- [18] L. KUČERA, *Expected complexity of graph partitioning problems*, Discrete Appl. Math., 57 (1995), pp. 193–212.

A CENSUS OF SMALL LATIN HYPERCUBES*

BRENDAN D. MCKAY[†] AND IAN M. WANLESS[‡]

Abstract. We count all latin cubes of order $n \leq 6$ and latin hypercubes of order $n \leq 5$ and dimension $d \leq 5$. We classify these (hyper)cubes into isotopy classes and paratopy classes (main classes). For the same values of n and d we classify all d -ary quasigroups of order n into isomorphism classes and also count them according to the number of identity elements they possess (meaning we have counted the d -ary loops). We also give an exact formula for the number of (isomorphism classes of) d -ary quasigroups of order 3 for every d . Then we give a number of constructions for d -ary quasigroups with a specific number of identity elements. In the process, we prove that no 3-ary loop of order n can have exactly $n - 1$ identity elements (but no such result holds in dimensions other than 3). Finally, we give some new examples of latin cuboids which cannot be extended to latin cubes.

Key words. Latin cube, Latin hypercube, Latin cuboid n -ary quasigroup, n -ary loop, identity elements, Jacobsthal sequence, total loop, MDS code

AMS subject classifications. 05B15, 20N05, 20N15, 94B25

DOI. 10.1137/070693874

1. Basic definitions. Let $[n]$ denote the set $\{1, 2, \dots, n\}$ and let $[n]^d$ denote the cartesian product $[n] \times [n] \times \dots \times [n]$ of d copies of $[n]$. By a *hypercube of order n and dimension d* we mean a d -dimensional array of n^d cells where the cells are indexed by $[n]^d$ and each cell contains an element of $[n]$ (which we will call a *symbol*). Suppose that H is such a hypercube and c is any cell of H . A *line through c* is the set of n cells of H whose coordinates match those of c except possibly in the k th coordinate (there is one line for each choice of k). A *hyperplane through c* is the set of n^{d-1} cells of H whose k th coordinate matches that of c (there is one hyperplane for each choice of k). Any hyperplane in a d -dimensional hypercube can be considered to be a $(d-1)$ -dimensional hypercube, simply by dropping the common coordinate. We use vector notation such as $\vec{v}$ for an element of $[n]^d$. In a hypercube H we denote the symbol in the cell with coordinates $\vec{v} = (v_1, v_2, \dots, v_d)$ by $H(\vec{v})$ or $H(v_1, v_2, \dots, v_d)$.

We say that a hypercube H is *latin* if the n cells in any line of H contain each of the symbols in $[n]$ exactly once. We define $\mathcal{H}_n^d$ to be the set of latin hypercubes of order n and dimension d . Given a hypercube $L \in \mathcal{H}_n^d$ and a cell $\vec{v} = (v_1, v_2, \dots, v_d) \in [n]^d$, we can define d permutations of $[n]$ corresponding to the lines through $\vec{v}$. Namely, for each $j \in [d]$ we define $\rho_j(\vec{v}, L)$ to be the permutation

$$k \mapsto L(i_1, \dots, i_{j-1}, k, i_{j+1}, \dots, i_d) \quad \text{for } k \in [n].$$

If $\rho_j((1, 1, \dots, 1), L)$ is the identity permutation for each $j \in [d]$, then we say that L is *reduced*. We define $\mathcal{R}_n^d \subseteq \mathcal{H}_n^d$ to be the set of reduced latin hypercubes of order n and dimension d .

For some purposes it is easier to think of each hypercube $H \in \mathcal{H}_n^d$ as, instead, a set $T_H \subseteq [n]^{d+1}$ where for each $\vec{v} = (v_1, v_2, \dots, v_d) \in [n]^d$ there is a $(d+1)$ -tuple $\langle v_1, v_2, \dots, v_d, H(\vec{v}) \rangle$ in T_H . (We adopt the convention of writing the tuples of a

*Received by the editors June 6, 2007; accepted for publication (in revised form) September 12, 2007; published electronically April 18, 2008.

<http://www.siam.org/journals/sidma/22-2/69387.html>

[†]Department of Computer Science, Australian National University, Canberra, ACT 0200, Australia (bdm@cs.anu.edu.au).

[‡]School of Mathematical Sciences, Monash University, Clayton, Vic 3800, Australia (ian.wanless@sci.monash.edu.au).

hypercube inside angle brackets $\langle \cdot \rangle$.) The latin property of the hypercube is equivalent to saying that no two of its tuples differ in exactly one coordinate.

Some of the interest in latin hypercubes stems from coding theory, given that a hypercube $L \in \mathcal{H}_n^d$ is equivalent to a maximum distance separable (MDS) code over an alphabet of size n , with length $d + 1$ and minimum distance 2. See Laywine and Mullen [16] for a good introduction to this area.

Another source of interest in latin hypercubes is in nonassociative algebra. A d -ary quasigroup of order n is a function $Q : [n]^d \rightarrow [n]$ such that $Q(\vec{u}) \neq Q(\vec{v})$ whenever $\vec{u}$ and $\vec{v}$ differ in exactly one coordinate. Such a quasigroup is clearly equivalent to a latin hypercube $L \in \mathcal{H}_n^d$, where $L(\vec{v}) = Q(\vec{v})$. Moreover L is reduced if and only if 1 is an identity element in the corresponding quasigroup Q . (An *identity element* of Q is an $x \in Q$ such that

$$Q(y, x, x, \dots, x) = Q(x, y, x, \dots, x) = \dots = Q(x, x, \dots, x, y) = y$$

for all $y \in Q$.) A quasigroup possessing an identity element is called a *loop*. One of the aims of this paper is to count d -ary quasigroups and d -ary loops of order n (up to isomorphism) for small d and n . We do this using the theory in section 2 and the algorithm in section 3. For $d > 2$ it transpires that identity elements in d -ary loops need not be unique. Hence, in section 5 we examine constructions for loops with various numbers of identity elements.

A third source of applications for latin hypercubes is in the design of statistical experiments. In this field the phrases “latin cube” and “latin hypercube” are used for a broader class of objects than we are allowing in our definition. See [6] or [24] in this regard. The statistical definition of latin hypercube seems unnatural in other contexts since it does not have the property that each hyperplane is necessarily a latin hypercube of one lower dimension. The confusion arising from these different definitions is unfortunate, but we choose to follow, for example, [4] and [15] in our use of “latin hypercube” and many papers (e.g., [8, 9, 14, 17, 22]) in our use of “latin cube”. The alternative terminology “permutation (hyper)cube” was advocated in [6] and used in [1, 3, 7, 11]. This terminology is better reserved for the higher dimensional analogues of permutation matrices (that is, multidimensional arrays with all entries 0 or 1 such that each line contains a unique 1). Interestingly, $(d+1)$ -dimensional permutation hypercubes of order n are combinatorial objects essentially equivalent to d -dimensional latin hypercubes of order n (both objects described in our preferred terminology). This simple observation was made by Gupta [9], who used the phrase “permutation cube” in our preferred sense.

In the cases $d = 1, 2, 3$ a latin hypercube of dimension d is a permutation, a latin square and a latin cube, respectively. Much of the theory of latin hypercubes is yet to be developed although some analogues of results for latin squares have been investigated for latin cubes and occasionally for hypercubes in general. Completion and embedding results have been obtained by Cruse [4], Kochol [14], and Lindner [17]. Fu [8] considered the range of values possible for the number of common entries shared by distinct latin cubes. (In other terminology, this equates to studying the possible sizes of trades in latin cubes.) A number of authors have considered sets of k mutually orthogonal latin hypercubes. Here again several different definitions are possible; see [6] and [16] for further information and references. The notions of intercalates, transversals, and prolongation were generalized by Heinrich [11] from latin squares to higher dimensions. The same generalization of transversals was used in [15] but a different generalization of transversals was given by Beljavskaja and Murathudjaev

[3]. Finally, strongly diagonal latin squares (also called Knut Vik designs) and totally symmetric latin squares were generalized to higher dimensions by Alavi, Lick, and Liu [1] and Bailey, Preece, and Zemroch [2], respectively.

The usual notions of isotopism, paratopism, and isomorphism generalize naturally from latin squares to higher dimension. Let S_n be the symmetric group on $[n]$ and let S_n^c denote the direct product of c copies of S_n . Then the natural action of S_n^{d+1} on $[n]^{d+1}$ induces an action on $\mathcal{H}_n^d$ (where, as discussed above, we associate each $H \in \mathcal{H}_n^d$ with a subset $T_H \subseteq [n]^{d+1}$). This action is called *isotopism* (or *isotopy*) and its orbits are called *isotopy classes*. Define Δ_n^{d+1} to be the diagonal subgroup of S_n^{d+1} , that is, $\Delta_n^{d+1} = \{(g, g, \dots, g) \in S_n^{d+1}\}$. An important special case of isotopism is the action of Δ_n^{d+1} on $\mathcal{H}_n^d$. This action is called *isomorphism* and its orbits are called *isomorphism classes*. If the hypercube is regarded as the table of values of a d -ary quasigroup on $[n]$, then isomorphisms of the hypercube correspond to standard isomorphisms of the quasigroup.

A further group action on a hypercube is provided by permutation of the elements of tuples. In this action, a permutation $\tau \in S_{d+1}$ maps the tuple $\langle v_1, v_2, \dots, v_{d+1} \rangle$ onto the tuple $\langle v_1, v_2, \dots, v_{d+1} \rangle^\tau = \langle w_1, w_2, \dots, w_{d+1} \rangle$, where $w_{i\tau} = v_i$ for $1 \leq i \leq d+1$. Here, and elsewhere, we use the superscript notation for the image of an object under a function; thus i^τ means $\tau(i)$, and L^τ is the image of L obtained by applying τ to each of its tuples. Such images are the *conjugates* (also called *parastrophes*) of L .

An arbitrary combination of a conjugacy and an isotopism is called a *paratopism* (or *paratopy*). The set of all paratopisms corresponds to the wreath product $S_n \wr S_{d+1}$ in its natural action on $[n]^{d+1}$. The orbits of its action on the set of all hypercubes are called *paratopy classes*, *main classes* or *species*.

The stabilizers of a latin hypercube L under isotopism, paratopism, and isomorphism are known respectively as the autotopism group, autoparatopism group, and automorphism group of L . We use, respectively, $\text{Is}(L)$, $\text{Par}(L)$, and $\text{Aut}(L)$ to denote these groups. For example, $\text{Aut}(L) = \{\sigma \in \Delta_n^{d+1} \mid L^\sigma = L\}$.

One of the main objectives of this paper is enumeration of latin hypercubes and associated structures (such as d -ary quasigroups). The enumeration of latin squares has a lengthy and well-known history for which we refer to the recent survey given in [18]. Considerably less work has been done for higher dimensions.

To count latin hypercubes it suffices to count reduced latin hypercubes and then apply the formula

$$(1) \quad |\mathcal{H}_n^d| = n! (n-1)!^{d-1} |\mathcal{R}_n^d|.$$

One of the first works was by Gupta [9], who enumerated latin cubes by hand and obtained the incorrect value $|\mathcal{R}_4^3| = 58$. The correct value $|\mathcal{R}_4^3| = 64$ was subsequently obtained by Mullen and Weber [22] and much later by Jia and Qin [13], with both pairs of researchers attempting to count and classify latin cubes of orders 1 up to 5.

Mullen and Weber [22] reported the numbers of reduced latin cubes of orders 1 up to 5 to be 1, 1, 1, 64, 40246 and the numbers of isomorphism classes to be 1, 1, 1, 19, 1860. They did not establish the number of isotopism classes for order 5 but found the number for orders 1 to 4, respectively, to be 1, 1, 1, 12. Zinoviev and Zinoviev [28] found that there are five paratopy classes of latin cubes of order 4. Our computations confirm the above numbers.

Apparently unaware of Mullen and Weber's work two decades earlier, Jia and Qin [13] attempted similar computations. They reported the same numbers of reduced latin cubes but gave the numbers of isotopism classes to be 1, 1, 1, 15, 479. These last two values are incorrect.

A patent application by Ito [12] included values of $|\mathcal{R}_4^d|$ for $d \leq 5$ that agree with our computations (see Table 1) and also with the values found by Potapov and Krotov [23]. Potapov and Krotov also proved that

$$3^{d+1}2^{2^d+1} \leq |\mathcal{H}_4^d| \leq (3^{d+1} + 1)2^{2^d+1}$$

for $d \geq 5$.

2. Counting equivalence classes. The numbers of isotopy and paratopy classes of latin hypercubes are related to the total number of latin hypercubes via the sizes of the autotopism and autoparatopism groups. The exact relationship is specified in the following lemma, whose proof is elementary.

LEMMA 1. *For any $L \in \mathcal{H}_n^d$, the paratopy class of L contains*

- (a) $\frac{(d+1)!}{|\text{Par}(L)|/|\text{Is}(L)|}$ isotopy classes,
- (b) $\frac{(d+1)!n!n^{d-1}}{|\text{Par}(L)|}$ reduced latin hypercubes, and
- (c) $\frac{(d+1)!(n!)^{d+1}}{|\text{Par}(L)|}$ latin hypercubes.

In the case of isomorphism classes, the natural setting is that of quasigroups and loops. The procedure for counting isomorphism classes of ordinary quasigroups given in [18, Theorem 4] applies equally well to the d -ary case, and the same proof applies with obvious adaptations, so we state our first theorem without proof.

Let $\mathcal{I}_n^d$ be a set containing one hypercube from each isotopy class of $\mathcal{H}_n^d$, and let $\mathcal{M}_n^d$ be a set containing one hypercube from each paratopy class of $\mathcal{H}_n^d$.

Define the *cycle structure* of a permutation γ to be the sequence $(n_1, n_2, \dots)$, where n_i is the number of cycles of length i in γ . If $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_{d+1})$ is a $(d+1)$ -tuple of permutations, define $\psi(\sigma)$ as follows:

- (i) If $\sigma_1, \sigma_2, \dots, \sigma_{d+1}$ have the same cycle structure, $(n_1, n_2, \dots)$, then $\psi(\sigma) = \prod_i n_i! i^{n_i}$.
- (ii) Otherwise, $\psi(\sigma) = 0$.

THEOREM 2. *The number of isomorphism classes of d -ary quasigroups of order n is*

$$\sum_{L \in \mathcal{I}_n^d} \frac{1}{|\text{Is}(L)|} \sum_{\sigma \in \text{Is}(L)} \psi(\sigma)^d = \sum_{L \in \mathcal{M}_n^d} \frac{(d+1)!}{|\text{Par}(L)|} \sum_{\sigma \in \text{Is}(L)} \psi(\sigma)^d.$$

As pointed out by a referee, some of these results are true in a more general setting. For some set X , let $\mathcal{F}_n^d(X)$ be the class of all functions from $[n]^{d+1}$ to X . A latin hypercube is such a function for $X = \{0, 1\}$, namely, the characteristic function of the set of $(d+1)$ -tuples it comprises. Isotopy, paratopy, and isomorphism are defined for $\mathcal{F}_n^d(X)$ by permutations of the domain $[n]^{d+1}$ in the same way as for latin hypercubes. We then have, with essentially the same proofs, that Lemma 1(a) and (c) and Theorem 2 are true for any subclass of $\mathcal{F}_n^d(X)$ which is closed under paratopy.

While counting d -ary quasigroups required no new theory, it turns out that counting d -ary loops is less straightforward. The “obvious” extension of [18, Theorem 5] to the d -ary case does not work, primarily due to the possibility that loops have more than one identity element (see section 5 for examples). Therefore, we need to develop a more complicated approach.

We say that two cells $\vec{u}$ and $\vec{v}$ of $L \in \mathcal{H}_n^d$ are *line-equivalent* if $\rho_j(\vec{u}, L) = \rho_j(\vec{v}, L)$ for all $j \in [d]$. If x, y are both identity elements of a d -ary loop, then $(x, x, \dots, x)$ is line equivalent to $(y, y, \dots, y)$.

For $1 \leq i \leq d+1$, define $L^{[i]}$ to be the conjugate of L obtained by exchanging the i th and $(d+1)$ th elements of tuples. This includes the case $L^{[d+1]} = L$. Also, for $h, s \in S_n^{d+1}$ we write $h^s = s^{-1}hs$.

THEOREM 3. *For $L \in \mathcal{H}_n^d$, choose $C(L)$ to be a set of cells that includes exactly one cell from each line-equivalence class of cells of L . For each cell $\vec{v}$, define*

$$N(\vec{v}, L) = |\{h \in \text{Is}(L) \mid h^{(\rho_1, \rho_2, \dots, \rho_d, 1)} \in \Delta_n^{d+1}\}|,$$

where $\rho_j = \rho_j(\vec{v}, L)$ for each j , and 1 is the identity of S_n .

Then the number of isomorphism classes of d -ary loops of order n is

$$(2) \quad \sum_{L \in \mathcal{I}_n^d} \frac{1}{|\text{Is}(L)|} \sum_{\vec{v} \in C(L)} N(\vec{v}; L) = \sum_{L \in \mathcal{M}_n^d} \frac{d!}{|\text{Par}(L)|} \sum_{t=1}^{d+1} \sum_{\vec{v} \in C(L^{[t]})} N(\vec{v}; L^{[t]}).$$

Proof. Let $\mathcal{L}_n^d$ consist of all the loops in $\mathcal{H}_n^d$. Note that we do not require any particular elements to be the identity elements. The number we seek is the number of equivalence classes of $\mathcal{L}_n^d$ under Δ_n^{d+1} . By the Burnside–Frobenius lemma, this is equal to the average number of loops fixed by an element of Δ_n^{d+1} . Our task is therefore to find the number of pairs (M, g) for $M \in \mathcal{L}_n^d$ and $g \in \Delta_n^{d+1}$, such that $M^g = M$.

We start by seeking the number of such pairs where M is in the isotopy class of some fixed $L \in \mathcal{H}_n^d$. This isotopy class consists of all L^σ for $\sigma \in S_n^{d+1}$, with the caveat that each hypercube appears for exactly $|\text{Is}(L)|$ values of σ . Leaving that factor for later, we need to find the number of pairs (σ, g) for $\sigma \in S_n^{d+1}$ and $g \in \Delta_n^{d+1}$, such that L^σ is a loop and $L^{\sigma g} = L^\sigma$. The last condition is equivalent to $\sigma g \sigma^{-1} \in \text{Is}(L)$.

Writing h for $\sigma g \sigma^{-1}$, our problem reduces to counting all pairs (σ, h) for $\sigma \in S_n^{d+1}$ and $h \in \text{Is}(L)$ such that

- (a) L^σ is a loop and
- (b) $h^\sigma \in \Delta_n^{d+1}$.

We proceed by parameterizing those σ satisfying (a). This is analogous to reducing a latin square, except that we do not require an identity element in the first position, and, indeed, we must remember that there can be more than one identity element.

- Choose a cell $\vec{v}$ and a permutation $\delta \in S_n$.
- Permute rows, columns, etc., so that the lines through cell $\vec{v}$ are all the identity permutation.
- Apply $\delta \in S_n$ to each of the $d+1$ components.

The resulting loop has $L(\vec{v})^\delta$ as an identity element. Symbolically, we have

$$(3) \quad \sigma = \sigma(\vec{v}, \delta) = (\rho_1(\vec{v}, L)\delta, \rho_2(\vec{v}, L)\delta, \dots, \rho_d(\vec{v}, L)\delta, \delta).$$

It is easy to see that every σ satisfying (a) has this form. However, this parameterisation has redundancies: $\sigma(\vec{u}, \delta') = \sigma(\vec{v}, \delta)$ if and only if $\vec{u}$ is line-equivalent to $\vec{v}$ and $\delta' = \delta$.

Finally, for each $\sigma = \sigma(\vec{v}, \delta)$, we need to know how many $h \in \text{Is}(L)$ satisfy (b). This does not depend on δ since $h^{\sigma(\vec{v}, \delta)} \in \Delta_n^{d+1} \Leftrightarrow h^{\sigma(\vec{v}, 1)} \in \Delta_n^{d+1}$.

TABLE 1

Number of reduced latin hypercubes, isotopy classes of latin hypercubes, paratopy classes of Latin hypercubes, and isomorphism classes of quasigroups for small orders n and dimensions d .

d	n	reduced	isotopy classes	paratopy classes	quasigroups
2	2	1	1	1	1
2	3	1	1	1	5
2	4	4	2	2	35
2	5	56	2	2	1411
2	6	9408	22	12	1130531
2	7	16942080	564	147	12198455835
2	8	535281401856	1676267	283657	2697818331680661
3	2	1	1	1	1
3	3	1	1	1	11
3	4	64	12	5	2589
3	5	40246	59	15	23192922
3	6	95909896152	5678334	264248	1381105636226980
4	2	1	1	1	1
4	3	1	1	1	21
4	4	7132	328	26	1565243
4	5	31503556	5466	86	435509352937
5	2	1	1	1	1
5	3	1	1	1	43
5	4	201538000	2133586	4785	263347981121
5	5	50490811256	1501786	3102	16751644838639300
6	2	1	1	1	1
6	3	1	1	1	85

We can now derive the first expression in (2). For each $L \in \mathcal{I}_n^d$, the number of pairs (σ, h) satisfying (a) and (b) can be found by testing, for one $\vec{v}$ from each line-equivalence class of cell, whether $\sigma(\vec{v}, 1)$ satisfies (b). For each of those that pass the test, we have $n!$ choices of δ . Then we divide by $|\text{Is}(L)|$ to account for the number of times each hypercube appears as L^σ and by $|\Delta_n^{d+1}| = n!$ as required by the Burnside–Frobenius lemma.

The summand of the outside summation in the first half of (2) is not invariant under conjugacy, so converting to a sum over paratopy classes is not just a matter of applying some factors. However, the summand is invariant under conjugacies that leave the $(d+1)$ th position fixed. Therefore, we get all the information we need from the conjugates $L^{[1]}, L^{[2]}, \dots, L^{[d+1]}$. Each of these represents $d! |\text{Is}(L)| / |\text{Par}(L)|$ isotopy classes, so the second half of (2) is obtained. $\square$

In Table 1 we give the results of our enumeration of latin hypercubes, classified according to various notions of equivalence. A more detailed count of hypercubes, classified according to group sizes, is presented in the appendix. A representative of each of the paratopy classes covered by Table 1 can be found in [21]. In Table 2 we give the results of our enumeration of loops (up to isomorphism) according to the number of identity elements they have. In section 5 we consider further the question of how many identity elements a loop can have. To save space, we have omitted a number of known values for $d = 2$ and $n \in \{9, 10, 11\}$; see [18, 20].

3. Construction method. To reduce the probability of error, all the computations were carried out independently by the two authors using slightly different algorithms. We now describe the first algorithm used.

Let $L \in \mathcal{H}_n^d$. For $1 \leq k \leq n$, the k th hyperplane of L is $L^{(k)} \in \mathcal{H}_n^{d-1}$ defined by $L^{(k)}(i_1, \dots, i_{d-1}) = L(i_1, \dots, i_{d-1}, k)$ for all $i_1, \dots, i_{d-1}$.

TABLE 2
Counts of loops according to their number of identity elements.

d	n	Number of identity elements						Total
		1	2	3	4	5	6	
2	2	1						1
2	3	1						1
2	4	2						2
2	5	6						6
2	6	109						109
2	7	23746						23746
2	8	106228849						106228849
2	9	9365022303540						9365022303540
3	2	0	1					1
3	3	1	0	0				1
3	4	3	15	0	1			19
3	5	1826	32	1	0	1		1860
3	6	797971315	1422290	525	90	0	6	799394226
4	2	1	0					1
4	3	0	0	1				1
4	4	2213	0	0	0			2213
4	5	1349704	9	0	0	0		1349713
5	2	0	1					1
5	3	1	0	0				1
5	4	34140	66060918	0	524800			66619858
5	5	2122915806	805040	280	0	70		2123721196
6	2	1	0					1
6	3	1	0	0				1

Next we define an ordering on latin hypercubes. The hypercube L can be specified by listing the symbols in each of the n^d cells in a particular order:

$$\begin{aligned} \Sigma(L) = [& L(1, 1, 1, \dots, 1), L(2, 1, 1, \dots, 1), \dots, L(n, 1, 1, \dots, 1), \\ & L(1, 2, 1, \dots, 1), L(2, 2, 1, \dots, 1), \dots, L(n, 2, 1, \dots, 1), \\ & L(1, 3, 1, \dots, 1), \dots, \\ & L(1, n, 1, \dots, 1), L(2, n, 1, \dots, 1), \dots, L(n, n, 1, \dots, 1), \\ & L(1, 1, 2, \dots, 1), \dots, \\ & L(1, n, n, \dots, n), L(2, n, n, \dots, n), \dots, L(n, n, n, \dots, n)]. \end{aligned}$$

The order of the cells (earlier indices varying faster) is important for what follows. We call L *isotopy-minimal* if $\Sigma(L) \leq \Sigma(L')$ for every L' in the isotopy class of L . Similarly, L is *paratopy-minimal* if $\Sigma(L) \leq \Sigma(L')$ for every L' in the paratopy class of L . In each case, the comparison $\leq$ is lexicographic order. The following properties follow easily from the definitions.

LEMMA 4. *Let L be a latin hypercube. Then the following hold.*

- (i) *There is exactly one isotopy-minimal hypercube in each isotopy class and exactly one paratopy-minimal hypercube in each paratopy class.*
- (ii) *If L is paratopy-minimal, then L is isotopy-minimal.*
- (iii) *If L is isotopy-minimal, then L is reduced.*
- (iv) *If L is paratopy-minimal in $\mathcal{H}_n^d$, then $L^{(1)}$ is paratopy-minimal in $\mathcal{H}_n^{d-1}$.*

The aim of the computation was to find all paratopy-minimal hypercubes $L \in \mathcal{H}_n^d$. As permitted by Lemma 4(iv), we took the paratopy-minimal hypercubes in $\mathcal{H}_n^{d-1}$ as $L^{(1)}$. Then we extended $L^{(1)}$ to L in all possible ways such that L was reduced. Finally, we rejected L if it was not paratopy-minimal.

To check that L is paratopy-minimal, we verified that $\Sigma(L) \leq \Sigma(L^\gamma)$ for each of the $n^{d-1}(d+1)!n!$ paratopisms γ that take L to a reduced hypercube. These paratopisms have the form $\gamma = \gamma(\vec{v}, \delta, L) = \sigma(\vec{v}, \delta)\tau$, where τ is one of the $(d+1)!$ conjugacies, and $\sigma(\vec{v}, \delta)$ is defined as in (3) with $\delta \in S_n$ such that $L(\vec{v})^\delta = 1$. In this case, the redundancy issue noted just after (3) does not occur: all these $n^{d-1}(d+1)!n!$ paratopisms are different. To see this, note that the condition $L(\vec{v})^\delta = 1$ implies that $\sigma(\vec{v}, \delta)$ maps position $\vec{v}$ onto position $(1, 1, \dots, 1)$. If $\gamma = \gamma(\vec{v}, \delta, L)$ was found such that $\Sigma(L^\gamma) < \Sigma(L)$, thereby rejecting L as not paratopy-minimal, the next candidate hypercube L' was first tested with $\gamma(\vec{v}, \delta, L')$. This often rejected L' as well, giving a useful speedup.

If the minimality test for L was passed, we had also found the autopatopism group: $\text{Par}(L)$ consists of those γ for which $\Sigma(L^\gamma) = \Sigma(L)$.

The second algorithm used to count $|\mathcal{H}_n^d|$ was similar to the first in that it generated a catalogue of paratopy class representatives by extending all paratopy class representatives from $\mathcal{H}_n^{d-1}$ in all ways that produced a reduced hypercube. However, the isomorphism testing was different. The reduced hypercubes were canonically labelled using nauty [19] in much the same way as latin squares were treated in [18]. Any hypercube was rejected if its canonical labelling was identical to that of a previously constructed hypercube. The autopatopism group of each hypercube was calculated by nauty while finding the canonical labelling.

It is clear that both these generation methods are quite crude, but the use of more sophisticated techniques, such as applying a minimality test after each hyperplane is added, would not help very much with these small values of n and d . In particular, we do not think they would improve the efficiency enough to make hypercubes with higher order and/or dimension enumerable with present computing power.

4. Number of quasigroups of order 3. In this section we count the d -ary quasigroups of order $n = 3$, for arbitrary d . Before developing that result we need to introduce a notion of linearity.

We say that $H \in \mathcal{H}_n^d$ is *linear* if there exist $s \in [n]$ and $c_i \in [n]$ for $i \in [d]$, such that

$$(4) \quad H(x_1, x_2, \dots, x_d) \equiv s + \sum_{i=1}^d c_i x_i \pmod{n}$$

for each $(x_1, x_2, \dots, x_d) \in [n]^d$. We say that a quasigroup/loop is *linear* if its corresponding hypercube is linear. For (4) to define a latin hypercube it is necessary and sufficient that each c_i should be relatively prime to n . Let ϕ denote Euler's phi function. It follows that there are exactly $n(\phi(n))^d$ linear hypercubes in $\mathcal{H}_n^d$, since we have n choices for s and $\phi(n)$ choices for each c_i .

For $n \leq 3$ and arbitrary d it is easy to prove by induction on d that every hypercube in $\mathcal{H}_n^d$ is linear. This conclusion also follows from the values of $|\mathcal{H}_n^d|$ as given, for example, by Finizio and Lewis [7] or Laywine and Mullen [16, p. 224].

In our next theorem we will use the fact that each of the $2^d 3$ hypercubes in $\mathcal{H}_3^d$ is linear to count the d -ary quasigroups of order 3. Our result is phrased in terms of the Jacobsthal sequence, which is defined by $a_n = \frac{1}{3}(2^n + (-1)^{n+1})$, or alternatively by the recurrence

$$(5) \quad a_n = a_{n-1} + 2a_{n-2} \quad \text{with} \quad a_1 = a_2 = 1.$$

This is sequence A001045 in Sloane's On-Line Encyclopedia of Integer Sequences [25],

which lists many combinatorial objects counted by the sequence. Here we extend that list.

THEOREM 5. *Let $\mathcal{Q}_{3,g}^d$ be the set of d -ary quasigroups of order 3 with an automorphism group of order g . Then $|\mathcal{Q}_{3,2}^d| = a_{d+1}$, $|\mathcal{Q}_{3,3}^d| = |\mathcal{Q}_{3,6}^d| = a_d$, and $|\mathcal{Q}_{3,g}^d| = 0$ for $g \notin \{2, 3, 6\}$. Hence $|\mathcal{Q}_3^d| = a_{d+2}$.*

Proof. First note that the result for $|\mathcal{Q}_3^d|$ follows from the other results and (5), since $|\mathcal{Q}_3^d| = \sum_g |\mathcal{Q}_{3,g}^d|$.

Let $\mathcal{H}_{3,g}^d$ be the set of d -dimensional hypercubes of order n for which the corresponding quasigroup is in $\mathcal{Q}_{3,g}^d$. Clearly, $|\mathcal{Q}_{3,g}^d| = |\mathcal{H}_{3,g}^d| g/3!$.

Consider the diagonal cells of an arbitrary $H \in \mathcal{H}_3^d$. Specifically, for $i \in \{1, 2, 3\}$ define u_i by $\langle i, i, i, \dots, i, u_i \rangle \in H$. Since H must be linear, we find that $u_2 - u_1 \equiv u_3 - u_2 \equiv u_1 - u_3 \pmod{3}$.

We now argue that $u_k = k$ if and only if H possesses an automorphism τ fixing k and swapping the other two elements of $\{1, 2, 3\}$.

Certainly if H has such an automorphism, then it must map $\langle k, k, \dots, k, u_k \rangle$ to $\langle k, k, \dots, k, \tau(u_k) \rangle$, from which we infer that $u_k = k$, the only fixed point of τ .

So assume that $u_k = k$ and define $\tau : [3] \rightarrow [3]$ by $\tau(x) \equiv -k - x \pmod{3}$. Note that τ fixes k and swaps the elements of $[3] \setminus \{k\}$. Since H is linear we can assume the existence of s, c_i satisfying (4). Given that $u_k = k$, this implies that $k = s + k \sum c_i$. Now if we apply the isomorphism τ to a general element $\langle x_1, x_2, \dots, x_d, s + \sum c_i x_i \rangle$ of H , it maps that element to

$$\left\langle -k - x_1, -k - x_2, \dots, -k - x_d, -k - s - \sum c_i x_i \right\rangle,$$

where all coordinates should be calculated mod 3. However, this is an element of H since

$$s + \sum c_i (-k - x_i) = s - k \sum c_i - \sum c_i x_i = s + (s - k) - \sum c_i x_i \equiv -k - s - \sum c_i x_i$$

mod 3. We conclude that τ is indeed an automorphism, as we contended.

At this point we separate into three cases. Define $\alpha \in [3]$ by $\alpha \equiv u_2 - u_1 \equiv u_3 - u_2 \equiv u_1 - u_3 \pmod{3}$.

Case 1. $\alpha \neq 1$.

By the definition of α there exists a unique $k \in \{1, 2, 3\}$ such that $u_k = k$. This means $\text{Aut}(H)$ contains a unique transposition and hence $|\text{Aut}(H)| = 2$.

Case 2. $\alpha = 1$ and there exists $k \in \{1, 2, 3\}$ such that $u_k = k$.

Since $\alpha = 1$ we know $u_k = k$ for all $k \in \{1, 2, 3\}$. This means $\text{Aut}(H)$ contains three transpositions and hence $|\text{Aut}(H)| = 6$.

Case 3. $\alpha = 1$ and there does not exist $k \in \{1, 2, 3\}$ such that $u_k = k$.

In this case $\text{Aut}(H)$ contains no transpositions. However, by applying the permutation $\pi : u_k \rightarrow k$ to the symbols of H we obtain a hypercube H' to which Case 2 applies. That means that π is an automorphism of H' , and hence also of H . We deduce that $|\text{Aut}(H)| = 3$.

Moreover, there must be exactly twice as many hypercubes in Case 3 as there are in Case 2. Any example from Case 2 can be turned into an example from Case 3 by applying either (123) or (321) to its symbols. Exactly one of these permutations will map an example from Case 3 to one from Case 2. It follows that $|\mathcal{H}_{3,3}^d| = 2|\mathcal{H}_{3,6}^d|$ and thus $|\mathcal{Q}_{3,3}^d| = |\mathcal{Q}_{3,6}^d|$.

We now consider how a $(d-1)$ -dimensional hypercube H can be embedded as the first hyperplane of a d -dimensional hypercube H' . Since our hypercubes are linear there is only one choice to make, namely, whether $c_d = 1$ or $c_d = 2$. If H is in Case 2 or 3, then for either choice H' must fall into Case 1. On the other hand, if H is in Case 1, then there is exactly one choice of c_d so that H' is also in Case 1. Compiling this information we get

$$\begin{aligned} |\mathcal{H}_{3,2}^d| &= 2|\mathcal{H}_{3,3}^{d-1} \cup \mathcal{H}_{3,6}^{d-1}| + |\mathcal{H}_{3,2}^{d-1}|, \\ |\mathcal{H}_{3,3}^d \cup \mathcal{H}_{3,6}^d| &= |\mathcal{H}_{3,2}^{d-1}| \end{aligned}$$

and hence

$$\begin{aligned} |\mathcal{Q}_{3,2}^d| &= 2|\mathcal{Q}_{3,3}^{d-1}| + |\mathcal{Q}_{3,2}^{d-1}|, \\ |\mathcal{Q}_{3,3}^d| &= |\mathcal{Q}_{3,2}^{d-1}|. \end{aligned}$$

The theorem now follows from (5) by induction on d . $\square$

5. Number of identity elements. It is well known that a binary quasigroup can have at most one identity element but that d -ary quasigroups for $d > 2$ can have multiple identity elements. In this section we investigate constructions which produce quasigroups with a specific number of identity elements.

It is beyond the scope of this paper to provide a complete answer to the question of for which n , d , and i there exists a quasigroup in $\mathcal{Q}_n^d$ with exactly i identity elements. We set the more modest goal of, at a minimum, constructing at least one quasigroup in all cases for which Table 2 claims existence. Our methods and results represent a step toward an eventual solution of the more general question. We begin with some notation that will be used throughout this section.

Suppose $A \in \mathcal{Q}_n^\alpha$ and $B \in \mathcal{Q}_n^\beta$. We define a new quasigroup in $\mathcal{Q}_n^{\alpha+\beta-1}$, which we call the *composition* of A and B and write as $A \vdash B$, by

$$(A \vdash B)(x_1, \dots, x_{\alpha+\beta-1}) = A(B(x_1, \dots, x_\beta), x_{\beta+1}, \dots, x_{\alpha+\beta-1}).$$

Also, for any quasigroup A and positive integer c we define A^c inductively by $A^1 = A$ and $A^c = A^{c-1} \vdash A$ for $c > 1$. In particular, if $\mathbb{Z}_n$ denotes the quasigroup in $\mathcal{Q}_n^2$ defined by $\mathbb{Z}_n(x, y) \equiv x + y \pmod{n}$, then $(\mathbb{Z}_n)^c$ is the quasigroup in $\mathcal{Q}_n^{c+1}$ which is evaluated simply by adding coordinates modulo n .

The next lemma makes some observations about the number of identity elements in compositions. The (entirely elementary) proof will be omitted. In part 3, the meaning of *total loop* is a loop in which every element is an identity element.

LEMMA 6.

1. $(\mathbb{Z}_n)^c \in \mathcal{Q}_n^{c+1}$ has $\gcd(n, c)$ identity elements.
2. If x is an identity element in both A and B , then x is an identity element in $A \vdash B$. The converse is false, as can easily be seen by considering observation 1.
3. Suppose A is a total loop. Then x is an identity element in $A \vdash B$ if and only if x is an identity element in B .

The following result proves, among other things, existence of total loops in many cases.

THEOREM 7. Suppose that n, d, i are positive integers satisfying

1. $n \notin \{3, 7\}$,
2. $d \geq 3$ and d is odd,
3. $i \equiv n \pmod{2}$ or $i = 2$.

Then there exists a loop in $\mathcal{Q}_n^d$ with exactly i identity elements.

Proof. For any odd $d > 1$ and any $n \notin \{3, 7\}$ Teirlinck [26] showed the existence in $\mathcal{Q}_n^d$ of a 2-idempotent quasigroup, namely, a quasigroup in which each subset of cardinality 1 or 2 is a subquasigroup. It is easy to see that every 2-idempotent quasigroup is a total loop, proving the case $i = n$.

Suppose $i \equiv n \pmod{2}$. In a 2-idempotent quasigroup, we can replace each of the subquasigroups on the pairs $\{i+1, i+2\}$, $\{i+3, i+4\}, \dots, \{n-1, n\}$ by the other possible subquasigroup on the same two elements. Doing so results in a new quasigroup in which none of $i+1, \dots, n$ is an identity element. However, no tuple containing any element less than $i+1$ has been altered, so the elements $1, \dots, i$ are all still identity elements.

It remains to show the $i = 2$ case. Again we start with a 2-idempotent quasigroup. Suppose after replacing the subquasigroup on the elements $\{1, 2\}$ by the other possible subquasigroup on those elements, we arrive at a quasigroup Q . Now define Q' by

$$Q'(\vec{x}) = \begin{cases} 1 & \text{if } Q(\vec{x}) = 2, \\ 2 & \text{if } Q(\vec{x}) = 1, \\ Q(\vec{x}) & \text{otherwise.} \end{cases}$$

Then Q' will be a quasigroup in which 1 and 2 are identity elements. However, if $x \notin \{1, 2\}$, then $Q'(1, x, x, \dots, x) = 2$ so x is not an identity element. $\square$

Teirlinck left open the case of existence when $n = 7$, but it is easy to check by computer that no total loop (and hence no 2-idempotent quasigroup) exists in $\mathcal{Q}_7^3$. Despite this, we will show that a total loop exists in $\mathcal{Q}_7^d$ for all sufficiently large d . Also, total loops (unlike 2-idempotent quasigroups) can exist in even dimensions. Consider $S \in \mathcal{Q}_7^3$ given by

$$\begin{aligned} S = & [5671234.6237415.7364152.1743526.2415763.3152647.4526371. \\ & 6237415.2345671.3471526.7514263.4652137.1726354.5163742. \\ & 7364152.3471526.6712345.4125637.1536274.5243761.2657413. \\ & 1743526.7514263.4125637.3456712.5267341.2631475.6372154. \\ & 2415763.4652137.1536274.5267341.7123456.6374512.3741625. \\ & 3152647.1726354.5243761.2631475.6374512.4567123.7415236. \\ & 4526371.5163742.2657413.6372154.3741625.7415236.1234567]. \end{aligned}$$

It is easy to check that $(\mathbb{Z}_7)^3 \vdash S$ is a total loop in $\mathcal{Q}_7^6$. By composing this total loop with itself we can then produce total loops in $\mathcal{Q}_7^{6+5k}$ for any integer $k \geq 0$.

THEOREM 8. For each odd $n > 3$ there exists a D such that $\mathcal{Q}_n^d$ contains a total loop for all $d \geq D$.

Proof. We first show that for $n = 7$ we can choose $D = 25$. For arbitrary positive integers k, c we have shown above the existence of a total loop $T \in \mathcal{Q}_7^{5k+1}$ and also noted in Lemma 6 that $(\mathbb{Z}_7)^{7c}$ is a total loop in $\mathcal{Q}_7^{7c+1}$. Forming $T \vdash (\mathbb{Z}_7)^{7c}$ then produces a total loop in $\mathcal{Q}_7^{7c+5k+1}$. Every integer exceeding 24 can be written in the form $7c + 5k + 1$ for nonnegative integers c, k . (See, for example, Theorem 3.15.1 in [27].)

The argument for odd $n \notin \{3, 7\}$ is similar, but we use a total loop $T \in \mathcal{Q}_n^{2k+1}$ instead (whose existence is guaranteed by Theorem 7). Using $T \vdash (\mathbb{Z}_n)^{nc}$ gives a total loop in $\mathcal{Q}_n^{nc+2k+1}$. Since n is odd, every integer exceeding $n-1$ can be written in the form $nc+2k+1$ for nonnegative integers c, k . Thus we may take $D = n$. $\square$

Next we consider quasigroups in which there is (only) one element which is *not* an identity element. It seems plausible, given the data from Table 2, that the following is the only general restriction on how many identity elements there can be in a ternary quasigroup.

LEMMA 9. *A 3-ary quasigroup of order n cannot have exactly $n-1$ identity elements.*

Proof. Let Q be a 3-ary quasigroup of order n on the symbols $[n]$ in which every $x \in [n] \setminus \{u\}$ is an identity element. Take an arbitrary $v \in [n] \setminus \{u\}$. Then there exists $x \in [n]$ such that $\langle v, x, u, v \rangle \in Q$. If $x \neq u$ this contradicts the fact that $\langle v, x, x, v \rangle \in Q$ since x is an identity element. We conclude that $\langle v, u, u, v \rangle \in Q$, and a similar argument shows that $\langle u, v, u, v \rangle \in Q$ and $\langle u, u, v, v \rangle \in Q$. Finally, there is some $x \in [n]$ such that $\langle x, u, u, u \rangle \in Q$. If $x \neq u$ this contradicts $\langle x, u, u, x \rangle \in Q$, so we must have $\langle u, u, u, u \rangle \in Q$. We have shown that u is an identity element, so if Q has at least $n-1$ identity elements then it must have n identity elements. $\square$

The data in Table 2 are consistent with a generalization of Lemma 9 to some higher dimensions. However, no such generalization is possible.

LEMMA 10. *For every $d > 3$ there exists an n and a $Q \in \mathcal{Q}_n^d$ such that Q has exactly $n-1$ identity elements.*

Proof. We first show that 5-ary quasigroups can have exactly $n-1$ identity elements. Let

$$\begin{aligned} E = & [78123456.85274163.12385674.27816345.34567812.41638527.56741238.63452781. \\ & 85274163.56781234.27416385.78123456.41638527.12345678.63852741.34567812. \\ & 12385674.27416385.34567812.81634527.56741238.63852741.78123456.45278163. \\ & 27816345.38527416.81634527.12345678.63452781.74163852.45278163.56781234. \\ & 34567812.41638527.56741238.63452781.78123456.85274163.12385674.27816345. \\ & 41638527.72143658.63852741.34567812.85274163.56781234.27416385.18325476. \\ & 56741238.63852741.78123456.45278163.12385674.27416385.34567812.81634527. \\ & 63452781.14365872.45278163.56781234.27816345.38527416.81634527.72143658] \end{aligned}$$

Now $Q \in \mathcal{Q}_8^5$ defined by $Q = (\mathbb{Z}_8)^2 \vdash E$ has exactly seven identity elements (the element 8 is not an identity element, since $Q(1, 8, 8, 8, 8) = 7$). By Theorem 7, there is a total loop $T \in \mathcal{Q}_8^{2k+1}$ for all $k \geq 1$. Then by Lemma 6, $T \vdash Q$ is a loop in $\mathcal{Q}_8^{2k+5}$ with exactly seven identity elements. Thus we have the required example whenever the dimension is odd.

For any even d , $(\mathbb{Z}_2)^{d-1}$ provides a (rather trivial) example. By Lemma 6, it is a d -ary quasigroup of order 2 with one identity element. $\square$

Our next result shows among other things that the $n=2$ examples used to prove Lemma 10 are the only linear examples of order n quasigroups with $n-1$ identity elements.

LEMMA 11. *The number of identity elements in a linear d -ary loop of order n divides n .*

Proof. Suppose L is a linear d -ary loop of order n , defined by (4). Since L is a loop it possesses at least one identity element, say, ι . Then $\langle y, \iota, \iota, \dots, \iota, y \rangle \in L$ for all $y \in [n]$ which implies that

$$(6) \quad s + \iota \sum_{i=2}^d c_i + c_1 y \equiv y \pmod{n}.$$

As (6) must hold for all y we conclude that $c_1 = 1$. A similar argument shows that $c_j = 1$ for all $j \in [d]$. Hence (6) reduces to $s + (d-1)\iota \equiv 0 \pmod{n}$. This recurrence either has no solution for ι (impossible since we know L has an identity element), or its number of solutions is a divisor of n . It is obvious that each solution will yield a (different) identity element. $\square$

Having considered loops with a high number of identity elements, we turn briefly to the other end of the spectrum. It is a trivial matter for any $n \geq 3$ and $d \geq 2$ to create a linear quasigroup with no identity elements. For example, take $c_i = -1$ for all i in (4), and consider the proof of Lemma 11 above. Building quasigroups with a unique identity element is also easy.

THEOREM 12. *For all $n \geq 4$ and $d \geq 2$ there exists a loop in $\mathcal{Q}_n^d$ with a unique identity element.*

Proof. The result for $d = 2$ is well known. For $d = 3$ and odd n we can use $(\mathbb{Z}_n)^2$, by Lemma 6.

Suppose n is even and $n > 2$. Let $Z'_n = \mathbb{Z}_n^{(\tau, \tau, \tau)}$ be the binary quasigroup isomorphic to $\mathbb{Z}_n$ by applying the transposition $\tau = (1 \frac{n}{2})$ uniformly to the triples. Now consider $Q = \mathbb{Z}_n \vdash Z'_n$. The element n is an identity element of $\mathbb{Z}_n$, Z'_n and hence also of Q , by Lemma 6. For $x \notin \{\frac{n}{2}, n\}$, we have $Q(x, n, x) \equiv Z'_n(x, n) + x \equiv 2x \not\equiv n \pmod{n}$, so x is not an identity element. Moreover, $\frac{n}{2}$ is not an identity element either, since $Q(\frac{n}{2}, \frac{n}{2}, n) = Z'(\frac{n}{2}, \frac{n}{2}) = 2 \neq n$.

We have thus shown the theorem for $d \leq 3$. From these base cases the theorem now follows for $d > 3$ and $n \neq 7$ by composition with a total loop whose existence is guaranteed by Theorem 7. For $n = 7$, we use the base cases $(\mathbb{Z}_7)^c$ for $c \in [6]$, each of which has a unique identity element, by Lemma 6. By composing these with the total loops from $\mathcal{Q}_7^{5k+1}$, we generate all the required examples for $n = 7$. $\square$

We now describe examples of quasigroups with each of the possible number of identity elements shown in Table 2.

If $n \leq 3$, then part 1 of Lemma 6 provides the examples we need. So we can assume henceforth that $n \in \{4, 5, 6\}$. In particular this means there exists a total loop in $\mathcal{Q}_n^3$, by Theorem 7, and a loop with a unique identity element in $\mathcal{Q}_n^d$ for all $d \geq 2$, by Theorem 12. This provides the required examples for $d = 4$, except that we need a quasigroup in $\mathcal{Q}_5^4$ with two identity elements. Let

$$\begin{aligned} Q_a &= [12345.23451.34512.45123.51234], \\ Q_b &= [12345.25413.34251.41532.53124], \\ Q_c &= [12345.25413.34521.43152.51234]. \end{aligned}$$

Then the quasigroup $Q(x_1, x_2, x_3, x_4) = Q_c(Q_a(x_1, x_2), Q_b(x_3, x_4))$, has two identity elements, as required.

By composition with a total loop we will obtain all required examples for $d = 5$ as soon as we find the examples for $d = 3$. Moreover, most of the examples needed for $d = 3$ can be found by applying Theorems 7 and 12. Only one case remains, namely,

we need an example in Q_6^3 with three identity elements. In fact, we give examples of all possible numbers of identity elements in Q_6^3 . Let

$$\begin{aligned} Q_0 &= [124653.213546.431265.652134.546312.365421], \\ Q_1 &= [123456.214365.632514.351642.465231.546123], \\ Q_2 &= [123456.214365.635214.356142.542631.461523], \\ Q_3 &= [123456.214365.632541.351624.546132.465213], \\ Q_4 &= [123456.214365.632514.351642.546231.465123], \\ Q_6 &= [123456.214365.632541.351624.546213.465132]. \end{aligned}$$

For each $i \in \{0, 1, 2, 3, 4, 6\}$, the quasigroup $Q_i \vdash Q_0$ has exactly i identity elements.

6. Incompletable latin cuboids. A natural method of building latin hypercubes is to add hyperplanes one at a time, as we did in section 3. We call the intermediate objects thereby created *latin hypercuboids*. In the three-dimensional case, Kochol [14] used the name “latin parallelepipeds.”

There is celebrated result due to Marshall Hall [10] that every latin rectangle is completable to a latin square. However, the equivalent statement is not true in higher dimensions. Kochol [14] proved that for any k and n satisfying $\frac{1}{2}n < k \leq n - 2$ there is an $n \times n \times k$ latin cuboid which cannot be completed to an $n \times n \times n$ latin cube. Although he did not say so, it is simple to use such examples to create noncompletable $n \times n \times \cdots \times n \times k$ latin hypercuboids in higher dimensions.

It is an open question how “thin” a noncompletable latin hypercuboid can be. In acknowledging that his theorem is not best possible, Kochol [14] alluded to an example of a noncompletable $5 \times 5 \times 2$ latin cuboid, although it seems this example might never have been published. Below we give examples of noncompletable $5 \times 5 \times 2$, $6 \times 6 \times 2$, $7 \times 7 \times 3$, and $8 \times 8 \times 4$ latin cuboids. In the latter two cases it remains open whether there are thinner examples. The $5 \times 5 \times 2$ cuboid cannot be extended even to three layers, while the other examples can be extended to $n \times n \times (n-2)$ latin cuboids, but no further.

$$\begin{aligned} &[12345.21453.34521.45132.53214. \\ &21453.13542.52134.34215.45321], \end{aligned}$$

$$\begin{aligned} &[123456.214365.345612.436521.561234.652143. \\ &214365.125643.456231.362154.643512.531426], \end{aligned}$$

$$\begin{aligned} &[1632745.6173452.2356174.5264317.3417526.7541263.4725631. \\ &2147653.7456231.3715462.6523174.4362715.1634527.5271346. \\ &3715426.1624573.6243751.4176235.2531647.5467312.7352164], \end{aligned}$$

$$\begin{aligned} &[14257638.68421375.26875413.57368241.85643127.31784562.43512786.72136854. \\ &26813457.47286531.38564172.71635824.13758246.62341785.54127368.85472613. \\ &31462785.14738256.62147538.83251467.78524613.57816324.25673841.46385172. \\ &42175863.86354712.51628347.35742186.27486531.14563278.78231654.63817425]. \end{aligned}$$

These examples show that Kochol’s theorem is not best possible for $n \in \{5, 6, 7, 8\}$ and hence leave wide open the question of what the best result might be in general. Cutler and Öhman [5] showed for all $a \leq m$ that every $(2mk) \times (2mk) \times a$ latin cuboid can be extended by one layer provided k is sufficiently large.

Appendix. In Tables 3, 4, and 5 we present the number of paratopy classes of latin hypercubes L with each combination of the parameters $d, n, I, P/I$, where d is the dimension, n is the order, $I = |\text{Is}(L)|$, and $P = |\text{Par}(L)|$. Lemma 1 explains how to compute the numbers of hypercubes, reduced hypercubes, and isotopy classes of hypercubes.

A representative of each of these paratopy classes can be found in [21].

TABLE 3
Counts of paratopy classes by $I = |\text{Is}(L)|$ and $P/I = |\text{Par}(L)|/|\text{Is}(L)|$ (part 1).

d	n	I	P/I	count	d	n	I	P/I	count	d	n	I	P/I	count
3	2	8	24	1	3	6	3	3	1	3	6	16	6	3
		total		1			3	4	337			16	8	24
3	3	54	24	1			3	6	1			16	24	5
		total		1			3	8	5			18	1	67
3	4	16	24	1			4	1	14855			18	2	54
		64	8	1			4	2	3898			18	4	28
		128	4	1			4	3	1			18	6	2
		128	24	1			4	4	1121			18	8	10
		384	24	1			4	6	12			20	1	1
		total		5			4	8	179			20	2	1
3	5	3	4	5			4	24	4			20	4	2
		4	4	1			5	4	1			24	1	67
		4	24	1			6	1	1302			24	2	75
		6	24	1			6	2	739			24	4	17
		10	4	1			6	4	244			24	6	1
		10	8	1			6	6	1			24	8	10
		12	4	1			6	8	13			27	1	5
		20	24	1			8	1	1118			27	2	11
		60	24	1			8	2	831			27	4	6
		100	8	1			8	3	1			27	8	3
		500	24	1			8	4	220			36	1	23
		total		15			8	6	6			36	2	26
3	6	1	1	75916			8	8	65			36	4	8
		1	2	17193			9	1	137			36	6	1
		1	3	29			9	2	90			36	8	6
		1	4	3049			9	4	30			48	1	2
		1	6	7			9	6	1			48	2	9
		1	8	309			9	8	13			48	4	8
		1	12	2			10	2	1			48	6	1
		1	24	3			10	8	3			48	8	3
		2	1	115256			12	1	316			54	1	6
		2	2	16921			12	2	208			54	2	6
		2	3	8			12	4	53			54	4	5
		2	4	2768			12	6	1			54	8	4
		2	6	15			12	8	7			54	24	1
		2	8	127			16	1	22			72	1	9
		3	1	4460			16	2	65			72	2	14
		3	2	1688			16	4	28			72	4	4

TABLE 4
Counts of paratopy classes by $I = |\text{Is}(L)|$ and $P/I = |\text{Par}(L)|/|\text{Is}(L)|$ (part 2).

d	n	I	P/I	count	d	n	I	P/I	count	d	n	I	P/I	count
3	6 (cont)	72	8	3	4	4	256	12	1	5	2	32	720	1
		81	8	1			512	8	2			total		1
		108	1	7			512	10	1	5	3	486	720	1
		108	2	2			512	12	2			total		1
		108	4	1			512	120	1	5	4	16	72	1
		108	8	2			1536	120	1			64	1	1909
		120	2	1			total		26			64	2	1574
		120	8	1			1	1	3			64	3	9
		144	2	3			1	2	5			64	4	572
		144	4	2			1	6	3			64	6	61
		144	8	3			2	2	2			64	8	164
		162	2	1			2	4	1			64	10	9
		216	1	2			2	6	1			64	12	88
		216	2	2			2	12	1			64	16	51
		216	8	1			3	1	24			64	20	1
		324	2	1			3	2	7			64	24	5
		324	4	1			3	6	3			64	36	8
		432	2	1			4	6	1			64	48	22
		432	24	1			5	2	9			64	60	1
		648	2	1			5	4	4			64	72	5
		720	8	1			10	2	1			64	120	4
		1296	8	1			10	4	2			64	720	3
		total		264248			10	6	2			128	1	16
4	2	16	120	1	4	5	10	12	1			128	2	24
		total		1			12	2	1			128	4	54
4	3	162	120	1			12	6	1			128	6	8
		total		1			20	2	2			128	8	13
4	4	32	8	2			20	12	1			128	12	15
		32	10	1			25	8	2			128	16	12
		32	12	3			50	4	1			128	36	2
		32	120	2			50	8	2			128	48	14
		64	4	3			50	12	2			128	60	1
		64	12	2			100	8	1			128	720	2
		64	24	2			100	12	1			256	1	4
		128	8	1			500	12	1			256	2	26
		256	4	1			2500	120	1			256	4	24
		256	8	1			total		86			256	6	8

TABLE 5
Counts of paratopy classes by $I = |\text{Is}(L)|$ and $P/I = |\text{Par}(L)|/|\text{Is}(L)|$ (part 3).

d	n	I	P/I	count	d	n	I	P/I	count	d	n	I	P/I	count
5	4 (cont)	256	8	18	5	5	2	8	1	5	5	25	6	9
		256	12	6			2	9	1			25	8	81
		256	16	6			2	12	3			25	12	4
		256	24	1			2	18	10			25	16	5
		256	36	2			2	36	6			50	4	4
		256	48	7			3	1	1128			50	6	1
		512	4	2			3	2	378			50	8	13
		512	12	1			3	3	31			50	12	3
		512	16	2			3	4	55			50	16	5
		1024	4	3			3	6	39			50	18	2
		1024	8	2			4	2	2			50	36	1
		1024	12	3			4	3	1			60	36	1
		1024	16	3			4	4	1			100	4	3
		1024	48	2			4	6	2			100	6	2
		1024	72	1			4	18	3			100	8	8
		2048	4	4			4	36	3			100	12	1
		2048	12	2			4	72	1			100	16	1
		2048	16	2			5	2	468			100	36	1
		2048	36	1			5	4	18			125	8	2
		2048	48	4			5	6	26			125	12	2
		2048	60	1			5	12	2			125	24	2
		2048	720	1			6	18	2			125	48	1
		6144	720	1			6	36	3			250	8	1
		total		4785			10	2	14			250	12	3
5	5	1	1	337			10	4	14			250	16	3
		1	2	105			10	6	10			250	36	1
		1	3	72			10	8	4			250	48	2
		1	4	28			10	12	5			500	12	1
		1	6	40			10	36	1			500	16	2
		1	12	4			12	2	2			500	36	1
		1	18	3			12	3	1			500	48	1
		1	36	4			12	4	1			2500	48	1
		2	1	3			12	6	1			2500	72	1
		2	2	6			20	2	2			12500	720	1
		2	3	7			20	12	1			total		3102
		2	4	7			20	72	1					
		2	6	5			25	4	70					

REFERENCES

- [1] Y. ALAVI, D. R. LICK, AND J. LIU, *Strongly diagonal Latin squares and permutation cubes*, Congr. Numer., 102 (1994), pp. 65–70.
- [2] R. A. BAILEY, D. A. PREECE, AND P. J. ZEMROCH, *Totally symmetric Latin squares and cubes*, Utilit. Math., 14 (1978), pp. 161–170.
- [3] G. B. BELJAVSKAJA AND S. MURATHUDJAEV, *About admissibility of N -ary quasigroups*, Combinatorics, Vol. I, Colloq. Math. Soc. János Bolyai, 18, North-Holland, Amsterdam, 1978, pp. 101–119.
- [4] A. B. CRUSE, *On the finite completion of partial Latin cubes*, J. Combin. Theory Ser. A, 17 (1974), pp. 112–119.
- [5] J. CUTLER AND L.-D. ÖHMAN, *Latin squares with forbidden entries*, Electron. J. Combin., 13 (2006), R47.
- [6] J. DÉNES AND A. D. KEEDWELL, *Latin Squares and Their Applications*, Akadémiai Kiadó, Budapest, 1974.
- [7] N. J. FINIZIO AND J. T. LEWIS, *Enumeration of maximal codes*, Congr. Numer., 102 (1994), pp. 139–145.
- [8] H.-L. FU, *On Latin cubes with prescribed intersections*, Ars Combin., 23 (1987), pp. 171–176.
- [9] H. GUPTA, *On permutation cubes and Latin cubes*, Indian J. Pure Appl. Math., 5 (1974), pp. 1003–1021.

- [10] M. HALL, *An existence theorem for Latin squares*, Bull. Amer. Math. Soc., 51 (1945), pp. 387–388.
- [11] K. HEINRICH, *Prolongation in m -dimensional permutation cubes*, in Algebraic and Geometric Combinatorics, North-Holland Math. Stud. 65, North-Holland, Amsterdam, 1982, pp. 229–238.
- [12] T. ITO, *Creation Method of Table, Creation Apparatus, Creation Program and Program Storage Medium*, U.S. Patent application 20040243621, Dec. 2, 2004.
- [13] X. W. JIA AND Z. P. QIN, *The number of Latin cubes and their isotopy classes*, J. Huazhong Univ. Sci. Tech., 27 (1999), pp. 104–106 (in Chinese).
- [14] M. KOCHOL, *Relatively narrow Latin parallelepipeds that cannot be extended to a Latin cube*, Ars Combin., 40 (1995), pp. 247–260.
- [15] C. F. LAYWINE AND G. L. MULLEN, *Latin cubes and hypercubes of prime order*, Fibonacci Quart., 23 (1985), pp. 139–145.
- [16] C. F. LAYWINE AND G. L. MULLEN, *Discrete Mathematics Using Latin Squares*, John Wiley, New York, 1998.
- [17] C. C. LINDNER, *A finite partial idempotent Latin cube can be embedded in a finite idempotent Latin cube*, J. Combin. Theory Ser. A, 21 (1976), pp. 104–109.
- [18] B. D. MCKAY, A. MEYNERT, AND W. MYRVOLD, *Small Latin squares, quasigroups and loops*, J. Combin. Des., 15 (2007), pp. 98–119.
- [19] B. D. MCKAY, *nauty* Graph Isomorphism Software, <http://cs.anu.edu.au/~bdm/nauty>.
- [20] B. D. MCKAY AND I. M. WANLESS, *On the number of Latin squares*, Ann. Comb., 9 (2005), pp. 335–344.
- [21] B. D. MCKAY AND I. M. WANLESS, *Latin Cubes and Hypercubes*, <http://cs.anu.edu.au/~bdm/data/latincubes.html>.
- [22] G. L. MULLEN AND R. E. WEBER, *Latin cubes of order ≤ 5* , Discrete Math., 32 (1980), pp. 291–297.
- [23] V. N. POTAPOV AND D. S. KROTOV, *Asymptotics for the number of n -quasigroups of order 4*, Siberian Math. J., 47 (2006), pp. 720–731.
- [24] D. A. PREECE, S. C. PEARCE, AND J. R. KERR, *Orthogonal designs for three-dimensional experiments*, Biometrika, 60 (1973), pp. 349–358.
- [25] N. J. A. SLOANE, *The On-Line Encyclopedia of Integer Sequences*, <http://www.research.att.com/~njas/sequences/>.
- [26] L. TEIRLINCK, *Generalized idempotent orthogonal arrays*, in Coding Theory and Design Theory, Part II, IMA Vol. Math. Appl. 21, Springer, New York, 1990, pp. 368–378.
- [27] H. S. WILF, *Generatingfunctionology*, 2nd ed., Academic Press, San Diego, 1994.
- [28] V. A. ZINOVIEV AND D. Z. ZINOVIEV, *Binary extended perfect codes of length 16 obtained by the generalized concatenated construction*, Probl. Inf. Transm., 38 (2002), pp. 296–322.

THE NINE MORSE GENERIC TETRAHEDRA*

D. SIERSMA[†] AND M. VAN MANEN[‡]

Abstract. There are two types of shapes for a generic triangle—acute and obtuse. These shapes are also distinguished by the (topological) Morse theory of the minimal distance function to the vertices. We can use the same method for a tetrahedron, and we show in this paper that there exist nine generic shapes. These can be described by a Morse poset or by a Gabriel graph. We also report on some computer experiments and compare our classification to another criterion used in computational geometry.

Key words. Morse theory, polyhedra, Voronoi diagram, tetrahedral shape

AMS subject classifications. Primary, 51M20; Secondary, 68U05

DOI. 10.1137/060656395

1. Introduction. Take N points $P_1, \dots, P_N$ in $\mathbb{R}^n$ and consider the function $d : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$d(X) = \min_{j=1, \dots, N} d(X, P_j).$$

We study the evolution of the sets $d_\epsilon = \{X \mid d(X) \leq \epsilon\}$ as ϵ increases. In particular, we are interested in the Euler characteristic χ of d_ϵ . In case ϵ is very small, d_ϵ consists of N small solid spheres. Thus $\chi = N$. If ϵ is very big, then d_ϵ is contractible, and hence $\chi = 1$.

For a generic set of points, d is a topological Morse function. In that case, as ϵ grows, d passes through a number of nondegenerate critical values. When d passes a critical value of index i , an i -cell gets attached.

The number of critical points of index i is s_i . From Morse theory we know that

$$\sum (-1)^i s_i = 1.$$

As an example, take a triangle with an obtuse angle in the plane. This is a special case of the above problem with $n = 2$ and $N = 3$. Assume further that the two legs that encompass the obtuse angle have different lengths. In that case, $s_0 = 3$, $s_1 = 2$, and $s_2 = 0$. For an acute triangle where the edges have different lengths, we obtain $s_0 = 3$, $s_1 = 3$, and $s_2 = 1$. In this sense, there are two different generic triangles.

Returning to the n -dimensional case, each critical point of index i corresponds to a subset of size $i + 1$ of $\{P_1, \dots, P_N\}$, but not every subset of size $i + 1$ corresponds to a critical point of index i . (With the obtuse triangle, the 2-face of the triangle, that is, the triangle itself, does not correspond to a local maximum of the function d .) Thus to the N -point set $P_1, \dots, P_N$ we can associate a set of subsets that correspond to critical points of d . By extension we will call these subsets and the geometric simplices that they span critical as well. This set of subsets is partially ordered by inclusion, and thus it is a poset. We will call it the *Morse poset*.

*Received by the editors April 5, 2006; accepted for publication (in revised form) December 10, 2007; published electronically April 18, 2008.

<http://www.siam.org/journals/sidma/22-2/65639.html>

[†]Mathematisch Instituut, Universiteit Utrecht, P.O. Box 80010, 3508 TA Utrecht, The Netherlands (D.Siersma@uu.nl).

[‡]Mathematical Institute, Hokkaido University, Kita 10, Nishi 8, Kita-Ku, Sapporo, Hokkaido, 060-0810, Japan (mvanmanen@gmail.com).

Question. In $\mathbb{R}^n$, for generic sets of N points, how many different Morse posets are there, up to combinatorial equivalence?

When $N = n + 1$ we thus ask how many “different” generic simplices there are. This is a natural first problem to consider.

In the plane the answer is two; see [Sie99, section 2]. For obtuse triangles, the Morse poset is

$$\{\{P_1\}, \{P_2\}, \{P_3\}, \{P_1, P_3\}, \{P_2, P_3\}\},$$

and for acute triangles, we get the Morse poset

$$\{\{P_1\}, \{P_2\}, \{P_3\}, \{P_1, P_2\}, \{P_1, P_3\}, \{P_2, P_3\}, \{P_1, P_2, P_3\}\}.$$

The main theorem of this article says that in $\mathbb{R}^3$, there are nine different generic tetrahedra (3-simplices).

In the first section, we recall the relevant Morse theory. Then we establish some notation and state what Voronoi diagrams and Gabriel graphs are. Next, we state and prove the main theorem. We also report on numerical experiments concerning volume data of the different compartments of the configuration space where the Morse poset is of a certain type. Finally, we compare our classification to the classification by shape types in [Ede01]. The Morse theoretic approach relates to alpha shapes as defined by Edelsbrunner [EM94].

2. Genericity conditions. We focus here on four points in $\mathbb{R}^3$, but most of the notation and definitions have straightforward extensions to the (n, N) general case.

We write P_{ij} for the middle of the interval P_iP_j , P_{ijk} for the center of the circumscribed circle of the triangle $P_iP_jP_k$, and P_{1234} for the center of the circumscribed sphere of the tetrahedron. For the tetrahedron itself, that is, the convex hull of $\{P_1, P_2, P_3, P_4\}$, we will use the notation $\mathbb{T}$.

We impose the following condition.

GENERICITY CONDITION 2.1. *We require the set of points P_1, P_2, P_3, P_4 to be in general position, so that the convex hull of P_1, P_2, P_3, P_4 is 3-dimensional. Moreover, the points P_{ijk} do not lie on one of the edges of the triangle $P_iP_jP_k$, and also P_{1234} does not lie on one of the planes of the triangles $P_iP_jP_k$.*

Genericity Condition 2.1 means that the function d is not too badly behaved. To express more carefully what that means we recall the definition of a topological Morse function; see [Mor59]. Let $P \in \mathbb{R}^n$, and let f be a continuous real-valued function on $\mathbb{R}^n$.

DEFINITION 2.2. *The function f is topologically regular at $P \in \mathbb{R}^n$ if there is some neighborhood U of P and a homeomorphism $\phi: U \rightarrow U$ such that the composition $f \circ \phi$ is a nonconstant affine function on U . The function f has a critical point at $P \in \mathbb{R}^n$ if f is not topologically regular at P . In that case, P is called a nondegenerate critical point of index i if there is a neighborhood U of P and a homeomorphism $\phi: U \rightarrow U$ such that*

$$f \circ \phi = f(P) - \sum_{j=1}^i x_j^2 + \sum_{j=i+1}^n x_j^2.$$

A topological Morse function is a continuous function that has only regular and nondegenerate critical points.

For topological Morse functions the two crucial statements that hold in the differentiable case—the regular interval theorem and the attachment of cells (see Chapter 5

in [Mil63])—are true as well, as Morse proves in [Mor73]. So, if we show that d is topologically regular, we can apply those theorems, just as was done in [Sie99].

We will need the following notation.

Let

$$\text{Terr}(P_i) = \{X \in \mathbb{R}^3 \mid d(X, P_i) \leq d(X, P_k) \text{ for all } k\}$$

and $V_i = \text{Terr}(P_i)$, $V_{ij} = V_i \cap V_j$ (i different from j), $V_{ijk} = V_i \cap V_j \cap V_k$ (i, j, k different from one another), and $V_{1234} = V_1 \cap V_2 \cap V_3 \cap V_4$.

PROPOSITION 2.3. *The function d is a topological Morse function if Genericity Condition 2.1 is fulfilled. In that case, d is topologically regular in all points of $\mathbb{R}^3$, except in P_1, P_2, P_3, P_4 , where d has a minimum, and (perhaps) in the points P_{ij} , P_{ijk} , and P_{1234} . Moreover, d has*

- a minimum exactly in the points P_1, P_2, P_3, P_4 ,
- a 1-saddle (saddle point of index 1) in P_{ij} iff $P_{ij} = V_{ij} \cap P_i P_j$,
- a 2-saddle (saddle point of index 2) in P_{ijk} iff $P_{ijk} = V_{ijk} \cap P_i P_j P_k$, and
- a maximum in P_{1234} iff $P_{1234} = V_{1234} \cap \mathbb{T}$ (equivalently, $P_{1234} \in \mathbb{T}$).

Proof. If $x = P_i$, then because the points lie in general position all P_j with $i \neq j$ lie at some positive distance from x , so d has a minimum there. If $x \in \text{Terr}(P_i)$ but $x \neq P_i$, then x is obviously topologically regular.

The function d restricted to the interior V_{ij} has a minimum if P_{ij} lies in that interior. This can happen only when $P_{ij} \neq P_{ijk}$, which is ensured by Genericity Condition 2.1.

In the directions, orthogonal to V_{ij} , d decreases, so we see that d has a critical point of index 1 at P_{ij} .

The function d restricted to the interior of V_{ijk} has a minimum at P_{ijk} if P_{ijk} lies in that interior. This can be the case only when $P_{ijk} \neq P_{1234}$. In the directions orthogonal to V_{ijk} , d decreases, so d has a critical point of index 2 at P_{ijk} .

Finally, if $P_{1234} \in \mathbb{T}$, d obviously has a local maximum there. $\square$

Remark 2.4. This statement is well known. One can find it in several places; see, for instance, [Ede04].

Remark 2.5. In the 2-dimensional case the distance function d is always a topological Morse function, whether the points are in general position or not. For a proof, see [Sie99]. It would be interesting to know whether the distance function would always be a topological Morse function in higher dimensions also, irrespective of general position considerations.

The conditions of Proposition 2.3 tell us exactly the positions of the critical points. As explained in the introduction, a critical point determines a subset of the points P_1, P_2, P_3, P_4 , which we call *critical subset*. The *Morse poset* is the set of critical subsets, partially ordered by inclusion. Two sets of points in $\mathbb{R}^3$ are called *combinatorially equivalent* if there exists a bijection between points that sends the critical subsets onto each other. We want to give a classification with respect to this equivalence relation.

Throughout this paper we deal with four points, generic in the sense of Genericity Condition 2.1 and contained in $\mathbb{R}^3$. However, it seems relevant to point out that our problem can be formulated in a more general setting. In the next few definitions we will deal with a point set $\{P_1, \dots, P_N\}$ in $\mathbb{R}^n$ without making any genericity assumptions.

DEFINITIONS 2.6. *A subset $\{P_i\}_{i \in I}$ belongs to the Delaunay triangulation if there are points $X \in \mathbb{R}^n$ with $d(X, P_i) = d(X, P_j)$ for all pairs $\{i, j\} \subset I$ and $d(X, P_i) < d(X, P_j)$ for all pairs $\{i, j\}$ with $i \in I$ and $j \notin I$. The Voronoi cell of such a subset $\{P_i\}_{i \in I}$ consists of the points $X \in \mathbb{R}^n$ that make the subset part of the Delaunay*

triangulation. The Delaunay cell of such a subset $\{P_i\}_{i \in I}$ is simply the convex hull of $\{P_i\}_{i \in I}$.

In particular, all subsets of cardinality 1 are part of the Delaunay triangulation. For each $1 \leq i \leq N$ we denote the closure of the Voronoi cell of P_i by $\text{Terr}(P_i)$, because such a Voronoi cell is best thought of as a territory with capital P_i .

The union of all $< n$ dimensional Voronoi cells is usually called the Voronoi diagram. In fact, the closure of any $(n - 1)$ -dimensional Voronoi cell is always the intersection of exactly two sets $\text{Terr}(P_i)$ and $\text{Terr}(P_j)$. Thus, dual to the Voronoi diagram we have the 1-skeleton

$$\{P_i P_j \mid \dim(\text{Terr}(P_i) \cap \text{Terr}(P_j)) = n - 1\}.$$

The Gabriel graph (see [BKOS97], where it is defined in $\mathbb{R}^2$) is a subset of this skeleton: it consists of those edges $P_i P_j$ such that the circumball with poles P_i and P_j does not contain any of the other points in the point set $\{P_1, \dots, P_N\}$. The Gabriel complex (see [AGJ00]) is the subcomplex of the Delaunay triangulation, generated by those $(n - 1)$ -dimensional Delaunay cells whose smallest circumball does not contain any of the other points.

Remark 2.7. General position is defined in this more general case as follows. We require that for each subset $\{P_i\}_{i \in I}$ in the Delaunay triangulation the possibly empty intersection of the Delaunay cell and the Voronoi cell is contained in both the relative interior of the Delaunay and the relative interior of the Voronoi cell. For more details, see [Ede01] and [EM94].

QUESTION 2.8. What is the relation between the Morse poset and the Gabriel graph and complex? The edge between P_i and P_j is part of the Gabriel graph iff the point P_{ij} is a saddle point (index 1) of d . These are exactly the critical subsets of size 2 of the Morse poset. In the 3-dimensional case a face between P_i , P_j , and P_k is part of the Gabriel complex iff the point P_{ijk} is a saddle point (index 2) of d . These are precisely the critical subsets of size 3 of the Morse poset. But the Gabriel complex also contains by definition the boundaries of those 2-simplices, while for the Morse poset this is not automatically the case. So the Morse poset is not a simplicial complex and can differ from the Gabriel complex. See examples of tetrahedrons of type $(4, 5, 3, 1)$ in the next section.

The Morse poset contains the information from both the Gabriel graph and the Gabriel complex. Two Morse posets can be combinatorially equivalent only if the underlying Gabriel graph and Gabriel complex are the same.

3. The main classification. From Proposition 2.3 we know the maximal number of critical points of each type. Moreover, the Euler characteristic is $+1$: $s_0 - s_1 + s_2 - s_3 = 1$. This gives a priori the following 9 possibilities:

s_0	s_1	s_2	s_3
4	6	4	1
4	5	3	1
4	4	2	1
4	3	1	1
4	2	0	1
4	6	3	0
4	5	2	0
4	4	1	0
4	3	0	0

But not all possibilities will occur. Since we start with four vertices and the result should be a connected space, we need at least three saddle points of index 1. This rules out the possibility $(4, 2, 0, 1)$.

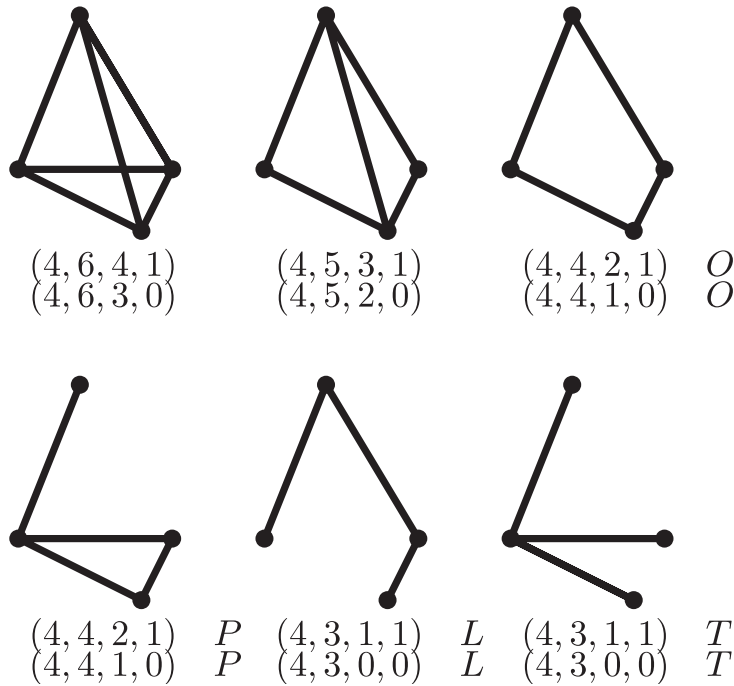


FIG. 1. List of possible Gabriel graphs. The vertices are the minima of the distance function. Each vertex of the tetrahedron $\mathbb{T}$ is a minimum of d . There are no other minima. The edges of the graph are the index 1 critical points of the distance function. Not each midpoint of an edge of $\mathbb{T}$ is an index 1 critical point of the distance function d . The graph $(4410)O$ can be laid out so as to form the letter “O.” The names of the other graphs are chosen similarly.

We list in Figure 1 the (a priori) possible Gabriel graphs for the above cases; they are the connected graphs with four vertices. Just as the cases $(4, 2, 0, 1)$ cannot occur, we will prove in this section by deriving a contradiction that the cases $(4, 4, 2, 1)P$, $(4, 3, 1, 1)T$, and $(4, 3, 1, 1)L$ do not occur. We will prove our main theorem.

THEOREM 3.1. *Up to combinatorial equivalence of their Morse posets there are nine generic tetrahedra. They are uniquely described by the nine Gabriel graphs $(4, 6, 4, 1)$, $(4, 6, 3, 0)$, $(4, 5, 3, 1)$, $(4, 5, 2, 0)$, $(4, 4, 2, 1)O$, $(4, 4, 1, 0)O$, $(4, 4, 1, 0)P$, $(4, 3, 0, 0)L$, and $(4, 3, 0, 0)T$, drawn in Figure 1.*

Proof. We have to exclude $(4, 3, 1, 1)T$, $(4, 3, 1, 1)L$, and $(4, 4, 2, 1)P$ from the list. In these cases M is critical. We first pay attention to the following.

3.1. Saddle points of index 1. We consider the plane E through $P_1P_2P_3$ (and set it as the “ground plane” in the picture). The half space containing P_4 is called “above,” and the other is called “below.”

We are going to consider the condition that the point $Y = P_{i4}$ is critical. We first make no assumptions about the triangle $P_1P_2P_3$ except that it does not have a right angle. We consider the point $X = P_4$ as a variable and denote its projection on the E -plane by X' . Let $Y = P_{14}$ and its projection be Y' . From the condition

that $Y = P_{14}$ is critical we get

$$d(P_1, Y) = d(P_4, Y) < \min \{d(P_2, Y), d(P_3, Y)\}.$$

The inequality is strict because of Genericity Condition 2.1.

For the projection, this means that in terms of Voronoi diagrams in E ,

$$Y' \in \text{Interior}(\text{Terr}(P_1)).$$

In terms of the projection X' , this means

$$X' \in \text{Interior}(2\text{Terr}(P_1)),$$

where by $2\text{Terr}(P_1)$ we mean scalar multiplication of $\text{Terr}(P_1)$ by a factor 2 from the point P_1 .

We can do the same for P_{24} and P_{34} . We get three subsets $2\text{Terr}(P_1)$, $2\text{Terr}(P_2)$, and $2\text{Terr}(P_3)$, which cover the plane E . They divide the plane into regions, where one, two, or three of the points P_{14} , P_{24} , or P_{34} are critical.

Inside the plane E , the border of $\text{Terr}(P_i)$ consists of two half lines that meet in P_{123} . The scalar multiplication by 2 maps P_{123} to P_i^* , the antipodal point of P_i on the circle through P_1 , P_2 , and P_3 .

Next, look at the case of an acute triangle drawn in Figure 2.

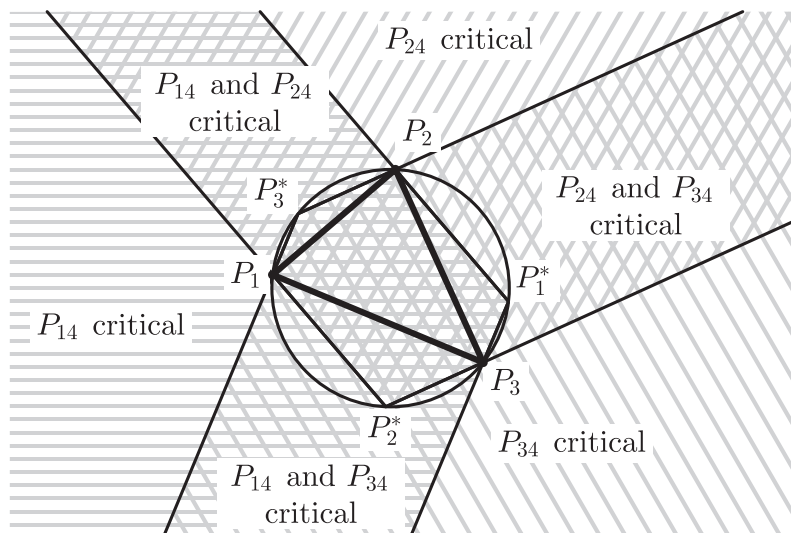


FIG. 2. The plane E when the triangle is acute.

The regions where only one P_{i4} is critical are all outside the disc D , which is bounded by the circumscribed circle of triangle $P_1P_2P_3$.

The picture for the obtuse case is in Figure 3.

Let P_2 be the vertex with the obtuse angle. We see that the region where only P_{24} is critical is outside D but that the regions where P_{14} or P_{34} is the only critical point can have some intersections with D .

3.2. The center M of the circumscribed sphere. We assume again that $P_1P_2P_3$ lies in the plane E . We know that $M = P_{1234}$ belongs to the axis of the triangle $P_1P_2P_3$, so its projection is $M' = P_{123}$. Fix M for the moment, and consider

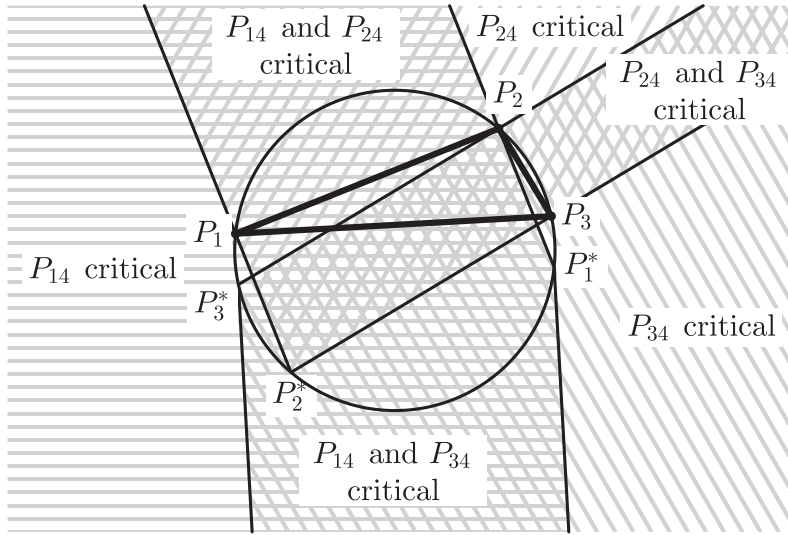


FIG. 3. The plane E when the triangle is obtuse. Indicated are the territories $2\text{Terr}(P_i)$, $i = 1, \dots, 3$, and their intersections.

$X = P_4$ as a variable. Possible positions of $X = P_4$ are on the sphere with center M and radius $r = d(P_1, M) = d(P_2, M) = d(P_3, M)$. Denote by $P_1^*P_2^*P_3^*$ the mirror image of $P_1P_2P_3$ after central reflection through P_{123} .

In the cases we wish to exclude, M is critical. In other words, the distance function has a local maximum in M , and M lies inside the tetrahedron. The point X must lie in the double cone with apex M and base the triangle $P_1P_2P_3$, but also on the sphere (M, r) . Consider the plane E'' , which is the image of E by central reflection through M . E'' is parallel to E . The cone intersects the plane in a triangle, whose orthogonal projection onto the plane E is the triangle $P_1^*P_2^*P_3^*$. X belongs to a part of the cap of the sphere, which lies “above” E'' (seen from M). It follows that X projects to a point X' which lies inside the disc D . This is true for any position of M . For the acute case this is all we need to conclude that it is not possible that M and only one P_{i4} are critical. In the obtuse case we have to use additional arguments to get to this same conclusion. Let P_2 be the obtuse angle. Recall that M is critical iff it belongs to the interior of the tetrahedron $P_1P_2P_3X$. Equivalently, X' lies in the interior of the reflected triangle, that is, in $P_1^*P_2^*P_3^*$. This triangle is contained in the sector $P_1^*M'P_3^*$, which implies that M is critical only if X' lies in the interior of this sector (the one which does not contain P_2). We combine the conditions for M to be critical and one single P_{i4} is critical. From the geometric observations above it follows that the combination of M being critical and a single P_{i4} being critical cannot occur together.

Next choose the appropriate labels for the vertices of the tetrahedron: In cases $(4, 3, 1, 1)T$ and $(4, 3, 1, 1)L$ the ground plane should contain two adjacent critical edges, and in case $(4, 4, 2, 1)P$ it should contain the triangle with the three critical edges. This observation rules out the graphs $(4, 3, 1, 1)T$, $(4, 3, 1, 1)L$, and $(4, 4, 2, 1)P$.

3.3. Positions of index 2 saddles. We are left with nine possibilities of Gabriel graphs. The Gabriel graphs do not give a complete picture of the combinatorics of the Morse points. The Gabriel graph, together with the information on whether M is a critical point or not, tells us only the number of saddles of index 2, but not the position.

LEMMA 3.2. *Generic tetrahedra with isomorphic Gabriel graphs have combinatorially equivalent Morse posets.*

Proof. We need to prove that in all cases there are unique positions for the 2-saddles.

(4, 6, 4, 1): Unique positions (no choices).

(4, 5, 3, 1): There are two triangles, where all three midpoints are critical. Both triangles must be acute. Take one of them in the “ground plane” and assume it to be $P_1P_2P_3$. Since P_{1234} is critical, P_{1234} lies on the same side as P_4 . It follows that V_{123} must intersect the ground plane in the point P_{123} , so that point is critical. The same reasoning applies to the second triangle; this fixes the position of two 2-saddles. The two positions left for the third saddle are combinatorially equivalent.

(4, 4, 1, 0) O : All positions equivalent.

(4, 4, 2, 1): Suppose P_{12} is not critical. Now P_3 or P_4 must be contained in the ball $B = B(P_{12}, r_{12})$, where $r_{12} = d(P_1, P_{12})$. If P_{123} is critical, then triangle $P_1P_2P_3$ must be acute; this means that P_3 is outside the ball B . If P_{124} is critical, then triangle $P_1P_2P_4$ must be acute; this means that P_4 is outside the ball B . It follows that the situation where two 2-saddles are separated by a noncritical edge cannot occur. This fixes the places of the 2-saddles up to permutation.

(4, 6, 3, 0): All positions equivalent.

(4, 5, 2, 0): Choose a triangle, say, $P_1P_2P_3$, where all three edges are critical. If P_{123} is not critical, then it follows that P_{1234} lies below the ground plane E . We look again at the projection P'_4 . First P'_4 must lie inside D . But since two of the P_{i4} (say, P_{14} and P_{24}) must be critical, we know that P'_4 also must lie in the region described in section 3.1. The intersection is a subset of D , which is contained in the region D^* bounded by the arc $P_1P_3^*P_2$ and by the edges $P_2P_3^*$ and $P_3^*P_1$. There are still other conditions to meet:

- P_4 must lie outside the ball $B(P_{12}, r_{12})$, since P_{12} is critical.
- P_4 must lie inside the ball $B(P_{123}, r_{123})$, where $r_{123} = d(P_1, P_{123})$.

This is not simultaneously possible if P'_4 lies in D^* . It follows that there is only one possibility: the 2-saddles are the centers of two triangles with all edges critical.

(4, 4, 1, 0) P : Let $P_1P_2P_3$ be the triangle with all three midpoints of the edges critical. This triangle is acute; take its plane as ground plane. If P_{1234} is above the ground plane, then P_{123} is critical. Assume now that P_{1234} is below the ground plane. It follows that P'_4 lies inside the disc D . But the fact that only P_{14} is critical means that P'_4 is outside. This is a contradiction, so only P_{123} can be critical.

(4, 3, 0, 0) T : All places equivalent.

(4, 3, 0, 0) L : All places equivalent. $\square$

3.4. Existence and statistics. We carried out some statistical experiments to see how the different types of tetrahedra are distributed among different 4-tuples of points. That all the cases do occur follows from the computer experiments described below. As a consequence all nine generic tetrahedrons are realized and the proof of Theorem 3.1 is complete. $\square$

Four points in $\mathbb{R}^3$ do not form a bounded space, and thus there lives no uniform probability distribution on it. As the Morse poset is invariant under translations, scaling, and rotations, without loss of generality we can assume that all four points lie on S^2 . By translation and scaling, this can always be achieved. We consider the configuration space of different points in S^2 . The nongeneric tetrahedra form a hypersurface (of measure zero) in this space—the discriminant hypersurface.

The uniform distribution on $[0, 1]^2$ leads (as Archimedes already used) to a uniform distribution on the 2-sphere by the map $\gamma : [0, 1]^2 \rightarrow S^2 \subset \mathbb{R}^3$ given by

$$\gamma(a_1, a_2) = (\sin(\arccos(2a_1 - 1)) \sin(2\pi a_2), \sin(\arccos(2a_1 - 1)) \cos(2\pi a_2), 2a_1 - 1).$$

We take two random numbers in $[0, 1]$ and map them to S^2 using γ .

For our experiment we used the Gnu Scientific Library; see [Gal06]. This library has an implementation of the apparently very reliable MT19937 random number generator. Two random numbers in $[0, 1]$ were mapped to S^2 using γ . We took samples of 10^8 tetrahedra. Here is one:

1	(4, 3, 0, 0) L	17,807,919
2	(4, 3, 0, 0) T	898,689
3	(4, 4, 1, 0) O	26,224,574
4	(4, 4, 1, 0) P	16,421,773
5	(4, 5, 2, 0)	24,350,101
6	(4, 6, 3, 0)	1,797,721
7	(4, 4, 2, 1)	3,266,345
8	(4, 5, 3, 1)	2,697,783
9	(4, 6, 4, 1)	6,535,095

Other samples gave approximately the same numbers, with a maximum difference of 3000.

For three random points on the circle, it is a simple exercise to see that chances are 3 out of 4 that one gets an obtuse triangle. These results indicate that for four random points on the unit sphere the chances are 7 out of 8 that the center of the unit sphere is a local maximum for the distance function. Indeed, the last three entries of the above table add up to 12,499,223. Doubtlessly, a more general statement can be proven here.

4. Notes and remarks.

4.1. Discrete Morse functions. A natural question is the existence of a discrete Morse function (in the sense of Forman) [For02] on the tetrahedron, which realizes the same effect as the minimal distance function. The question is discussed extensively in [vMS05]. It turns out that the corresponding discrete function contains more information than the Morse poset. It also includes, e.g., information about the critical values. Taking into account the critical values, one can make a much finer classification, which distinguishes, e.g., in the (4, 6, 4, 1) case, between an almost regular tetrahedron and one with a small acute base triangle and a top vertex far away. Another related concept is the alpha shape as discussed in, e.g., [EM94], which has applications in molecular shape analysis.

4.2. Higher dimensional results. We have not been able to prove a classification theorem in $\mathbb{R}^n$. However, upon request the authors will send interested readers a computer program that calculates the list in higher dimensions by just trying a lot of random point sets.

Except for the results in [Sie99] on four points in the plane we have no results on the number of Morse posets for N points when $N > n + 1$.

4.3. Edelsbrunner ratio.

DEFINITION 4.1. *The Edelsbrunner ratio ρ is the radius R of the circumsphere of $\mathbb{T}$ divided by the minimal edge length $\min_{i \neq j} d(P_i, P_j)$.*

The ratio is used by Edelsbrunner (see [Ede01, section 6.2]) to classify tetrahedra into “shape types.” This article has the same objective, so it is worthwhile to compare his criterion to ours.

For each of the nine generic types of tetrahedra the ratio ρ is bounded from below by the values in the table below. The infimum corresponds to the quadruple of points in the third column. These quadruples are not generic tetrahedra, except for the case $(4, 6, 4, 1)$, which corresponds to the global minimum of ρ .

Type	Ratio ρ	Infimum
$(4, 3, 0, 0) L$	$\frac{1}{2}\sqrt{3}$	$(0,0,0), (1,0,0), (1,1,0), (1,1,1)$
$(4, 3, 0, 0) T$	$\frac{1}{2}\sqrt{3}$	$(1,0,0), (0,1,0), (0,0,1), (0,0,0)$
$(4, 4, 1, 0) P$	$\sqrt{\frac{7}{12}}$	$(\frac{1}{2}\sqrt{3}, -\frac{1}{2}, 0), (-\frac{1}{2}\sqrt{3}, -\frac{1}{2}), (0,1,0), (0,1,\sqrt{3})$
$(4, 4, 1, 0) O$	$\frac{1}{2}\sqrt{2}$	$(0,0,0), (1,0,0), (1,1,0), (0,1,0)$
$(4, 5, 2, 0)$	$\frac{1}{2}\sqrt{2}$	$(0,0,0), (1,0,0), (1,1,0), (0,1,0)$
$(4, 6, 3, 0)$	$\frac{1}{2}\sqrt{2}$	$(\cos \alpha_j, \sin \alpha_j, 0) \ j = 1, \dots, 3, (0,0,1)$
$(4, 4, 2, 1)$	$\frac{1}{2}\sqrt{2}$	$(1,0,0), (0,1,0), (0,-1,0), (\cos \alpha, 0, \sin \alpha)$
$(4, 5, 3, 1)$	$\frac{1}{2}\sqrt{2}$	$(1,0,0), (0,1,0), (0,-1,0), (0,0,1)$
$(4, 6, 4, 1)$	$\frac{1}{4}\sqrt{6}$	$(0,1,0), (\frac{1}{2}\sqrt{3}, -\frac{1}{2}, 0), (-\frac{1}{2}\sqrt{3}, -\frac{1}{2}, 0), (0,0,\sqrt{2})$

For the case $(4, 6, 3, 0)$ the triangle $P_1P_2P_3$ is acute. For the case $(4, 4, 2, 1)$ the angle α should satisfy $-\frac{\pi}{2} \leq \alpha \leq 0$.

Hence, small values of ρ can be attained by all the types listed in Theorem 3.1. One can see from this that the two classifications are incomparable and conclude that they define different features.

REFERENCES

- [AGJ00] U. ADAMY, J. GIESSEN, AND M. JOHN, *New techniques for topologically correct surface reconstruction*, in Proceedings of the Conference on Visualization '00 (Salt Lake City, UT), IEEE Computer Society, Washington, DC, 2000, pp. 373–380.
- [BKOS97] M. DE BERG, M. VAN KREVELD, M. OVERMARS, AND O. SCHWARZKOPF, *Computational Geometry: Algorithms and Applications*, Springer-Verlag, Heidelberg, 1997.
- [Ede01] H. EDELSBRUNNER, *Geometry and Topology for Mesh Generation*, Cambridge Monographs on Applied and Computational Mathematics 7, Cambridge University Press, Cambridge, UK, 2001.
- [Ede04] H. EDELSBRUNNER, *Surface reconstruction by wrapping finite sets in space*, in Discrete and Computational Geometry: The Goodman-Pollack Festschrift, Springer-Verlag, New York, 2003, pp. 379–404.
- [EM94] H. EDELSBRUNNER AND E. P. MÜCKE, *Three-dimensional alpha shapes*, ACM Trans. Graphics, 13 (1994), pp. 43–72.
- [For02] R. FORMAN, *A user's guide to discrete Morse theory*, Sémin. Lothar. Combin., 48 (2002), Art. B48c.
- [Gal06] M. GALLASSI, J. THEILER, AND J. DAVIES, *Gnu Scientific Library Reference Manual*, 2nd ed., Network Theory Limited, Bristol, UK, 2006.
- [Mil63] J. MILNOR, *Morse Theory*, Princeton University Press, Princeton, NJ, 1963.
- [Mor59] M. MORSE, *Topologically non-degenerate functions on a compact n-manifold M*, J. Anal. Math., 7 (1959), pp. 189–208.
- [Mor73] M. MORSE, *F-deformations and F-tractions*, Proc. Natl. Acad. Sci. USA, 70 (1973), pp. 1634–1635.
- [Sie99] D. SIERSMA, *Voronoi diagrams and Morse theory of the distance function*, in Geometry in Present Day Science, World Scientific, Singapore, 1999, pp. 187–208.
- [vMS05] M. VAN MANEN AND D. SIERSMA, *Power diagrams and their applications*, <http://arxiv.org/abs/math/0508037> (2005).

WALKERS ON THE CYCLE AND THE GRID*

J. DÍAZ[†], X. PÉREZ[†], M. J. SERNA[†], AND N. C. WORMALD[‡]

Abstract. We present a model of the establishment and maintenance of communication between mobile agents. We assume that the agents move through a fixed environment modeled by a motion graph and are able to communicate if they are within distance at most d of each other. As the agents move randomly, we analyze the evolution in time of the connectivity between a set of w agents, asymptotically for a large number N of vertices, when w also grows large. The particular topologies of the environment we study here are the cycle and the toroidal grid.

Key words. mobile agents, ad hoc network, connectivity, distance graph

AMS subject classifications. 68R10, 60C05

DOI. 10.1137/060659855

1. Introduction. Consider a setting in which a large number of mobile agents can perform concurrent basic movements: either ahead/behind/left/right, determining a grid pattern, or left/right, describing a line. Each agent can communicate directly with any other agents which are within a given distance d . This enables communications with agents at a further distance using several intermediate agents. At each step in time there is an ad hoc network defined by the dynamic graph whose vertex set consists of the agents, with an edge between any two agents iff they are within distance d of each other.

Various aspects of such networks (connectivity, transport capacity, protocols, etc.) have been studied in the static case, i.e., the case in which the agents are randomly placed but fixed (see, for example, [XK06, H06] and references therein). Also, there has been quite a bit of experimental study dealing with the dynamic situation, i.e., the situation in which the agents are mobile: connections in the network are created and destroyed as the agents move further apart or closer together; see, for example, [AGE02, JBAS03, RMM01]. The paper [GHSZ] also deals with the problem of maintaining connectivity of mobile agents communicating by radio frequency, but from a perspective orthogonal to the one in the present paper. It describes a *kinetic data structure* to maintain the connected components of the union of unit-radius disks moving in the plane.

To the best of our knowledge, ours is the first work in which the dynamic features of such a network are studied theoretically. Moreover, we obtain much sharper results on the static properties than previously obtained for this type of network. In the static case, our results bear a strong resemblance to analogous results for similar problems of random points in continuous spaces (see Holst [H80] and Penrose [P97], [P99]). The

*Received by the editors May 14, 2006; accepted for publication (in revised form) December 20, 2007; published electronically April 18, 2008. An extended abstract of a preliminary version of this paper appeared in STACS [DPSW05].

<http://www.siam.org/journals/sidma/22-2/65985.html>

[†]Departament de Llenguatges i Sistemes, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain (diaz@lsi.upc.edu, xperez@lsi.upc.edu, mjserna@lsi.upc.edu). This research was supported by the EU 6th FP under contract 001907 (DELIS) and by the Spanish CICYT project TIC-2001-4917-E. The first author was also supported by the Distinció per a la Recerca de la Generalitat de Catalunya.

[‡]Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON N2L 3G1 Canada (nickwor@math.uwaterloo.ca). This author was supported by the Canada Research Chairs program and NSERC.

study of the behavior of multiple, simultaneous random walks has its own merits and could have further applications in computer science.

We propose what we call the *walkers model*, defined as follows. A connected graph $G = (V, E)$ with $|V| = N$ is given, as well as a number w of *walkers* (agents). Also given is a “distance” d . A set W of walkers, with $|W| = w$, is placed randomly and independently on the vertices of G (a vertex may contain more than one walker). Each walker has a range d for communication; that is, two walkers w_1 and w_2 can communicate in one hop if the distance, in G , between the position of the walkers is at most d . Two walkers can communicate if they can reach each other by a sequence of such hops. In addition, each walker takes an independent standard random walk on G , i.e., moves at each time step to a neighboring vertex, each neighbor chosen with equal probability.

The interesting features of the walkers model are encapsulated by the *random graph of walkers*, $\mathcal{W}(G, w, d)$. This random graph is generated using a random assignment $f : W \rightarrow V$ of walkers into the vertices of G . The vertices of $\mathcal{W}(G, w, d)$ are the vertices in G that receive at least one walker, two vertices in $\mathcal{W}(G, w, d)$ being joined by an edge iff they are at distance at most d in G . We are interested in the probability of $\mathcal{W}(G, w, d)$ being connected, or in the number of components and their sizes, with certain mild asymptotic restrictions on w and d .

Our primary goal with the walkers model is to characterize the dynamics of the connectivity of the communication network as represented by the random graph process $(\mathcal{W}_t(G, w, d))_{t \geq 0}$. Here $\mathcal{W}_t(G, w, d)$ denotes the graph of walkers at time $t = 0, 1, \dots$, and, at each time step, each walker simultaneously moves one step to a randomly selected neighbor vertex in G . In order to study this model, we first examine $\mathcal{W}(G, w, d)$, which we call the *static* model. This is a snapshot of the process at one point in time: we are interested in the distribution of the number of components, as well as some other information which helps to answer the dynamic questions. In particular, we are interested in studying the birth and death of components and the sudden connection and disconnection of $\mathcal{W}_t(G, w, d)$. We generally abbreviate $\mathcal{W}(G, w, d)$ to $\mathcal{W}(G)$ when w and d are understood and, similarly, $\mathcal{W}_t(G, w, d)$ to $\mathcal{W}_t(G)$.

We consider a sequence of graphs G with increasing numbers of vertices N for N tending to infinity. The parameters w and d are functions of N . We restrict our attention to the case $w \rightarrow \infty$ in order to avoid considering small-case effects. Of course we take $d \geq 1$. We make further restrictions on w and d in order to rule out noninteresting cases, such as values of the parameters in which the network is a.a.s. disconnected or a.a.s. connected. (Throughout this paper, a.a.s. will abbreviate “asymptotically almost surely,” which denotes “with probability tending to 1 as $N \rightarrow \infty$.”) We study the walkers model for two particular sequences of graphs G : the cycle C_N of length N and the $n \times n$ toroidal grid T_N of size $N = n^2$. (In the case of the grid, we use the ℓ^p distance for any $1 \leq p \leq \infty$.) The two cases have an essential difference that prevents a unified treatment: for the interesting values of w and d , disconnectedness of the graph of walkers for the cycle is basically caused by the presence of at least two large “gaps” between the walkers around the cycle, whereas for the grid, it is caused by the presence of one or more isolated walkers.

We now give a description of some of our main results. Throughout this paper, ϱ denotes w/N , which is the expected number of walkers at a vertex.

1.1. The cycle. First we consider the cycle C_N on N vertices. We will show that if $d = \Omega(N)$, then $\mathcal{W}(C_N)$ is a.a.s. connected, so all results here refer to $d = o(N)$.

To study connectedness in this case, we introduce the concept of a *hole*. There is a *hole* between two vertices u and v if u and v each contain at least one walker, but no vertex in the clockwise path from u to v contains a walker. We say that such a hole *follows* u or that u is the *start vertex* of the hole. The number of internal vertices in a hole is its *exact size*. A k -*hole* is a hole whose exact size is at least k .

Let H be the random variable counting the number of d -holes in the walkers model for C_N . Notice that at least two d -holes are needed to disconnect the walkers on C_N . To be precise,

$$(1) \quad \mathcal{W}(C_N) \text{ is connected iff } H \leq 1.$$

We define a new parameter $\mu = N(1 - e^{-\varrho})e^{-d\varrho}$, which plays a key role in characterizing the connectedness of $\mathcal{W}(C_N)$. It is straightforward to check that

$$\mu \sim \begin{cases} we^{-d\varrho} & \text{if } \varrho \rightarrow 0, \\ N(1 - e^{-\varrho})e^{-d\varrho} & \text{if } \varrho \rightarrow c, \\ Ne^{-d\varrho} & \text{if } \varrho \rightarrow \infty. \end{cases}$$

Here, and throughout the paper, asymptotics refer to $N \rightarrow \infty$ unless otherwise stated.

Regarding the behavior of H , and connectedness, we have the following.

THEOREM 1.1. *The expected number of holes satisfies $\mathbf{E}[H] \sim N(1 - e^{-\varrho})(1 - d/N)^w$. Furthermore,*

- (i) *if $\mu \rightarrow 0$, then $\mathcal{W}(C_N)$ is connected a.a.s.,*
- (ii) *if $\mu \rightarrow \infty$, then $\mathcal{W}(C_N)$ is disconnected a.a.s., and*
- (iii) *if $\mu = \Theta(1)$, then H is asymptotically Poisson with mean μ .*

The following corollary gives the asymptotic probability that $\mathcal{W}(C_N)$ is connected. It follows immediately from the theorem in view of (1).

COROLLARY 1.2. $\mathbf{P}(\mathcal{W}(C_N) \text{ is connected}) = e^{-\mu}(1 + \mu) + o(1)$.

We now turn to dynamic properties of the random graph of walkers on the cycle. One of our main end results concerns the expected time that the graph of walkers remains disconnected after a point in time at which it becomes disconnected. Define a *disconnected period* to be a maximal sequence of consecutive time steps for which the graph of walkers $\mathcal{W}_t(C_N)$ is disconnected. We define for $t > 0$

$$LD_t = \begin{cases} \max\{k \in \mathbb{N} \mid \mathcal{W}_t(C_N), \dots, \mathcal{W}_{t+k-1}(C_N) \text{ disconnected}\} & \text{if } \mathcal{W}_{t-1}(C_N) \\ & \text{is connected,} \\ 0 & \text{otherwise.} \end{cases}$$

This is the random variable counting the length of the disconnected period starting at time step t provided that it really starts then. Formally we define the *average length* of a disconnected period starting at time t to be

$$LD_{\text{av}} := \mathbf{E}[LD_t \mid LD_t > 0].$$

By symmetry, this is independent of t and so is a function of N , d , and w . The next result finds its size.

THEOREM 1.3. *For the walkers model on the cycle C_N , the average length of a disconnected period of $(\mathcal{W}_t(C_N))_{t \geq 0}$ satisfies*

$$LD_{\text{av}} \sim \begin{cases} 2 \frac{e^\mu - 1 - \mu}{\mu^2} \varrho^{-1} & \text{if } \varrho \rightarrow 0, \\ \frac{e^\mu - 1 - \mu}{1 + \mu - (1 + \mu + \lambda + \lambda^2)e^{-\lambda}} & \text{if } \varrho \rightarrow c, \\ \frac{e^\mu}{1 + \mu} & \text{if } \varrho \rightarrow \infty, \end{cases}$$

where $\lambda = (1 - \frac{3e^{-e} - e^{-\frac{3}{2}e}}{1 + e^{-\frac{1}{2}e}})\mu$. Here $0 < \lambda < \mu$ for $\varrho \rightarrow c$.

To convey a feeling of the complexity of the question of how long a network remains disconnected once it becomes so, we introduce the following *train paradox*. A visitor to Barcelona goes every morning to Catalunya station, where there are two metro lines meeting (red and green), though at different levels. He wishes to measure the average length of a train. Since there is plenty to see in Barcelona, each morning he chooses either the red line or the green line at random. He waits for the first train to leave on that line and records its length. He finds after many days that the average length recorded is 9 cars. But he notices that, restricted to the days that the train is already at the platform when he arrives, the average length is only 8 cars.

Could it be that the shorter trains wait longer for him? No, because the trains in Barcelona stop at stations for equal times. Moreover, on any given line they arrive regularly at equally spaced intervals, so the well-known bus paradox does not directly apply.

Which is a better measure—the length of the first train to arrive or the length of trains arriving at a prescribed time? The former, yielding the answer 9, might seem more natural at first. However, the explanation for the differing answers reveals the other to be meaningful, and perhaps even more so. The data above, in both versions of the paradox, arise if the red line has trains of average length 12 arriving every 10 minutes, and the green line has trains of average length 6 arriving at 5 minute intervals. In an extended time period, recording all the lengths of trains on all lines will yield 8 as the average.

Returning to the walkers model, as we shall see in section 3, when N is even there are many different configurations of walkers that cannot arise from a given initial configuration. As a Markov chain, the process is not ergodic; each configuration belongs to a class of mutually reachable configurations. The different classes of states correspond to different metro lines in the train paradox. The quantity LD_{av} in Theorem 1.3 is roughly equivalent to our traveler's measurement of the length of trains restricted to those days that a train is just arriving. However, the train paradox shows that this is not the only reasonable measure of average length. Moreover, the situation is even more complicated, as the train paradox would be if the intertrain time periods were variable on a given line. The average length of the first train to arrive would then be affected by any dependence between the length of a train and the time before the previous train. We wish to study the analogue of the average length of trains on a given metro line—the average length of a period of disconnection given the initial state. If N is divisible by 2 and the initial configuration is conditioned upon, the walkers process is “locked in” to a future in which the (conditional) average length of connected periods may be different from LD_{av} . However, we show that for almost all initial configurations this average is essentially asymptotically equal to LD_{av} . For $T \in \mathbb{Z}^+$, we define the *average disconnection time* of the graph of walkers in $[1, T]$ to be

$$LD_T = \frac{\sum_{t=1}^T LD_t}{|\{t \in \{1, \dots, T\} : LD_t > 0\}|}.$$

We show that LD_T converges in probability as $T \rightarrow \infty$ to a random variable which may depend on the class of the initial state but nothing else. (Actually, the value is in general different for different classes.) The notation $f \sim g$ a.a.s. used in the following theorem denotes that for all $\epsilon > 0$, a.a.s. $|f/g - 1| < \epsilon$ (see, for example, [W04]).

THEOREM 1.4. *For the walkers model on C_N , LD_T converges in probability as*

$T \rightarrow \infty$ (with N fixed) toward a random variable LD , uniquely determined almost everywhere. Furthermore, (as $N \rightarrow \infty$) we have $LD \sim LD_{\text{av}}$ a.a.s.

Intuitively, LD is the average length of the disconnected periods appearing in all the trajectories of the random process starting from the initial state.

In section 3.2 we give analogous results for maximal connected periods of $\mathcal{W}(C_N)$ (see Theorems 3.10 and 3.11).

During our examination of the cycle, we derive an incidental result (Corollary 3.7) on the expected time required before a certain infinite set of random walkers sprinkled on the nonnegative integers meets another such set sprinkled on the negative integers.

1.2. The grid. We turn now to the toroidal grid T_N with $N = n^2$ vertices, for which our results apply with any normed ℓ^p distance, for $1 \leq p \leq \infty$. We will show that if $d = \Omega(n)$, then $\mathcal{W}(T_N)$ is a.a.s. connected, so all results here refer to $d = o(n)$. We define h to be the number of grid points within distance d of any fixed point in T_N . This depends on the particular metric chosen, but in all cases it is easy to see that $h = \Theta(d^2)$.

We call a component of $\mathcal{W}(T_N)$ *simple* if it is formed by only one isolated vertex. Let X be the number of simple components. To examine the connectedness of $\mathcal{W}(T_N)$, we need to redefine the μ used for the cycle. For the following grid results, $\mu = N(1 - e^{-\varrho})e^{-h\varrho}$. Hence

$$\mu \sim \begin{cases} we^{-h\varrho} & \text{if } \varrho \rightarrow 0, \\ N(1 - e^{-\varrho})e^{-h\varrho} & \text{if } \varrho \rightarrow c, \\ Ne^{-h\varrho} & \text{if } \varrho \rightarrow \infty. \end{cases}$$

THEOREM 1.5.

- (i) For $\mu \rightarrow \infty$, $\mathcal{W}(T_N)$ is disconnected a.a.s.
- (ii) For $\mu = \Theta(1)$, a.a.s. all but one component of $\mathcal{W}(T_N)$ are simple, and the number X of simple components is asymptotically Poisson with expected value μ .
- (iii) For $\mu \rightarrow 0$, $\mathcal{W}(T_N)$ is connected a.a.s.

The theorem immediately gives the following.

COROLLARY 1.6. $\mathbf{P}(\mathcal{W}(T_N) \text{ is connected}) = e^{-\mu} + o(1)$.

For dynamic considerations, we also need a function $b = \Theta(d)$ which is related to the boundary of a ball of radius d in the grid. Like h , this function also depends on the metric. It is defined in section 4.2.

We define LD_{av} and LD_T for disconnected periods as for the cycle case above.

THEOREM 1.7. For the walkers model on the grid T_N with $d > 1$,

$$LD_{\text{av}} \sim \begin{cases} \frac{e^\mu - 1}{\mu b \varrho} & \text{if } d\varrho \rightarrow 0, \\ \frac{e^\mu - 1}{1 - e^{-\lambda}} & \text{if } d\varrho \rightarrow c, \\ e^\mu & \text{if } d\varrho \rightarrow \infty, \end{cases}$$

where $\lambda = (1 - e^{-b\varrho})\mu$. Here $0 < \lambda < \mu$ for $d\varrho \rightarrow c$. Furthermore, LD_T converges in probability for T growing large (N fixed) toward a random variable LD , uniquely determined almost everywhere, where $LD \sim LD_{\text{av}}$ a.a.s.

The case $d = 1$ is excluded for this result because of a technical complication that will be explained in the proofs, though our basic approach could be modified to deal with this case also. We also obtain exactly the same result as in Theorem 1.7 for connected periods as opposed to disconnected periods.

The remainder of the paper gives proofs of these theorems as well as stating and proving related ones. In section 2 we give basic definitions and technical lemmas to be used throughout the paper. In section 3 we deal with the cycle C_N , and in section 4 the toroidal grid T_N . One of the main differences between this case and the cycle is the need for a geometric lemma that may be of independent interest (Lemma 4.3). This bounds from below the number of nonoccupied vertices at distance at most d from the boundary of any connected component in $\mathcal{W}(T_N)$, as a function of a measure of the length of the boundary. The last section contains some discussion and related problems.

2. General definitions and basic results. We begin with some definitions and results which are common for all G . We call a component *simple* if it is formed by only one isolated vertex. Recall that $\varrho = w/N$ is the expected number of walkers at a vertex. For any $v \in V$, define h_v to be the number of vertices in G at distance at most d from v , and define $h = \min_{v \in V} h_v$. We say that a vertex or set of vertices is *empty of walkers* (*e.o.w.*), or simply *empty*, if it contains no walkers, and *occupied* otherwise. Note that there must be h empty vertices in G within distance d of a simple component.

By considering the coupon collector's problem, we observe that if $w = N \log N + \omega(N)$, then $\mathcal{W}(G, w, d)$ is trivially a.s. connected due to every vertex being occupied. Moreover, for the graphs G which we consider in this paper, if $h \in \Omega(N/\sqrt{w})$, then $\mathcal{W}(G, w, d)$ is a.s. connected as well. This last claim will be seen in Observations 3.1 and 4.1. Thus, we consider throughout the paper $w < N \log N + O(N)$ and $h = o(N/\sqrt{w})$. In fact, our proofs will just assume h to be $o(N)$. Note that, for the cycle, $h = 2d$.

We will need to compute the probability of certain configurations of walkers at two consecutive time steps t and $t+1$ in order to record the event that walkers jump to the appropriate place at that step. There is a convenient way to formulate this in terms of occupancy of directed edges. Let us regard G as a directed graph by considering each edge as a symmetric pair of directed edges. For any directed edge e going from vertex u to v , we say that a walker is on e between time steps t and $t+1$ if the walker is on u at time step t and jumps onto v within one step. Note that in this way we can encode dynamic transitions in terms of static configurations of walkers.

There is an alternative formulation in terms of cells. Each vertex is divided into as many cells as its degree, and each cell is associated with one of the directed edges stemming from the vertex. Then, the transition of the system between two consecutive time steps can be described by the placement of the walkers in the cells (see Figure 1). We use this representation particularly in our figures for the sake of simplicity and visual clarity.

Assign *size* 1 to all vertices in G . For a given directed edge stemming from a vertex v with degree δ_v , its *size* will be $1/\delta_v$. Given a set A of vertices or directed edges, we define the *size* of A to be the sum of the sizes of its elements.

The following technical lemma is used in most of our proofs to compute the probability of having a certain configuration of walkers (in terms of vertex or directed edge occupancy) at a given time step. The lemma applies to any graph G . Its proof, using an inclusion-exclusion argument, is fairly straightforward and so is omitted. (See [DPSW06] for details.)

LEMMA 2.1. *Let $A_0, \dots, A_m$ be pairwise disjoint sets of vertices (or directed edges) in G , with sizes $S_0, \dots, S_m$, respectively. Let $N = |V(G)|$. If $\sum_{i=0}^m S_i = o(N)$,*

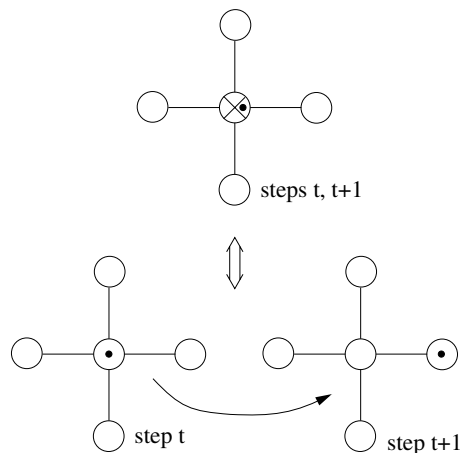


FIG. 1. The walker jumps to a neighbor according to which cell it is placed on.

then

$$\mathbf{P} \left(A_0 \text{ e.o.w.} \wedge \bigwedge_{i=1}^m (A_i \text{ not e.o.w.}) \right) \sim \left(1 - \frac{S_0}{N} \right)^w \prod_{i=1}^m (1 - e^{-S_i \varrho}).$$

To cover large sizes S , not necessarily $o(N)$, we need the following variation.

LEMMA 2.2. *Let A be a set of vertices in G of size S , and $v_1, \dots, v_m$ vertices not in A , with $m \geq 1$. Assume $N - S \rightarrow \infty$. The probability that no vertex in A is occupied and $v_1, \dots, v_m$ are all occupied is at most $p_0 p^{m-1} \alpha^w (1 + o(1))$, where $p_0 = 1 - e^{-\varrho/\alpha}$, $\alpha = 1 - S/N$, and $p = \min(1, \varrho/\alpha)$. Here the asymptotics are uniform over all m .*

Proof. The probability of the event $\mathcal{E}$ that all walkers avoid A is α^w . The probability that v_1 is occupied conditional upon $\mathcal{E}$ is $1 - (1 - (N - S)^{-1})^w$, which is asymptotic to p_0 . The lemma follows for $\varrho > \alpha$, i.e., $w > N - S$. Otherwise, conditional upon $\mathcal{E}$ and the event that $v_1, \dots, v_i$ are occupied, the probability that the next is occupied is clearly decreasing with i and is thus at most $w/(N - S) = \varrho/\alpha$. The lemma follows. $\square$

3. The cycle. Here $G = C_N$, the cycle with N vertices.

OBSERVATION 3.1. *Cover C_N with $\lceil \frac{N}{\lceil d/2 \rceil} \rceil$ paths of $\lceil d/2 \rceil$ vertices. If $d = \Omega(N/\sqrt{w})$, then the probability that some path is e.o.w. is at most*

$$\left\lceil \frac{N}{\lceil d/2 \rceil} \right\rceil \left(1 - \frac{\lceil d/2 \rceil}{N} \right)^w \leq O(\sqrt{w}) e^{-\Omega(\sqrt{w})} \rightarrow 0.$$

Thus, a.a.s. each of these paths is occupied (by at least one walker), and $\mathcal{W}(G, w, d)$ is connected.

In view of this observation, we assume for the rest of the section that $d = o(N)$. If $d = \Omega(N)$, then $\mathcal{W}(G, w, d)$ is a.a.s. connected.

Recall the definition of a hole from section 1 and that H is the random variable counting the number of d -holes. For this section, μ is defined as for Theorem 1.1.

3.1. Static properties. Here, we study the connectedness of the graph of walkers $\mathcal{W}(G, w, d)$ in the static situation by analyzing the behavior of H . In view of (1), if $\mathbf{E}[H] \rightarrow 0$, then $\mathcal{W}(G, w, d)$ is a.a.s. connected. Theorem 1.1 gives the asymptotic distribution of H .

Proof of Theorem 1.1. For any vertex v in C_N , let H_v be an indicator random variable such that $H_v = 1$ iff there is a d -hole following vertex v . Then,

$$(2) \quad H = \sum_{v \in V} H_v \quad \text{and} \quad \mathbf{E}[H] = \sum_{v \in V} \mathbf{P}(H_v = 1).$$

We shall compute the k th factorial moment:

$$(3) \quad \mathbf{E}[H]_k = \sum_{v_1 \neq \dots \neq v_k \in V} \mathbf{P}(H_{v_1} = 1 \wedge \dots \wedge H_{v_k} = 1).$$

Let $\mathcal{S}$ denote the set of tuples such that each v_i and v_j , $i \neq j$, have distance at least $d + 1$ around the cycle. For $(v_1, \dots, v_k) \notin \mathcal{S}$, the probability in (3) is 0 since one of the v_i “lies in” the hole following some v_j , and yet v_i must be occupied. For $(v_1, \dots, v_k) \in \mathcal{S}$, the probability in (3) is easily computed by applying Lemma 2.1:

$$\mathbf{P}(H_{v_1} = 1 \wedge \dots \wedge H_{v_k} = 1) \sim \left(1 - \frac{kd}{N}\right)^w (1 - e^{-\varrho})^k.$$

Since $d = o(N)$, we have $|\mathcal{S}| \sim N^k$, and thus from (3)

$$(4) \quad \mathbf{E}[H]_k \sim \left[N (1 - e^{-\varrho}) e^{-d\varrho + O(\frac{d^2 w}{N^2})} \right]^k.$$

In particular,

$$(5) \quad \begin{aligned} \mathbf{E}[H] &\sim N (1 - e^{-\varrho}) \left(1 - \frac{d}{N}\right)^w \\ &\sim N (1 - e^{-\varrho}) e^{-d\varrho + O(\frac{d^2 w}{N^2})}. \end{aligned}$$

In the case $\mu \rightarrow 0$, we have also $\mathbf{E}[H] \rightarrow 0$ since $(1 - d/N)^w \leq e^{-d\varrho}$. Then $\mathbf{P}(H = 0) \rightarrow 1$, and $\mathcal{W}(G, w, d)$ is connected a.a.s.. In the case that μ is bounded away from 0, taking logarithms,

$$(6) \quad d\varrho = \log N (1 - e^{-\varrho}) - \log \mu.$$

Considering separately the cases when $\varrho \rightarrow 0$ and when $\varrho \rightarrow c, \infty$, we get from (6) that $\frac{d^2 w}{N^2} = o(1)$, so we can ignore the term $O(\frac{d^2 w}{N^2})$ in the expression (4) and obtain

$$(7) \quad \mathbf{E}[H]_k \sim [N (1 - e^{-\varrho}) e^{-d\varrho}]^k = \mu^k.$$

Moreover, if μ is bounded, then it follows, from (7) and (for instance) from Theorem 1.23 in [Bol02], that H is asymptotically Poisson. For $\mu \rightarrow \infty$, we have that $\mathbf{E}[H]_2 \sim \mu^2$, and so it follows from Chebyshev’s inequality that a.a.s. $H > \mu/2$ (and $\mathcal{W}(G, w, d)$ is disconnected). $\square$

3.2. Dynamic properties. Assume that from an initial random placement f of the walkers, at each step, every walker moves from its current position to one of its neighbors, with probability $1/2$ of going either way. This is a standard random walk on the cycle for each walker. To study the connectivity properties of the dynamic graph of walkers we need to introduce some notation.

A *configuration* or *state* is an arrangement of the w walkers on the vertices of C_N . Consider the graph of configurations, where the vertices are the N^w different configurations. Each configuration can be represented by a vector $\mathbf{a} = (a_1, \dots, a_w) \in (\mathbb{Z}_N)^w$, where a_i indicates the label of the vertex being occupied by walker i . Given a configuration $\mathbf{a}$, there exists an edge between $\mathbf{a}$ and all configurations $(a_1 \pm 1, \dots, a_w \pm 1)$. Thus, any configuration has 2^w neighbors, and the relationship of being neighbors is symmetric. The dynamic process can be viewed as a random walk on the graph of configurations, in particular, a Markov chain $\mathcal{M}(N, w, d)$. We denote by $\mathcal{M}_t$ the state of the process at time step t .

For N even, given any two configurations $\mathbf{a}$ and $\mathbf{b}$, we say that they have the *same parity* if for all i and j , $a_i - a_j \equiv b_i - b_j \pmod{2}$. There are 2^{w-1} different parities. Note that the initial parity stays invariant during the dynamic process. The following lemma is straightforward, and the proof is left to the reader.

LEMMA 3.2. *Let $\mathbf{a}$ and $\mathbf{b}$ be any two configurations, and let $h_{\mathbf{a}, \mathbf{b}}$ denote the hitting time from $\mathbf{a}$ to $\mathbf{b}$. If N is odd, then $\mathbf{b}$ is reachable from $\mathbf{a}$ and $h_{\mathbf{a}, \mathbf{b}}$ is finite for any $\mathbf{a}$ and $\mathbf{b}$. If N is even, then $\mathbf{b}$ is reachable from $\mathbf{a}$ provided that $\mathbf{a}$ and $\mathbf{b}$ have the same parity, and in this case $h_{\mathbf{a}, \mathbf{b}}$ is finite.*

Then, if N is odd, $\mathcal{M}(N, w, d)$ consists of a single closed class of states, so it is irreducible and positive recurrent. It is trivial to verify aperiodicity, and hence we have that the chain is ergodic. However, if N is even, there are 2^{w-1} closed classes of states, where each class consists of all configurations with the same parity. We denote by $\mathcal{A}$ the set of classes. Let $A \in \mathcal{A}$ be any class of states, and let $\mathbf{a} \in A$ be a configuration. For this particular configuration, we can partition the set of walkers $W = W_1 \cup W_2$ so that the ones in W_1 lie in odd positions of the cycle and the ones in W_2 lie in even positions. Let A_1 be the set of all states which lead to this same partition, and A_2 the set of those which lead to the complementary one. Clearly, $A = A_1 \cup A_2$. Those states in A_1 are only reachable by an even number of steps from $\mathbf{a}$, and those in A_2 by an odd number of steps. Hence, if we restrict the Markov chain to any class of states, it is irreducible and positive recurrent but 2-periodic and hence not ergodic.

OBSERVATION 3.3. *Note that for any fixed t , the distribution of $\mathcal{W}_t(C_N)$ is just that of $\mathcal{W}(C_N)$ in the static case. That is, the initial uniform distribution stays invariant, even though when N is even the chain is not ergodic and there is no limit distribution. Hence, by Theorem 1.1, if $\mu \rightarrow 0$ or ∞ , then for any fixed t , $\mathcal{W}_t(C_N)$ is a.a.s. connected or a.a.s. disconnected, respectively.*

In view of this observation, we assume that $\mu = \Theta(1)$ for the remainder of the section, since we wish to study only the nontrivial dynamic situations. One thing this implies, in view of the definition of μ , is that

$$(8) \quad d\varrho \rightarrow \infty.$$

We define $H(t)$ to be the random variable that counts the number of d -holes at time step t . Then from section 3.1, $H(t)$ is asymptotically Poisson with expectation $\mu = \Theta(1)$.

For the dynamic properties of $\mathcal{W}_t(C_N)$, we are interested in the probability that

a new d -hole appears at a given time step. Moreover, we require knowledge of this probability conditional upon the number of d -holes already existing.

If there is a d -hole from u to v at time step t and all walkers at u and v move in the same direction on the next step, a new d -hole may appear following one of the neighbors of u (provided no new walkers move in to destroy this). These two d -holes, though different, are related, and the presence of the first makes the second more likely to occur. Similarly, the exact size of a d -hole following u may change in one step, making it technically a different d -hole but, again, related. In all these cases, the start vertex of the d -hole “moves” by at most 1; a d -hole at time step $t+1$ which does not follow u or a neighbor of u is not related to a d -hole following u at time step t . We need some definitions to make this loose description rigorous. Define a *d -hole line* to be a maximal sequence of pairs $(h_1, t_1), \dots, (h_l, t_l)$, where h_i is a d -hole existing at time step t_i for $1 \leq i \leq l$, and such that $t_i = t_{i-1} + 1$ and the start vertex of h_i is adjacent to, or equal to, the start vertex of h_{i-1} for $2 \leq i \leq l$. Fix two consecutive time steps t and $t+1$. If $t_1 = t+1$, we say that the line is *born* between t and $t+1$; if $t_l = t$, the line *dies* between t and $t+1$; and if $t = t_i$, $i \in \{1, \dots, l-1\}$, we say that the line *survives* during the interval $[t, t+1]$. Note that the time-reversal of the process has a d -hole line born at vertex u between $t+1$ and t iff the d -hole line dies at u between t and $t+1$.

We define random variables $S(t)$, $B(t)$, and $D(t)$ to be the number of d -hole lines surviving, being born, and dying between t and $t+1$. We have $D(t) + S(t) = H(t)$ and $B(t) + S(t) = H(t+1)$.

The proof of the following result is similar to that of Theorem 1.1 but rather more complicated, so we give only an abbreviated proof here. A full proof may be found in [DPSW06] or [P07]. Here “asymptotically independent Poisson” means that the joint distribution tends to that of independent Poisson variables.

PROPOSITION 3.4. *For any fixed t , the random variables $S(t)$, $B(t)$, and $D(t)$ are asymptotically independent Poisson, with the expectations*

$$\mathbf{E}[S(t)] \sim \begin{cases} \mu & \text{if } \varrho \rightarrow 0, \\ \mu - \lambda & \text{if } \varrho \rightarrow c, \\ 3\mu e^{-\varrho} & \text{if } \varrho \rightarrow \infty, \end{cases} \quad \text{and} \quad \mathbf{E}[B(t)] = \mathbf{E}[D(t)] \sim \begin{cases} \frac{1}{2}\mu\varrho & \text{if } \varrho \rightarrow 0, \\ \lambda & \text{if } \varrho \rightarrow c, \\ \mu & \text{if } \varrho \rightarrow \infty, \end{cases}$$

where λ is defined in Theorem 1.3.

Proof. We will estimate factorial moments of the appropriate variables. In C_N , let right denote “clockwise” and left “counterclockwise.” Moreover, for a vertex $v \in V$ and $i \geq 0$, let $v+i$ (respectively, $v-i$) denote the vertex i positions to the right (respectively, left) from v . All probabilities and events in this proof will involve two consecutive time steps t and $t+1$. We can describe these events from a static point of view, as explained in section 2.

We consider births in detail. There are four ways that a d -hole line can be born at v between time steps t and $t+1$ according to the following descriptions, as shown in Figure 2:

- b1. At time step t , there is a hole between $v+1$ and $v+d$ of exact size $d-2$. Then all walkers at v and $v+1$ move left and all walkers at $v+d$ and $v+d+1$ move right.
- b2. At time step t , there is a hole between $v+1$ and $v+d+1$ of exact size $d-1$. Then all walkers at v and $v+1$ move left and all walkers at $v+d+1$ and $v+d+2$ move right.

- b3. At time step t there is a hole between $v + 1$ and $v + d + 1$ of exact size $d - 1$, and $v + d + 2$ is occupied. Then all walkers at v and $v + 1$ move left, all walkers at $v + d + 1$ move right, and at least one walker at $v + d + 2$ moves left.
- b4. At time step t there is a hole between v and $v + d$ of exact size $d - 1$, and $v - 1$ is occupied. Then all walkers at v move left, all walkers at $v + d$ and $v + d + 1$ move right, and at least one walker at $v - 1$ moves right.

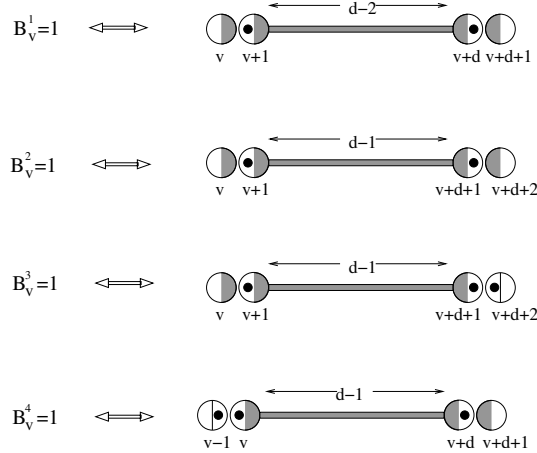


FIG. 2. Birth of a d -hole line at vertex v .

The expected number of d -hole lines born at a given time is N times the sum of the probabilities of these four events. Applying Lemma 2.1, together with the fact that $(1 - d/N)^w \sim e^{-d\varrho}$ (Theorem 1.1), we get (writing the four terms separately)

$$\begin{aligned} \frac{1}{N} \mathbf{E}[B(t)] &\sim \left(1 - e^{-\varrho/2}\right)^2 e^{-d\varrho} + \left(1 - e^{-\varrho/2}\right)^2 e^{-(d+1)\varrho} \\ &\quad + \left(1 - e^{-\varrho/2}\right)^3 e^{-(d+1/2)\varrho} + \left(1 - e^{-\varrho/2}\right)^3 e^{-(d+1/2)\varrho}. \end{aligned}$$

This leads to the expression for $\mathbf{E}[B(t)]$ stated in the theorem. The other expected values are obtained in a similar way (with three types of survival event and four types of death event). Computing the joint factorial moments of random variables counting births, deaths, and survivals gives the claim about Poisson distribution, similar to the proof of Theorem 1.1. We omit the remaining details. $\square$

As $H(t) = S(t) + D(t)$ and $H(t+1) = S(t) + B(t)$, the following is immediate.

COROLLARY 3.5. $H(t)$ and $B(t)$ are asymptotically independent, and so are $D(t)$ and $H(t+1)$.

It is natural to define the *lifespan* of a d -hole line as the number of time steps for which the line is alive. For any vertex v and time step t , the random variable $L_{v,t}$ is the lifespan of the d -hole line born at vertex v between time steps t and $t+1$. If no such birth takes place, $L_{v,t}$ is defined to be 0. Note that the random variables $L_{v,t}$ are identically distributed for all v and t .

In view of (8), it is easy to see that a state can be reached in which there are no holes. One way to do this is to force the walkers to move to positions in which they are almost equally spaced around the cycle. Then, by Lemma 3.2, for any initial state, the process will reach some such state within finite expected time. Thus the

expected lifespan of any given d -hole line, given the configuration of walkers at its birth, is bounded (and the bound is simply a function of N , d , and w).

Considering the train paradox discussed in section 1, we define the *average lifespan* of d -hole lines to be the expected time that a d -hole line will survive once born. Formally,

$$L_{\text{av}} := \mathbf{E}[L_{v,t} \mid L_{v,t} > 0].$$

By symmetry, this is independent of v and t and so is a function of N , d , and w . The next result finds its size.

THEOREM 3.6. *The average lifespan of d -hole lines satisfies*

$$L_{\text{av}} = \frac{\mathbf{E}[H(0)]}{\mathbf{E}[B(0)]} \sim \begin{cases} 2\rho^{-1} & \text{if } \rho \rightarrow 0, \\ \frac{\mu}{\lambda} & \text{if } \rho \rightarrow c, \\ 1 & \text{if } \rho \rightarrow \infty, \end{cases}$$

where λ is defined as in Theorem 1.3.

Proof. For any vertex v and time step t , we have

$$\begin{aligned} N\mathbf{E}[L_{v,t}] - \mathbf{E}[H(0)] &= \sum_{v \in V} \mathbf{E}[L_{v,t}] - \mathbf{E}[H(0)] \\ &= \frac{1}{T} \left(\sum_{v \in V} \sum_{t=0}^{T-1} \mathbf{E}[L_{v,t}] - \sum_{t=0}^{T-1} \mathbf{E}[H(t)] \right) \quad \text{for any } T \in \mathbb{Z}^+ \\ &= \frac{1}{T} \mathbf{E} \left[\sum_{v \in V} \sum_{t=0}^{T-1} L_{v,t} - \sum_{t=0}^{T-1} H(t) \right] \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \mathbf{E} \left[\sum_{v \in V} \sum_{t=0}^{T-1} L_{v,t} - \sum_{t=0}^{T-1} H(t) \right] \\ &= 0, \end{aligned}$$

since $H(t)$ is uniformly bounded (the bound depending on N) and so is the expected lifespan of any line. Thus, the portions of lifespans omitted in $\sum_{t=0}^{T-1} H(t)$ are bounded in expectation.

Now the rest follows easily:

$$\begin{aligned} \mathbf{E}[L_{v,t} \mid L_{v,t} > 0] &= \frac{\mathbf{E}[L_{v,t}]}{\mathbf{P}(L_{v,t} > 0)} \\ &= \frac{\mathbf{E}[H(0)]}{N\mathbf{P}(L_{v,t} > 0)} \\ &= \frac{\mathbf{E}[H(0)]}{\sum_v \mathbf{P}(L_{v,t} > 0)} \\ &= \frac{\mathbf{E}[H(0)]}{\mathbf{E}[B(0)]}. \end{aligned}$$

The final formula for this expression comes immediately from Proposition 3.4, noting that $H(t) = S(t) + D(t) \sim S(t)$ for $\rho \rightarrow 0$ and $D(t)$ dominates for $\rho \rightarrow \infty$. $\square$

The previous theorem has the following consequence of a rather different nature. Although this is entirely incidental to our main goals, it seems interesting enough to

mention. It is not clear how to establish this result without reference to our result on the cycle.

COROLLARY 3.7. *For constant $0 < p < 1$, randomly place walkers on the integers, each independently with probability p , but with probability 1 for integers 0 and 1. Let them all take the usual independent random walks. Then the expected time $g(p)$ before one of the initially nonpositive walkers meets or passes one of the initially positive walkers satisfies $g(p) \sim 2/p$ as $p \rightarrow 0$.*

Proof. We call the model introduced in the statement of the theorem the p -model. For fixed p , let $m = \lceil p^{-1} \rceil$. For fixed $i > 0$, consider the intervals $I_{-i} = [-im, -(i-1)m - 1]$ and $I_i = [(i-1)m + 1, im]$. Let H_i denote the event that in the first $i^2 m^2$ steps, some walker initially in I_i meets some walker initially in I_{-i} . Then $\mathbf{P}(H_i) > c > 0$ for some constant c independent of i (and p). This is because the probability that at least one site in I_i is initially occupied is at least $1 - (1-p)^m$, which is bounded below, and because after $i^2 m^2$ steps a walker's position is binomially distributed, subject to parity and relative to its starting point, with variance $i^2 m^2$, not forgetting that the walkers move independently.

Call a walker *positive* if it is initially positive and *negative* otherwise. For fixed M , the events $H_1, H_2, \dots, H_M$ are mutually independent, so the probability that none of them occurs is at most $(1-c)^M$. Thus the probability that none of the negative walkers meets or passes one of the positive walkers in the first $M i^2 m^2$ steps is at most $(1-c)^M$, i.e., exponentially small in M . It follows that $g(p)$ is finite and that a.s. the number of steps before the first positive walker passes the first negative walker is bounded in probability.

Still for fixed p , define a sequence of integers $d = d(N)$ satisfying $d = p^{-1} \log N + O(1)$, where the $O(1)$ term ensures that such an integer sequence can be found. Define $\tilde{g}(p)$ analogously to $g(p)$ but restricted to the walkers initially occupying the set $\mathbb{Z}(d) = \{k : |k| \leq d\}$. By the observation at the end of the previous paragraph, $\tilde{g}(p) \sim g(p)$.

Next consider our standard model of w walkers on a cycle of length $N + d$, where $w = \lfloor pN \rfloor$. Let $g_1(p)$ denote the expected lifespan of a d -hole line conditional upon it being born of size d . This is clearly independent of the time and place of birth, but let us fix the birth at a vertex v between time steps t and $t+1$. All possible configurations that are feasible at time steps t and $t+1$ and such that this birth occurs are equally likely. At time step t there is a walker at v and one at $v + d + 1 \pmod{N}$, with none in between. For any such configuration of walkers at time step $t+1$, we can map the part of the configuration near v (to its “left”) and near $v + d + 1$ (to its “right”) onto the integers, taking $[v-d, v]$ onto the integers $[-d, 0]$ and $[v+d+1, v+2d]$ onto $[1, d]$. Let V_d denote the set of sites mapped in this way to $\mathbb{Z}(d)$, and let f denote the mapping.

For the rest of the proof we omit some technical details. For any particular configuration in V_d , the probability that it arises is close to the probability of its image under f occurring in the p -model. In fact, it tends toward the same probability as $N \rightarrow \infty$. Similarly, the probability that two walkers occupy the same position within $\mathbb{Z}(d)$ or V_d will tend to 0 as $N \rightarrow \infty$ (in both cases). Hence $g_1(p) \rightarrow g(p)$ as $N \rightarrow \infty$.

Now let $g_2(p)$ denote the expected lifespan of a d -hole line conditional upon it being born of size $d+1$. The same argument as above gives $g_2(p) \rightarrow g(p)$ as $N \rightarrow \infty$.

It follows that for the walkers on the cycle with these parameters, $L_{\text{av}} \rightarrow g(p)$ as $N \rightarrow \infty$. Letting $p \rightarrow 0$ slowly and applying Theorem 3.6 now gives the theorem. $\square$

It is also interesting to ask how the average lifespan relates to the initial config-

uration of the walkers (see the train paradox in section 1). As noted earlier, if N is odd, then $\mathcal{M}(N, w, d)$ is ergodic and parity is immaterial. However, if N is even and the initial configuration is conditioned upon, the (conditional) average lifespan of d -hole lines may be different from L_{av} . Define $T \in \mathbb{Z}^+$,

$$L_T = \frac{\sum_{t=0}^{T-1} \sum_{v \in V} L_{v,t}}{|\{(v, t) : L_{v,t} > 0\}|},$$

where the denominator runs over all pairs $(v, t) \in V \times \{0, \dots, T-1\}$. If the denominator is zero (or the numerator is infinite, which happens with probability 0), the value is immaterial and may be defined as 0. Note that L_T is a function of a given trajectory of the process.

We next aim for an analogue of Theorem 1.4. Define

$$L = \frac{\mathbf{E}[H(0)|\Lambda]}{\mathbf{E}[B(0)|\Lambda]},$$

where Λ is the random variable which is the closed class in which the initial state $\mathcal{M}_0$ lies.

THEOREM 3.8. *For the walkers model on C_N , L_T converges in probability as $T \rightarrow \infty$ (with N fixed) toward a random variable L . Furthermore, (as $N \rightarrow \infty$) we have $L \sim L_{\text{av}}$ a.a.s.*

Proof. The proof proceeds by a renewal theory type of argument, showing a sharp concentration of the number of times that the associated Markov chain revisits a given state over a given long time period.

Let us define the truncated average lifespan of d -hole lines in $[0, T-1]$ to be

$$\overline{L}_T = \frac{\sum_{t=0}^{T-1} H(t)}{H(0) + \sum_{t=0}^{T-2} B(t)}$$

(defined by convention to be 0 if the denominator is 0). As we prove below, this is an approximation of L_T .

We first deal with the case that N is even. To prove the result we need to take into account the class of states containing the initial state, since different starting configurations of walkers may lead to different expected numbers of holes and births.

Let Λ be the random variable accounting for the closed class of states where the initial state $\mathcal{M}_0$ lies. Let A be any closed class of states. We condition on the event $(\Lambda = A)$, which we call A with some abuse of notation. By Lemma 3.2, the hitting time between any two states is finite.

We consider the Doob martingale $\Sigma_0, \dots, \Sigma_T$ defined by

$$\Sigma_i = \mathbf{E} \left[\sum_{t=0}^{T-1} H(t) \middle| A, \mathcal{M}_0, \dots, \mathcal{M}_{i-1} \right], \quad i = 0, \dots, T.$$

We have $\Sigma_0 = T\mathbf{E}[H(0)|A]$ and $\Sigma_T = \sum_{t=1}^T H(t)$. (We recall that this last expression is regarded in the probability space induced by restricting to the event A .)

From the fact that the expected hitting time between any two states is finite, we deduce that the differences $|\Sigma_{i+1} - \Sigma_i|$ are bounded above by a constant independent of T . Hence as an immediate consequence of Azuma's inequality we get that, conditional upon the event A ,

$$(9) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} H(t) = \mathbf{E}[H(0)|A] \quad \text{in probability}$$

and by a similar argument

$$(10) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \left(H(0) + \sum_{t=0}^{T-2} B(t) \right) = \mathbf{E}[B(0)|A] \quad \text{in probability.}$$

Then, by taking the ratio of (9) and (10), we get that

$$(11) \quad \lim_{T \rightarrow \infty} \overline{L_T} = L \quad \text{in probability.}$$

The case that N is odd is easier. There is just one closed class of states, and $\mathcal{M}(N, w, d)$ is ergodic. By the martingale argument as above, we get (noting Theorem 3.6)

$$(12) \quad \lim_{T \rightarrow \infty} \overline{L_T} = \frac{\mathbf{E}[H(0)]}{\mathbf{E}[B(0)]} = L_{\text{av}} \quad \text{in probability.}$$

Moreover, we obtain

$$(13) \quad \lim_{T \rightarrow \infty} L_T - \overline{L_T} = 0 \quad \text{in probability}$$

since $H(t)$ and $B(t)$ are uniformly bounded (the bound depending on N) and so is the expected lifespan of any line. Thus the portions of lifespans omitted in $\overline{L_T}$ are bounded in expectation, which is insignificant since the denominator grows linearly with T , as shown in (10).

In order to finish the proof, it suffices to show that as $N \rightarrow \infty$

$$L \sim L_{\text{av}} \quad \text{a.a.s.}$$

We note that the quantity $\mathbf{E}[H(0)|A]/\mathbf{E}[B(0)|A]$ may vary depending on the particular closed class of states A . Let us study this in more detail. Let $\mathbf{a} \in A$ be a configuration. As in section 3.2, let us partition the set of walkers W into W_1 and W_2 according to the parity of their positions in the cycle. Let us define the *imbalance* of the configuration as $\Delta(\mathbf{a}) = |w_1 - w_2|$, where $w_i = |W_i|$. It makes sense to define $\Delta(A) = \Delta(\mathbf{a})$ since it does not depend on the choice of $\mathbf{a}$.

We can compute the expectations of S , B , and D conditional upon $(\mathcal{M}_0 \in A)$ by proceeding the same way as in Proposition 3.4. The only difference is that w_1 walkers must go to $\frac{N}{2}$ of the vertices (say, those with odd position) and w_2 must go to the other $\frac{N}{2}$. We omit the details here since they are fairly tedious and completely analogous to the previous computations. We note that these expectations do not depend on the particular partition (W_1, W_2) but only on the imbalance $\Delta(A)$. In all cases we get

$$(14) \quad \frac{\mathbf{E}[H(0)|A]}{\mathbf{E}[B(0)|A]} = \Theta \left(\frac{\mathbf{E}[H(0)]}{\mathbf{E}[B(0)]} \right).$$

In fact, for $\frac{\Delta(A)}{N} = O(1)$ we have $\mathbf{E}[H(0)|A] = \Theta(\mathbf{E}[H(0)])$ and $\mathbf{E}[B(0)|A] = \Theta(\mathbf{E}[B(0)])$. For $\frac{\Delta(A)}{N} \rightarrow \infty$, these two statements are no longer true, but the extra factors in the numerator and denominator of (14) cancel out to within a factor of $\Theta(1)$.

However, not all imbalances are equally likely. In fact, for any $\epsilon > 0$, we have

$$(15) \quad \mathbf{P} \left(\Delta(\mathcal{M}_0) \geq w^{\frac{1+\epsilon}{2}} \right) = \mathbf{P} \left(\left| w_1(\mathcal{M}_0) - \frac{w}{2} \right| \geq \frac{1}{2} w^{\frac{1+\epsilon}{2}} \right) \leq \frac{w}{w^{1+\epsilon}} = o(1).$$

Moreover, for a (typical) class A such that $\Delta(A) < w^{\frac{1+\epsilon}{2}}$ and by the method explained above, we get

$$(16) \quad \frac{\mathbf{E}[H(0)|A]}{\mathbf{E}[B(0)|A]} \sim \frac{\mathbf{E}[H(0)]}{\mathbf{E}[B(0)]}.$$

From this last fact, together with (14) and (15), the theorem follows. $\square$

We turn now to connectivity issues, for which we use (1). The next lemma gives us the probability that there is one component at time step t but at least two at time step $t+1$.

LEMMA 3.9. *The probability that $\mathcal{W}_t(C_N)$ is connected and that $\mathcal{W}_{t+1}(C_N)$ is disconnected is given by*

$$\mathbf{P}\left(H(t+1) \geq 2 \wedge H(t) < 2\right) \sim \begin{cases} \frac{1}{2}\mu^2 e^{-\mu} \varrho & \text{if } \varrho \rightarrow 0, \\ e^{-\mu} (1 + \mu - (1 + \mu + \lambda + \lambda^2)e^{-\lambda}) & \text{if } \varrho \rightarrow c, \\ (1 + \mu)e^{-\mu}(1 - (1 + \mu)e^{-\mu}) & \text{if } \varrho \rightarrow \infty, \end{cases}$$

where λ is defined as in Theorem 1.3.

Proof. We can split the probability on the left-hand side according to the events $H(t) = 0$ and $H(t) = 1$. Noting that

$$\mathbf{P}\left(H(t+1) \geq 2 \wedge H(t) = 0\right) = \mathbf{P}\left(H(t) = 0 \wedge B(t) \geq 2\right),$$

and

$$\begin{aligned} \mathbf{P}\left(H(t+1) \geq 2 \wedge H(t) = 1\right) &= \mathbf{P}\left(S(t) + B(t) \geq 2 \wedge S(t) + D(t) = 1\right) \\ &= \mathbf{P}\left(S(t) = 1 \wedge B(t) \geq 1 \wedge D(t) = 0\right) + \mathbf{P}\left(S(t) = 0 \wedge B(t) \geq 2 \wedge D(t) = 1\right), \end{aligned}$$

the result follows from Proposition 3.4 and Corollary 3.5 provided $\varrho \not\rightarrow 0$. If $\varrho \rightarrow 0$, we need, for instance, to estimate the vanishingly small quantity $\mathbf{P}(S(t) = 1 \wedge B(t) \geq 1 \wedge D(t) = 0)$ asymptotically, but the moment calculations, as for the proof of Proposition 3.4, suffice for this, and we omit the routine details. $\square$

Recall the definition of a disconnected period from section 1. By Lemma 3.2, from a disconnected state the graph will always reach some connected one (for example, one in which all walkers occupy one of two adjacent sites) within finite expected time. Thus the expected length of any disconnected period is bounded (this bound depending on N).

Proof of Theorem 1.3. Arguing as in the proof of Theorem 3.6, and obtaining the asymptotic expressions for the numerator and the denominator from Theorem 1.1 and Lemma 3.9, we obtain

$$(17) \quad LD_{\text{av}} = \frac{\mathbf{P}(H(1) \geq 2)}{\mathbf{P}(H(0) < 2 \wedge H(1) \geq 2)} \sim \begin{cases} 2 \frac{e^\mu - 1 - \mu}{\mu^2} \varrho^{-1} & \text{if } \varrho \rightarrow 0, \\ \frac{e^\mu - 1 - \mu}{1 + \mu - (1 + \mu + \lambda + \lambda^2)e^{-\lambda}} & \text{if } \varrho \rightarrow c, \\ \frac{e^\mu}{1 + \mu} & \text{if } \varrho \rightarrow \infty. \end{cases} \quad \square$$

Proof of Theorem 1.4. Let the random variable $\Phi(t) = 1_{[H(t) \geq 2]}$ be the indicator of the event that $\mathcal{W}_t(C_N)$ is disconnected, and let $\Psi(t) = 1_{[H(t-1) < 2 \wedge H(t) \geq 2]}$ be the indicator of the event that a disconnected period starts at time step t .

We define the truncated average disconnection time of the graph of walkers in $[1, T]$ as

$$\overline{LD_T} = \frac{\sum_{t=1}^T \Phi(t)}{\sum_{t=1}^T \Psi(t)}.$$

Proceeding as in the proof of Theorem 3.8, we get that $\overline{LD_T}$ converges in probability for large T toward a random variable LD such that $LD \sim LD_{\text{av}}$. $\square$

From the definitions in section 1 relating to disconnected periods, by changing disconnected to connected, we define *connected periods* LC_t , LC_{av} , and LC_T . Using

$$LC_{\text{av}} = \frac{\mathbf{P}(H(1) < 2)}{\mathbf{P}(H(0) \geq 2 \wedge H(1) < 2)}$$

and an analogue of Lemma 3.9, we obtain the following.

THEOREM 3.10. *The average length of a connected period satisfies*

$$LC_{\text{av}} \sim \begin{cases} 2^{\frac{1+\mu}{\mu^2}} \varrho^{-1} & \text{if } \varrho \rightarrow 0, \\ \frac{1+\mu}{1+\mu-(1+\mu+\lambda+\lambda^2)e^{-\lambda}} & \text{if } \varrho \rightarrow c, \\ \frac{e^\mu}{e^\mu-(1+\mu)} & \text{if } \varrho \rightarrow \infty, \end{cases}$$

where λ is defined as in Theorem 1.3.

The result for connection times analogous to Theorem 1.4 is the following. The proof is almost identical and so is omitted.

THEOREM 3.11. *LC_T converges in probability as $T \rightarrow \infty$ (N fixed) toward a random variable LC , where $LC \sim LC_{\text{av}}$ a.a.s.*

4. The grid. In this section, we study the walkers model for $G = T_N$, the toroidal grid with $N = n^2$ vertices. We can refer to vertices by using coordinates in $\mathbb{Z}_n \times \mathbb{Z}_n$. For the grid we encounter significant new obstacles as compared to the cycle; see, for instance, the Geometric Lemma below.

For any $p \in [1, \infty]$ and any two vertices u and v in T_N , we define the distance $\text{dist}_{\ell^p}(u, v)$ as the minimal ℓ^p distance between any two points u' and v' in the square grid such that the coordinates of u' are congruent to those of u taken modulo n , and similarly for v and v' . Let us fix any such metric for our study of the connectedness of $\mathcal{W}(G, w, d)$ and write $\text{dist}(u, v)$ for short. We use dist to refer to this measure of distance, distinguishing it from Euclidean distance. Note that

$$(18) \quad \frac{1}{2} \cdot \text{dist}_{\ell^1}(u, v) \leq \text{dist}(u, v) \leq \text{dist}_{\ell^1}(u, v).$$

The number of vertices at distance at most d from any given vertex is $h = \Theta(d^2)$. The exact expression of h depends on the choice of the metric. Some examples are found in Table 1.

OBSERVATION 4.1. *Assume that $d < 2n$ (otherwise, the graph of walkers is always complete). For each $i, j < 4n/d$, let v_{ij} denote the vertex in T_N with coordinates $(\lfloor id/4 \rfloor, \lfloor jd/4 \rfloor)$. Let S_{ij} denote the set of grid points closer to v_{ij} than any of the other $v_{i'j'}$. Then there are $\Theta(N/d^2)$ disjoint sets S_{ij} each containing $\Theta(d^2)$ points. The probability that at least one of these S_{ij} is empty of walkers is at most*

$$\Theta(N/d^2)(1 - \Theta(d^2/N))^w = O(\sqrt{w})e^{-\Omega(\sqrt{w})},$$

TABLE 1
Number of vertices at distance at most d from a given vertex.

Metric	h
ℓ^1	$h = 2d(d+1)$
ℓ^2	$h \sim \pi d^2$ if $d \rightarrow \infty$
ℓ^∞	$h = 4d(d+1)$

which goes to 0 if $d^2 = \Omega(N/\sqrt{w})$. Thus, a.a.s. each of these pieces is occupied by at least one walker, and $\mathcal{W}(G, w, d)$ is connected.

In view of the observation, we assume for the rest of the section that $h = o(N)$, i.e., $d = o(n)$. If $d = \Omega(n)$, then $\mathcal{W}(G, w, d)$ is a.a.s. connected.

We wish to study the connection and disconnection of $\mathcal{W}(G, w, d)$ in a similar way to the cycle. For the grid, the notion of hole does not help, and we deal directly with components. Recall from the introduction that a simple component is one with just one vertex. These play a major role, and we shall prove that, for the interesting values of the parameters, a.a.s. there exist only simple components aside from one giant one.

Let $\mathcal{C}$ be any given component. The *edges* of $\mathcal{C}$ are the straight edges joining occupied vertices in $\mathcal{C}$ of distance at most d . The associated *forbidden region* $A_{\mathcal{C}}$ is the set of vertices not in $\mathcal{C}$ but at distance at most d from some vertex in $\mathcal{C}$ (i.e., those vertices which must be free of walkers for $\mathcal{C}$ to exist as a component). The *exterior* $\mathcal{E}_{\mathcal{C}}$ of $\mathcal{C}$ is the set containing all those vertices not in $\mathcal{C} \cup A_{\mathcal{C}}$. We partition $\mathcal{E}_{\mathcal{C}}$ into *external regions* as follows: two vertices belong to the same external region when they can be joined by a continuous arc not intersecting any edge of $\mathcal{C}$. Figure 3 shows a component with different external regions.

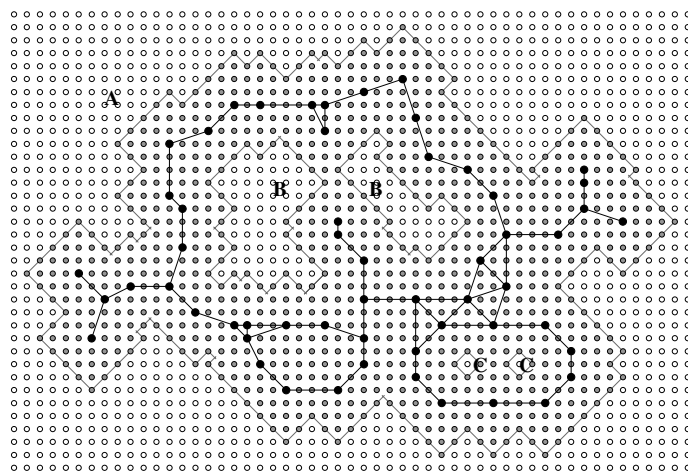


FIG. 3. Component (black), empty area (gray), external regions A, B, and C (white).

Recall that, in the terminology of planar maps, the *bounding cycle* of a face is a walk around the boundary of the face. Given an external region $\mathcal{E}_{\mathcal{C}}^i$, let $\mathcal{C}'$ be any connected subgraph of $\mathcal{C}$ that has no edges crossing and such that no vertices of $\mathcal{C}$ are contained in the face F of $\mathcal{C}'$ which contains $\mathcal{E}_{\mathcal{C}}^i$. Such subgraphs always exist: for instance, take the spanning tree of $\mathcal{C}$ whose length (sum of lengths

of edges) in terms of dist has been minimized and, subject to this, has the shortest Euclidean length. We refer to the bounding cycle of this face F as a *boundary walk* β in $\mathcal{C}$ with respect to $\mathcal{E}_C^i$. Such a walk is *maximal* if the face F does not properly contain a face of some other subgraph of $\mathcal{C}$ of the same type. In Figure 4, $\{1, 5, 3, 7, 9, 11, 13, 15, 13, 12, 14, 12, 10, 8, 6, 2, 4, 1\}$ is a nonmaximal boundary walk, and $\{1, 5, 3, 7, 9, 11, 13, 14, 12, 10, 8, 6, 2, 4, 1\}$ is a maximal one. A maximal walk always exists, because any nonmaximal walk can be diverted around any face that prevents it from being maximal.

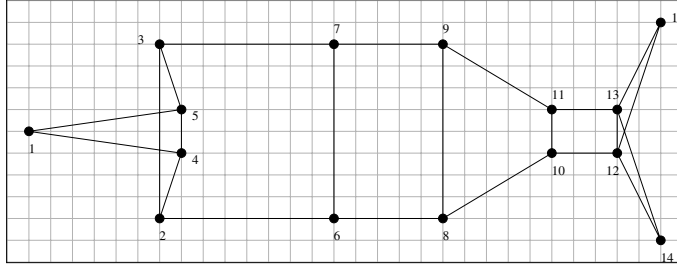


FIG. 4. The boundary walk.

We call a (directed) closed walk in $\mathcal{C}$ *regular* if, for each edge entering a vertex v , the next edge in the walk is the next edge in the clockwise direction around v .

For $i < n$, let us call a v -band of width i to any subset of T_N defined by $\{a, \dots, a + i - 1\} \times \mathbb{Z}_n$. Similarly, we define an h -band of height j . Define a *rectangle* of width i and height j to be the intersection of a v -band of width i and an h -band of height j . We can compare vertices in a rectangle according to their coordinates and use statements such as v_1 is more left than v_2 or v_3 is an uppermost vertex in the rectangle.

We say that a component $\mathcal{C}$ with at least 2 vertices is a *rectangular component* (*r-component*) if all of its vertices and edges and its forbidden region are contained in a rectangle as defined above. In particular, this implies that $\mathcal{C}$ contains no nonseparating cycle of the torus. Otherwise, $\mathcal{C}$ is an *nr-component*. For a given r -component $\mathcal{C}$, we define its *origin* as the leftmost of the lowermost vertices of $\mathcal{C}$. The *outside region* of an r -component is the only external region of the component having vertices outside any containing rectangle. We say that an nr -component is *type 1* if it is not the subgraph of some graph of walkers containing more than one nr -component and *type 2* otherwise. Note that this property is simply determined by the subgraph of the toroidal grid remaining when the component and the vertices of distance at most d from it are deleted. By definition, $\mathcal{W}(G, w, d)$ has at most one nr -component of type 1.

Let X , Y , and Z be the number of simple components, r -components, and nr -components, respectively, and note that $Z = Z_1 + Z_2$, where Z_i is the number of nr -components of type i . Then $\mathbf{E}[Z] = \mathbf{P}(Z_1 = 1) + \mathbf{E}[Z_2]$.

4.1. Static properties. In this section, we study the connectedness of $\mathcal{W}(G, w, d)$ in the static situation for the case $G = T_N$; in particular, we analyze the behavior of X , Y , and Z . For section 4, the value of μ is as defined for Theorem 1.5, rather than the value for the cycle.

PROPOSITION 4.2. *The expected number of simple components of $\mathcal{W}(G, w, d)$ for $G = T_N$ satisfies $\mathbf{E}[X] \sim N(1 - e^{-\mu})(1 - \frac{h}{N})^w$. Furthermore,*

- (i) *if $\mu \rightarrow 0$, then $\mathbf{E}[X] \rightarrow 0$ and there are no simple components a.s.,*

- (ii) if $\mu \rightarrow \infty$, then there exist simple components a.a.s. (and $\mathcal{W}(G, w, d)$ is disconnected), and
- (iii) if $\mu = \Theta(1)$, then X is asymptotically Poisson with mean μ .

Proof. We repeat the proof of Theorem 1.1 in the present context. To compute $\mathbf{E}[X]_k$, we focus on the set $\mathcal{S} = \{(v_1, \dots, v_k) | \text{dist}(v_i, v_j) > 2d \text{ if } i \neq j\}$, with size $|\mathcal{S}| \sim N^k$. Applying Lemma 2.1, we obtain

$$(19) \quad \mathbf{E}[X]_k \sim \left[N (1 - e^{-\ell}) e^{-h\ell + O(\frac{h^2 w}{N^2})} \right]^k.$$

Comparing with (4), the rest of the proof is as for Theorem 1.1. $\square$

From part (ii) of the proposition, if $h\varrho = O(1)$, then $\mu \rightarrow \infty$ and $\mathcal{W}(G, w, d)$ is disconnected a.a.s. In view of this, we may restrict our attention to the condition $h\varrho \rightarrow \infty$ in the study of r -components and nr -components.

Given a boundary walk $\beta = (v_1, \dots, v_k)$, we define $\text{length}(\beta) = \sum_{1 \leq i < k} \text{dist}(v_i, v_{i+1})$ as the sum of the distances (using the chosen metric) between consecutive vertices in β . We shall write $\text{length}_{\ell^p}(\beta)$ when we want specify that we are measuring distances in ℓ^p . Similarly, we define $\text{length}_v(\beta)$ (the vertical length) as the sum of the differences between y coordinates of consecutive vertices along the cycle and $\text{length}_h(\beta)$ (the horizontal length) using x coordinates in the same way.

The next lemma relates the size of the forbidden region outside a boundary cycle of a component to the length of the cycle and will play a key role in proving the main results.

LEMMA 4.3 (Geometric Lemma). *Let $\mathcal{C}$ be a component in T_N with β one of its maximal boundary walks and $l = \text{length}(\beta)$ its length. Assume that $\mathcal{C}$ has at least two occupied sites. Then the size of the forbidden region A_β outside β is bounded below by $|A_\beta| \geq dl/J$ for some sufficiently large constant J . Moreover, if $\mathcal{C}$ is rectangular, and β is a maximal boundary walk with respect to the outside region, we have $|A_\beta| \geq h + dl/J$.*

Proof. The main idea in this proof is to consider intervals in the grid “sticking outward” from the boundary walk and show that at least some small proportion of these intervals is on average free of walkers.

For convenience we take $J = 10^{10}$, though probably without large modifications the proof method will yield the result for $J = 1000$. (We do not attempt to optimize the constant J , since probably the theorem holds for J close to 2 when d is large, while for $d = 1$ in all but the case $p = 1$ we need $J \geq 4$. Some experiments suggest that $J = 4$ suffices even in this case.) Observe that in the 3×3 subgrid centered on the endpoint of an edge of β , there must be at least one vertex in the forbidden region outside β ; otherwise the boundary walk can be rerouted to contradict its maximality. Each point of the forbidden region can be counted in this way at most 8 times. Thus the forbidden region has size at least $l/(8d)$. This implies the first statement in the lemma provided that $8d^2 \leq 10^{10}$.

Throughout this proof, all distances referred to are measured using dist . For the second statement in the lemma, we begin with the fact that for a rectangular component, there are “caps” of empty region of size at least $h/2 - d$ on the top and bottom of the component, being the set of grid points within distance d , but above the leftmost point of $\mathcal{C}$ of the greatest vertical coordinate (or below the leftmost point of the least vertical coordinate). Without loss of generality, we assume that the component has vertical height greater than 1 (otherwise, we may interchange “vertical” and “horizontal”). Then there are also intervals of empty region of length

d projecting outward from the left- and rightmost vertices of the top level of $\mathcal{C}$. These caps and intervals account for an empty region of size h but use up some of the forbidden region counted in relation to ends of edges of β in the argument above. The number of endvertices of edges of β that are involved in the region of size h described is at most $4d + 2$, and there is still at least one endvertex of an edge of β not involved (the rightmost one in the bottom row of $\mathcal{C}$, say—we cannot claim two vertices here because it might be the same as the leftmost one), which means the above argument is valid if the bound on the size of the forbidden region is reduced by a factor of $4d + 3$. That is, the size of the forbidden region is at least $h + l/8d(4d + 3)$. So the second statement in the lemma holds provided that $8d^2(4d + 3) \leq 10^{10}$. Hence we may now assume that $d > 400$.

Choose $k = \lceil \frac{d}{100} \rceil$. Then $k \leq (d + 99)/100 \leq (d/100)(1 + 99/d) < d/80$. Assume without loss of generality that for the boundary walk β we have $\text{length}_v(\beta) \geq \text{length}(\beta)/2$. We place intervals of length k (note that all the ℓ_p metrics measure these intervals the same) horizontally along all grid lines from a maximal boundary cycle toward the outside. Those starting from a vertex of β point toward the outside according to the previous edge of β . (We assume β is oriented in some direction.) We delete any intervals that touch the boundary cycle in two or more places. Then each remaining interval will touch k vertices of the forbidden region outside β , and each such vertex will be touched by at most two different intervals. (Two intervals coming from opposite directions may touch the same vertex.)

We need to bound the number of intervals which were deleted. Call an edge of $\mathcal{C}$ *short* if the distance between its endpoints is less than $d/4$ and *long* otherwise. Suppose that an interval (that is to be deleted) touches a short edge at a point E and a point F on another edge of β , with the part between E and F in the exterior of β . Suppose that F has distance at most $d - d/4 - k$ from both endvertices of its edge. Then by the triangle inequality each endvertex (say, P and Q) of that edge is of distance at most d from each endvertex (R and T) of E 's edge. Thus the quadrilateral $PQRT$ (or triangle, if two points coincide) has diameter at most d , and thus the walk β can be changed to make a smaller face “outside” (meaning the side which was minimized). This can clearly be done also if other parts of β enter this quadrilateral, contradicting the maximality of β . Thus, F has distance at least $d - d/4 - k > 2d/3$ from one of the endvertices of its edge. So F 's edge is long, or, if F lies on more than one edge, they are all long.

We call a *middle* interval any interval originating from a long edge from point at least $1/8$ of the length of the edge from each end. Suppose that such a middle interval originating at a point E (called a *middle point*), on edge e of β , is deleted. Then it hits some other edge f at a point F of horizontal distance at most k from E . If f has an endvertex of distance less than $d/8 - k \geq d/8 - d/80 \geq d/9$ from F , we get a contradiction as above. Since e is long, its endvertices have distance at least $d/32$ from E . Since e and f are straight, either point on e of distance $rd/32$ from E , for any $r > 0$, has horizontal distance at most $(r + 1)k$ from f or the extension of f . (Furthermore, note for later that any two points on e and f at the same vertical coordinate have distance at most $32k$ apart.) It follows that the endvertices P and R of e and f above the line EF have distance less than d apart, as do the ends Q and T below (or on) the line EF . If either $P = R$ or $Q = T$, we now obtain a contradiction as before by rerouting β . A similar contradiction arises if either PT or QR has length at most d . So, picturing e and f with the line EF as making a near-perfect but very thin “H,” the distances along the H from P to T and from Q to R are both greater

than d . Thus the sum of the lengths of e and f is at least $2d - 2k$ (since the edge EF is included twice in these distances), and so each of e and f has length at least $d - 2k$, and the lengths of the arms of e and f above EF differ by at most k (as for the arms below EF). It follows that both PR and QT have length at most $34k$ (recalling the observation above about horizontal distance), which is at most $d/2$. Without loss of generality, PR is not an edge of β . This again contradicts the choice of β , either by shortening it at PR or, if another edge of β crosses PR , by joining P or R to an end of such an edge. We conclude that no middle intervals are deleted. Moreover, this shows that the point F in the previous paragraph cannot be a middle point. So every interval starting at a short edge that is deleted first hits a nonmiddle point of a long edge. Such a point can be hit by only one interval from a short edge unless it is a vertex of β , in which case it can be hit by two intervals. Thus if i_1 is the number of intervals starting at short edges that are deleted, j_1 is the number of long edges of β , and j_2 is the number of nonmiddle intervals originating at nonends of long intervals,

$$i_1 \leq 4j_1 + j_2.$$

Moreover, since no middle intervals are deleted, the number i_2 of intervals originating on long edges that are deleted similarly satisfies

$$i_2 \leq 4j_1 + j_2.$$

If j_3 is the total number of intervals before any deletions occur, we clearly have

$$j_2 \leq j_3/4$$

and also since long edges have length at least $d/4$,

$$(20) \quad j_3 \geq \text{length}_v(\beta) \geq \text{length}(\beta)/2 \geq j_1 d/8.$$

Combining these gives

$$i_1 + i_2 \leq 8j_1 + 2j_2 \leq j_3 \left(\frac{64}{d} + \frac{1}{2} \right) \leq 3j_3/4,$$

which shows that at least $j_3/4$ of all the intervals are not deleted. These intervals each cover $k \geq d/100$ vertices of the forbidden region, at most two covering any one such vertex, so the first part of the lemma follows for $d > 400$, using (20).

For the second claim of the lemma when $d > 400$, we may again add the caps of size $h - 2d$ but also two extra intervals of length d at the sides: assuming that β is oriented in the clockwise direction, the intervals of length k projecting from the leftmost vertex in the bottom level of $\mathcal{C}$, and from the rightmost vertex in the top level, are not used. It is here that extra intervals of the empty region of length d may be found. $\square$

The next lemma will be used to show that nr-components of type 2 are rare.

LEMMA 4.4. *Let $\mathcal{C}$ be an nr-component of type 2. Then $\mathcal{C}$ has a maximal boundary walk β with $\text{length}(\beta) \geq n - o(n)$.*

Proof. Let us divide T_N into $\lfloor \frac{n}{d+1} \rfloor$ v-bands $c_1, \dots, c_{\lfloor \frac{n}{d+1} \rfloor}$ of width $\geq d + 1$, and similarly into $\lfloor \frac{n}{d+1} \rfloor$ h-bands $r_1, \dots, r_{\lfloor \frac{n}{d+1} \rfloor}$ of width $\geq d + 1$.

Let $\mathcal{C}$ be an nr-component. Since $\mathcal{C}$ is connected, the v-bands (or h-bands) not containing vertices of $\mathcal{C}$ must be consecutive. If there were at least 2 consecutive v-bands and at least 2 consecutive h-bands without vertices of $\mathcal{C}$, then $\mathcal{C}$ would be

rectangular since $\mathcal{C}$ and $A_{\mathcal{C}}$ would be embeddable in the complements of the v-bands and h-bands. Hence, at most one v-band and some consecutive h-bands (or at most one h-band and some consecutive columns) may be without vertices of $\mathcal{C}$.

We assume now that $\mathcal{C}$ coexists with another nr-component $\mathcal{C}'$.

Case 1. Let us suppose first that $\mathcal{C}$ has no vertices in more than one v-band or h-band (v-band without loss of generality). Let c_1, c_2 be two consecutive v-bands not containing vertices of $\mathcal{C}$. Hence all h-bands, excepting at most one, contain vertices in $\mathcal{C}$. For each such h-band r_i , choose a vertex v_i in $\mathcal{C} \cap r_i$. We can also find some vertex w_i in $(c_1 \cup c_2) \cap r_i$ such that w_i is at distance $\geq d+1$ from any vertex in $\mathcal{C}$. By this construction, all w_i belong to the same external region of the component. Let β be any maximal boundary walk of $\mathcal{C}$ with respect to this external region. Then, the straight line joining v_i and w_i must intersect an edge of β , part of the edge contained in r_i . Hence, β crosses all h-bands except at most 3 and $\text{length}(\beta) \geq n - 4d - 3$.

Case 2. On the other hand, let us suppose that $\mathcal{C}$ has vertices in all v-bands and h-bands except for at most one of each. Without loss of generality, the other component, $\mathcal{C}'$, has vertices in all v-bands except for at most one. Thus, there are at least $\lfloor \frac{n}{d+1} \rfloor - 2$ v-bands with some vertices of both components. For each such v-band c_i , let us take vertices $v_i \in \mathcal{C} \cap c_i$, $w_i \in \mathcal{C}' \cap c_i$ and join them by a straight line. Notice that all the w_i belong to the same external region of $\mathcal{C}$, and let β be any maximal boundary walk with respect to this region. Then the line joining v_i and w_i must intersect an edge of β , part of the edge contained in c_i . Hence, β crosses all v-bands except at most 4 and $\text{length}(\beta) \geq n - 5d - 4$.

In any case, $\mathcal{C}$ has a maximal boundary walk β with $\text{length}(\beta) \geq n - o(n)$. $\square$

The next technical result shows that simple components are predominant a.a.s. in T_N . The proof uses the Geometric Lemma.

LEMMA 4.5. *If $h_Q \rightarrow \infty$, then $\mathbf{E}[Y] = o(\mathbf{E}[X])$ and $\mathbf{E}[Z_2] = o(\mathbf{E}[X])$.*

Proof. Let us first bound the expected number $\mathbf{E}[Y]$ of rectangular components with more than one vertex. The amount of area that is e.o.w., as guaranteed by the Geometric Lemma, is large enough to make such components rare.

Notice from (18) that the ℓ^1 -length of the edges of any boundary walk of a component are integers between 1 and $2d$.

Let $\mathcal{B}$ be the set of walks in T_N which are (for some configuration of the walkers) a maximal boundary walk of some r-component with respect to its outside region. For each $\beta \in \mathcal{B}$, choose a rooted spanning tree $T(\beta)$ of the graph induced by the edges of β . Note that given any such tree T of m vertices, we may recover β by joining certain pairs of vertices (with no edges crossing). The edges added are just diagonals added to a face of degree $2m - 2$. For each vertex $v \in V$, natural $m \geq 2$, and tuple $\mathbf{l} = (l_1, \dots, l_{m-1})$ of naturals $1 \leq l_i \leq 2d$, let $\mathcal{B}_{v,m,\mathbf{l}}$ be the set of all $\beta \in \mathcal{B}$ such that $T(\beta)$ has m vertices, is rooted at v , and has edges of ℓ^1 -lengths $l_1, \dots, l_{m-1}$. The number of such trees is at most $\prod_{j=1}^{m-1} 16l_j$, where a factor 4^m comes from the number of rooted plane trees, and each factor $4l_j$ is the number of vertices of ℓ^1 distance l_j from a given vertex. Therefore,

$$(21) \quad |\mathcal{B}_{v,m,\mathbf{l}}| \leq \prod_{j=1}^{m-1} Cl_j,$$

where C is constant, and also clearly

$$(22) \quad \mathbf{E}[Y] \leq \sum_{\substack{v \in V \\ m \geq 2 \\ 1 \leq l_1, \dots, l_{m-1} \leq 2d}} \sum_{\beta \in \mathcal{B}_{v,m,l_1,\dots,l_{m-1}}} \mathbf{P}(Y_\beta = 1),$$

where Y_β indicates the event of having some r -component with β one of its maximal boundary walks with respect to its outside region.

By Lemma 4.3, the size $|A_\beta|$ of the forbidden region outside β is an integer bounded below by $h + dl/J$, where $l = \text{length}(\beta)$. For technical purposes we consider a subset of A_β of size $h + \lceil dl/(2J+1) \rceil$, representing a region free of walkers. By Lemma 2.2, and noting that $\alpha > 1/2$ and hence $p < 2\varrho$, we obtain an upper bound for the probability of this emptiness occurring and the m vertices in β being occupied. Since this is necessary for the event Y_β to occur, we have

$$(23) \quad \begin{aligned} \mathbf{P}(Y_\beta = 1) &= O(1 - e^{-\varrho/\alpha})(2\varrho)^{m-1} \left(1 - \frac{S}{N}\right)^w \\ &= O(1 - e^{-\varrho})(2\varrho)^{m-1} \left(1 - \frac{h}{N}\right)^w e^{-\lceil dl/(2J+1) \rceil \varrho} \end{aligned}$$

since $\alpha \leq 1$. Furthermore, let $l' = l_1 + \dots + l_{m-1}$. Then since the spanning tree has length no more than the length of β , $\text{length}_{\ell^1}(\beta) > l'$. By (18), we have $l \geq l'/2$, and hence we get

$$(24) \quad \mathbf{P}(Y_\beta = 1) = O(1 - e^{-\varrho})(2\varrho)^{m-1} \left(1 - \frac{h}{N}\right)^w e^{-dl'\varrho/J'},$$

where $J' = 2(2J+1)$.

From (22), (21), and (24), we get

$$\mathbf{E}[Y] = O(1) \sum_{\substack{m \geq 2 \\ 1 \leq l_1, \dots, l_{m-1} \leq 2d}} N \left(\prod_{j=1}^{m-1} Cl_j \right) (1 - e^{-\varrho}) (2\varrho)^{m-1} \left(1 - \frac{h}{N}\right)^w e^{-dl'\varrho/J'}.$$

Therefore, using Proposition 4.2 for the asymptotic value of $\mathbf{E}[X]$,

$$(25) \quad \begin{aligned} \mathbf{E}[Y]/\mathbf{E}[X] &= O(1) \sum_{\substack{m \geq 2 \\ 1 \leq l_1, \dots, l_{m-1} \leq 2d}} \left(\prod_{j=1}^{m-1} Cl_j \right) (2\varrho)^{m-1} e^{-dl'\varrho/J'} \\ &= O(1) \sum_{m \geq 2} \left(\frac{C'}{d} \sum_{k=1}^{2d} kd\varrho e^{-kd\varrho/J'} \right)^{m-1}. \end{aligned}$$

In the case where $d\varrho \rightarrow \infty$, we have $\varrho e^{-c'd\varrho} = o(\max(1, \varrho)) = o(1)$ and hence $\mathbf{E}[Y] = o(\mathbf{E}[X])$. In the case where $d\varrho = O(1)$, we use $\sum_{k \geq 1} kc^{-\epsilon kd\varrho} = O((d\varrho)^{-2})$.

In the case where $d\varrho = O(1)$, we use $\sum_{k \geq 1} kc^{-kd\varrho} = O((d\varrho)^{-2})$ for $c < 1$, and (25) gives

$$\mathbf{E}[Y]/\mathbf{E}[X] = O(1) \sum_{m \geq 2} \left(\frac{C''}{d^2\varrho} \right)^{m-1} = o(1) \quad \text{as } d^2\varrho \rightarrow \infty.$$

To prove $\mathbf{E}[Z_2] = o(\mathbf{E}[X])$, from Lemma 4.4, each component counted by Z_2 has some maximal boundary walk β with $\text{length}(\beta) \geq n - o(n)$. If we apply Lemma 4.3 to this β , we have $|A_\beta| \geq ld/J$, where $l = \text{length}(\beta)$. Using $d = o(n)$ (since $h = o(N)$), we have $|A_\beta| \geq h + ld/2J$ for large N , and we then proceed similarly as for Y . $\square$

Proof of Theorem 1.5. From Proposition 4.2, if $\mu \rightarrow \infty$, then $\mathcal{W}(G, w, d)$ is disconnected a.a.s. In the other two cases, $\mu = O(1)$ and we must have $h\varrho \rightarrow \infty$. In this case we can apply Lemma 4.5 and get

$$\mathbf{P}(Y > 0) \leq \mathbf{E}[Y] = o(\mathbf{E}[X]) = o(1), \quad \mathbf{P}(Z_2 > 0) \leq \mathbf{E}[Z_2] = o(\mathbf{E}[X]) = o(1).$$

Thus, a.a.s. we have only simple components and at most one nr-component of type 1. The rest of the theorem follows from the asymptotic distribution of X given in Proposition 4.2. $\square$

4.2. Dynamic properties. According to the model, from an initial random placement f of the walkers, at each step, every walker moves from its current position to one of its neighbors, with a probability of $1/4$ of going either way. This is a standard random walk on the grid for each walker. We wish to study the connectivity properties of $\mathcal{W}_t(T_N)$. The analysis of the dynamic case is quite similar to that of the cycle, so we state the major results and point to the differing details in the proofs.

We define *states* (or *configurations*) and the graph of configurations in an analogous way to the cycle (see section 3.2). In this case, there are $N^w = n^{2w}$ different configurations of walkers, each one represented by a vector $\mathbf{a} = (a_1, \dots, a_w) \in (\mathbb{Z}_n \times \mathbb{Z}_n)^w$, where $a_i = (a_{i,x}, a_{i,y})$ indicates the label of the vertex being occupied by walker i . Given a configuration $\mathbf{a} = (a_1, \dots, a_w)$, there exists an edge between $\mathbf{a}$ and all configurations $\mathbf{b} = (b_1, \dots, b_w)$, such that $\text{dist}(a_i, b_i) = 1$. Thus, any configuration has 4^w neighbors, and the relationship of being neighbors is symmetric. As in the case of the cycle, the dynamic process can be seen as a random walk on the graph of configurations and thus as a Markov chain $\mathcal{M}(N, w, d)$.

For N even, given any two configurations $\mathbf{a}$ and $\mathbf{b}$, we say that they have the *same parity* if, for all i and j , $(a_{i,x} - a_{j,x}) + (a_{i,y} - a_{j,y}) \equiv (b_{i,x} - b_{j,x}) + (b_{i,y} - b_{j,y}) \pmod{2}$. With this definition of parity, Lemma 3.2 and its consequences also apply to the grid. Then, if N is odd, $\mathcal{M}(N, w, d)$ is ergodic, and if N is even, there are 2^{w-1} closed classes of states, where each class consists of all configurations with the same parity. The Markov chain restricted to any of these classes of states is irreducible and positive recurrent, but also 2-periodic, so we do not have ergodicity.

OBSERVATION 4.6. *Using the same argument as in Observation 3.3, for any fixed t , we can consider $\mathcal{W}_t(T_N)$ as a static $\mathcal{W}(T_N)$.*

In view of this observation and Theorem 1.5, we assume $\mu = \Theta(1)$ for the remainder of the section. This covers the nontrivial dynamic situations where $\mathcal{W}(G, w, d)$ is neither a.a.s. disconnected nor a.a.s. connected. Furthermore, in this case we need only to focus on the study of simple components. We also assume for the present subsection that $d \geq 2$. The case $d = 1$ is excluded for technical reasons.

We define $X = X(t)$ to be the random variable that counts the number of simple components at time step t . Given our assumptions about μ , for t in any fixed bounded time interval, $X(t)$ is asymptotically Poisson with expectation $\mu = \Theta(1)$, as studied in Proposition 4.2.

In analogy with d -hole lines in section 3.2, we define a *simple component line* to be a maximal sequence of pairs $(v_1, t_1), \dots, (v_l, t_l)$, where v_i is a simple component existing at time step t_i for $1 \leq i \leq l$, and such that $t_i = t_{i-1} + 1$ and the vertex v_i is adjacent to v_{i-1} , for $2 \leq i \leq l$. Birth, death, and survival of lines and the random

variables $B(t)$, $D(t)$, and $S(t)$ are defined analogously to the cycle case. For the case that $d = 1$, the concept of simple component lines is not quite adequate to describe what is going on unless $\varrho \rightarrow 0$, since otherwise it is reasonably likely that a vertex is occupied by several walkers that in one step jump to several different neighbors and create several component lines simultaneously. For this reason we exclude $d = 1$ at this point, though it is very easy to obtain most of the results claimed if $\varrho \rightarrow 0$. Presumably a suitable analysis of the simultaneous creation or destruction of several simple component lines gives results even for $\varrho \not\rightarrow 0$, but we have not done this.

We need one more definition referred to in section 1. Let v and v' be any two adjacent vertices in the grid. Let b be the number of directed edges whose origin is a vertex at distance strictly greater than d from v and whose destination is at distance at most d from v' . Note that this quantity b does not depend on the particular v and v' but just on parameter d . We have that $b = \Theta(d)$, but the exact expression of this b depends on the particular chosen metrics. Some examples are found in Table 2.

TABLE 2
Parameter b .

Metrics	b
ℓ^1	$b = 2d + 1/2$
ℓ^p ($p < \infty$)	$b \sim (1/\sqrt[p]{2} + 3/2) d$, if $d \rightarrow \infty$
ℓ^∞	$b = 3d + 1$

We next have a result analogous to Proposition 3.4.

PROPOSITION 4.7. *Let $d > 1$. For t in any fixed bounded time interval, the random variables $S(t)$, $B(t)$, and $D(t)$ are asymptotically independent Poisson, with the expectations*

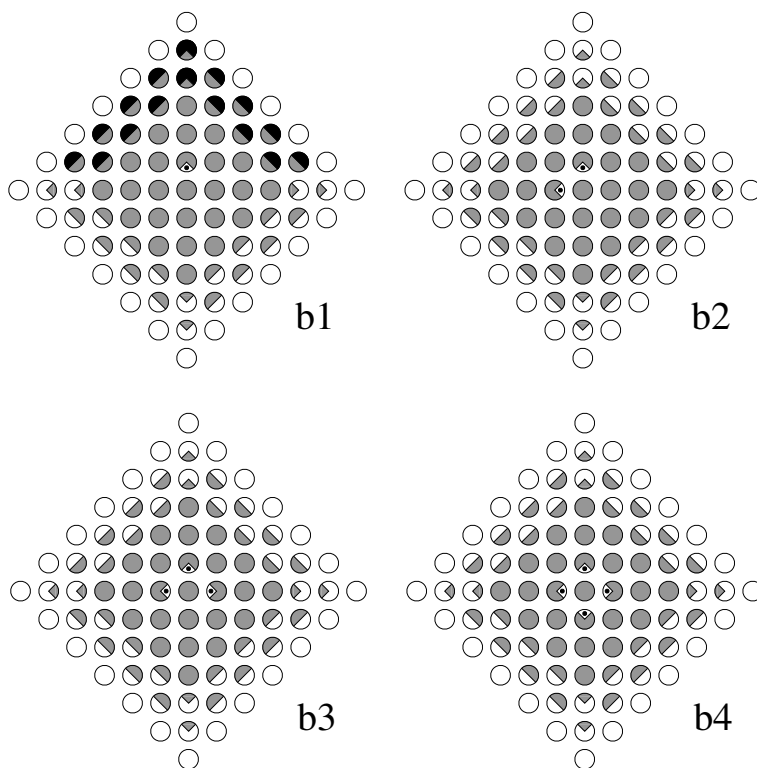
$$\mathbf{E}[S(t)] \sim \begin{cases} \mu & \text{if } d\varrho \rightarrow 0, \\ \mu - \lambda & \text{if } d\varrho \rightarrow c, \\ 4 \frac{1-e^{-\varrho/4}}{1-e^{-\varrho}} e^{-(b+3/4)\varrho} \mu & \text{if } d\varrho \rightarrow \infty, \end{cases}$$

$$\mathbf{E}[B(t)] = \mathbf{E}[D(t)] \sim \begin{cases} b\varrho\mu & \text{if } d\varrho \rightarrow 0, \\ \lambda & \text{if } d\varrho \rightarrow c, \\ \mu & \text{if } d\varrho \rightarrow \infty, \end{cases}$$

where $\lambda = (1 - e^{-b\varrho}) \mu$ as in Theorem 1.7.

Proof. For a vertex $v \in V$ with coordinates (x, y) , let v_N, v_S, v_E, v_W denote the vertices with coordinates $(x, y + 1), (x, y - 1), (x + 1, y), (x - 1, y)$. The proof is so similar to that of Proposition 3.4 that we discuss in detail only the ways that a simple component line can be born at v between time steps t and $t + 1$. We classify these events in 4 main classes, as shown in Figure 5.

- b1. At time step t , just one vertex of v_N, v_S, v_E, v_W is occupied and belongs to some bigger component. Then, all walkers there jump to v and stop communicating with other walkers.
- b2. At time step t , just two vertices of v_N, v_S, v_E, v_W are occupied. Then, all walkers there jump to v and do not communicate with other walkers.
- b3. At time step t , just three vertices of v_N, v_S, v_E, v_W are occupied. Then, all walkers there jump to v and do not communicate with other walkers.

FIG. 5. Birth of a simple component line at vertex v .

b4. At time step t , vertices v_N, v_S, v_E, v_W are occupied. Then, all walkers there jump to v and do not communicate with other walkers.

The first three classes have several subtypes (four for b1, six for b2, and four for b3) which correspond to choosing a subset of the four possible compass directions.

The rest of the proof follows that of Proposition 3.4. (See [DPSW06] for the full details.) $\square$

From Proposition 4.7, we can easily derive important consequences analogous to those of the cycle, always under the assumption stated after Observation 4.6. The first one gives us the probability that $\mathcal{W}_t(T_N)$ is connected but $\mathcal{W}_{t+1}(T_N)$ is disconnected. Note that this is a.s. equivalent to having no simple components at time step t and just one at $t+1$. The proof is immediate by the same argument as for Lemma 3.9.

LEMMA 4.8. *Let $d > 1$. The probability that $\mathcal{W}_t(T_N)$ is connected and that $\mathcal{W}_{t+1}(T_N)$ is disconnected is given by*

$$\mathbf{P}(X(t+1) \geq 1 \wedge X(t) = 0) \sim \begin{cases} \mu e^{-\mu} b \varrho & \text{if } d\varrho \rightarrow 0, \\ e^{-\mu}(1 - e^{-\lambda}) & \text{if } d\varrho \rightarrow c, \\ e^{-\mu}(1 - e^{-\mu}) & \text{if } d\varrho \rightarrow \infty, \end{cases}$$

where λ is defined as in Theorem 1.7.

In a similar way to the cycle case, we define the *lifespan* of a simple component line as the number of time steps for which the line is alive. For any vertex v and time t , the random variable $L_{v,t}$ counts the lifespan of the simple component line born at

vertex v between time steps t and $t + 1$. If this birth does not take place, $L_{v,t}$ is defined to be 0.

We finish by considering lifespans and connected periods. Define L_{av} (the average lifespan of simple component lines), L_T , and L as in section 3.2 but in terms of the new definition of $L_{v,t}$. Similarly, define LC_{av} to be the average length of the period from time t for which the graph of walkers on the grid is connected. We state the following theorem without proof because it is not difficult to adapt the argument used for Theorems 1.3, 1.4, 3.6, and 3.8. The proof is entirely straightforward in the case that L_{av} is bounded (i.e., $d\rho \not\rightarrow 0$). If L_{av} is unbounded, then there are two modifications to the argument. First, one needs to do the moment calculations more carefully (this is analogous to the situation at the end of the proof of Proposition 3.4; there is less independence when d is large, so this takes a bit of care). Second, for connectivity issues the observation that a complex component exists at any one time with probability tending to 0 does not show that we can ignore them, since the lifespan is no longer bounded. However, we may easily adapt the analysis to consider the creation or destruction of components (not just simple ones) in any given step to obtain the required result. We omit these details; they can be found in [P07].

THEOREM 4.9. *For the walkers model on T_N with $d > 1$,*

$$L_{\text{av}} \sim \begin{cases} \frac{1}{b\rho} & \text{if } d\rho \rightarrow 0, \\ \frac{\mu}{\lambda} & \text{if } d\rho \rightarrow c, \\ 1 & \text{if } d\rho \rightarrow \infty, \end{cases} \quad \text{and} \quad LC_{\text{av}} \sim \begin{cases} \frac{1}{\mu b\rho} & \text{if } d\rho \rightarrow 0, \\ \frac{1}{1-e^{-\lambda}} & \text{if } d\rho \rightarrow c, \\ \frac{1}{1-e^{-\mu}} & \text{if } d\rho \rightarrow \infty, \end{cases}$$

where λ is defined as in Theorem 1.7. Furthermore, L_T (LC_T) converges in probability for T growing large (N fixed) toward a random variable L (LC , respectively), uniquely determined almost everywhere, where $L \sim L_{\text{av}}$ and $LC \sim LC_{\text{av}}$ a.a.s.

The proof of Theorem 1.7 is a simple adjustment of the proof of Theorem 4.9 (as in the cycle case).

5. Conclusions and open problems. In this work we have characterized connectivity issues of a very large set of moving agents which move through a prescribed real or virtual graph. We believe it is the first time that these kinds of characterizations have been obtained, and it could open an interesting line of research. We gave characterizations for the cycle and the grid. The results obtained for the grid could easily be extended to the grid with diagonals. Also, an approach similar to ours should work with the k -dimensional toroidal grid, but a suitable substitute for the Geometric Lemma needs to be found.

Our results are based on results for the static problem, and the most difficult was to show that, for the crucial values of w and d , there is a.a.s. just one big component apart from simple ones. There is a corresponding argument in Penrose [P97], [P99] for the related problem in which w walkers are placed at random in a unit box in continuous Euclidean space (or the toroidal counterpart), and walkers of distance at most r can communicate directly. Our setting has one more parameter, but it is clear that the behavior for our problem should be similar to the continuous one with $r = d/n$, where our toroidal grid is $n \times n$. However, proving this equivalence would not be entirely straightforward in the case that d is small (constant or near-constant), since then the discrete ball of radius d is not well approximated by the continuous one.

In our model we use a fixed number w of walkers. One could alternatively place walkers randomly so that each vertex is occupied independently with probability p .

For example, they could be Poisson-distributed at each vertex (as suggested by an anonymous referee) with parameter λ such that $1 - e^{-\lambda} = p$. In the static case this alternative would bear the same relation to our model as the relation between the random graph models $G(n, m)$ and $G(n, p)$, and as for that case, one would expect similar properties when p is approximately w/N (or, more precisely, $1 - e^{-w/N}$, to capture the case that w is close to or greater than N). This would simplify some of our analysis (e.g., the proof of Lemma 2.1). However, it would be difficult to deduce all the results for our model in such a way. For one thing, some of the properties we study are not convex in the required sense (see [JLR00]). There are also other obstacles to using models with independent occupancy probabilities. For instance, if $w(s_0/N)^2 \neq o(1)$, then $(1 - (S_0/N))^w$, as in Lemma 2.1, is not asymptotic to $e^{-S_0 e}$.

Further planned work is the extension of the results to random geometric graphs, which will provide a model for omnidirectional radio communication in the setting described in [GHSZ]. Another interesting extension is the hypercube, as the number of neighbors of a vertex is then not constant. A further project is to study the connectivity of walkers when the underlying topology has obstacles which can interfere with communication.

REFERENCES

- [AGE02] P. K. AGARWAL, L. GUIBAS, H. EDELSBRUNNER, J. ERICKSON, M. ISARD, S. HARPELED, J. HERSHBERGER, C. JENSEN, L. KAVRAKI, P. KOEHL, M. LIN, D. MANOCHA, D. METAXAS, B. MIRTICH, AND D. MOUNT, *Algorithmic issues in modeling motion*, ACM Computing Surveys, 34 (2002), pp. 550–572.
- [Bol02] B. BOLLOBÁS, *Random Graphs*, 2nd ed., Academic Press, London, 2002.
- [DPSW05] J. DÍAZ, X. PÉREZ, M. SERNA, AND N. C. WORMALD, *Connectivity for wireless agents moving on a cycle or grid*, in Proceedings of STACS 2005, Lecture Notes in Comput. Sci. 3404, Springer-Verlag, Berlin, 2005, pp. 353–364.
- [DPSW06] J. DÍAZ, X. PÉREZ, M. SERNA, AND N. C. WORMALD, *Walkers on the Cycle and the Grid*, Research Report CORR 2006-09, Faculty of Mathematics, University of Waterloo, Waterloo, Canada, 2006.
- [GHSZ] L. GUIBAS, J. HERSHBERGER, S. SURI, AND LI ZHANG, *Kinetic Connectivity for Unit Disks*, Discrete Comput. Geom., 25 (2001), pp. 591–610.
- [H06] R. HEKMAT, *Ad-hoc Networks: Fundamental Properties and Network Topologies*, Springer-Verlag, New York, 2006.
- [H80] L. HOLST, *On multiple covering of a circle with random arcs*, J. Appl. Probab., 16 (1980), pp. 284–290.
- [JLR00] S. JANSON, T. ŁUCZAK, AND A. RUCIŃSKI, *Random Graphs*, Wiley, New York, 2000.
- [JBAS03] A. JARDOSH, E. BELDING-ROYER, K. ALMEROTH, AND S. SURI, *Towards realistic mobility models for mobile ad-hoc networks*, in Proceedings of ACM Mobicom, San Diego, CA, 2003, pp. 217–229.
- [P97] M. D. PENROSE, *The longest edge of the random minimal spanning tree*, Ann. Appl. Probab., 7 (1997), pp. 340–361.
- [P99] M. D. PENROSE, *On k-connectivity for a geometric random graph*, Random Structures Algorithms, 15 (1999), pp. 145–164.
- [P07] X. PÉREZ, *Aspects of Random Graphs*, Doctoral Dissertation, Universitat Politècnica de Catalunya, Barcelona, Spain, 2007.
- [RMM01] E. M. ROYER, P. M. MELLIAR-SMITH, AND L. E. MOSER, *An analysis of the optimum node density for ad hoc mobile networks*, in Proceedings of the IEEE International Conference on Communications, IEEE Communications Society, Piscataway, NJ, 2001, pp. 857–861.
- [W04] N. C. WORMALD, *Random graphs and asymptotics*, in Handbook of Graph Theory, J. L. Gross and J. Yellen, eds., CRC, Boca Raton, FL, 2004, pp. 817–836.
- [XK06] F. XUE AND P. R. KUMAR, *Scaling Laws for Ad Hoc Wireless Networks: An Information Theoretic Approach*, NOW Publishers, Delft, The Netherlands, 2006.

COMBINATORIAL PROPERTIES OF A ROOTED GRAPH POLYNOMIAL*

DAVID EISENSTAT[†], GARY GORDON[‡], AND AMANDA REDLICH[§]

Abstract. For a rooted graph G , let $EV(G; p)$ be the expected number of vertices reachable from the root when each edge has an independent probability p of operating successfully. We examine combinatorial properties of this polynomial, proving that G is k -edge connected if and only if $EV'(G; 1) = \dots = EV^{(k-1)}(G; 1) = 0$. We find bounds on the first and second derivatives of $EV(G; p)$; applications yield characterizations of rooted paths and cycles in terms of the polynomial. We prove reconstruction results for rooted trees and a negative result concerning reconstruction of more complicated rooted graphs. We also prove that the norm of the largest root of $EV(G; p)$ in $\mathbb{Q}[i]$ gives a sharp lower bound on the number of vertices of G .

Key words. expected rank, probabilistic graph

AMS subject classification. 90B15

DOI. 10.1137/06065951X

1. Introduction. Graph polynomials have a long history, dating to Birkhoff's use of the chromatic polynomial in an (unsuccessful) attempt to prove the four color theorem [7]. Two other polynomials, the reliability polynomial [12] and the 2-variable Tutte polynomial [8], also encode combinatorial data about the graph (the Tutte polynomial specializes to both the chromatic and reliability polynomials). While the original motivation for the study of these invariants is still important, much of the current interest in the Tutte polynomial is not related to any of its applications. See [9, 14] for some recent combinatorial applications.

It is in this spirit that we continue the study of the expected value polynomial $EV(G; p)$ applied to a rooted graph G , i.e., a graph with a distinguished vertex. The polynomial was introduced in [1] and [2], extended to antimatroids in [18], and applied to rooted graphs in [5, 19]. A closely related polynomial, called *pair connected reliability* by Amin, Siegrist, and Slater in [3, 4] and Siegrist in [21, 22] and *network resilience* by Colbourn in [13], is motivated by the reliability polynomial. A similar polynomial has also been defined for (nonrooted) graphs [6, 23].

In this paper, we concentrate on combinatorial properties of the rooted graph and their connection to the polynomial. In the sequel paper [15], we turn to applications, including practical questions about optimal location for the root for a given graph, randomness, and estimation.

Section 2 is concerned with bounds on the first and second derivatives of $EV(G; p)$. Applications of these bounds to the edge connectivity of the graph and graph reconstruction are given. The main result is Theorem 2.7.

THEOREM 2.7. *If G is a rooted graph, then G is k -edge connected if and only if $EV'(G; 1) = \dots = EV^{(k-1)}(G; 1) = 0$.*

*Received by the editors May 10, 2006; accepted for publication (in revised form) December 20, 2007; published electronically April 18, 2008. This work was supported by NSF grant DMS-0243763. <http://www.siam.org/journals/sidma/22-2/65951.html>

[†]Department of Comp. Sci., Princeton University, 35 Olden Street, Princeton, NJ 08540-5233 (deisenst@cs.princeton.edu).

[‡]Department of Mathematics, Lafayette College, Easton, PA 18042-1781 (gordong@lafayette.edu).

[§]Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139 (aredlich@mit.edu).

Section 3 examines the behavior of $EV(G; p)$ under standard graph-theoretic constructions. We give limits on the possibility of reconstructing G from $EV(G; p)$. When G is a rooted cycle or a rooted path, reconstruction is possible; in almost all other cases, it is not possible.

THEOREM 3.4.

- (1) $EV(G; p) = p + \cdots + p^n$ if and only if G is isomorphic to the rooted path P_{n-1} .
- (2) $EV(G; p) = 2p + \cdots + 2p^{n-1} + (n-1)p^n$ if and only if G is isomorphic to the rooted cycle C_n .

Conversely, Theorem 3.1 shows that *any* rooted graph is a subgraph of another rooted graph with a linear expected value polynomial. In a slightly different direction, Theorem 3.9 shows that rooted trees can be reconstructed from a *family* of expected rank polynomials.

In section 4 we give connections between the maximum norm of the zeros of $EV(G; p)$ in $\mathbb{C}$ and the number of vertices of G . This section is motivated by the study of the roots of the chromatic polynomial, which has connections to statistical physics [10, 20]. Our main result is Theorem 4.3.

THEOREM 4.3. *Let G be a connected rooted graph with $n > 1$ vertices. Suppose that the polynomial $EV(G; r) = 0$ for some $r \in \mathbb{Q}[i]$. Then $|r - 1| \leq n - 1$.*

While many of the proofs given here are straightforward (especially those concerning derivatives of the 1-variable polynomial), we believe the results are of sufficient interest to warrant further study. These results show that the polynomial encodes meaningful information about the rooted graph, but we also place bounds on how successful such an approach can be (Theorem 3.1).

2. Derivatives and edge connectivity. Let G be a connected rooted graph with edge set E , where each edge has the same independent probability p of being operational. We give two equivalent formulations of the expected rank polynomial, both of which will be important throughout this work. For a nonroot vertex v , let $Pr(v)$ denote the probability that v remains connected to the root. The following result appears explicitly in [5] and implicitly in [13] and [3, 4].

DEFINITION 2.1 (Proposition 2.7 of [5]). *Let G be a rooted graph, and let V be all the nonroot vertices of G . Then*

$$EV(G; p) = \sum_{v \in V} Pr(v).$$

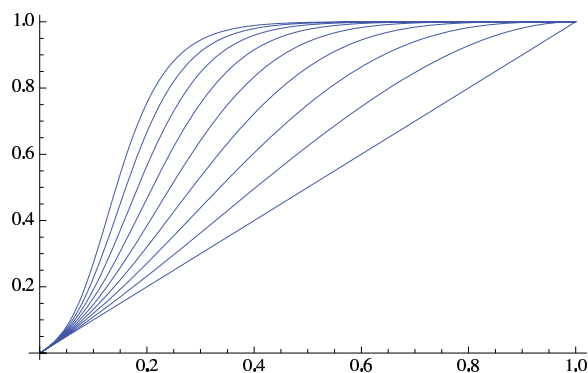
We will also need the following deletion-contraction expansion for the polynomial.

PROPOSITION 2.2 (deletion-contraction: Proposition 2.3 of [5]). *Let G be a rooted graph, and let e ($\neq$ loop) be an edge adjacent to the root. Then*

$$EV(G; p) = (1 - p) \cdot EV(G - e; p) + p \cdot EV(G/e; p) + p.$$

We will use the deletion-contraction formula of Proposition 2.2 repeatedly in this section. Throughout this section, we will assume an edge e incident to the root is not a loop (loops have no effect on the polynomial).

Example 2.3. We consider the expected rank polynomial $EV(K_n; p)$ for the rooted complete graph. This example is important since any random graph can be thought of as a subgraph of K_n . Obtaining closed form expressions for $EV(K_n; p)$ is difficult; see [4] for one approach. When applying deletion-contraction to K_n , multiple edges arise, and this gives rise to a recursive formula for $EV(K_n; p)$. (Replacing e by k

FIG. 1. Graphs of $EV(K_n; p)/(n-1)$ for $n = 2, \dots, 10$.

multiple edges can be thought of as increasing the probability that e succeeds from p to $1 - q^k$, where $q = 1 - p$.) If G' is obtained from the rooted graph G by replacing every edge of G by k parallel edges, then $EV(G') = EV(G; 1 - q^k)$. The proof of the next proposition follows from this observation.

PROPOSITION 2.4. *Let $q = 1 - p$. Then*

$$EV(K_n; p) = \sum_{k=1}^{n-1} \binom{n-1}{k} p^k q^{n-1-k} (EV(K_{n-k}; 1 - q^k) + k).$$

For fixed $p > 0$, we have $EV(K_n; p)/(n-1) \rightarrow 1$ as $n \rightarrow \infty$. This follows from a famous result of Erdős and Rényi [17]. If $p \gg \log n/n$, then the probability that G is connected approaches 1 as $n \rightarrow \infty$. Graphs of $EV(K_n; p)/(n-1)$ for $2 \leq n \leq 10$ appear in Figure 1.

We now turn our attention to the derivative $EV'(G; p)$, which is closely related to the connectivity of the rooted graph G . We begin by deriving sharp bounds on the size of $EV'(G; p)$ and $EV''(G; p)$.

We omit the straightforward proof of the lemma.

LEMMA 2.5. *Let G be a rooted graph with an edge e incident to the root. Then for all $p \in [0, 1]$, $1 + EV(G/e; p) \geq EV(G - e; p)$, with equality possible only when $p = 1$.*

It is easy to see that $EV'(G; p) \geq 0$ for all $0 \leq p \leq 1$: increasing p increases the expected number of vertices reachable from the root, so $EV(G; p)$ is an increasing function. We now give sharp upper and lower bounds on $EV'(G; p)$ and an upper bound on the second derivative $|EV''(G; p)|$.

PROPOSITION 2.6. *Let G be a rooted graph with n edges. Then for all $p \in [0, 1]$, the following hold:*

- (1) $EV'(G; p) \geq 0$. This inequality is strict if G has an edge incident to the root and $p < 1$.
- (2) $EV'(G; p) \leq n(n+1)/2$.
- (3) $|EV''(G; p)| \leq (n-1)n(n+1)/3$.

Proof. We prove (2)—the proofs of (1) and (3) are similar. We proceed by induction on n , and the base case $n = 0$ is trivial. When $n > 0$, we differentiate the formula 2.2:

$$\begin{aligned} EV'(G; p) &= (1 + EV(G/e; p)) - EV(G - e; p) \\ &\quad + pEV'(G/e; p) + (1 - p)EV'(G - e; p). \end{aligned}$$

If G has no edge e incident to the root, then $EV'(G; p) = 0$. Otherwise, we examine each term in this formula.

First, note that $1 + EV(G/e; p) \leq n$ since $EV(G/e; 1) = n - 1$ and $EV(G; p)$ is an increasing function. Also, $EV(G - e; p) \geq 0$ is clear. Finally,

$$pEV'(G/e; p) + (1 - p)EV'(G - e; p) \leq (n - 1)n/2$$

by induction. Putting the pieces together gives $EV'(G; p) \leq n(n + 1)/2$. $\square$

Notes.

1. The lower bound in Proposition 2.6(1) is sharp for all 2-edge connected graphs at $p = 1$. For the rooted cycle C_n ,

$$EV'(C_n; p) = \sum_{k=1}^{n-1} 2kp^{k-1} - n(n-1)p^{n-1},$$

so $EV'(C_n; 1) = 0$.

2. The upper bound in Proposition 2.6(2) is also sharp for rooted paths with n edges. If P_{n+1} denotes the rooted path on n edges with the root located at a leaf, then

$$EV(P_{n+1}; p) = \sum_{k=1}^n p^k,$$

so $EV'(P_{n+1}; 1) = \sum_{k=1}^n k = n(n + 1)/2$. The converse is also true when $p = 1$: If $EV'(G; 1) = n(n + 1)/2$, then G is a rooted path (Lemma 3.3(1)).

3. Note that it is not possible to bound $EV'(G; p)$ in terms of the number of vertices: if G is a graph with two vertices joined by k edges, then $EV(G; p) = 1 - (1 - p)^k$, so $EV'(G; 0) = k$.
4. The bounds of Proposition 2.6(3) are sharp for paths (upper bound) and cycles (lower bound) with n edges (P_{n+1} and C_n), again at $p = 1$. We will also prove a converse for the lower bound: If $EV''(G; 1) = -(n - 1)n(n + 1)/3$, then G is the rooted cycle C_n (Lemma 3.3(2)).

It is possible to derive bounds for even higher derivatives in this fashion, but these bounds will not be sharp, in general, with the exception of the lower bound for the third derivative. Further, the expected value polynomials of paths and cycles do not have maximal and minimal derivatives of all orders at $p = 1$.

Note that the second derivative of the deletion-contraction formula in Proposition 2.2 simplifies when $p = 1$:

$$EV''(G; 1) = 2EV'(G/e; 1) - 2EV'(G - e; 1) + EV''(G/e; 1).$$

This allows us to rewrite the upper bound in Proposition 2.6(3) in terms of the number of vertices of G , since G/e has one fewer vertex than G .

Recall that G is k -edge connected if removing fewer than k edges from G cannot disconnect G . The next result shows that k -edge connectivity is determined by $EV(G; p)$.

THEOREM 2.7. *If G is a rooted graph, then G is k -edge connected if and only if $EV'(G; 1) = \dots = EV^{(k-1)}(G; 1) = 0$.*

Proof. From Definition 2.1, we have

$$EV(G; p) = \sum_{v \in V} Pr(v),$$

where $Pr(v)$ is the probability that v is connected to the root. Fix v and let $S_1, \dots, S_n$ be the minimal subsets of E that, when removed, disconnect v from the root. Let $F(S)$ be the probability that all edges in S fail. Then we can compute $Pr(v)$ in terms of $F(S)$ via inclusion-exclusion:

$$Pr(v) = 1 - F(S_1) - \dots - F(S_n) + F(S_1 \cup S_2) + \dots \pm F(S_1 \cup \dots \cup S_n).$$

Now $F(S_i) = (1-p)^{|S_i|}$, so $Pr(v) = 1 + \sum_{j=1}^m a_j(1-p)^j$. Clearly, if k is the size of the smallest S_i , then $a_1 = \dots = a_{k-1} = 0$ and $a_k \neq 0$ (in fact, we must have $a_k < 0$). Finally, summing over all vertices gives the result. $\square$

As a quick check, note that $EV'(T; 1) > 0$ for any tree having $n > 0$ edges, so Theorem 2.7 shows that any tree is 1-edge connected. For the cycle C_n , we have $EV'(C_n; 1) = 0$ (see the remarks immediately following Proposition 2.6), but $EV''(C_n; 1) = -(n-1)n(n+1)/3$, so the theorem gives a verification that cycles are 2-edge connected.

When G is not rooted, Proposition 2.2 of [6] shows that $EV'(G; 1) = 0$ if and only if G is connected, where $EV(G; p)$ is defined using the matroid (cycle) rank function. In this case, the value of $|EV'(G; 1)|$ is just the number of isthmuses of G . Thus, Theorem 2.7 is a rooted generalization of this result.

3. Reconstructing graphs and an embedding theorem. $EV(G; p)$ is defined via connectivity; thus, it is not surprising that graph reconstruction is not possible using this invariant. The next theorem, one of the main results of this section, is a negative reconstruction result.

THEOREM 3.1. *Let G be a rooted graph. Then there is a rooted graph G' such that G is an induced subgraph of G' and $EV(G') = kp$ for some positive integer k .*

Proof. Let $EV(G; p) = a_1p + a_2p^2 + \dots + a_np^n$, where $a_i \in \mathbb{Z}$ and $a_n \neq 0$. We first show how to find a graph H_1 with the property that $EV(G \oplus H_1; p)$ has degree less than n . We then iterate this procedure, eventually producing a graph $G' = G \oplus H_1 \oplus \dots \oplus H_{n-1}$ so that $EV(G'; p)$ is a linear polynomial.

Case 1. $a_n < 0$. Let H_1 be the direct sum of a_n copies of the path P_{n+1} , the path with n edges, with each path rooted at a vertex of degree 1. Since $EV(P_{n+1}; p) = \sum_{k=1}^n p^k$, we find that the degree of $EV(G \oplus H_1)$ is at most $n-1$.

Case 2. $a_n > 0$. First attach k n -cycles to the root of G , where $k(n-1) > a_n$, and call the new graph G_1 . Now $EV(G_1; p) = b_np^n + \dots$ has degree n , and $b_n = a_n - k(n-1) < 0$, by construction. Now proceed as in Case 1.

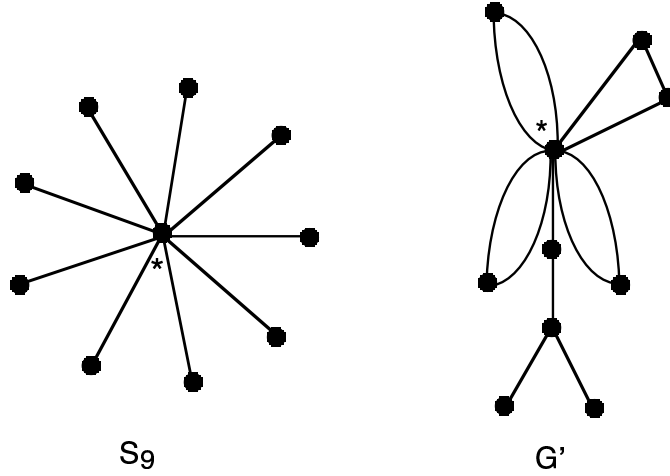
Now iterate this procedure to produce rooted graphs $H_2, H_3, \dots$ so that the degree of $EV(G \oplus H_1 \oplus \dots \oplus H_k)$ is at most $n-k$. This process will terminate when $k = n-1$. $\square$

Example 3.2. We apply (a variation of) the procedure described in the proof of Theorem 3.1 to the rooted cycle C_3 . First, note that $EV(C_3; p) = 2p + 2p^2 - 2p^3$. We attach a tree H_1 with $EV(H_1; p) = p + p^2 + 2p^3$, as in Figure 2. This gives $EV(C_3 \oplus H_1; p) = 3p + 3p^2$.

Now let $H_2 = C_2 \oplus C_2 \oplus C_2$, so $EV(H_2; p) = 6p - 3p^2$. Thus $EV(G'; p) = 9p$, where $G' = C_3 \oplus H_1 \oplus H_2$.

For example, if we let A_k denote the rooted graph formed by k parallel edges, then $EV(C_2 \oplus C_2 \oplus C_2 \oplus C_2; p) = EV(A_3 \oplus A_3 \oplus C_3; p) = 8p - 4p^2$.

Given Theorem 3.1, it is quite easy to construct nonisomorphic graphs with the same expected value polynomial. However, for certain classes of graphs, G can be uniquely reconstructed from $EV(G; p)$. We now show how rooted cycles are completely

FIG. 2. $EV(S_9; p) = EV(G'; p) = 9p$.

determined (within the class of all rooted graphs) by these polynomials. Recall that P_{n+1} denotes the rooted path with n edges.

LEMMA 3.3. *Let G be a rooted graph with n edges.*

- (1) $EV'(G; 1) = (n+1)n/2$ if and only if G is isomorphic to P_{n+1} .
- (2) $EV''(G; 1) = -(n+1)n(n-1)/3$ if and only if G is isomorphic to C_n .

Proof. We prove (1) and leave the similar proof of (2) to the reader. From the remarks following the proof of Proposition 2.6(2), $EV'(P_{n+1}; 1) = n(n+1)/2$.

For the converse, we use induction. If G has one edge, then there is nothing to prove. Now suppose $n > 1$ and $EV'(G; 1) = (n+1)n/2$. If G has no edges incident to the root, then $EV'(G; 1) = 0$, so we may assume that e is incident to the root. Then, as in the proof of Proposition 2.6(2), we have $EV'(G; 1) = 1 + EV(G/e; 1) - EV(G - e; 1) + EV'(G/e; 1)$.

Now $EV(G/e; 1) \leq n-1$ (since $G - e$ has $n-1$ edges) and $EV(G - e; 1) \geq 0$. Thus,

$$EV'(G/e; 1) = EV'(G; 1) - EV(G/e; 1) + EV(G - e; 1) - 1,$$

which gives $EV'(G/e; 1) \geq (n+1)n/2 - (n-1) - 1 = n(n-1)/2$. By Proposition 2.6(2), we have $EV'(G/e; 1) = n(n-1)/2$, which forces each of the inequalities given above to be equalities. Thus, $EV(G - e; 1) = 0$, so e is the only edge incident to the root of G , and $EV(G/e; 1) = n-1$, so G/e is connected. Furthermore, since $EV'(G/e; 1) = n(n-1)/2$, we have that G/e is isomorphic to the path P_n , by induction.

Now we have a rooted graph G with exactly one edge e incident to the root such that G/e is the path P_n . This forces G to be the path P_{n+1} . $\square$

An immediate consequence of Lemma 3.3 is the following.

THEOREM 3.4.

- (1) $EV(G; p) = p + \cdots + p^n$ if and only if G is isomorphic to the rooted path P_{n-1} .
- (2) $EV(G; p) = 2p + \cdots + 2p^{n-1} + (n-1)p^n$ if and only if G is isomorphic to the rooted cycle C_n .

Obviously, for any positive integer k , it is possible to produce k nonisomorphic rooted trees, all sharing the same expected value polynomial. On the other hand,

it is possible to uniquely reconstruct a rooted tree from a *sequence* of expected rank polynomials. We begin with a definition.

DEFINITION 3.5. *Let G be a rooted graph. Then the expected rank k polynomial is defined by*

$$EV_k(G; p) = \sum_{A \subseteq E: r(A)=k} p^{|A|} (1-p)^{|E|-|A|},$$

where $r(A)$ is the number of vertices in the component of A containing the root (not including the root).

Thus, $EV_k(G; p)$ is the probability that exactly k vertices are connected to the root. The proof of the next proposition is immediate.

PROPOSITION 3.6. *Let G be a rooted graph with $n+1$ vertices. Then*

- (1) $\sum_{k=1}^n EV_k(G; p) = 1$, and
- (2) $EV(G; p) = \sum_{k=1}^n k \cdot EV_k(G; p)$.

To keep track of this sequence of expected rank k polynomials, it is convenient to introduce a 2-variable generating function.

DEFINITION 3.7. *Let T be a rooted tree, and let $X(T)$ be a tree with a single edge e adjacent to the root such that $X(T)/e = T$. Then define $F(T; p, q)$ recursively as follows:*

$$F(\bullet) = 1,$$

$$F(X(T)) = q + pF(T),$$

$$F(T_1 \oplus T_2) = F(T_1)F(T_2).$$

The connection between the generating function $F(T; p, q)$ and the sequence of rank k expected rank polynomials $EV_k(T; p)$ is made explicit in the next proposition.

PROPOSITION 3.8. *$F(T; p, q)$ is uniquely recoverable from the sequence of polynomials $EV_0, \dots, EV_n$, where EV_k is the probability that exactly k vertices are connected to the root.*

Proof. Let $EV_k(p) = p^k g_k(p)$. Then set $q = 1 - p$, so $p = 1 - q$, and it is easy to show that $F(T) = \sum_k p^k g_k(1 - q)$. Thus, we can recover $F(T)$ from the sequence of polynomials, and this operation is easily invertible. $\square$

We now prove that rooted trees can be uniquely reconstructed from their sequence of rank k expected rank polynomials $\{EV_0, \dots, EV_n\}$.

THEOREM 3.9. *Let T be a rooted tree. Then $F(T(p, q))$ uniquely determines T up to isomorphism.*

Proof. It suffices to show that for all T , $F(X(T))$ is irreducible over $\mathbf{Z}[p, q]$. The result then follows by induction: if $F(T; p, q)$ factors, we reconstruct the rooted trees corresponding to the factors inductively. If $F(T; p, q)$ is irreducible, we will have $T = X(T')$ for some rooted tree T' , and $F(T'; p, q) = p^{-1}(F(T; p, q) - q)$, so we can reconstruct T' (and hence, T) inductively again.

Now write $F(T) = q + pG(p, q)$ for some polynomial $G(p, q)$, and suppose that $F(X(T)) = AB$. Then $F(X(T)) = (1 + pA')(q + pB')$ for some 2-variable polynomials A' and B' . If $q \mid B'$, then $q \mid (q + pB')$, so $q \mid F(X(T))$, which cannot be the case since $F(X(T))$ has exactly one pure p term, p^n , corresponding to all n edges operating successfully.

Hence q does not divide B' , and we let cp^α with $c \neq 0$ be the pure p term in B' of lowest degree. Then $F(X(T))$ contains a term cp^α that cannot be canceled by $p^2 A' B'$. As a result, $A' = 0$, and the factorization is trivial. $\square$

The sequence $\{EV_0, \dots, EV_n\}$ is equivalent to the (greedoid) Tutte polynomial of a rooted tree, which encodes information about the number of rooted subtrees of size k with exactly l leaves. More information about rooted tree reconstruction from this version of the Tutte polynomial can be found in [2, 11]. (Unrooted tree reconstruction is not possible in general—see [16].)

4. Zeros of the polynomial. Proposition 3.1 of [19] shows that the largest rational root of $EV(G; p)$ is a lower bound on the number of vertices (including the root). We generalize this result now, extending the bound to the absolute value of the largest rational root.

PROPOSITION 4.1. *Let G be a connected rooted graph with $n > 1$ vertices. Suppose the polynomial $EV(G; r) = 0$ for some $r \in \mathbb{Q}$. Then $|r - 1| \leq n - 1$.*

Proof. Let $f(p) = EV(G; p + 1)$. Then we can write

$$f(p) = (a - bp)g(p),$$

where $a/b = r - 1$ and $g \in \mathbb{Z}[p]$. $f(0) = a \cdot g(0) = n - 1$, so $a \mid (n - 1)$; and since $|a/b| \leq |a|$, $|r - 1| \leq n - 1$. $\square$

We can extend this result further, to the case where the root $r \in \mathbb{Q}[i]$.

PROPOSITION 4.2. *Let G be a connected rooted graph with $n > 1$ vertices. Suppose the polynomial $EV(G; r) = 0$ for some $r \in \mathbb{Q}[i] - \mathbb{Q}$. Then $|r - 1|^2 \leq n - 1$.*

Proof. Again, let $f(p) = EV(G; p + 1)$, and write $(a + bi)/c = r - 1$ for integers a, b , and c . Then

$$f(p) = (a^2 + b^2 - 2a(cp) + (cp)^2)g(p),$$

where $g(p) \in \mathbb{Z}[p]$. Since $n - 1 = f(0) = (a^2 + b^2)g(0)$ and $a^2 + b^2 \mid n - 1$, we have $(a^2 + b^2) = |c(r - 1)|^2 \geq |r - 1|^2$, so $|r - 1|^2 \leq (n - 1)$. $\square$

Putting Propositions 4.1 and 4.2 together gives the following.

THEOREM 4.3. *Let G be a connected rooted graph with $n > 1$ vertices. Suppose that the polynomial $EV(G; r) = 0$ for some $r \in \mathbb{Q}[i]$. Then $|r - 1| \leq n - 1$.*

All of these bounds are sharp. For the rational roots of Proposition 4.1, let G_1 be a graph with one vertex connected to the root by two edges and k vertices connected by one edge. Then $EV(G_1; p) = (k + 2)p - p^2$, which has a root at $p = k + 2$. For the lower bound, if we let G_2 be a tree with polynomial $kp + p^2$, then $EV(G_2; p)$ has a root at $p = -k$.

For the imaginary rational roots of Proposition 4.2, let $a^2 + b^2 = c^2$ be a Pythagorean triple, and let T be a tree with polynomial $((a - 1)^2 + b^2)p + 2(a - 1)p^2 + p^3$. Then $EV(T; p)$ has roots at $p = 1 - a \pm bi$, and T has $((a - 1) + 1)^2 + b^2 + 1 = c^2 + 1$ vertices and $|1 - a \pm bi - 1| = c$.

The next result is applicable to any polynomial $f(p)$ with positive integer coefficients and $f(0) = 0$.

PROPOSITION 4.4. *Let T be a tree with $n > 1$ vertices. Suppose the polynomial $EV(T; z) = 0$ for some $z \in \mathbb{C}$. Then $|z| \leq n - 2$.*

Proof. $EV(T; p) = n_1p + n_2p^2 + \dots + n_kp^k$ for positive integers n_j . Let $C = n_1 + \dots + n_{k-1}$. Then $C \leq n - 2$, since $n_1 + \dots + n_k = n - 1$.

When $C = 0$, $EV(T; p) = n_kp^k$, which has zeros only at $p = 0$. Otherwise, we can assume that $C \geq 1$. Suppose to the contrary that $|z| > C$. Then

$$|-n_kz^k| > C|z^{k-1}| > n_1|z| + \dots + n_{k-1}|z^{k-1}| > |n_1z + \dots + n_{k-1}z^{k-1}|,$$

and z is clearly not a zero. $\square$

When T is a tree, Proposition 4.4 allows us to drop the restriction that $r \in \mathbb{Q}[i]$.

COROLLARY 4.5. *Let T be a tree with $n > 1$ vertices. Suppose the polynomial $EV(T; p)$ has a zero at $p = z \in \mathbb{C}$. Then $|z - 1| \leq n - 1$.*

Unfortunately, this bound does not extend to all graphs and all complex zeros. For example, let $G = K_4 \oplus T$, where T is a tree such that $EV(T; p) = p + p^2 + p^3 + p^4 + p^5 + 5p^6$. Then $EV(G; p) = 4p + 7p^2 + p^3 - 20p^4 + 22p^5 - p^6$, which has a zero near $p = 21$. G , however, has only 14 vertices and 16 edges.

Similar constructions work with larger complete graphs, where we attach the smallest tree that will make the leading coefficient of $EV(G; p)$ equal to -1 .

It would be interesting to determine what other restrictions exist on zeros of the polynomial. This is similar to much of the current research on the chromatic polynomial of a graph [10, 20].

Acknowledgments. We thank Jennifer Feder and Greg Francos for useful discussions and the anonymous referee for several suggestions.

REFERENCES

- [1] M. AIVALIOTIS, *A Probabilistic Approach to Network Reliability in Graph Theory*, Honors thesis, Lafayette College, Easton, PA, 1998.
- [2] M. AIVALIOTIS, G. GORDON, AND W. GRAVEMAN, *When bad things happen to good trees*, J. Graph Theory, 37 (2001), pp. 79–99.
- [3] A. AMIN, K. SIEGRIST, AND P. SLATER, *Pair-connected reliability of a tree and its distance degree sequences*, Congr. Numer., 58 (1987), pp. 29–42.
- [4] A. AMIN, K. SIEGRIST, AND P. SLATER, *Exact formulas for reliability measures for various classes of graphs*, Congr. Numer., 58 (1987), pp. 43–52.
- [5] A. BAILEY, G. GORDON, M. PATTON, AND J. SCANCELLA, *Expected value expansions in rooted graphs*, Discrete Appl. Math., 128 (2003), pp. 555–571.
- [6] J. BENASHSKI, R. MARTIN, J. MOORE, AND L. TRALDI, *On the β -invariant for graphs*, Congr. Numer., 109 (1995), pp. 211–221.
- [7] G. D. BIRKHOFF, *A determinant formula for the number of ways of coloring a map*, Ann. of Math. (2), 14 (1912/1913), pp. 42–46.
- [8] T. BRYLAWSKI AND J. OXLEY, *The Tutte polynomial and its applications*, in Matroid Applications, Vol. 40, N. White, ed., Cambridge University Press, Cambridge, UK, 1992, pp. 123–225.
- [9] J. BONIN AND W. MILLER, *Characterizing combinatorial geometries by numerical invariants*, European J. Combin., 20 (1999), pp. 713–724.
- [10] J. BROWN, C. HICKMAN, A. SOKAL, AND D. WAGNER, *On the chromatic roots of generalized theta graphs*, J. Combin. Theory Ser. B, 83 (2001), pp. 272–297.
- [11] S. CHAUDHARY AND G. GORDON, *Tutte polynomials for trees*, J. Graph Theory, 15 (1991), pp. 317–331.
- [12] C. COLBOURN, *The Combinatorics of Network Reliability*, Oxford University Press, Oxford, UK, 1987.
- [13] C. COLBOURN, *Network resilience*, SIAM J. Algebraic Discrete Methods, 8 (1987), pp. 404–409.
- [14] A. DE MIER AND M. NOY, *On graphs determined by their Tutte polynomials*, Graphs Combin., 20 (2004), pp. 105–119.
- [15] D. EISENSTAT, J. FEDER, G. FRANCOS, G. GORDON, AND A. REDLICH, *Optimality and randomness in rooted graphs*, Discrete Appl. Math., 156 (2008), pp. 746–756.
- [16] D. EISENSTAT AND G. GORDON, *Non-isomorphic caterpillars with identical subtree data*, Discrete Math., 306 (2006), pp. 827–830.
- [17] P. ERDÖS AND A. RÉNYI, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl., 5 (1960), pp. 17–61.
- [18] G. GORDON, *Expected rank in antimatroids*, Adv. in Applied Math., 32 (2004), pp. 299–318.
- [19] G. GORDON AND E. JAGER, *Roots for rooted graph polynomials*, Networks, to appear.
- [20] R. SHROCK, *Chromatic polynomials and their zeros and asymptotic limits for families of graphs*, Discrete Math., 231 (2001), pp. 421–446.

- [21] K. SIEGRIST, *Expected value expansions in random subgraphs with applications to network reliability*, *Combin. Probab. Comput.*, 7 (1993), pp. 465–483.
- [22] K. SIEGRIST, A. AMIN, AND P. SLATER, *The optimal unicyclic graphs for pair-connected reliability*, *Discrete Appl. Math.*, 41 (1993), pp. 235–243.
- [23] L. TRALDI, *A note on reliability and expected value*, *Congr. Numer.*, 133 (1998), pp. 95–99.

TESTING TRIANGLE-FREENESS IN GENERAL GRAPHS*

NOGA ALON[†], TALİ KAUFMAN[‡], MICHAEL KRIVELEVICH[§], AND DANA RON[¶]

Abstract. In this paper we consider the problem of testing whether a graph is triangle-free and, more generally, whether it is H -free, for a fixed subgraph H . The algorithm should accept graphs that are triangle-free and reject graphs that are far from being triangle-free in the sense that a constant fraction of the edges should be removed in order to obtain a triangle-free graph. The algorithm is allowed a small probability of error. This problem has been studied quite extensively in the past, but the focus was on dense graphs, that is, when $d = \Theta(n)$, where d is the average degree in the graph and n is the number of vertices. Here we study the complexity of the problem in general graphs, that is, for varying d . In this model a testing algorithm is allowed to ask neighbor queries (i.e., “What is the i th neighbor of vertex v ?”), vertex-pair queries (i.e., “Is there an edge between vertices v and u ?”), and degree queries (i.e., “What is the degree of vertex v ?”). Our main finding is a lower bound of $\Omega(n^{1/3})$ on the necessary number of queries that holds for every $d < n^{1-\nu(n)}$, where $\nu(n) = o(1)$. Since when $d = \Theta(n)$ the number of queries sufficient for testing has been known to be independent of n , we observe an abrupt, *threshold-like* behavior of the complexity of testing around n . This lower bound holds for testing H -freeness of every nonbipartite subgraph H . Additionally, we provide sublinear upper bounds for testing triangle-freeness that are at most quadratic in the stated lower bounds, and we describe a transformation from certain one-sided error lower bounds for testing subgraph-freeness to two-sided error lower bounds. Finally, in the course of our analysis we show that dense random Cayley graphs behave like quasi-random graphs in the sense that relatively large subsets of vertices have the “correct” edge density. The result for subsets of this size cannot be obtained from the known spectral techniques that only supply such estimates for much larger subsets.

Key words. property testing, triangle-freeness, subgraph-freeness, randomized algorithms, approximation algorithms, Cayley graphs

AMS subject classifications. 68Q25, 68W20, 68W25, 68W40, 68Q17, 68R05, 68R10, 05C85, 05C25

DOI. 10.1137/07067917X

1. Introduction. *Property testing* is the study of the following type of computational tasks: Let P be some predetermined property, where we shall be interested in properties of a combinatorial nature. The task is to distinguish quickly and reliably between input objects possessing property P and input objects that are “far”

*Received by the editors January 4, 2007; accepted for publication (in revised form) November 24, 2007; published electronically April 18, 2008. An extended abstract of this paper appeared in the Proceedings of SODA, 2006.

<http://www.siam.org/journals/sidma/22-2/67917.html>

[†]Department of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel, and Institute for Advanced Study, Princeton, NJ 08540 (nogaa@post.tau.ac.il). This author’s research was supported in part by a USA Israel BSF grant, by a grant from the Israel Science Foundation, and by the Von Neumann Fund.

[‡]Computer Science and Artificial Intelligence Lab, MIT, Cambridge, MA 02138 (kaufmant@mit.edu). This work is part of the author’s Ph.D. thesis done at Tel Aviv University under the supervision of Prof. Noga Alon and Prof. Michael Krivelevich. This research was done in part while the author visited the Radcliffe Institute for Advanced Study, Harvard University.

[§]Department of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel (krivelev@post.tau.ac.il). This author’s research was supported in part by a USA Israeli BSF grant and by a grant from the Israel Science Foundation.

[¶]Department of Electrical Engineering-Systems, Tel Aviv University, Tel Aviv 69978, Israel (danar@eng.tau.ac.il). This research was done in part while the author was a fellow at the Radcliffe Institute for Advanced Study, Harvard University, and was supported in part by a grant from the Israel Science Foundation.

from having P , where distance is measured in some appropriately defined quantitative sense. In order to fulfill this task, the testing algorithm is given *query access* to a description of the input object O , adhering to some pre-agreed upon format; the algorithm's complexity is measured by the number of queries it asks before reaching a reliable decision. A testing algorithm is expected to query a sublinear portion of the input; hence, randomness plays an essential role in developing property testing algorithms. Property testing was defined by Rubinfeld and Sudan [18], who focused on testing algebraic properties. Goldreich, Goldwasser, and Ron [9] initiated the study of combinatorial property testing and in particular considered testing properties of graphs. Since then, the field has progressed enormously, with many papers devoted to it, where one of the main focuses is on testing graph properties. The interested reader is invited to consult the surveys [8, 17] for more details.

In this work we consider the problem of testing subgraph-freeness, and in particular triangle-freeness, in general graphs. Let n denote the number of vertices in the graph, let d denote the average degree, and let $d_{\max}$ denote the maximum degree. Given a distance parameter $\epsilon > 0$, we would like to design an algorithm that distinguishes with high probability between the case that the graph contains no triangles and the case in which more than $\epsilon \cdot dn$ edges should be removed so that no triangles remain. To this end we allow the algorithm query access to the graph. In particular, for any vertex of its choice, the algorithm may ask for the degree of the vertex, it may ask to get the i th neighbor of the vertex for every $i \leq n$ (if the vertex has less than i neighbors, then a null answer is returned), and it may ask whether there is an edge between any two vertices.¹

Subgraph-freeness, and more specifically, triangle-freeness, is one of the most basic problems studied in property testing. The interest in this problem is both due to the fact that triangle-freeness is a fundamental and simple graph property and due to the relation between triangle-freeness and the study of dense sets of integers with no three-term arithmetic progression.

Dense graphs. Most of the focus in previous works was on testing triangle-freeness in dense graphs, that is, when $d = \Theta(n)$. For this class of graphs the most appropriate input graph representation is the graph adjacency matrix, and a testing algorithm is allowed to query whether (u, v) is an edge of an input graph G , where $u, v \in V(G)$ (the so-called *vertex-pair* queries). The authors of [2] showed that it is possible to test triangle-freeness in dense graphs using a number of queries that are *independent* of n and have a tower-type behavior in $1/\epsilon$. Alon [1] proved that a superpolynomial dependence on $1/\epsilon$ is necessary for testing the subgraph-freeness of all nonbipartite subgraphs. When the fixed subgraph is bipartite, then $O(1/\epsilon)$ queries suffice [1]. It is also observed in [1] (and much earlier, though implicitly, in [19]) that the problem of testing triangle-freeness is intimately related to the famous (and very hard) problem of the existence of dense sets of integers without a three-term arithmetic progression. Alon's lower bound, which was proved for one-sided error algorithms, was extended in [6] to two-sided error algorithms. Other related results include [5].

Bounded-degree graphs. At the other extreme, an input graph is assumed to have its maximum degree bounded by an absolute constant $d_{\max} = O(1)$. In such a case, the input graph is usually represented by an array of incidence lists of all of its vertices

¹Clearly, a degree query to a vertex v can be implemented by performing at most $\log n$ queries concerning the neighbors of v . Therefore, an algorithm can do without degree queries and incur at most a logarithmic factor overhead. For simplicity of the presentation, we allow the algorithm to perform these queries.

(of length at most $d_{\max}$ each); accordingly, a testing algorithm queries the i th neighbor of a vertex v , where $1 \leq i \leq d_{\max}$ (the so-called *neighbor* queries). As was observed in [10], in this case $O(1/\epsilon)$ queries suffice for testing triangle-freeness. More generally, $O(d^\tau/\epsilon)$ queries suffice for testing H -freeness in graphs with maximum degree $O(d)$, where τ is the diameter of H .

General graphs. In this work we study the complexity of testing triangle-freeness of graphs that lie between the two extremes. Namely, we would like to understand the dependence of the query complexity on the average degree d , and we do not want to necessarily assume that $d_{\max} = O(d)$. In the latter aspect we follow the work [15] on testing the diameter of sparse, but unbounded-degree, graphs, and in both aspects we follow the work [14] on testing bipartiteness of general graphs. In particular, as in [14] we allow queries of both types—vertex-pair queries and neighbor queries—as well as degree queries (i.e., the algorithm may query the degree of any vertex). Note that the fact that the graph has varying degrees makes the task of testing triangle-freeness significantly harder. Consider, for example, sparse graphs, that is, graphs with average degree $d = O(1)$. As we mentioned before, when $d_{\max} = O(1)$, $O(1/\epsilon)$ queries suffice for testing triangle-freeness. However, our work shows that when $d_{\max} = \Theta(n)$ and $d = O(1)$, the number of queries required for testing triangle-freeness is $\Omega(\sqrt{n})$.

Our contributions. The main contributions of this paper, on a qualitative level, are as follows:

- We discover a threshold-type behavior in testing triangle-freeness: whenever $d = O(n^{1-\nu(n)})$, where $\nu(n)$ is a function that satisfies $\nu(n) = o(1)$, the number of queries necessary to test triangle-freeness is $\Omega(n^{1/3})$, while, as discussed above, for $d = \Theta(n)$ the query complexity is a function of ϵ only. This is in sharp contrast to the results of [14], where a smooth behavior of the complexity of testing bipartiteness as a function of d was described.
- We provide a transformation from lower bounds for testing triangle-freeness using one-sided error algorithms to those for two-sided error algorithms; though the suggested transformation is stated and proved for triangles and carries some technical restrictions, it is general enough to capture a variety of lower bounds of this sort.
- We give quantitative lower and upper bounds for testing triangle-freeness in general graphs.
- We show that the edge distribution in random Cayley graphs is close to that of truly random graphs of the same edge density. This is proven by direct combinatorial and probabilistic arguments, without relying on the eigenvalue machinery, which is incapable of proving such results for subsets that are too small. Although we need this result for property testing purposes, we feel it is of enough independent interest to be stated here.

1.1. A lower bound and a sharp threshold. Our main result is the following theorem.

THEOREM 1. *There is a lower bound of $\Omega(n^{1/3})$ for testing triangle-freeness in general graphs. The lower bound holds for algorithms that are allowed two-sided error, and for every d that is upper bounded by $O(n^{1-\nu(n)})$, where $\nu(n) = \frac{\log \log \log n + 4}{\log \log n}$. For some values of d the lower bound reaches $\Omega(n^{1/2})$.*

Theorem 1 is actually the union of three lower bounds (whose one-sided error versions are stated in Lemmas 1, 2, and 6), which are applied to different values of d .

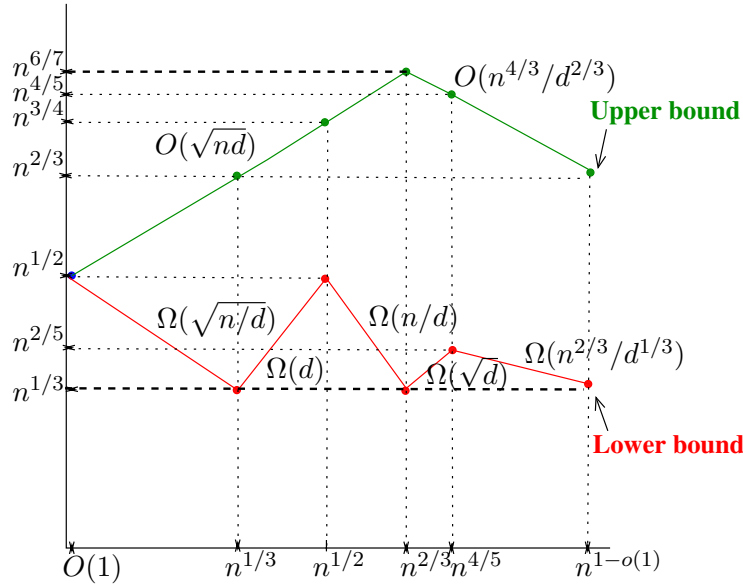


FIG. 1. A schematic illustration of our lower and upper bounds. The x -axis represents the average degree d , and the y -axis represents the bounds. For the sake of simplicity we ignore logarithmic factors in the bounds. Notice that the lower bound lies entirely above the horizontal line at height $n^{1/3}$, and the upper bound lies entirely below the horizontal line at height $n^{6/7}$.

The exact expression for the lower bound is

$$(1) \quad \Omega\left(\max\left\{\sqrt{n/d}, \min\{d, n/d\}, \min\left\{\sqrt{d}, n^{2/3}/d^{1/3}\right\} \cdot n^{-o(1)}\right\}\right).$$

For a schematic illustration see Figure 1.

Recall that when $d = \Theta(n)$ then testing can be performed using a number of queries that is independent of n [2]. Thus we observe a sharp transition between our lower bound of $\Omega(n^{1/3})$ that holds until $d = n^{1-\nu(n)}$ (recall that $\nu(n) = o(1)$) and the query complexity at $d = \Theta(n)$, which does not depend on n . The abrupt change of testing complexity around linear density can be partially explained by the fact that the celebrated regularity lemma of Szemerédi (whose relevance to testing triangle-freeness has been first indicated in [19] and was made explicit in [2]) starts yielding meaningful results for graphs whose number of edges is almost quadratic in the number of vertices. The exact behavior of the complexity of testing triangle-freeness when $n^{1-\nu(n)} \leq d \leq n$ remains open.

Remark. Using techniques that were previously applied in [1] it is possible to extend Theorem 1 to testing subgraph-freeness for other nonbipartite subgraphs and to show that for any fixed nonbipartite graph H there is a constant $a = a(H) > 0$ such that an algorithm for testing H -freeness in graphs on n vertices with average degree $d \leq n^{1-o(1)}$ has to ask at least n^a queries. This lower bound holds for two-sided error algorithms, using both vertex-pair and neighbor queries. We wish to note that the difference between the complexity of testing subgraph-freeness of bipartite and nonbipartite subgraphs is caused by the difference in the behavior of their Turán numbers—they are subquadratic for the former and quadratic for the latter. A more detailed discussion can be found in [1]. The technical details of the aforementioned general statement concerning testing H -freeness and its proof are quite cumbersome,

while the threshold-type behavior of the query complexity as a function of average degree d is observed already for the case when H is a triangle. Therefore, we prefer to concentrate on the basic case of testing triangle-freeness.

1.2. Upper bounds. We show that for every graph density, there exists an algorithm for testing triangle-freeness whose query complexity is sublinear in n . Furthermore, the upper bound is always at most quadratic in the corresponding lower bound (up to a factor of $\log n$).

THEOREM 2. *There is an upper bound of $\tilde{O}(n^{6/7})$ for testing triangle-freeness in general graphs for every value of d . The upper bound can go down to $\tilde{O}(n^{1/2})$ for some values of d . In all cases, up to logarithmic factors, the upper bound is at most quadratic in the lower bound that holds for that density. If $d_{\max} = O(d)$, then the upper bound is $O(n^{4/5})$ for all values of d .*

The exact expression for our upper bound is $\tilde{O}(\min\{\sqrt{nd}/\epsilon^{3/2}, (n^{4/3}/d^{2/3})/\epsilon^2\})$, where, in the case that $d_{\max} = O(d)$, the first term is replaced by d/ϵ . For a schematic illustration, see Figure 1.

1.2.1. Tight results. There are two cases in which our lower and upper bounds essentially match. The first case is graphs in which $d_{\max} = O(d)$ and $d \leq \sqrt{n}$. For this case the complexity is $\Theta(d)$ (for constant ϵ). The second case is general sparse graphs, that is, graphs for which $d = \Theta(1)$. For these graphs the complexity is $\tilde{\Theta}(\sqrt{n})$.

1.3. Our techniques.

Behrend graphs and Cayley graphs. In the proof of our third lower bound (Lemma 6), we build on graphs that are known as Behrend graphs, which were previously used in the context of testing triangle-freeness in [1]. Here we prove that random Behrend graphs have a certain property that we can exploit in order to obtain our lower bound. Behrend graphs are variants of the well-studied Cayley graphs, and our proof concerning properties of random Behrend graphs extends to Cayley graphs.

Specifically, we show that for dense random Cayley graphs the edge density in relatively large induced subgraphs is close to the edge density of the whole graph. It was previously shown [4] that random Cayley graphs are expanders and hence have the property that the density of every induced subgraph on sufficiently many vertices is very close to the density of the graph. However, the known techniques for proving this property are based on estimating the second eigenvalue of the graph's adjacency matrix and do not supply any informative bounds for sets of vertices that are much smaller than the number of vertices divided by the square root of the degree. Our results for Cayley graphs apply for Cayley graphs over both Abelian and non-Abelian groups, while Behrend graphs were considered only in an Abelian setting. Our techniques are somewhat reminiscent of those of [3, 11].

A reduction from one-sided error lower bounds to two-sided error lower bounds. We obtain our two main lower bounds by first establishing lower bounds that hold for one-sided error algorithms. We then prove a transformation from one-sided error lower bounds to two-sided error lower bounds that holds under certain assumptions and apply it to obtain our two-sided error lower bounds. This transformation may be of use in future lower bound proofs for subgraph freeness. We note that in [6] a transformation was given in the case of dense graphs, but it is not applicable in general.

1.4. Subsequent work. In [12] and [16] our lower bound and upper bound, respectively, were improved somewhat for particular ranges of d . Specifically, Guggemann [12] slightly modifies one of our constructions and obtains a lower bound of

$\Omega(\min\{(nd)^{1/3}, n/d\})$, instead of our lower bound of $\Omega(\min\{d, n/d\})$. Given our additional lower bound of $\Omega(\sqrt{n/d})$, this improves on our result for d that ranges between $\Omega(n^{1/5})$ and $O(n^{1/2})$.

In terms of upper bounds, Rast [16] describes an algorithm that combines ideas from our algorithms in a nontrivial manner and obtains an algorithm whose query complexity (in terms of its dependence on n and d) is $O(\max\{(nd)^{4/9}, n^{2/3}/d^{1/3}\})$. This improves on our result for d that ranges between $\Omega(n^{2/5})$ and $O(n^{4/5})$.

1.5. Paper organization. After giving some preliminary definitions in section 2, we proceed to lower bounds in sections 3, 4, and 5. In particular, section 5 describes Behrend graphs and proves a result on the edge density of random Behrend graphs, whose more general version (applicable to random Cayley graphs of Abelian groups) is discussed in section 8. Section 6 describes a reduction from lower bounds for two-sided testers to those for one-sided testers. Section 7 contains upper bounds for testing triangle-freeness.

2. Preliminaries. Let $G = (V, E)$ be an undirected graph with n vertices labeled $1, \dots, n$, and let d denote the average degree in G , where we assume that $d = \Omega(1)$.² For each vertex $v \in V$ let $\deg(v)$ denote the degree of vertex v . The edges incident to v (and their endpoints, the neighbors of v) are labeled from 1 to $\deg(v)$. Note that each edge has two, possibly different, labels, one with respect to each of its endpoints. For a graph G and a subset of vertices $U \subseteq V$, we refer to the edges in the subgraph of G that is induced by U as the edges *spanned* by U in G .

A graph G is said to be *triangle-free* if for every three vertices u, v, w in G , at least one of the three vertex-pairs (u, v) , (v, w) , or (w, u) is not an edge in G . A graph G is ϵ -far from (being) *triangle-free* if it is necessary to remove more than ϵdn edges from G in order to obtain a triangle-free graph. We note that since the number of edges in the graph is $(dn)/2$, the standard definition of ϵ -far would be that more than $(\epsilon dn)/2$ edges should be removed so that the graph becomes triangle-free. In order to simplify the presentation we slightly modify the definition.

A testing algorithm for triangle-freeness is required to accept with probability at least $2/3$ every graph that is triangle-free and to reject with probability at least $2/3$ every graph that is ϵ -far from being triangle-free, where ϵ is a given distance parameter. If the algorithm always accepts triangle-free graphs, then it has *one-sided error*; otherwise, it has *two-sided error*. In order to perform this task the testing algorithm is allowed the following types of queries:

- *Degree* queries: For any vertex u of its choice, the algorithm can obtain $\deg(u)$.
- *Neighbor* queries: For any vertex u and index $1 \leq i \leq \deg(u)$, the algorithm may obtain the i th neighbor of vertex u .
- *Vertex-pair* queries: For any pair of vertices (u, v) , the algorithm can query whether there is an edge between u and v in G .

3. A lower bound of $\Omega(\sqrt{n/d})$. In this section we establish our first and simplest lower bound.

LEMMA 1. *Every algorithm for testing triangle-freeness must perform $\Omega(\sqrt{n/d})$ queries. This lower bound holds for two-sided error algorithms as well.*

²Our results can be extended to the case that $d = o(1)$ (that is, very sparse graphs). However, for the sake of simplicity, and since we believe that the very sparse case is of less interest, we assume that $d = \Omega(1)$.

Proof. In order to prove a two-sided error lower bound of $\Omega(q)$ queries for testing triangle-freeness, it suffices to describe two families of graphs for which the following two conditions hold: (1) The graphs in the first family are all triangle-free, while the graphs in the second family are all $\Theta(1)$ -far from being triangle-free. (2) Any algorithm that distinguishes with constant probability between a graph selected uniformly in one family, and a graph selected uniformly in the second family, must perform $\Omega(q)$ queries.

In particular, consider the following two families of graphs over n vertices and with average degree d . Each family is determined by a single graph and consists of all possible $n!$ labelings of the vertices of the graph. Hence it suffices to describe the two graphs (one per family). In one graph there is a complete bipartite graph between two sets of vertices, each of size $\sqrt{nd}/2$. In the other graph there is a clique of size $\sqrt{nd}$. (We ignore rounding issues here.) In addition, in both graphs the remaining vertices are isolated. The first graph is clearly triangle-free, and it is not hard to verify that the second graph is $\Theta(1)$ -far from being triangle-free. However, in order to distinguish between the two graphs (or more precisely, in order to distinguish between graphs that are selected uniformly from each of the two families), the algorithm must obtain a vertex in the complete bipartite subgraph/clique. To this end the algorithm must perform $\Omega(n/\sqrt{nd}) = \Omega(\sqrt{n/d})$ queries. $\square$

4. A lower bound of $\Omega(\min\{d, n/d\})$. Our next lower bound improves on the lower bound in section 3 when $d > n^{1/3}$.

LEMMA 2. *Every one-sided error algorithm for testing triangle-freeness must perform $\Omega(\min\{d, n/d\})$ queries. This lower bound holds even when $d_{\max} = O(d)$.*

The lower bound in Lemma 2 is extended to two-sided error algorithms in section 6.

We prove Lemma 2 by describing a distribution on graphs such that the following hold: On one hand almost all of the support of the distribution is on graphs that are far from triangle-free. On the other hand, every algorithm that uses neighbor and/or vertex-pair queries must perform $\Omega(\min\{d, n/d\})$ queries before it views a triangle (with sufficiently high constant probability). Since we currently focus on testing algorithms that have one-sided error, this implies a lower bound on the query complexity of such algorithms.

We start by considering the case that $d = c \cdot \sqrt{n}$ for a particular constant $c < 1$. We later discuss how to deal with the case that $d > c \cdot \sqrt{n}$ and with the case that $d < c \cdot \sqrt{n}$.

4.1. Definition of the lower-bound distribution. The graphs we consider below are d -regular for $d = \Theta(\sqrt{n})$ but may have multiple edges. At the end of this section we discuss how to remove the multiple edges. Let D_Δ be a distribution over graphs with n vertices and degree $d = \frac{2}{3}\sqrt{n/3}$ that is defined as follows. A graph is generated by first partitioning the vertices into equal-size subsets of size $n' = n/3$ denoted V_1, V_2, V_3 . Next, between each pair of subsets, $d' = d/2 = \sqrt{n'}/3$ random perfect matchings are selected. In all that follows we assume that n' is sufficiently large ($n' > 100$ suffices).

LEMMA 3. *With probability $1 - o(1)$, a graph chosen uniformly according to the distribution D_Δ is $\Omega(1)$ -far from being triangle-free.*

Proof. We shall show that with high probability over the choice of a graph according to D_Δ , there are at least $\gamma \cdot nd$ edge-disjoint triangles in the graph for some constant γ .

Let the vertices in each set V_ℓ , $\ell \in \{1, 2, 3\}$, be denoted $v_{\ell,1}, \dots, v_{\ell,n'}$. For each pair of vertices $v_{\ell,i} \in V_\ell$, $v_{\ell',j} \in V_{\ell'}$, where $\ell \neq \ell'$, let $\alpha_{i,j}^{\ell,\ell'}$ be a 0/1 random variable that is 1 if there is an edge between $v_{\ell,i}$ and $v_{\ell',j}$ and is 0 otherwise. Then

$$(2) \quad \Pr \left[\alpha_{i,j}^{\ell,\ell'} = 1 \right] = 1 - \left(1 - \frac{1}{n'} \right)^{d'},$$

and so

$$(3) \quad \Pr \left[\alpha_{i,j}^{\ell,\ell'} = 1 \right] \leq \frac{d'}{n'}$$

and

$$(4) \quad \Pr \left[\alpha_{i,j}^{\ell,\ell'} = 1 \right] \geq \frac{d'}{n'} - \binom{d'}{2} \cdot \frac{1}{(n')^2} \geq \frac{d'}{n'} - \frac{1}{18n'} = \frac{d'}{n'} \cdot \left(1 - \frac{1}{18d'} \right).$$

Next consider any choice of three vertices $v_{1,i} \in V_1$, $v_{2,j} \in V_2$, and $v_{3,k} \in V_3$. Let $\Delta_{i,j,k}$ denote the 0/1 random variable that is 1 if and only if there is a triangle between the three vertices. Since the choice of the matchings between V_1 and V_2 is independent of the choice of the matchings between V_2 and V_3 and the choice of the matching between V_1 and V_3 ,

$$(5) \quad \Pr[\Delta_{i,j,k} = 1] \leq \left(\frac{d'}{n'} \right)^3$$

and

$$(6) \quad \Pr[\Delta_{i,j,k} = 1] \geq \left(\frac{d' - 1/18}{n'} \right)^3 = \left(\frac{d'}{n'} \right)^3 \cdot \left(1 - \frac{1}{18d'} \right)^3.$$

Let $\beta_{i,j,k}$ denote a 0/1 random variable that is 1 if and only if there is a triangle *different from* $(v_{1,i}, v_{2,j}, v_{3,k})$ that includes one of the edges $(v_{1,i}, v_{2,j})$ or $(v_{1,i}, v_{3,k})$ or $(v_{2,j}, v_{3,k})$. That is, there is a triangle of the form $(v_{1,i'}, v_{2,j}, v_{3,k})$, where $i' \neq i$, or $(v_{1,i}, v_{2,j'}, v_{3,k})$, where $j' \neq j$, or $(v_{1,i}, v_{2,j}, v_{3,k'})$, where $k' \neq k$. Since the probability of having an edge $(v_{\ell,r}, v_{\ell',s})$ conditioned on having the edge $(v_{\ell,r}, v_{\ell',s'})$ for any choice of $s' \neq s$ is at most $1 - (1 - \frac{1}{n'})^{d'-1}$,

$$\begin{aligned} \Pr[\Delta_{i,j,k} = 1 \ \& \ \beta_{i,j,k} = 1] &= \Pr[\Delta_{i,j,k} = 1] \cdot \Pr[\beta_{i,j,k} = 1 \mid \Delta_{i,j,k} = 1] \\ &\leq \left(\frac{d'}{n'} \right)^3 \cdot 3 \cdot (n' - 1) \cdot \left(\frac{d' - 1}{n'} \right)^2 \\ (7) \quad &\leq \left(\frac{d'}{n'} \right)^3 \cdot \frac{3(d' - 1)^2}{n'}. \end{aligned}$$

It follows that

$$\begin{aligned} \Pr[\Delta_{i,j,k} = 1 \ \& \ \beta_{i,j,k} = 0] &= \Pr[\Delta_{i,j,k} = 1] - \Pr[\Delta_{i,j,k} = 1 \ \& \ \beta_{i,j,k} = 1] \\ &\geq \left(\frac{d' - 1/18}{n'} \right)^3 - \left(\frac{d'}{n'} \right)^3 \cdot \frac{3(d' - 1)^2}{n'} \end{aligned}$$

$$\begin{aligned}
&= \left(\frac{d'}{n'}\right)^3 \cdot \left(\left(1 - \frac{1}{18d'}\right)^3 - 3 \cdot \left(1 - \frac{1}{d'}\right)^2 \cdot \frac{(d')^2}{n'} \right) \\
&\geq \frac{1}{2} \cdot \left(\frac{d'}{n'}\right)^3 \\
(8) \quad &= \frac{1}{54} \cdot (n')^{-3/2},
\end{aligned}$$

where we have used our assumption that n' and $d' = \sqrt{n'}/3$ are sufficiently large. Hence, the expected number of triples $(v_{1,i}, v_{2,j}, v_{3,k})$ that constitute a triangle that does not share an edge with any other triangle is at least

$$(9) \quad (n')^3 \cdot \frac{1}{54} \cdot (n')^{-3/2} = \frac{1}{54} \cdot (n')^{3/2}$$

$$(10) \quad = \frac{1}{54} \cdot n \cdot d.$$

Next, we shall use Chebyshev's inequality to show that with high probability, the number of such triples is not much smaller than this expected value.

For each triple $v_{1,i}, v_{2,j}, v_{3,k}$ let $\eta_{i,j,k}$ be the 0/1 random variable that is 1 if and only if $(v_{1,i}, v_{2,j}, v_{3,k})$ constitute a triangle, and there is no other triangle that shares any of its edges. Using our previous notation we have that

$$(11) \quad \Pr[\eta_{i,j,k} = 1] = \Pr[\Delta_{i,j,k} = 1 \ \& \ \beta_{i,j,k} = 0],$$

which implies (using (9)) that

$$(12) \quad \left(\text{Exp} \left[\sum_{i,j,k} \eta_{i,j,k} \right] \right)^2 \geq (n')^3 / c$$

for some constant $c > 1$. By Chebyshev's inequality,

$$\begin{aligned}
(13) \quad \Pr \left[\sum_{i,j,k} \eta_{i,j,k} < \frac{1}{2} \text{Exp} \left[\sum_{i,j,k} \eta_{i,j,k} \right] \right] &\leq \frac{4 \text{Var} \left[\sum_{i,j,k} \eta_{i,j,k} \right]}{\left(\text{Exp} \left[\sum_{i,j,k} \eta_{i,j,k} \right] \right)^2} \\
&< \frac{c' \text{Var} \left[\sum_{i,j,k} \eta_{i,j,k} \right]}{(n')^3}
\end{aligned}$$

for a constant $c' > 1$. We would like to bound the variance of $\sum_{i,j,k} \eta_{i,j,k}$:

$$(14) \quad \text{Var} \left[\sum_{i,j,k} \eta_{i,j,k} \right] = \text{Exp} \left[\left(\sum_{i,j,k} \eta_{i,j,k} \right)^2 \right] - \left(\text{Exp} \left[\sum_{i,j,k} \eta_{i,j,k} \right] \right)^2.$$

Now,

$$\text{Exp} \left[\left(\sum_{i,j,k} \eta_{i,j,k} \right)^2 \right] = \sum_{i,j,k} \sum_{i',j',k'} \text{Exp}[\eta_{i,j,k} \cdot \eta_{i',j',k'}]$$

$$\begin{aligned}
&= \sum_{i \neq i'} \sum_{j \neq j'} \sum_{k \neq k'} \text{Exp}[\eta_{i,j,k} \cdot \eta_{i',j',k'}] \\
&\quad + 3 \cdot \sum_i \sum_{j \neq j'} \sum_{k \neq k'} \text{Exp}[\eta_{i,j,k} \cdot \eta_{i,j',k'}] \\
&\quad + 3 \cdot \sum_{i,j} \sum_{k \neq k'} \text{Exp}[\eta_{i,j,k} \cdot \eta_{i,j,k'}] \\
&\quad + \sum_{i,j,k} \text{Exp}[\eta_{i,j,k}^2],
\end{aligned}
\tag{15}$$

while

$$\begin{aligned}
&\left(\text{Exp} \left[\sum_{i,j,k} \eta_{i,j,k} \right] \right)^2 = \sum_{i,j,k} \sum_{i',j',k'} \text{Exp}[\eta_{i,j,k}] \cdot \text{Exp}[\eta_{i',j',k'}] \\
&> \sum_{i \neq i'} \sum_{j \neq j'} \sum_{k \neq k'} \text{Exp}[\eta_{i,j,k}] \cdot \text{Exp}[\eta_{i',j',k'}].
\end{aligned}
\tag{16}$$

Therefore,

$$\begin{aligned}
\text{Var} \left[\sum_{i,j,k} \eta_{i,j,k} \right] &\leq \sum_{i \neq i'} \sum_{j \neq j'} \sum_{k \neq k'} \left(\text{Exp}[\eta_{i,j,k} \cdot \eta_{i',j',k'}] - \text{Exp}[\eta_{i,j,k}] \cdot \text{Exp}[\eta_{i',j',k'}] \right) \\
&\quad + 3 \cdot \sum_i \sum_{j \neq j'} \sum_{k \neq k'} \text{Exp}[\eta_{i,j,k} \cdot \eta_{i,j',k'}] \\
&\quad + 3 \cdot \sum_{i,j} \sum_{k \neq k'} \text{Exp}[\eta_{i,j,k} \cdot \eta_{i,j,k'}] \\
&\quad + \sum_{i,j,k} \text{Exp}[\eta_{i,j,k}^2].
\end{aligned}
\tag{17}$$

First, observe that, by definition of $\eta_{i,j,k}$ for every i, j and $k \neq k'$, since if both $(v_{1,i}, v_{2,j}, v_{3,k})$ and $(v_{1,i}, v_{2,j}, v_{3,k'})$ are triangles then they share an edge, we get that

$$\eta_{i,j,k} \cdot \eta_{i,j,k'} = 0.
\tag{18}$$

Next, observe that since $\Pr[\eta_{i,j,k} = 1] \leq \Pr[\Delta_{i,j,k} = 1]$, where by (5) $\Pr[\Delta_{i,j,k} = 1] \leq (d'/n')^3$,

$$\sum_{i,j,k} \text{Exp}[\eta_{i,j,k}^2] = \sum_{i,j,k} \text{Exp}[\eta_{i,j,k}] \leq (n')^3 \cdot \left(\frac{d'}{n'} \right)^3 = (d')^3.
\tag{19}$$

We turn to the sum over triangles that share a (single) vertex. Conditioned on there being a triangle $(v_{1,i}, v_{2,j}, v_{3,k})$, i.e., $\Delta_{i,j,k} = 1$, the probability of having the edge $(v_{1,i}, v_{2,j'})$ for $j' \neq j$, i.e., $\eta_{i,j}^{1,2} = 1$ (and similarly the edge $(v_{1,i}, v_{3,k'})$ for $k' \neq k$),

is upper bounded³ by $1 - (1 - \frac{1}{n'})^{d'-1} < \frac{d'}{n'}$. The probability of having the edge $(v_{2,j'}, v_{3,k'})$ (conditioned on the existence of the triangle $(v_{1,i}, v_{2,j}, v_{3,k})$) is upper bounded by $1 - (1 - \frac{1}{n-1})^{d'}$, which is at most $\frac{2d'}{n'}$ for $n' \geq 2$. Therefore,

$$\begin{aligned} \text{Exp}[\eta_{i,j,k} \cdot \eta_{i',j',k'}] &\leq \text{Exp}[\Delta_{i,j,k} \cdot \Delta_{i',j',k'}] \\ &= \Pr[\Delta_{i,j,k} = 1] \cdot \Pr[\Delta_{i',j',k'} = 1 \mid \Delta_{i,j,k} = 1] \\ (20) \quad &\leq 2 \left(\frac{d'}{n'} \right)^6, \end{aligned}$$

and so

$$(21) \quad \sum_i \sum_{j \neq j'} \sum_{k \neq k'} \text{Exp}[\eta_{i,j,k} \cdot \eta_{i',j',k'}] \leq (n')^5 \cdot \frac{2(d')^6}{(n')^6} = \frac{2(d')^6}{n'}.$$

To complete the proof we bound the difference between $\text{Exp}[\eta_{i,j,k} \cdot \eta_{i',j',k'}]$ and $\text{Exp}[\eta_{i,j,k}] \cdot \text{Exp}[\eta_{i',j',k'}]$ (for $i \neq i'$, $j \neq j'$, and $k \neq k'$). By definition,

$$\begin{aligned} &\text{Exp}[\eta_{i,j,k} \cdot \eta_{i',j',k'}] - \text{Exp}[\eta_{i,j,k}] \cdot \text{Exp}[\eta_{i',j',k'}] \\ (22) \quad &= \Pr[\eta_{i,j,k} = 1] \cdot (\Pr[\eta_{i',j',k'} = 1 \mid \eta_{i,j,k} = 1] - \Pr[\eta_{i',j',k'} = 1]). \end{aligned}$$

Since the proof of the following claim is a bit technical, we give it in the appendix.

CLAIM 4.

$$(23) \quad \Pr[\eta_{i',j',k'} = 1 \mid \eta_{i,j,k} = 1] = \left(1 + O\left(\frac{1}{d'}\right) \right) \cdot \Pr[\eta_{i',j',k'} = 1].$$

Since $\Pr[\eta_{i,j,k} = 1] = \Pr[\eta_{i',j',k'} = 1] \leq (\frac{d'}{n'})^3$, Claim 4 implies that

$$(24) \quad \text{Exp}[\eta_{i,j,k} \cdot \eta_{i',j',k'}] - \text{Exp}[\eta_{i,j,k}] \cdot \text{Exp}[\eta_{i',j',k'}] = O\left(\frac{(d')^5}{(n')^6}\right).$$

Hence,

$$\begin{aligned} \sum_{i \neq i'} \sum_{j \neq j'} \sum_{k \neq k'} \text{Exp}[\eta_{i,j,k} \cdot \eta_{i',j',k'}] - \text{Exp}[\eta_{i,j,k}] \cdot \text{Exp}[\eta_{i',j',k'}] &= O\left((n')^6 \cdot \frac{(d')^5}{(n')^6}\right) \\ (25) \quad &= O((d')^5). \end{aligned}$$

By combining (13) with (17)–(25), we get that

$$(26) \quad \Pr\left[\sum_{i,j,k} \eta_{i,j,k} < \frac{1}{2} \text{Exp}\left[\sum_{i,j,k} \eta_{i,j,k}\right]\right] = \frac{O((d')^5) + O((d')^3)}{(n')^3} = O\left(\frac{1}{d'}\right).$$

Because we have shown (see (10)) that $\text{Exp}[\sum_{i,j,k} \eta_{i,j,k}] = \Omega(nd)$, the lemma follows. $\square$

³The reason this is an upper bound and equality does not necessarily hold is that the triangle $(v_{1,i}, v_{2,j}, v_{3,k})$ may have parallel edges. That is, the event $\Delta_{i,j,k} = 1$ is a union of events in which each edge of the triangle may appear with different multiplicity.

The case $d > \frac{2}{3\sqrt{3}}\sqrt{n}$. The distribution in this case is the same as described at the start of this section: the graph vertices are partitioned into three equal parts V_1 , V_2 , and V_3 , and between every pair V_i and V_j we put $d/2$ random perfect matchings. We would like to show that with high probability the resulting graph contains at least $\frac{1}{c} \cdot nd$ edge-disjoint triangles in the graph for some constant c . To this end we think of the matchings as being selected in $k = \frac{3\sqrt{3}}{2} \cdot \frac{d}{\sqrt{n}}$ rounds, where in each round $d' = \frac{1}{3\sqrt{3}} \cdot \sqrt{n}$ matchings are selected between every pair $V_i, V_j, i \neq j$. For each round we can apply Lemma 3 and get that with high probability we have at least $\frac{1}{c} \cdot n \cdot d'$ edge-disjoint triangles. Observe that the triangles created in the different rounds are edge-disjoint. When $k = o(d')$ (i.e., $d = o(n)$), we can apply a union bound and get that with probability at least $1 - O(k/d') = 1 - o(1)$ we obtain $\frac{1}{c} \cdot n \cdot d$ edge-disjoint triangles. For larger k we can use the fact that the different rounds are independent. Therefore, with probability $1 - \exp(-k) = 1 - o(1)$ in at least $1/2$ of the rounds there are $\frac{1}{c} \cdot n \cdot d'$ edge-disjoint triangles, implying that there are at least $\frac{1}{2c} \cdot n \cdot d$ edge-disjoint triangles.

The case $d < \frac{2}{3\sqrt{3}}\sqrt{n}$. In this case we first partition the vertices into k parts $V^1, \dots, V^k$, where $|V^i| = \frac{27}{4}d^2$. We then apply the construction described at the start of this section to each V^i . In this case, too, for small k we can apply a union bound on the different V^i 's, and once k is sufficiently large we can use the fact that the different subgraphs are constructed independently. In either case we get that with high probability, for at least a half of the V^i 's there are $\Omega(|V^i| \cdot d)$ edge disjoint triangles within the subgraph induced by V^i .

4.2. The lower bound. Let A be any one-sided error algorithm for testing triangle-freeness, where A is allowed to perform both neighbor queries and vertex-pair queries. If A views a triangle in the tested graph, then clearly it can reject the graph. However, since A is a one-sided error algorithm, if it terminates before viewing a triangle, then it must accept. Suppose we run A on a graph chosen according to D_Δ , with some (sufficiently small) constant ϵ . Since we have shown that with high probability such a graph is ϵ -far from being triangle-free, the probability that A terminates before viewing a triangle must be small. Hence it remains to prove the following lemma.

LEMMA 5. *Any algorithm whose goal is to detect, with high constant probability, a triangle in a graph selected according to D_Δ must ask $\Omega(\min\{d, n/d\})$ queries.*

We note that the proof actually establishes the stronger statement: $\Omega(\min\{d, n/d\})$ queries are required to detect a cycle of any length.

Proof. We first present the argument for $d \geq \frac{2}{3\sqrt{3}}\sqrt{n}$ and later discuss the modifications required for $d < \frac{2}{3\sqrt{3}}\sqrt{n}$.

It will be convenient to view graphs in the support of D_Δ as being represented by “matchings over tables.” Namely, there are six tables: $T_{1,2}, T_{1,3}, T_{2,1}, T_{2,3}, T_{3,1}, T_{3,2}$, two for each set of vertices $V_b, b \in \{1, 2, 3\}$. Each table $T_{b,b'}$ is of size $(n/3) \times (d/2)$: there is a row for each vertex v in V_b , and each entry in v 's row corresponds to one of the $d/2$ edges that are incident to v and to vertices in $V_{b'}$. An edge between $u \in V_b$ and $v \in V_{b'}$ is represented by a pair of entries $(T_{b,b'}[u][i], T_{b',b}[v][j])$. Hence the $d/2$ perfect matchings between V_b and $V_{b'}$ correspond to a single perfect matching between the entries of the two tables $T_{b,b'}$ and $T_{b',b}$.

Let ALG be an algorithm that performs $Q = Q(n, d)$ queries and whose goal is to detect a triangle with probability at least $9/10$. The probability is taken over the choice of the graph G in the support of D_Δ and the coin flips of the algo-

rithm. Namely, ALG is a (possibly probabilistic) mapping from *query-answer histories* $\langle (q_1, a_1), \dots, (q_t, a_t) \rangle$ to q_{t+1} for every $t < Q$. A vertex-pair query is of the form $q_t = (u, v)$, where $u \in V_b$ and $v \in V_{b'}$ for some $b \neq b'$, $b, b' \in \{1, 2, 3\}$. The answer is either $a_t = (1, i, j)$, which denotes that there is an edge between u and v and it corresponds to the pair of entries $(T_{b,b'}[u][i], T_{b',b}[v][j])$, or $a_t = 0$, which denotes that there is no edge between u and v . A neighbor query is of the form $q_t = (u, b', i)$, where $u \in V_b$, $b' \neq b$. The answer is of the form $a_t = (v, j)$, where $v \in V_{b'}$, denoting that there is an edge between u and v and it corresponds to the pair of entries $(T_{b,b'}[u][i], T_{b',b}[v][j])$. We note that the mapping from query-answer histories to queries needs to be defined only on histories that are consistent with some graph in the support of D_Δ .

In what follows we define a process that answers the queries of the algorithm while generating a graph according to D_Δ . At any time t , the queries of the algorithm and the answers it is provided with determine the *knowledge graph* $G^t = (V^t, E^t, \bar{E}^t)$, where V^t are the vertices, E^t are the edges, and $\bar{E}^t$ are the nonedges. Namely, V^t consists of all vertices that appeared in queries of the algorithm or in answers to neighbor queries. Similarly E^t consists of all pairs u, v such that either (u, v) was a vertex-pair query that was answered positively or v was an answer to a neighbor query involving u , and $\bar{E}^t$ consists of all pairs u, v such that either (u, v) was a vertex-pair query that was answered negatively. For every edge $(u, v) \in E^t$ the knowledge graph will include the indices of the entries in the tables by which u and v are connected (that is, (i, j) such that $T_{b,b'}[u][i]$ is matched to $T_{b',b}[v][j]$).

Vertex-pair queries. Given a vertex-pair query $q_t = (u, v)$, where $u \in V_b$ and $v \in V_{b'}$, the process computes the probability, conditioned on the current knowledge graph, that (u, v) is an edge. Namely, it considers all graphs in the support of D_Δ that are consistent with G^{t-1} and answers positively with probability that is proportional to the number of these graphs in which there is an edge between u and v . Let $U_{t-1,b,b'}$ denote the number of unmatched entries in $T_{b,b'}$ at time t (before the t th query). Note that this number equals the number of unmatched entries in $T_{b',b}$ at the same time.

The conditional probability that there is an edge between u and v is upper bounded by

$$(27) \quad \frac{(d/2) \cdot (d/2)}{U_{t-1,b,b'} - |\bar{E}^{t-1}| \cdot (d/2) - (d/2)}.$$

To verify this consider an iterative process in which the (at most $d/2$) unmatched entries in u 's row in $T_{b,b'}$ are matched one by one to unmatched entries in $T_{b',b}$. After $0 \leq i < d/2$ steps, the probability of selecting any of the (at most $d/2$) unmatched entries in v 's row in $T_{b',b}$ is the current number of unmatched entries in v 's row, divided by $U_{t-1,b,b'} - i - |\bar{E}^{t-1}| \cdot (d/2)$. The term $U_{t-1,b,b'} - i$ is the number of unmatched entries after i steps, and the term $|\bar{E}^{t-1}| \cdot (d/2)$ is the maximum number of unmatched entries that cannot be matched to entries in u 's row because they correspond to vertices w such that $(u, w) \in \bar{E}^{t-1}$.

Since $U_{t-1,b,b'} \geq (n/3) \cdot (d/2) - (t-1)$, $|\bar{E}^{t-1}| \leq t-1$, and $t-1 < Q = o(n/d)$, the probability is $O(d/n)$. It follows that the probability that the algorithm gets a positive answer for *any* of the at most $Q = o(n/d)$ vertex-pair queries is $o(1)$. But if the algorithm always gets negative answers for its vertex-pair queries, it clearly cannot close a triangle with such a query.

Neighbor queries. Given a neighbor query (u, b', i) , where $u \in V_b$, the process gives an answer (v, j) , where $v \in V_{b'}$ according to the conditional probability that the entries $T_{b,b'}[u, i]$ and $T_{b',b}[v, j]$ are matched (where again, the conditioning is on the current knowledge graph G^{t-1}). Here we would like to upper bound the probability that v already belongs to the knowledge graph. Observe that the knowledge graph G^{t-1} contains at most $2t$ vertices and at most t edges and nonedges. Hence the conditional probability that $T_{b,b'}[u, i]$ is matched to $T_{b',b}[v, j]$ for some $v \in V_{b'}^{t-1} = V^{t-1} \cap V_{b'}$ and $j \in \{1, \dots, d/2\}$ is upper bounded by the following expression: $(d/2) \times (2t)$ (the maximum number of unmatched entries $T_{b',b}[v, j]$ for $v \in V_{b'}^{t-1}$) divided by the total number of unmatched entries in $T_{b',b}$ that do not belong to rows of vertices $u \in V_{b'}$ such that $(u, v) \in \bar{E}^{t-1}$, minus $d/2$ (where the argument is very similar to the one given for vertex-pair queries). By our assumption on Q (where $t-1 < Q$), the numerator in the above expression is $o(\min\{n, d^2\})$, and the denominator is at least $(n/3) \cdot (d/2) - (t-1) \cdot (d/2) - (d/2) = \Omega(n \cdot d)$. Hence, the probability that such an event occurs at a given particular time t is $o(\min\{(1/d), (d/n)\})$, and the probability that it occurs at any $t \leq Q$ is $o(1)$. But if the algorithm never gets a vertex in the knowledge graph as an answer to a neighbor query, it clearly cannot close a triangle, or any cycle, with such a query.

Finally, we address the case that $d < \frac{2}{3\sqrt{3}}\sqrt{n}$. Recall that in this case we first partition the vertices into k parts $V^1, \dots, V^k$, where $|V^i| = n' = \frac{27}{4}d^2$. We then apply the construction to each V^i . By the argument described above, here we can get a lower bound of the form $\Omega(\min\{d, n'/d\}) = \Omega(d)$. But for the current setting of d we also have that $\min\{d, n/d\} = d$, and so the lower bound holds in this case as well. $\square$ (Lemmas 5 and 2)

A remark about multiple edges. The lower bound proof stated above is valid for graphs that may contain multiple edges. In the distribution D_Δ the probability of a multiple edge between a pair of vertices is $O(\frac{d^2}{n^2})$. Thus, graphs created according to this distribution contain multiple edges with probability close to 1. However, with probability $1 - o(1)$, there are $O(d^2)$ multiple edges in the graphs created according to this distribution. We have shown in Lemma 3 that graphs created according to the distribution D_Δ are $\Omega(1)$ -far from being triangle-free with probability $1 - o(1)$. Hence we can deduce that by removing multiple edges from a graph G constructed according to the distribution D_Δ , the resulting graph is $\Omega(1)$ -far from being triangle-free with probability $1 - o(1)$. In addition, an algorithm ALG that interacts with the process detects a multiple edge with probability $o(1)$ due to the following reason: ALG does not detect any edges by vertex-pair queries with probability $1 - o(1)$. The probability of detecting a multiple edge in a neighbor query is at most the probability that such a query is answered by a vertex in the knowledge graph. This probability was shown in Lemma 3 to be $o(1)$. Thus, at the end of the interaction with ALG, the process can delete all the multiple edges from the resulting graph so that the resulting graph contains no multiple edges. The graph remains $\Omega(1)$ -far from being triangle-free after the deletion, and its average degree is $d - o(1)$. As a conclusion we get that our lower bound is valid also for graphs with no multiple edges.

5. An improved lower bound for high degrees. In this section we establish the following lemma, which improves on our previous lower bound of $\min\{d, n/d\}$ when the degree of the graph is at least $n^{2/3+o(1)}$.

LEMMA 6. *Every one-sided error testing algorithm for triangle-freeness must perform $\Omega(\min\{\sqrt{d}, \frac{n^{2/3}}{d^{1/3}}\} \cdot n^{-\nu(n)})$ queries, where $\nu(n) = \frac{\log \log \log n + 4}{\log \log n}$. This lower*

bound holds even for d -regular graphs.

In order to prove the lemma, here too we define a distribution over graphs that are far from being triangle-free. We then prove a lower bound on the number of queries that are required in order to detect a triangle with probability bounded away from zero in a graph that is generated according to the distribution. As we shall see, it will actually be convenient to consider graphs over $3n$ vertices and degree $2d$.

5.1. A variant of Behrend graphs. Our lower-bound distribution builds on graphs that are variants of what are known as *Behrend graphs* [7, 19, 20]. These graphs are defined by sets of integers that include no three-term arithmetic progression. Namely, these are sets $X \subset \{1, \dots, n\}$ such that for every three elements $x_1, x_2, x_3 \in X$, if $x_2 - x_1 = x_3 - x_2$ (i.e., $x_1 + x_3 = 2x_2$), then necessarily $x_1 = x_2 = x_3$. Below we describe a construction of such sets that are large (relative to n) and later explain how such sets determine Behrend graphs. Our construction of X uses ideas similar to those used in known constructions [7, 20] and gives a slightly weaker result. However, our alternative construction is somewhat simpler, and the size of the resulting set suffices for our purposes.

LEMMA 7. *For every sufficiently large n there exists a set $X \subset \{1, \dots, n\}$, $|X| \geq n^{1 - \frac{\log \log \log n + 4}{\log \log n}}$, such that X contains no three-term arithmetic progression.*

Proof. For simplicity we do not explicitly write floors (or ceilings). Let $b = \log n$ and $k = \log n / \log b - 2$. Since $\log n / \log b = \log n / \log \log n$, we have that $k < b/2$ for every $n \geq 8$. We arbitrarily select a subset of k different numbers $\{x_1, \dots, x_k\} \subset \{0, \dots, b/2 - 1\}$ and define $X = \{\sum_{i=1}^k x_{\pi(i)} b^i : \pi \text{ is a permutation of } \{1, \dots, k\}\}$. By the definition of X we have that $|X| = k!$. By using $z! > (z/e)^z$, we get that

$$(28) \quad |X| = k! \geq \left(\frac{\log \log n}{\log n} \right)^2 \cdot \left(\frac{\log n}{\log \log n} \right)! > n^{1 - \frac{\log \log \log n + 4}{\log \log n}}.$$

Consider any three elements $u, v, w \in X$ such that $u + v = 2w$. By definition of X , these elements are of the form $u = \sum_{i=1}^k x_{\pi_u(i)} b^i$, $v = \sum_{i=1}^k x_{\pi_v(i)} b^i$, and $w = \sum_{i=1}^k x_{\pi_w(i)} b^i \in X$, where π_u, π_v, π_w are permutations over $\{1, \dots, k\}$. Since $x_i < b/2$ for every $1 \leq i \leq k$, it must be the case that, for each i ,

$$(29) \quad x_{\pi_u(i)} + x_{\pi_v(i)} = 2x_{\pi_w(i)}.$$

This implies that, for every i ,

$$(30) \quad x_{\pi_u(i)}^2 + x_{\pi_v(i)}^2 \geq 2x_{\pi_w(i)}^2,$$

where the inequality in (30) is strict unless $x_{\pi_u(i)} = x_{\pi_v(i)} = x_{\pi_w(i)}$. If we sum over all i 's and there is at least one index i for which the inequality in (30) is strict, we get that

$$(31) \quad \sum_{i=1}^k x_{\pi_u(i)}^2 + \sum_{i=1}^k x_{\pi_v(i)}^2 > \sum_{i=1}^k 2x_{\pi_w(i)}^2,$$

which is a contradiction since we took permutations of the same numbers. Thus, we get that $u = v = w$. $\square$

Remark. In fact, the set constructed above is also three-term arithmetic progression free when all calculations are performed modulo n . We will use this observation below.

Behrend graphs. Given a set $X \subset \{1, \dots, n\}$ with no three-term arithmetic progression, we define the Behrend graph BG_X as follows. It has $3n$ vertices that are partitioned into three equal parts: V_1 , V_2 , and V_3 . For each $i \in \{1, 2, 3\}$ we associate with each vertex in V_i a different integer in $\{0, \dots, n-1\}$. The edges of the graph are defined as follows:

- The edges between V_1 and V_2 : For every $x \in X$ and $j \in \{0, \dots, n-1\}$ there is an edge between $j \in V_1$ and $(j+x) \bmod n \in V_2$.
- The edges between V_2 and V_3 : For every $x \in X$ and $j \in \{0, \dots, n-1\}$ there is an edge between $(j+x) \bmod n \in V_2$ and $(j+2x) \bmod n \in V_3$.
- The edges between V_1 and V_3 : For every $x \in X$ and $j \in \{0, \dots, n-1\}$ there is an edge between $j \in V_1$ and $(j+2x) \bmod n \in V_3$.

We shall say that an edge between $j \in V_1$ and $j' \in V_2$ or between $j \in V_2$ and $j' \in V_3$ is *labeled* by x if $j' = (j+x) \bmod n$, and we shall say that an edge between $j \in V_1$ and $j' \in V_3$ is *labeled* by x if $j' = (j+2x) \bmod n$.

The graph BG_X is $2|X|$ -regular and it contains $3|X|n$ edges. For every $j \in \{0, \dots, n-1\}$ and $x \in X$, the graph contains a triangle $(j, (j+x) \bmod n, (j+2x) \bmod n)$, where $j \in V_1$, $(j+x) \bmod n \in V_2$, and $(j+2x) \bmod n \in V_3$. There are $n \cdot |X|$ such edge-disjoint triangles, and every edge is part of one such triangle. Moreover, there are no other triangles in the graph. To verify this consider any three vertices j_1, j_2, j_3 , where $j_i \in V_i$ and such that there is a triangle between the three vertices. By definition of the graph, $j_2 = (j_1 + x_1) \bmod n$ for some $x_1 \in X$. $j_3 = (j_2 + x_2) \bmod n$ for some $x_2 \in X$. $j_3 = (j_1 + 2x_3) \bmod n$ for some $x_3 \in X$. Therefore, $(j_1 + x_1 + x_2) \bmod n = (j_1 + 2x_3) \bmod n$. That is, we get that $(x_1 + x_2) \bmod n = (2x_3) \bmod n$. Note that by the definition of X , for every $x \in X$, $x < n/2$, and so $x_1 + x_2 = 2x_3$. Since X contains no three-term arithmetic progression, the last implies that $x_1 = x_2 = x_3$, meaning that the triangle (j_1, j_2, j_3) is of the form $(j, (j+x) \bmod n, (j+2x) \bmod n)$.

5.2. The edge density of large sets in random Behrend graphs. In this subsection we prove the following lemma, which is central to the proof of Lemma 6. We shall use the following notation: For a subset $Y \subseteq X$ and a subset of vertices C in BG_Y , we let $e_Y(C)$ denote the number of edges spanned by C in BG_Y .

LEMMA 8. Let $0 < \beta < \frac{1}{2}$ and $0 < \alpha \leq 1$ be such that $\alpha - 2\beta > \frac{1}{\log \log n}$, and let $X \subset \{1, \dots, n\}$, $|X| \geq n^\beta$. Consider the random Behrend graph BG_Y obtained by choosing a random subset $Y \subseteq X$, $|Y| = d = \frac{|X|}{n^\beta}$. With high probability over the choice of Y , for every subset C of vertices in BG_Y , where $|C| = n^\alpha$, we have $e_Y(C) \leq \frac{90}{\alpha - 2\beta} \frac{n^{2\alpha}}{n^\beta}$ edges.

The lemma states that for sufficiently large subsets C (i.e., for $|C| = n^\alpha$, where $\alpha - 2\beta$ is a constant), the number of edges $e_Y(C)$ is not much larger than its expected value. Note that the smaller we choose β (i.e., the larger we choose Y), the smaller α can be. Thus, the lemma can be applied to sets of relatively small size.

Before proving the lemma we introduce some notation and prove two claims. For a subset $W \subseteq V_1 \cup V_2$, $|W| = s$, let $W_1 = W \cap V_1$, $W_2 = W \cap V_2$, and consider the subgraph of BG_X induced by W . Let

$$(32) \quad \begin{aligned} \Delta(W) = \{ & (j_2 - j_1) \bmod n : j_1 \in W_1, j_2 \in W_2, \\ & \text{and } (j_2 - j_1) \bmod n \in X \} \end{aligned}$$

denote the set of *differences* in W . That is, it is the set of labels of the edges between W_1 and W_2 in BG_X . Obviously, $|\Delta(W)| \leq |W|^2 = s^2$. For every difference $x \in \Delta(W)$, we define the *multiplicity* of x in W as the number of edges in BG_X between vertices in W_1 and vertices in W_2 that are labeled by x .

Let $k = \frac{5}{\alpha-2\beta}$. For β and α that satisfy the condition of the lemma ($\alpha - 2\beta > \frac{1}{\log \log n}$) we have that $k \leq 5 \log \log n$. We shall say that W is *good* if no difference in $\Delta(W)$ has multiplicity higher than k in W .

CLAIM 9. *With high probability over the choice of $Y \subseteq X$, for every good W such that $|W| = s \geq n^\beta \log n$, we have that $e_Y(W) \leq \frac{2ks^2}{n^\beta}$.*

Proof. Consider a fixed choice of a good W such that $|W| = s = n^\beta \log n$. By definition, $|\Delta(W)| \leq |W|^2 = s^2$. Since W is good, for every $x \in \Delta(W)$, if $x \in Y$, then the number of edges labeled by x in the subgraph of BG_X induced by W is at most k .

Since $Y \subseteq X$, $|Y| = d = \frac{|X|}{n^\beta}$ is a random subset selected uniformly from a set of size $|X|$, the expected size of $\Delta(W) \cap Y$ is $|\Delta(W)| \cdot n^{-\beta} \leq s^2 \cdot n^{-\beta}$. Using known bounds of the tail of the hypergeometric distribution (see, e.g., [13, p. 29]), the probability that $|\Delta(W) \cap Y| > 2 \cdot s^2 \cdot n^{-\beta}$ is upper bounded by $\exp(-\frac{cs^2}{n^{-\beta}})$ for some constant c . The claim follows by taking a union bound over all choices of W of size s . $\square$

We also need the following claim.

CLAIM 10. *Let C be a subset of $V_1 \cup V_2$ such that $|C| = n^\alpha$. Suppose we uniformly and independently select $W \subset C$, $|W| = n^\beta \log n$. Then the probability that W is not good is at most $\frac{1}{n^\beta}$.*

Proof. Note that by the definition of Behrend graphs, the edges between vertices in C that are labeled by a specific difference form a matching. When we choose a random subset $W \subset C$, the probability that there exists a single difference of C (i.e., an element of $\Delta(C)$) that has multiplicity at least $k+1$ in W is bounded by

$$(33) \quad |C|^2 |C|^{k+1} \binom{|C| - (2k+2)}{|W| - (2k+2)} \left(\frac{|C|}{|W|} \right)^{-1}.$$

To verify this expression, note first that there are at most $|C|^2$ possibilities from which to choose a difference from $\Delta(C)$. Since the edges of a specific difference form a matching of size at most $|C|$ over the edges of C , there are at most $|C|^{k+1}$ possibilities from which to choose the $k+1$ edges of this difference. The $k+1$ edges of the difference determine $2k+2$ of the vertices of W . Thus there are $\binom{|C| - (2k+2)}{|W| - (2k+2)}$ possibilities from which to choose the remaining vertices of W . Now,

$$(34) \quad \frac{|C|^2 |C|^{k+1} \binom{|C| - (2k+2)}{|W| - (2k+2)}}{\binom{|C|}{|W|}} \leq |C|^2 \cdot |C|^{k+1} \cdot \left(\frac{|W|}{|C|} \right)^{2k+2} \\ \leq n^{2\alpha} \left(\frac{\log^2 n}{n^{\alpha-2\beta}} \right)^{\frac{5}{\alpha-2\beta}} \leq \frac{1}{n^2}.$$

The last expression is upper bounded by $\frac{1}{n^\beta}$, as required. $\square$

Proof of Lemma 8. Consider a set C of vertices of BG_Y such that $|C| = n^\alpha$. Let $C_i = C \cap V_i$ for $1 \leq i \leq 3$. We will show that almost surely the number of edges between C_1 and C_2 is at most $\frac{30}{\alpha-2\beta} \frac{n^{2\alpha}}{n^\beta}$. The argument for the number of edges between C_2 and C_3 and between C_1 and C_3 is analogous, and hence the lemma

follows. We shall prove the claim for every C_1, C_2 such that $|C_1 \cup C_2| = n^\alpha$. Clearly this implies that it holds for every C_1, C_2 such that $|C_1 \cup C_2| \leq n^\alpha$.

By Claim 9, with high probability over the choice of Y , the following holds. For every good W , $W \subset C$, such that $|W| = s \geq n^\beta \log n$, the number of edges spanned by W in BG_Y is at most $2ks^2n^{-\beta}$. Assume from now on that the selected Y has this property. We shall use Claim 10 to derive an upper bound on the number of edges in BG_Y that are spanned by the vertices of C .

By our assumption on Y , if W is good and $|W| = s \geq n^\beta \log n$, then $e_Y(W) \leq 2ks^2n^{-\beta}$. Clearly, $e_Y(W) \leq s^2$. If we uniformly at random select $W \subset C$ such that $|W| = s$, then

$$\text{Exp}[e_Y(W)] \geq \frac{1}{2} \cdot e_Y(C) \cdot \frac{s^2}{n^{2\alpha}}.$$

We stress that the expectation is taken only over the choice of W and not over the choice of Y . Now,

$$\begin{aligned} & \text{Exp}[e_Y(W)] \\ &= \text{Exp}[e_Y(W) \mid W \text{ is good}] \cdot \Pr[W \text{ is good}] \\ & \quad + \text{Exp}[e_Y(W) \mid W \text{ is not good}] \cdot \Pr[W \text{ is not good}] \\ &\leq 2ks^2 \cdot n^{-\beta} + s^2 \cdot n^{-\beta} \\ (35) \quad &= (2k+1)s^2 \cdot n^{-\beta}. \end{aligned}$$

It follows that

$$(36) \quad e_Y(C) \leq (2k+1) \cdot 2|C|^2 \cdot n^{-\beta} \leq 5k|C|^2 \cdot n^{-\beta}.$$

Since $k = \frac{5}{\alpha-2\beta}$, the lemma follows. $\square$

As a corollary of Lemma 8 we get the following.

COROLLARY 11. *Let $0 < \beta < \frac{1}{2}$ and $X \subset \{1, \dots, n\}$, where $|X| \geq n^{1-\nu(n)}$ for $\nu(n) = \frac{\log \log \log n + 4}{\log \log n}$. Consider the random Behrend graph BG_Y obtained by choosing a random subset $Y \subseteq X$, $|Y| = d = \frac{|X|}{n^\beta}$. With high probability over the choice of Y , for every subset C of vertices in BG_Y such that $|C| \leq \min\{\sqrt{d}, \frac{n^{2/3}}{d^{1/3}}\} \cdot n^{-\nu(n)}$, the following bound applies: $|C| \cdot e_Y(C) \leq n^{1-\nu(n)}$.*

Proof. Note that $d = n^{1-\nu(n)-\beta}$, and so we need to bound the number of edges in sets C such that

$$(37) \quad |C| \leq \min \left\{ n^{\frac{1-\nu(n)-\beta}{2}}, n^{\frac{1+\nu(n)+\beta}{3}} \right\} \cdot n^{-\nu(n)} \leq n^{2/5-\nu(n)},$$

where $n^{2/5-\nu(n)}$ is obtained for $\beta = 1/5 - \nu(n)$.

Consider first the case that $\beta \leq \frac{1}{5} - \nu(n)$. In this case we need to bound the number of edges in sets C such that $|C| \leq \frac{n^{2/3-\nu(n)}}{d^{1/3}} = n^{\frac{1-2\nu(n)+\beta}{3}}$.

Let $\alpha = \frac{1-2\nu(n)+\beta}{3}$ and observe that $\alpha - 2\beta \geq \nu(n) > \frac{1}{\log \log n}$. Hence we can apply Lemma 8 and get that with high probability over the choice of Y , for every subset C such that $|C| = n^\alpha$, we have that $e_Y(C) \leq \frac{90}{\alpha-2\beta} \cdot \frac{n^{2\alpha}}{n^\beta}$. Clearly this upper bound also holds for every subset C such that $|C| \leq n^\alpha$. Hence,

$$(38) \quad |C| \cdot e_Y(C) \leq \frac{90}{\alpha-2\beta} \cdot \frac{n^{3\alpha}}{n^\beta} \leq \frac{90}{\nu(n)} \cdot n^{1-2\nu(n)} \leq n^{1-\nu(n)}.$$

Consider now the case that $\beta > \frac{1}{5} - \nu(n)$. Note that we need to bound the number of edges in sets C such that $|C| \leq n^{2/5 - \nu(n)}$. We have shown that the bound applies for the case that $\beta = \frac{1}{5} - \nu(n)$, and $|C| = n^{2/5 - \nu(n)}$. Hence, for $\beta > \frac{1}{5} - \nu(n)$, the sets C for which $|C| \leq n^{2/5 - \nu(n)}$ contain fewer edges, and therefore the previous bound applies. $\square$

5.3. The lower-bound distribution $BG(n, d)$. Let $X \subset [n]$ be a set with no three-term arithmetic progression, as constructed in subsection 5.1, such that $|X| = n^{1 - \nu(n)}$ (where $\nu(n) = \frac{\log \log \log n + 4}{\log \log n}$). Consider the Behrend graph, denoted BG_X , whose set of generators is X . Recall that BG_X , which is a graph over $3n$ vertices, contains $|X| \cdot n$ edge-disjoint triangles: every edge belongs to exactly one triangle, and every triangle corresponds to some $x \in X$.

For each subset $Y \subset X$ such that $|Y| = d$, we consider the subgraph of BG_X that contains all its vertices but only the edges labeled by differences $y \in Y$. This (sub)graph contains $n \cdot |Y| = nd$ edge-disjoint triangles and is hence ϵ -far from being triangle-free for any $0 < \epsilon < 1/3$. Next we apply a permutation π on the names of the vertices. More precisely, π consists of 3 permutations, $\pi^{(b)}$, $b \in \{1, 2, 3\}$, each over $\{0, \dots, n-1\}$. If we denote each vertex v in BG_X by a pair (b, i) , where $b \in \{1, 2, 3\}$ is the index of the subset that the vertex belongs to and $i \in \{0, \dots, n-1\}$, then $\pi(v) = \pi(b, i) = (b, \pi^{(b)}(i))$. We denote the resulting graph by $BG_{Y, \pi}$.

A graph is generated according to the distribution $BG(n, d)$ by uniformly selecting Y and π and outputting the resulting graph $BG_{Y, \pi}$. We also assume that the edges incident to a vertex v are ordered randomly in the incidence list of v . For the sake of simplicity, we do not include these random labelings in the notation.

5.4. Online generation of graphs according to $BG(n, d)$. In order to prove Lemma 6 we would like to show that any algorithm that is given query access to a graph generated according to $BG(n, d)$ must perform $\Omega(\min\{\sqrt{d}, \frac{n^{2/3}}{d^{1/3}}\} \cdot n^{-\nu(n)})$ queries in order to detect a triangle with sufficiently high constant probability. We wish to stress that the algorithm can be much more powerful/sophisticated potentially than just an algorithm that samples a random set of the input and looking for a triangle inside.

For the sake of our analysis, it will be convenient to define an online process, denoted P , that answers the algorithm's queries while generating a graph according to $BG(n, d)$. The process P will actually provide the algorithm with more information than required by neighbor and vertex-pair queries. Namely, whenever the algorithm asks a query involving a vertex $v \in \{(1, 0), \dots, (3, n-1)\}$ (in either type of query), the process will provide it with $\pi^{-1}(v)$. This will also be the case when the process answers a neighbor query (u, i) with a vertex v . Thus the algorithm is provided with the "identity" in BG_X of the vertices it has observed and can also derive the labels $y \in Y$ of the edges it has observed.

Clearly, a lower bound on the number of queries of an algorithm that is provided with the additional information described above constitutes a lower bound on an algorithm that is not provided with this additional information. It will actually be convenient to consider the following three types of queries:

1. Vertex queries: For any choice of a vertex $v \in \{(1, 0), \dots, (3, n-1)\}$ the algorithm is provided with $\pi^{-1}(v)$.
2. Random neighbor queries: For any vertex v the algorithm has already observed, it may ask for a new random neighbor u of v (together with $\pi^{-1}(u)$). The algorithm can indicate the subset to which the neighbor belongs.

3. Difference queries: For any $x \in X$ the algorithm can ask whether $x \in Y$.

Vertex queries are denoted $(VQ, (b, i))$, neighbor queries are denoted $(NQ, (b, i), b')$, and difference queries are denoted (DQ, x) , where $b \neq b' \in \{1, 2, 3\}$, $i \in \{0, \dots, n-1\}$, and $x \in X$.

Note that a vertex-pair query can be performed by asking at most two vertex queries and one difference query, and a neighbor query can be performed by asking a vertex query and a random neighbor query (recall that the edges adjacent to a vertex are labeled randomly in $BG_{Y,\pi}$). It follows that any lower bound on algorithms that perform these types of queries implies a lower bound that is at most a factor of 3 smaller on algorithms that perform vertex-pair and neighbor queries.

Let ALG be an algorithm that performs $Q = Q(n, d)$ queries of the above three types and whose goal is to detect a triangle with probability at least $9/10$. The probability is taken over the choice of the graph $BG_{Y,\pi}$ and the coin flips of the algorithm. Namely, ALG is a (possibly probabilistic) mapping from *query-answer histories* $\langle (q_1, t_1), \dots, (q_t, a_t) \rangle$ to q_{t+1} for every $t < Q$. The mapping needs to be defined only on histories that are consistent with some graph $BG_{Y,\pi}$.

As described above, a vertex query $q_t = (VQ, v_t)$ for $v_t \in \{(1, 0), \dots, (3, n-1)\}$ that has not yet been observed is answered by $\pi^{-1}(v_t)$. A random neighbor query $q_t = (NQ, v_t, b')$ for $v_t = (b, i)$ that has been observed is answered by a new random neighbor $U_t = (b', j)$ of v_t together with $\pi^{-1}(U_t)$. A difference query $q_t = (DQ, x)$ for $x \in X$ is answered by “1” (yes) or “0” (no), indicating whether $x \in Y$ or not.

Any query-answer history of length t can be used to define the *knowledge base* $K_t = (V_t, Y_t, \bar{Y}_t, \pi_t)$, where $V_t \subset \{(1, 0), \dots, (3, n-1)\}$, $Y_t, \bar{Y}_t \subset X$, and $\pi_t : V_t \rightarrow \{(1, 0), \dots, (3, n-1)\}$ (where π_t is one-to-one). Specifically, V_t consists of all vertices (b, j) such that $(b, j) = \pi^{-1}(v)$ for some v that appeared either in one of the first t queries of ALG or in one of the first t answers. The set Y_t consists of all $x \in X$ such that for some $t' \leq t$ there was either a query $q_{t'} = (DQ, x)$ that was answered by “1” or a query $q_{t'} = (NQ, v_{t'}, b')$ that was answered with $U_{t'}$, where the edge between $\pi^{-1}(v_{t'})$ and $\pi^{-1}(U_{t'})$ is labeled by x . The set $\bar{Y}_t$ consists of all $x \in X$ such that for some $t' \leq t$ there was a query $q_{t'} = (DQ, x)$ that was answered by “0.” Finally, for every $(b, j) \in V_t$, $\pi_t(b, j) = v$, where v is such that $\pi^{-1}(v) = (b, j)$.

Observe that V_t together with Y_t determine a subgraph of BG_X : the vertices of the subgraph are the vertices of V_t , and the edges are all pairs (u, v) , $u, v \in V_t$, such that there is an edge between u and v in BG_X and this edge is labeled by some $x \in Y_t$. For each $b \in \{1, 2, 3\}$ we let $V_{t,b} = \{(b, j) : (b, j) \in V_t\}$.

5.4.1. Definition of the process P . Let $R = R(n, d)$ denote the set of all graphs $BG_{Y,\pi}$ in the support of $BG(n, d)$. For $b \in \{1, 2, 3\}$ and $i, j \in \{0, \dots, n-1\}$ let $R_{b,i,j} \subset R$ denote the subset of graphs $BG_{Y,\pi} \in R$ such that $\pi(b, j) = (b, i)$. For $x \in X$ let $R_x \subset R$ denote the subset of graphs $BG_{Y,\pi} \in R$ such that $x \in Y$, and let $R_{\neg x} \subset R$ denote the subset of graphs $BG_{Y,\pi} \in R$ such that $x \notin Y$.

The process P answers ALG’s queries as follows, where we assume without loss of generality that ALG does not ask any query q_t whose answer can be deduced from the knowledge base K_{t-1} . In particular, for every vertex query $q_t = (VQ, v)$ we have that $v \notin V_{t-1}$, and for every difference query $q_t = (DQ, x)$ we have that $x \notin Y_{t-1} \cup \bar{Y}_{t-1}$. The process P initializes $R_0 = R$, and in general, for any $t \geq 0$, we have that R_t consists of all graphs in R that are consistent with the first t queries.

- To answer a vertex query $q_t = (VQ, v)$, where $v = (b, i)$, the process uniformly selects $(b, j) \in \{(b, 0), \dots, (b, n-1)\} \setminus V_{t-1,b}$ and sets $R_t = R_{b,i,j} \cap R_{t-1}$. Note that (b, j) is selected with probability $\frac{|R_{b,i,j} \cap R_{t-1}|}{|R_{t-1}|} = \frac{|R_t|}{|R_{t-1}|}$.

- Given a difference query (DQ, x) , with probability $\frac{d-|Y_{t-1}|}{|X|-|Y_{t-1}|-|\bar{Y}_{t-1}|}$ the process answers “1” and with probability $1 - \frac{d-|Y_{t-1}|}{|X|-|Y_{t-1}|-|\bar{Y}_{t-1}|} = \frac{|X|-d-|\bar{Y}_{t-1}|}{|X|-|Y_{t-1}|-|\bar{Y}_{t-1}|}$ it answers “0.” In the former case it sets $R_t = R_x \cap R_{t-1}$ and in the latter case it sets $R_t = R_{\neg x} \cap R_{t-1}$. Note that here the answer is “1” with probability $\frac{|R_x \cap R_{t-1}|}{|R_{t-1}|} = \frac{|R_t|}{|R_{t-1}|}$ and is “0” with probability $1 - \frac{|R_x \cap R_{t-1}|}{|R_{t-1}|} = \frac{|R_{\neg x} \cap R_{t-1}|}{|R_{t-1}|} = \frac{|R_t|}{|R_{t-1}|}$.
- To answer a random neighbor query $q_t = (NQ, v, b')$, the process performs two steps. First it selects $x \in X$ in the following manner. With probability $\frac{|Y_{t-1}|}{d}$ it selects x uniformly from Y_{t-1} , and with probability $1 - \frac{|Y_{t-1}|}{d}$ it selects x uniformly from $X \setminus (Y_{t-1} \cup \bar{Y}_{t-1})$. In either case the choice of x (together with b') determines the value of $j \in \{0, \dots, n-1\}$ such that there is an edge labeled x in BG_X between $\pi^{-1}(v)$ and (b', j) . If $(b', j) \in V_{t-1}$, then $u = \pi_t(b', j) = \pi(b', j)$ is known and $R_t = R_x \cap R_{t-1}$. Otherwise, $u = (b', i)$ is selected uniformly in $\{(b', 0), \dots, (b', n-1)\} \setminus V_{t-1, b'}$ and $R_t = R_x \cap R_{b', i, j} \cap R_{t-1}$.

Once P completes answering the Q queries of ALG it uniformly selects a graph in R_Q . The next lemma is easily derived from the definition of the process.

LEMMA 12. *For every algorithm ALG , the process P , when interacting with ALG , uniformly generates graphs in $BG(n, d)$*

Proof. Consider a specific graph G in $R_0 = BG(n, d)$. The probability that G is generated by P is

$$\begin{aligned}
 & \Pr[G \in R_1] \cdot \Pr[G \in R_2 | G \in R_1] \cdots \Pr[G \in R_Q | G \in R_{Q-1}] \cdot \frac{1}{|R_Q|} \\
 (39) \quad &= \frac{|R_1|}{|R_0|} \cdot \frac{|R_2|}{|R_1|} \cdots \frac{|R_Q|}{|R_{Q-1}|} \cdot \frac{1}{|R_Q|} = \frac{1}{|R_0|},
 \end{aligned}$$

and the lemma follows. $\square$

5.5. Proof of Lemma 6. Consider any algorithm ALG that interacts with the process P . We shall show that if ALG asks $Q = o(\min\{\sqrt{d}, \frac{n^{2/3}}{d^{1/3}}\} \cdot n^{-\nu(n)})$ queries, then the probability that it reveals a triangle is $o(1)$. By Corollary 11, with high probability over the choice of Y , $|Y| = d$, for every subset U of vertices such that $|U| = o(\min\{\sqrt{d}, \frac{n^{2/3}}{d^{1/3}}\} \cdot n^{-\nu(n)})$, we have $|U| \cdot e(U) = o(n^{1-\nu(n)})$, where $e(U)$ denotes the number of edges induced by the vertices of U . In what follows we assume that the graph generated by P in fact has this property, where we take into account the probability of $o(1)$ that this is not the case. Let E_t be the set of edges between vertices in V_t that are known to the algorithm after the first t queries. That is, E_t consists of all edges in BG_X whose labels are in Y_t and are between vertices in V_t . Therefore, for every $t = o(\min\{\sqrt{d}, \frac{n^{2/3}}{d^{1/3}}\} \cdot n^{-\nu(n)})$ we have that $t|E_t| = o(n^{1-\nu(n)})$.

Recall that every edge $(i, j) \in E_t$ participates in exactly one triangle. There are two ways by which ALG can close a triangle in its t th query. If the query is either a vertex query or a random neighbor query, the algorithm must receive as an answer one of the $|E_{t-1}|$ vertices that close triangles with (the known) edges between vertices in V_{t-1} . If the query is a difference query (DQ, x) (where $x \notin Y_{t-1} \cup \bar{Y}_{t-1}$), then there must be three vertices $(1, i), (2, j), (3, k) \in V_{t-1}$ that form a triangle in BG_X whose edges are labeled by x and the answer to the query is “1” (i.e., $x \in Y$). For the sake of simplicity we assume that whenever the algorithm obtains a vertex that

closes a triangle in BG_X , then it is also told whether this triangle is in $BG_{Y,\pi}$ or not (i.e., it gets a difference query “for free”). We now turn to bounding the probability of each of the above events by which a triangle is closed.

We first observe that since $|V_t| \leq t = o(n)$, whenever ALG asks a vertex query (VQ, v) , where $v = (b, i)$, the answer, (b, j) , is uniformly distributed in a subset of size $\Theta(n)$. Since $|Y_t| + |\bar{Y}_t| \leq t$, whenever ALG asks a random neighbor query $q_t = (NQ, v, b')$, with probability at least $1 - \frac{t}{d}$ we have that $(b', j) = \pi^{-1}(u)$ is uniformly distributed in a subset of size $|X| - t$. Since $|X| = \Omega(n^{1-\nu(n)})$ and $t = o(\sqrt{d}) = o(\sqrt{|X|})$, with probability $1 - o(1)$, for every neighbor query performed, the answer to the neighbor query is uniformly distributed in a subset of size $\Omega(n^{1-\nu(n)})$. Let us assume from this point on that this is in fact the case (where we take into account the $o(1)$ probability that this is not the case).

Given the above, the probability that ALG closes a triangle in the t th query when one of the edges of the triangle is already in E_t is $\frac{|E_t|}{\Omega(n^{1-\nu(n)})}$. It remains to bound the probability that ALG closes a triangle by obtaining a vertex (b, j) that closes a triangle in the subgraph of BG_X that is induced by V_{t-1} , and that the difference $x \notin Y_{t-1} \cup \bar{Y}_{t-1}$ corresponding to the triangle is determined to be in Y . Since the number of edges in the subgraph of BG_X that is induced by V_{t-1} is upper bounded by t^2 , the probability of the above event is at most

$$(40) \quad \frac{t^2}{\Omega(n^{1-\nu(n)})} \cdot \frac{d-t}{\Omega(n^{1-\nu(n)})} = \frac{t^2(d-t)}{\Omega(n^{2(1-\nu(n))})}.$$

Since $t = o(\min\{\sqrt{d}, n^{2/3}/d^{1/3}\} \cdot n^{-\nu(n)})$, we know that $t|E_t| = o(n^{1-\nu(n)})$ and $t^3 \cdot (d-t) = o(n^{2(1-\nu(n))})$. Hence, the probability that one of the above events occurs for any $t \leq Q$ is $o(1)$, as required. $\square$

6. From 1-sided error to 2-sided error. In this section we establish that under certain conditions, a one-sided error lower bound for triangle-freeness can be transformed into a two-sided error lower bound. Since these conditions hold for our one-sided error lower bounds, we obtain two-sided error lower bounds.

THEOREM 3. *Let D_Δ be a distribution over graphs with n vertices and average degree d , and let $q(n, d)$ be a function of these parameters. Assume that the following hold:*

- *With probability $1 - o(1)$ a graph selected according to D_Δ is ϵ -far from being triangle-free for some constant ϵ .*
- *One of the following two conditions holds:*
 1. *In all graphs in the support of D_Δ , the triangles are edge-disjoint, and for any algorithm A , the probability that A reveals a triangle in a graph selected according to D_Δ using $o(q(n, d))$ queries is less than $2/3$.*
 2. *For any algorithm A , the probability that A reveals a cycle (of any length) in a graph selected according to D_Δ using $o(q(n, d))$ queries is less than $2/3$.*

Then any two-sided error algorithm for testing triangle-freeness that has success probability at least $5/6$ must perform $\Omega(q(n/2, d))$ queries.

Since the distributions that are defined for our one-sided error lower bounds, which are stated in Lemmas 2 and 6, are as required by Theorem 3, we get the following corollary.

COROLLARY 13. *Any algorithm for testing triangle-freeness must perform*

$$\Omega\left(\max\left\{\min\{d, n/d\}, \min\left\{\sqrt{d}, n^{2/3}/d^{1/3}\right\} \cdot n^{-\nu(n)}\right\}\right)$$

queries. This lower bound holds even if the algorithm is allowed two-sided error and $d_{\max} = O(d)$.

6.1. Proof of Theorem 3: Condition 1. Given the distribution D_Δ we define two distributions over graphs that have $n' = 2n$ vertices and average degree d . One distribution, denoted D'_Δ , generates graphs that are ϵ -far from being triangle-free, and the other distribution, denoted $D_{\bar{\Delta}}$, generates graphs that are triangle-free. Assume, contrary to what is claimed in the theorem, that there exists a two-sided error algorithm A' for testing triangle freeness that performs $o(q(n'/2, d))$ queries and has success probability at least $5/6$. Then, in particular, using $o(q(n'/2, d)) = o(q(n/d))$ queries, A' should be able to distinguish with sufficiently high probability between graphs generated by D'_Δ and graphs generated by $D_{\bar{\Delta}}$. We shall show that we can then use A' to obtain an algorithm A that performs $o(q(n, d))$ queries and with probability at least $2/3$ reveals a triangle in a random graph generated according to D_Δ .

Defining the two distributions. In both distributions, a graph G' over $n' = 2n$ vertices is generated by first selecting a graph G from D_Δ . Every vertex v in G is replaced by two vertices, v_0 and v_1 . Every edge (u, v) in G is replaced by two edges—either the two edges (u_0, v_0) and (u_1, v_1) (so that they are “in parallel”) or the two edges (u_0, v_1) and (u_1, v_0) (so that they are “crossing”). If v is the j th neighbor of u and u is the ℓ th neighbor of v , then in both cases we maintain the ordering on neighbors. Namely, in the case of parallel edges we have that v_0 is the j th neighbor of u_0 and v_1 is the j th neighbor of u_1 , and u_0 is the ℓ th neighbor of v_0 and u_1 is the ℓ th neighbor of v_1 (an analogous correspondence holds for crossing edges). The difference between the distributions is in the choice (distribution on the choice) between the above two options.

Recall that the triangles in G are edge-disjoint. Hence, for each triangle in G , the edges between the corresponding vertices in G' can be determined independently from the edges that belong to other triangles. Consider a particular triangle (u, v, w) in G . There are $2^3 = 8$ ways to select the edges between the vertices $u_0, u_1, v_0, v_1, w_0, w_1$ (depending on whether we select parallel or crossing edges). In 4 of these ways we get 2 edge-disjoint triangles (e.g., (u_0, v_0, w_0) and (u_1, v_1, w_1)), and in 4 of these ways we get a single cycle of length 6 (e.g., $(u_0, v_0, w_0, u_1, v_1, w_1)$). The graph generated by D'_Δ simply selects one of the former 4 ways uniformly, and the graph generated by $D_{\bar{\Delta}}$ selects one of the latter 4 ways uniformly. For an illustration see Figure 2.

The basic but important observation is that for both distributions the following holds: If we consider any edge that belongs to a particular triangle in G , then the probability that the corresponding pair of edges in G' are parallel is equal to the probability that they are crossing. Moreover, this remains true if we condition it on any other (single) edge in the triangle being transformed to either parallel or crossing edges. Independence breaks down only when we consider all 3 edges in a triangle. We shall refer to this observation as the *independence observation*.

Using a two-sided error algorithm to find triangles. Let A' be a two-sided error algorithm for testing triangle freeness that performs $o(q(n'/2, d))$ queries when testing graphs over $n' = 2n$ vertices and has success probability at least $5/6$. We next show how to use it in order to detect triangles in a graph G over n vertices that is generated randomly according to D_Δ . The idea is that by performing queries to G and flipping some coins, we shall actually be emulating the execution of A' on graphs generated by either D'_Δ or $D_{\bar{\Delta}}$. Since A' is supposed to test graphs over $n' = 2n$ vertices, we denote the vertices in the queries it performs by $\{v_{1,0}, v_{1,1}, \dots, v_{n,0}, v_{n,1}\}$.

Let G be a graph generated according to D_Δ . Algorithm A (whose goal is to

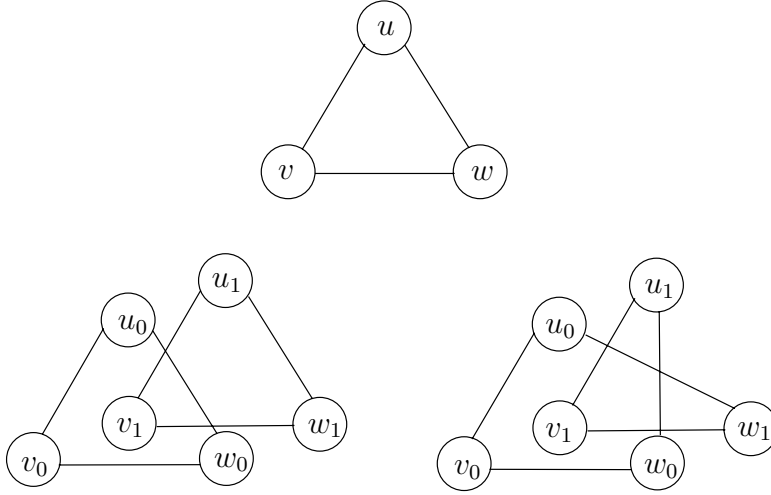


FIG. 2. On the top is a triangle (u, v, w) . On the bottom left is a transformation of this triangle into two disjoint triangles, and on the bottom right is a transformation into a cycle of length 6. In particular, on the left all pairs of edges are parallel, and on the right two are parallel and one is crossing.

detect a triangle in G) runs A' as a subroutine and answers its queries by performing queries to G and transforming the answers to the queries in an appropriate manner described below. In this process A maintains a *knowledge graph*, denoted $\hat{G}$, which contains all the edges it has observed in G as well as the “nonedges” (i.e., pairs (u, v) that do not have an edge between them). In addition, A records all the answers it has already given to A' .

Whenever A' performs a degree query for a vertex $v_{i,b}$ ($b \in \{0, 1\}$), algorithm A queries the degree of v_i and returns it as an answer. Whenever A' performs a vertex-pair query $(v_{i,b}, v_{j,b'})$ ($b, b' \in \{0, 1\}$), if (u, v) is an edge or a nonedge in the knowledge graph $\hat{G}$, then the answer to A' is determined. If this is not the case, then A performs the vertex-pair query (v_i, v_j) . If the answer is that there is no edge between the two vertices, then the answer given to A' is “no” as well. If the answer is that there is an edge, then there are two cases. If this edge closes a triangle with two other edges in $\hat{G}$, then A terminates successfully. Otherwise, with probability $1/2$ A answers that there is an edge between $v_{i,b}$ and $v_{j,b'}$ and with probability $1/2$ it answers that there is no such edge. In both cases the existence of the edge in G is recorded in the knowledge graph $\hat{G}$. In addition, in the former case A' is provided with the information concerning which neighbor is $v_{j,b'}$ of $v_{i,b}$.

Whenever A' performs a neighbor query $(v_{i,b}, \ell)$ (that does not correspond to an edge already in $\hat{G}$), algorithm A performs the neighbor query (v_i, ℓ) . Let the answer be (v_j, t) . Namely, there is an edge between v_i and v_j , where v_j is the ℓ th neighbor of v_i , and v_i is the t th neighbor of v_j . Here, too, if a triangle in G is detected, then A terminates successfully. Otherwise, it answers the query of A' in an analogous manner to the way a vertex-pair query is answered. If A' terminates before A has found a triangle, then A terminates unsuccessfully.

Completing the proof. Since A always terminates when or before it finds a triangle, by the independence observation, the distribution on the answers it gives to the queries of A' is exactly the same as the one we would get if the queries of A' were

answered by a graph that is selected either according to D'_Δ or according to $D_{\bar{\Delta}}$. We claim that this implies that the probability that A' terminates before A finds a triangle (thus causing A to terminate unsuccessfully) is less than $1/3$. Here the probability is taken over the choice of G , the coin flips of A , and the possible coin flips of A' .

Assume, contrary to the claim, that the probability that A' terminates before A finds a triangle is at least $1/3$. Consider the distribution over graphs that results from selecting with probability $1/2$ a graph G' according to D'_Δ , and with probability $1/2$ a graph G' according to $D_{\bar{\Delta}}$. By our counterassumption, the probability that A' terminates before it sees three edges of the form (v_{i,b_1}, v_{j,b_2}) , (v_{j,b_3}, v_{k,b_4}) , and (v_{k,b_5}, v_{i,b_6}) (where $b_1, \dots, b_6 \in \{0, 1\}$) is greater than $1/3$. In such a case, by the independence observation, the distribution on the answers to the queries of A' (and hence on its queries conditioned on these answers) is the same if the graph G' is selected according to D'_Δ or according to $D_{\bar{\Delta}}$. Therefore, the probability that A' terminates with an incorrect output is greater than $1/6 - o(1) > 1/9$, where the term $o(1)$ is due to the probability that G is not ϵ -far from triangle-free. But this contradicts our assumption on A' .

Since the number of queries performed by A before it terminates is upper bounded by the number of queries performed by A' , the theorem follows. $\square$

6.2. Proof of Theorem 3: Condition 2. Similarly to the proof of Theorem 3 under condition 1, given the distribution D_Δ we define two distributions over graphs that have $n' = 2n$ vertices and average degree d . One distribution, denoted D'_Δ , generates graphs that are $\Omega(1)$ -far from being triangle-free, and the other distribution, denoted $D_{\bar{\Delta}}$, generates graphs that are triangle-free. In fact, the graphs in the support of $D_{\bar{\Delta}}$ are all bipartite. We then show how it is possible to use an algorithm A' that can distinguish with high probability between graphs generated according to D'_Δ and graphs generated according to $D_{\bar{\Delta}}$ in order to obtain an algorithm A that with high probability reveals a cycle in a random graph generated according to D_Δ .

Defining the two distributions. In both distributions, a graph G' over $n' = 2n$ vertices is generated by first selecting a graph G from D_Δ . Here, too, each edge $(u, v) \in E(G)$ is replaced by two edges in G' : either the “parallel” edges (u_0, v_0) and (u_1, v_1) or the “crossing” edges (u_0, v_1) and (u_1, v_0) . The difference between the distributions is in the choice (distribution on the choice) between the above two options.

In the distribution D'_Δ , the decision whether an edge is transformed into parallel edges or into crossing edges is done independently, and with equal probability. Namely, for each function $\sigma : E(G) \rightarrow \{p, c\}$, there is a graph G_σ in the support of D'_Δ . For every edge $(u, v) \in E(G)$, if $\sigma(u, v) = p$, then in G_σ we have the parallel edges (u_0, v_0) and (u_1, v_1) , and if $\sigma(u, v) = c$, then in G_σ we have the crossing edges (u_0, v_1) and (u_1, v_0) . The function σ is selected uniformly at random.

On the other hand, each graph in the support of $D_{\bar{\Delta}}$ is determined by a function $\pi : V(G) \rightarrow \{\ell, r\}$. Here ℓ stands for “left” and r for “right” so that π defines a two-way partition of the vertices: If $\pi(v) = \ell$, then v_0 is put on the “left” side, and if $\pi(v) = r$, then v_0 is put on the “right” side. In either case, v_1 is put on the opposite side. The decision between parallel and crossing edges is done so that the resulting graph is bipartite. Specifically, for every edge $(u, v) \in E(G)$, if $\pi(u) = \pi(v)$, then in G_π we have the crossing edges (u_0, v_1) and (u_1, v_0) , and if $\pi(u) \neq \pi(v)$, then we have the parallel edges (u_0, v_0) and (u_1, v_1) . The function π is selected uniformly at random. We note that different choices of π may result in the same graph, but this is inconsequential to our argument.

Properties of the distributions. By definition of $D_{\bar{\Delta}}$, every graph in the support of $D_{\bar{\Delta}}$ is bipartite and hence is necessarily triangle-free. On the other hand, we claim that if the graph G (that is selected according to D_{Δ}) is ϵ -far from being triangle-free, then with probability $1 - \exp(-\Omega(\epsilon dn))$ over the choice of σ , the graph G_{σ} (in the support of D'_{Δ}) is at least $\epsilon/4$ -far from being triangle-free. To verify this claim observe that since G is ϵ -far from being triangle-free, it contains at least ϵdn edge-disjoint triangles. For each such triangle, the probability that it is transformed into two triangles in G_{σ} is $1/2$. By a multiplicative Chernoff bound, the probability that less than $(\epsilon dn)/4$ of the disjoint triangles in G are transformed into a pair of disjoint triangles is $\exp(-\Omega(\epsilon dn))$. That is, with probability $1 - \exp(-\Omega(\epsilon dn))$ there are $(\epsilon dn)/2$ disjoint triangles in the resulting graph, and since the graph contains $2n$ vertices, it is $(\epsilon/4)$ -far from being triangle-free. Since the probability that G is ϵ -far from triangle free for some constant ϵ is $1 - o(1)$, we get that with probability $1 - o(1)$, a graph generated according to D'_{Δ} is at least $\Omega(1)$ -far from being triangle-free.

Completing the proof. Most parts of the argument are almost identical to those in the proof of Theorem 3 subject to Condition 1, where the only difference is that “detecting a triangle” is replaced by “detecting a cycle.” The only essential difference is that the independence observation concerning the distribution on parallel/crossing edges conditioned on a triangle not being detected needs to be modified. Specifically, consider any fixed choice of edges in G that do not contain a cycle. We shall show that under both distributions, all mappings of these edges to parallel/crossing pairs of edges in G' have equal probability. This implies that if we fix any choice of parallel/crossing pairs of edges in G' , where the corresponding edges in G do not contain a cycle, and we consider a new edge that does not close a cycle with these edges, then under both distributions it has equal conditional probability to be mapped to a parallel/crossing pair of edges in G' .

Let $F = \{(u_1, v_1), \dots, (u_t, v_t)\}$ be any subset of edges in G that do not contain a cycle. For a graph G' either in the support of $D_{\bar{\Delta}}$ or in the support of D'_{Δ} , let $\sigma_{G'}(u_i, v_i) = p$ if the edge (u_i, v_i) is transformed into parallel edges in G' , and let $\sigma_{G'}(u_i, v_i) = c$ if it is transformed into crossing edges. By definition of the distributions, if G' is in the support of D'_{Δ} , so that $G' = G_{\sigma}$ for $\sigma : E(G) \rightarrow \{p, c\}$, then $\sigma_{G'}(u_i, v_i) = \sigma$, and if G' is in the support of $D_{\bar{\Delta}}$, so that $G' = G_{\pi}$ for $\pi : V(G) \rightarrow \{\ell, r\}$, then $\sigma_{G'}(u_i, v_i) = c$ when $\pi(u_i) = \pi(v_i)$, and $\sigma_{G'}(u_i, v_i) = p$ otherwise.

Consider the distribution over $\sigma_{G'}(u_1, v_1), \dots, \sigma_{G'}(u_t, v_t)$ when G' is chosen according to each of the two distributions. Then we claim that the distribution is uniform over $\{p, c\}^t$. This is clearly the case when G' is selected according to D'_{Δ} (as G' is determined by selecting $\sigma : E(G) \rightarrow \{p, c\}$ uniformly at random). It remains to verify that this is also true when $G' = G_{\pi}$, and $\pi : V(G) \rightarrow \{\ell, r\}$ is selected uniformly at random. Consider any vector $h \in \{p, c\}^t$ (where $h_i = p$ means that (u_i, v_i) is transformed into a parallel edge, and $h_i = c$ means that it is transformed into a crossing edge). Since F does not contain a cycle, it consists of a forest of trees, $T_1, \dots, T_s$. The total number of vertices in these trees is $t + s$. For each tree T_j , if the value of $\pi(v)$ is determined for some arbitrary vertex v in the tree, then the setting of h enforces a unique value $\pi(u)$ for every $u \in T_j$. In other words, the number of settings of π that induce a particular h is 2^s . Since the total number of settings of π is 2^{s+t} we get

$$(41) \quad \forall h \in \{p, c\}^t, \quad \Pr_{\pi} [\sigma_{G_{\pi}}(u_1, v_1), \dots, \sigma_{G_{\pi}}(u_t, v_t) = h] = \frac{2^s}{2^{t+s}} = \frac{1}{2^t},$$

and so the (modified) independence observation is established. The proof is completed analogously to the way it was completed under Condition 1 (using the modified independence observation and replacing detection of triangles with detection of cycles). $\square$

7. Upper bounds.

7.1. An upper bound of $\tilde{O}(\sqrt{nd}/\epsilon^{3/2})$ for general graphs. In this subsection we prove the next lemma.

LEMMA 14. *It is possible to test triangle-freeness by performing $\tilde{O}(\sqrt{nd}/\epsilon^{3/2})$ queries. If $d_{\max} = O(d)$, then $O(d/\epsilon)$ queries suffice.*

Proof. Let G be a graph with average degree d over n vertices that is ϵ -far from being triangle-free. By definition, G must contain at least ϵdn edges that belong to triangles. If $d_{\max} = O(d)$, then by uniformly selecting $\Theta(1/\epsilon)$ vertices and for each uniformly selecting an incident edge, with high probability we obtain an edge that belongs to a triangle. Conditioned on this event, if we now perform all $O(d)$ neighbor queries to the endpoints of each selected edge, we reveal a triangle.

If the maximum degree of the graph differs significantly from its average degree, then the above argument cannot be applied: First, the suggested edge selection process might not select with sufficiently high probability an edge that belongs to a triangle because the algorithm uniformly selects vertices rather than edges. Second, even if we obtain such an edge, its endpoints might have a very high degree. To address these issues, we first introduce some notation.

We say that a vertex has *high degree* if its degree is more than $c\sqrt{nd}$ (where we set c momentarily). We shall say that an edge is *covered* by these high degree vertices if *both* its endpoints have high degree. By definition, the high-degree vertices can cover at most $((1/c)\sqrt{nd})^2 = (1/c^2)nd$ edges. Hence, among the edges that belong to triangles, there are at least $(\epsilon - (1/c^2))nd$ edges that have at least one endpoint with degree at most $c\sqrt{nd}$. If we set $c = \sqrt{2/\epsilon}$, then we have at least $(\epsilon/2)nd$ such edges.

In order to obtain one of these edges, we would like to sample edges uniformly in G . In fact, it suffices to sample edges “almost uniformly” as defined in [14]. In [14] an algorithm is described that uses $\tilde{O}(\sqrt{n/\delta})$ queries to a graph G and for which the following holds: For all but at most $\delta/4$ -fraction of the edges of G the probability that the edge is selected is at least $\frac{1}{32nd}$. We refer to this algorithm as “Edge-Select.” By definition of the algorithm, if we set $\delta = \epsilon$, we get that there are at least $(\epsilon/4)nd$ edges that can be returned by “Edge-Select” such that these edges belong to triangles and have at least one endpoint with degree at most $\sqrt{2/\epsilon}\sqrt{nd}$. It follows that at a cost of $\tilde{O}(\sqrt{n}/\epsilon^{3/2})$ queries we obtain such an edge with a high constant probability. Thus the algorithm for detecting a triangle runs “Edge-Select” $\Theta(1/\epsilon)$ times. For each selected edge, if it has one endpoint with degree less than $\sqrt{2/\epsilon} \cdot \sqrt{nd}$, then it asks all neighbor queries for that vertex, and for each of them it asks all pair queries with the other endpoint. (If both endpoints have high degree, then the algorithm does nothing.) $\square$

7.2. An improved upper bound for relatively dense general graphs.

When the graph is relatively dense we can get the following improvement.

LEMMA 15. *It is possible to test triangle-freeness by performing $O(\max\{\frac{n^{4/3}}{\epsilon^{2/3}d^{2/3}}, \frac{d_{\max}^2}{\epsilon^2 d^2}\})$ queries.*

As a corollary we get the following.

COROLLARY 16. *It is possible to test triangle-freeness of graphs with average degree $d = \Omega(\sqrt{n})$ by performing $O(\frac{n^{4/3}}{d^{2/3}\epsilon^2})$ queries.*

Proof of Lemma 15. Let G be a graph over n vertices with average degree d and maximum degree $d_{\max}$ that is ϵ -far from being triangle-free. We shall show that if we take a uniform sample of $\Theta(\max\{\frac{n^{2/3}}{\epsilon^{1/3}d^{1/3}}, \frac{d_{\max}}{\epsilon d}\})$ vertices of G and ask vertex-pair queries between all pairs in the sample, then a triangle is detected with probability at least $2/3$.

Since G is ϵ -far from being triangle-free, it must contain at least ϵdn triples of vertices that form a triangle. This lower bound on the number of triangles implies that the expected number of triangles in a set of s uniformly selected vertices is at least $s^3 \cdot \frac{\epsilon dn}{n^3}$. It follows that for $s \geq n^{2/3}/(\epsilon d)^{1/3}$ the expected number of triangles spanned by the sample is at least 1. This unfortunately does not imply in general that a uniform sample of $s = \Omega(n^{2/3}/(\epsilon d)^{1/3})$ vertices spans a triangle with probability at least $2/3$. Rather, the size of the sample should depend on the ratio between $d_{\max}$ and d .

Let $s = c \cdot \max(\frac{n^{2/3}}{(\epsilon d)^{1/3}}, \frac{d_{\max}}{\epsilon d})$, where $c > 0$ is a sufficiently large constant. Since G is ϵ -far from being triangle-free, G must contain a family T of $(\epsilon dn)/3$ pairwise edge-disjoint triangles. Fix such a family, and for every $v \in V(G)$, let $d_T(v)$ be the number of triangles in T containing v ; obviously, $d_T(v) \leq d(v)/2 \leq d_{\max}/2$. We sample a set S of s vertices of G uniformly at random. Let X be the random variable counting the number of triangles of T spanned by S . Due to the Chebyshev inequality, it is enough to prove that $\text{Exp}[X]$ is at least a large constant, and the ratio $\text{Var}[X]/\text{Exp}^2[X]$ is at most a small enough constant. We will estimate both quantities.

Observe that each triangle of T falls into S with probability $(1 + o(1))s^3/n^3$. It follows that

$$(42) \quad \text{Exp}[X] = (1 + o(1)) \frac{s^3}{n^3} |T| = \Theta\left(\frac{\epsilon ds^3}{n^2}\right).$$

Thus, taking c large enough, we get that $\text{Exp}[X]$ is large enough, too. Also,

$$(43) \quad \begin{aligned} \text{Var}[X] &\leq \sum_{\substack{t, t' \in T \\ t \cap t' \neq \emptyset}} \Pr[t, t' \subset S] \\ &= \sum_{v \in V(G)} \binom{d_T(v)}{2} \frac{(1 + o(1))s^5}{n^5} \end{aligned}$$

(the latter estimate is due to the fact that, since T is pairwise edge-disjoint, for any $t, t' \in T$ with $t \cap t' \neq \emptyset$, the union $t \cup t'$ contains exactly five vertices). Recall that $d_T(v) \leq d_{\max}$. Due to convexity, we get

$$(44) \quad \text{Var}[X] = O\left(\frac{\epsilon dn}{d_{\max}} \cdot d_{\max}^2\right) \frac{s^5}{n^5}.$$

Using the assumption $s = \Omega(\frac{d_{\max}}{\epsilon d})$, we derive that $\text{Var}[X]/\text{Exp}^2[X]$ is small enough, as required. $\square$

8. Bounds on the edge density of random Cayley graphs with large sets of generators. In this section we consider random Cayley graphs. Behrend graphs are essentially variants of Cayley graphs. Using methods that were introduced

in subsection 5.2, we get bounds on the edge density of random Cayley graphs with large sets of generators. We start with a definition of Cayley graphs. For simplicity we consider only Cayley graphs over Abelian groups, but all arguments apply to the non-Abelian case as well.

Let H be a finite Abelian group. A set $X \subseteq H$ is *symmetric* if $X = -X$, where $-X = \{-x : x \in X\}$. The *Cayley graph* over H with respect to a symmetric set X , denoted CG_X , has H as its vertex set and distinct vertices $a, b \in H$ are connected by an edge if and only if $a - b$ (hence also $b - a$) is in X . We shall say in such a case that the edge *corresponds* to $x = a - b$. All operations involving vertices are performed in H . A *difference* of a set $T \subseteq H$ is an element $a - b$ such that $a, b \in T$. We let $\Delta(T)$ denote the set $\{a - b : a, b \in T\}$ of differences of T . By definition $\Delta(T)$ contains at most $|T|(|T| - 1)$ nonzero elements. The *multiplicity* in T of a difference $x \in \Delta(T)$ is $|\{(a, b) : a, b \in T \text{ and } a - b = x\}|$. Clearly, the differences that correspond to edges that are incident to a specific vertex are all distinct. It follows that the multiplicity in T of any specific difference in $\Delta(T)$ is at most $|T|$. Moreover, if we consider the complete graph over H , then the edges that correspond to a specific difference form a set of cycles that cover H .

In what follows we show that for dense random Cayley graphs the edge density in relatively large induced subgraphs is close to the edge density of the whole graph. It was previously shown [4] that random Cayley graphs are expanders and hence have the property that the density of every induced subgraph on sufficiently many vertices is very close to the density of the graph. However, the known techniques for proving this property are based on estimating the second eigenvalue of the graph's adjacency matrix and do not supply any informative bounds for sets of vertices that are much smaller than the number of vertices divided by the square root of the degree.

We shall use the following notation: For $X, C \subseteq H$, we let $e_X(C)$ denote the number of edges in the subgraph of CG_X that is induced by C .

THEOREM 4. *Let $0 < \beta < \frac{1}{2}$ and $0 < \alpha \leq 1$ satisfy $\alpha - 2\beta > \frac{1}{\log \log n}$, and let H be an Abelian group where $|H| = n$. Let $X \subset H$, $X = -X$, be determined as follows: Every pair $x, -x$ is chosen independently with probability $p(n) = \frac{1}{n^\beta}$ to be in X . With high probability over the choice of X , for every subset C of n^α vertices in CG_X , we have that $e_X(C) \leq \frac{90}{\alpha - 2\beta} \frac{n^{2\alpha}}{n^\beta}$.*

Similarly to Lemma 8, Theorem 4 shows that for sufficiently large subsets C (i.e., for $|C| = n^\alpha$, where $\alpha - 2\beta$ is a constant), the number of edges spanned by C in CG_X is close to its expected value. Here, too, the smaller we choose β (i.e., the larger we choose X), the smaller α can be. That is, the theorem can be applied to sets with smaller size. The proof of Theorem 4 is similar to the proof of Lemma 8. The difference lies in the proof of Claim 18, which is analogous to Claim 9. Here, too, we let $k = \frac{5}{\alpha - 2\beta}$ and note that for β and α as required in the theorem, $k \leq 5 \log \log n$. We shall say that W is *good* if no difference in $\Delta(W)$ has multiplicity higher than k in W . Theorem 4 follows from the next two claims using the same arguments that were applied in showing that Lemma 8 follows from Claims 9 and 10.

CLAIM 17. *With high probability over the choice of X , for every good W such that $|W| = s \geq n^\beta \log n$, we have that $e_X(W) \leq \frac{2ks^2}{n^\beta}$.*

Proof. Consider a fixed choice of a good W such that $|W| = s = n^\beta \log n$. By definition, $|\Delta(W)| \leq |W|^2 = s^2$. Since W is good, for every $x \in \Delta(W)$, if $x \in X$, then the number of edges labeled by x in the subgraph of CG_X induced by W is at most k . Since each $x \in \Delta(W)$ is chosen to be in X independently with probability $n^{-\beta}$, the expected size of $\Delta(W) \cap X$ is $|\Delta(W)| \cdot n^{-\beta} \leq s^2 \cdot n^{-\beta}$. By a multiplicative

Chernoff bound, the probability that $|\Delta(W)| \cap X > 2 \cdot s^2 \cdot n^{-\beta}$ is upper bounded by $\exp(-s^2 \cdot n^{-\beta})$. The claim follows by taking a union bound over all choices of W of size s . $\square$

CLAIM 18. *Let $C \subset H$ satisfy $|C| = n^\alpha$. Suppose we uniformly and independently select $W \subset C$, $|W| = n^\beta \log n$. Then the probability that W is not good is at most $\frac{1}{n^\beta}$.*

Proof. Consider any fixed difference $x \in \Delta(C)$, where there are at most $|C|^2$ such differences. Recall that there are at most $|C|$ edges in the subgraph of CG_H induced by C that correspond to x . We denote this set of edges by $E_x(C)$. Since $E_x(H)$ is a union of disjoint cycles, $E_x(C)$ is a union of disjoint cycles and paths. Therefore, every choice of $k+1$ edges in $E_x(C)$ is a union of r_1 (sub)paths and r_2 cycles where $1 \leq r_1 \leq k+1$ and $r_2 \leq (k+1-r_1)/3$ (the second inequality follows from the fact that the length of a cycle is at least 3). Note that when $r_1 = k+1$, then the $k+1$ edges constitute a matching, as was the case in the proof of Claim 10. In this case the number of vertices incident to them is $2(k+1)$.

More generally, the number of incident vertices is $k+1+r_1$. For each pair (r_1, r_2) , the number of choices of $k+1$ edges that constitute r_1 paths and r_2 cycles is at most $|C|^{r_1}(k+1)^{r_1} \cdot |C|^{r_2}$. Namely, to determine each of the r_1 paths, we select a starting vertex (out of $|C|$ vertices) and a length (between 1 and $k+1$). To determine each of the r_2 cycles, we select a vertex (that belongs to the cycle). Once the edges are selected, the number of choices of the remaining $|W| - (k+1+r_1)$ vertices in W is $\binom{|C|-(k+1+r_1)}{|W|-(k+1+r_1)}$. Therefore, for each fixed choice of r_1 and r_2 , the probability that W spans $k+1$ edges in $E_x(C)$ that constitute r_1 paths and r_2 cycles is at most

$$\begin{aligned}
 & \frac{|C|^{r_1}(k+1)^{r_1} |C|^{r_2} \binom{|C|-(k+1+r_1)}{|W|-(k+1+r_1)}}{\binom{|C|}{|W|}} \leq |C|^{r_1}(k+1)^{r_1} |C|^{r_2} \cdot \left(\frac{|W|}{|C|} \right)^{k+1+r_1} \\
 & = (k+1)^{r_1} \cdot \frac{|W|^{k+1+r_1}}{|C|^{k+1+r_2}} \\
 (45) \quad & < (k+1)^{r_1} \frac{|W|^{k+1+r_1}}{|C|^{(2/3)(k+1)+(1/3)r_1}}.
 \end{aligned}$$

The expression in (45) is maximized when the ratio between $|W|$ and $|C|$ is maximized (i.e., β is maximized with respect to α) and r_1 is maximized (i.e., $r_1 = k+1$). In this case we get an upper bound of $(k+1)^{k+1}(\log n)^{2(k+1)}n^{(2\beta-\alpha)(k+1)}$. Substituting $k = \frac{5}{\alpha-2\beta}$, where $k \leq 5 \log \log n$, summing over all r_1, r_2 , and taking a union bound over the at most $|C|^2 = n^{2\alpha} \leq n^2$ choices of $x \in \Delta(C)$, the claim follows. $\square$

It is worth noting that the technique here can be applied with other parameters as well. In particular, it can be shown, for example, that with high probability, in random Cayley graphs over groups of size n in which the number of generators is $(1+o(1))\frac{n}{2}$, every set X of at least some $\text{poly} \log(n)$ vertices spans $(1+o(1))\frac{1}{2}\binom{|X|}{2}$ edges. This is related to the results in [3] and [11].

Appendix. Proof of Claim 4. Recall that this claim is part of the proof of Lemma 3. Specifically, it is part of the probabilistic analysis of the number of disjoint triangles obtained when drawing a graph according to the distribution D_Δ , which is defined as follows. A graph is generated by first partitioning the vertices into equal-size subsets of size $n' = n/3$, denoted V_1, V_2, V_3 , where the vertices in each V_ℓ , $\ell \in \{1, 2, 3\}$ are denoted $\{v_{\ell,1}, \dots, v_{\ell,n'}\}$. Next, between each pair of subsets, $d' = d/2 = \sqrt{n}/3$

random perfect matchings are selected. Recall that $\eta_{i,j,k}$ is a 0/1 random variable that equals 1 if and only if the graph contains the triangle $(v_{1,i}, v_{2,j}, v_{3,k})$ and there is no other triangle that shares an edge with this triangle. The claim is that, for $i' \neq i, j' \neq j, k' \neq k$,

$$(46) \quad \Pr[\eta_{i',j',k'} = 1 \mid \eta_{i,j,k} = 1] = \left(1 + O\left(\frac{1}{d'}\right)\right) \cdot \Pr[\eta_{i',j',k'} = 1].$$

To prove the claim we break both $\Pr[\eta_{i',j',k'} = 1]$ and $\Pr[\eta_{i',j',k'} = 1 \mid \eta_{i,j,k} = 1]$ into a product of (conditional) probabilities and bound the ratio between each corresponding pair of probabilities. Recall that $\alpha_{i,j}^{\ell,\ell'}$ is a 0/1 random variable that indicates the existence of the edge $(v_{\ell,i}, v_{\ell',j})$, $\Delta_{i,j,k}$ is a 0/1 random variable that indicates the existence of the triangle $(v_{1,i}, v_{2,j}, v_{3,k})$, and $\beta_{i,j,k}$ is a 0/1 random variable that indicates whether there is a triangle different from $(v_{1,i}, v_{2,j}, v_{3,k})$ that contains one of the three edges $(v_{1,i}, v_{2,j})$, $(v_{1,i}, v_{3,k})$, $(v_{2,j}, v_{3,k})$. By definition,

$$(47) \quad \Pr[\eta_{i',j',k'} = 1] = \Pr[\Delta_{i',j',k'} = 1] \cdot \Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1]$$

and

$$(48) \quad \Pr[\eta_{i',j',k'} = 1 \mid \eta_{i,j,k} = 1] = \Pr[\Delta_{i',j',k'} = 1 \mid \eta_{i,j,k} = 1] \cdot \Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1 \ \& \ \eta_{i,j,k} = 1].$$

We first bound the ratio between $\Pr[\Delta_{i',j',k'} = 1 \mid \eta_{i,j,k} = 1]$ and $\Pr[\Delta_{i',j',k'} = 1]$, and then between $\Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1 \ \& \ \eta_{i,j,k} = 1]$ and $\Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1]$.

Before doing so we make a few observations (some of which were already used in the proof of Lemma 3). For illustrations of these observations, see Figure 3. Consider an event of the form $\alpha_{r,s}^{\ell,\ell'} = 1$. We know that $\Pr[\alpha_{r,s}^{\ell,\ell'} = 1] = 1 - \left(1 - \frac{1}{n'}\right)^{d'}$. How is the probability of this event influenced by the existence or nonexistence of other edges? Consider first conditioning on the event that $\alpha_{r',s'}^{\ell,\ell'} = 1$ for $r' \neq r$ and $s' \neq s$. This means that for at least one of the d' matchings between V_ℓ and $V_{\ell'}$, the vertex s' cannot be matched to r (and similarly, r' cannot be matched to s). In other words, in each of these matchings where the vertex s' cannot be matched to r , the probability that r is matched to s is $\frac{1}{n'-1}$ rather than $\frac{1}{n'}$. Therefore, for $r' \neq r, s' \neq s$,

$$(49) \quad \begin{aligned} \Pr\left[\alpha_{r,s}^{\ell,\ell'} = 1 \mid \alpha_{r',s'}^{\ell,\ell'} = 1\right] &\leq 1 - \left(1 - \frac{1}{n'-1}\right)^{d'} \\ &\leq \frac{d'}{n'-1} = \frac{d'}{n'} \cdot \frac{n'}{n'-1}. \end{aligned}$$

That is, the probability that $\alpha_{r,s}^{\ell,\ell'} = 1$ slightly increases when conditioning on $\alpha_{r',s'}^{\ell,\ell'} = 1$. Next, consider conditioning on there *not* being an edge between r' and s for $r' \neq r$, that is, $\alpha_{r',s}^{\ell,\ell'} = 0$. We claim that the probability that $\alpha_{r,s}^{\ell,\ell'} = 1$ increases as well in this case. This is true because $\alpha_{r',s}^{\ell,\ell'} = 0$ means that in each of the d' matchings, r' is matched to some vertex $s' \neq s$, so that the conditional probability that r is matched to s in every matching is $\frac{1}{n'-1}$. Therefore, for $r' \neq r$,

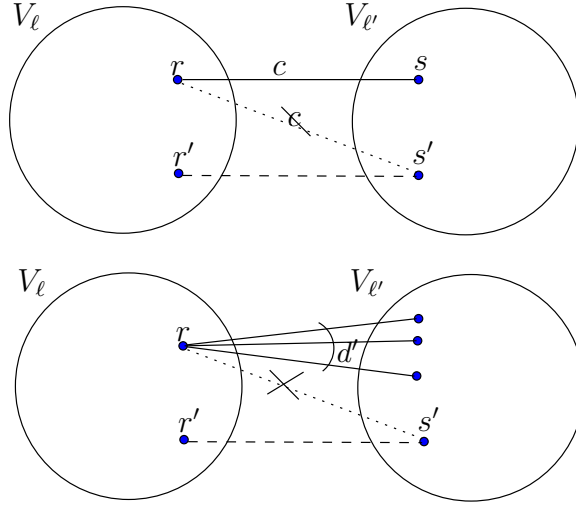


FIG. 3. An illustration of how the existence and nonexistence of edges influences the probability of obtaining another edge. On the top is an illustration of the case in which there is an edge between $v_{\ell,r}$ and $v_{\ell',s}$, and the event we are interested in is the existence of an edge between $v_{\ell,r'}$ and $v_{\ell',s'}$, where $r' \neq r$ and $s' \neq s$. In this illustration the label of the edge between $v_{\ell,r}$ and $v_{\ell',s}$ (the number of the matching among the d' matchings) is c , which implies that, in this matching, $v_{\ell,r}$ cannot be matched to $v_{\ell',s'}$, which increases the probability that $v_{\ell,r'}$ is matched to $v_{\ell',s'}$ in this matching. On the bottom is an illustration of the case that there is no edge between $v_{\ell,r}$ and $v_{\ell',s'}$. This means that in all d' matchings, $v_{\ell,r}$ is matched to other vertices in $V_{\ell'}$, which increases the probability, in every matching, that $v_{\ell,r'}$ is matched to $v_{\ell',s'}$.

$$\begin{aligned}
 \Pr \left[\alpha_{r,s}^{\ell,\ell'} = 1 \mid \alpha_{r',s}^{\ell,\ell'} = 0 \right] &= 1 - \left(1 - \frac{1}{n' - 1} \right)^{d'} \\
 (50) \qquad \qquad \qquad &\leq \frac{d'}{n' - 1} = \frac{d'}{n'} \cdot \frac{n'}{n' - 1}.
 \end{aligned}$$

Finally, we observe that for any $s' \neq s$

$$(51) \qquad \Pr \left[\alpha_{r,s}^{\ell,\ell'} = 1 \mid \alpha_{r',s}^{\ell,\ell'} = 0 \ \& \ \alpha_{r',s'}^{\ell,\ell'} = 1 \right] = \Pr \left[\alpha_{r,s}^{\ell,\ell'} = 1 \mid \alpha_{r',s}^{\ell,\ell'} = 0 \right].$$

This is true since the event $\alpha_{r',s}^{\ell,\ell'} = 0$ implies that $\alpha_{r',s'}^{\ell,\ell'} = 1$ for some $s' \neq s$, and the identity of s' is irrelevant when computing the probability that $\alpha_{r,s}^{\ell,\ell'} = 1$.

Given the above discussion, how does the conditioning on $\eta_{i,j,k} = 1$ (i.e., $\Delta_{i,j,k} = 1$ and $\beta_{i,j,k} = 0$) influence the probability of the event $\Delta_{i',j',k'} = 1$? Consider the event $\alpha_{i',j'}^{1,2} = 1$ (the other two events, $\alpha_{i',k'}^{1,3}$ and $\alpha_{j',k'}^{2,3}$, are analyzed similarly). Since $\Delta_{i,j,k} = 1$, we have that $\alpha_{i,j}^{1,2} = 1$. Since $\beta_{i,j,k} = 0$ we know that for every $j'' \neq j$ either $\alpha_{i,j''}^{1,2} = 0$ or $\alpha_{j'',k}^{2,3} = 0$. In particular, setting $j'' = j'$, we have that either $\alpha_{i,j'}^{1,2} = 0$ or $\alpha_{j',k}^{2,3} = 0$. Recall that we have argued above that conditioning on $\alpha_{i,j'}^{1,2} = 0$ can only increase the probability that $\alpha_{i',j'}^{1,2} = 1$ (assuming $i' \neq i$). Therefore,

$$(52) \qquad \Pr \left[\alpha_{i',j'}^{1,2} = 1 \mid \eta_{i,j,k} = 1 \right] \leq \Pr \left[\alpha_{i',j'}^{1,2} = 1 \mid \alpha_{i,j}^{1,2} = 1 \ \& \ \alpha_{i,j'}^{1,2} = 0 \right]$$

$$(53) \qquad \qquad \qquad \leq \Pr \left[\alpha_{i',j'}^{1,2} = 1 \mid \alpha_{i,j'}^{1,2} = 0 \right]$$

$$(54) \quad \leq \frac{d'}{n'} \cdot \frac{n'}{n' - 1}$$

$$(55) \quad \leq \frac{d'}{n'} \cdot \left(1 + \frac{2}{n'}\right).$$

In (53) we applied (51), in (54) we applied (50), and in (55) we used the assumption that $n' \geq 2$. The same upper bound holds for $\alpha_{j,k}^{2,3}$ and $\alpha_{i,k}^{1,3}$. Using the lower bound on $\Pr[\Delta_{i',j',k'} = 1]$ given in (6), we get

$$(56) \quad \frac{\Pr[\Delta_{i',j',k'} = 1 \mid \eta_{i,j,k} = 1]}{\Pr[\Delta_{i',j',k'} = 1]} \leq \frac{\left(1 + \frac{2}{n'}\right)^3}{\left(1 - \frac{1}{18d'}\right)^3} \leq 1 + O\left(\frac{1}{d'}\right).$$

It remains to upper bound the ratio between $\Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1 \ \& \ \eta_{i,j,k} = 1]$ and $\Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1]$. We start by observing that since the constraints imposed by the conditioning on $\beta_{i,j,k} = 0$ can only increase the probability of *having* a triangle that shares an edge with $(v_{1,i'}, v_{2,j'}, v_{3,k'})$,

$$(57) \quad \begin{aligned} & \Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1 \ \& \ \eta_{i,j,k} = 1] \\ & \leq \Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1 \ \& \ \Delta_{i,j,k} = 1]. \end{aligned}$$

Next, in order to upper bound the ratio between $\Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1 \ \& \ \Delta_{i,j,k} = 1]$ and $\Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1]$, we lower bound the ratio between $\Pr[\beta_{i',j',k'} = 1 \mid \Delta_{i',j',k'} = 1 \ \& \ \Delta_{i,j,k} = 1]$ and $\Pr[\beta_{i',j',k'} = 1 \mid \Delta_{i',j',k'} = 1]$. This suffices since $\Pr[\beta_{i',j',k'} = 1 \mid \Delta_{i',j',k'} = 1] < 1/2$, and for every $y \leq 1/2$, if we have that $x \geq (1 - \alpha)y$ for some $\alpha < 1$, then $(1 - x) \geq (1 + \alpha)(1 - y)$.

The event $\beta_{i',j',k'} = 1$ is a union of events of the form $\Delta_{i',j',k''} = 1$ for $k'' \neq k'$ (the other two types of events, $\Delta_{i',j'',k'} = 1$ and $\Delta_{i'',j',k'} = 1$, are analyzed analogously). The key observation is that for every $k'' \notin \{k', k\}$, the conditioning on the event $\Delta_{i,j,k} = 1$ (in addition to the conditioning on the event $\Delta_{i',j',k'} = 1$) can only increase the probability that $\Delta_{i',j',k''} = 1$ (since for at least one of the matchings i' cannot be matched to k). Therefore, the only case in which the conditioning on $\Delta_{i,j,k} = 1$ decreases the probability that $\Delta_{i',j',k''} = 1$ is for $k'' = k$. However, the weight of the event $\Delta_{i',j',k} = 1$ (and, similarly, $\Delta_{i,j',k'} = 1$ and $\Delta_{i',j,k'}$) relative to $\beta_{i',j',k'} = 1$ is $O(1/n')$, and so

$$(58) \quad \begin{aligned} & \Pr[\beta_{i',j',k'} = 1 \mid \Delta_{i',j',k'} = 1 \ \& \ \Delta_{i,j,k} = 1] \\ & \geq \left(1 - O\left(\frac{1}{n'}\right)\right) \cdot \Pr[\beta_{i',j',k'} = 1 \mid \Delta_{i',j',k'} = 1], \end{aligned}$$

from which it follows that

$$(59) \quad \frac{\Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1 \ \& \ \Delta_{i,j,k} = 1]}{\Pr[\beta_{i',j',k'} = 0 \mid \Delta_{i',j',k'} = 1]} \leq 1 + O\left(\frac{1}{n'}\right).$$

The claim follows by combining (47), (48), (56), (57), and (59).

Acknowledgments. The authors wish to thank Angelika Steger for bringing theses [12] and [16] to their attention. The authors would also like to thank the anonymous reviewer of this paper for helpful comments.

REFERENCES

- [1] N. ALON, *Testing subgraphs in large graphs*, Random Structures Algorithms, 21 (2002), pp. 359–370.
- [2] N. ALON, E. FISCHER, M. KRIVELEVICH, AND M. SZEGEDY, *Efficient testing of large graphs*, Combinatorica, 20 (2000), pp. 451–476.
- [3] N. ALON AND A. ORLITSKY, *Repeated communication and ramsey graphs*, IEEE Trans. Inform. Theory, 41 (1995), pp. 1276–1289.
- [4] N. ALON AND Y. ROICHMAN, *Random Cayley graphs and expanders*, Random Structures Algorithms, 5 (1994), pp. 271–284.
- [5] N. ALON AND A. SHAPIRA, *A characterization of easily testable induced subgraphs*, in Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), ACM, New York, SIAM, Philadelphia, 2004, pp. 942–951.
- [6] N. ALON AND A. SHAPIRA, *Testing subgraphs in directed graphs*, J. Comput. System Sci., 69 (2004), pp. 353–382.
- [7] F. A. BEHREND, *On sets of integers which contain no three terms in arithmetic progression*, Proc. Nat. Acad. Sci. USA, 32 (1946), pp. 331–332.
- [8] E. FISCHER, *The art of uninformed decisions: A primer to property testing*, Bullet. Eur. Assoc. Theoret. Comput. Sci. EATCS, 75 (2001), pp. 97–126.
- [9] O. GOLDBREICH, S. GOLDWASSER, AND D. RON, *Property testing and its connection to learning and approximation*, J. ACM, 45 (1998), pp. 653–750.
- [10] O. GOLDBREICH AND D. RON, *Property testing in bounded degree graphs*, Algorithmica, 32 (2002), pp. 302–343.
- [11] B. GREEN, *Counting sets with small sumset, and the clique number of random Cayley graphs*, Combinatorica, 25 (2005), pp. 307–326.
- [12] L. GUGELMANN, *Testing Triangle-Freeness in General Graphs: Lower Bounds*, Bachelor thesis, Dept. of Mathematics, ETH, Zurich, 2006.
- [13] S. JANSON, T. ŁUCZAK, AND A. RUCIŃSKI, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.
- [14] T. KAUFMAN, M. KRIVELEVICH, AND D. RON, *Tight bounds for testing bipartiteness in general graphs*, SIAM J. Comput., 33 (2004), pp. 1441–1483.
- [15] M. PARNAS AND D. RON, *Testing the diameter of graphs*, Random Structures Algorithms, 20 (2002), pp. 165–183.
- [16] T. RAST, *Testing Triangle-Freeness in General Graphs: Upper Bounds*, Bachelor thesis, Dept. of Mathematics, ETH, Zurich, 2006.
- [17] D. RON, *Property testing*, in Handbook of Randomized Computing, Volume II, S. Rajasekaran, P. M. Pardalos, J. H. Reif, and J. Rolim, eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, 2001, pp. 597–649.
- [18] R. RUBINFELD AND M. SUDAN, *Robust characterizations of polynomials with applications to program testing*, SIAM J. Comput., 25 (1996), pp. 252–271.
- [19] I. Z. RUZSA AND E. SZEMERÉDI, *Triple systems with no six points carrying three triangles*, Coll. Math. Soc. J. Bolyai., 18 (1976), pp. 939–945.
- [20] R. SALEM AND D. C. SPENCER, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. USA, 28 (1942), pp. 561–563.

A NOTE ON REED'S CONJECTURE*

LANDON RABERN†

Abstract. In [*J. Graph Theory*, 27 (1998), pp. 177–212], Reed conjectures that every graph satisfies $\chi \leq \lceil \frac{\omega + \Delta + 1}{2} \rceil$. We prove that this holds for graphs with disconnected complement. Combining this fact with a result of Molloy proves the conjecture for graphs satisfying $\chi > \lceil \frac{n}{2} \rceil$. Generalizing this we prove that the conjecture holds for graphs satisfying $\chi > \frac{n+3-\alpha}{2}$. It follows that the conjecture holds for graphs satisfying $\Delta \geq n + 2 - (\alpha + \sqrt{n+5-\alpha})$. In the final section, we show that if G is an even order counterexample to Reed's conjecture, then $\bar{G}$ has a 1-factor.

Key words. graph coloring, Reed's conjecture, chromatic number

AMS subject classification. 05C15

DOI. 10.1137/060659193

1. Introduction. Reed's conjecture states that every graph satisfies

$$(1) \quad \chi \leq \left\lceil \frac{\omega + \Delta + 1}{2} \right\rceil.$$

The weaker statement that every graph satisfies

$$(2) \quad \chi \leq \frac{\omega + n}{2}$$

follows easily by observing that in any optimal coloring the set of vertices in singleton color classes induces a clique.

We prove the following intermediate result.

KEY LEMMA. $\chi \leq \frac{\omega + \frac{n+\Delta+1}{2}}{2}$.

Using this result it is easy to prove Reed's conjecture for graphs with disconnected complement. Combining this fact with a result of Molloy tells us that Reed's conjecture holds for graphs satisfying $\chi > \lceil \frac{n}{2} \rceil$. Building on this result, we show that the conjecture holds for graphs satisfying $\chi > \frac{n+3-\alpha}{2}$. With a bit more work we determine that the conjecture holds for graphs satisfying $\Delta \geq n + 2 - (\alpha + \sqrt{n+5-\alpha})$.

Further analysis allows us to prove some related results.

In all that follows, *graph* will mean a finite simple graph with nonempty vertex set. Let $\mathbb{G}$ be the collection of all graphs. Let $R_t \subseteq \mathbb{G}$ be the graphs satisfying $\chi \leq \frac{1}{2}(\omega + \Delta + 1) + t$.

DEFINITION 1. Given graphs A and B , their join $A + B$ is the graph with vertex set $V(A) \cup V(B)$ and edge set $E(A) \cup E(B) \cup \{ab \mid a \in V(A), b \in V(B)\}$. Also, if X and Y are collections of graphs, we let $X + Y = \{A + B \mid A \in X, B \in Y\}$.

First, we present a few basic facts about joins.

LEMMA 1.1. Let A and B be graphs. Then

- (a) $|A + B| = |A| + |B|$,
- (b) $\omega(A + B) = \omega(A) + \omega(B)$,
- (c) $\chi(A + B) = \chi(A) + \chi(B)$,
- (d) $\Delta(A + B) = \max\{\Delta(A) + |B|, |A| + \Delta(B)\}$.

*Received by the editors May 7, 2006; accepted for publication (in revised form) September 4, 2007; published electronically April 25, 2008.

<http://www.siam.org/journals/sidma/22-2/65919.html>

†3200 Carbon Place, S-208, Boulder, CO 80301 (london.rabern@gmail.com).

Proof. These all follow immediately from the definitions. $\square$

2. Reed's conjecture for joins of graphs.

PROPOSITION 2.1. *Let A and B be graphs. Then $A + B \in R_0$.*

Proof. Applying the Key Lemma to A and B and adding the inequalities yields

$$\chi(A) + \chi(B) \leq \frac{1}{2} \left(\omega(A) + \omega(B) + \frac{\Delta(A) + |B| + |A| + \Delta(B) + 2}{2} \right).$$

Using Lemma 1.1(b), (c), and (d), this becomes

$$\chi(A + B) \leq \frac{1}{2} \left(\omega(A + B) + \frac{2\Delta(A + B) + 2}{2} \right) = \frac{1}{2} (\omega(A + B) + \Delta(A + B) + 1).$$

Hence $A + B \in R_0$. $\square$

LEMMA 2.2. $\mathbb{G} + R_t \subseteq R_t$ for all $t \in \mathbb{R}$.

Proof. Fix $t \in \mathbb{R}$. Let $G \in \mathbb{G}$ and $H \in R_t$. Applying (2) to G gives

$$\chi(G) \leq \frac{1}{2} (\omega(G) + |G|).$$

Also, since $H \in R_t$,

$$\chi(H) \leq \frac{1}{2} (\omega(H) + \Delta(H) + 1) + t.$$

Adding these inequalities and applying Lemma 2(b) and (c) gives

$$\chi(G + H) \leq \frac{1}{2} (\omega(G + H) + |G| + \Delta(H) + 1) + t.$$

Now Lemma 1.1(d) gives $|G| + \Delta(H) \leq \Delta(G + H)$ and the result follows. $\square$

Remark. We will use Lemma 2.2 in our proof of the Key Lemma; we note that we did not use the Key Lemma in its proof. It should also be noted that Lemma 2.2 just says that R_t is an ideal in the Abelian semigroup $(\mathbb{G}, +)$.

3. Proof of the Key Lemma. Combining Lemma 2.2 with the following two results of Molloy (see [2]) allows us to prove Reed's conjecture for graphs satisfying $\alpha = 2$. We note that these lemmas were proved by Gallai decades before Molloy rediscovered them (see [1]).

LEMMA 3.1. *Let G be a graph with $\chi(G) > \lceil \frac{|G|}{2} \rceil$. Then there exists $X \subseteq V(G)$ such that $\overline{G - X}$ is disconnected and $\chi(G - X) = \chi(G)$.*

As a corollary we get the following.

LEMMA 3.2. *If G is a vertex critical graph with $\chi(G) > \lceil \frac{|G|}{2} \rceil$, then $\overline{G}$ is disconnected.*

PROPOSITION 3.3. *If G is a graph with $\alpha(G) \leq 2$, then $G \in R_{\frac{1}{2}}$.*

Proof. Assume this is not the case and let G be a counterexample with the minimum number of vertices, say $|G| = n$. Since $\alpha(G) \leq 2$, we see that $V(G) \setminus N(v) \cup \{v\}$ induces a clique for each $v \in V(G)$. Hence $\omega(G) \geq n - \delta(G) - 1$ which gives

$$\Delta(G) + \omega(G) + 1 \geq n.$$

So, since $G \notin R_{\frac{1}{2}}$, we have $\chi(G) > \lceil \frac{n}{2} \rceil$.

Now, using minimality of G , we see that G is vertex critical. Thus $\overline{G}$ is disconnected by Lemma 3.2. Hence we have $m \geq 2$ and (nonempty) graphs $C_1, \dots, C_m$ such

that $G = C_1 + \cdots + C_m$. But, for $1 \leq i \leq m$, minimality of G gives $C_i \in R_{\frac{1}{2}}$ since $\alpha(C_i) \leq \alpha(G) \leq 2$ and $|C_i| < n$. Hence $G = C_1 + \cdots + C_m \in R_{\frac{1}{2}}$ by Lemma 2.2. This contradiction completes the proof. $\square$

To prove our Key Lemma, we will also need the following result from [3].

THEOREM 3.4. *Let $I_1, \dots, I_m$ be disjoint independent sets in a graph G . Then*

$$(3) \quad \chi(G) \leq \frac{1}{2} \left(\omega(G) + |G| - \sum_{j=1}^m |I_j| + 2m - 1 \right).$$

Remark. If we increase the bound in this theorem by $\frac{1}{2}$ we get a corollary of (2).

To apply Theorem 3.4 and Proposition 3.3 we need a definition.

DEFINITION 2. *Let G be a graph and r a positive integer. A collection of disjoint independent sets in G , each with at least r vertices, will be called an r -greedy partial coloring of G . A vertex of G is said to be missed by a partial coloring just in case it appears in none of the independent sets.*

We use the following consequence of Theorem 3.4.

COROLLARY 3.5. *Let G be a graph which is not complete and let C be an r -greedy partial coloring of G . Then*

$$\chi(G) \leq \frac{1}{2}(\omega(G) + |G| - (r-2)|C| - 1).$$

Combining Corollary 3.5 for $r = 3$ with the following lemma yields the Key Lemma.

LEMMA 3.6. *Let G be a graph, and of all 3-greedy partial colorings of G , let C be one that misses the minimum number of vertices. Then*

$$\chi(G) \leq \frac{1}{2}(\omega(G) + \Delta(G) + 1) + \frac{|C|+1}{2}.$$

Proof. The first case to consider is when C misses zero vertices. In this case, C is a proper coloring of G , and hence $\chi(G) \leq |C|$. Thus

$$\chi(G) \leq \frac{1}{2}(\chi(G) + |C|) \leq \frac{1}{2}(\Delta(G) + 1 + |C|) \leq \frac{1}{2}(\omega(G) + \Delta(G) + 1) + \frac{|C|+1}{2}.$$

Otherwise, C misses at least one vertex, and by the minimality condition placed on C , each vertex missed by C must be adjacent to at least one vertex in each element of C . Hence $\Delta(G - \cup C) \leq \Delta(G) - |C|$. In addition, $\alpha(G - \cup C) \leq 2$. Thus applying Proposition 3.3 to $G - \cup C$ yields

$$\begin{aligned} \chi(G) &\leq |C| + \chi(G - \cup C) \\ &\leq |C| + \frac{1}{2}(\omega(G - \cup C) + \Delta(G - \cup C) + 1) + \frac{1}{2} \\ &\leq |C| + \frac{1}{2}(\omega(G) + \Delta(G) - |C| + 1) + \frac{1}{2} \\ &= \frac{1}{2}(\omega(G) + \Delta(G) + 1) + \frac{|C|+1}{2}. \quad \square \end{aligned}$$

4. Further details. Proposition 2.1 allows us to prove Reed's conjecture for graphs satisfying $\chi > \lceil \frac{n}{2} \rceil$.

COROLLARY 4.1. *If G is a graph with $\chi(G) > \lceil \frac{|G|}{2} \rceil$, then $G \in R_0$.*

Proof. Let G be a graph with $\chi(G) > \lceil \frac{|G|}{2} \rceil$. Then, by Lemma 3.1, we have $X \subseteq V(G)$ such that $\overline{G-X}$ is disconnected and $\chi(G-X) = \chi(G)$. Since $\overline{G-X}$

is disconnected, there exist graphs A and B such that $G - X = A + B$. Hence, by Proposition 2.1,

$$\chi(G) = \chi(G - X) \leq \frac{1}{2}(\omega(G - X) + \Delta(G - X) + 1) \leq \frac{1}{2}(\omega(G) + \Delta(G) + 1),$$

whence $G \in R_0$. $\square$

We can generalize the above corollary by passing from R_0 to $R_{\frac{1}{2}}$.

COROLLARY 4.2. *If G is a graph with $\chi(G) > \frac{|G|+3-\alpha(G)}{2}$, then $G \in R_{\frac{1}{2}}$.*

Proof. Let G be a graph with $\chi(G) > \frac{|G|+3-\alpha(G)}{2}$ and I an independent set in G with $\alpha(G)$ vertices. Put $H = G \setminus I$. Then

$$\begin{aligned} \chi(H) &\geq \chi(G) - 1 \\ &> \frac{|G| + 3 - \alpha(G)}{2} - 1 \\ &= \frac{|G| + 1 - \alpha(G)}{2} \\ &= \frac{|H| + 1}{2}. \end{aligned}$$

Hence, by Corollary 4.1, we have

$$\chi(H) \leq \frac{\omega(H) + \Delta(H) + 1}{2}.$$

But I is a maximal independent set, and hence each vertex of H is adjacent to at least one vertex in I . In particular, $\Delta(H) \leq \Delta(G) - 1$, whence

$$\chi(G) \leq \chi(H) + 1 \leq \frac{\omega(H) + \Delta(G) - 1 + 1}{2} + 1 \leq \frac{\omega(G) + \Delta(G) + 1}{2} + \frac{1}{2}.$$

The theorem follows. $\square$

Remark. After noting that the conditions on the graph order in both Corollaries 4.1 and 4.2 are sufficient to force the existence of a doubly critical edge in a minimum counterexample to Reed's conjecture, it is natural to ask if similar results can be proved under the weaker condition that the graph contains a doubly critical edge. This is indeed the case, as is shown in [4] and [5]. The former paper generalizes the method of proof using singleton color classes that was sketched in the introduction as justification of (2). It is proved that a graph with more than $\frac{n}{2}$ singleton color classes satisfies Reed's conjecture. The latter paper proves Reed's conjecture for claw-free graphs containing a doubly critical edge and shows that general graphs containing a doubly critical edge satisfy $\chi \leq \frac{1}{3}\omega + \frac{2}{3}(\Delta + 1)$.

COROLLARY 4.3. *Let G be a graph and $t \geq \frac{1}{2}$. If $G \notin R_t$, then $\Delta(G) + 1 \leq |G| - 2t - \omega(G) - \alpha(G) + 2$.*

Proof. This is an immediate consequence of the previous corollary. $\square$

LEMMA 4.4. *Let G be a graph with $\alpha(G) \leq 2$. Then $\omega(G)^2 + \omega(G) \geq |G|$.*

Proof. Let K be a maximal clique in G . Then each vertex of $G - K$ is nonadjacent to at least one vertex in K , and hence some vertex $v \in K$ is nonadjacent to at least $\frac{|G-K|}{|K|}$ vertices. Since $\alpha(G) \leq 2$, the vertices nonadjacent to v form a clique, whence $\omega(G) \geq \frac{|G-K|}{|K|} = \frac{|G|-\omega(G)}{\omega(G)}$, which yields

$$\omega(G)^2 + \omega(G) \geq |G|. \quad \square$$

PROPOSITION 4.5. *If G is a graph with $\Delta(G) \geq |G| + 2 - (\alpha(G) + \sqrt{|G| + 5 - \alpha(G)})$, then $G \in R_{\frac{1}{2}}$.*

Proof. We prove the contrapositive. Let G be a graph with n vertices, maximal degree Δ , clique number ω , and independence number α such that $G \notin R_{\frac{1}{2}}$. Let I be a maximal independent set in G . Let S be a maximal collection of disjoint 3-vertex independent sets of $G - I$. Since $\alpha(G - (\cup S) \cup I) \leq 2$, we may apply Lemma 4.4 to get $\omega^2 + \omega \geq |G - (\cup S) \cup I| = n - \alpha - 3|S|$. Hence

$$(4) \quad |S| \geq \frac{n - \alpha - (\omega^2 + \omega)}{3}.$$

Now, combining the fact that $G \notin R_{\frac{1}{2}}$ with Corollary 3.5, we have $n - \alpha - |S| + 1 > \Delta + 2$. Putting this together with (4), we have

$$n - \alpha - \Delta - 1 > |S| \geq \frac{n - \alpha - (\omega^2 + \omega)}{3},$$

which implies that

$$(5) \quad \Delta < \frac{2n + \omega^2 + \omega - 2\alpha - 3}{3}.$$

By Corollary 4.3, $\omega \leq n - \Delta - \alpha$. Plugging this into (5) and doing a little algebra, we find that $\Delta < n - \alpha + 2 - \sqrt{n - \alpha + 5}$. This completes the proof of the contrapositive. $\square$

COROLLARY 4.6. *If G is a graph with $\Delta(G) \geq |G| - (1 + \sqrt{|G| + 2})$, then $G \in R_{\frac{1}{2}}$.*

Proof. Let G be a graph with $\Delta(G) \geq |G| - (1 + \sqrt{|G| + 2})$. If $\alpha(G) \leq 2$, then $G \in R_{\frac{1}{2}}$ by Proposition 3.3. Otherwise, $\alpha(G) \geq 3$ and $G \in R_{\frac{1}{2}}$ by Proposition 4.5. $\square$

5. Graphs without perfect matchings in their complements. The bound on Δ in Corollary 4.3 works only for $t \geq \frac{1}{2}$. In this section we need bounds on Δ that work for $t = 0$ as well. Using a single independent set of maximal order in Theorem 3.4, we deduce the following.

COROLLARY 5.1. *Let G be a graph. Then*

$$\chi(G) \leq \frac{1}{2}(\omega(G) + |G| - \alpha(G) + 1).$$

COROLLARY 5.2. *Let G be a graph and $t \in \frac{1}{2}\mathbb{Z}$. If $G \notin R_t$, then $\Delta(G) + 1 \leq |G| - 2t - \alpha(G)$.*

Proof. Assume $G \notin R_t$. Applying Corollary 5.1 gives

$$\frac{1}{2}(\omega(G) + \Delta(G) + 1) + t < \chi(G) \leq \frac{1}{2}(\omega(G) + |G| - \alpha(G) + 1).$$

The corollary follows. $\square$

COROLLARY 5.3. *Let G be a graph and $t \geq 0$. If $G \notin R_t$, then $\Delta(G) + 1 \leq |G| - 2t - \omega(G)$.*

Proof. This is an immediate consequence of Corollary 4.1. $\square$

Note that for $t \geq \frac{1}{2}$ in Corollary 5.2, we must have $\alpha(G) \geq 3$ by Proposition 3.3, so the lemma gives $\Delta(G) + 1 \leq |G| - 2t - 3$.

LEMMA 5.4. *If $k \geq 2$ and $G_1, \dots, G_k$ are graphs with $\Delta(G_i) + 1 \leq |G_i| - 3$ for each i , then*

$$G_1 + \dots + G_k \in R_{2-k}.$$

Proof. Assume this is not the case and let $G_1, \dots, G_k$ constitute a counterexample with the smallest k . Then, by Proposition 2.1, $k > 2$. Set $D = G_1 + \dots + G_{k-1}$. Note that $D \in R_{2-(k-1)} = R_{3-k}$ by the minimality of k . Let $t \in \frac{1}{2}\mathbb{Z}$ be minimal such that $G_k \in R_t$. Since $G_k \notin R_{t-\frac{1}{2}}$, using Corollary 5.2 for $t \geq 1$ and the fact that $\Delta(G_k) + 1 \leq |G_k| - 3$ for $t \leq \frac{1}{2}$, we find that $\Delta(G_k) + 1 \leq |G_k| - 2t - 2$. We have

$$\begin{aligned} \chi(D + G_k) &= \chi(D) + \chi(G_k) \\ &\leq \frac{1}{2}(\omega(D) + \omega(G_k) + \Delta(D) + \Delta(G_k) + 2) + 3 - k + t \\ &= \frac{1}{2}(\omega(D + G_k) + \Delta(D) + \Delta(G_k) + 2) + 3 - k + t \\ &= \frac{1}{2}(\omega(D + G_k) + \Delta(D) + |G_k| + 1 + \Delta(G_k) - |G_k| + 1) + 3 - k + t \\ &\leq \frac{1}{2}(\omega(D + G_k) + \Delta(D + G_k) + 1) + \frac{1}{2}(\Delta(G_k) - |G_k| + 1) + 3 - k + t \\ &\leq \frac{1}{2}(\omega(D + G_k) + \Delta(D + G_k) + 1) + \frac{1}{2}(-2t - 3 + 1) + 3 - k + t \\ &= \frac{1}{2}(\omega(D + G_k) + \Delta(D + G_k) + 1) + 2 - k. \end{aligned}$$

Hence $G_1 + \dots + G_k \in R_{2-k}$, contradicting our assumption. $\square$

The hypotheses of this lemma can be weakened, but we do not use the following stronger lemma in what follows.

LEMMA 5.5. *If $k \geq 2$ and $G_1, \dots, G_k$ are graphs, which are not 5-cycles, with $\Delta(G_i) + 1 \leq |G_i| - 2$ for each i , then*

$$G_1 + \dots + G_k \in R_{2-k}.$$

Proof. This is similar to the proof of Lemma 5.4. Graphs with $\Delta(G_i) + 1 \leq |G_i| - 3$ matter only if $G_i \in R_{\frac{1}{2}} \setminus R_0$. Corollary 5.3 shows that such graphs have $\omega(G_i) \leq 2$ and Corollary 5.2 shows they have $\alpha(G) \leq 2$. Thus they have order less than 6 and we see that the only one that breaks the lemma is C_5 . $\square$

DEFINITION 3. *The matching number of a graph G , denoted $\nu(G)$, is the number of edges in a maximum matching of G .*

PROPOSITION 5.6. *Let G be a graph. If $\nu(\overline{G}) < \lfloor \frac{|G|}{2} \rfloor$, then $G \in R_0$.*

Proof. Assume $\nu(\overline{G}) < \lfloor \frac{|G|}{2} \rfloor$. Then, by Tutte's theorem, we have $X \subseteq V(G)$ such that $\overline{G} - X$ has at least m odd components, where $m \geq |X| + 2$ if $|G|$ is even and $m \geq |X| + 3$ if $|G|$ is odd. Hence we have graphs $G_1, \dots, G_m$ such that $G - X = G_1 + \dots + G_m$. Note that by picking one vertex from each component we induce a clique. Hence $\omega(G) \geq m$. To get a contradiction, assume $G \notin R_0$. First, assume there is some G_i for which $\Delta(G_i) + 1 \geq |G_i| - 2$; then

$$\begin{aligned} \Delta(G) + 1 &\geq \Delta(G - X) + 1 \geq |G_1| + \dots + |G_{i-1}| + \Delta(G_i) + |G_{i+1}| + \dots + |G_m| + 1 \\ &\geq |G| - |X| - 2. \end{aligned}$$

Since $G \notin R_0$,

$$\chi(G) > \frac{1}{2}(\omega(G) + \Delta(G) + 1) \geq \frac{1}{2}(m + (|G| - |X| - 2)) \geq \left\lceil \frac{|G|}{2} \right\rceil.$$

Hence $G \in R_0$ by Corollary 4.1! Thus we may assume $\Delta(G_i) + 1 \leq |G_i| - 3$ for each i . Now Lemma 5.4 yields $G - X \in R_{-|X|}$, whence $G \in R_0$. This contradiction completes the proof. $\square$

COROLLARY 5.7. *Let G be an even order graph. If $G \notin R_0$, then $\overline{G}$ has a 1-factor.*

DEFINITION 4. A graph is called *matching covered* if every edge participates in a perfect matching.

COROLLARY 5.8. Let G be an even order graph with $G \notin R_1$. Then $\overline{G}$ is matching covered.

Lemma 5.4 can be generalized.

LEMMA 5.9. Let $m \in \mathbb{N}$. Let $k \geq 2$ and $G_1, \dots, G_k$ be graphs such that $|G_i| < r(m, m) \Rightarrow G_i \in R_{\frac{1}{2}}$. If $\Delta(G_i) + 1 \leq |G_i| - m$ for each i , then

$$G_1 + \dots + G_k \in R_{(m-1)(1-\frac{k}{2})}.$$

Proof. Assume this is not the case and let $G_1, \dots, G_k$ constitute a counterexample with the smallest k . Then, by Proposition 2.1, $k > 2$. Set $D = G_1 + \dots + G_{k-1}$. Note that $D \in R_{(m-1)(1-\frac{k-1}{2})}$ by the minimality of k . Let $t \in \frac{1}{2}\mathbb{Z}$ be minimal such that $G_k \in R_t$. We would like to have $\Delta(G_k) + 1 \leq |G_k| - 2t - (m-1)$. If $t \leq \frac{1}{2}$, then we are all good since $\Delta(G_k) + 1 \leq |G_k| - m$. So, to get a contradiction, assume $t \geq 1$ and $\Delta(G_k) + 1 > |G_k| - 2t - (m-1)$. Then, by Corollary 5.3, $\omega(G_k) \leq (m-1)$. Also, by Corollary 5.2, $\alpha(G_k) \leq (m-1)$. Hence $|G_k| < r(m, m)$, contradicting the fact that $G_k \notin R_{\frac{1}{2}}$. Hence we do indeed have $\Delta(G_k) + 1 \leq |G_k| - 2t - (m-1)$.

We have

$$\begin{aligned} \chi(D + G_k) &= \chi(D) + \chi(G_k) \\ &\leq \frac{1}{2}(\omega(D) + \omega(G_k) + \Delta(D) + \Delta(G_k) + 2) + (m-1)(1 - \frac{k-1}{2}) + t \\ &= \frac{1}{2}(\omega(D + G_k) + \Delta(D) + \Delta(G_k) + 2) + (m-1)(1 - \frac{k-1}{2}) + t \\ &= \frac{1}{2}(\omega(D + G_k) + \Delta(D) + |G_k| + 1 + \Delta(G_k) - |G_k| + 1) \\ &\quad + (m-1)(1 - \frac{k-1}{2}) + t \\ &\leq \frac{1}{2}(\omega(D + G_k) + \Delta(D + G_k) + 1) + \frac{1}{2}(\Delta(G_k) - |G_k| + 1) \\ &\quad + (m-1)(1 - \frac{k-1}{2}) + t \\ &\leq \frac{1}{2}(\omega(D + G_k) + \Delta(D + G_k) + 1) + \frac{1}{2}(-2t - 4 + 1) \\ &\quad + (m-1)(1 - \frac{k-1}{2}) + t \\ &= \frac{1}{2}(\omega(D + G_k) + \Delta(D + G_k) + 1) + (m-1)(1 - \frac{k}{2}). \end{aligned}$$

Hence $G_1 + \dots + G_k \in R_{(m-1)(1-\frac{k}{2})}$, contradicting our assumption. $\square$

We can do a bit better than Lemma 5.9 in the following special case.

LEMMA 5.10. If $k \geq 2$ and $G_1, \dots, G_k$ are noncomplete graphs, then

$$G_1 + \dots + G_k \in R_{1-\frac{k}{2}}.$$

Proof. Assume this is not the case and let $G_1, \dots, G_k$ constitute a counterexample with the smallest k . Then, by Proposition 2.1, $k > 2$. Set $D = G_1 + \dots + G_{k-1}$. Note that $D \in R_{1-\frac{(k-1)}{2}}$ by the minimality of k . Let $t \in \frac{1}{2}\mathbb{Z}$ be minimal such that $G_k \in R_t$. Since $G_k \notin R_{t-\frac{1}{2}}$, if $t \geq \frac{1}{2}$, then, by Corollary 5.3, $\Delta(G_k) + 1 \leq |G_k| - 2t - 1$. If $t \leq 0$ and $\Delta(G_k) + 1 > |G_k| - 2t - 1$, then $t = 0$ and $\Delta(G_k) + 1 = |G_k|$; however, Corollary 5.1 shows that the only such graphs are complete graphs, which we have excluded,

whence $\Delta(G_k) + 1 \leq |G_k| - 2t - 1$. We have

$$\begin{aligned}
 \chi(D + G_k) &= \chi(D) + \chi(G_k) \\
 &\leq \frac{1}{2}(\omega(D) + \omega(G_k) + \Delta(D) + \Delta(G_k) + 2) + 1 - \frac{(k-1)}{2} + t \\
 &= \frac{1}{2}(\omega(D + G_k) + \Delta(D) + \Delta(G_k) + 2) + 1 - \frac{(k-1)}{2} + t \\
 &= \frac{1}{2}(\omega(D + G_k) + \Delta(D) + |G_k| + 1 + \Delta(G_k) - |G_k| + 1) + 1 - \frac{(k-1)}{2} + t \\
 &\leq \frac{1}{2}(\omega(D + G_k) + \Delta(D + G_k) + 1) + \frac{1}{2}(\Delta(G_k) - |G_k| + 1) + 1 - \frac{(k-1)}{2} + t \\
 &\leq \frac{1}{2}(\omega(D + G_k) + \Delta(D + G_k) + 1) + \frac{1}{2}(-2t - 2 + 1) + 1 - \frac{(k-1)}{2} + t \\
 &= \frac{1}{2}(\omega(D + G_k) + \Delta(D + G_k) + 1) + 1 - \frac{k}{2}.
 \end{aligned}$$

Hence $G_1 + \cdots + G_k \in R_{1-\frac{k}{2}}$, contradicting our assumption. $\square$

Combining Reed's conjecture with Lemma 5.9 proves the following conjecture.

CONJECTURE 5.11. *Let $m \in \mathbb{N}$. If $k \geq 2$ and $G_1, \dots, G_k$ are graphs with $\Delta(G_i) + 1 \leq |G_i| - m$ for each i , then*

$$G_1 + \cdots + G_k \in R_{(m-1)(1-\frac{k}{2})}.$$

Acknowledgments. The referee's comments helped to greatly improve this paper's layout and readability. Thanks to Michael Molloy for pointing out that Lemmas 3.1 and 3.2 had already been proven by Gallai long before he rediscovered them.

REFERENCES

- [1] T. GALLAI, *Kritische Graphen. II*, Magyar Tud. Akad. Mat. Kutató Int. Közl., 8 (1963), pp. 373–395.
- [2] M. MOLLOY, *Chromatic neighborhood sets*, J. Graph Theory, 31 (1999), pp. 303–311.
- [3] L. RABERN, *On graph associations*, SIAM J. Discrete Math., 20 (2006), pp. 529–535.
- [4] L. RABERN, *Coloring and The Lonely Graph*, <http://www.arxiv.org/abs/0707.1069> (2007).
- [5] L. RABERN, *Coloring graphs containing a doubly critical edge*, J. Graph Theory, submitted.
- [6] B. REED, ω , Δ , and χ , J. Graph Theory, 27 (1998), pp. 177–212.

THE MINIMUM NUMBER OF DISTINCT AREAS OF TRIANGLES DETERMINED BY A SET OF n POINTS IN THE PLANE*

ROM PINCHASI†

Abstract. We prove a conjecture of Erdős, Purdy, and Straus on the number of distinct areas of triangles determined by a set of n points in the plane. We show that if P is a set of n points in the plane, not all on one line, then P determines at least $\lfloor \frac{n-1}{2} \rfloor$ triangles with pairwise distinct areas. Moreover, one can find such $\lfloor \frac{n-1}{2} \rfloor$ triangles all sharing a common edge.

Key words. triangles, distinct areas, Euler's formula

AMS subject classification. 52C10

DOI. 10.1137/07068299X

1. Introduction. Let P be a set of n points in the plane. We consider all the triangles whose vertices are any three noncollinear points of P . We regard these triangles as triangles determined by the set P . Denote by $g(P)$ the number of distinct areas of triangles determined by P . Clearly, $g(P)$ can be as large as $\binom{n}{3}$ if P is in a general enough position. We are interested in the minimum possible value of $g(P)$. Denote by $g(n)$ the minimum possible value of $g(P)$ over all noncollinear sets P of n points in the plane.

The function $g(n)$ was first investigated by Erdős and Purdy [EP77]. They showed that there exist positive absolute constants c_1, c_2 such that for every n we have

$$c_1 n^{3/4} \leq g(n) \leq c_2 n.$$

It is easy to observe an $\lfloor \frac{n-1}{2} \rfloor$ upper bound for $g(n)$ by taking P to be a set of equally spaced $\lfloor \frac{n}{2} \rfloor$ points on a line ℓ together with $\lceil \frac{n}{2} \rceil$ equally spaced points on a line ℓ' parallel to ℓ (see [BP79] and [S78]).

Erdős, Purdy, and Straus conjectured the following.

CONJECTURE 1.1. *For every n , $g(n) \geq \lfloor \frac{n-1}{2} \rfloor$.*

The problem of bounding the function $g(n)$ from below is listed, for instance, in the problem collection by Croft, Falconer, and Guy [CFG91] as well as in the problem collection by Brass, Moser, and Pach [BMP05].

A linear lower bound for $g(n)$ was obtained by Burton and Purdy in 1979 [BP79]. They showed $g(n) \geq 0.32n$ by relating this problem to another problem on the minimum number of distinct directions determined by a noncollinear point set of size n . A complete solution of the latter problem was found only a little later in 1982 by Ungar [U82]. Combining the result in [U82] with the technique of [BP79], one obtains the lower bound $g(n) \geq (\sqrt{2} - 1)n - \mathcal{O}(1)$.

Recently, Dumitrescu and Tóth [DT] improved the lower bound and showed $g(n) \geq \frac{17}{38}n - \mathcal{O}(1) \approx 0.4473n$.

In this paper we prove Conjecture 1.1.

THEOREM 1.2. *The number of distinct areas of triangles determined by a noncollinear point set of size n is at least $\lfloor \frac{n-1}{2} \rfloor$.*

*Received by the editors February 18, 2007; accepted for publication (in revised form) September 1, 2007; published electronically April 25, 2008.

<http://www.siam.org/journals/sidma/22-2/68299.html>

†Israel Institute of Technology, Technion, Haifa, Israel (room@math.technion.ac.il).

We note that all the previously mentioned results, as well as the result in Theorem 1.2, are in fact a little stronger. The lower bounds refer even to triangles that have distinct areas but that all share a common edge.

2. The proof. We show the existence of a line γ determined by at least two points of the set P such that there are at least $\lfloor \frac{n-1}{2} \rfloor$ distinct positive distances of points of P from γ . This will clearly prove the theorem.

It will be more convenient for us to consider the dual problem in the plane. That is, we perform a standard duality transform which takes a point (a, b) to the line $y + ax + b = 0$ in the plane, and it takes a nonvertical line $y + a'x + b' = 0$ to the point (a', b') .

It is easy to observe that if a point p lies on a line L , then the dual of p is a line containing the point which is the dual of L . It is also easy to see that two points with the same x -coordinate in the dual plane represent the dual of two parallel lines in the primal plane. Finally, observe that if a line $y + a'x + b' = 0$ lies above a point (a, b) (that is, if $b + a'a + b' < 0$), then the point (a', b') , dual to the line, lies below the line $y + ax + b = 0$, which is the dual of (a, b) .

Therefore, we rotate the set P , if necessary, so that no two points of P have the same x -coordinate, and then the dual of P is a collection $\mathcal{M}$ of n lines, no two of which are parallel and not all passing through the same point (for otherwise the points of P are all collinear). Consider now an intersection point q of two or more lines in $\mathcal{M}$. q is the dual of a line γ containing two or more points of P . Draw a vertical line r through q and let $x_1, \dots, x_k$ be all the intersection points of r with lines in $\mathcal{M}$, which lie above q on r . For each $i = 1, \dots, k$ let ℓ_i be a line in $\mathcal{M}$ passing through x_i . Assume that for every $1 \leq i \leq k$, ℓ_i is the dual of the point $p_i \in P$. Then the points $p_1, \dots, p_k$ lie strictly below γ , because the lines $\ell_1, \dots, \ell_k$ lie strictly above q . Moreover, for every $1 \leq i < j \leq k$, the line through p_i and p_j cannot be parallel to γ , because ℓ_i and ℓ_j do not meet on r . It follows that the distances from $p_1, \dots, p_k$ to γ are distinct. A similar scenario holds if we consider the intersection points of r with lines in $\mathcal{M}$, which are all strictly below q on r .

Theorem 1.2 therefore follows from the following theorem.

THEOREM 2.1. *Let $\mathcal{A}$ be an arrangement of n blue nonvertical lines in the plane such that no two of the lines are parallel, and not all the lines pass through the same point. Draw vertical red lines such that each red line passes through an intersection point of (at least) two blue lines, and every intersection point of two (or more) blue lines lies on a red line. Then there is an intersection point p of at least two blue lines, and a red line ℓ through p such that either there are at least $\lfloor \frac{n-1}{2} \rfloor$ different intersection points of ℓ with blue lines such that all are above p on ℓ , or there are at least $\lfloor \frac{n-1}{2} \rfloor$ different intersection points of ℓ with blue lines, all which lie below p on ℓ .*

Proof. Let $\mathcal{L}$ denote the collection of red lines, and let $m = |\mathcal{L}|$.

For every line $\ell \in \mathcal{L}$ let $b(\ell)$ denote the intersection point of (at least) two blue lines on ℓ with the least y -coordinate. Similarly, let $t(\ell)$ denote the intersection point of (at least) two blue lines on ℓ with the largest y -coordinate (it is possible that $b(\ell) = t(\ell)$). Denote by $S(\ell)$ the set of all intersection points of ℓ with blue lines that are either strictly below $b(\ell)$ or strictly above $t(\ell)$. Denote by $Q(\ell)$ the set of all intersection points of ℓ with a single blue line that are strictly above $b(\ell)$ and strictly below $t(\ell)$. Finally, denote by $R(\ell)$ the set of all intersection points of (at least) two blue lines on ℓ .

Observe that there are at least $S(\ell)/2 + Q(\ell) + R(\ell) - 1$ intersection points on ℓ that are either all above $b(\ell)$ or all below $t(\ell)$.

Let $S = \cup_{\ell \in \mathcal{L}} S(\ell)$, $Q = \cup_{\ell \in \mathcal{L}} Q(\ell)$, and $R = \cup_{\ell \in \mathcal{L}} R(\ell)$.

Let p and q be two intersection points in $R(\ell)$ for some red line ℓ . We say that p and q are adjacent if there is no point $s \in R(\ell)$ on the open line segment between p and q (on ℓ). Since there are exactly m lines and each includes at least one point from R , we deduce that there are exactly $|R| - m$ pairs of adjacent intersection points from R . Let p and q be a pair of adjacent intersection points on a red line ℓ . We say that the pair (p, q) is *nice* if there is no blue line separating p from q . Consequently, if p and q are adjacent but (p, q) is not a nice pair, then there must be a point in Q on the open line segment of ℓ between p and q . Therefore, the number of nice pairs of adjacent intersection points is at least $|R| - m - |Q|$.

Consider now the arrangement $\mathcal{A}$ of the blue lines only. We denote, as usual, by f_k the number of faces in the arrangement with precisely k edges. It is important to note that we consider the arrangement $\mathcal{A}$ on the *projective* plane, and so we identify the two ends at infinity for each of the blue lines. For $k \geq 2$ denote by t_k the number of intersection points through which there are precisely k blue lines. Observe that since not all the blue lines pass through the same point, there is no face with two edges in $\mathcal{A}$. Recall (see for example [F04, p. 73]) that from Euler's formula it follows that

$$(1) \quad 3 + \sum_{k \geq 2} (k-3)t_k + \sum_{k \geq 3} (k-3)f_k = 0.$$

One crucial observation is that if (p, q) is a nice pair, then p and q are vertices of the same face in $\mathcal{A}$. Moreover, the line segment between p and q is a diagonal of that face. For a face F with k edges the maximum possible number of diagonals in the vertical direction is $\lfloor \frac{k-2}{2} \rfloor$. Observe that this number is always less than or equal to $k-3$, as long as $k \geq 3$, which is true in our case. It follows that the maximum number of nice pairs (p, q) is at most $\sum_{k \geq 3} (k-3)f_k$. This, together with the lower bound on the number of nice pairs, gives

$$(2) \quad |R| - m - |Q| \leq \sum_{k \geq 3} (k-3)f_k.$$

Observe also that by definition,

$$(3) \quad |R| = \sum_{k \geq 2} t_k.$$

Equation (2), together with (3), gives

$$(4) \quad |Q| \geq \sum_{k \geq 2} t_k - m - \sum_{k \geq 3} (k-3)f_k.$$

We are now ready to complete the proof. Since every red line crosses each of the n blue lines we conclude that $|S| + |Q| = mn - \sum_{k \geq 2} kt_k$.

Therefore,

$$\begin{aligned} \frac{|S|}{2} + |Q| + |R| &= \frac{|S| + |Q|}{2} + \frac{|Q|}{2} + |R| \\ &\geq \frac{1}{2} \left(mn - \sum_{k \geq 2} kt_k + \sum_{k \geq 2} t_k - m - \sum_{k \geq 3} (k-3)f_k \right) + \sum_{k \geq 2} t_k \\ &= \frac{1}{2} \left(m(n-1) - \left(\sum_{k \geq 2} (k-3)t_k + \sum_{k \geq 3} (k-3)f_k \right) \right) = \frac{m(n-1) + 3}{2}. \end{aligned}$$

Since there are precisely m red lines, it follows by averaging that there exists a red line ℓ such that

$$\frac{|S(\ell)|}{2} + |Q(\ell)| + |R(\ell)| \geq \frac{n-1}{2} + \frac{3}{2m}.$$

It follows that either there are at least $\lceil \frac{n-1}{2} + \frac{3}{m} - 1 \rceil = \lfloor \frac{n-1}{2} \rfloor$ intersection points of ℓ with blue lines, all of which lie below $t(\ell)$, or there are at least this many intersection points of blue lines with ℓ , all of which lie above $b(\ell)$. $\square$

REFERENCES

- [BMP05] P. BRASS, W. MOSER, AND J. PACH, *Research Problems in Discrete Geometry*, Springer, New York, 2005.
- [BP79] G. R. BURTON AND G. B. PURDY, *The directions determined by n points in the plane*, J. London Math. Soc. (2), 20 (1979), pp. 109–114.
- [CFG91] H. CROFT, K. FALCONER, AND R. GUY, *Unsolved Problems in Geometry*, Springer-Verlag, New York, 1991.
- [DT] A. DUMITRESCU AND CS. D. TÓTH, *Distinct triangle areas in a planar point set*, in Integer Programming and Combinatorial Optimization, Lecture Notes in Comput. Sci. 4513, Springer, Berlin, 2007, pp. 119–129.
- [EP77] P. ERDŐS AND G. PURDY, *Some extremal problems in geometry*. V, in Proceedings of the Eighth Southeastern Conference on Combinatorics, Graph Theory and Computing (Louisiana State Univ., Baton Rouge, LA, 1977), Congressus Numeration XIX, Utilitas Math., Winnipeg, Man., 1977, pp. 569–578.
- [F04] S. FELSNER, *Geometric Graphs and Arrangements. Some Chapters from Combinatorial Geometry*, Adv. Lectures Math., Friedr. Vieweg & Sohn, Wiesbaden, 2004.
- [S78] E. G. STRAUS, *Some extremal problems in combinatorial geometry*, in Combinatorial Mathematics (Proc. Internat. Conf. Combinatorial Theory, Australian Nat. Univ., Canberra, 1977), Lecture Notes in Math. 686, Springer, Berlin, 1978, pp. 308–312.
- [U82] P. UNGAR, *$2N$ noncollinear points determine at least $2N$ directions*, J. Combin. Theory Ser. A, 33 (1982), pp. 343–347.

ON THE COMPLEXITY OF ORDERED COLORINGS*

ARVIND GUPTA[†], JAN VAN DEN HEUVEL[‡], JÁN MAŇUCH[†], LADISLAV STACHO[§],
AND XIAOHONG ZHAO[†]

Abstract. We introduce two variants of proper colorings with imposed partial ordering on the set of colors. One variant shows very close connections to some fundamental problems in graph theory, e.g., directed graph homomorphism and list colorings. We study the border between tractability and intractability for both variants.

Key words. graph coloring, ordered coloring, directed graph homomorphism, complexity, NP-completeness, polynomial algorithm

AMS subject classifications. 05C15, 68Q25, 05C85

DOI. 10.1137/060676222

1. Introduction. We introduce two variants of proper colorings with imposed partial ordering on the set of colors. Vertices of all considered graphs G are labeled with integers from 1 to $|V(G)|$, and we will use the normal order of the integers. In particular, the vertices form a completely ordered set. The set of colors forms a partially ordered set. The following ordered colorings are proper colorings satisfying additional requirements.

In the first coloring problem, we require that every vertex colored A is smaller than every vertex colored B for every two colors A, B for which A is smaller than B in the partial order. We will show that this problem is in P if the set of colors contains at most two incomparable colors by reduction to 2-SAT; otherwise, it is intractable.

In the second coloring problem, we require that for every edge whose end vertices are colored by A and B that the vertex with color A is smaller than the vertex with color B , for every two colors A, B for which A is smaller than B in the partial order. Note that a vertex colored A can be greater than a vertex colored B , provided they are not adjacent. We show that this problem is NP-complete if the partial ordering on colors contains at least two incomparable pairs. Otherwise, the problem is in P.

2. Basic definitions. Most of our terminology and notation will be standard and can be found in any textbook on graph theory such as [2], and on computational complexity such as [4].

Throughout this paper a graph $G = (V, E)$ is finite, simple, and loopless. We will always use n for the number of vertices and assume that the vertices of the input

*Received by the editors November 29, 2006; accepted for publication (in revised form) November 13, 2007; published electronically April 25, 2008. The first author's research was supported in part by an NSERC (Natural Science and Engineering Research Council of Canada) grant. The third author's research was supported in part by PIMS (Pacific Institute for Mathematical Sciences) and by Science and Technology Assistance Agency grant APVT-20-018902 (Slovak grant agency). The fourth author's research was supported in part by an NSERC (Natural Science and Engineering Research Council of Canada) grant, and VEGA grant 2/3164/23 (Slovak grant agency).

<http://www.siam.org/journals/sidma/22-2/67622.html>

[†]School of Computing Science, Simon Fraser University, Burnaby, BC V5A 1S4 Canada (arvind@sfu.ca, jmanuch@sfu.ca, xzhao2@sfu.ca).

[‡]Department of Mathematics, London School of Economics, London, WC2A 2AE, UK (jan@maths.lse.ac.uk).

[§]Department of Mathematics, Simon Fraser University, Burnaby, BC V5A 1S6 Canada (lstacho@sfu.ca).

graph G are integers from 1 to n (in other words, we set $V = [1, n]$). With this assumption we can consider the vertex set as a linearly ordered set $(V, \leq)$.

The set of available colors is denoted by Col . A *coloring* of a graph G is a function $\varphi : V \rightarrow Col$ so that for every edge $uv \in E$ we have $\varphi(u) \neq \varphi(v)$. We additionally assume that $\preceq$ is an antisymmetric, reflexive, and transitive relation on Col ; i.e., $\mathcal{C} = (Col, \preceq)$ is a partially ordered set. We sometimes call $(Col, \preceq)$ the *color-poset*. We will study the complexity of two coloring problems in which we will color vertices of the input graph G with colors in Col satisfying further requirements. We will refer to a coloring that satisfies these requirements as a *feasible coloring*.

Let $C \in Col$, where the set V_C of vertices in G colored by C will be called a *color class*.

If $(P, \preceq)$ is a poset, then an *antichain* is a set $S \subseteq P$ so that no two elements from S are comparable (i.e., for all $x, y \in S$ we have $x \not\preceq y$ and $y \not\preceq x$).

Given $X, Y \subseteq V$, an edge with one end-vertex in X and the other in Y is called an X, Y -edge. Let $E(X, Y)$ be the set of all X, Y -edges.

3. Coloring with ordered color classes. In this section we study the complexity of the following decision problem.

PROBLEM 3.1. *Fix a color-poset $(Col, \preceq)$. Given a graph $G = (V, E)$ with $V = [1, n]$, determine whether G can be colored with Col such that for any two colors A, B with $A \preceq B$ and for any two vertices $u \in V_A$ and $v \in V_B$ we have $u \leq v$.*

If a coloring as in Problem 3.1 exists, then we say that the graph G can be *feasibly colored with $(Col, \preceq)$* .

Note that if $\preceq$ is the empty relation on Col , then Problem 3.1 is the well-known graph coloring problem.

THEOREM 3.1. *If the poset $(Col, \preceq)$ does not contain an antichain of size 3, then Problem 3.1 can be solved in polynomial time. Otherwise, the problem is NP-complete.*

The proof that Problem 3.1 is in P if the longest antichain in $(Col, \preceq)$ has length at most 2 is quite technical. For this reason we first give the proof for a specific small poset with all of the details. The proof of the general case will be done with less detail after that. The fairly straightforward proof that the problem is NP-complete if $(Col, \preceq)$ contains an antichain of length 3 will be the final proof in this section.

LEMMA 3.2. *Let $Col = \{A, B, C\}$ and suppose that the only relation between the colors is $A \preceq B$. Problem 3.1 can be solved in polynomial time for this poset $(\{A, B, C\}, \preceq)$.*

Proof. Suppose a graph G can be feasibly colored with $(\{A, B, C\}, \preceq)$. It then follows that there exists a vertex $v \in V$ such that vertices in the set $V_{AC} = [1, v-1]$ are colored A or C , and vertices in the set $V_{BC} = [v, n]$ are colored B or C . Therefore, the subgraphs G_{AC} and G_{BC} of G , induced by V_{AC} and V_{BC} , respectively, are bipartite. Suppose G_{AC} has k components and G_{BC} has ℓ components. Let $\{(X_i, \overline{X}_i)\}_{i=1}^k$ be the set of bipartitions of connected components $\{\mathcal{X}_i\}_{i=1}^k$ of G_{AC} , and let $\{(Y_j, \overline{Y}_j)\}_{j=1}^\ell$ be the set of bipartitions of connected components $\{\mathcal{Y}_j\}_{j=1}^\ell$ of G_{BC} .

Every bipartite component $\mathcal{X}_i$ (respectively, $\mathcal{Y}_j$) has exactly two 2-colorings. Thus, we can associate each $\mathcal{X}_i$ with a Boolean variable x_i as follows: If the bipartition X_i of $\mathcal{X}_i$ is colored C , then $x_i = 0$; otherwise, $x_i = 1$. Similarly, we associate each $\mathcal{Y}_j$ with a Boolean variable y_j : If the bipartition Y_j of $\mathcal{Y}_j$ is colored C , then $y_j = 0$; otherwise, $y_j = 1$.

Every (V_{AC}, V_{BC}) -edge of G imposes a constraint on a feasible coloring of G . This constraint can be equivalently expressed in terms of the Boolean variables as follows:

- For every (X_i, Y_j) -edge of G , the clause $x_i \vee y_j$ should be satisfied.
- For every $(X_i, \overline{Y_j})$ -edge of G , the clause $x_i \vee \overline{y_j}$ should be satisfied.
- For every $(\overline{X_i}, Y_j)$ -edge of G , the clause $\overline{x_i} \vee y_j$ should be satisfied.
- For every $(\overline{X_i}, \overline{Y_j})$ -edge of G , the clause $\overline{x_i} \vee \overline{y_j}$ should be satisfied.

Let f_v be the conjunction of the above 2-literal disjunctions for all (V_{AC}, V_{BC}) -edges of G ; see Figure 3.1 for an example.

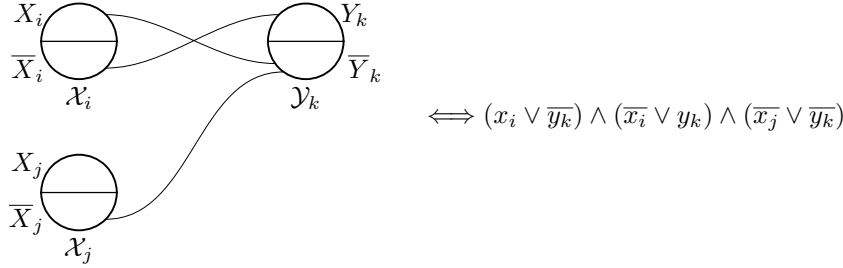


FIG. 3.1. An example of the transformation of a graph G into a 2-CNF formula.

Obviously, there exists a feasible coloring of G if and only if the formula f_v is satisfiable. An assignment of Boolean values to variables in f_v can be translated into a coloring of G in linear time. Hence, to find a coloring for G feasible with $(\{A, B, C\}, \preceq)$, we can use the following algorithm:

- For each $v \in V$, check if the subgraphs induced on $[1, v-1]$ and on $[v, n]$ are bipartite. If so, construct the formula f_v and use a 2-SAT solver to check the satisfiability of the formula. The graph G has a feasible coloring if and only if for at least one $v \in V$ the subgraphs induced on $[1, v-1]$ and on $[v, n]$ are bipartite and the formula f_v is satisfiable.

In [3], a 2-SAT solver working in linear time in the size of the formula is described. Note that each f_v has $O(n)$ variables and $O(n^2)$ clauses. Since there are n choices for v , Problem 3.1 can be solved in $O(n^3)$ steps. $\square$

We next give the general proofs of the two parts of Theorem 3.1.

THEOREM 3.1 A. *If the poset $(Col, \preceq)$ does not contain an antichain of size 3, then Problem 3.1 can be solved in polynomial time.*

Proof. If $\mathcal{P} = (P, \preceq)$ is a poset and $u \in P$, then by $(P - u, \preceq)$ we denote the poset with the same ordering $\preceq$ on the set $P \setminus \{u\}$. For a poset $\mathcal{P}$ with $|P| = p$ elements that does not contain an antichain of size 3, we will construct a collection $\Lambda_{\mathcal{P}}$ of p -tuples $\bar{\lambda} = (\lambda_1, \dots, \lambda_p)$, where each λ_i is a set of one or two elements from P . The concatenation of two tuples $\bar{\lambda} = (\lambda_1, \dots, \lambda_k)$ and $\bar{\mu} = (\mu_1, \dots, \mu_\ell)$ is $\bar{\lambda} * \bar{\mu} = (\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_\ell)$.

The construction of $\Lambda_{\mathcal{P}}$ is recursive, as follows:

- If P has one element, say $P = \{a\}$, then $\Lambda_{\mathcal{P}} = \{\{a\}\}$.
- If $|P| \geq 2$ and P has a maximum element a , then $\Lambda_{\mathcal{P}} = \{\bar{\lambda} * (\{a\}) \mid \bar{\lambda} \in \Lambda_{\mathcal{P}-a}\}$.
- If $|P| \geq 2$ and P has no maximum element, then, since the longest antichain has length 2, there are two maximal elements a, b (so $a \not\preceq b$ and $b \not\preceq a$, while $p \preceq a$ or $p \preceq b$ for all $p \in P$). In that case we set $\Lambda_{\mathcal{P}} = \{\bar{\lambda} * (\{a, b\}) \mid \bar{\lambda} \in \Lambda_{\mathcal{P}-a}\} \cup \{\bar{\lambda} * (\{a, b\}) \mid \bar{\lambda} \in \Lambda_{\mathcal{P}-b}\}$.

Note that the number of tuples in $\Lambda_{\mathcal{P}}$ is finite.

Now we go back to a color poset $\mathcal{C} = (Col, \preceq)$ with no antichain longer than 3. Set $c = |Col|$. Consider a graph $G = (V, E)$ with $V = [1, n]$. Suppose G can be feasibly

colored with $\mathcal{C}$. With any such coloring we can associate a c -tuple $\bar{\lambda} = (\lambda_1, \dots, \lambda_c) \in \Lambda_{\mathcal{C}}$ and a $(c+1)$ -tuple of integers $(v_1, v_2, \dots, v_{c+1})$ so that the following hold:

1. We have $1 = v_1 \leq v_2 \leq \dots \leq v_c \leq v_{c+1} = n+1$.
2. For all $i = 1, \dots, c$, the subgraph of G induced by $[v_i, v_{i+1} - 1]$ is an independent set or a bipartite graph.
3. For all $i = 1, \dots, c$, the colors appearing on the vertices $[v_i, v_{i+1} - 1]$ are exactly the colors in λ_i .

The existence of these tuples follows directly from the fact that each antichain in $\mathcal{C}$ has length 1 or 2, and hence we cannot “mix” three or more colors from Col when coloring the graph.

Given a $(c+1)$ -tuple of integers $(v_1, v_2, \dots, v_{c+1})$, it is easy to check if this tuple satisfies conditions 1 and 2 above. If it does, then we can choose a c -tuple $\bar{\lambda} = (\lambda_1, \dots, \lambda_c) \in \Lambda_{\mathcal{C}}$ and use the same technique as applied in the proof of Lemma 3.2 to transform the problem of the existence of a feasible coloring of G in accordance with condition 3 into the existence of a solution of a certain 2-SAT problem. Hence, the question if there exists a coloring of G in accordance with two chosen tuples satisfying conditions 1–3 can be done in time polynomial in n .

The number of $(c+1)$ -tuples $(v_1, v_2, \dots, v_{c+1})$ with $1 = v_1 \leq v_2 \leq \dots \leq v_c \leq v_{c+1} = n+1$ is $O(n^c)$. And the number of tuples in $\Lambda_{\mathcal{C}}$ is $O(2^c)$. Therefore, to check if a graph G on n ordered vertices $[1, n]$ can be feasibly colored with $\mathcal{C}$, we need to consider at most $O(n^c)$ 2-SAT problems. Using estimates for the numbers of variables and clauses in each of the 2-SAT problems similar to those in the proof of Lemma 3.2, we can conclude that Problem 3.1 can be solved in polynomial time. $\square$

THEOREM 3.1 B. *If the poset $(Col, \preceq)$ contains an antichain of size 3 or more, then Problem 3.1 is NP-complete.*

Proof. Let $\mathcal{C} = (Col, \preceq)$ be a poset containing an antichain of size 3 or more. It is obvious that Problem 3.1 is in NP. To prove the problem is NP-complete we give a reduction from the proper K -coloring problem, which is well known to be NP-complete for any fixed $K \geq 3$.

Let $S \subseteq Col$ be an antichain in $\mathcal{C}$ with $|S| \geq 3$, chosen such that S is *maximal*; i.e., for all $C \in Col \setminus S$ we have that either $C \preceq A$ for some $A \in S$, or $A' \preceq C$ for some $A' \in S$, but not both (since S is an antichain). Let $Col_D \subseteq Col$ be the set of colors $C \in Col \setminus S$ such that $C \preceq A$ for some $A \in S$, and define $Col_U \subseteq Col$ similarly for colors in $Col \setminus S$ which are larger than some color in S . Set $K = |S|$, $n_D = |Col_D|$, and $n_U = |Col_U|$.

Given a graph $G' = (V', E')$ on n' vertices, we construct a graph G with ordered vertex set $[1, n]$, where $n = n_D + n' + n_U$, so that G' has a K -coloring if and only if G can be feasibly colored with $\mathcal{C}$. Let the vertices of G be $[1, n_D] \cup [n_D + 1, n_D + n'] \cup [n_D + n' + 1, n_D + n' + n_U]$. Add edges so that the graph on $[n_D + 1, n_D + n']$ is isomorphic to G' , and vertices in $[1, n_D] \cup [n_D + n' + 1, n_D + n' + n_U]$ are universal (joined to all other vertices).

Suppose G' has a K -coloring. Then we can color G in accordance with $\mathcal{C}$ as follows:

- Give each vertex in $[1, n_D]$ its own color from Col_D , using some linear extension of the order imposed by $\mathcal{C}$ on Col_D .
- Color the vertices in $[n_D + 1, n_D + n']$ with colors from S , according to the K -coloring possible on G' .
- Give each vertex in $[n_D + n' + 1, n_D + n' + n_U]$ its own color from Col_U , using some linear extension of the order imposed by $\mathcal{C}$ on Col_U .

It is easy to check that this coloring of G is feasible with $\mathcal{C}$, where the crucial observation is that S is an antichain in $\mathcal{C}$, and hence every proper coloring with colors from S is always in accordance with the poset order.

Next suppose that G has a feasible coloring with $\mathcal{C}$. Such a coloring must have the following properties:

- Each vertex in $[1, n_D]$ has a distinct color from Col_D , and this color is smaller than any color appearing on $[n_D + 1, n_D + n'] \cup [n_D + n' + 1, n_D + n' + n_U]$.
- Each vertex in $[n_D + n' + 1, n_D + n' + n_U]$ has a distinct color from Col_U , and this color is larger than any color appearing on $[1, n_D] \cup [n_D + 1, n_D + n']$.

From this it follows that the vertices in $[n_D + 1, n_D + n']$ are colored with colors from S . Since S is an antichain, the only requirement to color those vertices with S is that it must be a proper coloring. Such a proper coloring immediately gives a K -coloring of G' . $\square$

4. Coloring with ordered edges. In this section we study the complexity of the following decision problem.

PROBLEM 4.1. *Fix a color-poset $(Col, \preceq)$. Given a graph $G = (V, E)$ with $V = [1, n]$, determine whether G can be colored with Col such that for any two colors A, B with $A \preceq B$ and for any edge $(u, v) \in E$ with $u \in V_A$ and $v \in V_B$ we have $u \leq v$.*

We will transform this problem into a directed graph homomorphism problem. In this section, by (u, v) we mean a directed edge (arc) from u to v . We will use E to denote the set of directed edges of a directed graph (digraph) as well.

The semicomplete digraphs depicted in Figure 4.1 play a crucial role in our characterization.

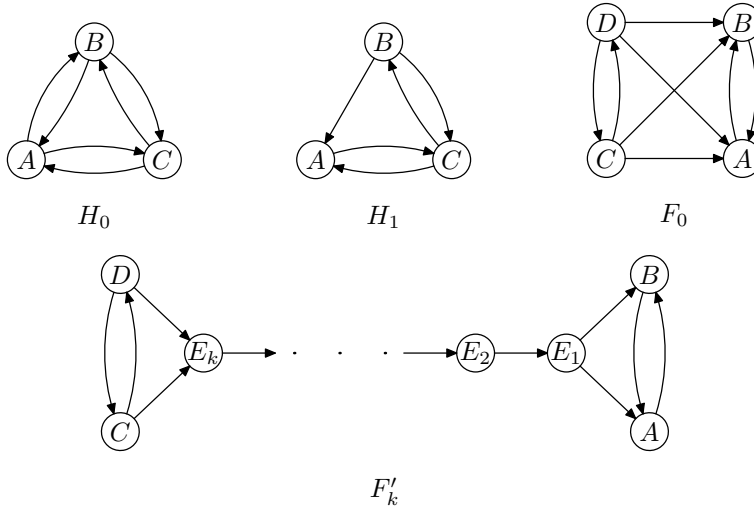


FIG. 4.1. The crucial semicomplete digraphs in the dichotomy characterization of Problem 4.1. F'_k is obtained by taking the transitive closure of F_k .

DEFINITION 4.1. *Let H be a fixed digraph. An H -coloring of a digraph G is a homomorphism from G to H (a mapping φ from $V(G)$ to $V(H)$ such that if $(x, y) \in E(G)$, then $(\varphi(x), \varphi(y)) \in E(H)$). The acyclic H -coloring problem asks whether for a given acyclic digraph G there is an H -coloring of G .*

DEFINITION 4.2. A semicomplete digraph is a digraph H which contains a spanning tournament; i.e., between any two vertices there is either one arc or both opposite arcs. We say that a semicomplete digraph is quasi-acyclic if after removal of all pairs of oppositely directed arcs we obtain a transitive acyclic subdigraph of H .

An alternative definition of a quasi-acyclic semicomplete digraph is as a semicomplete digraph in which every directed cycle contains at least two arcs that come from pairs of opposite arcs. It is easy to see that there is a one-to-one correspondence between posets and quasi-acyclic semicomplete digraphs. Given a color-poset $\mathcal{C} = (Col, \preceq)$, to obtain the corresponding digraph take $V = Col$. And for every two colors A and B , $A \neq B$, if $A \prec B$ add an arc (A, B) , if $B \prec A$ add an arc (B, A) , and, otherwise, add a pair of opposite arcs (A, B) and (B, A) . The resulting semicomplete digraph is quasi-acyclic and will be denoted $H_{\mathcal{C}}$.

Similarly, a graph G with ordered vertex set $[1, n]$ can be considered as a digraph G' if we replace all edges $uv \in E(G)$ with $u < v$ by arcs (u, v) .

Without a proof we note that, with this notation, we have that G has a coloring by $\mathcal{C}$ feasible for Problem 4.1 if and only if there is an $H_{\mathcal{C}}$ -coloring of G' .

THEOREM 4.3. Let H be a quasi-acyclic semicomplete digraph. The acyclic H -coloring problem can be solved in polynomial time if H contains at most one directed 2-cycle (a pair of oppositely directed arcs). Otherwise, the problem is NP-complete.

This is exactly the same dichotomy as for the general H -coloring problem where H is any semicomplete digraph; cf. [1]. In particular, the first part of the theorem follows from the polynomiality of this problem. But for convenience we give a short proof below as well (Lemma 4.4).

The second part of the theorem follows from Lemmas 4.5–4.10 that make up most of the remainder of this section. In those lemmas we first prove the NP-completeness of the acyclic H -coloring problem for three crucial quasi-acyclic semicomplete digraphs and then for any quasi-acyclic semicomplete digraph containing any of those three as an induced subdigraph.

LEMMA 4.4. Let H be a quasi-acyclic semicomplete digraph. The acyclic H -coloring problem can be solved in polynomial time if H contains at most one directed 2-cycle.

Proof. The problem is trivial if $|V(H)| = 1$. If $|V(H)| = 2$ and there is a 2-cycle in H , then an acyclic digraph G has a homomorphism from G to H if and only if the underlying graph of G (obtained by ignoring the directions on the arcs) is bipartite. Similarly, if $|V(H)| = 2$ and there is no 2-cycle in H , then an acyclic digraph G has a homomorphism from G to H if and only if we can partition $V(G)$ into $G_D \cup G_U$ such that every arc in G has its tail in G_D and its head in G_U .

So assume $|V(H)| \geq 3$. Since H has an acyclic spanning tournament, there are both a universal sink and a universal source in H , i.e., there is a vertex $u \in V(H)$ so that for all $v \in V(H)$, $v \neq u$, we have $(v, u) \in E(H)$; and there is a vertex $u' \in V(H)$ so that for all $v \in V(H)$, $v \neq u'$, we have $(u', v) \in E(H)$. Since H has at most one 2-cycle $\{(p, q), (q, p)\}$ and after removing that 2-cycle we obtain a transitive acyclic subdigraph of H , it is easy to check that we can choose at least one of u, u' different from both p and q . Suppose that the universal sink u in H is different from p and q . Let G be an acyclic digraph, and let G_i be the set of sinks in G . Then there is a homomorphism from G to H if and only if there is a homomorphism from $G - G_i$ to $H - u$. Indeed, homomorphisms from $G - G_i$ to $H - u$ are easily extendable to homomorphisms from G to H . The converse follows since in any homomorphism from G to H , the neighbors of vertices in G_i are not mapped to u , and hence we can

remap the vertices in G_i to u if necessary.

A similar observation holds if the universal source in H is different from p and q . We use induction on $|V(H)|$ to get a straightforward polynomial algorithm to check if there is a homomorphism from H to G . $\square$

The following lemma follows immediately from the NP-completeness of the ordinary 3-coloring problem.

LEMMA 4.5. *Let H_0 be the quasi-acyclic semicomplete digraph depicted in Figure 4.1. The acyclic H_0 -coloring problem is NP-complete.*

LEMMA 4.6. *Let H_1 be the quasi-acyclic semicomplete digraph depicted in Figure 4.1. The acyclic H_1 -coloring problem is NP-complete.*

Proof. The problem is obviously in NP. We will show that given a Boolean formula $f(x_1, x_2, \dots, x_m)$ in conjunctive normal form (CNF), we can construct an acyclic digraph $\tilde{G}(f)$ which has an H_1 -coloring φ if and only if f is satisfiable. For every Boolean variable a of f (respectively, every auxiliary variable encoding the Boolean value of a subformula of f), we will construct the gadget $G_a = (V_a, E_a)$ defined as follows: $V_a = \{v_a^1, \dots, v_a^6\}$ and $E_a = \{(v_a^2, v_a^1), (v_a^3, v_a^1), (v_a^3, v_a^2), (v_a^4, v_a^2), (v_a^5, v_a^3), (v_a^5, v_a^4), (v_a^6, v_a^1), (v_a^6, v_a^4), (v_a^6, v_a^5)\}$; see Figure 4.2.

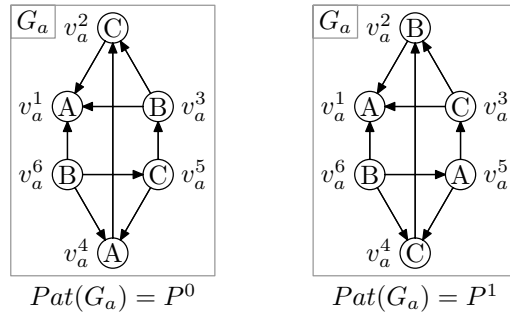


FIG. 4.2. The two possible homomorphisms (coloring patterns) from G_a to H_1 : On the left-hand side $Pat(G_a) = P^0$; On the right-hand side $Pat(G_a) = P^1$.

Each gadget G_a is acyclic and has only two H_1 -colorings, called *color patterns*; see Figure 4.2. Let the color pattern P^0 of G_a with $\varphi(v_a^1) = \varphi(v_a^4) = A$, $\varphi(v_a^3) = \varphi(v_a^6) = B$, and $\varphi(v_a^2) = \varphi(v_a^5) = C$ represent the fact that the value of the Boolean variable a is 0. Similarly, let the color pattern P^1 of G_a with $\varphi(v_a^1) = \varphi(v_a^5) = A$, $\varphi(v_a^2) = \varphi(v_a^6) = B$, and $\varphi(v_a^3) = \varphi(v_a^4) = C$ represent the fact that $a = 1$.

Now let us give an overview of the main ideas in the construction of $\tilde{G}(f)$. First, we construct a gadget G_{x_i} for every variable x_i occurring in the formula f . Second, for every disjunctive subformula d_i of $f = d_1 \wedge d_2 \wedge \dots \wedge d_k$, we construct an acyclic digraph $G(d_i)$ containing a gadget G_{d_i} and using the gadgets $G_{x_1}, \dots, G_{x_m}$ such that for every H_1 -coloring of $G(d_i)$, $Pat(G_{d_i}) = P^1$ if and only if color patterns of $G_{x_1}, \dots, G_{x_m}$ correspond to a true assignment for d_i . Finally, we extend each $G(d_i)$ to a digraph $\tilde{G}(d_i)$ such that the gadget G_{d_i} will have the color pattern P^1 in every H_1 -coloring of $\tilde{G}(d_i)$. Since the gadgets $G_{x_1}, \dots, G_{x_m}$ are common for every $\tilde{G}(d_i)$, it will follow that there is an H_1 -coloring of $\tilde{G}(f) = \bigcup_{i=1}^k \tilde{G}(d_i)$ if and only if there is a true assignment for f , i.e., if f is satisfiable.

In the following we describe a recursive construction of the digraph $G(d)$, where $d = d_i$. If d has no $\vee$'s, then either $d = x_j$ or $d = \overline{x_j}$, for some $j = 1, \dots, m$. As

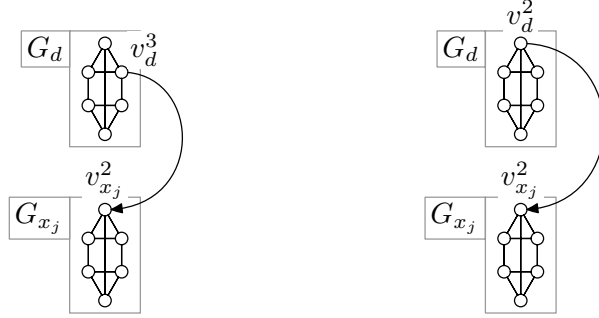


FIG. 4.3. On the left: $G(d)$ for the Boolean formula $d = x_j$; on the right: $G(d)$ for the Boolean formula $d = \bar{x}_j$.

described earlier, add a gadget G_d to $G(d)$. To complete the construction add the arc $(v_d^3, v_{x_j}^2)$ if $d = x_j$, and the arc $(v_d^2, v_{x_j}^2)$ if $d = \bar{x}_j$; see Figure 4.3. First, suppose $d = x_j$. If the color pattern of G_{x_j} is P^0 , then because of the arc $(v_d^3, v_{x_j}^2)$, the only possible color pattern for G_d is P^0 , and similarly if the color pattern of G_{x_j} is P^1 , then the only possible color pattern for G_d is P^1 . Since every true assignment of d has $x_j = 1$ and the remaining variables can have arbitrary values, for every H_1 -coloring of $G(d)$, $\text{Pat}(G_d) = P^1$ if and only if color patterns of $G_{x_1}, \dots, G_{x_m}$ correspond to such a true assignment. If $d = \bar{x}_j$, then the argument is similar.

Assume now that d contains an $\vee$. Then there are disjunctive subformulas d' and d'' in f such that $d = d' \vee d''$. We recursively construct digraphs $G(d')$ and $G(d'')$ (in polytime). Construct $G(d)$ as follows. Take a union of $G(d')$ and $G(d'')$ and add a new variable gadget G_d and a disjunction gadget consisting of three new vertices w_d^1, w_d^2, w_d^3 and the arcs $(w_d^2, w_d^1), (w_d^3, w_d^2)$. Furthermore, add the arcs $(w_d^1, v_d^2), (w_d^1, v_d^5), (w_d^3, v_d^2), (w_d^3, v_d^5)$, and (v_d^4, w_d^2) connecting the disjunction gadget with the variable gadgets; see Figure 4.4.

It is easy to check that in every H_1 -coloring of $G(d')$ (respectively, $G(d'')$), the gadget $G_{d'}$ in $G(d')$ (respectively, $G_{d''}$ in $G(d'')$) has the color pattern P^1 if and only if the color patterns of gadgets $G_{x_1}, \dots, G_{x_m}$ represent a true assignment for d' (respectively, d''). Otherwise it has the color pattern P^0 . Since the color patterns of gadgets $G_{x_1}, \dots, G_{x_m}$ represent a true assignment for d if and only if they represent a true assignment for d' or d'' , it is enough to show that in any H_1 -coloring of $G(d)$, $\text{Pat}(G_d) = P^1$ if and only if $\text{Pat}(G_{d'}) = P^1$ or $\text{Pat}(G_{d''}) = P^1$.

Consider an H_1 -coloring of $G(d)$. If $\text{Pat}(G_{d'}) = \text{Pat}(G_{d''}) = P^1$, then $\varphi(w_d^1) = \varphi(w_d^3) = C$ and $\varphi(w_d^2) = A$ or B ; hence $\text{Pat}(G_d) = P^1$. (Note that this is true even if $\varphi(w_d^2) = B$; the vertex v_d^4 cannot be colored A because of the arc (v_d^4, w_d^2) .) If $\text{Pat}(G_{d'}) = P^1$ and $\text{Pat}(G_{d''}) = P^0$, then $\varphi(w_d^1) = C$, and since (w_d^3, w_d^2) is an arc, $\varphi(w_d^2) = A$ and $\varphi(w_d^3) = B$. Hence, $\text{Pat}(G_d) = P^1$. The case $\text{Pat}(G_{d'}) = P^0$ and $\text{Pat}(G_{d''}) = P^1$ is analogous to the previous one. If $\text{Pat}(G_{d'}) = \text{Pat}(G_{d''}) = P^0$, then $\varphi(w_d^1)$ and $\varphi(w_d^3)$ is either A or B . Since $G(d)$ contains the arcs (w_d^2, w_d^1) and (w_d^3, w_d^2) , we must have $\varphi(w_d^2) = C$; hence $\text{Pat}(G_d) = P^0$. This verifies the construction of the digraph $G(d)$.

Finally, we extend each $G(d)$ to a digraph $\tilde{G}(d)$ such that the gadget G_d will have the color pattern P^1 in every H_1 -coloring of $\tilde{G}(d)$. Then, for every H_1 -coloring of $\tilde{G}(d)$, the color patterns of $G_{x_1}, \dots, G_{x_m}$ must represent a true assignment for d .

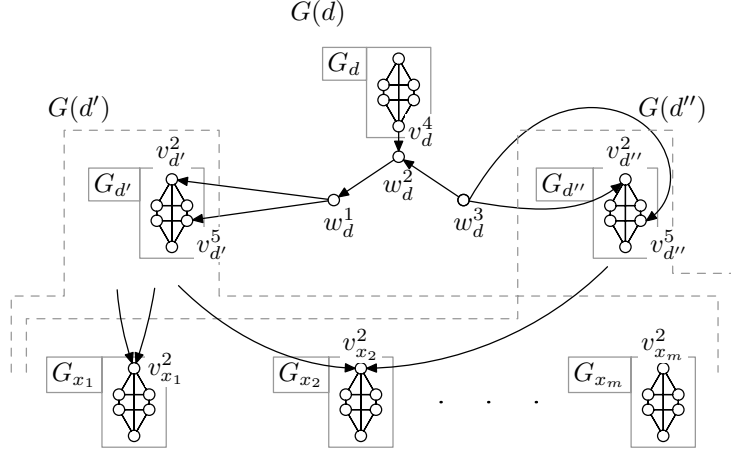


FIG. 4.4. The construction of $G(d)$ for the Boolean formula $d = d' \vee d''$ using the union of the digraphs $G(d')$ and $G(d'')$. In the example above we assume that d' contains both x_1 and $\bar{x}_1$, which is indicated by two arcs entering the gadget G_{x_1} . Similarly, the arcs entering the gadget G_{x_2} indicate that the variable x_2 (or its complement) occurs in both d' and d'' . Finally, no arc entering the gadget G_{x_m} indicates that neither d' nor d'' contain x_m or $\bar{x}_m$.

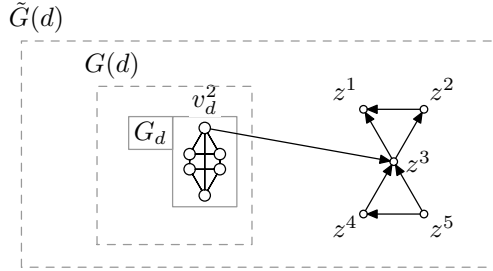
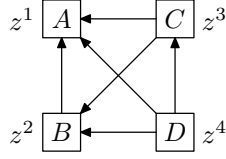


FIG. 4.5. The construction of $\tilde{G}(d)$.

It follows that $\tilde{G}(f)$ has an H_1 -coloring if and only if f is satisfiable.

Let Z be the digraph with $V(Z) = \{z^1, \dots, z^5\}$, and $E(Z) = \{(z^2, z^1), (z^3, z^2), (z^4, z^3), (z^5, z^4), (z^1, z^5)\}$. Since z^3 has two incoming (outgoing) arcs from vertices connected by an arc, it cannot be colored B (respectively, A). Hence, z^3 must be colored C . Consequently, $\varphi(z^1) = \varphi(z^4) = A$ and $\varphi(z^2) = \varphi(z^5) = B$. Now, we are ready to construct $\tilde{G}(d)$. Take the union of $G(d)$ and Z and add an arc (v_d^2, z^3) ; see Figure 4.5. Obviously, in any H_1 -coloring of $\tilde{G}(d)$, $\text{Pat}(G_d) = P^1$.

Finally, let us deduce that the digraph $\tilde{G}(f)$ is acyclic. First, by induction, we show that for every subformula d used in the construction, $G(d)$ is acyclic. This is certainly true for every $d = x_j$ or $d = \bar{x}_j$. Now, consider the formula $d = d' \vee d''$; see Figure 4.4. By induction, $G(d')$ and $G(d'')$ are acyclic and since all arcs incident with the common part of $G(d')$ and $G(d'')$ (the input variable gadgets) end in the common part, the union $G(d') \cup G(d'')$ is acyclic as well. The digraph $G(d)$ contains this union and two new gadgets, and all arcs connecting these two parts start in the new gadgets and end in the union. Hence, $G(d)$ is also acyclic. Thus, for every d_i in f , $G(d_i)$ and, obviously, also $\tilde{G}(d_i)$ are acyclic. By the same argument as earlier their union $\tilde{G}(f)$

FIG. 4.6. The digraph Z .

is also acyclic. $\square$

LEMMA 4.7. Let F_0 be the quasi-acyclic semicomplete digraph depicted in Figure 4.1. The acyclic F_0 -coloring problem is NP-complete.

Proof. The problem is obviously in NP. The proof of NP-hardness follows the lines of the proof of Lemma 4.6. We will again show that given a Boolean formula $f(x_1, x_2, \dots, x_m)$ in CNF, we can construct an acyclic digraph $\tilde{G}(f)$ which has a F_0 -coloring if and only if f is satisfiable.

The main differences with the proof of Lemma 4.6 is that the new variable gadgets will share a four-vertex subdigraph Z with $V(Z) = \{z^1, \dots, z^4\}$ and $E(Z) = \{(z_i, z_j) \mid i > j\}$; see Figure 4.6. Note that the digraph Z has a unique F_0 -coloring up to swapping A and B , or C and D . Without loss of generality, it is enough to consider only those F_0 -colorings φ of $\tilde{G}(f)$ which have $\varphi(z^1) = A$, $\varphi(z^2) = B$, $\varphi(z^3) = C$, and $\varphi(z^4) = D$. This assumption will reduce the number of color patterns of variable gadgets to two.

For every Boolean variable a of f (input or auxiliary), we will construct the following acyclic gadget $G_a = (V_a, E_a)$ with $V_a = \{v_a^1, \dots, v_a^6\}$ and $E_a = \{(v_a^2, v_a^1), (v_a^3, v_a^2), (v_a^4, v_a^3), (v_a^5, v_a^4), (v_a^6, v_a^5), (v_a^6, v_a^3), (v_a^6, v_a^5)\}$. As proposed above, we complete the construction of the gadget by connecting it to the common digraph Z by arcs (v_a^3, z^1) , (v_a^4, z^1) , (v_a^5, z^1) , (v_a^6, z^1) , (v_a^2, z^2) , (z^3, v_a^3) , and (z^3, v_a^4) ; see Figure 4.7. The figure shows two coloring patterns of G_a : P^0 with $\varphi(v_a^1) = \varphi(v_a^3) = B$, $\varphi(v_a^2) = A$, $\varphi(v_a^4) = \varphi(v_a^6) = D$, and $\varphi(v_a^5) = C$, which will represent the fact that $a = 0$; and P^1 with

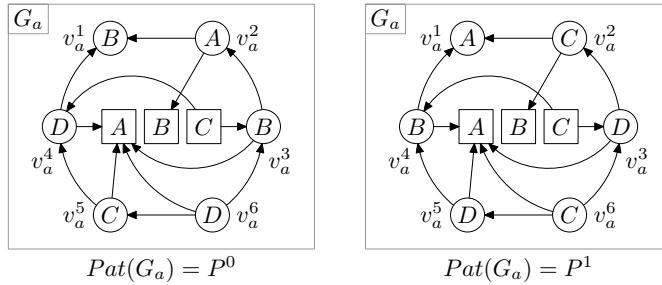


FIG. 4.7. Two coloring patterns of the gadget G_a : $\text{Pat}(G_a) = P^0$ (on the left) and $\text{Pat}(G_a) = P^1$ (on the right). The squares labeled with A , B , and C represent vertices z^1 , z^2 , and z^3 of the digraph Z common to all gadgets, respectively. For simplicity, we omitted the fourth vertex of Z and the arcs inside Z .

$\varphi(v_a^1) = A$, $\varphi(v_a^2) = \varphi(v_a^6) = C$, $\varphi(v_a^3) = \varphi(v_a^5) = D$, and $\varphi(v_a^4) = B$ corresponding with $a = 1$.

For each disjunctive subformula $d = d_i$ of $f = d_1 \wedge \dots \wedge d_k$, the construction of $\tilde{G}(d)$ is analogous to the construction in the proof of Lemma 4.6. Therefore, we describe only the main ingredients of the construction: the base step $d = x_j$ or $d = \bar{x}_j$,

the inductive step $d = d' \vee d''$, and forcing the color pattern of G_d to P^1 . In the base step, add the arc $(v_{x_j}^6, v_d^5)$, if $d = x_j$; or the arc $(v_{x_j}^6, v_d^6)$, if $d = \overline{x_j}$, joining the input variable gadget G_{x_j} to the variable gadget G_d ; see Figure 4.8.

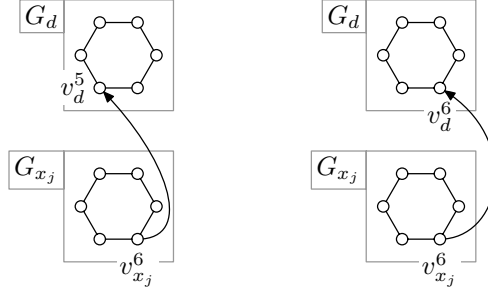


FIG. 4.8. On the left: $G(d)$ for the Boolean formula $d = x_j$; on the right: $G(d)$ for the Boolean formula $d = \overline{x_j}$.

One can easily check that $G(d)$ is acyclic and has the required property.

In the inductive step, construct $G(d)$ as follows. Take the union of $G_{d'}$ and $G_{d''}$ and add one new gadget G_d and three new vertices w_d^1, w_d^2, w_d^3 . Furthermore, add the arcs $(v_{d'}^2, w_d^1)$, $(v_{d'}^5, w_d^2)$, $(v_{d'}^5, v_d^4)$, $(w_d^3, v_{d''}^5)$, $(v_{d''}^5, v_d^4)$, (w_d^2, w_d^1) , (w_d^2, v_d^1) , and (w_d^3, w_d^2) ; see Figure 4.9.

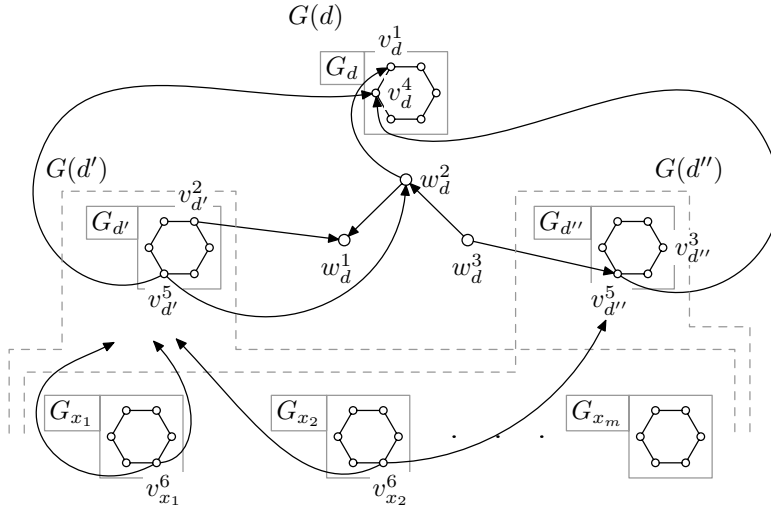


FIG. 4.9. The construction of $G(d)$ for the Boolean formula $d = d' \vee d''$ using the union of the digraphs $G(d')$ and $G(d'')$.

We show that in any feasible coloring of $G(d)$, $\text{Pat}(G_d) = P^1$ if and only if $\text{Pat}(G_{d'}) = P^1$ or $\text{Pat}(G_{d''}) = P^1$. Notice that the arcs $(v_{d'}^2, w_d^1)$ and $(v_{d'}^5, v_d^4)$ guarantee that if one of the gadgets $G_{d'}$ or $G_{d''}$ has the color pattern P^1 , then gadget G_d must have the color pattern P^1 . Now, it is enough to show that there is a feasible coloring for the three w -vertices in each of these three cases. Indeed, we can color the w -vertices as follows: if $\text{Pat}(G_{d'}) = P^1$ and $\text{Pat}(G_{d''}) = P^0$, $\varphi(w^1) = A$, $\varphi(w^2) = B$, $\varphi(w^3) = D$;

if $\text{Pat}(G_{d'}) = P^0$ and $\text{Pat}(G_{d''}) = P^1$, $\varphi(w^1) = B$, $\varphi(w^2) = D$, $\varphi(w^3) = C$; and if $\text{Pat}(G_{d'}) = P^1$ and $\text{Pat}(G_{d''}) = P^1$, $\varphi(w^1) = A$, $\varphi(w^2) = B$, $\varphi(w^3) = C$. It remains to consider the case when $\text{Pat}(G_{d'}) = P^0$ and $\text{Pat}(G_{d''}) = P^0$. We have $\varphi(v_{d'}^2) = A$ and $\varphi(v_{d'}^5) = \varphi(v_{d''}^5) = C$. The arc $(v_{d'}^2, w_d^1)$ forces $\varphi(w_d^1) = B$, and the arc $(w_d^3, v_{d''}^5)$ forces $\varphi(w_d^3) = D$. Furthermore, the arcs $(v_{d'}^2, w_d^1)$, $(w_d^3, v_{d''}^5)$, and $(v_{d''}^5, w_d^2)$ force $\varphi(w_d^2) = A$ because of the arc (w_d^2, v_d^1) , $\varphi(v_d^1) = B$. Hence, $\text{Pat}(G_d) = P^0$. This verifies the construction of the digraph $G(d)$.

Finally, to extend $G(d)$ to $\tilde{G}(d)$ forcing G_d to the color pattern P^1 , add the arc (v_d^1, z^2) . Let us observe that the digraph $\tilde{G}(f)$ is acyclic. Note that if a digraph contains a sink (only incoming arcs) or a source (only outgoing arcs), then it can be removed from the digraph without affecting its acyclicity. Therefore, it is enough to show that the digraph $\tilde{G}(f)'$ obtained from $\tilde{G}(f)$ by removing the vertices of Z and the vertices w_d^1 and w_d^3 , for every subformula d used in the construction, is acyclic, which can be easily seen using a similar argument as in the proof of Lemma 4.6. $\square$

LEMMA 4.8. *Let $k \geq 1$, and let F_k be the quasi-acyclic semicomplete digraph obtained by forming the transitive closure of the digraph F'_k depicted in Figure 4.1. The acyclic F_k -coloring is NP-complete.*

Proof. The reduction from SAT is similar to the reduction described in the proof of Lemma 4.7. In fact, all of the gadgets and connections among them require just slight modifications. The digraph Z in the proof of Lemma 4.7 can be viewed as the transitive closure of the directed path (z^4, z^3, z^2, z^1) . In this proof the digraph Z will be the transitive closure of the path $(z^4, z^3, q^k, \dots, q^1, z^2, z^1)$. Again, Z has a unique F^k -coloring up to swapping A and B , or C and D . In any coloring, $\varphi(q^i) = E_i$ and, without loss of generality, let $\varphi(z^1) = A$, $\varphi(z^2) = B$, $\varphi(z^3) = C$, and $\varphi(z^4) = D$.

There are $k + 4$ colors available to color a variable gadget G_a depicted in Figure 4.7, which would result in many different color patterns for the gadget. In order to restrict the number of color patterns of G_a to two, we replace the arcs connecting the vertices of the variable gadget with vertices of the digraph Z with the following ten arcs: (q^1, v_a^1) , (v_a^2, z^2) , (q^2, v_a^2) , (v_a^3, q^{k-1}) , (z^3, v_a^3) , (v_a^4, z^1) , (q^2, v_a^4) , (v_a^5, q^{k-1}) , (z^4, v_a^5) , and (v_a^6, q^k) , and replace the arcs (v_a^3, v_a^2) and (v_a^5, v_a^4) by two directed paths with $k - 1$ internal nodes (with all arcs on each path having the same direction as the original arc); see Figure 4.10(a).

In the case $k = 1$, there are no vertices q^2 and q^{k-1} . Instead of the arcs (q^2, v_a^2) , (q^2, v_a^4) , (v_a^3, q^{k-1}) , and (v_a^5, q^{k-1}) , we add arcs (v_a^3, z^1) , (v_a^5, z^1) , (v_a^3, z^2) , (v_a^5, z^2) , (z^3, v_a^2) , (z^3, v_a^4) , (z^4, v_a^2) , and (z^4, v_a^4) . Note also that in this case the vertices q^1 and q^k are the same (the squares labeled E_1 and E_k in Figure 4.10(a)), and that the directed paths $(v_a^3, \dots, v_a^2)$ and $(v_a^5, \dots, v_a^4)$ contain only one arc each, as in the original variable gadget depicted in Figure 4.7.

It is easy to see that the arcs connecting vertices of Z to vertices of the gadget force exactly two color choices for each vertex v_a^i : A or B for v_a^1 ; A or E_1 for v_a^2 ; E_k or D for v_a^3 ; B or E_1 for v_a^4 ; E_k or C for v_a^5 ; and C or D for v_a^6 . Now, assume that $\varphi(v_a^1) = A$. Then $\varphi(v_a^2) = E_1$. Since the path $(v_a^3, \dots, v_a^2)$ contains $k - 1$ internal nodes, $\varphi(v_a^3) \neq E_k$; i.e., $\varphi(v_a^3) = D$. Then $\varphi(v_a^4) = C$ and $\varphi(v_a^5) = E_k$. Similarly, v_a^4 cannot be colored with E_1 ; i.e., $\varphi(v_a^4) = B$. Note also that $\varphi(v_a^4) = B$ implies that $\varphi(v_a^1) = A$. This defines the color pattern P^1 of the variable gadget G_a . The only other possible color pattern is the complementary one. One can easily check that this complementary color pattern P^0 is valid as well; see Figure 4.10(b).

In the inductive construction of $\tilde{G}_d$, there are no differences in the base step in the case $d = \overline{x_j}$ (see the right-hand side picture in Figure 4.8). In the case $d = x_j$,

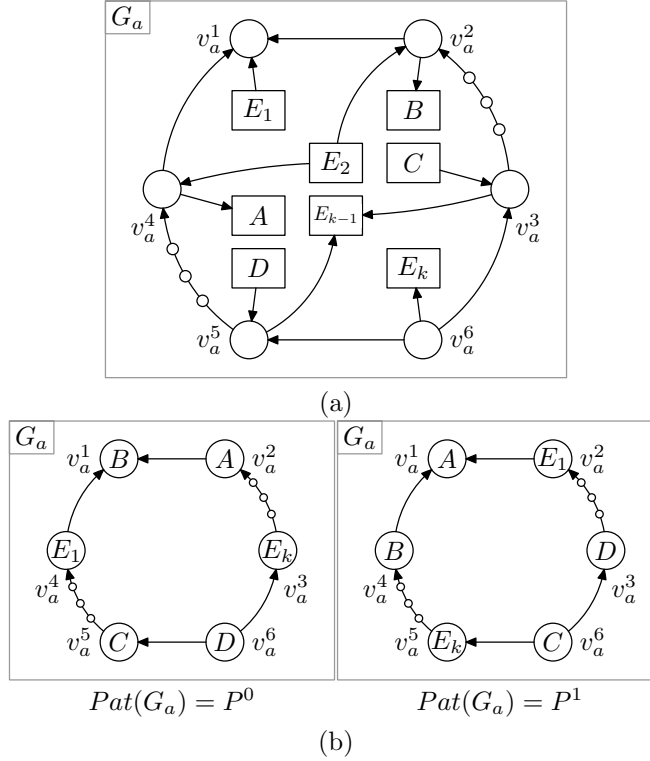


FIG. 4.10. (a) A variable gadget used in the proof that F_k -coloring is NP-complete for $k > 1$. The squares labeled with $A, B, C, D, E_1, E_2, E_{k-1}$, and E_k represent vertices $z^1, z^2, z^3, z^4, q^1, q^2, q^{k-1}$, and q^k of the common digraph Z , respectively. The arrows with circles connecting v_a^3 with v_a^2 and v_a^5 with v_a^4 depict directed paths $(v_a^3, \dots, v_a^2)$ and $(v_a^5, \dots, v_a^4)$ with $k-1$ internal nodes.

(b) Two color patterns of the variable gadget: $\text{Pat}(G_a) = P^0$ (on the left) and $\text{Pat}(G_a) = P^1$ (on the right).

we add an arc $(v_{x_j}^6, v_d^3)$; see Figure 4.11. It is easy to see that both gadgets must be in the same state.

In the inductive step $d = d' \vee d''$, we construct $G(d)$ as follows. Take the union of $G_{d'}$ and $G_{d''}$ and add one new gadget G_d and one new vertex w_d . Furthermore, add the arcs $(v_{d'}^6, v_d^5), (v_{d''}^6, v_d^5), (w_d, v_d^1), (w_d, v_{d'}^1)$ and a directed path from $v_{d''}^3$ to w_d with $k-1$ internal nodes; see Figure 4.11. If $\text{Pat}(G_{d'}) = P^1$ (respectively, $\text{Pat}(G_{d''}) = P^1$), then the arc $(v_{d'}^6, v_d^5)$ (respectively, $(v_{d''}^6, v_d^5)$) forces $\text{Pat}(G_d) = P^1$. On the other hand, if $\text{Pat}(G_{d'}) = P^0$, then $\varphi(w_d)$ cannot be B ; and if $\text{Pat}(G_{d''}) = P^0$, then $\varphi(v_{d''}^3) = E_k$. Hence, $\varphi(w_d)$ is either A or B . Thus, if both $\text{Pat}(G_{d'}) = \text{Pat}(G_{d''}) = P^0$, then $\varphi(w_d) = A$. Hence, $\varphi(v_d^1) = B$; i.e., $\text{Pat}(G_d) = P^0$. Note that only $\varphi(w_d) = A$ forces $\text{Pat}(G_d) = P^0$ and that w_d is not forced to have color A if either $\text{Pat}(G_{d'})$ or $\text{Pat}(G_{d''})$ is P^1 . Thus, we can see that $\text{Pat}(G_d) = P^1$ if and only if $\text{Pat}(G_{d'}) = P^1$ or $\text{Pat}(G_{d''}) = P^1$. The rest of the proof is analogous to the proof of Lemma 4.7. $\square$

The following lemma gives a simple characterization when a digraph contains at least two directed 2-cycles.

LEMMA 4.9. *A quasi-acyclic semicomplete digraph H contains at least two directed 2-cycles if and only if it contains either H_0, H_1 , or F_0 (see Figure 4.1) as an induced subdigraph.*

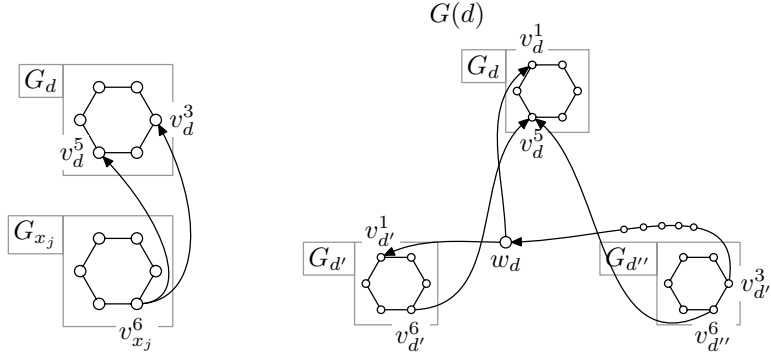


FIG. 4.11. On the left: $G(d)$ for the Boolean formula $d = x_j$. On the right: $G(d)$ for the Boolean formula $d = d' \vee d''$. The path connecting $v_{d'}^3$ with v_d^5 contains $k - 1$ internal nodes.

Proof. Sufficiency is trivial. For necessity, suppose that H does not contain H_0 . Now, if there are two 2-cycles in H which share a vertex, then the three vertices on these two 2-cycles induce H_1 . Otherwise, take any two 2-cycles. Since they are vertex-disjoint, the four vertices on them induce F_0 . $\square$

Therefore, to prove the NP-completeness part of Theorem 3.1, it is enough to show the following lemma.

LEMMA 4.10. Suppose H is a quasi-acyclic semicomplete digraph. If H contains either H_0 , H_1 , or F_0 (see Figure 4.1) as an induced subdigraph, then the acyclic H -coloring problem is NP-complete.

Proof. We will distinguish three cases.

Case 1. We first prove that when H contains H_0 as an induced subdigraph, the acyclic H -coloring problem is NP-complete by reduction from the proper 3-coloring problem. In particular, given a graph G , we will construct an acyclic digraph K which has an H -coloring if and only if G is 3-colorable.

Let A , B , and C be the vertices of H_0 . Find a directed Hamilton path P in H on which vertices A , B , and C are consecutive (by using a topological sort on the poset corresponding to H). Let P_ℓ (respectively, P_u) denote the subpath of P containing all vertices preceding (respectively, following) the three vertices A , B , and C on P . Note that P has the property that its transitive closure is H -colorable. This will be used later in the proof.

We construct an acyclic digraph K as follows. Start with subpaths P_ℓ and P_u and vertices in $V(G)$. Add an arc from the last vertex of P_ℓ to every vertex in $V(G)$ and from every vertex in $V(G)$ to the first vertex of P_u ; see Figure 4.12.

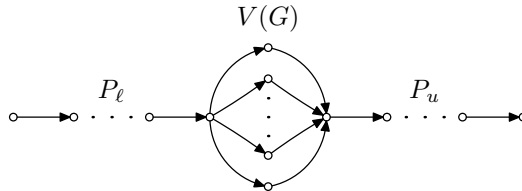


FIG. 4.12. First step in the construction of the acyclic digraph $\vec{K}$.

Next, take the transitive closure of this digraph. Observe that the resulting

digraph has many H -colorings in which vertices on P_ℓ and P_u are mapped to corresponding vertices in H , and vertices in $V(G)$ are arbitrarily mapped to vertices A , B , and C . Consider any acyclic orientation of G and add every arc of this orientation to the constructed digraph joining corresponding vertices in $V(G)$. The resulting acyclic digraph is K .

First, suppose that G is 3-colorable with colors A , B , and C . For an H -coloring of K choose the H -coloring from the previous paragraph that agrees on vertices in $V(G)$ with the 3-coloring.

Second, suppose that K has an H -coloring. Since $V_P = V(P_\ell) \cup V(P_u)$ induces a transitive tournament in K , any such H -coloring must use all but three vertices of H . Since every vertex of $V(G)$ is incident (is a tail or a head of an arc) with every vertex in V_P , in any H -coloring of K the vertices in $V(G)$ are mapped to the remaining three vertices of H ; i.e., G is 3-colorable.

Case 2. Next, we prove that when H contains H_1 but not H_0 as an induced subdigraph, the H -coloring problem is NP-complete by reduction from the acyclic H_1 -coloring problem (see Lemma 4.6). Let G be an acyclic digraph for which we want to decide whether it is H_1 -colorable. We will construct an acyclic digraph K which is H -colorable if and only if G is H_1 -colorable. The construction of K is exactly as in the previous case with the only difference that G already fixes an acyclic orientation used in the construction. However, it is not obvious that there exists a directed Hamilton path in H with three consecutive vertices B , C , and A inducing H_1 .

Consider the poset $(V(H), \preceq)$ corresponding to H . We say that a triple $[B, A, C]$ of different vertices in H is *nice* if $B \preceq A$ and the pairs A, C and B, C are both incomparable. Obviously, there is a nice triple since H_1 is an induced subdigraph of H . Take a directed Hamilton path P in H on which there is a nice triple $[B, A, C]$ such that the distance between B and A along P is the smallest possible. We will show that the distance between B and A along P is one. Suppose by contradiction that $E \neq B$ is the immediate predecessor of A on P . Note that for any two X, Y such that X precedes Y on P , H contains the arc (X, Y) . (This follows since H is semicomplete.) Hence, if H contains both arcs (A, E) and (E, A) (a double arc (A, E)), then by exchanging A and E we obtain another directed Hamilton path in H with smaller distance of A and B , a contradiction. Thus, we may assume that H contains the arc (E, A) but not (A, E) , i.e., $E \prec A$. Moreover, either $E \prec C$ or they are incomparable, since otherwise $C \prec A$.

Suppose that $B \prec E$. It follows that C and E are incomparable; hence the triple $[B, E, C]$ is nice and the distance between B and E is smaller along P . Hence, B and E must be incomparable. Let F be an immediate successor of B on P . By a similar argument as above, it follows that $B \prec F$, and A and F are incomparable. Therefore, $E \neq F$ and they are incomparable. Now, $[B, F, E]$ is a nice triple with distance one between B and F on P , a contradiction.

We say that a nice triple $[B, A, C]$ on a directed Hamilton path is *very nice* if the distance between B and A along P is one. By the above argument, there is a directed Hamilton path with a very nice triple. Take such a path P and a very nice triple $[B, A, C]$ such that the sum of the distances between B and C and between A and C on P is the smallest possible. We will show that this sum of distances is three. Without loss of generality, suppose that C follows A on P . Suppose by contradiction that $E \neq C$ is the immediate successor of A on P . If E is incomparable with both A and B , then $[B, A, E]$ is a very nice triple with a smaller sum of distances. It follows that C and E must be incomparable (otherwise, by transitivity, at least one of A, C

or B, C would be comparable). If A and E are incomparable as well, then A, C, E induce H_0 in H , a contradiction. Hence, $[A, E, C]$ is a very nice triple with a smaller sum of distances on P , a contradiction.

Obviously, if G is H_1 -colorable, then we can use this coloring to construct an H -coloring of K as above. Conversely, suppose now that K has an H -coloring. As before, it follows that $V(G)$ is mapped to some three vertices of H . Since H_0 is not an induced subdigraph of H , the three vertices induce a subdigraph of H_1 . Therefore, G is H_1 -colorable.

Case 3. Finally, we prove that when H contains F_0 but neither H_0 nor H_1 as an induced subdigraph, the H -coloring problem is NP-complete by reduction from one of the following problems: the acyclic F_0 -coloring problem (Lemma 4.7) or the acyclic F_k -coloring problem for some $k \geq 1$ (Lemma 4.8).

Take a directed Hamilton path P in H . Since H_0 and H_1 are not induced subdigraphs of H , it is easy to see that we can make sure that any two incomparable vertices in the corresponding poset $(V(H), \preceq)$ are consecutive on P . Since F_0 is an induced subdigraph of H there are at least two incomparable pairs. Take two such pairs D, C and B, A which are in this order on P and are closest to each other. Similarly, it is easy to see that for any vertex E in between C and B on P , $D \prec E$, $C \prec E$, $E \prec B$, and $E \prec A$. Therefore, vertices of the subpath P_m of P from D to A induce F_k , where k is the number of vertices between C and D .

We will show that the H -coloring problem is NP-complete by reduction from the acyclic F_k -coloring problem. Let P_ℓ (respectively, P_u) denote the subpath of P containing all vertices preceding D (respectively, following A) on P . Given G , construct the acyclic digraph K similarly as in the previous cases. Obviously, if G is F_k -colorable, then we can use this coloring to construct an H -coloring of K as in the previous cases. Conversely, suppose now that K has an H -coloring. We will show that vertices of $V(G)$ in K are colored with colors on P_m in H . By contradiction, suppose $x \in V(G)$ is colored by Y not on P_m . Without loss of generality, let $Y \in P_\ell$. Since for every vertex z on P_ℓ there is an arc (z, x) in K , the color of z must be either incomparable with Y or a predecessor of Y on P . Since there is at most one incomparable vertex with Y and it must lie on P_ℓ of H , we conclude that the color of z is on P_ℓ in H and different from Y . Since each vertex on P_ℓ of K must have a different color, there are not enough colors for them, a contradiction. Therefore, G is F_k -colorable. $\square$

Acknowledgment. We would like to thank the anonymous referees for their careful reading and extremely helpful comments.

REFERENCES

- [1] J. BANG-JANSEN, P. HELL, AND G. MACGILLIVRAY, *The complexity of colouring by semicomplete digraphs*, SIAM J. Discrete Math., 1 (1988), pp. 281–298.
- [2] R. DIESTEL, *Graph Theory*, 3rd ed., Springer-Verlag, Berlin, 2005.
- [3] S. EVEN, A. ITAI, AND A. SHAMIR, *On the complexity of timetable and multicommodity flow problems*, SIAM J. Comput., 5 (1976), pp. 691–703.
- [4] M. GAREY AND D. JOHNSON, *Computers and Intractability. A Guide to the Theory of NP-Completeness*, W. H. Freeman and Co., San Francisco, 1979.

HEREDITARY DOMINATION IN GRAPHS: CHARACTERIZATION WITH FORBIDDEN INDUCED SUBGRAPHS*

ZSOLT TUZA†

Abstract. The leaf graph of a connected graph is obtained by joining a new vertex of degree one to each noncutting vertex. We prove that if a connected graph G is not dominated by any of its induced paths, then G is dominated by a connected induced subgraph whose leaf graph, too, is an induced subgraph of G . It follows that, for every nonempty class $\mathcal{D}$ of connected graphs, all of the minimal graphs not dominated by any induced subgraph isomorphic to some $D \in \mathcal{D}$ are cycles (of well-determined lengths) and leaf graphs of some graphs $H \notin \mathcal{D}$. In particular, if $\mathcal{D}$ is closed under the operation of taking connected induced subgraphs, then the hereditarily $\mathcal{D}$ -dominated graphs are characterized by the following family of forbidden induced subgraphs: leaf graphs of the connected graphs that are not in $\mathcal{D}$, but all of their connected induced subgraphs are in $\mathcal{D}$, and the cycle C_{t+2} , where t is the length of the shortest path not in $\mathcal{D}$ (if $\mathcal{D}$ does not contain all paths). This solves a problem that was open since the 1980s. A solution for the case of induced-hereditary classes $\mathcal{D}$ has been found simultaneously by Bacsó by applying a different method.

Key words. dominating set, forbidden induced subgraph, induced-hereditary property, leaf graph, structural domination

AMS subject classifications. 05C69, 05C75

DOI. 10.1137/070699482

1. Introduction. Let $G = (V, E)$ be a *finite* simple graph. A vertex subset $X \subseteq V$ is called a *dominating set* of G if every $v \in V \setminus X$ is adjacent to some $x \in X$. A *dominating (induced) subgraph* is a(n) (induced) subgraph whose vertex set is dominating in G . We shall use the notation $G' \sqsubseteq G$ to express that G' is an *induced* subgraph of G (allowing $G' = G$). Further definitions and notation will be introduced in section 1.1.

The goal of this short note is to give a complete solution to the long-standing problem of characterizing the *structure* of dominating subgraphs in terms of forbidden induced subgraphs. More precisely, a necessary and sufficient condition will be described for the graphs in which every connected induced subgraph admits a dominating induced subgraph of a given type.

The first paper relating the structure of dominating subgraphs with that of the graphs themselves appears to be [7], where it was proved that every connected graph without induced $2K_2$ (two disjoint edges) has a dominating set that induces a complete subgraph or a P_3 (path of length two). Inspired by this result, a systematic study of this phenomenon with respect to hereditary domination was started in [2]. From this perspective, the characterization of “trivially perfect” graphs in the early paper [9] from 1962 can be viewed as the very first member of a long sequence of theorems of this kind, even if it deals just with a single dominating vertex.

Although several powerful related methods have been developed in earlier papers (the star-cutset lemma [2], the induced path lemma [3], the cut-point lemma, and the sequential construction of nondominating subgraphs [5]), the current proof is

*Received by the editors August 7, 2007; accepted for publication (in revised form) January 14, 2008; published electronically May 7, 2008.

<http://www.siam.org/journals/sidma/22-3/69948.html>

†Computer and Automation Institute of the Hungarian Academy of Sciences, Budapest, Hungary, and Department of Computer Science, University of Pannonia, Veszprém, Hungary (tuza@sztaki.hu). This work was supported by Hungarian Scientific Research Fund grant OTKA T-049613.

self-contained and unexpectedly short. The heart of the present method is the leaf subgraph theorem. This structural result on connected graphs not dominated by any induced paths is proved in section 2. The complete solution for the graph-class equation $\text{Dom}(\mathcal{D}) = \text{Forb}(\mathcal{F})$ is derived from it in section 3. In the same section, a lemma of independent interest is presented, too, namely, that the operation of creating so-called leaf graphs preserves the “induced subgraph” relation between connected graphs. In the last section we illustrate the method with two applications.

Bacsó [1] has found a different proof of the main theorem simultaneously and also obtained for nonhereditary $\mathcal{D}$ that all 2-connected minimal forbidden subgraphs are cycles. We note here, without entering into details, that, on the other hand, the “king theorem” of [1] can be derived shortly from our leaf subgraph theorem. We refer to the longer paper [1] for a survey on the literature of structural domination.

It should be noted that the situation for *infinite* graphs is different. Some results in that direction have been obtained jointly by Bacsó, Jung, and the author.

1.1. Leaf graphs and the main theorem. Let $\mathcal{D}$ and $\mathcal{F}$ be given (finite or infinite) classes of graphs; it will be assumed throughout that $\mathcal{D} \neq \emptyset$ and all $D \in \mathcal{D}$ are *connected*. We denote by $\text{Forb}(\mathcal{F})$ the class of graphs containing no induced subgraphs isomorphic to any $F \in \mathcal{F}$.

A graph G is said to be *hereditarily $\mathcal{D}$ -dominated* if every *connected* $G' \subseteq G$ has a dominating $D \subseteq G'$, with $D \in \mathcal{D}$. The class of all such G will be denoted by $\text{Dom}(\mathcal{D})$.

Given a graph H with at least two vertices, the *leaf graph* of H —denoted by $F(H)$ —is obtained by attaching a leaf (i.e., joining a “pendant” vertex of degree one) to each of its *noncutting* vertices. This notion was introduced in [5]. In the next section we shall prove the following notion.

LEAF SUBGRAPH THEOREM. *If a connected graph G is not dominated by any induced paths, then it has a connected dominating induced subgraph H that can be extended to its leaf graph $F(H)$ as an induced subgraph of G .*

From this result, the complete solution of the graph-class equation $\text{Dom}(\mathcal{D}) = \text{Forb}(\mathcal{F})$ will be derived in section 3. We say that a graph H is *minimal not in $\mathcal{D}$* if H is connected, $H \notin \mathcal{D}$, and every *connected* proper induced subgraph of H is in $\mathcal{D}$. As usual, let P_n and C_n denote the path and cycle, respectively, on n vertices. The characterization for *hereditary* classes $\mathcal{D}$ is as follows.

MAIN THEOREM. *If $\mathcal{D} \neq \emptyset$ is a class of connected graphs, closed under connected induced subgraphs, then the minimal forbidden induced subgraphs for $\text{Dom}(\mathcal{D})$ are the following:*

- (i) *the leaf graphs $F(H)$ for the graphs H minimal not in $\mathcal{D}$ and*
- (ii) *the cycle C_{t+2} if $P_t \notin \mathcal{D}$ but $P_{t-1} \in \mathcal{D}$.*

In particular, if $\mathcal{D}$ contains all paths, then (ii) is void and there do not exist any 2-connected minimal forbidden induced subgraphs. On the other hand, if $\mathcal{D}$ does not contain all paths, then the value of t above is uniquely determined, since $\mathcal{D}$ is assumed to be a hereditary class.

One direction of the *main theorem*—namely, that the graphs listed in (i) and (ii) are minimal forbidden ones—is easy to see, and therefore the message of the result is that no others exist. For the non-2-connected ones, (i) was already proved in the cut-point lemma of [5]; but the present proof for both (i) and (ii) is entirely self-contained.

In the proofs, the following standard notion will be applied. Given a subgraph H of a graph G and a vertex v of H , a *private neighbor* of v is a vertex $x \in V(G) \setminus V(H)$ whose *unique* neighbor in H is v .

2. Proof of the leaf subgraph theorem. Let H be a connected graph, and suppose that it has $k + 1$ noncutting vertices $v_0, v_1, \dots, v_k$; $k = k(H)$. The case of $k(H) = 1$ characterizes the paths with at least one edge (and $k = 0$ would mean a single vertex, but we do not need to consider it). If H is not a path—that is, if $k(H) \geq 2$ —we assign a value $m(H)$ to H as follows.

First, introduce a positive integer ℓ_i for $0 \leq i \leq k$ as follows. If v_i is not a leaf of H , let $\ell_i = 1$; and if v_i has degree one, let $\ell_i - 1$ be equal to the distance of v_i from the nearest vertex of degree greater than two (i.e., $\ell_i > 1$ in the latter case). This ℓ_i is well-defined because H is supposed to be connected and not a path. We assume that the noncutting vertices are indexed in an order satisfying $\ell_0 \geq \ell_1 \geq \dots \geq \ell_k$. Let

$$m(H) = \sum_{i=0}^k \ell_i \cdot 2^i.$$

We prove next that the following property is valid:

- (1) For each $0 \leq i \leq k$, either $k(H - v_i) < k(H)$ (where $H - v_i$ possibly is a path, for which no m -value is assigned) or $k(H - v_i) = k(H)$ and $m(H - v_i) \leq m(H) - 2^i$.

Indeed, if a cut-vertex x of H becomes noncutting in $H - v_i$, then v_i is an isolated vertex in $H - x$, and thus the ℓ -value for x in $H - v_i$ is precisely $\ell_i - 1$; there can be at most one such x for each v_i , implying that $k(H - v_i) \leq k(H)$. If $\ell_{i+1} < \ell_i$ in H , then x keeps the position of v_i in the noncutting vertex order and yields $m(H - v_i) = m(H) - 2^i$. But otherwise, to make the ℓ -sequence of $H - v_i$ nonincreasing, it is necessary to place x somewhere after ℓ_{i+1} , and then $m(H)$ decreases with a higher power of two.

Now, among all connected dominating induced subgraphs of G , we choose H^* with the smallest k -value and, under this condition, with the smallest m -value, too. We are going to prove that suitable vertices x_i ($i = 0, 1, \dots, k$) can be sequentially selected in G , such that each x_i is a private neighbor of v_i and those x_i are mutually nonadjacent; i.e., the subgraph induced by $V(H^*) \cup \{x_i \mid 0 \leq i \leq k\}$ is isomorphic to $F(H^*)$.

Let i be arbitrary in the range $0 \leq i \leq k$, and suppose that x_j has been selected for all $0 \leq j < i$. In order to find x_i , consider the subgraph H_i of G induced by

$$(V(H^*) \cup \{x_j \mid 0 \leq j < i\}) \setminus \{v_i\}.$$

This H_i is connected (since v_i is noncutting in H^*) and $k(H_i) = k(H^* - v_i)$; moreover, the ℓ -value of each x_j ($j < i$) in H_i is precisely $\ell_j + 1$, where ℓ_j is meant for v_j in H^* . Hence, $m(H_i \cup \{v_i\}) = m(H^*) + 2^i - 1$ holds, because $x_0, \dots, x_{i-1}, v_i, \dots, v_k$ keep the nonincreasing order of ℓ -values. Thus, by (1) for $H = H^*$, either $k(H_i) < k(H^*)$ or $k(H_i) = k(H^*)$ and $m(H_i) < m(H^*)$ is obtained. We conclude, by the choice of H^* , that this H_i cannot be dominating in G , so that the next private neighbor x_i for v_i can be selected.

3. Solution of $\text{Dom}(\mathcal{D}) = \text{Forb}(\mathcal{F})$. Before proving the *main theorem*, we present a useful auxiliary result.

INDUCED LEAF SUBGRAPH LEMMA. *If $H' \subseteq H$ and H' is connected, then $F(H') \subseteq F(H)$. Moreover, if $F(H) \subseteq G$ and $H' \subseteq H$, then also $F(H') \subseteq G$.*

Proof. Let $v_1, \dots, v_s$ be the noncutting vertices of H' . If v_i (for any $1 \leq i \leq s$) is a noncutting vertex of H as well, then it has a private neighbor x_i in $F(H) - H$; and if v_i is a cut-vertex of H , then it has a neighbor x_i in some component of $H - v_i$ that

contains no vertices of H' . This proves that $F(H') \subseteq F(H)$. From here, the second assertion follows by the transitivity of " $\subseteq$." $\square$

Proof of the main theorem. Let a connected graph G and a class $\mathcal{D} \neq \emptyset$ of connected graphs be given, with $\mathcal{D}$ closed under connected induced subgraphs. It suffices to prove that G contains either a dominating induced subgraph in $\mathcal{D}$ or an induced subgraph of type (i) or (ii). Choose either a *shortest* dominating path $P_t \subseteq G$, if there exists one, or a dominating $H \subseteq G$ with the properties described in the *leaf subgraph theorem*. The former, P_t , either is in $\mathcal{D}$ —that surely holds in the particular case of $t = 1$ —or yields $P_{t+2} \subseteq G$ or $C_{t+2} \subseteq G$, since both endpoints of the dominating P_t have private neighbors in G (by the minimality of t). The latter, H , either is in $\mathcal{D}$ or contains a connected $H' \subseteq H$ minimal not in $\mathcal{D}$. On applying the *induced leaf subgraph lemma*, we then obtain $F(H') \subseteq G$. Concerning C_{t+2} , the side condition (ii) is needed because if $P_t, P_{t-1} \notin \mathcal{D}$, then $P_{t+1} \notin \mathcal{D}$ by the assumptions on $\mathcal{D}$, and so P_{t+1} has no dominating subgraph in $\mathcal{D}$; hence, C_{t+2} is not *minimal* non- $\mathcal{D}$ -dominated. $\square$

3.1. Unrestricted $\mathcal{D}$. If $K_1 \notin \mathcal{D}$, then no graph (having at least one vertex) can be *hereditarily* $\mathcal{D}$ -dominated. To avoid this trivial negative case, let $K_1 \in \mathcal{D}$. It is still assumed that all $D \in \mathcal{D}$ are *connected*. Then, the following facts are easy to see:

- A cycle C_{t+2} is not $\mathcal{D}$ -dominated if and only if $P_t, P_{t+1}, C_{t+2} \notin \mathcal{D}$.
- The leaf graph $F(H)$ of a connected graph H is not $\mathcal{D}$ -dominated if and only if there doesn't exist any $D \in \mathcal{D}$ such that $H \subseteq D \subseteq F(H)$.

The second case includes the non- $\mathcal{D}$ -dominated *paths*, too, since $P_{t+2} = F(P_t)$ for all $t \geq 2$.

If a graph G is not (hereditarily) $\mathcal{D}$ -dominated and $G \subseteq G'$, then G' is not hereditarily $\mathcal{D}$ -dominated either. Thus, in characterizing $\text{Dom}(\mathcal{D}) = \text{Forb}(\mathcal{F})$, also for arbitrary $\mathcal{D}$, the class $\mathcal{F}$ is determined by its members *minimal under induced inclusion*. Hence, along the lines of the proof of the *main theorem*, the following solution is obtained.

CHARACTERIZATION THEOREM. *Let $\mathcal{D}$ be a family of graphs, where all $D \in \mathcal{D}$ are connected. Then the minimal forbidden subgraphs for $\text{Dom}(\mathcal{D})$ are those of the cycles and leaf graphs described above that are minimal under the partial order " $\subseteq$."*

To obtain $\mathcal{F}$ for a given $\mathcal{D}$, the following procedure can be designed:

1. Construct $\mathcal{F}^*$ by considering all connected $H \notin \mathcal{D}$ and selecting
 - (a) $L(H)$, if there exist no $D \in \mathcal{D}$ such that $H \subseteq D \subseteq L(H)$, and also
 - (b) C_{t+2} , if H is the path P_t , and $P_t, P_{t+1}, C_{t+2} \notin \mathcal{D}$.
2. Determine $\mathcal{F} = \{\text{minimal members of } \mathcal{F}^* \text{ under } \subseteq\}$.

It is important to note that, in the case of nonhereditary $\mathcal{D}$, first the leaf graphs of the connected graphs $H \notin \mathcal{D}$ and the non- $\mathcal{D}$ -dominated cycles should be selected, and only after that should the minimal ones among them be kept; the opposite order (selecting the minimal graphs $H \notin \mathcal{D}$ and then taking their leaf graphs and cycles) may yield a different—and hence incorrect—output.

4. Applications. Here we present two applications of the method. The first one is an alternative derivation of a result proved in [4] on dominating subgraphs of bounded diameter (what is a nonhereditary property). The second one illustrates with a theorem of [8] on triangle-free graphs how domination in a restricted graph class can be handled.

The proofs apply the steps and notation of the procedure described after the *characterization theorem* of the previous section without further reference to it.

Dominating with bounded diameter. Let $k \geq 0$ be an integer and $\mathcal{D} = \mathcal{D}_k$ the family of graphs of diameter at most k . This $\mathcal{D}$ is not an induced-hereditary class, because, e.g., every wheel graph belongs to it for all k , but removing the central vertex from a wheel yields a cycle, which has a dominating subgraph of diameter at most k only if its length does not exceed $2k + 1$.

The graphs $H \notin \mathcal{D}$ are precisely those containing at least two vertices at distance $k + 1$ apart, which necessarily yields P_{k+2} as an induced subgraph. Thus, $F \in \mathcal{F}^*$ holds if and only if $F = C_t$ for some $t \geq k + 4$ or F is a leaf graph and $P_{k+4} \subseteq F$. Since $P_{k+4} \subseteq C_t$ for all $t > k + 4$, the only candidates to be minimal members of $\mathcal{F}^*$ are P_{k+4} and C_{k+4} . Taking into account that $\text{diam}(C_{k+4}) \leq k$ if and only if $k \geq 3$, we obtain the following:

$$\mathcal{F} = \{P_{k+4}, C_{k+4}\} \text{ if } k \leq 2 \text{ and } \mathcal{F} = \{P_{k+4}\} \text{ if } k \geq 3.$$

Dominating triangle-free graphs with complete bipartite subgraphs.

Let $\mathcal{D}$ be the family of complete bipartite graphs, excluding edgeless graphs with more than one vertex but including K_1 . The following theorem was proved in [8]:

*A triangle-free graph G is hereditarily $\mathcal{D}$ -dominated if and only if
 $P_6 \not\subseteq G$ and $C_6 \not\subseteq G$.*

To derive this from our results, we first observe that P_4 is a graph minimal not in $\mathcal{D}$. Hence, $P_6, C_6 \in \mathcal{F}^*$, and, moreover, the *induced leaf subgraph lemma* yields that $P_6 \subseteq L(H)$ for all H with $P_4 \subseteq H$. On the other hand, if a connected graph H' has more than one vertex and $P_4 \not\subseteq H'$, then $V(H')$ can be partitioned into two nonempty subsets A and B that are completely joined [6]. Since our attention is restricted now to triangle-free graphs, both A and B must be independent sets, so that $H' \in \mathcal{D}$ holds and such an H' cannot generate any further members of $\mathcal{F}^*$. Consequently, $\mathcal{F} = \{P_6, C_6\}$.

REFERENCES

- [1] G. BACSÓ, *Complete description of forbidden subgraphs in the structural domination problem*, Discrete Math., to appear.
- [2] G. BACSÓ AND ZS. TUZA, *Dominating cliques in P_5 -free graphs*, Period. Math. Hungar., 21 (1990), pp. 303–308.
- [3] G. BACSÓ AND ZS. TUZA, *A characterization of graphs without long induced paths*, J. Graph Theory, 14 (1990), pp. 455–464.
- [4] G. BACSÓ AND ZS. TUZA, *Dominating subgraphs of small diameter*, J. Combin. Inform. Syst. Sci., 22 (1997), pp. 51–62.
- [5] G. BACSÓ AND ZS. TUZA, *Structural domination in graphs*, Ars Combin., 63 (2002), pp. 235–256.
- [6] D. G. CORNEIL, H. LERCHS, AND L. S. BURLINGTON, *Complement reducible graphs*, Discrete Appl. Math., 3 (1981), pp. 163–174.
- [7] M. EL-ZAHAR AND P. ERDŐS, *On the existence of two non-neighboring subgraphs*, Combinatorica, 5 (1985), pp. 295–300.
- [8] J. LIU AND H. ZHOU, *Dominating subgraphs in graphs with some forbidden structures*, Discrete Math., 135 (1994), pp. 163–168.
- [9] E. S. WOLK, *The comparability graph of a tree*, Proc. Amer. Math. Soc., 3 (1962), pp. 789–795.

MATROID PARITY AND JUMP SYSTEMS: A SOLUTION TO A CONJECTURE OF RECSKI*

JÁCINT SZABÓ†

Abstract. In 1981 András Recski conjectured that, if given a number $q \in \mathbb{N}$, a linearly represented matroid M , a partition $S_1 \dot{\cup} \dots \dot{\cup} S_n$ of a subset of its ground set S into classes of size k , and a prescription $A \subseteq \{0, 1, \dots, k\}$ without two consecutive gaps, then one can find in polynomial time an independent set F of M of size q such that $|F \cap S_i| \in A$ for all $1 \leq i \leq n$, if one exists. In this paper we prove this conjecture. The proof is based on Lovász' result on the polynomial solvability of the matroid parity problem for linearly represented matroids and on an important technique about jump systems, proved by Sebő. We give an application to rigidity theory and another one to the unique solvability of linear networks containing memoryless multiports.

Key words. matroid parity, jump systems, rigidity

AMS subject classifications. 05B35, 05B99

DOI. 10.1137/07068494X

1. Introduction. Matroid parity is a problem of high importance in combinatorial optimization. It is a common generalization of the matching problem of graphs and the matroid intersection problem. If M is a matroid on ground set S , and some (not necessarily disjoint) pairs are designated in S , then we call a set of k pairs a *matching* in M , if their union has rank $2k$ (that is, maximum possible). Given a matroid together with some designated pairs, the matroid parity problem is to find a maximum matching in it. Jensen and Korte [5] and Lovász [8] have shown that, in general, the matroid parity problem is of exponential complexity under the independence oracle framework. However, as a groundbreaking result, Lovász [7] gave a polynomial algorithm for finding a maximum matching in a matroid which is linearly represented over $\mathbb{Q}$. The diversity and the importance of solvable special cases of the matroid parity problem make it a vivid and intriguing field of research. For more references and background, see [10, 16].

DEFINITION 1.1. We denote by $\mathbb{N}$ the set of nonnegative integers. If $H \subseteq \mathbb{N}$, then $i \in \mathbb{N} - H$ is called a *gap* of H if $\min H < i < \max H$.

In [14] Recski studied the following slight generalization of the matroid parity problem. For $k \in \mathbb{N}$ and $A \subseteq \{0, \dots, k\}$ we define the problem $P(A)$ as follows. Given a number $q \in \mathbb{N}$, a matroid $M = (S, r)$, and a partition $\pi = S_1 \dot{\cup} \dots \dot{\cup} S_n$ of a subset of S into classes of size k , decide whether M has an independent set F of size q such that $|F \cap S_i| \in A$ for all $1 \leq i \leq n$. This problem was motivated by engineering applications; see [12, 13]. It is easy to see that, if π is a partition into pairs and $A = \{0, 2\}$, we get back the matroid parity problem. Recski [14] proved that, under the independence oracle model, $P(A)$ is polynomial if A is an interval and is of exponential complexity otherwise. He also proved that, if we are given a representation of M over $\mathbb{Q}$, then $P(A)$ is NP-complete if A has two consecutive gaps, and for the prescriptions A without two consecutive gaps he conjectured $P(A)$ to be

*Received by the editors March 11, 2007; accepted for publication (in revised form) January 16, 2008; published electronically May 7, 2008.

<http://www.siam.org/journals/sidma/22-3/68494.html>

†MTA-ELTE Egerváry Research Group (EGRES), Institute of Mathematics, Eötvös University, Budapest, Pázmány P. s. 1/C, H-1117, Hungary (jacint@elte.hu). Research is supported by OTKA grants K60802 and TS049788 and by European MCRTN Adonet contract grant 504438.

polynomial [14]. In this paper we verify a slight extension of this conjecture, Theorem 1.3 below.

DEFINITION 1.2. Let S be a finite set, $S = \dot{\bigcup}_{t \in T} S_t$ be a partition, and $H(t) \subseteq \mathbb{N}$ be prescriptions for all $t \in T$. We say that a set $F \subseteq S$ is H -conformal if $|F \cap S_t| \in H(t)$ for all $t \in T$.

THEOREM 1.3. The following problem is solvable in polynomial time. Given $q \in \mathbb{N}$, a matroid $M = (S, r)$ represented over $\mathbb{Q}$, a partition $S = \dot{\bigcup}_{t \in T} S_t$, and for all $t \in T$ a prescription $H(t) \subseteq \mathbb{N}$ without two consecutive gaps, find an H -conformal independent set F of M of size q , if one exists.

Just like in the linear matroid parity algorithm of Lovász [7], the field $\mathbb{Q}$ can be replaced by any fields where algebraic operations can be performed in constant time.

We show two applications. First we prove a result on a three-dimensional pinning problem of bar-and-joint frameworks. Then we show that Theorem 1.3 yields a polynomially verifiable sufficient condition to the unique solvability of linear networks with memoryless multiports, a question raised by Recski and Takács [13].

2. Proof of the conjecture.

DEFINITION 2.1. For $F \subseteq T$ we denote by $\chi_F \in \{0, 1\}^T$ the characteristic vector of F . For $x, y \in \mathbb{Z}^T$ we say that x' is a step from x to y , if either $x' = x + \chi_t$ and $x(t) < y(t)$ or $x' = x - \chi_t$ and $x(t) > y(t)$ for some $t \in T$. We call $J \subseteq \mathbb{Z}^T$ a jump system if, for all $x, y \in J$ and step x' from x to y , either $x' \in J$ or some step from x' to y is contained in J .

Let $J \subseteq \mathbb{Z}^S$ be a jump system and $\varphi : S \rightarrow T$ a function. For $x \in \mathbb{Z}^S$ let $\varphi x \in \mathbb{Z}^T$ be the vector, with components $\varphi x(t) = \sum \{x(s) : \varphi(s) = t\}$ for $t \in T$. Finally, let $\varphi(J) = \{\varphi x : x \in J\}$, the homomorphic image of J .

Jump systems were introduced by Bouchet and Cunningham [1], who justified their importance by claiming that the following sets are jump systems:

- the set of degree sequences of all subgraphs of an undirected graph $G = (V, E)$, denoted by $J_G \subseteq \mathbb{N}^V$,
- the set of characteristic vectors of bases of a matroid $M = (S, r)$, denoted by $J_M \subseteq \mathbb{N}^S$.

The notion of the homomorphic image was introduced by Cunningham [3], who proved that the homomorphic image of a jump system is a jump system again.

DEFINITION 2.2. Let T be a finite set, $J \subseteq \mathbb{Z}^T$ be a jump system, and $H(t) \subseteq \mathbb{N}$ be a prescription for every $t \in T$. For $x \in \mathbb{Z}^T$ and $t \in T$ define

$$\delta_H^x(t) = \min\{|x(t) - i| : i \in H(t)\},$$

and let $\delta_H^x = \sum \{\delta_H^x(t) : t \in T\}$. Finally, $\delta_H = \min\{\delta_H^x : x \in J\}$. A vector $x \in J$ minimizing δ_H^x is called H -optimal in J .

A set $H \subseteq \mathbb{N}$ of the form $\{l, l+2, \dots, u-2, u\}$ is called a parity interval.

The membership problem of jump systems is to determine δ_H for a jump system J under the assumption that no $H(t)$ contains two consecutive gaps. An involved formula to the membership problem was worked out by Lovász [9]. We do not need this deep result here, only a simpler but very important tool, Lemma 2.4, which was developed by Sebő. This lemma appeared in Cornuéjols [2], where it was originally used in the construction of the first polynomial algorithm to the so-called *degree prescribed subgraph problem*, that is, the problem of determining δ_H for a jump system J_G , if $G = (V, E)$ is an undirected graph. Building on this lemma, it was possible to reduce the improvement of δ_H^x to the much easier problem of determining $\delta_{H'}$, where $H'(v)$ is a parity interval for all $v \in V$. This is exactly what we shall do but now for a jump system $\varphi(J_M)$, with M a matroid.

DEFINITION 2.3 (Sebő [2]). Let $J \subseteq \mathbb{Z}^T$ be a jump system and $H(t) \subseteq \mathbb{Z}$ for $t \in T$ be prescriptions without two consecutive gaps, and let $x \in J$. Define $I(t) \subseteq \mathbb{Z}$ to be the maximal interval such that $x(t) \in I(t)$ and the elements of

$$H^x(t) := I(t) \cap H(t)$$

have the same parity. (Observe that either $I(t)$ is unbounded from above or $\max I(t)$ and $\max I(t) + 1$ belong to $H(t)$. Similarly from below.) Now we define at most $2|T| + 1$ collections of prescriptions:

- H^x .
- ${}^w H^x$ for $w \in T$ with $I(w)$ bounded from above: ${}^w H^x(t) = H^x(t)$ for $t \neq w$ and ${}^w H^x(w) = \{\max I(w) + 1\}$.
- ${}_w H^x$ for $w \in T$ with $I(w)$ bounded from below: ${}_w H^x(t) = H^x(t)$ for $t \neq w$ and ${}_w H^x(w) = \{\min I(w) - 1\}$.

LEMMA 2.4 (Sebő [2]). Let $J \subseteq \mathbb{Z}^T$ be a jump system and $H(t) \subseteq \mathbb{Z}$ for $t \in T$ be prescriptions without two consecutive gaps. Let $x \in J$ be a vector which is not H -optimal in J . Then there exists an $x_0 \in J$ such that $\delta_{H_0}^{x_0} < \delta_H^x$ for either $H_0 = H^x$, or $H_0 = {}^w H^x$, or $H_0 = {}_w H^x$ for some $w \in T$.

Proof. Let $x^* \in J$ with $\delta_H^{x^*} < \delta_H^x$. Choose x' minimizing $\sum \{|x'(t) - x^*(t)| : t \in T\}$ among all vectors $x' \in J$ for which $\delta_H^{x'} \leq \delta_H^x$ and $x'(t) \in I(t)$ for all $t \in T$. If $\delta_H^{x'} < \delta_H^x$, then we are done with $x_0 = x'$ and $H_0 = H^x$. Otherwise there exists a coordinate $z \in T$ with $\delta_H^{x'}(z) > \delta_H^{x^*}(z)$. Assume that $x'(z) > x^*(z)$ (one can argue analogously if $x'(z) < x^*(z)$). Let $x'' = x' - \chi_z$ be a step from x' to x^* . It is easy to see that $\delta_H^{x''} < \delta_H^{x'}$ and $x''(t) \in I(t)$ for all $t \in T$. Thus $x'' \in J$ is impossible by the choice of x' . So we assume that there exists a step $x''' = x'' + \chi_w \in J$ from x'' to x^* (the case $x''' = x'' - \chi_w$ can be treated analogously). Clearly, $\delta_H^{x'''} \leq \delta_H^{x''} + 1 \leq \delta_H^{x'}$; hence by the choice of x' we get $x'''(w) \notin I(w)$. This can happen only if $x''(w) \in H(w)$ and $x''(w) + 1 = x'''(w) \in H(w)$, and thus $\delta_H^{x'''} = \delta_H^{x''} < \delta_H^x$. This means that we are done with $x_0 = x'''$ and $H_0 = {}^w H^x$. $\square$

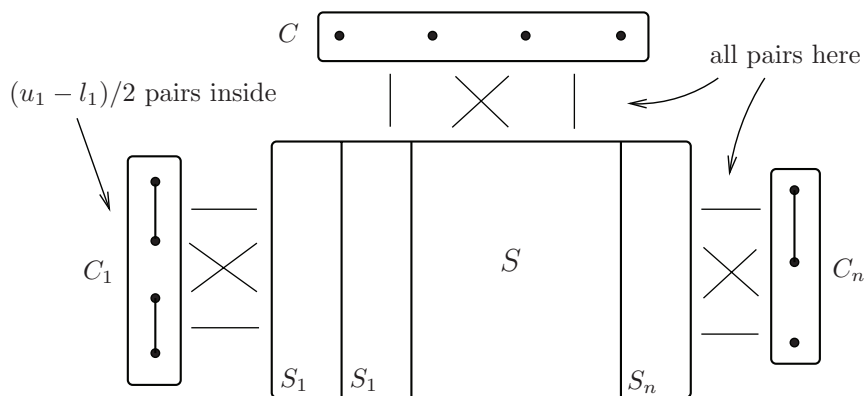
Proof of Theorem 1.3. Denote $r(M)$ simply by r , and let $l_t = \min H(t)$ and $u_t = \max H(t)$ for $t \in T$. We can clearly assume that $q \leq r$ and that $u_t \leq |S_t|$ for all $t \in T$. We already mentioned that the set of characteristic vectors of bases of a matroid is a jump system. As the q -truncation of M is a matroid, we have that

$$J' = \{\chi_F : F \subseteq S \text{ is independent in } M \text{ of size } q\} \subseteq \{0, 1\}^S$$

is a jump system. We define a function $\varphi : S \rightarrow T$, which maps $s \in S_t$ to $t \in T$ for all $s \in S$, and define $J = \varphi(J') \subseteq \mathbb{N}^T$. J is a jump system, and our goal is to decide whether $\delta_H = 0$ holds in J . Choose a vector $x \in J$. By Lemma 2.4 it is enough to construct a polynomial algorithm which finds an H_0 -optimal vector in J , where H_0 runs over the prescription sets listed in Definition 2.3. Fix such a prescription set H_0 , and recall that $H_0(t)$ is a parity interval for all $t \in T$.

Next we construct a representation over $\mathbb{Q}$ of an auxiliary linear matroid $\hat{M}$; see Figure 1. First, for all $t \in T$ add a set C_t of u_t coloops to M , and then add a set C of $r - q$ coloops to M . The representation of this new linear matroid $\hat{M}$ is trivial to construct, its rank is $\hat{r} = r + \sum u_t + r - q$, and its ground set $\hat{S}$ has size $|S| + \sum u_t + r - q$. We designate some pairs in $\hat{M}$ as follows: For all $t \in T$ take all $|C_t||S_t|$ pairs between C_t and S_t , together with $(u_t - l_t)/2$ disjoint pairs inside C_t , and then take all pairs between C and S .

Using the matroid parity algorithm of Lovász [7] we can find a maximum matching of $\hat{M}$ in polynomial time. Let us denote the size of this matching by ν and the union

FIG. 1. The auxiliary represented matroid $\hat{M}$.

of pairs in it by $K \subseteq \hat{S}$. Recall that K is independent in $\hat{M}$ and $|K| = 2\nu$. Denote by $K' \subseteq K$ the union of those pairs which are disjoint from C , and let $F' = K' \cap S$. Note that F' is independent in M . For a set $F \subseteq S$ we denote $\varphi\chi_F \in \mathbb{N}^T$ by $v(F)$. We will be done if we prove that if $F_0 \subseteq S$ is an independent set of M of size q , minimizing $|F_0 \Delta F'|$ (that is, if $|F_0 \Delta F'| = ||F'| - q|$), then $v(F_0)$ is H_0 -optimal in J .

First, it is not hard to see that

$$\delta_{H_0}^{v(F')}(t) \leq |C_t \setminus K| = u_t - |C_t \cap K| \quad \text{for all } t \in T.$$

Trivially, $r \geq |S \cap K|$ and $r - q \geq |C \cap K|$. Clearly,

$$q - |F'| \leq r - |S \cap K| \quad \text{and} \quad |F'| \leq r - |C \cap K|,$$

and thus

$$||F'| - q| \leq (r - |S \cap K|) + (r - q - |C \cap K|).$$

Summarizing,

$$\begin{aligned} \delta_{H_0}^{v(F_0)} &\leq \delta_{H_0}^{v(F')} + |F_0 \Delta F'| = \sum_{t \in T} \delta_{H_0}^{v(F')}(t) + ||F'| - q| \\ &\leq \sum_{t \in T} u_t - \sum_{t \in T} |C_t \cap K| + 2r - q - |(S \cup C) \cap K| = \hat{r} - |K|. \end{aligned}$$

On the other hand, if F is an independent set of M of size q such that $v(F)$ is H_0 -optimal in J , then it is straightforward to construct a matching of $\hat{M}$ with the union of size $\hat{r} - \delta_{H_0}^{v(F)}$. Thus $\delta_{H_0} = \hat{r} - |K|$, and $v(F_0)$ is indeed H_0 -optimal in J . $\square$

3. Applications to rigidity and to unique solvability of linear networks.

For a thorough background on rigidity we refer to Jackson and Jordán [4]. If $G = (V, E)$ is a graph and $p : V \rightarrow \mathbb{R}^d$ is an embedding into the Euclidean d -space, then (G, p) is said to be a *framework*. We think of the edges of G as rigid bars with flexible joints at the vertices. An *infinitesimal motion* means an assignment of velocity $x(v) \in \mathbb{R}^d$ to each vertex $v \in V$ such that the bar lengths are preserved, that is, $(p(u) - p(v)) \perp (x(u) - x(v))$ for all $uv \in E$. Given a subspace $U_v \leq \mathbb{R}^d$ for all $v \in V$ we say that (G, p) is *fixed* by the collection $(U_v : v \in V)$ if 0 is the only

infinitesimal motion under the constraint $x(v) \in U_v$ for all $v \in V$. We say that $Z \subseteq V$ is a *pinning set* if (G, p) is fixed by the collection

$$U_v = \begin{cases} \{0\} & \text{if } v \in Z, \\ \mathbb{R}^d & \text{if } v \notin Z. \end{cases}$$

We recollect the notion of the *rigidity matrix* $A \in \mathbb{R}^{E \times dV}$ of the framework (G, p) , where $dV = \{(v, i) : v \in V, 1 \leq i \leq d\}$: For each edge $uv \in E$, in the row indexed by uv , the entries in the d columns of u and v contain the d coordinates of $p(v) - p(u)$ and $p(u) - p(v)$, respectively, and the other entries are 0. The rigidity matrix has the important and well-known property that the infinitesimal motions of (G, p) are exactly the vectors $x \in \mathbb{R}^{dV}$ with $Ax = 0$. Let $Z \subseteq V$, and let A_Z be the submatrix of A we get when deleting the $d|Z|$ columns indexed by Z . Observe that $Z \subseteq V$ is a pinning set if and only if $x = 0$ is the only solution of $A_Z x = 0$. This is tantamount to saying that the columns of A_Z are linearly independent or, in other words, that they form an independent set in the column matroid M_A of A . Thus in dimension $d = 2$, the problem of finding a minimum pinning set leads to the parity problem of a linearly represented matroid, which can be solved by the algorithm of Lovász [7]. A combinatorial formula to this problem was subsequently given by Lovász in [6]. For $d = 3$, finding a minimum pinning set is NP-complete by Mansfield [11].

However, with the following modification we get a tractable problem for $d = 3$. Take a partition of V into two subsets V_1 and V_2 , and define $H(v) = \{0, i, 3\}$ for $v \in V_i$. $H(v)$ has no two consecutive gaps for all $v \in V$, so Theorem 1.3 implies that we can find in polynomial time the maximum q for which there exists an H -conformal independent set of M_A of size q . If $J \subseteq 3V$ is such an independent set and $U_v \leq \mathbb{R}^3$ is the subspace spanned by $\{e_i : (v, i) \in 3V \setminus J\}$ ($e_i \in \mathbb{R}^3$ is the i th unit vector), then by the above considerations (G, p) is fixed by $(U_v : v \in V)$. We obtain the following.

THEOREM 3.1. *The following problem is polynomial time solvable. Given a three-dimensional framework (G, p) and a partition $V = V_1 \dot{\cup} V_2$, find a collection $(U_v \leq \mathbb{R}^3 : v \in V)$ minimizing $\sum_{v \in V} \dim U_v$ such that (G, p) is fixed by $(U_v : v \in V)$ and*

- for $v \in V_1$ $U_v = \{0\}$, $U_v = \mathbb{R}^3$, or U_v is a coordinate line,
- for $v \in V_2$ $U_v = \{0\}$, $U_v = \mathbb{R}^3$, or U_v is a coordinate plane.

More generally, assign to every vertex $v \in V$ a set of vectors $\{a_i^v \in \mathbb{R}^d : i \in S_v\}$, where the S_v 's are disjoint finite sets and $S = \bigcup_{v \in V} S_v$. Then define a matrix $A \in \mathbb{R}^{E \times S}$ as follows. For all $uv = e \in E$ and $i \in S_v$ let

$$A_{e,i} = (p(u) - p(v))^T a_i^v,$$

and let the other entries be 0. We point out that if a_i^v is the i th unit vector for all $v \in V$ and $1 \leq i \leq d$, then A is simply the rigidity matrix. Again, if $y \in \mathbb{R}^S$, with $Ay = 0$, then $x \in \mathbb{R}^{dV}$, with $x(v) = \sum \{a_i^v y_i : i \in S_v\}$ is an infinitesimal motion of (G, p) . Thus given a framework (G, p) , and for each $v \in V$ a set of vectors $\{a_i^v \in \mathbb{R}^d : i \in S_v\}$ and prescription $H(v) \subseteq \mathbb{N}$ without two consecutive gaps, we can find in polynomial time a collection $(U_v \leq \mathbb{R}^d : v \in V)$ minimizing $\sum_{v \in V} \dim U_v$ with the following properties:

- (G, p) is fixed by $(U_v : v \in V)$,
- U_v is generated by some vectors a_i^v for $v \in V$,
- $\dim U_v \in H(v)$ for $v \in V$.

Another application of Theorem 1.3 is that we can check in polynomial time a weaker version of a sufficient condition obtained by Recski and Takács [12, 13] to the

unique solvability of linear networks with memoryless multiports. It is well known how to assign a *network graph* $G = (V, F)$ to an electric network containing voltage and current sources and linear memoryless multiports (e.g., resistors and gyrators) in such a way that V corresponds to the wires, disjoint n -tuples of F correspond to the n terminal-pairs of the n -ports, while the remaining edges of F correspond to the voltage and current sources (for a survey on electric networks, see Recski [15]). Recski and Takács [12, 13] introduced the notion of *potentially singular* multiports, which are the only possible devices in the network which can contribute to singularities. If D is a potentially singular n -port and $F_D \subseteq F$ is the set of n corresponding edges in G , then they introduce a collection of k_D partitions of F_D : $\pi_D^j = \{L^j, C^j, N_1^j, \dots, N_{l_j}^j\}$ for $1 \leq j \leq k_D$. They extend the well-known notion of a normal tree as follows.

DEFINITION 3.2 (see [13]). *A normal tree of the network graph $G = (V, F)$ is a spanning tree T of G satisfying the following conditions:*

1. T contains all edges corresponding to voltage sources.
2. T contains no edges corresponding to current sources.
3. If D is a potentially singular n -port, then for all $1 \leq j \leq k_D$ at least one of the following statements is met:
 - (a) $E(T) \cap L^j \neq \emptyset$,
 - (b) $C^j \not\subseteq E(T)$,
 - (c) $|E(T) \cap N_i^j| \neq 1$ for at least one index i with $1 \leq i \leq l_j$.

They proved the following.

THEOREM 3.3 (Recski and Takács [13]). *If the network graph has a normal tree, then the network has a unique solution.*

We show how to decide the existence of a normal tree *under the assumption that* $k_D = 1$ *for all potentially singular multiports* D . First recall that the edge sets of the spanning trees of a graph form the bases of the cycle matroid M_G of $G = (V, F)$. We define a hypergraph $H = (W, \mathcal{E})$, with $F \subseteq W$, and a collection of prescriptions $(H(e) \subseteq \mathbb{N} : e \in \mathcal{E})$ as follows. For every voltage source (resp., current source) and corresponding edge $f \in F$ add a hyperedge $e = \{f\}$ to $\mathcal{E}$ with prescription $H(e) = \{1\}$ (resp., $H(e) = \{0\}$). Then do the following for each potentially singular n -port D with partition $\pi_D = \{L, C, N_1, \dots, N_l\}$ of F_D . First add a set $X = \{x_D, y_D\}$ consisting of two coloops to M_G , and let $\tilde{L} = L \cup \{x_D\}$. In addition, add the following hyperedges to $\mathcal{E}$ with prescriptions as indicated:

1. the classes of $\pi_D - L + \tilde{L}$, with $H(\tilde{L}) = \{1, 2, \dots, |\tilde{L}|\}$, $H(C) = \{0, 1, \dots, |C| - 1\}$, and $H(N_i) = \{0, 2, 3, \dots, |N_i|\}$ for $1 \leq i \leq l$,
2. $U := \bigcup_{1 \leq i \leq l} N_i$, with $\{0, 1, \dots, l - 1\}$,
3. X with $\{0, 2\}$,
4. and $Y := F_D \cup \{y_D\}$ with $\{0, 1, 2\}$.

Denote the resulting matroid, which is just M_G plus some coloops, by M . Finally, construct a matroid M' from M as follows: For every element f of M which is incident to the hyperedges $e_1, \dots, e_d \in \mathcal{E}$ for $d = \deg_H(f)$, replace f by a parallel class $\{f^{e_1}, \dots, f^{e_d}\}$. For $e \in \mathcal{E}$ let $e' = \{f^e : f \in e\}$. In this way $\mathcal{E}' := \{e' : e \in \mathcal{E}\}$ is a partition of the new ground set W' .

CLAIM 3.4. *If B' is an H -conformal base of M' , and B is the corresponding base of M then $B \cap F$ is the edge set of a normal tree in G .*

Proof. If $f \in F$ corresponds to a voltage source (resp., current source) then, as $|B \cap \{f\}| \in H(\{f\})$, we have $f \in B$ (resp., $f \notin B$). Now let D be a potentially singular n -port.

- (a) If $B \cap L = \emptyset$, then necessarily $B' \cap \tilde{L}' = \{x_D^{\tilde{L}}\}$; thus $B' \cap X' = \emptyset$ and so $y_D^Y \in B' \cap Y'$.

- (b) If $C \subseteq B$, then $B' \cap Y'$ contains an element s^Y , with $s \in C$.
- (c) If $|B \cap N_i| = 1$ for all $1 \leq i \leq l$, then $B' \cap N'_i = \emptyset$. As $|B' \cap U'| \leq l - 1$ we can conclude that $B' \cap Y'$ contains an element s^Y , with $s \in U$.

Thus if all of (a), (b), and (c) are violated in Definition 3.2, condition 3, then $|B' \cap Y'| \geq 3$, which is impossible by the hyperedge defined in 4. $\square$

Using similar arguments, from a normal tree of G it is not hard to construct an H -conformal base of M' . As a representation of M' over $\mathbb{Q}$ is trivial to construct using the well-known representation of M_G , we can decide in polynomial time whether M' has an H -conformal base, and we are done.

Acknowledgments. I would like to thank Tibor Jordán, for his suggestions and for calling my attention to the application in rigidity, and András Recski for helpful discussions.

REFERENCES

- [1] A. BOUCHET AND W. H. CUNNINGHAM, *Delta-matroids, jump systems, and bisubmodular polyhedra*, SIAM J. Discrete Math., 8 (1995), pp. 17–32.
- [2] G. CORNUÉJOLS, *General factors of graphs*, J. Combin. Theory Ser. B, 45 (1988), pp. 185–198.
- [3] W. H. CUNNINGHAM, *Improved bounds for matroid partition and intersection algorithms*, SIAM J. Comput., 15 (1986), pp. 948–957.
- [4] B. JACKSON AND T. JORDÁN, *The d -dimensional rigidity matroid of sparse graphs*, J. Combin. Theory Ser. B, 95 (2005), pp. 118–133.
- [5] P. M. JENSEN AND B. KORTE, *Complexity of matroid property algorithms*, SIAM J. Comput., 11 (1982), pp. 184–190.
- [6] L. LOVÁSZ, *Matroid matching and some applications*, J. Combin. Theory Ser. B, 28 (1980), pp. 208–236.
- [7] L. LOVÁSZ, *Selecting independent lines from a family of lines in a space*, Acta Sci. Math. (Szeged), 42 (1980), pp. 121–131.
- [8] L. LOVÁSZ, *The matroid matching problem*, in Algebraic Methods in Graph Theory, Vols. I and II, Colloquia Mathematica Societatis János Bolyai 25, North-Holland, Amsterdam, 1981, pp. 495–517.
- [9] L. LOVÁSZ, *The membership problem in jump systems*, J. Combin. Theory Ser. B, 70 (1997), pp. 45–66.
- [10] M. MAKAI, *Parity Problems of Combinatorial Polymatroids*, Ph.D. thesis, Eötvös University, Budapest, Hungary, 2007; also available online from www.cs.elte.hu/~marci.
- [11] A. MANSFIELD, *On the computational complexity of a rigidity problem*, IMA J. Appl. Math., 27 (1981), pp. 423–429.
- [12] A. RECSKI, *Sufficient conditions for the unique solvability of linear networks containing memoryless 2-ports*, Internat. J. Circuit Theory Appl., 8 (1980), pp. 95–103.
- [13] A. RECSKI AND J. TAKÁCS, *On the combinatorial sufficient conditions for linear network solvability*, Internat. J. Circuit Theory Appl., 9 (1981), pp. 351–354.
- [14] A. RECSKI, *On the generalization of the matroid parity and the matroid partition problems, with applications*, in Combinatorial Mathematics, North-Holland Math. Stud. 75, Elsevier, Amsterdam, 1981, pp. 567–573.
- [15] A. RECSKI, *Matroid Theory and Its Applications in Electric Network Theory and in Statics*, Akadémiai Kiadó, Budapest, Hungary, 1989.
- [16] A. SCHRIJVER, *Combinatorial Optimization. Polyhedra and Efficiency*, Algorithms Combin. 24, Springer-Verlag, Berlin, 2003.

CYCLES INTERSECTING EDGE-CUTS OF PRESCRIBED SIZES*

TOMÁŠ KAISER[†] AND RISTE ŠKREKOVSKI[‡]

Abstract. We prove that every cubic bridgeless graph G contains a 2-factor which intersects all (minimal) edge-cuts of size 3 or 4. This generalizes an earlier result of the authors, namely that such a 2-factor exists provided that G is planar. As a further extension, we show that every graph contains a cycle (a union of edge-disjoint circuits) that intersects all edge-cuts of size 3 or 4. Motivated by this result, we introduce the concept of a coverable set of integers and discuss a number of questions, some of which are related to classical problems of graph theory such as Tutte's 4-flow conjecture and the Dominating Cycle Conjecture.

Key words. graph, edge-cut, coverable set, dominating cycle

AMS subject classification. 05C38

DOI. 10.1137/070683635

1. Introduction. We study the existence of cycles intersecting all edge-cuts of prescribed sizes in a graph. Throughout this paper, a *cycle* in a graph G is a subgraph all of whose vertex degrees are even, and an *edge-cut* (in short, a *cut*) is an inclusionwise minimal set of edges whose removal increases the number of components of G . Our graphs are undirected and contain no loops, but they may contain parallel edges.

Our starting point is the main result of [8].

THEOREM 1.1. *For any planar graph G , there exists a (not necessarily proper) 2-coloring of $V(G)$ such that there is no monochromatic circuit of length 3 or 4.*

In an equivalent dual form, Theorem 1.1 states that every bridgeless planar cubic graph has a 2-factor intersecting all cuts of size 3 or 4. (A graph is *bridgeless* if it is connected and has no bridges.) In the present paper, we extend the latter result to all bridgeless cubic graphs. Furthermore, we lift the regularity assumption, proving the following.

THEOREM 1.2. *Every graph G has a cycle intersecting all cuts of size 3 or 4.*

We note the following direct consequence of this result.

COROLLARY 1.3. *Every 3-edge-connected graph contains a cycle C such that the graph G/C (obtained by contracting each component of C to a vertex) is 5-edge-connected.*

Motivated by Theorem 1.2, we introduce a concept which plays a central role in this paper. Let $\mathbb{N}$ be the set of positive integers and $A \subseteq \mathbb{N}$. We say that a cycle C in a graph G is *A-covering* if it intersects all cuts T with $|T| \in A$. If $\mathcal{Q}$ is a class of graphs, then A is *coverable in $\mathcal{Q}$* if every graph from $\mathcal{Q}$ contains an A -covering cycle. A set that is coverable in the class of all graphs is just said to be *coverable*.

*Received by the editors February 25, 2007; accepted for publication (in revised form) December 20, 2007; published electronically May 14, 2008. The research was supported by project 1M0545 and Research Plan MSM 4977751301 of the Czech Ministry of Education.

<http://www.siam.org/journals/sidma/22-3/68363.html>

[†]Department of Mathematics and Institute for Theoretical Computer Science, University of West Bohemia, Univerzitní 8, 306 14 Plzeň, Czech Republic (kaisert@kma.zcu.cz).

[‡]Department of Mathematics, University of Ljubljana, Jadranska 19, 1111 Ljubljana, Slovenia, and Institute for Theoretical Computer Science, Charles University, Malostranské náměstí 25, 118 00 Praha 1, Czech Republic.

Thus, an equivalent version of Theorem 1.2 is that the set $\{3, 4\}$ is coverable. Which other sets are coverable? $\mathbb{N}$ itself is not; clearly, a connected graph has an $\mathbb{N}$ -covering cycle if and only if it has a spanning Eulerian subgraph (spanning connected cycle), which is not the case, for instance, for the graph $K_{2,3}$ (or for any graph with a bridge). In fact, $K_{2,3}$ shows that even the set $\{2\}$ is not coverable.

For a less trivial example of a noncoverable set, consider $A = \{3, 5\}$ and the Petersen graph P_{10} . For any vertex v of P_{10} , the edges incident with v constitute a cut as P_{10} is 3-edge-connected. Since $3 \in A$, any A -covering cycle is a 2-factor. Every 2-factor F of P_{10} is formed by two circuits of length 5. The complement of F is a cut of size 5 that is not intersected by F . It follows that P_{10} has no A -covering cycle.

On the other hand, it may well be that the presence of P_{10} in a graph G (as a minor) is the only obstruction to the existence of a $\{3, 5\}$ -covering cycle in G . Recall that a graph H is a *minor* of a graph G if H can be obtained from a subgraph of G by a sequence of edge contractions. The graph G is *Petersen-minor-free* (or *P_{10} -free*) if P_{10} is not a minor of G . Petersen-minor-free graphs are the subject of the famous *4-flow conjecture* of Tutte [15]. Since we will not need to go into the details of integer flows (which can be found in [16]), let us state Tutte's conjecture in a form that does not refer to 4-flows.

CONJECTURE 1.4. *The edges of any P_{10} -free bridgeless graph can be covered by two cycles.*

Observe that if $E(G)$ is covered by two cycles C_1 and C_2 , then C_1 is $(2\mathbb{N} + 1)$ -covering, where $2\mathbb{N} + 1 = \{3, 5, 7, \dots\}$. Indeed, any odd cut not intersected by C_1 cannot be covered by C_2 , for the intersection of a cycle and a cut has even size. Conversely, it is not difficult to prove that if G has a $(2\mathbb{N} + 1)$ -covering cycle, then $E(G)$ can be covered by two cycles. Thus, Conjecture 1.4 can equivalently be stated in terms of coverability.

CONJECTURE 1.5. *The set $2\mathbb{N} + 1 = \{3, 5, 7, \dots\}$ is coverable in the class of P_{10} -free graphs.*

Observe that Conjecture 1.5 is not restricted to bridgeless graphs. The reason is that any set $A \subseteq \mathbb{N}$ with $1 \notin A$ is coverable in the class of bridgeless graphs if and only if it is coverable in the class of all graphs.

Conjecture 1.4 is well known to be true for planar graphs. Indeed, this special case is equivalent to the Four Color Theorem (see, e.g., [16]). It follows that $2\mathbb{N} + 1$ is coverable in the class of planar graphs.

To conclude this section, we point out a relation to another long-standing conjecture. A *dominating cycle* in G is a circuit C such that each edge of G is incident with a vertex of C . (Note that in our terminology, “dominating circuit” would be a more appropriate term.) The *Dominating Cycle Conjecture* has several equivalent forms [3, 10, 13]; we state the one due to Fleischner and Jackson [4] (see section 2 for a definition of cyclically k -connected graphs).

CONJECTURE 1.6. *Every cyclically 4-edge-connected cubic graph has a dominating cycle.*

By Tutte's theorem [14], Conjecture 1.6 is true for planar graphs. Note that if G is a cyclically 4-edge-connected cubic graph, then a circuit is a dominating cycle in G if and only if it is $(\mathbb{N} + 3)$ -covering, where $\mathbb{N} + 3 = \{4, 5, 6, \dots\}$. Thus, the following is a generalization of the Dominating Cycle Conjecture.

CONJECTURE 1.7. *The set $\mathbb{N} + 3$ is coverable.*

A result of Thomassen [12, Theorem 4.1] implies that $\mathbb{N} + 3$ is coverable in the class of planar graphs. Further questions related to coverable sets are asked in section 6.

2. Notation and definitions. Let us review a few definitions. As mentioned above, all the graphs we consider are loopless multigraphs. The vertex set and the edge set of a graph G are denoted by $V(G)$ and $E(G)$, respectively. If $E = E(G)$, we write $E(v)$ for the set of edges incident with the vertex v . For a subset $X \subseteq V(G)$, let $G[X]$ denote the subgraph of G induced by the vertices of X . If V_1, V_2 are vertex-disjoint subsets of $V(G)$, let $[V_1, V_2]$ denote the set of edges of G with one endvertex in V_1 and the other endvertex in V_2 . Recall that a *cut* in a connected graph G is a subset $C \subseteq E(G)$ such that $G - C$ is disconnected and C is minimal with this property. Note that $G - C$ has two components, say, G_1 and G_2 , and that $C = [V(G_1), V(G_2)]$. If any of the graphs G_i consists of a single vertex, then C is a *trivial* cut; otherwise C is called *nontrivial*. Similarly, if any of the graphs G_i is a tree, then C is an *acyclic* cut; otherwise C is *cyclic*.

We write $G_1(C)$ ($G_2(C)$, respectively) for the graph obtained from G by contracting all of G_2 (G_1 , respectively) to a new vertex and removing any loops which arise. Note that in both graphs thus constructed, C corresponds to a trivial cut.

Let v be a vertex of degree 2 in a graph. *Suppressing* v consists in removing v along with the incident edges, and joining the former neighbors of v by an edge if they are distinct. Note that if v is incident with two parallel edges, then suppressing v is the same as removing it. To *contract* an edge means to identify its endvertices and remove all the resulting loops. By definition, the contraction of a subgraph is the contraction of all of its edges.

A graph G is *cyclically k -edge-connected* if it contains no cyclic cut of size at most $k - 1$. A cycle H is *spanning* in G if each vertex of G is incident with an edge of H .

We refer to edge cuts of size k as *k -cuts*. Similarly, a set of size n is referred to as an *n -set*. A set of even size is an *even set*. The terms *n -subset* and *even subset* are defined in an analogous way.

3. Interlaced cuts. Before proving Theorem 1.2 (in sections 4 and 5), we need to obtain some information on the possible configurations of cuts of size 3 and 4. Throughout this section, G is a connected graph and $C_1 = [X_1, X_2]$ and $C_2 = [Y_1, Y_2]$ are two cuts in G . For $i, j \in \{1, 2\}$, set $A_{ij} = X_i \cap Y_j$. (See Figure 1 for an illustration.) The sets A_{ij} are called the *regions* corresponding to C_1 and C_2 . We say that C_1 *interlaces* C_2 if each of $G[X_1]$ and $G[X_2]$ contains an edge of C_2 . In this section, we study the interlacement relation for small cuts in connected graphs.

PROPOSITION 3.1. *With the above definitions,*

$$C_1 \cap C_2 = [A_{11}, A_{22}] \cup [A_{12}, A_{21}].$$

Proof. An edge of C_1 joins a vertex of X_1 to a vertex of X_2 , and an edge of C_2 joins a vertex of Y_1 to a vertex of Y_2 . Thus, if $e \in C_1 \cap C_2$, then either e has one endvertex in $X_1 \cap Y_1 = A_{11}$ and the other one in $X_2 \cap Y_2 = A_{22}$, or else e has one endvertex in $X_1 \cap Y_2 = A_{12}$ and the other one in $X_2 \cap Y_1 = A_{21}$. $\square$

By the following proposition, C_1 interlaces C_2 if and only if C_2 interlaces C_1 . If these equivalent conditions hold, then we say that C_1 and C_2 are *interlaced* or that they form an *interlacing pair*.

PROPOSITION 3.2. *The following three claims are equivalent (under our assumption that G is connected):*

- (a) C_1 interlaces C_2 ,
- (b) C_2 interlaces C_1 ,
- (a) (c) all the sets A_{ij} ($i, j \in \{1, 2\}$) are nonempty.

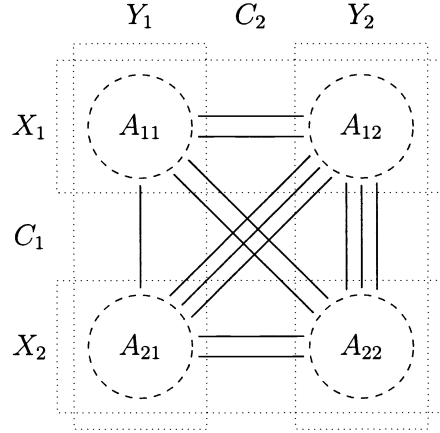


FIG. 1. The regions A_{ij} corresponding to cuts $C_1 = [X_1, X_2]$ and $C_2 = [Y_1, Y_2]$.

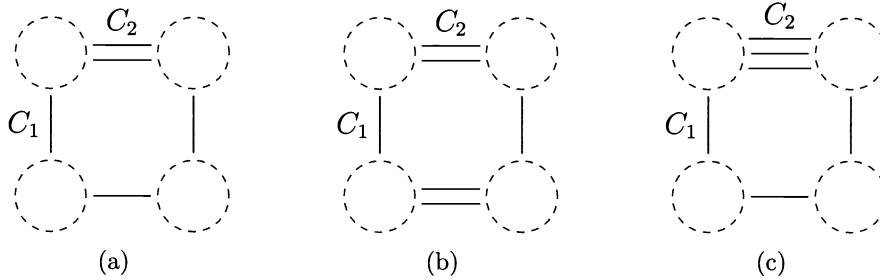


FIG. 2. A 2-cut C_1 (vertical edges) interlaced with a cut C_2 of size 3 or 4 (horizontal edges).

Proof. By symmetry, it is enough to prove that (a) $\Rightarrow$ (c) $\Rightarrow$ (b).

We first show that (a) $\Rightarrow$ (c). Since C_1 interlaces C_2 , there exists an edge of C_2 with both endvertices in X_1 . Notice that one of these two endvertices is in $X_1 \cap Y_1$ and the other one is in $X_1 \cap Y_2$. Thus, $A_{11} \neq \emptyset$ and $A_{12} \neq \emptyset$. Similarly, one can show that $A_{21} \neq \emptyset$ and $A_{22} \neq \emptyset$.

It remains to show that (c) $\Rightarrow$ (b). The graph G has an edge e_1 with one endvertex in A_{11} and the other one in A_{21} . Otherwise, since G is connected, the assumptions $A_{11} \neq \emptyset$ and $A_{21} \neq \emptyset$ imply that the set $C^* = [A_{11}, V(G) \setminus A_{11}]$ is nonempty and that it is a proper subset of C_2 . Thus, C^* contradicts the minimality of C_2 . Since $e_1 \in [A_{11}, A_{21}]$ belongs to Y_1 , it follows that $e_1 \in C_1$ and $e_1 \in G[Y_1]$. Similarly, one can show that C_1 has an edge e_2 in $G[Y_2]$. Thus, C_2 interlaces C_1 . $\square$

From the above two propositions, we easily obtain the following:

PROPOSITION 3.3. *Suppose that the cuts C_1 and C_2 are interlaced. Then there exist edges $e_1, e_2 \in C_1 \setminus C_2$ and $f_1, f_2 \in C_2 \setminus C_1$ such that for each $i \in \{1, 2\}$, it holds that $e_i \in [A_{1i}, A_{2i}]$ and $f_i \in [A_{i1}, A_{i2}]$.*

PROPOSITION 3.4. *Let C_1 and C_2 be interlaced cuts in a graph G . If $|C_1| = 2$, then $C_1 \cap C_2 = \emptyset$.*

Proof. Apply Propositions 3.1 and 3.3. $\square$

Observe that by Propositions 3.1–3.3, all the possible mutual positions of the cuts C_1, C_2 from Proposition 3.4, where $|C_2| \in \{3, 4\}$, are as shown in Figure 2.

PROPOSITION 3.5. *If G is a 3-edge-connected graph, then it has no interlacing pair of 3-cuts.*

Proof. Suppose that C_1 and C_2 are two such cuts. Propositions 3.1–3.3 imply that one of the sets $A_{11}, A_{12}, A_{21}, A_{22}$ is connected to the rest of G by at most 2 edges. This contradicts the 3-edge-connectedness of G . $\square$

Propositions 3.1 and 3.3 imply the following:

PROPOSITION 3.6. *Suppose that G is a 3-edge-connected graph, C_1 is a 3-cut, C_2 is a 4-cut, and the cuts C_1 and C_2 are interlaced. Then $C_1 \cap C_2 = \emptyset$.*

The structure of G with respect to the cuts C_1, C_2 from Proposition 3.6 is necessarily the one shown in Figure 3.

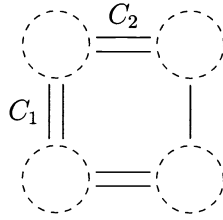


FIG. 3. A 3-cut C_1 interlaced with a 4-cut C_2 .

PROPOSITION 3.7. *Suppose that G is a 3-edge-connected graph without nontrivial 3-cuts, C_1 and C_2 are interlaced 4-cuts, and C_1 is cyclic. Then the following hold:*

- (a) $C_1 \cap C_2 = \emptyset$;
- (b) if C_2 is acyclic, then one component of $G - C_2$ consists of a single edge;
- (c) if C_2 is cyclic, then each component of $G - C_1 - C_2$ is incident with exactly two edges of C_1 and two edges of C_2 .

Proof. (a) Suppose that $C_1 \cap C_2 \neq \emptyset$. If the set $[A_{11}, A_{22}]$ contains at least two edges, then Proposition 3.3 implies that each of the sets A_{12}, A_{21} is connected to the rest of G by at most 2 edges; this contradicts the edge-connectivity assumption on G . Thus, $[A_{11}, A_{22}]$ and (by a symmetric argument) $[A_{12}, A_{21}]$ contain at most one edge each.

Suppose now that the set $[A_{12}, A_{21}]$ is empty. Then $[A_{11}, A_{22}]$ contains precisely one edge. Using the assumption that all 3-cuts are trivial, one can easily show that either one of the sets A_{ij} ($i, j \in \{1, 2\}$) is connected to the rest of the graph by at most 2 edges, or for some $i, j \in \{1, 2\}$, the subgraphs $G[X_i]$ and $G[Y_j]$ are isomorphic to K_2 . The former possibility contradicts the edge-connectivity assumption on G , while the latter one contradicts the assumption that C_1 or C_2 is cyclic.

By the above, each of the sets $[A_{11}, A_{22}]$ and $[A_{12}, A_{21}]$ contains precisely one edge. Proposition 3.3 implies that each A_{ij} is connected to the rest of the graph by exactly 3 edges. Since G contains no nontrivial 3-cuts, it follows that each A_{ij} consists of a single vertex, and so G is isomorphic to K_4 . Hence, both of the cuts C_1 and C_2 are acyclic, a contradiction.

(b) Given an acyclic 4-cut C_2 , let K be a component of $G - C_2$ that contains no cycle. Thus, K is a tree, and since K contains an edge of C_1 (because C_1 and C_2 are interlaced), we have $|V(K)| \geq 2$. It follows that K has at least two leaves. Since G is 3-edge-connected, each leaf of K is incident with at least two edges of C_2 . This implies that K has exactly two leaves, and hence is a path. Any internal vertex of the path would also have to be incident with an edge of C_2 , which is impossible. We conclude that K consists of a single edge.

(c) Assume that both C_1 and C_2 are cyclic and let K be a component of $G - C_1 - C_2$. Also let L denote the component of $G - C_2$ containing K . By symmetry, all we need is to show that K is incident with at least two edges of C_1 . Suppose the contrary. Then either K or the other component K' of $G - C_1 - C_2$ contained in L is incident with at most 3 edges of $C_1 \cup C_2$. Our assumption that G is 3-edge-connected implies that, in fact, both K and K' are incident with exactly two edges of C_2 and one edge of C_1 . Since G has no nontrivial 3-cuts, both K and K' consist of a single vertex. We conclude that L consists of a single edge, which contradicts the assumption that $G - C_2$ is cyclic. $\square$

The structure of G with respect to the cuts C_1, C_2 from Proposition 3.7 is as depicted in Figure 4.

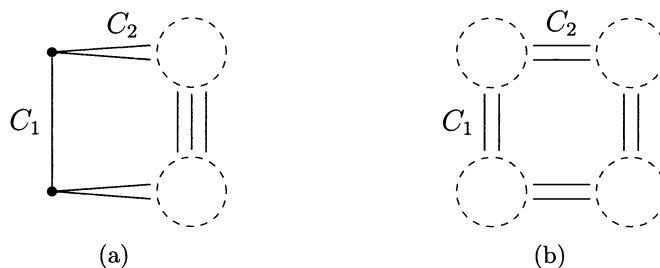


FIG. 4. Two interlaced 4-cuts from Proposition 3.7.

PROPOSITION 3.8. If C_1 and C_2 are distinct noninterlaced cuts with $C_1 \cap C_2 \neq \emptyset$, then precisely one of the sets $A_{11}, A_{12}, A_{21}, A_{22}$ is empty.

Proof. Note that by Proposition 3.2, at least one of the sets A_{ij} is empty. Since the cuts are distinct, no more than one of these sets can be empty. $\square$

The structure of two distinct noninterlaced cuts with nonempty intersection is shown in Figure 5. Note that between any two of the 3 parts, there must be at least one edge.

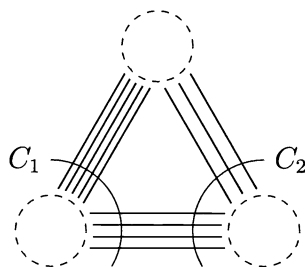


FIG. 5. Noninterlaced cuts C_1 and C_2 with common edges.

4. Graphs with small degrees. By a well-known theorem of Petersen, every bridgeless cubic graph G has a 2-factor. In this section, we prove a result which implies that in fact, G has a 2-factor which is a $\{3, 4\}$ -covering cycle. To this end, we shall make use of the following extension of Petersen's theorem, due to Schönberger [11].

THEOREM 4.1. Let G be a cubic bridgeless multigraph and $e, f \in E(G)$. Then G has a 2-factor containing both e and f .

Let v be a vertex of degree 4 in a graph $G = (V, E)$. Let $Y \subset E(v)$ be a 2-set and let $X \subseteq E(v)$ be a nonempty even set. Note that if $|X| \neq 2$, then $X = E(v)$. We say

that X crosses Y if $X \cap Y \neq \emptyset$ and $Y \neq X$. Thus, $E(v)$ crosses each of its 2-subsets. For example, if $E(v) = \{a, b, c, d\}$ and $Y = \{a, b\}$, then the even sets which cross Y are $\{a, c\}$, $\{a, d\}$, $\{b, c\}$, $\{b, d\}$, and $\{a, b, c, d\}$.

Let $w \in V(G)$. We say that a subgraph $H \subseteq G$ extends a set $X \subseteq E(w)$ if $E(H) \cap E(w) = X$. For brevity, we call a cycle *good* if it is $\{3, 4\}$ -covering.

For a partition E_1, E_2 of $E(w)$, let $G(w, E_1, E_2)$ denote the graph constructed from G by splitting w into two new adjacent vertices w_1 and w_2 so that $E(w_1) = E_1 \cup \{w_1 w_2\}$ and $E(w_2) = E_2 \cup \{w_1 w_2\}$. Notice that G can be obtained from $G(w, E_1, E_2)$ by contracting the edge $w_1 w_2$. Furthermore, any cut in G is a cut in $G(w, E_1, E_2)$. Thus, a good cycle in $G(w, E_1, E_2)$ gives rise to a good cycle in G .

If C is a cut in a graph G , then we identify C with the corresponding edge-sets in $G_1(C)$ and $G_2(C)$. We say that a subgraph F_1 of $G_1(C)$ and a subgraph F_2 of $G_2(C)$ agree on C if they contain precisely the same edges of C (with respect to this identification). For such a pair of subgraphs, $F_1 \cup F_2$ denotes the subgraph of G consisting of all edges corresponding to those in F_1 or in F_2 .

LEMMA 4.2. *Suppose that C is a minimum size cut in a bridgeless graph G and $|C| \leq 3$. If each graph $G_i(C)$ ($i = 1, 2$) has a good cycle F_i such that F_1 and F_2 agree on C , then $F_1 \cup F_2$ is a good cycle of G .*

Proof. We need to verify that $F_1 \cup F_2$ intersects each cut of size 3 or 4 in G . Let D be such a cut. Let A_{ij} ($i, j \in \{1, 2\}$) be the regions corresponding to the cuts $C_1 = C$ and $C_2 = D$ in G as defined at the beginning of section 3. If C and D are not interlaced, then D corresponds to a cut D' of the same size in some $G_i(C)$ (in $G_1(C)$, say). Since F_1 is a good cycle, it intersects D' , which implies that $F_1 \cup F_2$ intersects D .

We can thus assume that C and D are interlaced. By Propositions 3.4–3.6, the structure of G with respect to C and D is shown in Figure 2 or 3. Observe that in each of the possible cases, there is a region (say, A_{11}) incident with a single edge of C and 2 or 3 edges of D . Let R be the set of edges of $G_1(C)$ with exactly one endvertex in A_{11} . Since R is a cut of size 3 or 4 in $G_1(C)$, it is intersected by F_1 . As $|R \cap C| = 1$, F_1 must use at least one edge of D . It follows that $F_1 \cup F_2$ intersects D . $\square$

The following theorem deals with $\{3, 4\}$ -covering cycles in graphs with maximum degree at most 4, the focus being on cubic graphs where any such cycle is a 2-factor.

THEOREM 4.3. *Let G be a 2-connected graph and let v be a vertex of G . Assume that v is of degree at most 4 and all the other vertices of G are of degree at most 3.*

- (a) *If v is of degree at most 3, then each 2-set $Y \subset E(v)$ can be extended to a good cycle of G .*
- (b) *If v is of degree 4, then there exists a 2-set $X \subset E(v)$ such that every even set $Y \subseteq E(v)$ which crosses X can be extended to a good cycle of G .*

Before proving Theorem 4.3, let us consider an example. Let v be a vertex of degree 4 and write $E(v) = \{a, b, c, d\}$ and $Y = \{a, b\}$. Part (b) of the theorem claims that each of the sets $\{a, c\}$, $\{a, d\}$, $\{b, c\}$, $\{b, d\}$, and $\{a, b, c, d\}$ can be extended to a $\{3, 4\}$ -covering cycle of G .

Proof of Theorem 4.3. Suppose that the theorem is false and G is a counterexample with the minimum number of vertices. Since G is 2-connected, $|V(G)| > 2$ and G has no vertices of degree 1. In a series of claims, we show that G is cyclically 5-edge-connected.

Whenever we consider a cut $C = [X_1, X_2]$ and the graphs $G_1(C)$ and $G_2(C)$, we write v_1 for the vertex of $G_2(C)$ obtained by contracting X_1 . Similarly, we let v_2 denote the vertex of $G_1(C)$ obtained by contracting X_2 .

CLAIM 1. *The graph G contains no vertex of degree 2.*

If $w \neq v$ is a vertex of degree 2, then use the induction hypothesis to find a good cycle in the graph obtained by suppressing w and note that this yields a suitable good cycle in G . We may thus assume that the only vertex of degree 2 is v . Let z be a neighbor of v . Since G is 2-connected, v and z are joined by a single edge. Consider the graph G' obtained by suppressing v and let e be the newly created edge. The degree of z in G' is 3. Let $Y \subset E(G')$ be a set consisting of e and one other edge incident with z . Use the induction hypothesis to find a good cycle of G' extending Y . This yields a suitable good cycle of G .

CLAIM 2. *The degree of v is 4.*

Assume that v has degree 3. Let $Y = \{a_1, a_2\} \subset E(v)$ be the given 2-set, and let b be the edge in $E(v) \setminus Y$. Let w be the endvertex of a_1 distinct from v . If the edges a_1, a_2 are parallel, then since G is bridgeless, $|V(G)| > 2$, and the degree of w is 3, there is a unique edge $c \in E(w) \setminus Y$. This implies that $\{b, c\}$ is a 2-cut. Let G' be the graph obtained from $G \setminus \{v, w\}$ by adding an edge e' joining the endvertex $v' \neq v$ of b to the endvertex $w' \neq w$ of c . The edge e' is not a loop, as otherwise G would contain a bridge incident with v' .

Note that the degree of v' in G' is 3. Let Y' be the set of edges incident with v' distinct from e' . By the induction hypothesis (in which v' plays the role of v), G' has a good cycle F' that extends Y' and, therefore, does not contain e' . Since any cut of size 3 or 4 in G corresponds to a cut of size 3 or 4 in G' (although not necessarily a cut of the same size), it follows that adding the edges a_1 and a_2 to F' , we obtain a good cycle of G extending Y .

It remains to consider the case that a_1 and a_2 are not parallel. Let G'' be the graph obtained by contracting a_1 . Clearly, we may assume that b is not parallel to a_1 , so the contraction does not create a loop. Note that G'' contains a unique vertex v'' of degree 4. Write E'' for the set of edges incident with v'' . By the induction hypothesis (applied to v''), find a 2-set $X'' \subset E''$ with the property stated in the theorem. Since a subset of E'' crosses X'' if and only if it crosses its complement in E'' , we may assume that $a_2 \in X''$. Observe that there are exactly two 2-subsets of E'' containing a_2 and crossing X'' . It follows that at least one of them—call it Y'' —is different from $\{a_2, b\}$. By the assumption, Y'' can be extended to a good cycle F'' of G'' . The corresponding edges in G , together with a_1 , comprise a good cycle of G .

CLAIM 3. *The graph G is 3-edge-connected and cyclically 4-edge-connected.*

Assume the claim false. Since G is bridgeless, there is a cut C that is either a 2-cut or a cyclic 3-cut. Consider the graphs $G_1(C)$ and $G_2(C)$. We may assume that $v \in V(G_1(C))$. Let E_1 be the set of edges of $G_1(C)$ incident with v . By the minimality of G , there is a 2-set $X \subset E_1$ such that any even set $Y \subset E_1$ crossing X can be extended to a good cycle of $G_1(C)$. We assert that X , as a subset of $E(G)$, has the property stated in the theorem.

Indeed, let $Y \subset E(v) \subset E(G)$ be an even set crossing X . In $G_1(C)$, Y can be extended to a good cycle F_1 . Let Y_2 be the set of edges of $F_1 \subset G_1(C)$ containing the vertex v_2 of $G_1(C)$. (Recall that this is the vertex representing a contracted component of $G - C$.) The size of Y_2 is either 0 or 2.

If $|Y_2| = 2$, we can use the minimality of G to extend Y_2 to a good cycle F_2 of $G_2(C)$. The cycles F_1 and F_2 agree on C . By Lemma 4.2, $F_1 \cup F_2$ is a good cycle of G .

In the remaining case, Y_2 is empty (and, necessarily, $|C| = 2$). Let z be a neighbor of v_1 in $G_2(C)$. By Claim 1, the degree of z in $G_2(C)$ is at least 3. Let Y'_2 be a set of

two edges incident with z but not with v_2 . (Note that v_2 is not incident with a pair of parallel edges by the 2-connectedness of G .) If we extend Y'_2 to a good cycle F_2 in $G_2(C)$, then F_1 and F_2 may be combined as above to a good cycle in G .

CLAIM 4. *Every cyclic 4-cut is interlaced with some other cyclic 4-cut.*

Suppose that the claim is false. We may thus assume that a cyclic 4-cut $C = \{a, b, c, d\}$ is not interlaced with any other cyclic 4-cut. It follows that every cyclic 4-cut corresponds to a 4-cut in either $G_1(C)$ or $G_2(C)$. Clearly, a trivial cut in G corresponds to a cut of the same size in G_1 or G_2 . Since, by Claim 3, G has no nontrivial 3-cuts, we conclude the following:

(*) *Each 3-cut in G corresponds to a 3-cut in $G_1(C)$ or $G_2(C)$, and each cyclic 4-cut in G corresponds to a 4-cut in $G_1(C)$ or $G_2(C)$.*

Note that $G_1(C)$ and $G_2(C)$ are bridgeless and v_2 and v_1 are vertices of degree 4.

We may assume that $v \in V(G_1)$. Let $E_i = E(G_i)$ ($i = 1, 2$). By the minimality of G , there is a 2-set $X_2 \subset E_2(v_1)$ such that any even subset of $E_2(v_1)$ crossing X_2 can be extended to a good cycle of G_2 .

Let $G_1^* = G_1(v_2, X_2, C \setminus X_2)$ be the result of the splitting operation defined at the beginning of this section. Thus, v_2 is split into two adjacent vertices. Let us write v_2^1 for the vertex incident with the edges in X_2 and v_2^2 for the other vertex. Furthermore, we let E_1^* denote the set $E(G_1^*)$. Note that G_1^* has fewer vertices than G . By the minimality of G , G_1^* contains a 2-set $X_1 \subset E_1^*(v)$ such that any even set $Y_1 \subset E_1^*(v)$ crossing X_1 can be extended to a good cycle of G_1 . We claim that $X = X_1$ satisfies the analogous condition for G in place of G_1 . Let $Y \subset E(v)$ be an even set. Extend the corresponding set of edges of G_1^* to a good cycle F_1^* of G_1^* . This determines a good cycle F_1 of G_1 . Consider the set Y_2 of edges of $E_2(v_1)$ corresponding to the edges in $F_1 \cap C$.

Due to the way we split v_2 , the even set Y_2 crosses X_2 , for otherwise F_1^* would not pass through either v_2^1 or v_2^2 . By the choice of X_2 , Y_2 can be extended to a good cycle F_2 of G_2 . The good cycles F_1 and F_2 agree on C . By (*), the cycle $F_1 \cup F_2$ intersects all 3-cuts and cyclic 4-cuts of G .

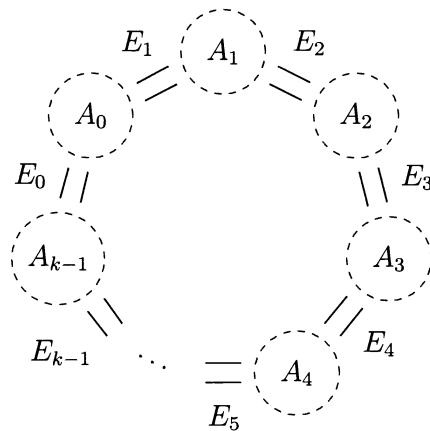
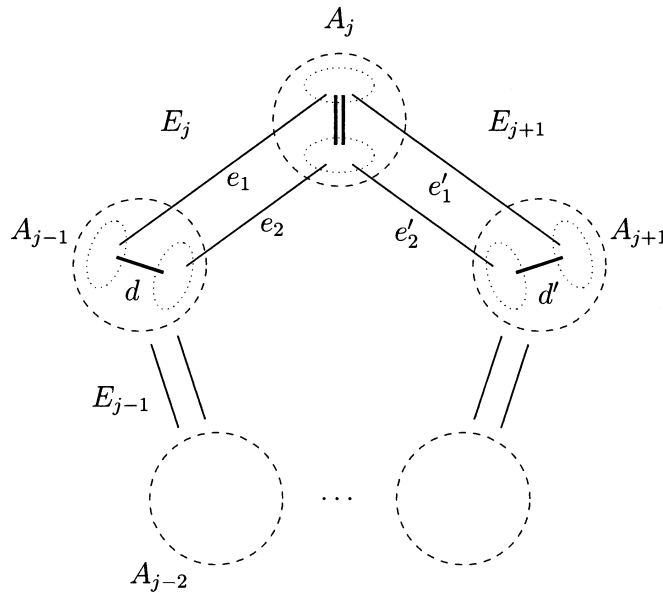
We claim that if C' is an acyclic 4-cut, then it is also intersected by $F_1 \cup F_2$. This is clear if C' is not interlaced with C . In the opposite case, Proposition 3.7 (b) implies there is a 3-cut C'' in G such that C' intersects C'' in two edges. Since $F_1 \cup F_2$ intersects C'' , it must also intersect C' . We conclude that $F_1 \cup F_2$ is a good cycle of G .

CLAIM 5. *The graph G is cyclically 5-edge-connected.*

Assume, to the contrary, that G contains a cyclic 4-cut C . By Claim 4, C is interlaced with a cyclic 4-cut C' . By Proposition 3.7 (c), each component of $G - C_1 - C_2$ is incident with exactly two edges from C_1 and two edges from C_2 ; moreover, these edges clearly form a 4-cut in G (see Figure 4 (b)). We may thus define $\mathcal{E} = (E_0, E_1, \dots, E_{k-1})$ to be an inclusion-maximal collection of pairwise disjoint 2-subsets of $E(G)$ with the following properties:

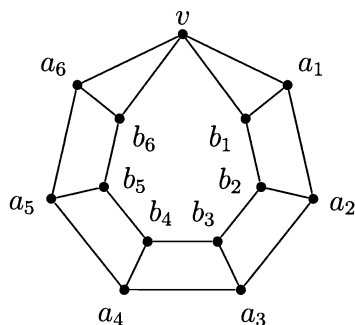
- (1) $C = E_r \cup E_s$ and $C' = E_{r'} \cup E_{s'}$ for some $r, s, r', s' \in \{0, 1, \dots, k-1\}$,
- (2) $E_i \cup E_j$ is a 4-cut for all distinct $i, j \in \{0, 1, \dots, k-1\}$.

Notice that $k \geq 4$ (since r, s, r' , and s' are all distinct). Let E^* be the union of all the sets E_i . We may assume that $E_0, E_1, \dots, E_{k-1}$ are enumerated in such a way that for each $j \in \{0, 1, \dots, k-1\}$, the graph $G - (E_j \cup E_{j+1})$ has a component A_j containing no edge from E^* (the indices being taken modulo k throughout this proof). Clearly, the structure of G with respect to the sets E_i is “cyclic” as shown in Figure 6.

FIG. 6. The structure of G with respect to the sets E_i .FIG. 7. The cut C^* in the proof of Claim 5 (shown in bold lines).

We now prove that for each j , $E_j \cup E_{j+1}$ is an acyclic 4-cut. Assume the opposite. By Claim 4, $E_j \cup E_{j+1}$ is interlaced with a cyclic 4-cut C^* . Let B be a component of $G - (E_j \cup E_{j+1}) - C^*$ contained in A_j . Proposition 3.7 (c) implies that B is incident with two edges (e_1 and e'_1) of $E_j \cup E_{j+1}$ and two edges (c and c') of C^* . If $E_j = \{e_1, e'_1\}$, then $\{c, c'\} \cup E_j$ and $\{c, c'\} \cup E_{j+1}$ are 4-cuts and, more generally, the collection $(\{c, c'\}, E_0, \dots, E_{k-1})$ satisfies (1) and (2) above. This leads to a contradiction with the maximality of k . Since a symmetric argument applies if $E_{j+1} = \{e_1, e'_1\}$, we conclude that $e_1 \in E_j$ and $e'_1 \in E_{j+1}$. Let e_2 be the edge of E_j distinct from e_1 , and let e'_2 be the edge of E_{j+1} distinct from e'_1 . The situation is illustrated in Figure 7.

Since e_1 and e_2 are in different components of $G - C^*$, their endvertices in A_{j-1} must be in different components of $A_{j-1} - C^*$. Thus, A_{j-1} contains an edge d of C^* .

FIG. 8. The prism $C_7 \square K_2$ with a contracted edge.

By a similar argument, A_{j+1} contains an edge d' of C^* . Since $|C^*| = 4$, each of A_{j-1} and A_{j+1} contain exactly one edge of C^* , and each A_i with $i \notin \{j-1, j, j+1\}$ is disjoint from C^* . Note that there is at least one such set A_i , since we observed above that $k \geq 4$.

Each component of $A_{j-1} - d$ is incident with an edge of E_{j-1} , for otherwise $\{e_1, d\}$ or $\{e_2, d\}$ would be a 2-cut in G . But A_{j-2} contains no edge of C^* , and hence all of $A_{j-1} - d$ is contained in a single component of $G - C^*$. By the same token, this component also contains $A_{j+1} - d'$. It follows that C^* does not separate G , which is a contradiction. We have shown that each $E_j \cup E_{j+1}$ is an acyclic 4-cut as claimed.

This implies that each A_j is either a single vertex or a copy of K_2 . If A_j is a single vertex, then this vertex is v and it is adjacent to each of the four vertices of $A_{j-1} \cup A_j$. In the other case, each of the two vertices $\{a_j, a'_j\}$ of A_j is adjacent to a vertex from A_{j-1} and a vertex from A_{j+1} . Finally, a_j and a'_j have no common neighbor except possibly for v . Thus, we have completely determined the structure of G . It can be described as follows. Let B_k be the k -prism (the Cartesian product of a cycle C_k with K_2). There are two copies of C_k in B_k ; let $a_0, \dots, a_{k-1}$ be the vertices of one copy and $b_0, \dots, b_{k-1}$ be the vertices of the other copy, with a_i adjacent to b_i in B_k for each i . By the above, G is (up to isomorphism) the graph obtained from B_k by contracting the edge $a_0 b_0$ (see Figure 8).

Let us determine which even subsets of $E(v)$ can be extended to a good cycle. Let $Y \subseteq E(v)$ be an even subset. First, if $Y = E(v)$, then the cycle consisting of all the edges $a_i a_{i+1}$ and $b_i b_{i+1}$ (indices modulo k) is good and extends Y . A good cycle extending $\{va_1, vb_1\}$ is $va_1 a_2 \dots a_{k-1} b_{k-1} b_{k-2} \dots b_1 v$, while a good cycle extending $\{va_{k-1}, vb_{k-1}\}$ is obtained in a symmetric manner. The remaining cases depend on the parity of k .

Assume that k is odd and $Y = \{va_1, va_{k-1}\}$. A good cycle extending Y is $va_1 b_1 b_2 a_2 \dots b_{k-1} a_{k-1} v$. The case $Y = \{vb_1, vb_{k-1}\}$ is symmetric. Thus, if we define $X = \{va_1, vb_{k-1}\}$, then any even subset of $E(v)$ crossing X can be extended to a good cycle.

If k is even, then $Y = \{va_1, vb_{k-1}\}$ extends to the good cycle $va_1 b_1 b_2 a_2 \dots a_{k-1} b_{k-1} v$, and the case $Y = \{vb_1, va_{k-1}\}$ is symmetric again. We put $X = \{vb_1, vb_{k-1}\}$ and observe that, as above, every even subset of $E(v)$ that crosses X extends to a good cycle. This contradiction with the assumption that G is a counterexample establishes Claim 5.

By Claim 5, each cut C of size 3 or 4 in G is acyclic. Thus, one component of $G - C$ is a copy of K_2 (if $|C| = 4$) or a single vertex (if $|C| = 3$).

Let $E(v) = \{a, b, c, d\}$. If every even subset Y' of $E(v)$ can be extended to a good cycle of G , then we are done. Therefore, suppose that for some Y' , this is not the case.

If $Y' = E(v)$, then consider the cubic bridgeless graph $G(v, \{a, b\}, \{c, d\})$. By Theorem 4.1, $\{a, b\}$ can be extended to a 2-factor F of $G(v, \{a, b\}, \{c, d\})$. Note that F contains c and d . Thus, F corresponds to a good cycle of G that contains all the edges of Y' , a contradiction.

We may therefore assume that Y' is a 2-subset of $E(v)$, say $Y' = \{a, b\}$, and define $X = Y'$. Let Y be an even subset of $E(v)$ which crosses X . Since we have already observed that $E(v)$ extends to a good cycle of G , it may be assumed that $|Y| = 2$, say (without loss) $Y = \{a, d\}$. Consider the graph $G(v, \{a, c\}, \{b, d\})$. Let v_1, v_2 be the adjacent vertices into which v is split. Using Theorem 4.1, extend $\{a, v_1v_2\}$ to a 2-factor F of $G(v, \{a, c\}, \{b, d\})$. Notice that F contains precisely one of the edges b, d . It cannot be b , for otherwise we obtain a good cycle of G with $F(v) = X$, and this contradicts the choice of X . Thus, F must contain d . Since $F(v) = \{a, d\}$ and F is a good cycle of G , we obtain the right extension of Y , and we have established the theorem. $\square$

COROLLARY 4.4. *Every graph with maximum degree at most 3 has a $\{3, 4\}$ -covering cycle.*

Theorem 4.3 implies the following strengthening of Theorem 4.1.

COROLLARY 4.5. *Every cubic bridgeless graph has a 2-factor which intersects all cuts of size 3 and 4. Moreover, any two incident edges can be extended to such a 2-factor.*

5. An extension to arbitrary graphs. In this section, we generalize Corollary 4.4 to graphs with unrestricted degrees which, moreover, do not need to be bridgeless. We first prove a lemma that can be used to reduce coverability questions to the case of cubic graphs.

Let us call a class $\mathcal{Q}$ of graphs *splittable* if every graph $G \in \mathcal{Q}$ of maximum degree $\Delta(G) \geq 4$ contains a vertex v such that for some partition (E_1, E_2) of the set of edges incident with v , it holds that $G(v, E_1, E_2) \in \mathcal{Q}$ and $|E_1|, |E_2| \geq 2$. (Recall that the graph $G(v, E_1, E_2)$ is defined at the beginning of section 4.) As usual, the class $\mathcal{Q}$ is called *minor-closed* if every minor of a graph in $\mathcal{Q}$ is in $\mathcal{Q}$.

LEMMA 5.1. *Let $\mathcal{Q}$ be a splittable, minor-closed class of graphs and $A \subset \{3, 4, \dots\}$ a set of integers. Then A is coverable in $\mathcal{Q}$ if and only if A is coverable in the class of all cubic bridgeless graphs in $\mathcal{Q}$.*

Proof. The “only if” direction is clear. To prove the converse, suppose that a set A is coverable in the class of cubic bridgeless graphs in $\mathcal{Q}$ but not in $\mathcal{Q}$ itself. For a graph H , let $\Delta(H)$ denote the maximum degree of H and let $V_\Delta(H)$ be the set of vertices of degree $\Delta(H)$. Choose a graph $G \in \mathcal{Q}$ with no A -covering cycle and with the property that the triple $(\Delta(G), |V_\Delta(G)|, |V(G)|)$ is minimal in the lexicographic ordering. By our assumptions, G contains either a bridge or a vertex of degree different from 3. We distinguish three possible cases. In each case, we consider a suitable reduction of G to a “smaller” graph that preserves cuts whose cardinality is in A , which enables us to reach a contradiction.

Assume first that G contains a bridge e . Let G_1 and G_2 be the components of $G - e$. Since each G_i is in $\mathcal{Q}$, the induction hypothesis implies that it contains an A -covering cycle F_i . Furthermore, every cut in G except $\{e\}$ corresponds to a cut of the same size in G_1 or G_2 . Since $1 \notin A$, $F_1 \cup F_2$ is A -covering in G .

Next, assume that v is a vertex of degree 2 and let $E(v) = \{e_1, e_2\}$. Let G' be the graph obtained by suppressing v (and removing the resulting loop if e_1 and e_2 are parallel). Noting that $2 \notin A$, it is not difficult to see that A -covering cycles in G' (which exist by the induction hypothesis since G' is a minor of G) correspond to A -covering cycles in G .

Finally, let v be a vertex of degree at least 4. Using the assumption that $\mathcal{Q}$ is splittable, choose a partition of $E(v)$ into sets E_1 and E_2 with $|E_1|, |E_2| \geq 2$ such that $G^* = G(v, E_1, E_2)$ is in $\mathcal{Q}$. As above, the minimality of G implies that G^* contains an A -covering cycle F^* . Since every cut of G corresponds to a cut of the same size in G^* , F^* determines an A -covering cycle in G , a contradiction. $\square$

We can now derive Theorem 1.2, stated in section 1.

THEOREM 1.2. *Every graph G has a cycle intersecting all cuts of size 3 or 4.*

Proof. We apply Lemma 5.1 to the class $\mathcal{Q}$ of all graphs and $A = \{3, 4\}$. The result then follows from Corollary 4.4. $\square$

6. Concluding remarks. By Theorem 1.2, both the sets $\{3\}$ and $\{4\}$ are coverable. On the other hand, $\{1\}$ and $\{2\}$ are not. How about the other single-element sets?

QUESTION 6.1. *Is it true that for all $k \geq 3$, $\{k\}$ is coverable?*

We are unable to say anything for $k \geq 5$, except that Conjecture 1.7 clearly implies an affirmative answer to this question. On the other hand, since (as we noted in section 1) the conjecture is true for planar graphs, any set $\{k\}$ is coverable in the class of planar graphs (which will be denoted by $\mathcal{P}$ throughout this section).

Having determined which sets of size 1 are coverable in $\mathcal{P}$, we may attempt the same for sets of size 2. Let $A = \{a, b\}$ be a pair of positive integers with $a < b$. If $a \leq 2$, then A is not coverable in $\mathcal{P}$, and if $a \geq 4$, then the planar case of Conjecture 1.7 implies that A is coverable in $\mathcal{P}$. Thus, we may assume that $a = 3$. Since the set $2\mathbb{N} + 1 = \{3, 5, 7, \dots\}$ is coverable in $\mathcal{P}$ (by the Four Color Theorem), we may assume that b is even and $b \geq 6$.

QUESTION 6.2. *Let $k \geq 3$. Is $\{3, 2k\}$ coverable in $\mathcal{P}$?*

In fact, this is an equivalent form of a question posed by Broersma et al. [2] in connection with Theorem 1.1: For which $k \geq 3$ can one 2-color the vertices of every planar graph in such a way that there is no monochromatic circuit of length 3 or $2k$? To our knowledge, the question is open. It may even be that $\{3, 2k\}$ ($k \geq 3$) is coverable in the class of all graphs.

One might speculate that even the set consisting of 3 and all the numbers $2k$ ($k \geq 2$) is coverable. We show that this is not the case.

PROPOSITION 6.3. *The set $A = \{3, 4, 6, 8, 10, \dots\}$ is not coverable.*

Proof. Let G be a 3-connected, non-Hamiltonian, cubic bipartite graph. (Such graphs were first constructed by J. D. Horton; see [1, Exercise 4.2.14] or [6].) Let C be an A -covering cycle in G . As $3 \in A$, C is a 2-factor. Since C has more than one component, there exists a cut K contained in the complement of C . Let G_1 denote a component of $G \setminus K$. The size of K must be odd since $|K| \neq 2$ and C is A -covering. Thus, the number of vertices of G_1 is odd. However, C covers the vertices of G_1 by disjoint circuits of even length, a contradiction. $\square$

A computer-assisted investigation of a smallest known non-Hamiltonian bipartite cubic graph on 50 vertices constructed in [5] (see also [9]) shows that this graph in fact admits no $\{3, 4, 6, 8, 10\}$ -covering cycle. We are indebted to Ondřej Rucký for this information.

If we restrict our attention to the class $\mathcal{P}$, then the argument used in the proof of Proposition 6.3 no longer applies, for there is no known example of a 3-connected planar cubic bipartite graph which is not Hamiltonian. Indeed, a well-known conjecture of D. Barnette [7, section 2.12] states that there is no such graph. Thus, we conclude our paper with the following question.

QUESTION 6.4. *Is $\{3, 4, 6, 8, 10, \dots\}$ coverable in $\mathcal{P}$?*

Acknowledgment. We thank the anonymous referees for comments that helped improve the presentation of our results.

REFERENCES

- [1] J. A. BONDY AND U. S. R. MURTY, *Graph Theory with Applications*, Macmillan, New York, 1976.
- [2] H. BROERSMA, F. V. FOMIN, J. KRATOCHVÍL, AND G. J. WOEGINGER, *Planar graph coloring avoiding monochromatic subgraphs: Trees and paths make it difficult*, *Algorithmica*, 44 (2006), pp. 343–361.
- [3] H. FLEISCHNER, *Cycle decompositions, 2-coverings, removable cycles, and the four-color-disease*, in *Progress in Graph Theory*, J. A. Bondy and U. S. R. Murty, eds., Academic Press, New York, 1984, pp. 233–246.
- [4] H. FLEISCHNER AND B. JACKSON, *A note concerning some conjectures on cyclically 4-edge-connected 3-regular graphs*, in *Graph Theory in Memory of G. A. Dirac*, *Ann. Discrete Math.* 41, L. D. Andersen et al., eds., North-Holland, Amsterdam, 1989, pp. 171–178.
- [5] J. P. GEORGES, *Non-Hamiltonian bicubic graphs*, *J. Combin. Theory Ser. B*, 46 (1989), pp. 121–124.
- [6] J. D. HORTON, *On two-factors of bipartite regular graphs*, *Discrete Math.*, 41 (1982), pp. 35–41.
- [7] T. R. JENSEN AND B. TOFT, *Graph Coloring Problems*, J. Wiley & Sons, New York, 1995.
- [8] T. KAISER AND R. ŠKREKOVSKI, *Planar graph colorings without short monochromatic cycles*, *J. Graph Theory*, 46 (2004), pp. 25–38.
- [9] A. K. KELMANS, *Cubic bipartite cyclically 4-connected graphs with no Hamiltonian cycles*, *Uspekhi Mat. Nauk*, 43 (1988), no. 3 (261), pp. 181–182 (in Russian). *Russian Math. Surveys*, 43 (1988), pp. 205–206 (in English).
- [10] M. M. MATTHEWS AND D. P. SUMNER, *Hamiltonian results in $K_{1,3}$ -free graphs*, *J. Graph Theory*, 8 (1984), pp. 139–146.
- [11] T. SCHÖNBERGER, *Ein Beweis des Petersenschen Graphensatzes*, *Acta Litt. Sci. Szeged*, 7 (1934), pp. 51–57.
- [12] C. THOMASSEN, *Decomposing a planar graph into degenerate graphs*, *J. Combin. Theory Ser. B*, 65 (1995), pp. 305–314.
- [13] C. THOMASSEN, *Reflections on graph theory*, *J. Graph Theory*, 10 (1986), pp. 309–324.
- [14] W. T. TUTTE, *A theorem on planar graphs*, *Trans. Amer. Math. Soc.*, 82 (1956), pp. 99–116.
- [15] W. T. TUTTE, *On the algebraic theory of graph colourings*, *J. Combin. Theory*, 1 (1966), pp. 15–50.
- [16] C.-Q. ZHANG, *Integer Flows and Cycle Covers of Graphs*, Dekker, New York, 1997.

CHAINS IN THE NONCROSSING PARTITION LATTICE*

NATHAN READING†

Abstract. We establish recursions counting various classes of chains in the noncrossing partition lattice of a finite Coxeter group. The recursions specialize a general relation which is proven uniformly (i.e., without appealing to the classification of finite Coxeter groups) using basic facts about noncrossing partitions. We solve these recursions for each finite Coxeter group in the classification. Among other results, we obtain a simpler proof of a known uniform formula for the number of maximal chains of noncrossing partitions and a new uniform formula for the number of edges in the noncrossing partition lattice. All of our results extend to the m -divisible noncrossing partition lattice.

Key words. absolute order, Catalan number, chain enumeration, Coxeter element, Coxeter group, generalized cluster complex, noncrossing partition, parking function

AMS subject classifications. Primary, 20F55; Secondary, 05A18, 05E15

DOI. 10.1137/07069777X

1. Introduction. The lattice of noncrossing partitions was defined and studied in 1972 by Kreweras [25]. For surveys of results on this lattice and on its mathematical applications, see [26, 30]. Through the results of [5, 6, 11, 29], the noncrossing partition lattice was recognized as a special case of a construction valid for an arbitrary finite Coxeter group. The notation L_W will stand for the noncrossing partition lattice of a finite Coxeter group W . In particular, when W is the symmetric group, L_W is the usual noncrossing partition lattice. Detailed enumeration of chains in L_W for various Coxeter groups W has been carried out in [1, 4, 7, 14, 15, 24, 25, 29, 32].

The key result of this paper is a formula relating certain rank-selected chain numbers for L_W to chain enumerations arising in parabolic subgroups. For a sequence $(j_1, j_2, \dots, j_{k+1})$ summing to $n = \text{rank}(W)$, let $C_{(j_1, j_2, \dots, j_{k+1})}(W)$ count multichains $x_1 \leq x_2 \leq \dots \leq x_k$ in L_W with $\ell_T(x_1) = j_1$, $\ell_T(x_k) = n - j_{k+1}$ and $\ell_T(x_i) = \ell_T(x_{i-1}) + j_i$ for $i = 2, \dots, k$. Here ℓ_T is the rank function of L_W . For each simple reflection $s \in S$, let $W_{\langle s \rangle}$ denote the parabolic subgroup generated by $S \setminus \{s\}$. The Coxeter number h is the order of a Coxeter element of W . The sequence $(j_1, j_2, \dots, \hat{j}_i, \dots, j_{k+1})$ is obtained by deleting j_i from $(j_1, j_2, \dots, j_{k+1})$.

THEOREM 1.1. *If (W, S) is a finite irreducible Coxeter system and $j_i = 1$, then*

$$C_{(j_1, j_2, \dots, j_{k+1})}(W) = \frac{h}{2} \sum_{s \in S} C_{(j_1, j_2, \dots, \hat{j}_i, \dots, j_{k+1})}(W_{\langle s \rangle}).$$

The theorem is proved uniformly in section 2 by a method similar to that used by Fomin and Zelevinsky to prove a recursive formula [18, Proposition 3.7] counting the facets of the cluster complex. In that context, one “rotates” a root by a modified Coxeter element (of order $h + 2$) until one obtains the negative of a simple root. This allows one to pass to a parabolic subgroup. Here, we rotate a reflection by an

*Received by the editors July 19, 2007; accepted for publication (in revised form) February 8, 2008; published electronically May 14, 2008. The author was partially supported by NSF grant DMS-0202430.

<http://www.siam.org/journals/sidma/22-3/69777.html>

†Department of Mathematics, North Carolina State University, Raleigh, NC 27695 (nathan.reading@ncsu.edu).

unmodified Coxeter element (of order h) until we obtain a simple reflection, which allows us to pass to a parabolic subgroup.

Theorem 1.1 is a broad generalization of what appears to be the only nontrivial enumerative fact about L_W previously known to have a uniform proof: the formula $nh/2$ for the number of atoms (or coatoms) of L_W . There are, however, uniform bijections to other sets, namely clusters [2, 23, 27] and sortable elements [23, 27], but no uniform proof is known for the enumeration of these other sets. (The results of [23] apply only to the crystallographic case. The proofs in [27] are made uniform by the results of [28].) There is also a uniform determination [3, Corollary 4.4] of the Möbius function of L_W in terms of positive clusters, but no uniform proof is known for the enumeration of positive clusters.

In section 3, we specialize Theorem 1.1 to provide recursions for some important classes of chains. In some cases, the recursion leads to a uniform formula. However, even in those cases, deriving the formula from the recursion requires a type-by-type approach. We now briefly summarize the results obtained.

We first consider the number $\text{MC}(W)$ of maximal chains in L_W . We obtain a recursion for $\text{MC}(W)$ and, by solving the recursion type-by-type, a uniform formula for $\text{MC}(W)$. As pointed out in [14], this uniform formula follows from previous type-by-type determinations of the zeta polynomial of L_W . In the exceptional types, the recursion on $\text{MC}(W)$ can be solved without a computer, thus providing the first verification of the formula for $\text{MC}(W)$ without brute-force computer counting.

We next give a recursion on the number of reduced words (in the alphabet of reflections) for elements of L_W . The recursion implies a relationship between the number of such words and the face numbers of the generalized cluster complexes of [17]. Another specialization of Theorem 1.1 leads to a uniform formula, which appears to be new, for the number of edges in L_W . More generally, we consider saturated chains in L_W of a fixed length. The number of such chains appears to exhibit the same odd behavior observed in [17] for the f - and h -numbers of generalized cluster complexes.

Theorem 1.1 and its corollaries are statements about h -fold symmetry. It is intriguing that one of the corollaries can be obtained, by taking leading coefficients, from a similar statement about the $(mh + 2)$ -fold symmetry of a generalized cluster complex, as explained in section 3.

We conclude the paper with a brief discussion, in section 4, of generalizations to m -divisible noncrossing partitions.

2. Proof of the main theorem. In this section, we define L_W and gather the simple facts about Coxeter groups and noncrossing partitions that are necessary to prove Theorem 1.1. We then prove the theorem and comment on the case where W is reducible. Much of the background material on L_W is due to Armstrong [1], Bessis [5], and Brady and Watt [12, 13].

We assume basic background on Coxeter groups and root systems, which is found, for example, in [8, 9, 22]. Let (W, S) be a finite Coxeter system of rank n . We fix a representation of W as a real reflection group acting on a Euclidean space V and make no distinction between an element of W and its action on V .

Let T be the set of reflections in W . For any reflection $t \in T$, let H_t be the reflecting hyperplane associated to t . Any element $w \in W$ can be written as a T -word—a word in the alphabet T . A *reduced* T -word for w is a T -word which has minimal length among all T -words for w . The *absolute length* of w is the length of a reduced T -word for w . This should not be confused with the more common notion

of length in W : the length of a reduced word for w in the alphabet S . The *absolute order* on W sets $u \leq v$ if and only if u has a reduced T -word that is a prefix of some reduced T -word for v . Note that since T is fixed as a set by conjugation, for any $u \leq v$ and any $w \in W$, the interval $[u, v]$ in the absolute order is isomorphic to the interval $[wuw^{-1}, wvw^{-1}]$.

Suppose $t_1 t_2 \cdots t_k$ is a reduced T -word for w . Then, for any $i \in [k-1]$, another reduced T -word for w can be obtained by replacing t_i with t_{i+1} and t_{i+1} with the reflection $(t_{i+1} t_i t_{i+1})$, while leaving all other letters of the word unchanged. Similarly, for $i \in [2, k]$, one can replace t_i with t_{i-1} and t_{i-1} with $(t_{i-1} t_i t_{i-1})$. This implies that $u \leq v$ if and only if u has a reduced T -word that is a subword of some reduced T -word for v . Furthermore, $u \leq v$ if and only if u has a reduced T -word that is a postfix of some reduced T -word for v .

A Coxeter element c is any element of W of the form $c = s_1 s_2 \cdots s_n$ with each element of S occurring exactly once. The order of c is the *Coxeter number* h . The primary object of study in this paper is the interval $[1, c]$ in the absolute order, often called the noncrossing partition lattice, and denoted here by L_W . Any two Coxeter elements are conjugate in W , so the isomorphism type of L_W does not depend on the choice of Coxeter element c . The fact that L_W is a lattice was given a uniform proof in [13] and later, with slightly less generality in [23].

The Coxeter diagram of W is a tree and thus a bipartite graph. Let $S = S_+ \cup S_-$ be a bipartition of the diagram. Define involutions

$$c_+ = \prod_{s \in S_+} s \quad \text{and} \quad c_- = \prod_{s \in S_-} s$$

so that $c = c_- c_+$ is a Coxeter element with $c_+ c c_+ = c_- c c_- = c_+ c_- = c^{-1}$. Thus conjugation by c_+ and conjugation by c_- are isomorphisms from $[1, c]$ to $[1, c^{-1}]$. Let $c_+^{(d)}$ be the $|d|$ -fold product $c_{(-1)^d} \cdots c_+ c_- c_+$ if $d \geq 0$ or $c_+ c_- c_+ \cdots c_{(-1)^d}$ if $d < 0$. We have $c_+^{(d)} c_+^{(-d)} = 1$ for any $d \in \mathbb{Z}$ and, furthermore, $c_+^{(2h)} = c^h = 1$.

The key to the proof of Theorem 1.1 is a result from Steinberg's 1959 paper [34] on finite reflection groups. (See also [22, sections 3.16–3.20] or [9, section V.6.2].)

PROPOSITION 2.1 (see Steinberg [34]). *Let (W, S) be an irreducible finite Coxeter system. The orbit of any reflection under the conjugation action of the dihedral group $\langle c_+, c_- \rangle$ either:*

- (i) *has $h/2$ elements and intersects S in a single element, or*
- (ii) *has h elements and intersects S in a two-element set.*

The proof of Theorem 1.1 uses some fundamental facts about absolute order which we now quote as Theorem 2.2. A clean seven-page exposition (with complete proofs) of these results can be obtained by reading Brady and Watt's paper [12] followed by section 2 of their paper [11]. We phrase these properties in terms of fixed spaces $F_w = \ker(w - I)$ of elements $w \in W$, rather than the *moved spaces* of [12]. This change is harmless because the moved space is the orthogonal complement of F_w . Note that $H_t = F_t$ for each $t \in T$.

THEOREM 2.2 (see Brady and Watt [11, 12]).

- (i) *If $t_1 t_2 \cdots t_k$ is a reduced T -word for $w \in W$, then $F_w = H_{t_1} \cap H_{t_2} \cap \cdots \cap H_{t_k}$.*
- (ii) *If $x, y \in [1, c]$, then $x \leq y$ if and only if $F_y \subseteq F_x$.*
- (iii) *If $t \in T$, then $t \leq c$.*

One useful consequence of Theorem 2.2 is the following lemma (cf. [11, Lemma 2.3]). Recall that $W_{\langle s \rangle}$ is the parabolic subgroup generated by $S \setminus \{s\}$, and let c' be the Coxeter element for $W_{\langle s \rangle}$ obtained by deleting s from the defining word for c . The

subword characterization of absolute order implies that $c' < c$, so $[1, c'] \subset [1, c]$. We write $[u, v]_{\langle s \rangle}$ for an interval in the absolute order on $W_{\langle s \rangle}$ and continue to write $[u, v]$ for an interval in the absolute order on W .

LEMMA 2.3. *The inclusion $W_{\langle s \rangle} \hookrightarrow W$ restricts to an isomorphism from $[1, c']_{\langle s \rangle}$ to $[1, c']$.*

Proof. Theorem 2.2(ii) implies that the inclusion $W_{\langle s \rangle} \hookrightarrow W$ restricts to an isomorphism from $[1, c']_{\langle s \rangle}$ to its image and, furthermore, that this image is contained in $[1, c']$. The parabolic subgroup $W_{\langle s \rangle}$ is the set of elements of W which fix the subspace $\cap_{s' \in \langle s \rangle} H_{s'}$. But $\cap_{s' \in \langle s \rangle} H_{s'} = F_{c'}$ by Theorem 2.2(i), so Theorem 2.2(ii) implies that $[1, c'] \subseteq W_{\langle s \rangle}$. Thus the inclusion of $[1, c']_{\langle s \rangle}$ in W is the entire interval $[1, c']$. $\square$

To simplify the proof of Theorem 1.1, we employ a basic result about counting multichains in L_W . (Cf. [1, Lemma 3.1.2].)

PROPOSITION 2.4. *If $(j'_1, \dots, j'_{k+1})$ is a permutation of $(j_1, \dots, j_{k+1})$, then*

$$C_{(j'_1, \dots, j'_{k+1})}(W) = C_{(j_1, \dots, j_{k+1})}(W).$$

Proof. It is sufficient to prove the proposition in the case where the two sequences agree except that $j'_i = j_{i+1}$ and $j'_{i+1} = j_i$ for some i . Setting $x_0 = 1$ and $x_{k+1} = c$, a multichain $x_1 \leq x_2 \leq \dots \leq x_k$ is uniquely encoded by the sequence

$$(\delta_0, \dots, \delta_k) = (x_0^{-1}x_1, x_1^{-1}x_2, \dots, x_k^{-1}x_{k+1}).$$

In [1], this is called the *delta sequence* of $x_1 \leq x_2 \leq \dots \leq x_k$. A sequence of elements of L_W is a delta sequence for some multichain in L_W if and only if the absolute lengths of the elements of the sequence sum to $n = \text{rank}(W)$ and the product, in order, of the sequence is c . A multichain is counted by $C_{(j_1, \dots, j_{k+1})}(W)$ if and only if its delta sequence has $\ell_T(\delta_{i-1}) = j_i$ for all i . Given a delta sequence $\delta = (\delta_0, \dots, \delta_k)$ with this property, define a new sequence $\delta' = (\delta'_0, \dots, \delta'_k)$ agreeing with δ except that $\delta'_{i-1} = \delta_{i-1}\delta_i\delta_{i-1}^{-1}$ and $\delta'_i = \delta_{i-1}$. It is immediate that δ' is the delta sequence for a multichain counted by $C_{(j'_1, \dots, j'_{k+1})}(W)$ and, furthermore, that the map $\delta \mapsto \delta'$ defines a bijection between the two sets of chains. $\square$

We now prove the main theorem.

Proof of Theorem 1.1. We continue to fix a particular Coxeter element $c = c_-c_+$. Proposition 2.4 implies that it is enough to consider the case where $i = k+1$. A multichain counted by $C_{(j_1, \dots, j_{k+1})}(W)$ consists of an element x covered by c and a multichain in $[1, x]$ with rank-differences given by $(j_1, \dots, j_k)$. In light of Theorem 2.2(iii) and the prefix/postfix characterization of absolute order, an element x is covered by c if and only if $x = ct$ for some $t \in T$.

For each $t \in T$, consider the orbit of $[1, ct]$ under the conjugation action of $\langle c_+, c_- \rangle$. Since conjugation by c_- is an involutive isomorphism from $[1, c]$ to $[1, c^{-1}]$, all intervals in this orbit are isomorphic. In light of Proposition 2.1, we can define $d = d(t)$ to be the smallest $d \geq 0$ such that $c_+^{(d)}tc_+^{(-d)} = c_+^{(d+1)}tc_+^{(-d-1)} = s$ for some $s \in S$. (Cf. the proof of [19, Lemma 4.1].) Let $s = s(t) = c_+^{(d)}tc_+^{(-d)}$. If d is even, then $c_+sc_+ = s$, or in other words, $s \in S_+$. Thus in this case

$$c_+^{(d)}(ct)c_+^{(-d)} = cs = \prod_{s' \in S_-} s' \prod_{s'' \in (S_+ \setminus \{s\})} s''.$$

If d is odd, then $s \in S_-$ and

$$c_+^{(d)}(ct)c_+^{(-d)} = c^{-1}s = \prod_{s' \in S_+} s' \prod_{s'' \in (S_- \setminus \{s\})} s''.$$

In either case $[1, ct]$ is isomorphic to $[1, c']$, where c' is some Coxeter element for $W_{\langle s \rangle}$.

By Lemma 2.3, we conclude that $C_{(j_1, \dots, j_k, 1)}(W)$ counts triples (t, s, μ) such that $t \in T$, $s = s(t)$ and μ is a multichain in $L_{W_{\langle s \rangle}}$ with rank-differences given by $(j_1, \dots, j_k)$. Alternately, we can count such pairs by first specifying s and applying Proposition 2.1. Each s belonging to an orbit of size $\frac{h}{2}$ contributes $\frac{h}{2} C_{(j_1, \dots, j_k)}(W_{\langle s \rangle})$ to the count. For each pair s, s' belonging to an orbit of size h , the lattices $L_{W_{\langle s \rangle}}$ and $L_{W_{\langle s' \rangle}}$ are isomorphic because they are isomorphic to conjugate intervals in L_W . Together s and s' contribute $h \cdot C_{(j_1, \dots, j_k)}(W_{\langle s \rangle}) = h \cdot C_{(j_1, \dots, j_k)}(W_{\langle s' \rangle})$ to the count. $\square$

In general, the parabolic subgroups $W_{\langle s \rangle}$ appearing in Theorem 1.1 are not irreducible. When W is reducible, L_W is a direct product. Thus by basic chain-counting techniques we have the following proposition.

PROPOSITION 2.5. *If $W = W_1 \times W_2$ with $\text{rank}(W_1) = n_1$, then*

$$C_{(j_1, \dots, j_{k+1})} = \sum C_{(j'_1, \dots, j'_{k+1})}(W_1) \cdot C_{(j_1 - j'_1, \dots, j_{k+1} - j'_{k+1})}(W_2),$$

where the sum is over all sequences $(j'_1, \dots, j'_{k+1})$ summing to n_1 and having $j'_i \leq j_i$ for all $i \in [k+1]$.

Remark 2.6. Stembridge [33] pointed out that Theorem 1.1 and Proposition 2.5 can be replaced by a single formula. Let W be a finite Coxeter group, not necessarily irreducible. For each $s \in S$, let h_s denote the Coxeter number of the irreducible component of W containing s . One factors c_+ as $(c_+)_1(c_+)_2 \cdots (c_+)_l$, where $(c_+)_i$ is in the i th irreducible component of W , and similarly for c_- . For each $s \in S$, let $(c_+)_s$ be $(c_+)_i$ if s is in the i th irreducible component of W , and similarly $(c_-)_s$. Replacing $c_{\pm}$ and h by $(c_{\pm})_s$ and h_s in the proof of Theorem 1.1, we obtain

$$C_{(j_1, j_2, \dots, j_{k+1})}(W) = \frac{1}{2} \sum_{s \in S} h_s \cdot C_{(j_1, j_2, \dots, \hat{j}_i, \dots, j_{k+1})}(W_{\langle s \rangle}).$$

Remark 2.7. We have seen that Proposition 2.1 describes a fundamental symmetry of L_W . In fact, the defining symmetry of cluster complexes also ultimately rests on Proposition 2.1. Specifically, [18, Theorem 2.6], which establishes the dihedral symmetry of the cluster complex, is a corollary of [18, Proposition 2.5], which in turn uses [18, Lemma 2.1], cited to [9, Exercise V.6.2]. Proposition 2.1 is not stated explicitly in [9], but is the key to the results which can be applied to solve [9, Exercise V.6.2]. And in fact, [18, Theorem 2.6] is an easy corollary of Proposition 2.1.

Remark 2.8. The concepts involved in the proof of Theorem 1.1 shed light on another similarity between noncrossing partitions and clusters. The definition of the cluster complex rests on a “compatibility” relation on certain roots. A negative simple root $-\alpha$ is compatible with a root β if and only if β belongs to the parabolic root subsystem obtained by deleting α . The rest of the compatibility relation is defined by requiring that compatibility be invariant under the dihedral action of $\langle \tau_+, \tau_- \rangle$, where $\tau_{\pm}$ is a modification of $c_{\pm}$. A similar approach can be made to noncrossing partitions. For $s \in S_-$ and $t \in T$, we have $st \in [1, c]$ if and only if $t \in W_{\langle s \rangle}$. When $s \in S_+$, we have $ts \in [1, c]$ if and only if $t \in W_{\langle s \rangle}$. This observation, together with the fact that the conjugation action of $\langle c_+, c_- \rangle$ acts by automorphisms, completely determines L_W .

3. Applications of the main theorem.

Maximal chains. The maximal chains in L_W are of particular interest for several reasons: For any finite Coxeter group, the maximal chains in L_W index the maximal faces in a CW-complex that is an Eilenberg–MacLane space (or “ $K(\pi, 1)$ ”) for the associated Artin group [10, 11]. Furthermore, maximal chains of classical noncrossing partitions are in bijection with parking functions [32]. In fact, the combinatorics of parking functions encodes rank-selected chain enumeration in the classical case [32, Proposition 3.2]. Generalizations to the case $W = B_n$ have been studied in [7, 21]. The number of maximal chains of classical noncrossing partitions also coincides with the dimension of the ring of diagonal coinvariants. (See, for example, [20, Theorem 4.2.4].)

Let $\text{MC}(W)$ denote the number of maximal chains in L_W . Since $\text{MC}(W) = C_{(1,1,\dots,1)}(W)$, Theorem 1.1 has the following corollary.

COROLLARY 3.1. *If (W, S) is a finite irreducible Coxeter system, then*

$$\text{MC}(W) = \frac{h}{2} \sum_{s \in S} \text{MC}(W_{\langle s \rangle}).$$

Proposition 2.5 becomes much simpler in this special case.

PROPOSITION 3.2. *If $W = W_1 \times W_2$ with $\text{rank}(W_1) = n_1$ and $\text{rank}(W_2) = n_2$, then*

$$\text{MC}(W) = \text{MC}(W_1) \text{MC}(W_2) \binom{n_1 + n_2}{n_1}.$$

The recursions in Corollary 3.1 and Proposition 3.2 can be solved to give formulas or values for each finite Coxeter group. The results are tabulated below, followed by examples illustrating how they were obtained.

A_n	B_n	D_n	E_6	E_7	E_8	F_4	H_3	H_4	$I_2(m)$
$(n+1)^{n-1}$	n^n	$2(n-1)^n$	41472	1062882	37968750	432	50	1350	m

Example 3.3. There is one maximal chain in L_W when W has rank zero. For $W = A_1$, since $h = 2$ we have $\text{MC}(A_1) = \frac{2}{2} \cdot 1 = 1$. For $W = I_2(m)$ we have $h = m$, so $\text{MC}(I_2(m)) = \frac{m}{2}(1+1) = m$.

Example 3.4. For $W = H_3$, $h = 10$ and the maximal parabolic subgroups are $I_2(5)$, $A_1 \times A_1$, and A_2 . Corollary 3.1 and Proposition 3.2 say that

$$\text{MC}(H_3) = \frac{10}{2} \left(5 + 1 \cdot 1 \cdot \binom{2}{1} + 3 \right) = 50.$$

Example 3.5. In each classical case, the formula for $\text{MC}(W)$ is proved by induction, applying Abel’s identity (see [16]). For $W = A_n$, the inductive step is

$$\sum_{i=0}^{n-1} \binom{n-1}{i} (i+1)^{i-1} (n-i)^{n-i-2} = 2(n+1)^{n-2}.$$

This is proved by rewriting the binomial coefficient in the left side as a sum of two binomial coefficients, splitting into two sums, and reversing the order of summation in one of the sums. The two summations are then identical and, by Abel’s identity, each equals $(n+1)^{n-2}$. For $W = B_n$, the inductive step is

$$\sum_{i=0}^{n-1} \binom{n-1}{i} i^i (n-i)^{n-i-2} = n^{n-1},$$

which is proved by reversing the order of summation and applying Abel's identity. For $W = D_n$, the inductive step is

$$\sum_{i=2}^{n-1} \binom{n-1}{i} (i-1)^i (n-i)^{n-i-2} = (n-1)^{n-1} - n^{n-2}.$$

This is proved by evaluating the left side from $i = 0$ to $n - 1$, reversing the order of summation, applying Abel's identity, and then subtracting off the $i = 0$ term.

The results tabulated above constitute a proof, without brute-force computer counting, of a uniform formula for $\text{MC}(W)$ pointed out in [14, Proposition 9].

THEOREM 3.6. *If W is a finite irreducible¹ Coxeter group, then*

$$\text{MC}(W) = \frac{n! h^n}{|W|}.$$

As explained in [14], Theorem 3.6 follows from a more general fact that has been verified type-by-type: For irreducible W , the zeta polynomial of L_W is the Fuss–Catalan number $\text{Cat}^{(m)}(W)$:

$$(3.1) \quad Z(L_W, m+1) = \text{Cat}^{(m)}(W) = \prod_{i=1}^n \frac{mh + e_i + 1}{e_i + 1}.$$

The numbers e_i are fundamental numerical invariants called the exponents of W . The zeta polynomial $Z(P, q)$ of a poset P counts, for each q , multichains $p_1 \leq p_2 \leq \dots \leq p_{q-1}$ in P . See [31, section 3.11] for details on zeta polynomials. By [31, Proposition 3.11.1], the leading term of $Z(P, q)$ is $(Mq^d)/d!$, where d is the length of the longest chain in P and M is the number of chains of length d . Maximal chains in L_W have length $n = \text{rank}(W)$, so the theorem follows by taking the coefficient of m^n in (3.1) and applying the fact that $|W| = \prod_{i=1}^n (e_i + 1)$.

The proof of Theorem 3.6 by zeta polynomials suggests an alternate proof of Corollary 3.1. By [17, Proposition 8.4], $\text{Cat}^{(m)}(W)$ also counts the facets of the generalized cluster complex associated to an irreducible W . A “rotation” of order $mh + 2$ on the generalized cluster complex leads to a recursion [17, Proposition 8.3] on $\text{Cat}^{(m)}(W)$ and thus on zeta polynomials of L_W :

$$(3.2) \quad Z(L_W, m+1) = \frac{mh+2}{2n} \sum_{s \in S} Z(L_{W_{\langle s \rangle}}, m+1).$$

Corollary 3.1 arises by extracting the coefficient of m^n in (3.2). The juxtaposition of this alternate proof with the proof via Theorem 1.1 is striking in that two different dihedral symmetries are related by passing to leading coefficients.

The proof of Corollary 3.1 via (3.2) can presumably be made uniform. The fact that $Z(L_W, 2)$ counts facets of the cluster complex (the case $m = 1$ of the generalized cluster complex) is proven uniformly in [2], based on results of [13]. Recently the results of [13] were extended [35] to the case $m \geq 1$, and presumably the analogous extension of [2] will eventually be undertaken. The recursion counting facets of the generalized cluster complex was proven uniformly, except for [17, Theorems 3.4 and 3.7]. However, the extension of the results of [2] to the case $m \geq 1$ can be expected

¹When W is reducible, the formula holds with h^n replaced by $\prod_{s \in S} h_s$. (See Remark 2.6.)

to provide uniform proofs of [17, Theorems 3.4 and 3.7]. It should be stressed that a uniform proof of (3.1), or even Theorem 3.6, is completely lacking. Indeed, a uniform proof of (3.1) would specialize to a uniform proof that the number of elements of L_W equals $\text{Cat}^{(1)}(W)$. This is an important open problem in W -Catalan combinatorics.

Reduced T -words. Let $\text{TW}_k(W)$ be the number of reduced T -words for elements of absolute length k in L_W . These are chains $x_0 < x_1 < x_2 < \cdots < x_k$ in L_W with $\ell_T(x_i) = i$ for each i , counted by $C_{(1, \dots, 1, n-k)}(W)$. Theorem 1.1 implies the following corollary.

COROLLARY 3.7. *If (W, S) is a finite irreducible Coxeter system, then*

$$\text{TW}_k(W) = \frac{h}{2} \sum_{s \in S} \text{TW}_{k-1}(W_{\langle s \rangle}).$$

Inspired by the alternate proof of Corollary 3.1, we notice a connection between TW_k and the generalized cluster complex. Let $f_k(W, m)$ be the number of k -vertex (i.e., $(k-1)$ -dimensional) simplices in the generalized cluster complex associated to W . When W is irreducible, [17, Proposition 8.3] says that

$$(3.3) \quad f_k(W, m) = \frac{mh + 2}{2k} \sum_{s \in S} f_{k-1}(W_{\langle s \rangle}, m).$$

Taking leading coefficients and interpreting the result as a recursion on $k!$ times the leading coefficient $\langle m^k | f_k(W, m) \rangle$, we obtain a recursion identical to Corollary 3.7. Using Proposition 2.5 and another assertion of [17, Proposition 8.3], one easily checks that $\text{TW}_k(W)$ also behaves like $k! \langle m^k | f_k(W, m) \rangle$ when W is reducible and when $k = 0$. Thus by induction on the rank of W , we have the following theorem.

THEOREM 3.8. *For any finite Coxeter group W ,*

$$\text{TW}_k(W) = k! \langle m^k | f_k(W, m) \rangle.$$

In particular, (nonuniform) formulas for $\text{TW}_k(W)$ can be obtained from [17, Theorem 8.5]. In light of the alternate proof of Theorem 3.6 via zeta polynomials, Theorem 3.8 suggests that $Z(L_W^{[0, k]}, m + 1) = f_k(W, m)$, where $L_W^{[0, k]}$ is the restriction of L_W to ranks $0, \dots, k$. However, this fails even in the smallest examples.

Edges. Let $E(W)$ be the number of edges in the Hasse diagram of L_W . Theorem 1.1 specializes to a recursion on cover relations $x < y$ in L_W with $\ell_T(x) = i$. The right side is a sum over $s \in S$ of the number of elements of $L_{W_{\langle s \rangle}}$ at rank i . Summing from $i = 0$ to $i = n - 1$, we obtain the following recursion on $E(W)$.

COROLLARY 3.9. *If (W, S) is a finite irreducible Coxeter system, then*

$$E(W) = \frac{h}{2} \sum_{s \in S} \text{NC}(W_{\langle s \rangle}).$$

Here $\text{NC}(W)$ stands for the number of elements of L_W . Setting $m = 1$ in (3.1), we see that for W irreducible,

$$(3.4) \quad \text{NC}(W) = \text{Cat}^{(1)}(W) = \prod_{i=1}^n \frac{h + e_i + 1}{e_i + 1}.$$

Furthermore, $\text{Cat}^{(1)}(W)$ counts facets of the cluster complex associated to W . By a recursion [18, Proposition 3.7] counting facets of the cluster complex,

$$(3.5) \quad \text{NC}(W) = \frac{h+2}{2n} \sum_{s \in S} \text{NC}(W_{\langle s \rangle}).$$

Comparing (3.5) with Corollary 3.9, we obtain a uniform formula for the number of edges of L_W .

THEOREM 3.10. *If (W, S) is a finite irreducible Coxeter system, then*

$$E(W) = \frac{nh}{h+2} \text{NC}(W) = \frac{nh}{|W|} \prod_{i=2}^n (h + e_i + 1).$$

The proof of Theorem 3.10 appears to be uniform, but is not, since there is no known uniform proof that $\text{NC}(W) = \text{Cat}^{(1)}(W)$.

Remark 3.11. Equation (3.5) can also be interpreted in terms of L_W . The equation can be rearranged to state that

$$(3.6) \quad \sum_{s \in S} (\text{NC}(W) - \text{NC}(W_{\langle s \rangle})) = \frac{h}{2} \sum_{s \in S} \text{NC}(W_{\langle s \rangle}).$$

The right side of the equation is $E(W)$. The left side of (3.6) counts pairs $(s, x) \in S \times L_W$ with $x \notin W_{\langle s \rangle}$. (Cf. the proof of Lemma 2.3.) A comparison of the expression for $E(W_1 \times W_2)$ arising from Proposition 2.5 with the left side of (3.6) in the case $W = W_1 \times W_2$ shows that, even in the reducible case, $E(W)$ counts pairs (s, x) as above. We have no explanation, within the combinatorics of noncrossing partitions, for the coincidence between these two counts.

Saturated chains. Theorem 1.1 specializes to a recursion on saturated chains $x_0 < x_1 < \cdots < x_k$ with $\ell_T(x_0) = i$. Summing over all possible i , we obtain the following recursion on the total number $\text{SC}_k(W)$ of saturated chains of length k (i.e., $k+1$ elements) in L_W .

COROLLARY 3.12. *If (W, S) is a finite irreducible Coxeter system and $k > 0$, then*

$$\text{SC}_k(W) = \frac{h}{2} \sum_{s \in S} \text{SC}_{k-1}(W_{\langle s \rangle}).$$

Corollaries 3.1 and 3.9 are special cases of Corollary 3.12 which lead to uniform formulas, and (3.4) is a uniform formula for $\text{SC}_0(W)$. One more special case leads to a uniform formula. When $k = n-1$, the recursion of Corollary 3.12 is solved in essentially the same manner as the recursion of Corollary 3.1 to obtain, for W irreducible,

$$(3.7) \quad \text{SC}_{n-1} = \frac{2n! h^n}{|W|}.$$

This is also a trivial corollary of Theorem 3.6, since L_W has a unique minimal and maximal element. Naturally, one seeks a formula for SC_k that generalizes Theorems 3.6 and 3.10 and (3.4) and (3.7). However, the obvious generalization

$$(3.8) \quad n(n-1) \cdots (n-k+1) \frac{h^k}{|W|} \prod_{i=k+1}^n (h + e_i + 1)$$

does not work beyond the cases $k \in \{0, 1, n-1, n\}$. It appears to fail in a way that is exactly analogous to the situation for f -numbers and h -numbers of cluster complexes, as explained in section 4.

4. m -divisible noncrossing partitions. This section is an extended remark about generalizing the results of this paper to the poset of m -divisible noncrossing partitions, as defined by Armstrong in [1]. We call this poset $L_W^{(m)}$. Armstrong names it $\text{NC}_{(m)}(W)$ and also considers the dual poset, under the name $\text{NC}^{(m)}(W)$. The m -divisible noncrossing partitions are the delta sequences of m -element multichains in L_W . (See the proof of Proposition 2.4.) In particular, $L_W^{(1)} = L_W$.

Let $\delta = (\delta_0, \dots, \delta_k)$ and $\delta' = (\delta'_0, \dots, \delta'_k)$ be delta sequences of multichains in L_W . The definition of $L_W^{(m)}$ sets $\delta \leq \delta'$ if and only if $\delta_i \leq \delta'_i$ for all $i = 1, \dots, k$. The poset $L_W^{(m)}$ is a graded meet-semilattice [1, Theorem 3.4.4] with rank function $\sum_{i=1}^n \ell_T(\delta_i)$. Let $C_{(j_1, j_2, \dots, j_{k+1})}^{(m)}(W)$ count multichains $x_1 \leq x_2 \leq \dots \leq x_k$ in $L_W^{(m)}$ with rank-differences given by $(j_1, j_2, \dots, j_{k+1})$. Theorem 1.1 generalizes as follows.

THEOREM 4.1. *If (W, S) is a finite irreducible Coxeter system and $j_i = 1$, then*

$$C_{(j_1, j_2, \dots, j_{k+1})}^{(m)}(W) = \frac{mh}{2} \sum_{s \in S} C_{(j_1, j_2, \dots, \hat{j}_i, \dots, j_{k+1})}^{(m)}(W_{\langle s \rangle}).$$

Theorem 4.1 is proved by a straightforward but notationally cumbersome generalization of the proof of Theorem 1.1, replacing the action of $\langle c_+, c_- \rangle$ with the action of $\langle L^*, R^* \rangle$, as defined in [1, section 3.4.6]. We omit the details.

Theorem 4.1 implies a formula for the number $\text{MC}^{(m)}(W)$ of maximal chains in $L_W^{(m)}$:

$$(4.1) \quad \text{MC}^{(m)}(W) = \frac{mh}{2} \sum_{s \in S} \text{MC}^{(m)}(W_{\langle s \rangle}).$$

The powers of m in the formula can be factored out trivially, so that this recursion is solved exactly as in the case $m = 1$. Thus for W irreducible we have $\text{MC}^{(m)}(W) = \frac{n!(mh)^n}{|W|}$, as was pointed out earlier by Armstrong [1, Corollary 3.6.10].

The analogue of Corollary 3.7 behaves similarly: powers of m can be factored out. Thus saturated chains $x_0 < x_1 < x_2 < \dots < x_k$ such that x_0 is the unique minimal element of $L_W^{(m)}$ are counted by $m^k \text{TW}_k(W)$. It is an easy exercise, given some familiarity with $L_W^{(m)}$, to construct an (m^k) -to-1 map from these “lower-saturated” chains in $L_W^{(m)}$ to lower-saturated chains in L_W . There are also m -analogues of the formulas given in Corollary 3.9, Theorem 3.10, and Remark 3.11, for the number $E^{(m)}(W)$ of edges in $L_W^{(m)}$. In each case, with h replaced by mh and with superscripts “ (m) ” in the appropriate places, the analogous proof works.

Corollary 3.12 generalizes, with similar modifications, to a recursion counting saturated chains in $L_W^{(m)}$. The resulting type-by-type formulas for $\text{SC}_k^{(m)}(W)$ are polynomials in m . Factoring these into irreducible factors reveals multiplicative formulas which seem to be loosely based on (3.8), with h replaced by mh throughout. These factorizations exhibit the odd phenomena first observed in formulas [17, Theorems 8.5 and 10.2] for the f - and h -numbers of generalized cluster complexes, including “levels” of exponents and single “mysterious” factors. The formulas for $\text{SC}_k^{(m)}(W)$ appear to be badly behaved exactly when $f_k(W, m)$ and $h_k(W, m)$ are badly behaved, and the bad behaviors take the same form.

Acknowledgments. I thank Drew Armstrong, Christos Athanasiadis, Jon McCammond, and John Stembridge for helpful comments on earlier versions of this paper.

REFERENCES

- [1] D. ARMSTRONG, *Generalized Noncrossing Partitions and Combinatorics of Coxeter Groups*, Mem. Amer. Math. Soc., to appear.
- [2] C. A. ATHANASIADIS, T. BRADY, J. MCCAMMOND, AND C. WATT, *h -vectors of generalized associahedra and noncrossing partitions*, Int. Math. Res. Not., 2006, Art. ID 69705, 28 pp.
- [3] C. A. ATHANASIADIS, T. BRADY, AND C. WATT, *Shellability of noncrossing partition lattices*, Proc. Amer. Math. Soc., 135 (2007), pp. 939–949.
- [4] C. A. ATHANASIADIS AND V. REINER, *Noncrossing partitions for the group D_n* , SIAM J. Discrete Math., 18 (2004), pp. 397–417.
- [5] D. BESSIS, *The dual braid monoid*, Ann. Sci. École Norm. Sup. (4), 36 (2003), pp. 647–683.
- [6] P. BIANE, *Some properties of crossings and partitions*, Discrete Math., 175 (1997), pp. 41–53.
- [7] P. BIANE, *Parking functions of types A and B*, Electron. J. Combin., 9 (2002), Note 7, 5 pp.
- [8] A. BJÖRNER AND F. BRENTI, *Combinatorics of Coxeter Groups*, Graduate Texts in Mathematics 231, Springer, New York, 2005.
- [9] N. BOURBAKI, *Lie Groups and Lie Algebras, Chapters 4–6*, in Elements of Mathematics. Translated from the 1968 French original by Andrew Pressley, Springer-Verlag, Berlin, 2002.
- [10] T. BRADY, *Artin groups of finite type with three generators*, Michigan Math. J., 47 (2000), pp. 313–324.
- [11] T. BRADY AND C. WATT, *$K(\pi, 1)$'s for Artin groups of finite type*, in Proceedings of the Conference on Geometric and Combinatorial Group Theory, Part I (Haifa, 2000), Geom. Dedicata, 94 (2002), pp. 225–250.
- [12] T. BRADY AND C. WATT, *A partial order on the orthogonal group*, Comm. Algebra, 30 (2002), pp. 3749–3754.
- [13] T. BRADY AND C. WATT, *Non-crossing partition lattices in finite real reflection groups*, Trans. Amer. Math. Soc., 360 (2008), pp. 1983–2005.
- [14] F. CHAPOTON, *Enumerative properties of generalized associahedra*, Sémin. Lothar. Combin., 51 (2004/05), Art. B51b, 16 pp.
- [15] P. H. EDELMAN, *Chain enumeration and noncrossing partitions*, Discrete Math., 31 (1980), pp. 171–180.
- [16] S. B. EKHAD AND J. E. MAJEWICZ, *A short WZ-style proof of Abel's identity*, Electron. J. Combin., 3 (1996), no. 2, Research Paper 16, approx. 1 p.
- [17] S. FOMIN AND N. READING, *Generalized cluster complexes and Coxeter combinatorics*, Int. Math. Res. Not., 2005, no. 44, pp. 2709–2757.
- [18] S. FOMIN AND A. ZELEVINSKY, *Y -systems and generalized associahedra*, Ann. of Math. (2), 158 (2003), pp. 977–1018.
- [19] S. FOMIN AND A. ZELEVINSKY, *Cluster algebras. II. Finite type classification*, Invent. Math., 154 (2003), pp. 63–121.
- [20] M. HAIMAN, *Combinatorics, symmetric functions, and Hilbert schemes*, in Current Developments in Mathematics, 2002, Int. Press, Somerville, MA, 2003, pp. 39–111.
- [21] P. HERSH, *Deformation of chains via a local symmetric group action*, Electron. J. Combin., 6 (1999), Research Paper 27, 18 pp.
- [22] J. E. HUMPHREYS, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics 29, Cambridge University Press, Cambridge, UK, 1990.
- [23] C. INGALLS AND H. THOMAS, *Noncrossing partitions and representations of quivers*, preprint, 2006 (math.RT/0612219).
- [24] C. KRATTENTHALER AND T. MÜLLER, *Decomposition numbers for finite Coxeter groups and generalised non-crossing partitions*, preprint, 2007 (arXiv:0704.0199).
- [25] G. KREWERAS, *Sur les partitions non croisées d'un cycle*, Discrete Math., 1 (1972), pp. 333–350.
- [26] J. MCCAMMOND, *Noncrossing partitions in surprising locations*, Amer. Math. Monthly, 113 (2006), pp. 598–610.
- [27] N. READING, *Clusters, Coxeter-sortable elements and noncrossing partitions*, Trans. Amer. Math. Soc., 359 (2007), pp. 5931–5958.
- [28] N. READING AND D. SPEYER, *Sortable elements in infinite Coxeter groups*, preprint, 2008 (arXiv:0803.2722).
- [29] V. REINER, *Non-crossing partitions for classical reflection groups*, Discrete Math., 177 (1997), pp. 195–222.

- [30] R. SIMION, *Noncrossing partitions*, in Formal Power Series and Algebraic Combinatorics (Vienna, 1997), Discrete Math., 217 (2000), pp. 367–409.
- [31] R. P. STANLEY, *Enumerative Combinatorics*, Vol. 1, Cambridge Studies in Advanced Mathematics 49, Cambridge University Press, Cambridge, UK, 1997.
- [32] R. P. STANLEY, *Parking functions and noncrossing partitions*, The Wilf Festschrift, Electron. J. Combin., 4 (1997), no. 2, Research Paper 20, approx. 14 pp.
- [33] J. R. STEMBRIDGE, personal communication, 2004.
- [34] R. STEINBERG, *Finite reflection groups*, Trans. Amer. Math. Soc., 91 (1959), pp. 493–504.
- [35] E. TZANAKI, *Faces of generalized cluster complexes and noncrossing partitions*, SIAM J. Discrete Math., 22 (2008), pp. 15–30.

ON THE FIRST-FIT CHROMATIC NUMBER OF GRAPHS*

JÓZSEF BALOGH[†], STEPHEN G. HARTKE[‡], QI LIU[†], AND GEXIN YU[§]

Abstract. The *first-fit chromatic number* of a graph is the number of colors needed in the worst case of a greedy coloring. It is also called the *Grundy number*, which is defined to be the maximum number of classes in an ordered partition of the vertex set of a graph G into independent sets $V_1, V_2, \dots, V_k$ so that for each $1 \leq i < j \leq k$ and for each $x \in V_j$ there exists a $y \in V_i$ such that x and y are adjacent. In this paper, we study the first-fit chromatic number of outerplanar and planar graphs as well as Cartesian products of graphs, and in particular we give asymptotically tight results for outerplanar graphs.

Key words. first-fit chromatic number, Grundy number, Grundy coloring, greedy coloring, random graph, planar graph, Cartesian product

AMS subject classifications. 05C15, 05C35, 05C85

DOI. 10.1137/060672479

1. Introduction. Given an ordering of the vertices of a simple graph G , a greedy (proper) coloring of G assigns to each vertex the first available color not used on a neighboring vertex earlier in the order. The first-fit chromatic number is the number of colors needed in the worst case of a greedy coloring. An equivalent definition of first-fit chromatic number is given below.

The *first-fit chromatic number* of G , written as $\chi_{FF}(G)$, is defined to be the maximum number of classes in an ordered partition of the vertex set of G into independent sets $V_1, V_2, \dots, V_k$ so that for each $1 \leq i < j \leq k$ and for each $x \in V_j$ there exists a $y \in V_i$ such that x and y are adjacent. A partition with this property is called a *first-fit partition*.

Historically, first-fit partitions and the first-fit chromatic number are also called Grundy colorings and the Grundy number, respectively. The study of Grundy coloring dates back to the 1930s [13], when Grundy used them in the study of kernels of directed graphs. Many researchers have studied this coloring under different names; see [8] for details. It is believed that Christen and Selkow [5] were the first to define and study the Grundy number as a graph parameter. Some recent results about it can be found, for example, in Füredi et al. [9], Zaker [21], Erdős et al. [7], and Germainian and Kheddouci [10].

Stemming from a natural connection with the classical dynamic storage allocation problem in computer science, Woodall [20] studied the first-fit chromatic number of an interval graph I_k with maximum clique size k and showed $\chi_{FF}(I_k) = O(k \log k)$. After a series of papers by Kierstead [15], Kierstead and Qin [16], and Pemmaraju, Raman, and Varadarajan [19], Brightwell, Kierstead, and Trotter [4] showed that

*Received by the editors October 16, 2006; accepted for publication (in revised form) January 3, 2008; published electronically May 16, 2008.

<http://www.siam.org/journals/sidma/22-3/67247.html>

[†]Department of Mathematics, University of Illinois, Urbana, IL 61801 (jobal@math.uiuc.edu, qiliu@math.uiuc.edu). The research of the first author was supported in part by NSF grants DMS-0302804, DMS-0600303, and DMS-0603769, UIUC Campus Research Board #06139 and #07048, and OTKA 049398.

[‡]Department of Mathematics, University of Nebraska, Lincoln, NE 68588 (shartke2@math.unl.edu).

[§]Department of Mathematics, Vanderbilt University, Nashville, TN 37240 (gexin.yu@vanderbilt.edu). The research of this author was supported in part by NSF grant DMS-0652306.

$\chi_{FF}(I_k) \leq 8k$. Kierstead and Trotter [17] improved the lower bound by showing that for every $\varepsilon > 0$, $\chi_{FF}(I_k) > (5 - \varepsilon)k$ when k is sufficiently large. We study the first-fit chromatic numbers of more general families of graphs.

THEOREM 1. *Let $\mathcal{F}$ be a family of graphs closed under taking subgraphs. Suppose that there exists a constant $d \geq 1$ such that $e(G) \leq dn(G)$ for every $G \in \mathcal{F}$. Then for every $G \in \mathcal{F}$,*

$$\chi_{FF}(G) \leq \log_{\frac{d+1}{d}}(n) - \log_{\frac{d+1}{d}}(d+4) + (2d+2).$$

Here, $e(G)$ denotes the number of edges of G , and $n(G)$ the number of vertices of G .

Irani [14] proved a similar result, showing that for the graphs above, $\chi_{FF}(G) = O(d \log n)$. However, our constants are slightly better and the proof techniques are different.

As a corollary, we obtain upper bounds of the first-fit chromatic numbers of planar and outerplanar graphs.

THEOREM 2. (i) *Let G be an n -vertex planar graph. Then*

$$\chi_{FF}(G) \leq \log_{4/3}(n) + 8 - \log_{4/3}(7).$$

(ii) *If G is additionally outerplanar, then*

$$\chi_{FF}(G) \leq \log_{3/2}(n) + 6 - \log_{3/2}(6).$$

For infinitely many integers n we construct planar graphs on n vertices with first-fit chromatic number $\log_{6^{1/6}} n$ (note that $6^{1/6} \approx 1.348$), and we show that the upper bound for outerplanar graphs is asymptotically best possible.

We also study the first-fit chromatic number of random graphs. The first results were proved by Grimmett and McDiarmid [11] and then improved by Bollobás and Erdős [2], who showed among other results that if $G \sim G_{n,1/2}$ is a random graph with edge density $1/2$, then with high probability (as $n \rightarrow \infty$),

$$\chi_{FF}(G) \geq \left(1 - 3\sqrt{\frac{\ln \ln n}{\ln n}}\right) \frac{n}{\log_2 n}.$$

Bollobás and Erdős [2] and McDiarmid [18] also showed that with high probability,

$$\chi_{FF}(G) \leq \left(1 + \frac{5 \ln \ln n}{\ln n}\right) \frac{n}{\log_2 n}.$$

In [11] and [2] much more was proven, including results about $G_{n,p}$. An excellent summary of these results appears in Bollobás's book *Random Graphs* [3, pp. 294–298].

We improve the lower bound on the first-fit chromatic number of $G_{n,1/2}$ by showing the following theorem.

THEOREM 3. *Let $G \sim G_{n,1/2}$ be a random graph with edge density $1/2$. Then with high probability (as $n \rightarrow \infty$),*

$$\chi_{FF}(G) \geq \left(1 - \frac{30(\ln \ln n)^2}{\log_2 n}\right) \frac{n}{\log_2 n}.$$

Note that the chromatic number of a random graph $G \sim G_{n,1/2}$ is asymptotically $n/(2 \log_2 n)$ almost surely (see Alon and Spencer [1]), and the ratio of the two chromatic numbers is asymptotically 2.

We discuss the first-fit chromatic number of Cartesian products of graphs in the last section. Effantin and Kheddouci [6] proved that if G is bipartite, then $\chi_{FF}(G \square P_3) \geq \chi_{FF}(G) + 2$. Using this, they proved that $\chi_{FF}(C_{m_1} \square C_{m_2} \square \cdots \square C_{m_k}) = 2k + 1$, where $k \geq 2$, $m_i \geq 4$, and m_i is even. (The case when all cycles are C_4 requires additional work.) Note that this result gives a short proof for the first-fit chromatic number of hypercubes which was originally obtained by Hoffman and Johnson [12]: A k -dimensional hypercube Q_k has first-fit chromatic number $k + 1$ for $k \geq 3$, and $\chi_{FF}(Q_2) = 2$.

Here we extend the above result to include odd cycles in the product and prove the same lower bound for $G \square P_3$ and $G \square P_4$, but we allow G to be nonbipartite as long as $\chi_{FF}(G)$ is even. As a result, $\chi_{FF}(C_{m_1} \square C_{m_2} \square \cdots \square C_{m_k}) = 2k + 1$ for all $k \geq 2$ and $m_i \geq 4$. For contrast, we construct graphs G with the property that $\chi_{FF}(G \square P_2)$ is as large as $2\chi_{FF}(G)$.

This paper is organized as follows: We prove the results on planar graphs in section 2, the lower bound of random graphs in section 3, and the results on products of graphs in section 4.

2. First-fit chromatic number of planar graphs. We begin with a technical, purely combinatorial lemma about sequences of integers. The proof is delayed until the end of the section.

LEMMA 4. *Given a sequence $a_1, a_2, \dots, a_t$ of positive integers and a constant $d \geq 1$ which satisfy*

- (i) $a_i \geq 1$ for $i = 1, \dots, t$,
- (ii) $\sum_{\ell=1}^t a_\ell = n$,
- (iii) $\sum_{\ell=i}^t (\ell - i) a_\ell \leq d(\sum_{\ell=i}^t a_\ell)$ for $i = 1, \dots, t - d$.

Then

$$t \leq \log_{\frac{d+1}{d}}(n) - \log_{\frac{d+1}{d}}(d+4) + (2d+2).$$

We now use Lemma 4 to prove Theorems 1 and 2.

Proof of Theorem 1. Suppose that $\chi_{FF}(G) = t$. Then there exists a first-fit partition $\langle V_1, V_2, \dots, V_t \rangle$ of $V(G)$. For $1 \leq i \leq t$, let G_i be the subgraph $G[V_i, V_{i+1}, \dots, V_t]$ induced by the last $t - i + 1$ parts. Since $\langle V_1, V_2, \dots, V_t \rangle$ is a first-fit partition,

$$e(G_i) \geq \sum_{\ell=i}^t (\ell - i) |V_\ell|.$$

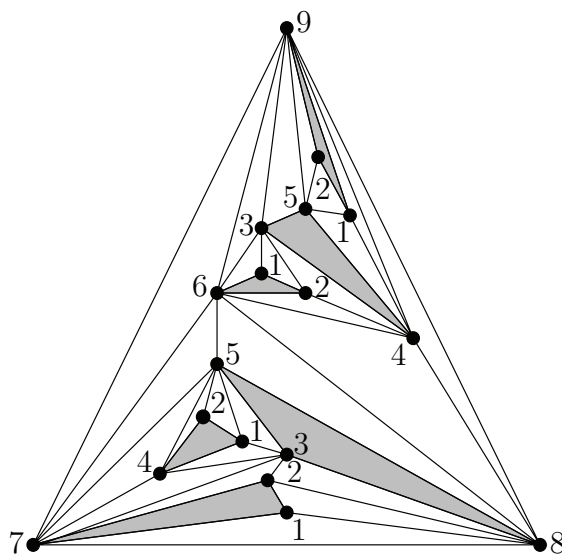
Since G_i is a member of $\mathcal{F}$, G_i satisfies the edge bound

$$e(G_i) \leq d \left(\sum_{\ell=i}^t |V_\ell| \right).$$

Hence, the sequence $|V_1|, |V_2|, \dots, |V_t|$ satisfies the conditions of Lemma 4 where $a_i = |V_i|$, and so

$$\chi_{FF}(G) \leq \log_{\frac{d+1}{d}}(n) - \log_{\frac{d+1}{d}}(d+4) + (2d+2). \quad \square$$

Proof of Theorem 2. (i) If G is a planar graph with $n \geq 3$ vertices, then $e(G) \leq 3n - 6 < 3n$. Since planarity is preserved when taking subgraphs, the result then follows by applying Theorem 1.

FIG. 1. An 18-vertex graph H with first-fit chromatic number 9.

(ii) Since outerplanar graphs with n vertices have at most $2n-3$ edges, we similarly obtain the result. $\square$

PROPOSITION 5. *For every n_0 , there is an $n > n_0$ such that there exists an n -vertex planar graph with first-fit chromatic number at least $\log_{6^{1/6}} n$.*

Proof. A planar graph is called a *triangulation* if all of its faces are triangles. There exists a triangulation H (see Figure 1) with 18 vertices that has first-fit chromatic number at least 9. (The vertices with label i belong to the U_i in a first-fit partition $\cup_{i=1}^9 U_i$ of H .) The graph H has the additional property that the vertices of H can be covered by exactly 6 independent triangles (shown in gray in Figure 1).

Given a triangulation G_1 and a set $\mathcal{F}$ of triangles that cover every vertex of G_1 , we replace every triangle of $\mathcal{F}$ with a copy of H . This forms a new triangulation G_2 , where $n(G_2) \leq 6n(G_1)$. Suppose that $V(G_1)$ has an optimal first-fit partition $\langle V_1, V_2, \dots, V_t \rangle$. Then $V(G_2)$ has a first-fit partition $\langle U_1, U_2, \dots, U_{t+6} \rangle$, where $U_i = V_{i-6}$ when $i \geq 7$ and, for $i \leq 6$, U_i is the set of vertices labeled i from copies of H . Hence, $\chi_{FF}(G_2) \geq \chi_{FF}(G_1) + 6$. Note that G_2 has a set of $6|\mathcal{F}|$ triangles that cover every vertex of G_1 .

By iterating this procedure starting with $G_1 = C_3$, we obtain a sequence G_i of graphs with $3 \cdot 6^{i-1}$ vertices and first-fit chromatic number at least $3 + 6(i-1)$. Hence, we have constructed a series of planar graphs whose asymptotic first-fit chromatic number is at least $\log_{6^{1/6}} n$, where n is the number of the vertices of the graph. $\square$

By considering larger graphs H as a pattern, we may produce a better lower bound, but we do not believe this method can show that the upper bound is tight.

The following construction gives a lower bound for the first-fit chromatic number of outerplanar graphs.

PROPOSITION 6. *For every n_0 , there is an $n > n_0$ such that there exists an n -vertex outerplanar graph with first-fit chromatic number at least $\log_{3/2}(n+2) - \log_{3/2}(5) + 3$.*

Proof. We construct a sequence of outerplanar graphs as follows:

1. $G_0 = C_3$.

2. Suppose the outer face of G_i is $v_1 v_2 \dots v_{n_i}$ in order, where n_i is the number of vertices of G_i . Generate G_{i+1} by adding $k = \lceil n_i/2 \rceil$ vertices $u_1, \dots, u_k$ such that u_j is adjacent to v_{2j-1} and v_{2j} , where $j = 1, \dots, k$. (If n_i is odd, then u_k is adjacent to v_{n_i} and v_1 .)

Then G_{i+1} is also outerplanar and $\chi_{FF}(G_{i+1}) \geq \chi_{FF}(G_i) + 1$. As $\chi_{FF}(G_0) = 3$, we have $\chi_{FF}(G_i) \geq i + 3$. Since $n_i = n_{i-1} + \lceil n_{i-1}/2 \rceil \leq 3n_{i-1}/2 + 1$ and $n_0 = 3$, we have $n_i \leq 5(3/2)^i - 2$. Thus $\chi_{FF}(G_i) \geq i + 3 \geq \log_{3/2}(n_i + 2) - \log_{3/2}(5) + 3$. $\square$

So the upper bound given by Theorem 2 (ii) for outerplanar graphs is asymptotically sharp.

We conclude this section with the proof of Lemma 4.

Proof of Lemma 4. Given a sequence $a_1, a_2, \dots, a_t$ that satisfies the hypotheses of the lemma, we define a “shift” operation S_j for $1 \leq j \leq t-1$ on such a sequence as follows:

$$S_j(a_1, a_2, \dots, a_t) = a_1, \dots, a_{j-1}, a_j + 1, a_{j+1} - 1, a_{j+2}, \dots, a_t.$$

Clearly, if $a_1, a_2, \dots, a_t$ satisfies (i) and (ii), then $S_j(a_1, a_2, \dots, a_t)$ does as well if $a_{j+1} \geq 2$.

We repeatedly apply S_j to the sequence, where j is the largest index such that $a_{j+1} > 1$ and (iii) holds for $S_j(a_1, a_2, \dots, a_t)$. We stop when no such index j exists. After the process terminates, we obtain a new sequence $b_1, b_2, \dots, b_t$ that satisfies the conditions of the lemma for the same d , but, for each $j = 2, \dots, t$, either $b_j = 1$ or $b_j \geq 2$ and (iii) does not hold for $S_{j-1}(b_1, b_2, \dots, b_t)$.

Given the final sequence $b_1, b_2, \dots, b_t$, let i_0 be the unique index $j \geq 1$ such that

- (a) for any i with $j < i \leq t$, $b_i = 1$, and
- (b) if $j > 1$, then $b_j \geq 2$ and $S_{j-1}(b_1, b_2, \dots, b_t)$ violates (iii).

Note that the definition of i_0 includes the possibility that $i_0 = 1$. However, the next claim implies that $i_0 > 1$, since we may assume that $t > 2d + 2$. (Otherwise the conclusion of the lemma is trivially satisfied.)

Claim A. $2 \leq i_0 = t - (2d + 1)$.

Proof. Condition (iii) for $b_1, \dots, b_t$ with $i \geq i_0$ implies that

$$\left(\frac{t-i+1}{2} - d \right) (t-i) \leq db_i,$$

since $b_i = 1$ for $i > i_0$. Thus,

$$(1) \quad \text{if } \left(\frac{t-i+1}{2} - d \right) (t-i) \geq d+1, \text{ then } b_i > 1.$$

Define

$$f(x) = \left(\frac{x+1}{2} - d \right) x = \frac{1}{2}x^2 - (d-1/2)x.$$

Note that $f(x)$ is increasing on the interval $[d-1/2, \infty)$.

Since (iii) imposes no condition when $i > t-d$, we may always apply S_j for $j \geq t-d$ as long as $a_j \geq 2$. Thus we have $i_0 \leq t-d-1$, and so $t-i_0 \geq d+1$.

Since $b_i = 1$ for $i > i_0$ and by (1), we have that $f(t-i) < d+1$ for $i > i_0$. Since $f(2d+1) = 2d+1 > d+1$, we have that $i_0 \geq t-(2d+1) \geq 2$. Note that this rules out the possibility that $i_0 = 1$.

For the upper bound on i_0 , we first note that if any sequence $a'_1, a'_2, \dots, a'_t$ satisfies (iii), then $S_j(a'_1, a'_2, \dots, a'_t)$ satisfies (iii) for $i \neq j+1$. Thus, if $S_j(b_1, b_2, \dots, b_t)$ violates (iii), then the violation occurs for $i = j+1$:

$$(2) \quad 0 \cdot (b_{j+1} - 1) + \sum_{\ell=j+2}^t (\ell - j - 1)b_\ell \geq d \left(b_{j+1} - 1 + \sum_{\ell=j+2}^t b_\ell \right) + 1.$$

Since $i_0 > 1$, $b_{i_0} \geq 2$, and $S_{i_0-1}(b_1, b_2, \dots, b_t)$ does not satisfy (iii), and by (2), we have

$$(3) \quad \left(\frac{t - i_0 + 1}{2} - d \right) (t - i_0) = f(t - i_0) \geq db_{i_0} - d + 1 \geq d + 1,$$

where we use the fact that $b_i = 1$ for all $i_0 < i \leq t$. This implies

$$(4) \quad \left(\frac{t - i_0 + 1}{2} \right) (t - i_0) \geq d(t - i_0 + 1) + db_{i_0} - 2d + 1 > d(t - i_0 + 1),$$

where the first inequality is equivalent to (3). Hence, $i_0 < t - 2d$, and since i_0 is an integer, $i_0 \leq t - (2d + 1)$. $\square$

Claim B. $b_i \geq 2$ for all $i \leq i_0$, and as a consequence, $S_{i-1}(b_1, b_2, \dots, b_t)$ violates (iii) for all $i \leq i_0$.

Proof. We use downward induction on i . The base case is when $i = i_0$, which follows from the definition of i_0 .

Thus, we assume the inductive hypothesis is true for i , where $1 < i \leq i_0$. Hence

$$(5) \quad \sum_{\ell=i}^t (\ell - i)b_\ell \geq d \left(\sum_{\ell=i}^t b_\ell \right) - d + 1.$$

Condition (iii) for $i - 1$ states that

$$(6) \quad \sum_{\ell=i-1}^t (\ell - i + 1)b_\ell \leq d \left(\sum_{\ell=i-1}^t b_\ell \right).$$

Subtracting (5) from (6), we obtain

$$(7) \quad \sum_{\ell=i}^t b_\ell \leq db_{i-1} + d - 1,$$

which implies

$$\begin{aligned} db_{i-1} &\geq \sum_{\ell=i}^t b_\ell + 1 - d \geq (t - i + 2) + 1 - d && \text{since } b_i \geq 2, \\ &\geq t - d + 3 - i_0 && \text{since } i \leq i_0, \\ &\geq d + 4 && \text{since } i_0 = t - (2d + 1) \text{ by Claim A.} \end{aligned}$$

Hence $b_{i-1} \geq 2$, and additionally $S_j(b_1, b_2, \dots, b_t)$ violates (iii) when $j = i - 2$. This completes the induction. $\square$

By Claims A and B, for $b_1, b_2, \dots, b_t$ we have that (iii) holds and $b_i = 1$ for $i > i_0$ and $b_i \geq 2$ for $i \leq i_0$. We know from (7) that for $i = 1, \dots, i_0 - 1$,

$$db_i \geq \sum_{\ell=i+1}^t b_\ell - (d - 1).$$

Let $B_j := \sum_{\ell=j}^t b_\ell$. Then for $j \leq i_0 - 1$,

$$(8) \quad d(B_j - (d-1)) \geq (d+1)(B_{j+1} - (d-1)).$$

Thus, iterating the recurrence (8) starting with $j = 1$ until $j + 1 = i_0$, we have (note that $B_1 = n$)

$$n - (d-1) = B_1 - (d-1) \geq \left(\frac{d+1}{d}\right)^{i_0-1} (B_{i_0} - (d-1)).$$

Taking logarithms of both sides and solving for i_0 ,

$$i_0 \leq \log_{\frac{d+1}{d}}(n - (d-1)) - \log_{\frac{d+1}{d}}(B_{i_0} - (d-1)) + 1.$$

Note that $i_0 = t - (2d+1)$ and $b_{i_0} \geq 2$ implies $B_{i_0} = t - i_0 + b_{i_0} \geq 2d+3$, and from $i_0 = t - (2d+1)$ we have

$$\begin{aligned} t = i_0 + (2d+1) &\leq \log_{\frac{d+1}{d}}(n - (d-1)) - \log_{\frac{d+1}{d}}(d+4) + 1 + (2d+1) \\ &\leq \log_{\frac{d+1}{d}}(n) - \log_{\frac{d+1}{d}}(d+4) + (2d+2). \quad \square \end{aligned}$$

3. The first-fit chromatic number of random graphs. Recall that if G is a random graph on n labeled vertices in which every edge is chosen randomly and independently with probability p , then we say that $G \sim G_{n,p}$. Since n is large, for the sake of clarity of exposition, we will ignore the issue of integrality of the quantities.

For a vertex x in a graph G , let $\bar{N}(x)$ denote the set of nonneighbors of x in G . Let $\bar{N}(A) = \cap_{x \in A} \bar{N}(x) - A$. In our proof of Theorem 3 the key statements are Lemmas 8 and 9, whose proofs are based on repeated use of Chernoff's inequality. For a variant of Chernoff's inequality, see [1, p. 263]; we will use the following version.

LEMMA 7. *Let $G \sim G_{n,1/2}$ be a random graph, T a subset of $V(G)$ of size t , and x a vertex in $V(G) - T$. Then*

$$\mathbb{P} \left[|\bar{N}(x) \cap T| > \left(\frac{1}{2} + \gamma \right) t \right] < e^{-2\gamma^2 t}.$$

LEMMA 8. *Let $G \sim G_{n,1/2}$ be a random graph, and let $\delta_1 = 18/\sqrt{\log_2 n}$. Then for large n , G contains a maximal independent set C of size at most $(1 + \delta_1) \log_2 n$ with probability at least $1 - n^{-2}$. Note that during the process of choosing C no information is used about the graph spanned by $V(G) - C$.*

Proof. We construct a maximal independent set S iteratively in three phases. We start with $S_0 = \emptyset$. In each step of the first phase we extend this set by adding a vertex v_i so that $S_i = S_{i-1} \cup \{v_i\}$.

Phase I. This phase lasts for r iterations, until $|\bar{N}[S_r]| \leq (\ln n)^3$.

Let $\alpha := \alpha(n) = 2/\ln n$. Let x_1 be an arbitrarily chosen vertex of G ; set $S_1 := \{x_1\}$. We pick a sequence of vertices $x_2, x_3, \dots, x_r$ inductively as follows: let $A_i = \bigcap_{j=1}^i \bar{N}[x_j]$. (For convenience, let $A_0 = V(G)$.) While $|A_i| > (\ln n)^3$, arbitrarily choose x_{i+1} in A_i and set $S_i = S_{i-1} \cup \{x_i\}$.

Note that the selection of the vertices $x_1, x_2, \dots, x_i$ is independent of the edges in the subgraph of G induced by A_i . Hence, the subgraph induced by A_i is a random graph with distribution $G_{|A_i|,1/2}$, and we expect that x_{i+1} is adjacent to about half of the vertices in A_i . By Lemma 7 and using that $|A_i| > (\ln n)^3$,

$$(9) \quad \mathbb{P} [|\bar{N}(x_{i+1}) \cap A_i| > (1/2 + \alpha)|A_i|] < e^{-2\alpha^2|A_i|} < e^{-2\alpha^2(\ln n)^3} = e^{-8 \ln n} = n^{-8}.$$

Thus,

$$\mathbb{P} \left[|\overline{N}(x_{i+1}) \cap A_i| \leq (1/2 + \alpha)|A_i| \text{ for all } 1 \leq i < r \right] > 1 - rn^{-8} > 1 - n^{-7}.$$

As $|A_0| = n$ and for every $i \leq r$ we have $|A_{i+1}| \leq (1/2 + \alpha)|A_i|$, we conclude that

$$(10) \quad r \leq \log_{\frac{1}{\frac{1}{2} + \alpha}} n < \left[1 + \frac{2}{(\ln 2)^2} \frac{1}{\log_2 n} \right] \log_2 n < \left[1 + \frac{5}{\log_2 n} \right] \log_2 n.$$

The middle inequality follows from the inequalities $1 - x/2 > \frac{1}{1+x} > 1 - x$ and $\ln(1 - x) > -2x$.

Phase II. In the second phase we have the same process, but because $|\overline{N}[S_r]| = o(n)$, we shall have an estimate similar to (9) only after adding $m := \ln \ln n$ vertices at each iteration to the independent set. We proceed with this phase until $|\overline{N}[S_i]| \leq 180 \frac{\ln n}{\ln \ln n}$.

We will add m independent vertices $y_1^i, y_2^i, \dots, y_m^i$ to S_i , one at a time.

Subphase IIa. In this subphase we add the i th group of m vertices to S_{i-1} . Let

$$S_i(j) := S_{i-1} \cup \{y_1^i, y_2^i, \dots, y_j^i\}.$$

As before, pick y_j^i arbitrarily in $B^i(j-1) := \overline{N}[S_i(j-1)]$ (note that $S_i(0) = S_{i-1}$). While $j \leq m$ and $|B^i(j-1)| > \frac{180 \ln n}{\ln \ln n}$, arbitrarily choose y_j^i in $B^i(j-1)$, and set $S_i = S_{i-1} \cup \{y_1^i, y_2^i, \dots, y_m^i\}$. By Lemma 7 and using the fact that $|B^i(j-1)| > \frac{180 \ln n}{\ln \ln n}$,

$$\mathbb{P} \left[|B^i(j)| > \frac{2}{3} |B^i(j-1)| \right] < \exp \left[-\frac{1}{18} |B^i(j-1)| \right] \leq \exp \left[-10 \frac{\ln n}{\ln \ln n} \right].$$

These events are independent for different pairs of i and j . Thus, the probability that each pair (i, j) has more than $\frac{2}{3} |B^i(j-1)|$ nonneighbors in $B^i(j-1)$ is

$$\mathbb{P} \left[|\overline{N}(y_j^i) \cap B^i(j-1)| > \frac{2}{3} |B^i(j-1)| \text{ for all } 1 \leq i \leq m \right] < \left(e^{-\frac{10 \ln n}{\ln \ln n}} \right)^m = n^{-10}.$$

Since $|\overline{N}[S_i]| > \frac{2}{3} |\overline{N}[S_{i-1}]|$ implies that $|\overline{N}(y_j^i) \cap B^i(j-1)| > \frac{2}{3} |B^i(j-1)|$ for all $1 \leq j \leq m$, we have

$$(11) \quad \mathbb{P} \left[|\overline{N}[S_i]| \leq \frac{2}{3} |\overline{N}[S_{i-1}]| \right] > 1 - n^{-10}.$$

We repeat Subphase IIa while $|\overline{N}[S_i]| > \frac{180 \ln n}{\ln \ln n}$. If in the last iteration of Subphase IIa we cannot choose m vertices, then we simply add the vertices chosen so far to S . Suppose that Subphase IIa is repeated q times. By (11), $q \leq \log_{\frac{2}{3}} ((\ln n)^3) + 1$. The “+1” comes from the last possibly incomplete iteration of Subphase IIa. Hence, $(3 \log_{3/2} \ln n + 1) \ln \ln n$ vertices are added to S during Phase II.

Phase III. This phase occurs while $|\overline{N}[S]| \leq \frac{180 \ln n}{\ln \ln n}$. We choose any maximal independent set C in $\overline{N}[S]$ to add to S . As $\overline{N}[S]$ spans a random graph with edge-density $1/2$, it does not have an independent set of size $10\sqrt{\log_2 n}$ with probability at least

$$1 - \left(\frac{|\overline{N}[S]|}{10\sqrt{\log_2 n}} \right) 2^{-(\frac{10\sqrt{\log_2 n}}{2})} > 1 - n^{-3}.$$

Thus, $|C| \leq 10\sqrt{\log_2 n}$ with high probability.

Thus, after the completion of the three phases, S is a maximal independent set of G of size at most $\log_{1/(1/2+\alpha)} n + (3 \log_{3/2} \ln n + 1) \ln \ln n + 10\sqrt{\log_2 n}$ with probability greater than $(1 - n^{-3})^3 > 1 - n^{-2}$. Note that

$$\log_{\frac{1}{\frac{1}{2}+\alpha}} n + \left(3 \log_{3/2} \ln n + 1\right) \ln \ln n + 10\sqrt{\log_2 n} \leq (1 + \delta_1) \log_2 n$$

when $\delta_1 = 18/\sqrt{\log_2 n}$ and n is large. Hence the lemma is proved. $\square$

LEMMA 9. Let $G \sim G_{n,1/2}$ be a random graph, and let $\delta_2 = 10(\ln \ln n)^2 / \log_2 n$. Then for large n , G contains a maximal independent set C of size at most $(1 + \delta_2) \log_2 n$ with probability at least $(1 - n^{-3})^2(1 - p(n))$, where

$$2^{-(10 \log_2^2 \log_2 n)} \leq p(n) \leq \left(\frac{c}{\log_2 n}\right)^{30 \log_2 \log_2 n}$$

and c is a constant. Note that during the process of choosing C no information is used about the graph spanned by $V(G) - C$.

Proof. We construct a maximal independent set S iteratively in three phases, as in the proof of Lemma 8. Phases I and II are identical to those of Lemma 8.

Phase III. This phase occurs while $|\overline{N}[S]| \leq \frac{180 \ln n}{\ln \ln n}$. We choose any maximal independent set C in $\overline{N}[S]$ to add to S . As $\overline{N}[S]$ spans a random graph with edge-density $1/2$, it does not have an independent set of size $10 \log_2 \log_2 n$ with probability at least

$$1 - \binom{|\overline{N}[S]|}{10 \log_2 \log_2 n} 2^{-(10 \log_2^2 \log_2 n)}.$$

Thus, $|C| \leq 10 \log_2 \log_2 n$ with high probability. Let $p(n) := \binom{|\overline{N}[S]|}{10 \log_2 \log_2 n} 2^{-(10 \log_2^2 \log_2 n)}$.

First,

$$\begin{aligned} \binom{|\overline{N}[S]|}{10 \log_2 \log_2 n} 2^{-(10 \log_2^2 \log_2 n)} &\leq \left(\frac{e \cdot \frac{180 \ln n}{\ln \ln n}}{10 \log_2 \log_2 n}\right)^{10 \log_2 \log_2 n} \cdot 2^{-(10 \log_2^2 \log_2 n)} \\ &\leq \left[\frac{e \cdot \frac{180 \ln n}{\ln \ln n}}{10 \log_2 \log_2 n} \cdot 2^{-4 \log_2 \log_2 n}\right]^{10 \log_2 \log_2 n} \\ &= \left[\frac{e \cdot \frac{180 \ln n}{\ln \ln n}}{10 \log_2 \log_2 n} \cdot \frac{1}{(\log_2 n)^4}\right]^{10 \log_2 \log_2 n} \leq \left(\frac{c}{\log_2 n}\right)^{30 \log_2 \log_2 n}. \end{aligned}$$

Second,

$$\binom{|\overline{N}[S]|}{10 \log_2 \log_2 n} 2^{-(10 \log_2^2 \log_2 n)} \geq 2^{-(10 \log_2^2 \log_2 n)}.$$

$$\text{So } 2^{-(10 \log_2^2 \log_2 n)} \leq p(n) \leq \left(\frac{c}{\log_2 n}\right)^{30 \log_2 \log_2 n}.$$

Thus, after the completion of the three phases, S is a maximal independent set of G of size at most $\log_{1/(1/2+\alpha)} n + (3 \log_{3/2} \ln n + 1) \ln \ln n + 10 \log_2 \log_2 n$ with probability greater than $(1 - n^{-3})^2(1 - p(n))$. Note that

$$\log_{\frac{1}{1/2+\alpha}} n + (3 \log_{3/2} \ln n + 1) \ln \ln n + 10 \log_2 \log_2 n \leq (1 + \delta_2) \log_2 n$$

when $\delta_2 = 10(\ln \ln n)^2 / \log_2 n$ and n is large. Hence the lemma is proved. $\square$

Proof of Theorem 3. We create a first-fit partition by iteratively removing small maximal independent sets S_i generated by Lemmas 8 and 9 from G . Set $G_1 = G$, and for $i > 1$, $G_i = G - \cup_{j \leq i} S_j$. Let $n_i = |V(G_i)|$. We use the following important observation that G_i is a random graph distributed as $G_{n_i, 1/2}$ since the edges of G_i were never considered when the independent sets $S_1, S_2, \dots, S_i$ were constructed.

We iterate the process until $|V(G_t)| < n^{1/2} \log_2 n$. After that point, we greedily partition the remaining vertices into independent sets. Clearly, $\chi_{FF}(G) \geq t$.

By Lemma 8, there exists a maximal independent set S_i in G_i of size at most $(1 + \delta_1) \log_2 n_i$ with probability greater than $1 - n_i^{-2}$. Note that

$$\begin{aligned} \mathbb{P}[|S_i| < (1 + \delta_1) \log_2 n_i \text{ for all } 1 \leq i \leq t] &\geq 1 - t \left(n^{1/2} \log_2 n \right)^{-2} \\ &\geq 1 - n \left(\frac{1}{n (\log_2 n)^2} \right) = 1 - \frac{1}{(\log_2 n)^2}. \end{aligned}$$

Hence, with high probability,

$$t \geq \frac{n - n^{1/2} \log_2 n}{(1 + \delta_1) \log_2 n} = \frac{1}{1 + \delta_1} \left(\frac{n}{\log_2 n} \right) - \frac{1}{1 + \delta_1} n^{1/2} > (1 - 2\delta_1) \frac{n}{\log_2 n} > \frac{n}{2 \log_2 n}.$$

We will improve this lower bound on t by using Lemma 9 to show that many of the S_i are of a smaller size. However, since the probability in Lemma 9 is smaller, we can no longer use the union-sum bound as above. Instead, we will use Chernoff's inequality.

Let X_i be the indicator variable for the event that in Phase III during the construction of our maximal independent set S_i , G_i contains a maximal independent set of size larger than $10 \log_2 \log_2 n_i$. Then $X_i = 1$ with probability at most $p_i = p(n_i)$.

Next, we want to apply the following version of Chernoff's inequality to prove that, with high probability, $\sum_{i=1}^t X_i \leq 2pt$, where $p = \frac{\sum_{i=1}^t p_i}{t}$.

LEMMA 10 (see [1, pp. 263–265]). *Let $p_1, \dots, p_n \in [0, 1]$, and $\mathbb{P}[X_i = 1 - p_i] = p_i, \mathbb{P}[X_i = -p_i] = 1 - p_i$. If $X = \sum_{i=1}^n X_i$, then*

$$\mathbb{P}[X > a] < \exp(-2a^2/n).$$

Note that if $\mathbb{P}[X_i = 1]$ is smaller, then (10) remains true. So to get an upper bound of $\sum_{i=1}^t X_i$ with high probability, in the following proof we may assume that $\mathbb{P}[X_i = 1] = p_i = p(n_i)$.

Proof. Let $Y_i = X_i - p_i$ for $i = 1, \dots, t$. Then $\mathbb{P}[Y_i = 1 - p_i] = p_i$ and $\mathbb{P}[Y_i = -p_i] = 1 - p_i$. By Lemma 10 we have $\mathbb{P}[\sum_{i=1}^t Y_i > pt] < \exp(-2(pt)^2/t) = \exp(-2p^2t)$. Hence $\mathbb{P}[X > 2pt] < \exp(-2p^2t)$.

Since

$$p = \frac{\sum_{i=1}^t p_i}{t} \geq \frac{\sum_{i=1}^t 2^{-(10 \log_2 \log_2 n_i)}}{t} \geq 2^{-(10 \log_2 \log_2 n)} \geq 2^{-50(\log_2 \log_2 n)^2}$$

and $t \geq \frac{n}{2 \log_2 n}$,

$$\begin{aligned} \mathbb{P}[X > 2pt] &< \exp(-2p^2t) < \exp\left(-2 \cdot 2^{-100(\log_2 \log_2 n)^2} \cdot \frac{n}{2 \log_2 n}\right) \\ &= \exp(-2^{-100(\log_2 \log_2 n)^2 + \log_2 n - \log_2 \log_2 n}) \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

So $\sum_{i=1}^t X_i \leq 2pt$ with high probability. $\square$

By Lemma 9, there exists a maximal independent set S_i in G_i of size at most $(1 + \delta_2) \log_2 n_i$ with probability greater than $(1 - n_i^{-3})^2(1 - p(n))$.

We have proved that with high probability, each S_i has size at most $(1 + \delta_1) \log_2 n_i$. Meanwhile, by the proof above we know that among these S_i 's, at most $2pt$ of them can have size greater than $(1 + \delta_2) \log_2 n_i$. Thus,

$$2pt \cdot (1 + \delta_1) \log_2 n + (t - 2pt)(1 + \delta_2) \log_2 n \geq n - n^{1/2} \log_2 n.$$

Hence,

$$\begin{aligned} \chi_{FF}(G) \geq t &\geq \frac{n - n^{1/2} \log_2 n}{[2p(1 + \delta_1) + (1 - 2p)(1 + \delta_2)] \log_2 n} \\ &= \left[\frac{1}{1 + \delta_2} \left(\frac{n}{\log_2 n} \right) - \frac{1}{1 + \delta_2} n^{1/2} \right] \frac{1 + \delta_2}{1 + \delta_2 + 2p(\delta_1 - \delta_2)} \\ &> (1 - 2\delta_2) \frac{n}{\log_2 n} \left(\frac{1 + \delta_2}{1 + \delta_2 + 2p\delta_1} \right) \\ &> (1 - 2\delta_2)(1 - 2p\delta_1) \frac{n}{\log_2 n} > (1 - 2p\delta_1 - 2\delta_2) \frac{n}{\log_2 n}. \end{aligned}$$

Since each $n_i \geq n^{1/2} \log_2 n > n^{1/2}$, we have

$$p = \frac{\sum_{i=1}^t p_i}{t} \leq \left(\frac{c}{\log_2(\sqrt{n})} \right)^{30 \log_2 \log_2(\sqrt{n})}.$$

When n is large, we have

$$\left(\frac{c}{\log_2(\sqrt{n})} \right)^{30 \log_2 \log_2(\sqrt{n})} < \frac{\log_2 \log_2 n}{\sqrt{\log_2 n}}$$

and so $2p\delta_1 < 2 \cdot \log_2 \log_2 n / \sqrt{\log_2 n} \cdot (18 / \sqrt{\log_2 n}) < \delta_2$.

Hence $\chi_{FF}(G) > (1 - 3\delta_2) \frac{n}{\log_2 n}$ for $\delta_2 = 10(\ln \ln n)^2 / \log_2 n$ and as $n \rightarrow \infty$. $\square$

3.1. Concentration. We next consider the concentration of the first-fit chromatic number of random graphs. We omit the proof of the following easy statement.

PROPOSITION 11. (i) Let G be a graph and H be a subgraph of G formed by deleting one edge from G . Then $|\chi_{FF}(G) - \chi_{FF}(H)| \leq 1$.

(ii) Let G be a graph and H be a subgraph of G formed by deleting one vertex from G . Then $\chi_{FF}(H) \leq \chi_{FF}(G) \leq \chi_{FF}(H) + 1$.

Note that there are graphs whose first-fit chromatic number drops when an edge is added. The smallest such example is P_4 , which has $\chi_{FF}(P_4) = 3$, while $\chi_{FF}(C_4) = 2$.

Since the first-fit chromatic number satisfies the vertex Lipschitz condition, we obtain a tight concentration via the vertex exposure martingale and Azuma's inequality (see [1, pp. 95–96]).

THEOREM 12. Let $G \sim G_{n,p}$ for any probability p . Then

$$\mathbb{P} [|\chi_{FF}(G) - \mathbb{E}[\chi_{FF}(G)]| > \lambda \sqrt{n-1}] < 2e^{-\lambda^2/2}.$$

This concentration result is analogous to the result of Shamir and Spencer [1] about the tight concentration of the chromatic number of the random graph.

4. First-fit chromatic number of product of graphs. In this section we discuss the first-fit chromatic number of $G \square P_2$, $G \square P_3$, and $G \square P_4$.

THEOREM 13. *Let $\chi_{FF}(G)$ be even. Then*

$$\chi_{FF}(G \square P_3) \geq \chi_{FF}(G) + 2.$$

Proof. For simplicity let $V(P_3) := \{1, 2, 3\}$. Suppose $\chi_{FF}(G) = 2k$. Then there exists a first-fit partition of $V(G)$ into $2k$ nonempty sets $\langle V_1, V_2, \dots, V_{2k} \rangle$. We construct a partition $\langle U_1, \dots, U_{2k+2} \rangle$ of $V(G \square P_3)$ as follows:

1. For $1 \leq i \leq k-1$, let

$$\begin{aligned} U_{2i-1} &= \{(u, 1) \mid u \in V_{2i-1}\} \cup \{(v, 2) \mid v \in V_{2i}\} \cup \{(u, 3) \mid u \in V_{2i-1}\}, \\ U_{2i} &= \{(v, 1) \mid v \in V_{2i}\} \cup \{(u, 2) \mid u \in V_{2i-1}\} \cup \{(v, 3) \mid v \in V_{2i}\}. \end{aligned}$$

2. For $k \leq i \leq k+1$, choose two adjacent vertices $u \in V_{2k-1}$ and $v \in V_{2k}$. Let $U_{2k-1} = \{(u, 3), (v, 1)\}$, $U_{2k} = \{(u, 1), (v, 3)\}$, $U_{2k+1} = \{(u, 2)\}$, $U_{2k+2} = \{(v, 2)\}$.

3. Add the remaining vertices to the previous parts greedily.

Since the V_i ($1 \leq i \leq 2k$) are all independent, it is easy to see that all U_i ($1 \leq i \leq 2k+2$) are also independent.

For each $1 \leq j < i \leq k-1$ and each $u \in V_{2i-1}$, $v \in V_{2i}$, there exist vertices $x_1, y_1 \in V_{2j-1}$ and vertices $x_2, y_2 \in V_{2j}$ such that x_1, x_2 are neighbors of u and y_1, y_2 are neighbors of v . Thus by the construction, (u, a) is adjacent to $(x_1, a) \in U_{2j-1}$ and $(x_2, a) \in U_{2j}$ for $a = 1, 3$, and $(u, 2)$ is adjacent to $(x_2, 2) \in U_{2j-1}$ and $(x_1, 2) \in U_{2j}$.

Similarly, (v, a) is adjacent to $(y_1, a) \in U_{2j-1}$ and $(y_2, a) \in U_{2j}$ for $a = 1, 3$, and $(v, 2)$ is adjacent to $(y_2, 2) \in U_{2j-1}$ and $(y_1, 2) \in U_{2j}$.

Similarly, we can check that any vertex in $U_{2k-1}, U_{2k}, U_{2k+1}, U_{2k+2}$ has a neighbor in the lower indexed subsets.

Thus $\langle U_1, \dots, U_{2k+2} \rangle$ is a first-fit partition of $G \square P_3$, so $\chi_{FF}(G \square P_3) \geq \chi_{FF}(G) + 2$ if $\chi_{FF}(G)$ is even. $\square$

We next consider the first-fit chromatic number of $G \square P_4$.

THEOREM 14. *For every graph G ,*

$$\chi_{FF}(G \square P_4) \geq \chi_{FF}(G) + 2.$$

Proof. Suppose $\chi_{FF}(G) = t$.

If t is even, since $G \square P_3$ is an induced subgraph of $G \square P_4$, then $\chi_{FF}(G \square P_4) \geq \chi_{FF}(G \square P_3) \geq \chi_{FF}(G) + 2$.

If $t = 2k-1$ is odd, then we can make a partition similar to the one in the proof of Theorem 13 as follows. Let $\{1, 2, 3, 4\}$ be the vertex set of P_4 . Then we have the following.

1. For $1 \leq i \leq k-1$, let

$$\begin{aligned} U_{2i-1} &= \{(u, 1) \mid u \in V_{2i-1}\} \cup \{(v, 2) \mid v \in V_{2i}\} \\ &\quad \cup \{(u, 3) \mid u \in V_{2i-1}\} \cup \{(v, 4) \mid v \in V_{2i}\}, \\ U_{2i} &= \{(v, 1) \mid v \in V_{2i}\} \cup \{(u, 2) \mid u \in V_{2i-1}\} \\ &\quad \cup \{(v, 3) \mid v \in V_{2i}\} \cup \{(u, 4) \mid u \in V_{2i-1}\}. \end{aligned}$$

2. Choose a vertex $u \in V_{2k-1}$. Let $U_{2k-1} = \{(u, 1), (u, 4)\}$, $U_{2k} = \{(u, 3)\}$, $U_{2k+1} = \{(u, 2)\}$.

3. For the remaining vertices, add them to the previous parts greedily.

It is easy to check that the above partition is a first-fit partition following an argument similar to the one in the proof of Theorem 13. $\square$

Effantin and Kheddouci [6] proved the following theorem for products of even cycles. We extend this result to include odd cycles in the product.

THEOREM 15 (see [6]). $\chi_{FF}(C_{m_1} \square C_{m_2} \square \dots \square C_{m_k}) = 2k + 1$, where $k \geq 2$, $m_i \geq 4$, and m_i is even.

COROLLARY 16. Let $C_{m_1}, C_{m_2}, \dots, C_{m_k}$ be cycles of size $m_1, m_2, \dots, m_k$, respectively, where $k \geq 2$ and $m_i \geq 4$. Then

$$\chi_{FF}(C_{m_1} \square C_{m_2} \square \dots \square C_{m_k}) = 2k + 1.$$

Proof. The upper bound follows from the obvious degree bound, $\chi_{FF}(G) \leq \Delta(G) + 1$. For the lower bound, we first reorder m_i 's to place the even m_i 's first, followed by the odd ones.

Case 1. If no m_i is odd, then the corollary follows from Theorem 15.

Case 2. If we have $k_1 \geq 2$ even m_i 's, then the product graph of k_1 even cycles has first-fit chromatic number $2k_1 + 1$ by Theorem 15. Then for the following odd cycles, since P_4 is an induced subgraph of any odd cycle of length at least 4, we can repeatedly apply Theorem 14.

If we have at most one even cycle which is not C_4 , then the fact that $\chi_{FF}(C_n) = 3$ for $n \geq 5$ and repeatedly applying Theorem 14 prove the corollary. If the sole even cycle is C_4 , then we use the fact that $\chi_{FF}(C_4 \square P_4) = 5$ and again repeatedly apply Theorem 14. $\square$

Note that a statement for $G \square P_2$ analogous to Theorems 13 and 14 does not hold because $\chi_{FF}(P_2 \square P_2) = 2 = \chi_{FF}(P_2)$. How large can $\chi_{FF}(G \square P_2)$ be? In the following example we show that $\chi_{FF}(G \square P_2)$ can be as large as $2\chi_{FF}(G)$. Since $G \square P_2$ is an induced subgraph of $G \square P_3$ and $G \square P_4$, $\chi_{FF}(G \square P_3)$ and $\chi_{FF}(G \square P_4)$ can be as large as $2\chi_{FF}(G)$ as well.

PROPOSITION 17. Let G be a graph formed by removing a perfect matching M from K_{2m} . Then $\chi_{FF}(G \square P_2) = 2\chi_{FF}(G)$.

Proof. Suppose $V(K_{2m}) = \{u_1, \dots, u_{2m}\}$ and the deleted matching is $M = \{u_1 u_2, \dots, u_{2m-1} u_{2m}\}$. Consider a first-fit partition $\langle V_1, \dots, V_t \rangle$ of $V(G)$, where $\chi_{FF}(G) = t$. Without loss of generality, suppose $u_1 \in V_1$. Then except u_2 , all other vertices are adjacent to u_1 , so they cannot be in V_1 . Then u_2 must be in V_1 ; otherwise it does not have any neighbor in V_1 . So $V_1 = \{u_1, u_2\}$. Inductively, we can see that $V_i = \{u_{2i-1}, u_{2i}\}$. So $\chi_{FF}(G) = t = m$. For $G \square P_2$, we can make a partition as follows:

For $1 \leq i \leq t$, $U_{2i-1} = \{(u_{2i-1}, 1), (u_{2i}, 2)\}$, $U_{2i} = \{(u_{2i-1}, 2), (u_{2i}, 1)\}$, where $\{1, 2\}$ is the vertex set of P_2 .

So $\chi_{FF}(G \square P_2) \geq 2\chi_{FF}(G)$. For the upper bound, $\chi_{FF}(G \square P_2) \leq \Delta(G \square P_2) + 1 = (2m - 1) + 1 = 2m$. Hence $\chi_{FF}(G \square P_2) = 2\chi_{FF}(G)$ in this case. $\square$

We believe that $\chi_{FF}(G \square P_2) \leq 2\chi_{FF}(G)$ for every graph G .

Acknowledgments. The authors thank Zoltán Füredi for his encouragement and helpful discussions and Weiting Cao for helpful discussions during the early stages of this project. The authors also thank the anonymous referee for an improvement to Lemma 4 and for comments that improved the readability of the paper, and Alan Frieze for pointing out reference [2].

REFERENCES

- [1] N. ALON AND J. H. SPENCER, *The Probabilistic Method*, 2nd ed., John Wiley and Sons, New York, 2000.
- [2] B. BOLLOBÁS AND P. ERDŐS, *Cliques in random graphs*, Math. Proc. Cambridge Philos. Soc., 80 (1976), pp. 419–427.
- [3] B. BOLLOBÁS, *Random Graphs*, 2nd ed., Cambridge Stud. Adv. Math. 73, Cambridge University Press, Cambridge, UK, 2001.
- [4] G. R. BRIGHTWELL, H. A. KIERSTEAD, AND W. T. TROTTER, *A Note on First-Fit Coloring of Interval Graphs*, preprint, 2003.
- [5] C. A. CHRISTEN AND S. M. SELKOW, *Some perfect coloring properties of graphs*, J. Combin. Theory Ser. B, 27 (1979), pp. 49–59.
- [6] B. EFFANTIN AND H. KHEDDOUCI, *Grundy number of graphs*, Discuss. Math. Graph Theory, 27 (2007), pp. 5–18.
- [7] P. ERDŐS, W. R. HARE, S. T. HEDETNIEMI, AND R. LASKAR, *On the equality of Grundy and ochromatic number of graphs*, J. Graph Theory, 11 (1987), pp. 157–159.
- [8] P. ERDŐS, S. T. HEDETNIEMI, R. C. LASKAR, AND C. E. PRINS, *On the equality of the partial Grundy and upper ochromatic numbers of graphs*, Discrete Math., 272 (2003), pp. 53–64.
- [9] Z. FÜREDI, A. GYÁRFÁS, G. N. SÁRKÖZY, AND S. SELKOW, *Inequalities for the first-fit chromatic number*, J. Graph Theory, to appear.
- [10] C. GERMAINIAN AND H. KHEDDOUCI, *Grundy coloring for power graphs*, Electron. Notes Discrete Math., 15 (2003), pp. 65–67.
- [11] G. R. GRIMMETT AND C. J. H. MCDIARMID, *On colouring random graphs*, Math. Proc. Cambridge Philos. Soc., 77 (1975), pp. 313–324.
- [12] D. G. HOFFMAN AND P. D. JOHNSON, JR., *Greedy colorings and the Grundy chromatic number of the n -cube*, Bull. Inst. Combin. Appl., 26 (1999), pp. 49–57.
- [13] P. M. GRUNDY, *Mathematics and games*, Eureka, 2 (1939), 6C-8.
- [14] S. IRANI, *Coloring inductive graphs on-line*, Algorithmica, 11 (1994), pp. 53–72.
- [15] H. A. KIERSTEAD, *The linearity of first-fit coloring of interval graphs*, SIAM J. Discrete Math., 1 (1988), pp. 526–530.
- [16] H. A. KIERSTEAD AND J. QIN, *Coloring interval graphs with first-fit*, Discrete Math., 144 (1995), pp. 47–57.
- [17] H. A. KIERSTEAD AND W. T. TROTTER, *private communication*.
- [18] C. J. H. MCDIARMID, *Colouring random graphs badly*, in Graph Theory and Combinatorics, Res. Notes in Math. 34, R. J. Wilson, ed., Pitman, Boston, 1979, pp. 76–86.
- [19] S. PEMMARAJU, R. RAMAN, AND K. VARADARAJAN, *Buffer minimization using max-coloring*, in Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms, New Orleans, LA, SIAM, Philadelphia, 2004, pp. 562–571.
- [20] D. R. WOODALL, *Problem No. 4*, in Proceedings of the British Combinatorial Conference 1973, London Math. Soc., Lecture Note Series 13, Cambridge University Press, 1974.
- [21] M. ZAKER, *Results on the Grundy chromatic number of graphs*, Discrete Math., 306 (2006), pp. 3166–3173.

NONLINEAR MATROID OPTIMIZATION AND EXPERIMENTAL DESIGN*

Yael Bernstein[†], Jon Lee[‡], Hugo Maruri-Aguilar[§], Shmuel Onn[†],
Eva Riccomagno[¶], Robert Weismantel^{||}, and Henry Wynn[§]

Abstract. We study the problem of optimizing nonlinear objective functions over matroids presented by oracles or explicitly. Such functions can be interpreted as the balancing of multicriteria optimization. We provide a combinatorial polynomial time algorithm for arbitrary oracle-presented matroids, that makes repeated use of matroid intersection and an algebraic algorithm for vectorial matroids. Our work is partly motivated by applications to minimum-aberration model-fitting in experimental design in statistics, which we discuss and demonstrate in detail.

Key words. matroid, matroid intersection, Binet–Cauchy identity, spanning tree, matching, exact matching, experimental design, aberration, algebraic statistics, learning, interpolation, regression, nonlinear discrete optimization, combinatorial optimization, integer programming

AMS subject classifications. 05A, 15A, 51M, 52A, 52B, 52C, 62H, 62K, 65C, 68Q, 68R, 68U, 68W, 90B, 90C

DOI. 10.1137/070696465

1. Introduction. In this article, partly motivated by applications to minimum-aberration model-fitting in experimental design, which will be discussed briefly at the end of this introduction and in detail in section 5, we study the problem of optimizing an arbitrary nonlinear function over a matroid as follows.

Nonlinear matroid optimization. Given matroid M on ground set $N := \{1, \dots, n\}$, integer weight vectors $w_i = (w_{i,1}, \dots, w_{i,n}) \in \mathbb{Z}^n$ for $i = 1, \dots, d$, and function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$, find a matroid base $B \in \mathcal{B}(M) \subset 2^N$ minimizing the “balancing” by f of

*Received by the editors July 6, 2007; accepted for publication (in revised form) February 11, 2008; published electronically May 16, 2008.

<http://www.siam.org/journals/sidma/22-3/69646.html>

[†]Technion - Israel Institute of Technology, 32000 Haifa, Israel (yaelber@tx.technion.ac.il, onn@ie.technion.ac.il). The first author was supported by an Irwin and Joan Jacobs Scholarship and by a scholarship from the Graduate School of the Technion. The fourth author was partially supported by the Joan and Reginald Coleman-Cohen Exchange Program during a stay of Henry Wynn at the Technion, was partially supported by the Mathematisches Forschungsinstitut Oberwolfach during a stay within the Research in Pairs Programme, and was also supported by the ISF - Israel Science Foundation.

[‡]IBM T.J. Watson Research Center, Yorktown Heights, NY 10598 (jonlee@us.ibm.com). This author was partially supported by the Mathematisches Forschungsinstitut Oberwolfach during a stay within the Research in Pairs Programme.

[§]London School of Economics, London WC2A 2AE, UK (h.maruri-aguilar@lse.ac.uk, h.wynn@lse.ac.uk). The third and last authors were supported by the Research Councils UK (RCUK) Basic Technology grant “Managing Uncertainty in Complex Models.” The last author was partially supported by the Joan and Reginald Coleman-Cohen Exchange Program during a stay of his at the Technion.

[¶]Università degli Studi di Genova, 16146 Genova, Italy (riccomagno@dima.unige.it).

^{||}Otto-von-Guericke Universität Magdeburg, D-39106 Magdeburg, Germany (weismantel@imo.math.uni-magdeburg.de). This author was partially supported by the Mathematisches Forschungsinstitut Oberwolfach during a stay within the Research in Pairs Programme and was also supported by the European TMR Network ADONET 504438.

the d weights $w_i(B) := \sum_{j \in B} w_{i,j}$ of base B ,

$$f(w_1(B), \dots, w_d(B)) = f\left(\sum_{j \in B} w_{1,j}, \dots, \sum_{j \in B} w_{d,j}\right).$$

(All necessary basics about matroid theory are provided in section 2.1. For more details consult [14] or [17].)

Nonlinear matroid optimization can be interpreted as *multicriteria matroid optimization*: the d given weight vectors $w_1, \dots, w_d$ represent d different criteria, where the value of base $B \in \mathcal{B}(M)$ under criterion i is its i th weight $w_i(B) = \sum_{j \in B} w_{i,j}$, and where the objective is to minimize the “balancing” $f(w_1(B), \dots, w_d(B))$ of the d given criteria by the given function f .

In fact, we have a hierarchy of problems of increasing generality, parameterized by the number d of weight vectors. At the bottom of the hierarchy lies standard linear matroid optimization, recovered with $d = 1$ and f being the identity on $\mathbb{Z}$. At the top of the hierarchy lies the problem of maximizing an arbitrary function over the set of bases, with $d = n$ and $w_i = \mathbf{1}_i$, the i th standard unit vector in $\mathbb{Z}^n$ for all i ; see Proposition 2.4.

It will often be convenient to collect the weight vectors in a $d \times n$ matrix W . Thus, the i th row of this matrix is the i th weight vector $w_i = (w_{i,1}, \dots, w_{i,n})$. For each subset $S \subseteq N$ we define its W -image to be

$$W(S) := (w_1(S), \dots, w_d(S)) := \left(\sum_{j \in S} w_{1,j}, \dots, \sum_{j \in S} w_{d,j}\right) \in \mathbb{Z}^d,$$

which is the vector giving the value of S under each of the d weight vectors. The nonlinear matroid optimization then asks for a matroid base $B \in \mathcal{B}(M)$ minimizing the objective function $f(W(B))$.

The computational complexity of nonlinear matroid optimization depends on the number d of weight vectors, on the representation of weights (binary $\langle w_{i,j} \rangle$ versus unary $|w_{i,j}|$; see section 2.1), on the type of function f and its presentation, and on the type of matroid and its presentation. We will be able to handle an arbitrary function f presented by a *comparison oracle* that, queried on $u, v \in \mathbb{Z}^d$, asserts whether or not $f(u) \leq f(v)$, and an arbitrary matroid presented by an *independence oracle*, that, queried on $I \subseteq N$, asserts whether or not I is *independent* in M , that is, whether $I \subseteq B$ for some base $B \in \mathcal{B}(M)$; see section 2.1. These are very broad presentations that reveal little information (per query) on the function and matroid, making our problem setting rather expressive but difficult for achieving strong complexity results.

Standard linear matroid optimization is well known to be efficiently solvable by the greedy algorithm. Nonlinear matroid optimization with d fixed and f *concave* turns out to be solvable in polynomial time as well [8, 11]. In fact, using sophisticated geometric methods, the nonlinear optimization problem with d fixed and f concave has been recently shown to be efficiently solvable for bipartite matching [4] and for broader classes of combinatorial optimization problems [12]. Therein, maximization rather than minimization form is used; hence, convex rather than concave functions are considered.

However, generally, nonlinear matroid optimization is intractable, even for uniform matroids. In particular, if d is variable, then exponential time is needed even for $\{0, 1\}$ -valued weights, and if the weights are encoded in binary, then exponential

time is needed even for fixed dimension $d = 1$. See Propositions 2.3, 2.4, and 2.5 in what follows for various intractability statements.

In spite of these difficulties, we establish here the efficient solvability of the problem as follows.

THEOREM 1.1. *For every fixed d and p , there is an algorithm that, given a matroid M presented by an independence oracle on the n -element ground set N , integers $a_1, \dots, a_p \in \mathbb{Z}$, weight vectors $w_1, \dots, w_d \in \{a_1, \dots, a_p\}^n$, and function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ presented by a comparison oracle, solves the nonlinear matroid optimization problem in time that is polynomial in n and $\max\langle a_i \rangle$.*

We also state the following natural immediate corollary concerning $\{0, 1, \dots, p\}$ -valued weights.

COROLLARY 1.2. *For every fixed d and p , there is an algorithm that, given n -element matroid M presented by an independence oracle, $w_1, \dots, w_d \in \{0, 1, \dots, p\}^n$ and function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ presented by a comparison oracle, solves the nonlinear matroid optimization problem in time polynomial in n .*

The algorithm establishing Theorem 1.1 is combinatorial and makes repeated use of matroid intersection (see, e.g., [9] or [10], and see [7] for another recent interesting application of matroid intersection). In fact, it invokes the matroid intersection algorithm roughly n^{p^d} times, and hence it is quite heavy. However, most matroids appearing in practice, including graphic matroids and those arising in the applications to experimental design to be discussed later, are vectorial. Therefore, we also provide another more efficient, linear-algebraic algorithm for vectorial matroids. Moreover, this algorithm applies to weights with an unlimited number (rather than a fixed number p) of different values $w_{i,j}$ of entries.

THEOREM 1.3. *For every fixed d , there is an algorithm that, given integer $m \times n$ matrix A , weight vectors $w_1, \dots, w_d \in \mathbb{Z}^n$, and function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ presented by a comparison oracle, solves the nonlinear optimization problem over the (real) vectorial matroid of A in time polynomial in $\langle A \rangle$ and $\max |w_{i,j}|$.*

A specific application that can be solved by either the combinatorial algorithm underlying Theorem 1.1 or the more efficient linear-algebraic algorithm underlying Theorem 1.3 is the following example.

Example 1.4 (minimum-norm spanning tree). Fix any positive integer d . Let G be any connected graph with edge set $E = \{e_1, \dots, e_n\}$, and let $w_1, \dots, w_d \in \mathbb{Z}^n$ be weight vectors with $w_{i,j}$ representing the values of edge e_j under the i th criterion. Let A be the vertex-arc incidence matrix of an arbitrary orientation of G . Then the vectorial matroid of A is the graphic matroid of G whose bases are the spanning trees of G . Now fix also any q that is either a positive integer or ∞ , and let $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ be the l_q norm given by $\|u\|_q := (\sum_{i=1}^d |u_i|^q)^{\frac{1}{q}}$ for finite q and $\|u\|_\infty := \max_{i=1}^d |u_i|$. Note that a comparison oracle for $f(u) = \|u\|_q$ is easily and efficiently realizable. Then Theorems 1.1 and 1.3 assure that a spanning tree T of G minimizing the l_q norm of the multicriteria vector, given by

$$\|(w_1(T), \dots, w_d(T))\|_q = \|W(T)\|_q$$

is computable in time polynomial in n and $\max |w_{i,j}|$. Note that if $P \neq NP$, then the problem is *not* solvable in time polynomial in the binary length $\langle w_{i,j} \rangle$ even for the graph obtained from a path by replacing every edge by two parallel copies; see Proposition 2.5.

Experimental design. We conclude the introduction with a brief discussion of the application to experimental design, elaborated in detail in section 5. The general

framework is as follows. We are interested in learning an unknown system whose output y is an unknown function Φ of a multivariate input $x = (x_1, \dots, x_k) \in \mathbb{R}^k$. It is customary to call the input variables x_j *factors* of the system. In order to learn the system, we perform several experiments. Each experiment i is determined by a point $p_i = (p_{i,1}, \dots, p_{i,k})$ and consists of feeding the system with input $x := p_i$ and measuring the corresponding output $y_i = \Phi(p_i)$. Based on these experiments, we wish to *fit a model* for the system, namely, determine an estimation $\hat{\Phi}$ of the function Φ , that satisfies the following properties:

- It lies in a prescribed class of functions.
- It is consistent with the outcomes of the experiments.
- It minimizes the *aberration*—a suitable criterion—among models in the class.

More detailed discussion and precise definitions will be given in section 5. We have the following broad corollary of Theorems 1.1 and 1.3; see section 5 for the precise statement and its various practical specializations to concrete aberration criteria useful in optimal model-fitting in experimental design.

COROLLARY 1.5. *An aberration-minimum multivariate-polynomial model is polynomial time computable.*

The article proceeds as follows. In section 2 we set some notation and preliminaries, make some preparations for the algorithms in the following sections, and demonstrate various intractability limitations on nonlinear matroid optimization. In section 3 we discuss arbitrary matroids presented by oracles, and provide the combinatorial algorithm for nonlinear matroid optimization, thereby establishing Theorem 1.1. In section 4 we provide the more efficient algebraic algorithm for nonlinear optimization over vectorial matroids, thereby proving Theorem 1.3. We conclude in section 5 with a detailed discussion of the experimental design application and prove Corollary 5.2 (a refined version of Corollary 1.5) and its various practical specializations. Readers interested mostly in experimental design can go directly to section 5, where the minimum-aberration model-fitting problem is reduced to nonlinear optimization over a suitable matroid, and where each of the algorithms developed in sections 3 and 4 can be invoked as a black box.

2. Preliminaries, preparation, and limitations.

2.1. Preliminaries. We use $\mathbb{R}$ for the reals, $\mathbb{Z}$ for the integers, and $\mathbb{N}$ for the nonnegative integers. The i th standard unit vector in $\mathbb{R}^n$ is denoted by $\mathbf{1}_i$. The *support* of $x \in \mathbb{R}^n$ is the index set $\text{supp}(x) := \{j : x_j \neq 0\}$ of nonzero entries of x . The integer lattice $\mathbb{Z}^n$ is naturally embedded in $\mathbb{R}^n$. Vectors will be interpreted as either row or column vectors interchangeably—this will be relevant only when such vectors are multiplied by matrices—in which case the correct interpretation will be clear from the context. We often collect a sequence of vectors designated by a lower-case letter as the rows of a matrix designated by the corresponding upper case letter. Thus, our weight vectors $w_i = (w_{i,1}, \dots, w_{i,n})$, $i = 1, \dots, d$ are arranged in a $d \times n$ matrix W , and our design points $p_i = (p_{i,1}, \dots, p_{i,k})$, $i = 1, \dots, m$ are arranged in an $m \times k$ matrix P . The space $\mathbb{R}^n$ is endowed with the standard inner product which, for $w, x \in \mathbb{R}^n$, is $w \cdot x := \sum_{i=1}^n w_i x_i$. Vectors w in $\mathbb{R}^n$ are also regarded as linear functions on $\mathbb{R}^n$ via the inner product $w \cdot x$. Therefore, we refer to elements of $\mathbb{R}^n$ as points, vectors, or linear functions, as is appropriate from the context. We write $\mathbf{1}$ for the vector with all entries equal to 1, with the dimension being clear from the context.

Our algorithms are analyzed for rational data only, and the time complexity is as in the standard Turing machine model; see, e.g., [1, 6, 16]. The input typically

consists of rational (usually integer) numbers, vectors, matrices, and finite sets of such objects. The *binary length* of an integer number $z \in \mathbb{Z}$ is defined to be the number of bits in its binary representation, $\langle z \rangle := 1 + \lceil \log_2(|z| + 1) \rceil$ (with the extra bit for the sign). The length of a rational number presented as a fraction $r = \frac{p}{q}$ with $p, q \in \mathbb{Z}$ is $\langle r \rangle := \langle p \rangle + \langle q \rangle$. The length of an $m \times n$ matrix A (or a vector) is the sum $\langle A \rangle := \sum_{i,j} \langle a_{i,j} \rangle$ of the lengths of its entries. Note that the length of A is no smaller than the number of entries, $\langle A \rangle \geq mn$. Therefore, when A is, say, part of an input to an algorithm, with m and n not fixed, the length $\langle A \rangle$ already incorporates mn , and so we will typically not account additionally for m, n directly. But sometimes we will also emphasize n as part of the input length. Similarly, the length of a finite set S of numbers, vectors, or matrices is the sum of lengths of its elements and hence, since $\langle S \rangle \geq |S|$, automatically accounts for its cardinality. Some input numbers affect the running time of some algorithms through their unary presentation, resulting in so-called “pseudo-polynomial” running time. The *unary length* of an integer number $z \in \mathbb{Z}$ is the number $|z| + 1$ of bits in its unary representation (again, an extra bit for the sign). The unary length of a rational number, vector, matrix, or finite set of such objects is defined again as the sums of lengths of their numerical constituents, and is again no smaller than the number of such numerical constituents. Often part of the input is presented by oracles. Then the running time counts also the number of oracle queries. An oracle algorithm is *polynomial time* if its running time, including the number of oracle queries, is polynomial in the length of the input.

Next, we make some basic definitions concerning matroids and set some associated notation; for matroid theory, see [14] or [17]. A *matroid* M is described by giving a finite *ground set* $\mathcal{E}(M)$ and a nonempty set $\mathcal{B}(M)$ of subsets of $\mathcal{E}(M)$ called the set of *bases* of M , such that for every $B, B' \in \mathcal{B}(M)$ and $i \in B \setminus B'$ there is an $i' \in B'$ such that $B \setminus \{i\} \cup \{i'\} \in \mathcal{B}(M)$. A subset I of $\mathcal{E}(M)$ is *independent* in M if $I \subseteq B$ for some base $B \in \mathcal{B}(M)$. The set of independent sets of M is denoted by $\mathcal{I}(M)$. Often it is convenient to let $\mathcal{E}(M) = N = \{1, 2, \dots, n\}$ for some positive integer n . The *rank* $\text{rank}(S)$ of a subset $S \subseteq \mathcal{E}(M)$ is the maximum cardinality $|I|$ of an $I \in \mathcal{I}(M)$ that is contained in S . The *rank* of the matroid is $\text{rank}(M) := \text{rank}(\mathcal{E}(M))$ and is equal to the cardinality $|B|$ of every base $B \in \mathcal{B}(M)$. An *independence oracle* for M is one that, queried on $I \subseteq \mathcal{E}(M)$, asserts whether I is in $\mathcal{I}(M)$. An independence oracle allows us to compute the rank of every $S \subseteq \mathcal{E}(M)$ as follows. Start with $I := \emptyset$. For each $j \in S$ (in any order) do: if $I \cup \{j\}$ is in $\mathcal{I}(M)$, then set $I := I \cup \{j\}$. Output $\text{rank}(S) := |I|$.

An important example of a matroid is the *graphic matroid* $M(G)$ of a graph G with $\mathcal{E}(M(G))$ being the edge set of G and $\mathcal{B}(M(G))$ equal to the set of edge sets of spanning forests of G . A further key example is the *vectorial matroid* of an $m \times n$ matrix A (over a field $\mathbb{F}$) with $\mathcal{B}(M)$ the set of subsets of indices of maximal linearly-independent subsets of columns of A . Throughout, when treating complexity issues, we assume that $\mathbb{F}$ is the field $\mathbb{Q}$ of rationals, that A has integer components, and that we do arithmetic over $\mathbb{Q}$. A *uniform matroid* is any matroid that is isomorphic to the matroid $\mathcal{U}_{m,n}$ having ground set $N = \{1, \dots, n\}$ and having all m -subsets of N as bases. If M_1 and M_2 are matroids with disjoint ground sets, then their *direct sum* $M_1 \oplus M_2$ has $\mathcal{E}(M_1 \oplus M_2) := \mathcal{E}(M_1) \uplus \mathcal{E}(M_2)$ and $\mathcal{B}(M_1 \oplus M_2) := \{B_1 \uplus B_2 : B_1 \in \mathcal{B}(M_1), B_2 \in \mathcal{B}(M_2)\}$. A *partition matroid* is any matroid that is a direct sum of uniform matroids $\oplus_{i=1}^r \mathcal{U}_{m_i, n_i}$ (with the ground sets of the $\mathcal{U}_{m_i, n_i}$ labeled to be pairwise disjoint). If $m_i = m$ and all $n_i = k$ for all i , then we write $\mathcal{U}_{m,k}^r$ for this r -fold sum of $\mathcal{U}_{m,k}$.

It is well known that all graphic matroids and uniform matroids are vectorial. Furthermore, any partition matroid that is the direct sum $\oplus_{i=1}^r \mathcal{U}_{1,n_i}$ of rank-1 uniform matroids is graphic: it is the matroid of the graph obtained from any forest on r edges by replacing the i th edge by n_i parallel copies for $i = 1, \dots, r$. In particular, $\mathcal{U}_{1,k}^r$ is a graphic matroid for any positive k and r .

2.2. Preparation. Here we provide preparatory ingredients which will be used in algorithms in sections 3 and 4. First, we record for later the following statement, which follows directly from the definitions.

PROPOSITION 2.1. *Consider n -element matroid M , weights $w_1, \dots, w_d \in \mathbb{Z}^n$, and $f : \mathbb{Z}^d \rightarrow \mathbb{R}$. Put*

$$U := \{W(B) : B \in \mathcal{B}(M)\} = \{u \in \mathbb{Z}^d : u = W(B) \text{ for some base } B \in \mathcal{B}(M)\}.$$

Then the optimal objective value of the corresponding nonlinear matroid optimization problem satisfies

$$f^* = \min \{f(u) : u \in U\}.$$

In light of this viewpoint, with respect to the definitions of U and f^* from Proposition 2.1, an *optimal W -image* is any $u \in U$ satisfying $f(u) = f^*$.

Thus, the problem of computing f^* reduces to that of constructing the set U of W -images of bases, and then extracting an optimal W -image from U . This really is the crucial component of the solution of the nonlinear matroid optimization problem. While the cardinality of $U = \{W(B) : B \in \mathcal{B}(M)\}$ may be polynomial under suitable assumptions on the data, its direct computation by computing $W(B)$ for every base is prohibitive since the number of matroid bases is typically exponential in n . Instead, we will construct a finite superset Z of “potential” W -images of bases, satisfying $U \subseteq Z \subseteq \mathbb{Z}^d$, and then filter U out of Z . However, this is not an easy task either: deciding if a given $u \in \mathbb{Z}^d$ satisfies $u = W(B)$ for some base B is NP-complete already for fixed $d = 1$ and uniform matroids or partition matroids that are the direct sums of rank-1 uniform matroids; see Proposition 2.5. The way the filtration of U out of Z is done is precisely the key difference between our two algorithms for nonlinear matroid optimization in sections 3 and 4.

Next, we demonstrate that finding an optimal base for a nonlinear matroid optimization problem can be reduced to finding an optimal W -image for a small number of subproblems. Consider data for a nonlinear matroid optimization problem, consisting of a matroid M , weight vectors $w_1, \dots, w_d \in \mathbb{Z}^n$, and function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$. Each subset $S \subseteq N$ gives a *subproblem* of nonlinear matroid optimization as follows. The matroid of the subproblem is the *restriction* of M to S ; that is, the matroid $M.S$ on ground set S in which a subset $I \subseteq S$ is independent if and only if it is independent in M . Note that an independence oracle for the restriction matroid $M.S$ is realizable at once from that of M . The weight vectors of the subproblem are the restrictions of the original weight vectors to S . The function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ in the subproblem is the same as in the original problem. We have the following useful statement.

LEMMA 2.2. *The nonlinear matroid optimization problem of finding an optimal base of an n -element matroid is reducible in time polynomial in n to finding an optimal W -image for $n + 1$ subproblems.*

Proof. Denote by $u^*(S)$ the optimal W -image for the subproblem on $S \subseteq N$. Now compute an optimal base for the original problem, by computing an optimal W -image for $n + 1$ such subproblems as follows.

```

Start with  $S := N$ ;
Compute  $m := \text{rank}(S)$ ;
Compute an optimal  $W$ -image  $u^* := u^*(N)$  of the original problem;
for  $j=1, 2, \dots, n$  do
    Compute  $\text{rank}(S \setminus \{j\})$ ;
    Compute an optimal  $W$ -image  $u^*(S \setminus \{j\})$ ;
    if  $\text{rank}(S \setminus \{j\}) = m$  and  $f(u^*(S \setminus \{j\})) = f(u^*)$  then set  $S := S \setminus \{j\}$ ;
end
return  $B := S$ ;

```

It is not hard to verify that the set B obtained is indeed an optimal base for the original problem. $\square$

2.3. Limitations. Here we provide various intractability statements about the nonlinear matroid optimization problem and some of its relatives. In particular, we show that, if the weights are encoded in binary, then even the 1-dimensional problem, namely with fixed $d = 1$, over any matroid given explicitly or by an oracle, requires examining the objective value of *every* base, and hence cannot be solved in polynomial time.

It is convenient to define a certain class of rank k matroids on $2k$ elements, so as to generalize the uniform matroid $\mathcal{U}_{k,2k}$ and the partition matroids $\mathcal{U}_{1,2}^k$. For some r , $1 \leq r \leq k$, partition $K = \{1, \dots, k\}$ into r parts as $\uplus_{i=1}^r K_i$. Correspondingly, let $\bar{K}_i := \{\bar{j} : j \in K_i\}$. Let $k_i = |K_i| = |\bar{K}_i|$. For $i = 1, \dots, r$, let $\mathcal{M}_i$ be a uniform matroid of rank k_i on ground set $K_i \cup \bar{K}_i$ (so $\mathcal{M}_i \cong \mathcal{U}_{k_i,2k_i}$). Let $\mathcal{M}_{r,k} := \oplus_{i=1}^r \mathcal{M}_i$. Observe that $\mathcal{M}_{1,k} \cong \mathcal{U}_{k,2k}$ (which is uniform) and $\mathcal{M}_{k,k} \cong \mathcal{U}_{1,2}^k$ (which is graphic), so *hardness results with respect to the matroid classes $\mathcal{M}_{r,k}$ apply to uniform and graphic matroids*. We note and will soon use that for any r , $1 \leq r \leq k$, $|\mathcal{B}(\mathcal{M}_{r,k})|$ is not bounded by any polynomial in $|\mathcal{E}(\mathcal{M}_{r,k})| = 2k$.

PROPOSITION 2.3. *Computing an optimal solution of the 1-dimensional nonlinear matroid optimization problem over any matroid M , given explicitly or by an oracle, on n -element ground sets, and a univariate function f presented by a comparison oracle, cannot be done in polynomial time. In particular, with the single weight vector $w := (1, 2, 4, \dots, 2^{n-1})$, solution of the nonlinear matroid optimization problem requires examining $f(W(B))$ for each of the $|\mathcal{B}(M)|$ bases of M . In particular, for each r , $1 \leq r \leq k$, the problem cannot be solved in polynomial time for the class of matroids $\mathcal{M}_{r,k}$.*

Proof. The weights $w(S) = \sum_{j \in S} 2^{j-1}$ of the 2^n subsets $S \subseteq N$ attain precisely all 2^n distinct values $0, 1, \dots, 2^n - 1$. Since the function f is arbitrary, this implies that the objective value $f(W(B))$ of each base B can be arbitrary. Therefore, if the value $f(W(B))$ of some base B is not compared against, it may be that this value is the unique minimum one and the nonlinear matroid optimization problem cannot be correctly solved. The final remark of the proposition follows since, for every r , $|\mathcal{B}(\mathcal{M}_{r,k})|$ is not bounded by any polynomial in $|\mathcal{E}(\mathcal{M}_{r,k})| = 2k$. $\square$

PROPOSITION 2.4. *Computing an optimal solution of the nonlinear matroid optimization problem in variable dimension $d = n$, over any matroid M , with $\{0, 1\}$ -valued weights, the i th weight being the standard unit vector $w_i := \mathbf{1}_i$ in $\mathbb{Z}^n$ for all i , and with $f : \mathbb{Z}^n \rightarrow \mathbb{R}$ a function presented by a comparison oracle, requires examining*

$f(W(B))$ for each of the $|\mathcal{B}(M)|$ bases of M . In particular, for each r , $1 \leq r \leq k$, the problem cannot be solved in polynomial time for the class of matroids $\mathcal{M}_{r,k}$.

Proof. The W -images $W(B) = (w_1(B), \dots, w_n(B))$ of the 2^n subsets $B \subseteq N$ attain precisely all 2^n distinct vectors in $\{0, 1\}^n$. Since the function f is arbitrary, this implies that the objective value $f(W(B))$ of each base can be arbitrary. The rest of the argument is as in the proof of Proposition 2.3. $\square$

Consider nonlinear matroid optimization with a matroid M , weights $w_1, \dots, w_d \in \mathbb{Z}^n$, and function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$. As explained in section 2.2, a crucial component in solving the problem is to identify W -images of bases; that is, points $u \in \mathbb{Z}^d$ satisfying $u = W(B)$ for some $B \in \mathcal{B}(M)$. The following proposition shows that, with binary-encoded weights, both the nonlinear matroid optimization problem with explicitly given univariate convex quadratic function f and the problem of deciding if a given u is a W -image of some base, are intractable already for fixed $d = 1$ and uniform matroids or partition matroids that are the direct sums of rank-1 uniform matroids.

PROPOSITION 2.5. *Given matroid M , a single nonnegative weight vector $w \in \mathbb{N}^n$, and nonnegative integer $u \in \mathbb{N}$, encoded in binary, the following problems are NP-complete, when already restricted to the class of matroids $\mathcal{M}_{r,k}$, for any r , $1 \leq r \leq k$:*

1. *Determining whether $u = W(B) = \sum_{j \in B} w_j$ for some base $B \in \mathcal{B}(M)$.*
2. *Determining whether the optimal objective value is zero for the 1-dimensional nonlinear matroid optimization problem over M , with the explicit convex univariate function $f(y) := (y - u)^2$.*

Proof. The NP-complete *subset-sum problem* is to decide, given $a_0, a_1, \dots, a_k \in \mathbb{N}$, whether there is a subset $S \subseteq K = \{1, \dots, k\}$ with $\sum_{j \in S} a_j = a_0$. Given such a_i , let $u := a_0$, and let $w := (a_1, \dots, a_k, 0, \dots, 0) \in \mathbb{N}^K \times \mathbb{N}^{\bar{K}}$. Then, for any r , a base B of $\mathcal{M}_{r,k}$ satisfies $W(B) = u$ if and only if $S := B \cap K$ satisfies $\sum_{j \in S} a_j = a_0$. This reduces the subset-sum problem to the problem considered in the first part of the proposition, showing that it is indeed NP-complete.

For the second part, note that the objective value $f(W(B)) = (W(B) - u)^2$ of every base B is nonnegative, and B has $f(W(B)) = 0$ if and only if $W(B) = u$. Thus, the optimal objective value is zero if and only if there is a base with $W(B) = u$. So the problem in the first part of the proposition reduces to the problem in the second part, showing the latter to be NP-complete as well. $\square$

3. Arbitrary matroids. In this section we develop a combinatorial algorithm for nonlinear matroid optimization that runs in polynomial time for any matroid presented by an independence oracle, provided that the number p of distinct values taken by the entries $w_{i,j}$ of the weight vectors is fixed. In particular, the algorithm applies to $\{0, 1\}$ -valued weight vectors as well as to $\{0, 1, \dots, p\}$ -valued weight vectors for any fixed p .

As explained in section 2.2, we will filter the set $U = \{W(B) : B \in \mathcal{B}(M)\}$ of W -images of bases out of a suitable superset Z . For this, we next show how to efficiently decide if a given $u \in \mathbb{Z}^d$ satisfies $u = W(B)$ for some $B \in \mathcal{B}(M)$. We start with $\{0, 1\}$ -valued W -images with pairwise-disjoint supports.

LEMMA 3.1. *There is an algorithm that, given matroid M presented by an independence oracle on the n -element ground set N , weight vectors $w_1, \dots, w_d \in \{0, 1\}^n$ with pairwise-disjoint supports, and $u \in \mathbb{N}^d$, determines if M has a base B with W -image $W(B) = u$, in time polynomial in n and $\langle u \rangle$.*

Proof. For each base B of M and for each $i = 1, \dots, d$ we have $w_i(B) = |B \cap \text{supp}(w_i)|$. Therefore, a base B has $W(B) = u$ if and only if $|B \cap \text{supp}(w_i)| = u_i$ for $i = 1, \dots, d$. So we may and do assume $\sum_{i=1}^d u_i \leq \text{rank}(M)$ and $u_i \leq |\text{supp}(w_i)|$ for

all i else M has no base with $W(B) = u$. Let

$$\mathcal{B}' := \left\{ B \subseteq N : |B \cap \text{supp}(w_i)| = u_i, \quad i = 1, \dots, d, \right. \\ \left. \left| B \cap \left(N \setminus \bigcup_{i=1}^d \text{supp}(w_i) \right) \right| = \text{rank}(M) - \sum_{i=1}^d u_i \right\}.$$

It is easy to see that $\mathcal{B}' = \mathcal{B}(M')$ is the set of bases of a partition matroid M' on ground set N , for which an independence oracle is efficiently realizable. Moreover, M has a base B with $W(B) = u$ if and only if $\mathcal{B}(M) \cap \mathcal{B}(M')$ is nonempty. These observations justify the following algorithm:

```

if  $\sum_{i=1}^d u_i \leq \text{rank}(M)$  or  $u_i > |\text{supp}(w_i)|$  for some  $i = 1, \dots, d$  then return
    NO;
    Determine if  $\mathcal{B}(M) \cap \mathcal{B}(M')$  is nonempty by computing a max-cardinality
     $S \in \mathcal{I}(M) \cap \mathcal{I}(M')$ ;
    if  $|S| = \text{rank}(M)$ ;
    then
        return YES and  $B := S$ ;
    else
        return NO;
    end
    
```

Computing a max-cardinality $S \in \mathcal{I}(M) \cap \mathcal{I}(M')$ can be efficiently carried out using a maximum-cardinality matroid-intersection algorithm; see, e.g., [9] or [10] and the references therein. $\square$

Next, we consider weight vectors for which the number p of distinct $w_{i,j}$ values is fixed. So, we assume that $w_1, \dots, w_d \in \{a_1, \dots, a_p\}^n$ for arbitrary given integer numbers $a_1, \dots, a_p$. Note that the a_i can vary and be very large, since they affect the running time through their binary length $\langle a_i \rangle$.

LEMMA 3.2. *For every fixed d and p , there is an algorithm that, given matroid M presented by an independence oracle on ground set N , integers $a_1, \dots, a_p$, weight vectors $w_1, \dots, w_d \in \{a_1, \dots, a_p\}^n$, and $u \in \mathbb{N}^d$, decides if M has a base B with $W(B) = u$ in time polynomial in n , $\max \langle a_i \rangle$, and $\langle u \rangle$.*

Proof. Let $V := \{a_1, \dots, a_p\}^d$, and let Q be the $d \times p^d$ pattern matrix having columns that are all of the p^d points in V . Let, as usual, W be the $d \times n$ matrix with rows $w_1, \dots, w_d$. For $j = 1, \dots, n$, let w^j denote the j th column of W . We will exploit the fact that, no matter how large n is, the columns w^j of W all lie in the fixed set V —so the number of distinct columns w^j of W is limited.

Define a $p^d \times n$ selector matrix $\widehat{W}$, having rows $\widehat{w}_v$ indexed by V , columns $\widehat{w}^j$ indexed by N , and

$$\widehat{w}_{v,j} := \begin{cases} 1 & \text{if } w^j = v, \\ 0 & \text{otherwise,} \end{cases}$$

for $v \in V$, $j \in N$. Note that each column $\widehat{w}^j$ of $\widehat{W}$ is a standard unit vector, selecting the unique pattern (i.e., column) of V that agrees with the column w^j of W . It should be clear that the rows of $\widehat{W}$, namely the $\widehat{w}_v$, lie in $\{0, 1\}^n$ and have pairwise-disjoint supports—this will enable us to appeal to Lemma 3.1. We observe and will make use of the fact that the weight matrix W factors as $W = Q\widehat{W}$. Therefore, the W -image and

$\widehat{W}$ -image of a base B satisfy $W(B) = Q\widehat{W}(B)$. This implies that there exists a base $B \in \mathcal{B}(M)$ with $W(B) = u$ if and only if for some p^d -dimensional vector $\widehat{u}$ satisfying $u = Q\widehat{u}$ there exists a base $B \in \mathcal{B}(M)$ satisfying $\widehat{W}(B) = \widehat{u}$. Since $\widehat{W}$ is $\{0, 1\}$ -valued, any such vector $\widehat{u} = \widehat{W}(B)$ must lie in $\{0, \dots, m\}^{p^d}$, where $m := \text{rank}(M)$. Therefore, checking if there is a base $B \in \mathcal{B}(M)$ with $W(B) = u$ reduces to going over all vectors $\widehat{u} \in \{0, \dots, m\}^{p^d}$, and for each vector $\widehat{u}$ checking if $Q\widehat{u} = u$ and if there is a base B satisfying $\widehat{W}(B) = \widehat{u}$. This justifies the following algorithm:

```

let  $Q$  be the  $d \times p^d$  pattern matrix, and let  $\widehat{W}$  be the selector matrix (both
determined by  $W$ );
for  $\widehat{u} \in \{0, 1, \dots, m\}^{p^d}$  do
  if  $Q\widehat{u} = u$  then
    if there is a base  $B \in \mathcal{B}(M)$  with  $\widehat{W}(B) = \widehat{u}$  then return  $B$ ;
  end
end
return NO;

```

Since d and p are fixed and $m \leq n$, the number $(m+1)^{p^d}$ of such potential vectors $\widehat{u}$ is polynomial in the data. For each such vector $\widehat{u}$, checking if $Q\widehat{u} = u$ is easily done by direct multiplication; and checking if there exists a base $B \in \mathcal{B}(M)$ satisfying $\widehat{W}(B) = \widehat{u}$ can be done in polynomial time using the algorithm of Lemma 3.1 applied to the matroid M , the $\widehat{d} := p^d$ weight vectors $\widehat{w}_v$, $v \in V$ (the $\{0, 1\}$ -valued rows of the matrix $\widehat{W}$, having pairwise-disjoint supports), and the vector $\widehat{u}$. $\square$

We are now in position to solve the nonlinear optimization problem over a matroid that is presented by an independence oracle.

THEOREM 1.1. *For every fixed d and p , there is an algorithm that, given a matroid M presented by an independence oracle on the n -element ground set N , integers $a_1, \dots, a_p \in \mathbb{Z}$, weight vectors $w_1, \dots, w_d \in \{a_1, \dots, a_p\}^n$, and function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ presented by a comparison oracle, solves the nonlinear matroid optimization problem in time that is polynomial in n and $\max\langle a_i \rangle$.*

Proof. We have three major steps.

1. First, under the hypotheses of the theorem, assume that in polynomial time we can calculate an optimal W -image for the problem. Then, for every subset $S \subseteq N$, an optimal W -image for the subproblem on S can be computed in polynomial time, since the entries of the restrictions of $w_1, \dots, w_d$ to S also attain values in $\{a_1, \dots, a_p\}$. By Lemma 2.2 this implies that an optimal base can be found and the nonlinear matroid optimization problem solved in polynomial time. So, it remains to show that in polynomial time we can calculate an optimal W -image for the problem.
2. To accomplish this, we first show how to compute the set U of W -images of bases of M , by working with an appropriately defined superset Z of U . Let $m := \text{rank}(M)$. Consider any $i = 1, \dots, d$ and any m -subset B of N . Then, since $w_{i,j} \in \{a_1, \dots, a_p\}$ for all j , we have that, for some nonnegative integers $\lambda_{i,1}, \dots, \lambda_{i,p}$ with $\sum_{k=1}^p \lambda_{i,k} = m$,

$$w_i(B) = \sum_{j \in B} w_{i,j} = \sum_{k=1}^p \lambda_{i,k} a_k.$$

Therefore, we find that the set U of W -images of bases satisfies

$$\begin{aligned} U &= \{W(B) : B \in \mathcal{B}(M)\} \\ &\subseteq \{W(B) : B \subseteq N, |B| = m\} \\ &\subseteq Z := \left\{ \sum_{k=1}^p \lambda_k a_k : \lambda \in \{0, 1, \dots, m\}^p, \sum_{k=1}^p \lambda_k = m \right\}^d. \end{aligned}$$

These observations justify the following algorithm to compute the set U of W -images of bases:

```

Compute  $m := \text{rank}(M)$  and let  $a := (a_1, \dots, a_p)$ ;
Start with  $Z := \emptyset$ ;
for  $\Lambda \in \{0, 1, \dots, m\}^{d \times p}$  do
  if  $\Lambda \mathbf{1} = m \mathbf{1}$  then let  $Z := Z \cup \{\Lambda a\}$ ;
end
Start with  $U := \emptyset$ ;
for  $u \in Z$  do
  if there is a base  $B \in \mathcal{B}(M)$  with  $W(B) = u$  then let
     $U := U \cup \{u\}$ ;
end
return  $U$ ;

```

Observe that, since d and p are fixed, $|Z| \leq |\{0, 1, \dots, m\}^{d \times p}| = (m+1)^{pd}$ is polynomially bounded in $m \leq n$ and hence so are the numbers of iterations in each of the “for” loops of the algorithm. Also note that, in each iteration of the second loop, we can apply the algorithm of Lemma 3.2 to determine, in polynomial time, whether there is a base $B \in \mathcal{B}(M)$ with $W(B) = u$. Therefore, we can efficiently determine U .

3. By repeatedly querying the comparison oracle of f on $|U| - 1$ suitable pairs of points in U , we obtain a $u^* \in U$ that satisfies $f(u^*) = \min\{f(u) : u \in U\}$, which by Proposition 2.1 is an optimal W -image.

Finally, by embedding steps 2 and 3 above as subroutines to solve the $n + 1$ restrictions of the problem to suitable subsets $S \subseteq N$ in step 1, we obtain our desired algorithm. $\square$

4. Vectorial matroids. In this section we develop an algebraic algorithm for nonlinear matroid optimization over vectorial matroids. It applies to any matroid that is vectorial over an ordered field. For matroids that are vectorial over the rationals $\mathbb{Q}$, it runs in time polynomial in the binary length $\langle A \rangle$ of the matrix A representing the matroid and in the unary length $\max |w_{i,j}|$ of the weights. It is much more efficient than the combinatorial algorithm of section 3 and applies to weights with an unlimited number of different values $w_{i,j}$ of entries.

First, we show that it suffices to deal with nonnegative weight vectors.

LEMMA 4.1. *The nonlinear matroid optimization problem with arbitrary integer weight vectors $w_1, \dots, w_d \in \mathbb{Z}^n$ is polynomial-time reducible to the special case of nonnegative vectors $w_1, \dots, w_d \in \mathbb{N}^n$.*

Proof. Consider a matroid M , integer weights $w_1, \dots, w_d \in \mathbb{Z}^n$, and function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$. Let $m := \text{rank}(M)$ and $\omega := \max |w_{i,j}|$. Define nonnegative weights by $w'_{i,j} := w_{i,j} + \omega$ for all i, j , and define a new function $f' : \mathbb{Z}^d \rightarrow \mathbb{R}$ by $f'(u_1, \dots, u_d) := f(u_1 - m\omega, \dots, u_d - m\omega)$ for every $u = (u_1, \dots, u_d) \in \mathbb{R}^d$. Note that the unary length

of each new weight $w'_{i,j}$ is at most twice the maximum unary length of the original weights $w_{i,j}$, and a comparison oracle for f' is easily realizable from a comparison oracle for f . Then for every base B and for each $i = 1, \dots, d$, we have

$$w'_i(B) = \sum_{j \in B} w'_{i,j} = \sum_{j \in B} (w_{i,j} + \omega) = \left(\sum_{j \in B} w_{i,j} \right) + m\omega = w_i(B) + m\omega,$$

implying the following equality between the new and original objective function values:

$$f'(w'_1(B), \dots, w'_d(B)) = f'(w_1(B) + m\omega, \dots, w_d(B) + m\omega) = f(w_1(B), \dots, w_d(B)).$$

Therefore, a base $B \in \mathcal{B}(M)$ is optimal for the nonlinear matroid optimization problem with data $M, w_1, \dots, w_d$ and f if and only if it is optimal for the problem with $M, w'_1, \dots, w'_d$ and f' . $\square$

So we assume henceforth that the weights are nonnegative. As explained in section 2.2 and carried out in section 3, we will filter the set $U = \{W(B) : B \in \mathcal{B}(M)\}$ of W -images of bases out of a suitable superset Z . However, instead of checking if $u \in U$ for one point $u \in Z$ after the other, we will filter here the entire set U out of Z at once. We proceed to describe this procedure.

Let A be an $m \times n$ integer matrix of full row rank m , and let M be the vectorial matroid of A . Note that $m = \text{rank}(M)$. Let $w_1, \dots, w_d \in \mathbb{N}^n$ be nonnegative integer weight vectors, and let $\omega := \max w_{i,j}$. Then for each $i = 1, \dots, d$ and each m -subset B of N , we have $w_i(B) \in \{0, 1, \dots, m\omega\}$, and therefore

$$\begin{aligned} U &= \{W(B) : B \in \mathcal{B}(M)\} \subseteq \{W(B) : B \subseteq N, |B| = m\} \\ &\subseteq Z := \{0, 1, \dots, m\omega\}^d \subseteq \mathbb{N}^d. \end{aligned}$$

We will show how to filter the set U out of the above superset Z of potential W -images of bases. For each base $B \in \mathcal{B}(M)$, let $A_{.B}$ denote the nonsingular $m \times m$ submatrix of A consisting of those columns indexed by $B \subseteq N$. Define the following polynomial in d variables $x_1, \dots, x_d$:

$$(1) \quad g = g(x) := \sum_{u \in Z} g_u x^u := \sum_{u \in Z} g_u \prod_{k=1}^d x_k^{u_k},$$

where the coefficient g_u corresponding to $u \in Z$ is the nonnegative integer

$$(2) \quad g_u := \sum \{\det^2(A_{.B}) : B \in \mathcal{B}(M), W(B) = u\}.$$

Now, $\det^2(A_{.B})$ is positive for every base $B \in \mathcal{B}(M)$. Thus, the coefficient g_u corresponding to $u \in Z$ is nonzero if and only if there exists a matroid base $B \in \mathcal{B}(M)$ with $W(B) = u$. So the desired set U is precisely the set of exponents of monomials x^u with nonzero coefficient in g . We record this for later use.

PROPOSITION 4.2. *Let M be the vectorial matroid of an $m \times n$ matrix A of rank m , let $w_1, \dots, w_d \in \mathbb{N}^n$, and let $g(x)$ be the polynomial in (1). Then $U := \{W(B) : B \in \mathcal{B}(M)\} = \{u \in Z : g_u \neq 0\}$.*

By Proposition 4.2, to compute U , it suffices to compute all coefficients g_u . Unfortunately, they cannot be computed directly from the definition (2), since this involves again checking exponentially many $B \in \mathcal{B}(M)$ —precisely what we are trying to avoid!

Instead, we will compute the g_u by interpolation. However, in order to do so, we need a way of evaluating $g(x)$ under numerical substitutions. We proceed to show how this can be efficiently accomplished.

Let X be the $n \times n$ diagonal matrix whose j th diagonal component is the monomial $\prod_{i=1}^d x_i^{w_{i,j}}$ in the variables $x_1, \dots, x_d$, that is, the matrix of monomials defined by

$$X := \text{diag} \left(\prod_{i=1}^d x_i^{w_{i,1}}, \dots, \prod_{i=1}^d x_i^{w_{i,n}} \right).$$

The following lemma will enable us to compute the value of $g(x)$ under numerical substitutions.

LEMMA 4.3. *For any $m \times n$ matrix A of rank m and nonnegative weights $w_1, \dots, w_d \in \mathbb{N}^n$ we have*

$$g(x) = \det(AXA^T).$$

Proof. By the classical Binet–Cauchy identity, for any two $m \times n$ matrices C, D of rank m we have $\det(CD^T) = \sum \{\det(C_{\cdot B}) \det(D_{\cdot B}) : B \in \mathcal{B}(M)\}$. Applying this to $C := AX$ and $D := A$, we obtain

$$\begin{aligned} \det(AXA^T) &= \sum_{B \in \mathcal{B}(M)} \det((AX)_{\cdot B}) \det(A_{\cdot B}) = \sum_{B \in \mathcal{B}(M)} \prod_{j \in B} \prod_{i=1}^d x_i^{w_{i,j}} \det(A_{\cdot B}) \det(A_{\cdot B}) \\ &= \sum_{B \in \mathcal{B}(M)} \prod_{i=1}^d x_i^{w_i(B)} \det^2(A_{\cdot B}) = \sum_{u \in Z} \sum_{\substack{B \in \mathcal{B}(M): \\ W(B)=u}} \det^2(A_{\cdot B}) \prod_{i=1}^d x_i^{u_i} \\ &= \sum_{u \in Z} g_u x^u = g(x). \quad \square \end{aligned}$$

Lemma 4.3 paves the way for computing the coefficients of the polynomial $g(x) = \sum_{u \in Z} g_u x^u$ by interpolation. We will choose sufficiently many suitable points on the moment curve in $\mathbb{R}^Z$, substitute each point into x , and evaluate $g(x)$ using the lemma. We will then solve the system of linear equations for the coefficients g_u . The next lemma describes the details and shows that this can be done efficiently.

LEMMA 4.4. *For every fixed d , there is an algorithm that, given any $m \times n$ matrix A of rank m and weights $w_1, \dots, w_d \in \mathbb{N}^n$, computes all coefficients g_u of $g(x)$ in time polynomial in $\max w_{i,j}$ and $\langle A \rangle$.*

Proof. Let $\omega := \max w_{i,j}$ and $s := m\omega + 1$. Then a superset of potential W -images of bases is $Z := \{0, 1, \dots, m\omega\}^d$ and satisfies $|Z| = s^d$. For $t = 1, 2, \dots, s^d$, let $X(t)$ be the numerical matrix obtained from X by substituting $t^{s^{i-1}}$ for x_i , $i = 1, \dots, d$. By Lemma 4.3 we have $g(x) = \det(AXA^T)$, and therefore we obtain the following system of s^d linear equations in the s^d variables g_u , $u \in Z$:

$$\begin{aligned} \det(AX(t)A^T) &= \det \left(A \text{diag}_j \left(\prod_{i=1}^d t^{w_{i,j} s^{i-1}} \right) A^T \right) = \sum_{u \in Z} g_u \prod_{i=1}^d t^{u_i s^{i-1}} \\ &= \sum_{u \in Z} t^{\sum_{i=1}^d u_i s^{i-1}} g_u, \quad t = 1, 2, \dots, s^d. \end{aligned}$$

As u runs through Z , the sum $\sum_{i=1}^d u_i s^{i-1}$ attains precisely all $|Z| = s^d$ distinct values $0, 1, \dots, s^d - 1$. This implies that, under the total order of the points u in Z by increasing value of $\sum_{i=1}^d u_i s^{i-1}$, the vector of coefficients of the g_u in the equation corresponding to t is precisely the point $(t^0, t^1, \dots, t^{s^d-1})$ on the moment curve in $\mathbb{R}^Z \cong \mathbb{R}^{s^d}$. Therefore, the equations are linearly independent and hence the system can be uniquely solved for the g_u .

These observations justify the following algorithm to compute the g_u , $u \in Z$:

```

Compute  $m := \text{rank}(A)$ ;
let  $\omega := \max w_{i,j}$ , and let  $s := m\omega + 1$ ;
let  $X := \text{diag}_j \left( \prod_{i=1}^d x_i^{w_{i,j}} \right)$ ;
for  $t = 1, 2, \dots, s^d$  do
    let  $X(t)$  be the numerical matrix obtained by substituting  $t^{s^{i-1}}$  for  $x_i$ ,
     $i = 1, 2, \dots, d$ , in  $X$ ; Compute  $\det(AX(t)A^T)$ ;
end
let  $Z := \{0, 1, \dots, m\omega\}^d$ ;
Compute and return the unique solution  $g_u$ ,  $u \in Z$ , of the square linear
system:

$$\det(AX(t)A^T) = \sum_{u \in Z} t^{\sum_{i=1}^d u_i s^{i-1}} g_u, \quad t = 1, 2, \dots, s^d.$$


```

We now show that this system can be solved in polynomial time. First, the number of equations and indeterminates is $s^d = (m\omega + 1)^d$ and hence polynomial in the data. Second, for each $i, j = 1, \dots, n$ and $t = 1, 2, \dots, s^d$, it is easy to see that the (i, j) th entry of $AX(t)A^T$ satisfies

$$\left| \sum_{h=1}^n a_{i,h} \prod_{k=1}^d t^{s^{k-1} w_{k,h}} a_{j,h} \right| \leq \sum_{h=1}^n |a_{i,h} a_{j,h}| p^{ds^d \max w_{k,h}},$$

implying that the binary length $\langle AX(t)A^T \rangle$ of $AX(t)A^T$ is polynomially bounded in the data as well.

It follows that $\det(AX(t)A^T)$ can be computed in polynomial time by Gaussian elimination for all t , and the system of equations can indeed be solved for the g_u in polynomial time. We further note that the system of equations is a Vandermonde system, so the number of arithmetic operations needed to solve it is just quadratic in its dimensions. $\square$

We can now efficiently solve the nonlinear optimization problem over vectorial matroids with unary weights.

THEOREM 1.3. *For every fixed d , there is an algorithm that, given integer $m \times n$ matrix A , weight vectors $w_1, \dots, w_d \in \mathbb{Z}^n$, and function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ presented by a comparison oracle, solves the nonlinear optimization problem over the (real) vectorial matroid of A in time polynomial in $\langle A \rangle$ and $\max |w_{i,j}|$.*

Proof. Let M be the vectorial matroid of A . Recall that linear-algebraic operations on A can be done in polynomial time, say by Gaussian elimination. Dropping some rows of A if necessary without changing M , we may assume that A has rank m . An independence oracle for M is readily realizable since $S \subseteq N$ is independent in M precisely when the columns of A indexed by S are linearly independent. Applying, if necessary, the procedure of Lemma 4.1 and adjusting the weights while at most doubling the unary length of the maximum weight, we may also assume that the weights are nonnegative.

We will show how to compute an optimal W -image u^* in polynomial time. This will also imply that, for every subset $S \subseteq N$, an optimal W -image for the subproblem on S can be computed in polynomial time. By Lemma 2.2 this will show that an optimal base can be found and the nonlinear matroid optimization problem solved in polynomial time.

Let $\omega := \max w_{i,j}$ and consider the superset $Z := \{0, 1, \dots, m\omega\}^d$ of potential W -images of bases and the polynomial $g(x) = \sum_{u \in Z} g_u x^u$ as defined in (1) and (2). By Proposition 4.2 we have

$$U = \{W(B) : B \in \mathcal{B}(M)\} = \{u \in Z : g_u \neq 0\}.$$

Applying now the algorithm of Lemma 4.4, we can compute in polynomial time the right-hand side and hence the left-hand side, providing the filtration of the set U of W -images of bases out of Z . By repeatedly querying the comparison oracle of f on suitable pairs of points in U , we obtain a $u^* \in U$ satisfying $f(u^*) = \min\{f(u) : u \in U\}$, which by Proposition 2.1 is the desired optimal W -image. $\square$

5. Experimental design. We now discuss applications of nonlinear matroid optimization to experimental design. For general information on experimental design, see, e.g., the monograph [15] and the references therein. As outlined in the introduction, we consider the following rather general framework. We wish to learn an unknown system whose output y is an unknown function Φ of a multivariate input $x = (x_1, \dots, x_k) \in \mathbb{R}^k$. It is customary to call the input variables x_i *factors* of the system. We perform several experiments. Each experiment i is determined by a point $p_i = (p_{i,1}, \dots, p_{i,k})$ and consists of feeding the system with input $x := p_i \in \mathbb{R}^k$ and measuring the corresponding output $y_i := \Phi(p_i) \in \mathbb{R}$. Based on these experiments, we wish to *fit a model* for the system, namely, determine an estimation $\hat{\Phi}$ of Φ , that:

- lies in a prescribed class of functions;
- is consistent with the outcomes of the experiments;
- minimizes the *aberration*—a suitable criterion—among models in the class.

We concentrate on (multivariate) *polynomial models* defined as follows. Each nonnegative integer vector $\alpha \in \mathbb{N}^k$ serves as an *exponent* of a corresponding monomial $x^\alpha := \prod_{h=1}^k x_h^{\alpha_h}$ in the system input $x \in \mathbb{R}^k$. Each finite subset $B \subset \mathbb{N}^k$ of exponents provides a *model* for the system, namely a polynomial *supported on* B , i.e., having monomials with exponents in B only,

$$\Phi_B(x) = \sum_{\alpha \in B} c_\alpha x^\alpha,$$

where the c_α are real coefficients that need to be determined from the measurements by interpolation.

We assume that the set of design points $\{p_1, \dots, p_m\} \subset \mathbb{R}^k$ is prescribed. Indeed, in practical applications, it may be impossible or too costly to conduct experiments involving arbitrarily chosen points. The problem of choosing the design (termed the *inverse problem* in the statistics literature; see [2] and the references therein), is of interest in its own right, and its computational aspects will be considered elsewhere. We collect the design points in an $m \times k$ *design* matrix P . Thus, the i th row of this matrix is the i th design point p_i . A model $B \subset \mathbb{N}^k$ is *identifiable* by a design P if for any possible measurement values $z_i = \Phi(p_i)$ at the design points, there is a unique polynomial $\Phi_B(x)$ supported on B that interpolates Φ , that is, satisfies $\Phi_B(p_i) = z_i = \Phi(p_i)$ for every design point $p_i = (p_{i,1}, \dots, p_{i,k})$.

Among models identifiable by a given design, we wish to determine one that is best under a suitable criterion. Roughly speaking, common criteria ask for *low degree polynomials*. To make this precise, for each identifiable model B , consider the following *total degree vector* whose i th entry is the total degree of variable x_i over all monomials supported on B :

$$\sum_{\alpha \in B} \alpha = \left(\sum_{\alpha \in B} \alpha_1, \dots, \sum_{\alpha \in B} \alpha_k \right).$$

Now, given any function $f : \mathbb{Z}^k \rightarrow \mathbb{R}$, the *aberration* of model B induced by f is defined to be

$$\mathcal{A}(B) := f \left(\sum_{\alpha \in B} \alpha \right).$$

The term *aberration* is the one used in the statistics literature in this context; see, e.g., [5, 18] and the references therein. We now give some concrete examples of functions providing useful aberrations.

Example 5.1 (some concrete useful aberrations).

- Consider the function $f(u) := \frac{1}{|B|}(u_1 + \dots + u_k)$. Then the aberration of model B is

$$\mathcal{A}(B) = \frac{1}{|B|} \sum_{\alpha \in B} \alpha_1 + \dots + \frac{1}{|B|} \sum_{\alpha \in B} \alpha_k = \frac{1}{|B|} \sum_{\alpha \in B} \sum_{i=1}^k \alpha_i,$$

which is the average total degree of monomials supported on B .

- Consider $f(u) := \frac{1}{|B|}(\pi_1 u_1 + \dots + \pi_k u_k)$ for some real weights $\pi_1, \dots, \pi_k$. Then

$$\mathcal{A}(B) = \pi_1 \frac{1}{|B|} \sum_{\alpha \in B} \alpha_1 + \dots + \pi_k \frac{1}{|B|} \sum_{\alpha \in B} \alpha_k$$

is the weighted average degree, allowing for preferences of some variables over others.

- Consider the function $f(u) := \frac{1}{|B|}(\max\{u_1, \dots, u_k\})$. Then the aberration of B is

$$\mathcal{A}(B) = \max \left\{ \frac{1}{|B|} \sum_{\alpha \in B} \alpha_1, \dots, \frac{1}{|B|} \sum_{\alpha \in B} \alpha_k \right\},$$

and is the maximum over variables of the average variable degree of monomials supported on B .

- More generally, consider $f(u) := \frac{1}{|B|}(\|\pi \cdot u\|_q) = \frac{1}{|B|} \left(\left(\sum_{i=1}^k |\pi_i u_i|^q \right)^{\frac{1}{q}} \right)$. Then the aberration of B is

$$\mathcal{A}(B) = \left\| \pi \cdot \left(\frac{1}{|B|} \sum_{\alpha \in B} \alpha \right) \right\|_q,$$

which is the l_q -norm of the weighted average degree vector of monomials supported on B .

We can now formally define the minimum-aberration model-fitting problem.

Minimum-aberration model-fitting problem. Given a design $P = \{p_1, \dots, p_m\}$ of m points in $\mathbb{R}^k$, a set $N = \{\beta_1, \dots, \beta_n\}$ of n potential exponents in $\mathbb{N}^k$, and a function $f : \mathbb{Z}^k \rightarrow \mathbb{R}$, find a model $B \subseteq N$ that is identifiable by P and is of minimum aberration

$$\mathcal{A}(B) = f \left(\sum_{\beta_j \in B} \beta_j \right).$$

Next, we demonstrate how to formulate the minimum-aberration model-fitting problem as a nonlinear matroid optimization problem. Consider the following $m \times n$ matrix A having rows indexed by P and columns indexed by N , defined by

$$a_{i,j} := p_i^{\beta_j} = \prod_{h=1}^k p_{i,h}^{\beta_{j,h}}, \quad i = 1, \dots, m, j = 1, \dots, n.$$

It can be verified that a model $B \subseteq N$ is identifiable by a design P if and only if the $m \times m$ matrix A_B is invertible. (In the terminology of algebraic geometry, the model B is identifiable if the congruence classes of the monomials $x^{\beta_1}, \dots, x^{\beta_m}$ form a basis for the quotient of the algebra of polynomials $\mathbb{R}[x_1, \dots, x_k]$ modulo the ideal of polynomials vanishing on the design points; see [3, 13] and the references therein for more on this.) If B is identifiable, then, given any vector of measurements $y \in \mathbb{R}^m$ at the design points, the vector of coefficients $c \in \mathbb{R}^m$ of the unique polynomial $\Phi_B(x) = \sum_{j=1}^m c_j \prod_{h=1}^k x_h^{\beta_{j,h}}$ supported on B that is consistent with the measurements is given by $c := (A_B)^{-1} y$.

If the rank of A is less than m , then no $B \subseteq N$ is identifiable, and the set N of potential exponents should be augmented with more exponents. So assume that A has rank m . Let M be the vectorial matroid of A , so that

$$\mathcal{B}(M) := \{B \subseteq N : B \text{ is identifiable by } P\}.$$

Now define k weight vectors $w_1, \dots, w_k \in \mathbb{N}^n$ by $w_{i,j} := \beta_{j,i}$ for $i = 1, \dots, k, j = 1, \dots, n$. Then the aberration of model B is

$$\mathcal{A}(B) = f \left(\sum_{\beta_j \in B} \beta_j \right) = f \left(\sum_{\beta_j \in B} w_{1,j}, \dots, \sum_{\beta_j \in B} w_{k,j} \right).$$

Thus, the aberration of a model B identifiable by the design P is precisely the objective function value of the base B in the nonlinear matroid optimization problem over the matroid M above, with $d := k$ and the weights $w_1, \dots, w_k \in \mathbb{N}^n$ as above, and with the given function $f : \mathbb{Z}^k \rightarrow \mathbb{R}$. Assuming that a comparison oracle for the function f can be realized, which is practically always true, and that the design points are rational so that they can be input and processed on a digital computer, we obtain the following corollary of Theorems 1.1 and 1.3.

COROLLARY 5.2. *For every fixed k , there is an algorithm that, given a rational design $P = \{p_1, \dots, p_m\}$ in $\mathbb{R}^k$, a set $N = \{\beta_1, \dots, \beta_n\}$ in $\mathbb{N}^k$, and a function $f : \mathbb{Z}^k \rightarrow \mathbb{R}$ presented by a comparison oracle, solves the minimum-aberration model-fitting problem in time polynomial in $m, n, \langle P \rangle$, and $\max \beta_{i,j}$.*

It is very natural and common in practice to consider *hierarchical models*; that is, models B with the property that $\beta \leq \alpha \in B$ implies $\beta \in B$. In [3, 13], in the

context of the theory of Gröbner bases in commutative algebra, it was shown that the smallest set containing all m -point hierarchical models B (termed *staircases* therein) in $\mathbb{N}^k$ is the following set, consisting of roughly $O(m \log m)$ points,

$$N := \left\{ \alpha \in \mathbb{N}^k : \prod_{h=1}^k (\alpha_h + 1) \leq m \right\}.$$

Thus, Corollary 5.2 will be typically applied with this set N as the set of potential monomial exponents.

We proceed to describe a more general useful class of aberrations that can be treated, which is naturally suggested by the nonlinear matroid optimization formulation. As before, we are given a design $P = \{p_1, \dots, p_m\}$ in $\mathbb{R}^k$ and a set $N = \{\beta_1, \dots, \beta_n\}$ of potential exponents in $\mathbb{N}^k$. But now we are also given d weight vectors $w_1, \dots, w_d \in \mathbb{Z}^n$. The function f is now defined on $\mathbb{R}^d$ rather than $\mathbb{R}^k$. The aberration induced by the weights and the function is now simply the objective function of the nonlinear matroid optimization problem, which for an identifiable model $B \subseteq N$ is given by

$$\mathcal{A}(B) := f(W(B)),$$

where W is the matrix with rows w_i . Note that aberrations of the type considered before can be recovered as a special case with $d := k$ and $w_{i,j} := \beta_{j,i}$ for all i, j . Here are a few useful examples.

Example 5.3 (some concrete useful generalized aberrations). Let $N = \{\beta_1, \dots, \beta_n\} \subset \mathbb{N}^k$ be any set of exponents. Let θ be a small positive integer, say $\theta = 1$ or $\theta = 2$, that will serve as a desired bound on the degrees of variables in monomials of the sought after models B contained in N .

- Let $d := 1$, and define the single weight vector $w_1 \in \mathbb{N}^n$ by

$$w_{1,j} := \begin{cases} 0 & \text{if } \beta_{j,i} \leq \theta \text{ for every } i = 1, \dots, k, \\ 1 & \text{otherwise,} \end{cases}$$

for $j = 1, \dots, n$. Then $\sum_{\beta_j \in B} w_{1,j}$ is the number of monomials supported on $B \subseteq N$ that do *not* meet the degree bound; in particular, for $\theta = 1$ it is the number of non square-free monomials. Taking $f : \mathbb{Z} \rightarrow \mathbb{R}$ to be the identity $f(u) := u$, the aberration $\mathcal{A}(B)$ of model B is the number of undesired monomials. In particular, an optimal model B has $\mathcal{A}(B) = 0$ if and only if the design admits an identifiable model with all variables in all monomials having degree at most θ .

- Now let $d := k$, and define weight vectors $w_1, \dots, w_k \in \mathbb{N}^n$ by

$$w_{i,j} := \begin{cases} 0 & \text{if } \beta_{j,i} \leq \theta, \\ 1 & \text{otherwise,} \end{cases}$$

for $i = 1, \dots, k$, $j = 1, \dots, n$. Then $\sum_{\beta_j \in B} w_{i,j}$ is the number of monomials supported on model $B \subseteq N$ for which variable x_i violates the degree bound θ . Defining $f : \mathbb{Z}^k \rightarrow \mathbb{R}$ by $f(u) := \max_{i=1}^k u_i$, we get that the aberration $\mathcal{A}(B)$ of model B is the maximum over variables of the number of monomials having x_i violating the degree bound θ . The optimal model will minimize the maximum violation.

We have the following generalized minimum-aberration model-fitting problem and corollary.

Generalized minimal-aberration model-fitting problem. Given a design $P = \{p_1, \dots, p_m\}$ in $\mathbb{R}^k$, a set $N = \{\beta_1, \dots, \beta_n\}$ of potential exponents in $\mathbb{N}^k$, weight vectors $w_1, \dots, w_d \in \mathbb{Z}^n$, and a function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$, find a model $B \subseteq N$ that is identifiable by P and is of minimum aberration

$$\mathcal{A}(B) = f(W(B)) = f\left(\sum_{\beta_j \in B} w_{1,j}, \dots, \sum_{\beta_j \in B} w_{d,j}\right).$$

COROLLARY 5.4. *For every fixed k and d , there is an algorithm that, given any rational design $P = \{p_1, \dots, p_m\}$ in $\mathbb{R}^k$, any set $N = \{\beta_1, \dots, \beta_n\}$ of potential exponents in $\mathbb{N}^k$, weight vectors $w_1, \dots, w_d \in \mathbb{Z}^n$, and function $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ presented by a comparison oracle, solves the generalized minimum-aberration model-fitting problem in time polynomial in $m, n, \langle P \rangle, \max \beta_{j,i}$, and $\max |w_{i,j}|$.*

REFERENCES

- [1] A. V. AHO, J. E. HOPCROFT, AND J. D. ULLMAN, *The Design and Analysis of Computer Algorithms*, Addison-Wesley, Reading, MA, 1975.
- [2] A. C. ATKINSON, A. DONEV, AND R. TOBIAS, *Optimum experimental designs, with SAS*, Oxford Statistical Science Series, 34, Oxford University Press, Oxford, 2007.
- [3] E. BABSON, S. ONN, AND R. THOMAS, *The Hilbert zonotope and a polynomial time algorithm for universal Gröbner bases*, Adv. in Appl. Math., 30 (2003), pp. 529–544.
- [4] Y. BERSTEIN AND S. ONN, *Nonlinear bipartite matching*, Discrete Optim., 5 (2008), pp. 53–65.
- [5] A. FRIES AND W. G. HUNTER, *Minimum aberration 2^{k-p} designs*, Technometrics, 22 (1980), pp. 601–608.
- [6] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability*, W. H. Freeman, San Francisco, 1979.
- [7] R. HASSIN AND A. LEVIN, *An efficient polynomial time approximation scheme for the constrained minimum spanning tree problem using matroid intersection*, SIAM J. Comput., 33 (2004), pp. 261–268.
- [8] R. HASSIN AND A. TAMIR, *Maximizing classes of two-parameter objectives over matroids*, Math. Oper. Res., 14 (1989), pp. 362–375.
- [9] J. LEE, *A First Course in Combinatorial Optimization*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, UK, 2004.
- [10] J. LEE AND J. RYAN, *Matroid applications and algorithms*, INFORMS (formerly ORSA) J. Comput., 4 (1992), pp. 70–98.
- [11] S. ONN, *Convex matroid optimization*, SIAM J. Discrete Math., 17 (2003), pp. 249–253.
- [12] S. ONN AND U. G. ROTHBLUM, *Convex combinatorial optimization*, Discrete Comput. Geom., 32 (2004), pp. 549–566.
- [13] S. ONN AND B. STURMFELS, *Cutting corners*, Adv. in Appl. Math., 23 (1999), pp. 29–48.
- [14] J. G. OXLEY, *Matroid Theory*, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1992.
- [15] G. PISTONE, E. RICCOMAGNO, AND H. P. WYNN, *Algebraic Statistics*, Monographs on Statistics and Applied Probability 89, Chapman & Hall/CRC, Boca Raton, FL, 2001.
- [16] A. SCHRIJVER, *Theory of Linear and Integer Programming*, John Wiley and Sons, Chichester, UK, 1986.
- [17] D. J. A. WELSH, *Matroid Theory*, Academic Press, London, 1976.
- [18] H. WU AND C. F. J. WU, *Clear two-factor interactions and minimum aberration*, Ann. Statist., 30 (2002), pp. 1496–1511.

SPANNING TREES WITH MANY LEAVES IN GRAPHS WITH MINIMUM DEGREE THREE*

PAUL S. BONSMAT[†]

Abstract. We present two lower bounds for the maximum number of leaves over all spanning trees of a graph. For connected, triangle-free graphs on n vertices, with minimum degree at least three, we show that a spanning tree with at least $(n+4)/3$ leaves exists. For connected graphs with minimum degree at least three, without diamonds induced by vertices of degree three, we show that a spanning tree with at least $(2n+12)/7$ leaves exists. (A diamond is a K_4 minus one edge.) The proofs use the fact that spanning trees with many leaves correspond to small connected dominating sets. Both of these bounds are best possible for their respective graph classes. For both bounds, simple polynomial time algorithms that find spanning trees satisfying the bounds are given.

Key words. spanning tree, max leaf, connected dominating set, extremal graph theory

AMS subject classifications. 05C69, 05C85, 05C35

DOI. 10.1137/060664318

1. Introduction. For standard graph theoretic terminology we refer to [2]. A *tree* is a connected graph without cycles. A *spanning subgraph* of a graph G is a subgraph containing all vertices of G . A *spanning tree* of G is a spanning subgraph of G that is a tree. A *leaf* is a vertex with degree one. A subset $S \subseteq V(G)$ of the vertices of graph G is a *dominating set* if every vertex of G is in S or is adjacent to a vertex in S . $S \subseteq V(G)$ is a *connected set* if the graph it induces, denoted by $G[S]$, is connected. A set that is both a dominating set and a connected set will be called a *CD-set* for short. The minimum vertex degree of G is denoted by $\delta(G)$, or δ if there is no cause for confusion. See section 2 for additional definitions.

In this paper we study spanning trees with many leaves and the closely related concept of small CD-sets. The well-known relation between these concepts is as follows: a graph $G \neq K_2$ on n vertices has a spanning tree with at least k leaves if and only if G has a CD-set of size at most $n - k$.

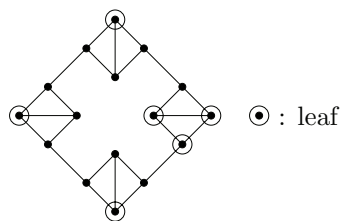
Much research has been done on the corresponding optimization problems, where the goal is to maximize the number of leaves of a spanning tree (resp., to minimize the number of vertices in a CD-set). The problems are known to be $\mathcal{NP}$ -hard [9]. For both problems approximation algorithms exist. For finding spanning trees with many leaves, approximation algorithms have been found by Lu and Ravi [15] and Solis-Oba [17], with approximation guarantees of 3 (resp., 2). When the input is restricted to cubic graphs, a $5/3$ -approximation recently has been presented by Correa et al. [7]. Approximation algorithms for finding small CD-sets have been found by Guha and Khuller [12] and Ruan et al. [16]. These algorithms do not have constant factor approximation guarantees, and in [12] it is shown that it is unlikely that such algorithms exist.

A second branch of results on these problems is in the area of extremal graph theory. We are interested in statements of this type: every graph on n vertices that

*Received by the editors July 6, 2006; accepted for publication (in revised form) December 10, 2007; published electronically May 21, 2008.

<http://www.siam.org/journals/sidma/22-3/66431.html>

[†]Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany (bonsma@math.tu-berlin.de). This research was done at the University of Twente, Enschede, The Netherlands.

FIG. 1.1. *Theorem 1 is best possible.*

satisfies a certain set of properties has a spanning tree with at least $\alpha n + \beta$ leaves ($0 < \alpha < 1$). Even though the focus in this area of research is often not on algorithms, many of these results are constructive and correspond to polynomial time algorithms.

Linial conjectured that every graph on n vertices with minimum degree δ has a spanning tree with at least $n - 3n/(\delta + 1) + c_\delta$ leaves, for some appropriate constant c_δ . For sufficiently large δ , it was observed in [6] that a result by Alon [1] disproves this conjecture. However, for small values of δ the conjecture has been shown to be true. For $\delta = 3$, Linial and Sturtevant [14] and Kleitman and West [13] independently proved the conjecture.

THEOREM 1 (Linial and Sturtevant, Kleitman and West). *Every connected graph on n vertices with $\delta \geq 3$ has a spanning tree with at least $\lceil \frac{1}{4}n + 2 \rceil$ leaves.*

For $\delta = 4$, a proof appears in [13]. For $\delta = 5$, Griggs and Wu [11] gave a proof. For small values of δ greater than five, little is known. For more information, see [6].

The result in Theorem 1 is best possible for every n . For $n = 4k$, an extremal example is constructed as follows; see Figure 1.1. Start with k disjoint *diamonds*, which is a K_4 minus an edge. Add k edges such that the graph becomes connected and cubic. A *cubic graph* is a graph in which all vertices have degree three. The resulting graph has no spanning tree with more than $k + 2$ leaves. For other values of n , similar constructions can be found. An immediate question is whether the bound can be improved when diamonds are forbidden. One result in this direction is by Griggs, Kleitman, and Shastri [10].

THEOREM 2 (see Griggs, Kleitman, and Shastri). *Every connected cubic graph without diamonds on n vertices has a spanning tree with at least $(n + 4)/3$ leaves.*

The bound in Theorem 2 is the best possible linear bound for this graph class: There exist cubic graphs without diamonds that admit no more than $n/3 + 2$ leaves, for every $n = 6k$ [10]. The single graph Q_3 shows that the additive term $\frac{4}{3}$ cannot be increased. (Q_3 is the 3-dimensional cube graph, which has eight vertices and no spanning tree with more than four leaves.) However, for other values of n no examples that show that this bound is sharp are known. See section 8 for a discussion on the sharpness of this bound and the other bounds in this paper.

New results. Our first question is what kind of bounds can be obtained when we consider graphs with $\delta \geq 3$ instead of cubic graphs but still forbid (certain types of) diamonds. In this paper, we present one best possible bound for these graphs. In addition, even though the proof of Kleitman and West for Theorem 1 is very short and elegant, the proof of Theorem 2 uses the same techniques but consists mainly of a very long and intricate case study. Therefore it is useful to develop new techniques for proving results of this kind. In section 3, the ideas behind our new method are introduced.

As a first application of these techniques, in section 4 we first prove that the bound from Theorem 2 also holds for graphs with $\delta \geq 3$ without triangles.

THEOREM 3. *Every connected, triangle-free graph on n vertices with $\delta \geq 3$ has a spanning tree with at least $(n + 4)/3$ leaves.*

Our second result is that for graphs with $\delta \geq 3$, the following statement holds. Here a *cubic diamond* is a diamond induced by vertices of degree three. (Similarly, a *cubic triangle* is a triangle induced by vertices of degree three.)

THEOREM 4. *Every connected graph on n vertices with $\delta \geq 3$, without cubic diamonds, has a spanning tree with at least $(2n + 12)/7$ leaves.*

In fact we prove the following stronger version of the statement.

THEOREM 5. *Every connected graph on n vertices with $\delta \geq 3$ that contains D cubic diamonds has a spanning tree with at least $(2n - D + 12)/7$ leaves.*

Because of space considerations, we do not give the details of the proof of Theorem 5 in this paper, but a proof sketch is given in section 5. Details can be found in [3]. See also section 5 for a slightly sharper formulation of Theorem 5. The bounds in Theorem 3 and Theorem 5 are best possible, in the same sense as explained above. This is shown in section 6.

Our proofs are constructive and correspond to efficient and simple algorithms: Spanning trees that satisfy the bound from Theorem 3 can be found with an algorithm that repeatedly adds vertices or pairs of vertices to a leaf set (according to some simple rules) until no more leaves can be added, and then constructs a spanning tree with those leaves. Spanning trees that satisfy the bound from Theorem 5 can be found using a similar algorithm that also involves local search on these leaf sets. These algorithms are discussed in section 7. We end in section 8 with a discussion of the results.

One important application of our bounds lies in the area of *fixed parameter tractable (FPT) algorithms*. These algorithms answer the following decision problem: does a graph G on n vertices have a spanning tree with at least k leaves? k is often chosen as the *parameter*, and then an algorithm for this problem is an FPT algorithm if its time complexity is bounded by a function $g(n)f(k)$, where $g(n)$ is a polynomial in the input size n , and the *parameter function* $f(k)$ may be any function. For designing fast FPT algorithms, the main goal is to find an algorithm where the parameter function $f(k)$ grows slowly. After a long line of improvements in this direction, two of the current best FPT algorithms in this respect appear in [4] and [8]. Bonsma, Brueggemann, and Woeginger [4] used the bound from Theorem 1 to construct an FPT algorithm with $f(k) \in O(9.49^k) \approx O(\binom{4k}{k})$. The factor 4 here comes from the factor $\frac{1}{4}$ in Theorem 1. Since our new Theorem 4 improves this factor to $\frac{2}{7}$, while adding only a minor extra restriction on the graph class, this can easily be used to improve the FPT algorithm given in [4]. The resulting algorithm has $f(k) \in O(\binom{3.5k}{k}) \approx O(8.12^k)$ and is the new fastest FPT algorithm. For details we refer to [3].

Another application of bounds of this type lies in the area of approximation algorithms, as demonstrated recently by Correa et al. [7]. They used the bound from Theorem 2 to deduce a bound similar to the one in Theorem 5: their bound states that every cubic graph has a spanning tree with at least $(3n - 2c + 14)/10$ leaves, where c is the number of components that result from deleting all internal vertices of diamonds (vertices that have all of their neighbors in the diamond itself). Combined with a suitable upper bound, this gives the $5/3$ -approximation algorithm for finding leafy trees in cubic graphs.

We also remark that apart from a small difference of $\frac{2}{7}$ in the constant, Theorem 5 can be seen as a generalization of Theorem 1: since $D \leq \frac{1}{4}n$ (cubic diamonds are vertex disjoint), it follows directly from Theorem 5 that connected graphs with $\delta \geq 3$ have a spanning tree with at least $(2n - D + 12)/7 \geq (\frac{7}{4}n + 12)/7 = \frac{1}{4}n + \frac{12}{7}$ leaves.

2. Preliminaries. We distinguish between $\subseteq$ and $\subset$: subset and proper subset, respectively. A graph G is called k -connected if $|V(G)| > k$ and there is no set $S \subset V(G)$ with $|S| < k$ such that $G - S$ is disconnected. For a nonempty set $S \subset V(G)$, $[S, \bar{S}]$ denotes the set of edges with exactly one end vertex in S . A set of this form is called an *edge cut*. If the single edge e forms an edge cut, e is called a *bridge*. K_n denotes a complete graph on n vertices. K_3 is called a *triangle*. A graph with no triangles as subgraphs is called *triangle-free*.

In this paper, the *blocks* of a graph are maximal 2-connected subgraphs. So unlike in the usual definition, a K_2 is not a block, and bridges are not contained in any block of the graph. The following lemma is an easy-to-prove variant of a well-known lemma.

LEMMA 6. *A connected graph without leaves and with cut vertices has at least two blocks that contain exactly one cut vertex.*

A set S is called minimal (maximal) for property ϕ if it satisfies property ϕ and no elements can be removed (added) such that ϕ is still satisfied. Minimal and maximal subgraphs are defined similarly. For a set $S \subseteq V(G)$, $G[S]$ denotes the subgraph of G induced by S ; this is the unique maximal subgraph of G with vertex set S . We will write $S - x$ and $S + x$ instead of $S \setminus \{x\}$ and $S \cup \{x\}$, respectively. If w is a function that assigns values to elements of S , then for $T \subseteq S$, $w(T) = \sum_{v \in T} w(v)$.

If $S \subseteq V(G)$, then a vertex $v \in S$ will be called an S -vertex. A neighbor v of u with $v \in S$ is called an S -neighbor of u . Similarly, a neighbor that is part of a CD-set that is currently considered is called a CD-set neighbor, etc.

Leafy spanning trees and small CD-sets. The following close relation between spanning trees and CD-sets is well known. If T is a spanning tree for $G = (V, E)$ with leaf set L , then $V \setminus L$ is a CD-set, unless $G = K_2$. Similarly, for every CD-set S , we can find in polynomial time a spanning tree for which $V \setminus S$ is a subset of the leaf set.

In the remainder of the paper, we will consider CD-sets instead of leaf sets of spanning trees. We feel that this is more practical for our purposes, which is illustrated by our methods, and discussed in section 8. If S is a CD-set, we will still call the vertices in $V \setminus S$ the *leaves* of S . If S is a CD-set for G and $G[S - v]$ is not connected, v is called a *connector*. If S is a CD-set for G and $S - v$ is not dominating, v is called a *dominator*. Observe that if S and $S' \subset S$ are CD-sets for G , and $v \in S$ is a dominator or connector, then $v \in S'$. A leaf v is called an *i-leaf* if it has i CD-set neighbors. A CD-set is called a *2-CD-set* if all its leaves are 1-leaves or 2-leaves.

Since we will consider CD-sets in the remainder of this paper, we state our main results in terms of CD-sets. Theorem 3 is equivalent to the following theorem.

THEOREM 7. *Every connected, triangle-free graph on n vertices with $\delta \geq 3$ has a CD-set S with $|S| \leq (2n - 4)/3$.*

Theorem 7 will be proved in section 4. Theorem 5 is equivalent to the following theorem.

THEOREM 8. *Every connected graph on n vertices with $\delta \geq 3$ that contains D cubic diamonds has a CD-set S with $|S| \leq (5n - 12 + D)/7$.*

In section 5, a slightly stronger variant of Theorem 8 is stated and proved.

3. Introduction to the method. We give a simple example to explain our proof technique.

DEFINITION 9. *$S \subseteq V(G)$ is a standard CD-set for graph G if the following properties hold:*

1. S is a 2-CD-set for G .
2. Every $v \in S$ has $d_G(v) \geq 3$.

Suppose that S is a standard CD-set for graph G and that $G[S]$ is a tree. We show that $|S|$ is roughly bounded from above by $\frac{2}{3}n$. For this it is sufficient to show that $|S| \leq 2|\bar{S}|$. This is easy when $G[S]$ is a tree: vertices with degree one or two in $G[S]$ have at least two, resp., one $\bar{S}$ -neighbors in G , since these vertices have degree at least three in G . A tree has average degree less than two, so there are at least $|S|$ edges in $[S, \bar{S}]$. Combining this with the fact that vertices in $\bar{S}$ are incident with at most two of these edges (S is a 2-CD-set), we obtain $|S| \leq 2|\bar{S}|$. This is the basic counting method behind the bounds we deduce. The next question is of course how to find a CD-set with the above properties, or similar properties.

Since we consider only graphs with minimum degree at least three, the property that vertices in S have degree at least three is easily satisfied. Finding a small 2-CD-set is not always possible, however. Fortunately, for our counting techniques, we may ignore edges between two vertices of high degree. More precisely, when u is an i -leaf with $i \geq 3$, we may delete incident edges to make it a 2-leaf, as long as this does not reduce the minimum degree of vertices in S below three. This leads to the following definition of *potential standard CD-sets*.

DEFINITION 10. $S \subseteq V(G)$ is a potential standard CD-set for G if there exists a spanning subgraph G' of G such that S is a standard CD-set for G' . Such a graph G' is called a realization of S . G' is a maximal realization of S if there is no other realization of S of which G' is a subgraph.

Observe that for a given potential standard CD-set S , every maximal realization has the same number of edges.

In the counting method example above, we required one additional property, namely that $G[S]$ is a tree. We cannot always find a potential standard CD-set S such that $G[S]$ is a tree, but we will show below that as long as we require S to be a *minimal* potential standard CD-set, the blocks of $G[S]$ have useful properties that can be used in our counting method to give similar bounds. In the next section we extend the above counting method to prove Theorem 7. In section 5, we prove Theorem 8. To prove this second bound, we impose three additional restrictions on the minimal potential standard CD-set we consider, and we slightly change the counting method.

4. Small CD-sets for graphs without triangles. In this section we state three properties for a pair consisting of a potential standard CD-set S and a realization G' . Below we will prove a bound for pairs S and G' that satisfy these properties. But first we will show that such a pair S and G' exists for every connected graph G with $\delta(G) \geq 3$.

Theorem 12 below states the essential properties of the pair S and G' that are needed to prove the bound. Property 1 is added to allow an easier formulation of properties 2 and 3 and to facilitate the later analysis, although it is not necessary for the method. Properties 2 and 3 are needed to ensure that S is relatively small.

To prove the existence of a pair S and G' with these properties, we can simply consider any *minimal* potential standard CD-set S ; for every such S a realization G' can be found such that the three properties are satisfied. However, it is not trivial to prove that a minimal potential standard CD-set can easily be found, and therefore we prove the statement using a potential standard CD-set S that is defined differently. This will be more convenient when we consider constructive algorithms in section 7.

DEFINITION 11. A potential standard CD-set S for graph G is 2-optimal if there is no potential standard CD-set $S' \subset S$ with $|S'| \geq |S| - 2$.

Clearly, any minimal potential standard CD-set is 2-optimal. It can also be shown that any 2-optimal potential standard CD-set is a minimal potential standard CD-set, but since we do not need this statement and the proof is not trivial, we omit this proof. Note that for Theorem 12, we do not yet need to assume that G contains no triangles.

THEOREM 12. *A connected graph G with $\delta(G) \geq 3$ has a potential standard CD-set S , with realization G' , with the following properties. Degrees, neighbors, and all other graph-related notions are taken with respect to G' .*

1. *If $u, v \in S$, $d(u) \geq d(v) \geq 4$, and $uv \in E(G')$, then uv is a bridge of $G'[S]$.*
2. *Every vertex $v \in S$ that is neither a dominator nor a connector has at least three neighbors in S , and all of its neighbors in S have degree three.*
3. *If $\{u, v\} \subseteq S$, $uv \in E(G')$, and u and v are both neither dominators nor connectors, then $G'[S] - u - v$ is not connected.*

Proof. Consider a 2-optimal potential standard CD-set S for graph G , with realization G'' . Consider a realization G' that is obtained from G'' as follows. First let $G' = G''$. If there is an edge $uv \in E(G'[S])$ that is not a bridge in $G'[S]$, with $d_{G'}(u) \geq 4$ and $d_{G'}(v) \geq 4$, remove uv from G' . Repeat this until no such edges exist anymore.

This procedure of removing edges does not destroy the connectivity of $G'[S]$, and it maintains the property that all vertices in S have degree at least three in S . So clearly G' is again a realization of S . By removing edges like this, it is now obvious that property 1 holds for S and G' .

To prove the remaining two properties, we will use the fact that S is a 2-optimal potential standard CD-set for G . Let v be a vertex in S that is not a dominator or connector. If v has at most two neighbors in S , then $S - v$ is again a potential standard CD-set, with realization G' , a contradiction with the 2-optimality of S . Now suppose v has at least three neighbors in S , and one of these neighbors $u \in S$ has $d_{G'}(u) \geq 4$. Since v is not a connector, uv is not a bridge of $G'[S]$, so by property 1, $d(v) = 3$. Then $S - v$ is a potential standard CD-set with realization $G' - uv$, again a contradiction. We conclude that any vertex v that is not a dominator or connector has at least three S -neighbors in G' and that all of these neighbors have degree three in G' , which proves property 2.

Now we will prove property 3. Suppose that u and v in S are neighbors in G' , and both are neither connectors nor dominators. By property 2, u and v have at least three neighbors in S , and all of those neighbors have degree three. It follows that u and v both have degree three, and therefore they have no neighbors outside of S . So $S - u - v$ is again a dominating set. If $G'[S] - u - v$ is connected, then u and v are 2-leaves with respect to $S - u - v$ and G' , so $S - u - v$ is a potential standard CD-set, a contradiction with the 2-optimality of S . We conclude that $G'[S] - u - v$ is not connected, which proves property 3. $\square$

Now that we have proved the existence of a CD-set with these properties, we can use them to establish upper bounds on the size of such a CD-set, assuming that the given graph G has no triangles.

To prove the bound in Theorem 13 below, we consider a potential standard CD-set S and realization G' having the properties stated in Theorem 12, and use an extension of the counting technique sketched before in section 3: leaves distribute a weight of 1 to their CD-set neighbors, and we show that these weights can be redistributed over the CD-set such that every vertex receives a weight of at least $\frac{1}{2}$. Previously we assumed that $G'[S]$ was a tree. Now we will allow $G'[S]$ to have blocks, and we

require only that these blocks do not contain triangles.

Observe that when $G'[S]$ is not a tree, there may be vertices in S that are not connectors or dominators. In this case a minimal CD-set $S' \subset S$ exists. It is for this set S' that we prove the bound. To improve the clarity of the exposition, we first state the theorem which yields this bound and then prove the two lemmas that are used in its proof.

THEOREM 13. *Let $G = (V, E)$ be a connected graph on n vertices with $\delta(G) \geq 3$. Let S be a potential standard CD-set for G with realization G' that satisfy the properties from Theorem 12 such that in addition no block of $G'[S]$ contains a triangle. If $S' \subseteq S$ is a minimal CD-set, then $|S'| \leq (2n - 4)/3$.*

Proof. In order to prove this theorem, we assign weights w to the vertices of S such that the total weight $w(S)$ is $|V \setminus S| + \frac{3}{2}|S \setminus S'|$. We show that $w(S) \geq \frac{1}{2}|S| + 2$. Then we can combine these equations to obtain the result.

The weight assignment is as follows: all leaves of S distribute a weight of 1 equally among their neighbors in S' (vertices in $S \setminus S'$ receive no weight yet). This means we have assigned a total weight of $|V \setminus S|$. Next, we assign a weight of $\frac{3}{2}$ to each vertex in $S \setminus S'$.

In order to show that $w(S) \geq \frac{1}{2}|S| + 2$, we prove three lower bounds for this weight assignment w .

1. A vertex v with degree two in $G'[S]$ is a dominator or connector (property 2 from Theorem 12), so $v \in S'$. Since v has at least one leaf neighbor $u \notin S$ (S is a standard CD-set for G'), and u has at most two neighbors in $G'[S']$, u assigns at least $\frac{1}{2}$ to v . This shows that for a vertex v with degree two in $G'[S]$, $w(v) \geq \frac{1}{2}$.
2. A vertex v with degree one in $G'[S]$ is not a connector, so it is a dominator (property 2). From the 1-leaf adjacent to it, it receives a weight of 1. In addition, since v has degree at least three and leaves have at most two CD-set neighbors, it gains an additional weight of at least $\frac{1}{2}$. So for a vertex v with degree one in $G'[S]$, $w(v) \geq \frac{3}{2}$.
3. Let $G'[B]$ be a block of $G'[S]$; let L denote the set of vertices in B with two neighbors in B , and $H = B \setminus L$ denote the set of vertices in B with at least three neighbors in B . Then $|L| \geq |H| + 1$ (which is shown below in Lemma 14). Let $C \subseteq S$ denote the set of connectors, which are the cut vertices of $G'[S]$. We will show that $w(B \setminus C) \geq \frac{1}{2}|B| + 2 - |C \cap B|$. A vertex in L that is not a connector also has degree two in $G'[S]$, and therefore must be a dominator (property 2). A dominator in B has weight at least 1. We consider two cases:

- (a) Suppose $|(S \setminus S') \cap B| \geq 1$. Then one of the vertices in H has an additional weight of at least $\frac{3}{2}$. We have $w(B \setminus C) \geq \frac{3}{2} + |L \setminus C| \geq \frac{3}{2} + |L| - |C \cap B|$. Since $|L| \geq |H| + 1$, we have $|L| \geq \frac{1}{2}|B| + \frac{1}{2}$. Hence $w(B \setminus C) \geq 2 + \frac{1}{2}|B| - |C \cap B|$.
- (b) Suppose $(S \setminus S') \cap B = \emptyset$. Suppose B contains a vertex v that is not a dominator or connector with respect to S . Since $v \in S'$, it is a connector or dominator with respect to S' . But $(S \setminus S') \cap B = \emptyset$, so v is also not a connector for S' . Therefore, in S , v is adjacent to a 2-leaf u , and the other CD-set neighbor of u is in $S \setminus S'$. In this case, v receives a weight of 1 from u . Recall that dominators also receive a weight of at least 1. So all vertices in $B \setminus C$ receive a weight of at least 1. It follows that $w(B \setminus C) \geq |B \setminus C| = |B| - |C \cap B|$. Since $G'[B]$ is not a triangle, $|B| \geq 4$. Hence $w(B \setminus C) \geq \frac{1}{2}|B| + 2 - |C \cap B|$.

If $G'[S] \neq K_1$, Lemma 15 below now shows that $w(S) \geq \frac{1}{2}|S| + 2$. In that case we have

$$\frac{3}{2}|S \setminus S'| + |V \setminus S| = w(S) \geq \frac{1}{2}|S| + 2 \Leftrightarrow$$

$$\frac{3}{2}|S| - \frac{3}{2}|S'| + |V| - |S| \geq \frac{1}{2}|S| + 2 \Leftrightarrow$$

$$|V| - 2 \geq \frac{3}{2}|S'| \Leftrightarrow |S'| \leq (2|V| - 4)/3.$$

If $G'[S] = K_1$, then this final inequality follows easily since $|V| \geq 4$. $\square$

We continue to prove the two lemmas that were used in the above proof. We adopt the notation from the proof. In particular, for a block $G'[B]$ of $G'[S]$, the set L is the set of vertices in B that have two neighbors in B , and $H = B \setminus L$ is the set of vertices with at least three neighbors in B . The first lemma gives a lower bound for the number of L -vertices in a block of $G'[S]$.

LEMMA 14. *Let CD -set S and graph G' have the properties stated in Theorem 12. For every block $G'[B]$ in $G'[S]$, $|L| \geq |H|$. If $G'[B]$ contains no triangles, then $|L| \geq |H| + 1$.*

Proof. Let $G'' = G'[B]$. Consider two neighbors $u, v \in H$. If $d_{G'}(u) \geq 4$, then $d_{G'}(v) = 3 = d_{G''}(v)$, since uv is not a bridge of $G'[S]$ (property 1 from Theorem 12). Then v is not a dominator or connector since all its neighbors are in B . This is a contradiction with property 2. So both u and v have degree three in G' and therefore are not dominators or connectors. So by property 3, they form a 2-cut in $G'[S]$ and also in G'' . This shows that all neighbor pairs $u, v \in H$ form a 2-cut in G'' and have degree three in G' .

We show that it follows that every vertex in H has at most one neighbor in H . Suppose $u \in H$ has two neighbors v and w in H . $G'' - u - v$ is disconnected, and u and v both have exactly one neighbor in both components. Therefore, if u and v share a neighbor, this must be a vertex in L since G'' is 2-connected. So w is not a neighbor of v . Now it can be checked that $G'' - u - w$ is still connected, a contradiction with the statement from the previous paragraph.

So every vertex in H has at most one neighbor in H and therefore at least two neighbors in L . Clearly, every vertex in L has at most two neighbors in H . It follows that $|L| \geq |H|$.

In addition, if $|L| = |H|$, then every vertex in H has exactly one neighbor in H , and every vertex in L has two neighbors in H . We show that it follows that G'' contains a triangle. Consider two neighbors $u, v \in H$. Let C_1 and C_2 be the two components of $G'' - u - v$, such that $|V(C_1)| \leq |V(C_2)|$. If C_1 contains vertices of H , then let u' and v' be two neighbors in $H \cap V(C_1)$. Let C'_1 and C'_2 be the components of $G'' - u' - v'$, such that $|V(C'_1)| < |V(C_1)|$. We can continue to find new neighbor pairs in H like this, until the smallest component no longer contains vertices of H . Since every vertex in L has two neighbors in H , this component consists of a single vertex, and we have found a triangle. $\square$

In the proof of Theorem 13, we started with a certain weight assignment and proved three properties for this. We anticipated in the proof of Theorem 13 that this yields that the total weight of S is at least $\frac{1}{2}|S| + 2$, so slightly more than $\frac{1}{2}$ on average. The following lemma proves this statement.

LEMMA 15. Let $G = (V, E) \neq K_1$ be a connected graph with nonnegative weights w on the vertices such that the following hold:

1. If $d(v) = 2$, then $w(v) \geq \frac{1}{2}$.
2. If $d(v) = 1$, then $w(v) \geq \frac{3}{2}$.
3. If $G[B]$ is a block of G , and C is the set of cut vertices of G , then $w(B \setminus C) \geq \frac{1}{2}|B| + 2 - |C \cap B|$.

For this graph, $w(V) \geq \frac{1}{2}|V| + 2$.

Proof. We prove the statement by induction on $|V|$. If G is 2-connected, then the statement follows immediately from property 3. If $G = K_2$, then the statement is clearly true.

In any other case, G has at least one leaf u such that $G - u$ is not K_1 , or it has at least one block with exactly one cut vertex (Lemma 6). We consider these two cases.

Suppose first that u is a leaf of G , with neighbor v . Consider the graph $G' = G - u$ with weight function $w'(v) = w(v) + w(u) - \frac{1}{2}$, and $w'(x) = w(x)$ for all other vertices. We prove that for G' and w' , the three properties hold again, so we can use induction. Note that since $w(u) \geq \frac{3}{2}$, $w'(v) \geq w(v) + 1$.

If v has degree two in G' , then since $w'(v) \geq 1$, the first property holds for v . If v has degree one in G' , then $w(v) \geq \frac{1}{2}$ (since v has degree two in G), so $w'(v) \geq \frac{3}{2}$, and the second property holds for v . For all other vertices, the weights and degrees do not change, so these two properties hold for all vertices.

For a block $G'[B]$, it is obvious that property 3 still holds, unless $v \in B$ and v is not a cut vertex anymore in G' . In this case, let C_G (resp., $C_{G'}$) denote the set of cut vertices of G and G' ($C_G = C_{G'} + v$):

$$w'(B \setminus C_{G'}) \geq w(B \setminus C_G) + 1 \geq \frac{1}{2}|B| - |C_G \cap B| + 3 = \frac{1}{2}|B| - |C_{G'} \cap B| + 2,$$

so property 3 holds.

We conclude that if G has a leaf u , then we can construct a new graph $G' = G - u$ with weight function w' such that $w'(V(G')) = w(V(G)) - \frac{1}{2}$, for which the three properties hold. $G' \neq K_1$ and G' is connected, so by induction, $w(V(G)) = w'(V(G')) + \frac{1}{2} = \frac{1}{2}|V(G')| + \frac{5}{2} = \frac{1}{2}|V(G)| + 2$, which proves the lemma for this case.

Now we consider the case that G has no leaves. Then G has a block B which contains only a single cut vertex u . Consider the graph $G' = G - (B - u)$ with weight function $w'(u) = \frac{3}{2}$ and $w'(x) = w(x)$ for all other vertices. $G' \neq K_1$, so if we can show that the three properties hold for G' and w' , we can use induction.

If $d_{G'}(u) = 1$ or $d_{G'}(u) = 2$, then w' clearly satisfies the corresponding properties. For a block $G'[B']$, property 3 clearly holds when $u \notin B'$ or u is still a cut vertex in G' . In the other case, let C_G (resp., $C_{G'}$) denote the set of cut vertices of G and G' . Now

$$w'(B' \setminus C_{G'}) = w(B' \setminus C_G) + \frac{3}{2} \geq \frac{1}{2}|B'| - |C_G \cap B'| + \frac{7}{2} > \frac{1}{2}|B'| - |C_{G'} \cap B'| + \frac{5}{2},$$

which is even better than necessary.

Now for G' all weight properties are satisfied and we can use induction: let $V = V(G)$ and $V' = V(G')$:

$$\begin{aligned} w(V) &\geq w'(V') + w(B - u) - \frac{3}{2} \geq \left(\frac{1}{2}|V'| + 2 \right) + \left(\frac{1}{2}|B| + 1 \right) - \frac{3}{2} \\ &= \frac{1}{2}|V'| + \frac{1}{2}|B| + \frac{3}{2} = \frac{1}{2}(|V| - |B| + 1) + \frac{1}{2}|B| + \frac{3}{2} = \frac{1}{2}|V| + 2. \end{aligned}$$

This proves the lemma for the case where G has no leaves, so the lemma is true in all cases. $\square$

In Theorem 13, there is an additional condition that no block of $G'[S]$ contains triangles. Unfortunately, such a G' and S do not always exist for every connected graph G with $\delta(G) \geq 3$. In the following text a potential standard CD-set S that has a realization G' such that $G'[S]$ contains no triangles will be called a *potential standard CD-set without triangles*.

First consider a graph G that contains a cubic triangle consisting of three cut vertices. Every potential standard CD-set S for G contains these three vertices, and every realization contains all incident edges, so G has no potential standard CD-set without triangles. We can even construct planar, 3-connected, cubic graphs that have no potential standard CD-set without triangles. In fact, we prove a stronger statement; we show that it is $\mathcal{NP}$ -complete to decide whether a given graph with these properties has a potential standard CD-set without triangles.

THEOREM 16. *Deciding whether a planar, cubic, 3-connected graph G has a potential standard CD-set S with realization G' such that $G'[S]$ contains no triangles is $\mathcal{NP}$ -complete.*

Proof. The problem of deciding whether a given graph has a spanning subgraph that is a path is called the Hamilton path problem. This problem is $\mathcal{NP}$ -complete even when restricted to planar, cubic, 3-connected graphs [9]. Let $G = (V, E)$ be such a Hamilton path instance. Replacing every vertex with a cubic triangle gives a graph G' , which is again planar, cubic, and 3-connected. Suppose S is a CD-set for G' such that $G'[S]$ contains no triangles (a *CD-set without triangles*). Define $F \subset E$ as follows: if the edge corresponding to $uv \in E$ is part of $G'[S]$, add uv to F . Since S contains at most two vertices of every triangle in G' , $\Delta((V, F)) \leq 2$. Suppose a vertex v with $d(v) = 0$ exists in (V, F) . Then, because S is a dominating set, every neighbor of v has degree one in (V, F) . So (V, F) has maximum degree two and at least three vertices with degree one and therefore has multiple nontrivial components, contradicting the fact that $G'[S]$ is connected. Therefore (V, F) has minimum degree one and maximum degree two and is connected, which gives a Hamilton path for G . Similarly, a Hamilton path for G can be used to find a CD-set without triangles for G' . This shows that it is $\mathcal{NP}$ -complete to decide whether a planar, cubic, 3-connected graph G' has a CD-set without triangles.

For the constructed graph G' , it can be checked that a CD-set without triangles exists if and only if a potential standard CD-set without triangles exists, which proves the $\mathcal{NP}$ -completeness of the given problem. $\square$

So we cannot hope to improve our CD-set construction method to always find potential standard CD-sets without triangles. However, for graphs without triangles, Theorem 12 and Theorem 13 together immediately prove Theorem 7.

THEOREM 7. *Every connected, triangle-free graph on n vertices with $\delta \geq 3$ has a CD-set S with $|S| \leq (2n - 4)/3$.*

5. Small CD-sets for graphs with few cubic diamonds. In order to give a clear overview of the new ideas in this section, we will be a little sloppy; we will, for instance, omit the constants from the bounds we discuss. The correct and detailed reasoning can be found in the proof sketches.

In the previous section we showed that every connected graph with $\delta \geq 3$ without triangles, on n vertices, has a CD-set containing at most $\frac{2}{3}n$ vertices. This bound was proved by letting leaves distribute a weight of 1 among their CD-set neighbors, and showing that these weights can be distributed over the CD-set such that every CD-set

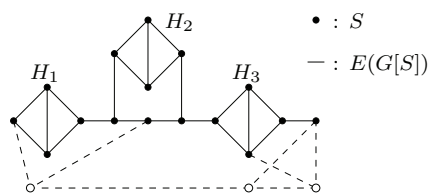
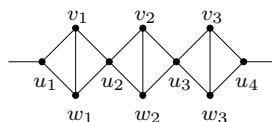
FIG. 5.1. $G[S]$ contains one cubic diamond block.

FIG. 5.2. A diamond necklace.

vertex receives a weight of at least $\frac{1}{2}$. Triangles in the given graph G caused a problem only when they appeared in blocks of $G'[S]$, where S is the potential standard CD-set that was considered, and G' its realization.

In this section we will construct a CD-set with at most $\frac{5}{7}n$ vertices for graphs that may have triangles. For this, we use the same counting method, but instead we have only to assign a weight of $\frac{2}{5}$ to each CD-set vertex. With this more relaxed goal, only two possible blocks of $G'[S]$ cannot be handled with (a slightly more detailed analysis of) the method from the previous section: we have a weight deficiency for blocks of $G'[S]$ that are cubic triangles in G' and blocks of $G'[S]$ that are cubic diamonds in G' .

DEFINITION 17. A subgraph H of G is a cubic diamond if H is a diamond that is induced by four vertices of degree three in G .

Let $S \subseteq V(G)$. Subgraph H of $G[S]$ is a cubic diamond block of $G[S]$ if H is a cubic diamond with respect to G and a block with respect to $G[S]$. (Vertices of H may have degree two in $G[S]$.)

Figure 5.1 shows an illustration of cubic diamond blocks: H_1 is a cubic diamond block of $G[S]$ that contains one vertex with degree two in $G[S]$, but H_2 and H_3 are not cubic diamond blocks.

In the previous section, we used an arbitrary minimal potential standard CD-set S to find a CD-set that satisfied the bound, using three properties we deduced for such a set S , and the chosen realization G' . In this section, we consider a potential standard CD-set S that has these properties plus some additional properties, which enables us to account for triangle and diamond blocks in $G'[S]$.

We consider again the examples from section 1 that show that Theorem 1 is best possible. Recall that these are graphs that have no CD-set with fewer than $\frac{3}{4}n$ vertices, and that these examples contain many cubic diamonds. So when G contains many cubic diamonds, we cannot prove the desired bound of $\frac{5}{7}n$. Instead we show that a CD-set exists that contains at most $\frac{5}{7}n + \frac{1}{7}D$ vertices, where D is the number of cubic diamonds in the given graph G . This is equal to $\frac{3}{4}n$ when D is maximum (when $D = n/4$).

To prove this bound, we need to ensure that for the potential standard CD-set S of G and realization G' we consider, cubic diamond blocks in $G'[S]$ only occur when they cannot be avoided. It turns out that we can always avoid cubic diamond blocks in $G'[S]$, except those that are part of a *diamond necklace* of G (see Figure 5.2).

DEFINITION 18. An induced subgraph H of G is a diamond necklace in G if

for some $k \geq 1$, the vertices of H can be labeled $u_1, \dots, u_{k+1}, v_1, \dots, v_k, w_1, \dots, w_k$, such that the edge set of H consists exactly of $u_i v_i, u_i w_i, v_i w_i, v_i u_{i+1}, w_i u_{i+1}$ for $i = 1, \dots, k$, and such that $u_2, \dots, u_k$ have degree four in G , and all other vertices of H have degree three in G .

Note that a cubic diamond is also a diamond necklace. The new property 4 in Theorem 21 below shows that the only cubic diamond blocks that occur in $G'[S]$ are part of a diamond necklace of G .

The next new property in Theorem 21 is introduced to solve the weight deficiency of triangle blocks of $G'[S]$ in our counting method. Observe that for every triangle block in $G'[S]$, there is one 1-leaf in $\bar{S}$; the number of 1-leaves is at least $T + 2$, where T is the number of triangle blocks in $G'[S]$. If these additional 1-leaves assign a weight of $\frac{6}{5}$ instead of 1 to their CD-set neighbors, the weight deficiency for triangle blocks can be solved. A more detailed analysis of our counting method shows that not all 1-leaves need to assign $\frac{6}{5}$, only those directly related to blocks in some way. These will be called *block-leaves* and are defined as follows.

DEFINITION 19. A vertex $v \in S$ is a simple CD-set vertex with respect to the CD-set S for graph G if $d(v) = 3$ and

- $G[S - v]$ has two components, and $S - v$ is a dominating set for G , or
- $G[S - v]$ is connected, and v has exactly one 1-leaf neighbor.

Simple CD-set vertices are those vertices that do not receive additional weight in our counting method.

DEFINITION 20. A 1-leaf v is a block-leaf with respect to the CD-set S for graph G if there is a path P in $G[S]$ with end vertices x and y ($x = y$ is possible), consisting only of simple CD-set vertices, such that

- x is a neighbor of v .
- y is part of a block of $G[S]$, but none of the other vertices of P are part of a block.

Of course we still need to ensure that the total weight assigned is at most equal to the number of leaves, and therefore we compensate this additional weight by letting non-block-leaves (a *non-block-leaf* is a leaf but not a block-leaf) assign only a weight of $\frac{4}{5}$ and choosing a CD-set S such that the number of block-leaves (that assign $\frac{6}{5}$) is at most half of the total number of leaves. This is ensured by property 5 of Theorem 21.

THEOREM 21. A connected graph G with $\delta(G) \geq 3$ has a potential standard CD-set S , with realization G' , with the following properties. Degrees, neighbors, and all other graph related notions are taken with respect to G' .

1. If $u, v \in S$, $d(u) \geq d(v) \geq 4$ and $uv \in E(G')$, then uv is a bridge of $G'[S]$.
2. Every vertex $v \in S$ that is neither a dominator nor a connector has at least three neighbors in S , and all of its neighbors in S have degree three.
3. If $\{u, v\} \subseteq S$, $uv \in E(G')$, and u and v are both neither dominators nor connectors, then $G'[S] - u - v$ is not connected.
4. The number of cubic diamond blocks in $G'[S]$ is at most the number of diamond necklaces in G .
5. The number of block-leaves is at most half the total number of leaves.

In addition, if $|S| = 2$, then $|\bar{S}| \geq 4$.

Proof sketch. The main idea is to consider a potential standard CD-set S and realization G' that are optimal according to the following priorities:

- Minimize $|S|$.
- Maximize the number of 2-leaves in $\bar{S}$.
- Minimize the number of cubic diamond blocks in $G'[S]$.

- Minimize $|E(G'[S])|$.
- Maximize $|E(G')|$.

By this we mean that among all pairs of S and G' that minimize $|S|$, we consider pairs that maximize the number of 2-leaves. Among all these pairs, we choose one that minimizes the number of cubic diamond blocks in $G'[S]$, etc. For S and G' chosen this way, it can be shown that the above properties hold.

The proof of properties 1–3 is very similar to that in the proof of Theorem 12. The proof of the last two properties is more involved; unlike the other properties, these properties do not hold for just any minimal potential standard CD-set. However it can be shown that if these properties do not hold, a small change can be made to the standard CD-set S and realization G' , such that an improved S_2 and G'_2 are found, according to the priorities. By “small change” we mean that $|S_2 \setminus S| \leq 3$ in every case. For details, we refer to [3]. $\square$

Analogous to the use of Theorems 12 and 13, now that the existence of a CD-set with the properties stated in Theorem 21 has been established, we will use these properties to determine an upper bound on the size of such a CD-set.

THEOREM 22. *Let S be a potential standard CD-set for graph $G = (V, E)$, with realization G' , satisfying the properties from Theorem 21. If $S' \subseteq S$ is a minimal CD-set, then $|S'| \leq (5|V| - 12 + D)/7$, where D is the number of cubic diamond blocks in $G'[S]$.*

Proof sketch. In order to prove this theorem, we assign weights w to the vertices of S such that the total weight is at most $|V \setminus S| + \frac{7}{5}|S \setminus S'| + D/5$, and then show that $w(S) \geq \frac{2}{5}|S| + \frac{12}{5}$. Then we can combine the two inequalities to obtain the result.

The weight assignment is as follows: block-leaves assign a weight of $\frac{6}{5}$ to their CD-set neighbor, and all other leaves distribute a weight of $\frac{4}{5}$ equally among their neighbors in S' (vertices in $S \setminus S'$ receive no weight yet). Since at most half of the leaves are block-leaves (property 5 of Theorem 21), we have assigned a total weight of not more than $|V \setminus S|$. Next, we assign a weight of $\frac{7}{5}$ to each vertex in $S \setminus S'$. Vertices in $S \setminus S'$ that are part of a cubic diamond block of $G'[S]$ receive an additional weight of $\frac{1}{5}$.

When $G'[S]$ is not a tree, it can be shown that $w(S) \geq \frac{2}{5}|S| + \frac{12}{5}$ for this weight assignment, using a proof similar to that of Lemma 15. For details we refer to [3]. It follows that

$$\begin{aligned} \frac{7}{5}|S \setminus S'| + D/5 + |V \setminus S| &\geq w(S) \geq \frac{2}{5}|S| + \frac{12}{5} \Leftrightarrow \\ \frac{7}{5}|S| - \frac{7}{5}|S'| + D/5 + |V| - |S| &\geq \frac{2}{5}|S| + \frac{12}{5} \Leftrightarrow \\ |V| + D/5 - \frac{12}{5} &\geq \frac{7}{5}|S'| \Leftrightarrow |S'| \leq (5|V| - 12 + D)/7. \end{aligned}$$

If $G'[S]$ is a tree, then this last inequality can be proved by assigning the weights as we did in the proof of Theorem 13: every leaf distributes a weight of 1 equally among its CD-set neighbors. For details we again refer to [3]. $\square$

In the bound from Theorem 22, the term D is equal to the number of cubic diamond blocks in $G'[S]$. Using property 4 from Theorem 21, we know that for the S and G' we consider, all cubic diamond blocks are part of diamond necklaces in G , so the number of diamond necklaces is an upper bound for D . In the theorem

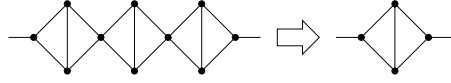


FIG. 5.3. Reducing a necklace containing multiple diamonds.

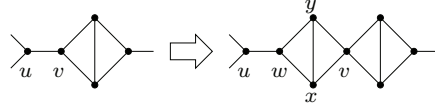


FIG. 5.4. Reversing the necklace reduction.

below, we show that diamond necklaces can be reduced when they consist of more than one diamond, and therefore this term D can even be bounded by the number of cubic diamonds in G (diamond necklaces consisting of one diamond). This yields the desired theorem. Theorem 23 is actually formulated slightly stronger than Theorem 8, where D is the number of cubic diamonds in G . The stronger formulation was used to obtain the improved FPT algorithm [3] (see section 1).

THEOREM 23. *Let G be a connected graph on n vertices with $\delta \geq 3$. G has a CD-set S with $|S| \leq (5n - 12 + D)/7$, where D is the number of cubic diamonds in G that contain three vertices of S .*

Proof. First we reduce diamond necklaces that are not cubic diamonds in G : replace every diamond necklace containing at least two diamonds by a single cubic diamond as shown in Figure 5.3. These are called the *new* cubic diamonds, and the other cubic diamonds are called the *original* cubic diamonds. The resulting graph G_1 is again a simple, connected graph with $\delta(G_1) \geq 3$.

Consider a potential standard CD-set S of G_1 and realization G' that satisfy the properties from Theorem 21. Let D and D' denote the number of original (resp., new) cubic diamonds of G_1 that are a block of $G'[S]$. Since all diamond necklaces are now cubic diamonds, the number of cubic diamond blocks in $G'[S]$ is $D + D'$ (property 4). By Theorem 22, any minimal CD-set $S' \subseteq S$ has $|S'| \leq (5|V(G_1)| - 12 + D + D')/7$. Note that for any minimal CD-set $S' \subseteq S$, the cubic diamonds in G' that contain three vertices of S' are exactly those that are blocks in $G'[S]$.

From S' we can construct a CD-set for G : we can reconstruct G from G_1 by applying the operation illustrated in Figure 5.4 a number of times for every new cubic diamond. This operation introduces three new vertices. At the same time, we can maintain a CD-set by adding at most two new vertices to the CD-set every time we apply the operation: if u and v in Figure 5.4 are not in the previous CD-set, then we add v and x . If one of u and v is in the previous CD-set, then we add w and x . Let k be the number of times this operation has to be applied in order to reconstruct G from G_1 , and construct the corresponding CD-set S'' . Using $k \geq D'$, we can bound $|S''|$ as follows:

$$\begin{aligned} |S''| &\leq |S'| + 2k \leq (5|V(G_1)| - 12 + D + D')/7 + 2k \\ &\leq (5(|V(G)| - 3k) - 12 + D + k)/7 + 2k = (5|V(G)| - 12 + D)/7. \end{aligned}$$

The reconstruction does not affect the original cubic diamonds and also does not affect which vertices of these diamonds are in the CD-set. So D is the number of cubic diamonds in G that contain three vertices of S'' . $\square$

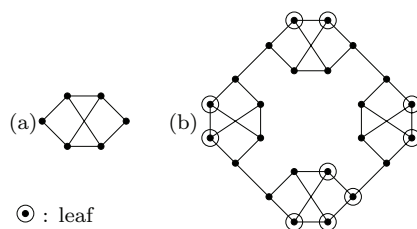


FIG. 6.1. The bound of Theorem 7 is best possible.

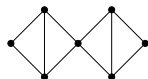


FIG. 6.2. Building block for the worst case examples for Theorem 8.

6. Worst case examples for the CD-set size. In this section we show that the bounds in Theorems 7 and 8 are best possible linear bounds. Throughout this section, n denotes the number of vertices of G .

Theorem 7 shows that every connected, triangle-free graph on n vertices with $\delta \geq 3$ has a CD-set S with $|S| \leq (2n - 4)/3$. We use examples very similar to those mentioned by Griggs, Kleitman, and Shastri [10] to show that this bound is best possible: for $n = 6k$, take k copies of the graph shown in Figure 6.1(a) and connect them in a cyclic form as shown in Figure 6.1(b). The resulting graph is cubic and connected and contains no triangles. It can be checked that it has no CD-set with fewer than $\frac{2}{3}n - 2 = \lfloor (2n - 4)/3 \rfloor$ vertices. This shows that for a linear bound $\alpha n - \beta$, $\alpha = \frac{2}{3}$ is best possible. When we consider the graph Q_3 on eight vertices which has no CD-set with fewer than four vertices, we see that $\beta = \frac{4}{3}$ cannot be increased.

Theorem 8 shows that every connected graph G with $\delta \geq 3$ has a CD-set S with $|S| \leq (5n - 12 + D)/7$, where D is the number of cubic diamonds in G . To prove that this bound is best possible we use the graph shown in Figure 6.2 as a building block, together with diamonds. For any D and $n = 4D + 7k$, take D diamonds and k copies of the graph in Figure 6.2, and connect them in a cyclic form similar to before. The resulting graph is connected, has $\delta = 3$, and has D cubic diamonds. It can be checked that it has no CD-set with fewer than $5k + 3D - 2 = \frac{5}{7}n + \frac{1}{7}D - 2 = \lfloor (5n + D - 12)/7 \rfloor$ vertices. It follows that for a bound of the form $\alpha n + \gamma D - \beta$, the choices $\alpha = \frac{5}{7}$ and $\gamma = \frac{1}{7}$ are best possible. Q_3 again shows that $\beta = \frac{12}{7}$ cannot be improved.

7. An algorithmic viewpoint. Let S be a potential standard CD-set for graph G , with realization G' , that satisfy the properties from Theorem 12, such that $G'[S]$ contains no triangles, and let $S' \subseteq S$ be a minimal CD-set. Theorem 13 then states that $|S'| \leq (2n - 4)/3$ holds, with $n = |V(G)|$. In the proof of Theorem 12, it is shown that for every 2-optimal potential standard CD-set S with realization, a new realization G' can be found such that the properties from Theorem 12 hold. Now we will show that in polynomial time, we can find a 2-optimal potential standard CD-set S in a connected graph with minimum degree at least three. First we show that potential standard CD-sets can be recognized in polynomial time.

LEMMA 24. *It can be checked in polynomial time whether $S \subseteq V(G)$ is a potential standard CD-set for G , and if so, a maximal realization G' of S can be found in polynomial time.*

Proof. We construct an auxiliary bipartite graph H with vertex set $A \cup B$: for every vertex $v \in S$, we add $d(v) - 3$ vertices to A (only vertices with degree at least four will correspond to vertices in A). For every i -leaf in $\bar{S}$ with $i > 2$, we add $i - 2$ vertices to B . We join vertices in A to vertices in B if in G edges exist between the corresponding vertices.

We show that a realization G' of S exists if and only if H has a matching M that saturates every vertex in B . If M is such a matching, delete every edge in G that corresponds to an edge of M , to obtain G' . For every former i -leaf (with $i > 2$), $i - 2$ incident edges are deleted since M saturates B , so only 1-leaves and 2-leaves remain. For every vertex in S with original degree d , at most $d - 3$ incident edges are deleted, so vertices in S still have degree at least three. Clearly, S is still a CD-set in G' .

Similarly, every maximal realization of S corresponds to a B -saturating matching in H .

The existence of polynomial time algorithms now follows from the well-known fact that polynomial time algorithms exist for finding a maximum matching (in fact, for bipartite graphs a specialized algorithm exists) and that if a B -saturating matching exists, every maximum matching saturates B . $\square$

THEOREM 25. *For a connected, triangle-free graph G on n vertices with $\delta(G) \geq 3$, in polynomial time a CD-set S' for G with $|S'| \leq (2n - 4)/3$ can be constructed.*

Proof. The algorithm is as follows. Start with $S = V(G)$. Since G is connected and $\delta(G) \geq 3$, this S is a potential standard CD-set for G . Repeatedly remove one or two vertices from S while maintaining a potential standard CD-set. Since we remove only at most two vertices at a time, and potential standard CD-sets can be recognized in polynomial time (Lemma 24), this can be done in polynomial time. When we have a 2-optimal potential standard CD-set, set $S' = S$ and then repeatedly remove vertices from S until a minimal CD-set S' is obtained.

At this point, S is a 2-optimal potential standard CD-set for G , and $S' \subseteq S$ is a minimal CD-set for G . In the proof of Theorem 12 it is shown that a realization G' can be constructed for S such that S and G' satisfy the properties stated in Theorem 12. Since G contains no triangles, $G'[S]$ contains no triangles. Then by Theorem 13, $|S'| \leq (2n - 4)/3$. $\square$

A CD-set that satisfies the bound from Theorem 23 can also be found in polynomial time, since the proof of Theorem 21 is constructive. However, in this case a local search procedure is needed. The main idea of this local search procedure is as follows: for a given standard CD-set S and realization G' , we can check in polynomial time whether there is an improved standard CD-set S_2 and realization G'_2 with $|S_2 \setminus S| \leq 3$, according to the priorities stated in the proof of Theorem 21. When considering the detailed proof of Theorem 21, it can be seen that considering this neighborhood is enough to find a standard CD-set and realization that satisfy the desired properties. For details we again refer to [3].

THEOREM 26. *Let G be a connected graph with $\delta(G) \geq 3$. We can find in polynomial time a CD-set S for G with $|S| \leq (5|V(G)| - 12 + D)/7$, where D is the number of cubic diamonds in G that contain three vertices of S .*

8. Discussion. In this paper, we introduced new techniques to find small CD-sets, and spanning trees with many leaves. We feel that our methods show that considering small CD-sets is more practical for this purpose than looking at leafy spanning trees, which was previously done to deduce bounds of this kind. For instance, the concept of minimality of a CD-set proves to be very useful, but no similar concept exists for spanning trees. In addition, there are simple but powerful ways of doing

local search on minimal CD-sets, which again have no easy and equally powerful equivalent in the world of leafy trees. To put it differently, spanning trees contain too much information for our purposes: it does not matter which leaves are connected to which CD-set vertices or which tree is chosen to connect the CD-set vertices. These techniques may be useful to prove other similar results: for instance the conjecture of Linial (see section 1) for other small values of δ , or the conjecture from [7] that for cubic graphs a $3/2$ -approximation exists for finding leafy trees.

We mentioned that for our two bounds (Theorem 7 and 8) and the bound in Theorem 2, Q_3 is the only known example that prevents us from finding a stronger linear bound. If we exclude this graph and a few other small graphs (in [10], $K_2 \times C_5$ and another graph on ten vertices are mentioned), it may be possible to show that the bounds become $|S'| \leq \frac{2}{3}n - 2$ (resp., $|S'| \leq \frac{5}{7}n + \frac{1}{7}D - 2$). We have shown in section 6 that these two bounds cannot be improved any further unless we exclude infinitely many graphs. Again we do not expect an easy proof for these improved bounds.

Addendum. Recently, together with Florian Zickfeld we have been able to improve the results from the current paper; see [5]. Previously we conjectured that all graphs with $\delta \geq 3$ without diamond necklaces have a spanning tree with at least $(2n - 4)/3$ leaves [3]. In [5] it is shown that this is not true, but the bound does hold when one more special subgraph is forbidden, which is called a *2-blossom*. It is shown that graphs without diamond necklaces and 2-blossoms have a spanning tree with at least $(2n_{\geq 3} - 4)/3$ leaves, where $n_{\geq 3}$ is the number of vertices in the graph with degree at least three. This result is used to find an FPT algorithm for the problem mentioned in the introduction, with parameter function $f(k) = O(\binom{3k}{k}) \approx O(6.75^k)$.

REFERENCES

- [1] N. ALON, *Transversal numbers of uniform hypergraphs*, Graphs Combin., 6 (1990), pp. 1–4.
- [2] J. A. BONDY AND U. S. R. MURTY, *Graph Theory with Applications*, American Elsevier, New York, 1976.
- [3] P. S. BONSMMA, *Sparse Cuts, Matching-Cuts and Leafy Trees in Graphs*, Ph.D. thesis, University of Twente, Enschede, the Netherlands, 2006. <http://purl.org/utwente/57117>.
- [4] P. S. BONSMMA, T. BRUEGGEMANN, AND G. J. WOEGINGER, *A faster FPT algorithm for finding spanning trees with many leaves*, in MFCS 2003 (Bratislava), Lecture Notes in Comput. Sci. 2747, Springer, Berlin, 2003, pp. 259–268.
- [5] P. S. BONSMMA AND F. ZICKFELD, *Spanning trees with many leaves in graphs without diamonds and blossoms*, <http://arxiv.org/pdf/0707.2760> (ext. abs. accepted for LATIN 2008), 2007.
- [6] Y. CARO, D. B. WEST, AND R. YUSTER, *Connected domination and spanning trees with many leaves*, SIAM J. Discrete Math., 13 (2000), pp. 202–211.
- [7] J. R. CORREA, C. G. FERNANDES, M. MATAMALA, AND Y. WAKABAYASHI, *A $5/3$ -approximation for finding spanning trees with many leaves in cubic graphs*, in Approximation and Online Algorithms, Proceedings of WAOA 2007, Lecture Notes in Comput. Sci. 4927, C. Kaklamanis and M. Skutella, eds., Springer, Berlin, 2008, pp. 184–192.
- [8] V. ESTIVILL-CASTRO, M. R. FELLOWS, M. A. LANGSTON, AND F. A. ROSAMOND, *FPT is P-time extremal structure I*, in ACiD 2005 (Durham), vol. 4 of Texts in Algorithmics, London, 2005, King's College Publications, pp. 1–41.
- [9] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability*, W. H. Freeman, San Francisco, 1979.
- [10] J. R. GRIGGS, D. J. KLEITMAN, AND A. SHASTRI, *Spanning trees with many leaves in cubic graphs*, J. Graph Theory, 13 (1989), pp. 669–695.
- [11] J. R. GRIGGS AND M. WU, *Spanning trees in graphs of minimum degree 4 or 5*, Discrete Math., 104 (1992), pp. 167–183.
- [12] S. GUHA AND S. KHULLER, *Approximation algorithms for connected dominating sets*, Algorithmica, 20 (1998), pp. 374–387.
- [13] D. J. KLEITMAN AND D. B. WEST, *Spanning trees with many leaves*, SIAM J. Discrete Math., 4 (1991), pp. 99–106.

- [14] N. LINIAL AND D. STURTEVANT, 1987, unpublished.
- [15] H. LU AND R. RAVI, *Approximating maximum leaf spanning trees in almost linear time*, J. Algorithms, 29 (1998), pp. 132–141.
- [16] L. RUAN, H. DU, X. JIA, W. WU, Y. LI, AND K. KO, *A greedy approximation for minimum connected dominating sets*, Theoret. Comput. Sci., 329 (2004), pp. 325–330.
- [17] R. SOLIS-OBA, *2-approximation algorithm for finding a spanning tree with maximum number of leaves*, in Algorithms—ESA '98 (Venice), Lecture Notes in Comput. Sci. 1461, Springer, Berlin, 1998, pp. 441–452.

NEAR-UNANIMITY FUNCTIONS AND VARIETIES OF REFLEXIVE GRAPHS*

RICHARD C. BREWSTER[†], TOMAS FEDER[‡], PAVOL HELL[§], JING HUANG[¶], AND
GARY MACGILLIVRAY[¶]

Abstract. Let H be a graph and $k \geq 3$. A *near-unanimity function of arity k* is a mapping g from the k -tuples over $V(H)$ to $V(H)$ such that $g(x_1, x_2, \dots, x_k)$ is adjacent to $g(x'_1, x'_2, \dots, x'_k)$ whenever $x_i x'_i \in E(H)$ for each $i = 1, 2, \dots, k$, and $g(x_1, x_2, \dots, x_k) = a$ whenever at least $k - 1$ of the x_i 's equal a . Feder and Vardi proved that, if a graph H admits a near-unanimity function, then the homomorphism extension (or retraction) problem for H is polynomial time solvable. We focus on near-unanimity functions on reflexive graphs. The best understood are reflexive chordal graphs H : they always admit a near-unanimity function. We bound the arity of these functions in several ways related to the size of the largest clique and the leafage of H , and we show that these bounds are tight. In particular, it will follow that the arity is bounded by $n - \sqrt{n} + 1$, where $n = |V(H)|$. We investigate substructures forbidden for reflexive graphs that admit a near-unanimity function. It will follow, for instance, that no reflexive cycle of length at least four admits a near-unanimity function of any arity. However, we exhibit nonchordal graphs which do admit near-unanimity functions. Finally, we characterize graphs which admit a conservative near-unanimity function. This characterization has been predicted by the results of Feder, Hell, and Huang. Specifically, those results imply that, if $P \neq NP$, the graphs with conservative near-unanimity functions are precisely the so-called bi-arc graphs. We give a proof of this statement without assuming $P \neq NP$.

Key words. graph homomorphism, near unanimity function, homomorphism extension, retraction, dichotomy

AMS subject classifications. 05Cxx, 68Q10, 68W25

DOI. 10.1137/S0895480103436748

1. Introduction. We consider finite undirected graphs without multiple edges, but with loops allowed. A graph in which no vertex has a loop is called *irreflexive*, and a graph in which every vertex has a loop is called *reflexive*. When we say a graph satisfies a property, such as being connected, a tree, a cycle, etc., we mean that the underlying irreflexive graph (i.e., the graph obtained from it by deleting all loops if there are any) has the property.

Given graphs G and H , with lists $L(v) \subseteq V(H)$, for each $v \in V(G)$, a *list homomorphism* of G to H with respect to the lists L is a function $f : V(G) \rightarrow V(H)$ which satisfies the following two properties:

- (i) $f(v) \in L(v)$ for all $v \in V(G)$;
- (ii) $f(u)f(v) \in E(H)$ for all $uv \in E(G)$.

Note that a list homomorphism can map two adjacent vertices of G to the same vertex of H only if the vertex of H has a loop, and, in particular, it must map any vertex

*Received by the editors October 28, 2003; accepted for publication (in revised form) January 4, 2008; published electronically May 21, 2008. This work was partially supported by the Natural Science and Engineering Research Council of Canada.

<http://www.siam.org/journals/sidma/22-3/43674.html>

[†]Department of Math. and Stats., Thompson Rivers University, Kamloops, BC, Canada, V2C 3H5 (rbrewster@tru.ca).

[‡]Department of Computer Science, Stanford University, 268 Waverley St., Palo Alto, CA, 94301 (taras@theory.stanford.edu).

[§]School of Computing Science, Simon Fraser University, Burnaby, BC, Canada, V5A 1S6 (pavol@cs.sfu.edu).

[¶]Department of Math. and Stats., University of Victoria, Victoria, BC, Canada, V8W 3P4 (jing@math.uvic.ca, gmacgill@uvic.ca).

of G with a loop to a vertex of H with a loop. List homomorphisms are introduced in [9].

For a fixed graph H , the *list homomorphism problem*, LIST-HOM H , asks whether an input graph G , together with lists L , admits a list homomorphism to H with respect to the given lists. The complexity of all list homomorphism problems has recently been classified [11]: LIST-HOM H is polynomial time solvable when H is a *bi-arc graph* and is NP-complete when H is not a bi-arc graph. (The definition of bi-arc graphs appears in section 5.)

In the case when the input lists $L(v) = V(H)$ for all $v \in V(G)$, the list homomorphism problem is the *homomorphism problem*, HOM H , or the *H -coloring problem*. When $H = K_n$, the irreflexive complete graph on n vertices, the homomorphism problem HOM H becomes the *n -coloring problem*, which is polynomial time solvable when $n \leq 2$ and NP-complete when $n \geq 3$. The complexity of all HOM H problems has been classified by Hell and Nešetřil [16]: HOM H is polynomial time solvable when H is bipartite or contains a loop and is NP-complete if H is irreflexive and not bipartite.

The *homomorphism extension problem*, EXTH, another special case of list homomorphisms, is of particular interest. In EXTH the inputs are restricted so that each list is either a singleton set or the entire set $V(H)$. Extension problems obviously correspond to questions of extending a given partial mapping (“precoloring”) to a homomorphism and have been historically studied under an equivalent formulation called *retract problems*, RETH; cf. [7, 3, 15, 21, 30]. The *retract problem*, RETH, for a fixed graph H takes as an input a graph G containing H as a subgraph and asks whether or not there is a homomorphism f of G to H such that $f(v) = v$ for all $v \in V(H)$. Such a homomorphism f is called a *retraction* of G to H . If there is a retraction of G to H , then H is called a *retract* of G .

It seems difficult to classify the complexity of all extension problems. In particular, Feder and Vardi [13] have shown that extension problems capture the complexity of the much larger class of all constraint satisfaction problems (CSPs) in the following sense: for each CSP, say, Π , there exists a reflexive graph H such that Π and EXTH are polynomially equivalent. This means that even proving that each extension problem is NP-complete or is solvable in polynomial time would answer a difficult open question in complexity theory [13]. Recall that, by contrast, for list homomorphism problems, we know the exact classification of the complexity [9, 10, 11]. In particular, by techniques similar to [13] it can be seen that if a graph H admits a conservative near-unanimity function (as defined below), then LIST-HOM H is polynomial time solvable. In turn, this implies that RETH and EXTH are also polynomial time solvable. See [17].

Let H be a graph and $k \geq 3$ be an integer. A *near-unanimity function of arity k* (or NUF $_k$ for short) on H is a mapping $g : V(H)^k \rightarrow V(H)$ which satisfies the following properties:

- (i) $g(x_1, x_2, \dots, x_k)$ is adjacent to $g(x'_1, x'_2, \dots, x'_k)$ whenever $x_i x'_i \in E(H)$ for each $i = 1, 2, \dots, k$, and
- (ii) $g(x_1, x_2, \dots, x_k) = a$ whenever at least $k - 1$ of the x_i 's equal a .

Early papers on near-unanimity functions include [1, 20]. Near-unanimity functions of arity 3, also called *majority functions*, are much studied [2, 17, 31]. A near-unanimity function g of arity k on H is a *conservative near-unanimity function* if $g(x_1, x_2, \dots, x_k) \in \{x_1, x_2, \dots, x_k\}$ for all vertices of H^k . It is shown in [11] that all bi-arc graphs admit a conservative near-unanimity function, implying that the corresponding list homomorphism problems can be solved in polynomial time via the

results mentioned above. It is also shown in [11] that graphs H which are not bi-arc graphs have NP-complete list homomorphism problems. Hence, if we assume that $P \neq NP$, then bi-arc graphs are precisely the graphs which admit a conservative near-unanimity function. We will prove this is the case without the assumption $P \neq NP$.

The *categorical product* of a family of graphs, $\{H_i\}_{i \in I}$, denoted $\prod_{i \in I} H_i$, has as its vertex set the Cartesian product $\prod_{i \in I} V(H_i)$. (We restrict our attention to finite products, i.e., $|I| < \infty$.) Two vertices $(g_i)_{i \in I}$ and $(h_i)_{i \in I}$ are adjacent if g_i and h_i are adjacent in each H_i , $i \in I$. We may write $H_1 \times H_2 \times \cdots \times H_k$ for the product of the k graphs, $H_1, H_2, \dots, H_k$, and H^k for the product of k copies of H . Thus, a near-unanimity function of arity k is a homomorphism of H^k to H that is *nearly unanimous*, i.e., satisfies condition (ii) above. Finally, a *graph variety* is a class $\mathcal{V}$ of graphs which contains all products and all retracts of members of $\mathcal{V}$. Given a class of graphs C , the *variety generated by C* is the smallest variety containing all of C . (The intersection of two varieties is itself a variety, and thus the concept of smallest is well defined.)

By abuse of notation we also denote by NUF_k the *class of all graphs that admit a NUF_k* . We let $\text{NUF} = \bigcup_{k=1}^{\infty} \text{NUF}_k$, i.e., the class of graphs each of which admits a near-unanimity function of some arity. We show that, for each fixed $k \geq 3$, the class NUF_k is a variety and that this collection of varieties is strictly monotone, i.e., $\text{NUF}_3 \subset \text{NUF}_4 \subset \cdots$ (with strict inclusions). We show the class of chordal graphs is contained in NUF . It follows that the *variety generated by chordal graphs*, i.e., the smallest variety containing all chordal graphs, is also contained in NUF ; the variety generated by chordal graphs has been further investigated in [27]. The variety generated by *cop-win* or *dismantable* reflexive graphs (see [29]) contains the variety generated by chordal graphs. We give an example of a dismantable graph which does not belong to NUF . An extended examination of the inclusions described here (and of other varieties) is developed in [27]. In particular, NUF is strictly contained in the variety generated by dismantable graphs, and the variety generated by chordal graphs is strictly contained in the variety NUF [27, 25]. A polynomial time algorithm for recognizing graphs in NUF based on dismantability is given in [25].

We give two bounds on the arity of a near-unanimity function of a chordal graph with n vertices, in terms of its clique-size and its *leafage* (defined in section 3.1), respectively. It follows, in particular, that the arity is at most $n - \sqrt{n} + 1$. We present some forbidden substructures for graphs to have a NUF . It follows from these conditions that no reflexive cycle of length at least four admits a near-unanimity function. However, we shall exhibit nonchordal graphs which do admit near-unanimity functions. Finally, we give a proof, without assuming $P \neq NP$, of the result predicted by [11], that the graphs which admit a conservative near-unanimity function are precisely the bi-arc graphs. See also [5, 23].

2. Basic properties.

PROPOSITION 2.1. *For each $k \geq 3$, a graph H admits a NUF_k if and only if each connected component of H admits a NUF_k .*

Proof. Let $H_1, H_2, \dots, H_p$ be the connected components of H . Suppose that $g : H^k \rightarrow H$ is a NUF_k on H . We begin by defining for each H_i a NUF_k , say, f_i , on H_i . Note that by definition g is a homomorphism of H^k to H , and hence any restriction $h = g|_X$ of g to a subgraph X of H^k is a homomorphism $h : X \rightarrow H$.

A vertex of H^k with at least $k - 1$ coordinates equal is called a *nearly unanimous* vertex. For each connected component C of H_i^k , we define f_i on C as follows. If C has a nearly unanimous vertex $x = (x_1, x_2, \dots, x_k)$, define $f_i = g|_C$. By near-unanimity, we

have $f_i(x) \in H_i$. Since $f_i(C)$ is also connected, $f_i(C)$ is a subgraph of H_i . On the other hand, if C contains no such vertex, then define $f_i : C \rightarrow H_i$ by $f_i(z_1, z_2, \dots, z_k) = z_1$. It is easily seen that each f_i is a NUF_k on H_i .

Conversely, suppose $g_i : H_i^k \rightarrow H_i$ is a NUF_k on H_i for each $i = 1, 2, \dots, p$. Let C be a component of H^k , and let $x = (x_1, x_2, \dots, x_k)$ and $y = (y_1, y_2, \dots, y_k)$ be two vertices in C . Since C is connected, there is a walk from x to y in C and, thus, a walk from x_s to y_s in H for each s , where $1 \leq s \leq k$. That is, for each component C of H^k and each coordinate s , there exists a component H_j of H such that $x_s \in H_j$ if and only if $y_s \in H_j$. Hence exactly one of the following conditions holds for all of the vertices of C :

- (i) all k coordinates belong to the same H_j ; i.e., C is a subgraph of H_j^k ;
- (ii) exactly $k - 1$ coordinates belong to the same H_j , and the other coordinate belongs to H_m , $m \neq j$; or
- (iii) at most $k - 2$ coordinates belong to any H_j .

Let $x = (x_1, x_2, \dots, x_k)$ be a vertex in C . In case (i), define $g(x) = g_j(x)$. In case (ii), choose some coordinate t such that $x_t \in H_j$. Define $g(x) = x_t$. (We fix t for the entire component C ; thus, g is simply the projection of C onto its t th coordinate.) Finally, in case (iii), let $g(x) = x_1$.

In all cases, g is a homomorphism. Moreover, g satisfies the near-unanimity condition. In case (i), g inherits the property from g_j . In case (ii), any nearly unanimous vertex in C must have all $k - 1$ coordinates from H_j , including x_t , equal. Finally, in case (iii) there are no nearly unanimous vertices. $\square$

PROPOSITION 2.2. *For each $k \geq 3$, if a graph H admits a NUF_k , then H admits a NUF_{k+1} .*

Proof. Let $g : H^k \rightarrow H$ be a NUF_k on H . Then the function $h : H^{k+1} \rightarrow H$ defined as $h(x_1, x_2, \dots, x_k, x_{k+1}) = g(x_1, x_2, \dots, x_k)$ is a NUF_{k+1} on H . $\square$

The argument above also proves the following.

COROLLARY 2.3. *For each $k \geq 3$, if a graph H admits a conservative NUF_k , then H admits a conservative NUF_{k+1} .*

We provide examples of graphs in $\text{NUF}_{k+1} - \text{NUF}_k$ for each $k \geq 3$ in section 4.1. Thus the converse of Proposition 2.2 does not hold. On the other hand, we will show that the converse of Corollary 2.3 does hold; i.e., H admits a conservative NUF_k for some $k \geq 3$ if and only if H admits a conservative NUF_3 .

PROPOSITION 2.4. *Let G be a graph that admits a NUF_k , and let H be a retract of G . Then H also admits a NUF_k .*

Proof. Let g be a NUF_k on G , and let $r : G \rightarrow H$ be a retraction. Then it is easy to verify that $g' = r \circ (g|_{V(H)^k})$ is a NUF_k on H . $\square$

The above result shows that the class of graphs which admit a NUF_k are closed under retractions. In the case of conservative functions, the class is also closed under taking induced subgraphs.

PROPOSITION 2.5. *Suppose the graph G admits a conservative NUF_k and H is an induced subgraph of G . Then H admits a conservative NUF_k .*

Proof. Let g be a conservative NUF_k for G . The restriction of g to H^k is a near-unanimity homomorphism of H^k to G , and to H as well since g is a conservative NUF . $\square$

PROPOSITION 2.6. *Let $X_1, X_2, \dots, X_n$ be graphs in NUF_k . Then the product $X_1 \times X_2 \times \dots \times X_n$ is also in NUF_k .*

Proof. By associativity of the product, it suffices to verify the claim for two graphs. It is easy to check that if g_X and g_Y are near-unanimity functions of arity k for X

and Y , respectively, then

$$g((x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)) = (g_X(x_1, x_2, \dots, x_k), g_Y(y_1, y_2, \dots, y_k))$$

is a NUF_k for $X \times Y$. $\square$

COROLLARY 2.7. *For each $k \geq 3$, the class NUF_k is a variety.*

A *dominating vertex* of a graph G is one adjacent to all other vertices of G . We make the following observation.

PROPOSITION 2.8. *If the reflexive graph H has a dominating vertex v , then H admits a NUF_k for all $k \geq 3$.*

Proof. By Proposition 2.2 it suffices to show that H admits a NUF_3 . The function

$$g(a, b, c) = \begin{cases} x & \text{if at least two of } a, b, c \text{ equal } x, \\ v & \text{otherwise} \end{cases}$$

is a NUF_3 . $\square$

A graph is *chordal* if it does not contain an induced cycle of length greater than three. We shall show that each chordal graph belongs to NUF , and we shall show that each reflexive cycle of length at least four does not. However, Proposition 2.8 allows us to find examples of nonchordal graphs with a NUF . The *wheel* w_n is the graph obtained from a cycle of length n by adjoining one (new) vertex adjacent to all other vertices.

COROLLARY 2.9. *Each reflexive wheel $w_n, n \geq 4$, is a nonchordal graph in NUF_3 .*

Since NUF_3 is the variety generated by finite paths [2, 17, 19], we see that each reflexive wheel is in the variety generated by finite paths and thus in the variety generated by chordal graphs. In [27] it is shown that the variety generated by chordal graphs is in fact strictly contained in the variety NUF .

3. Reflexive chordal graphs. In this section and the next section, all graphs are assumed to be reflexive unless otherwise stated. We show in this section that every reflexive chordal graph admits a near-unanimity function. (By contrast, in section 4.2 we shall show that no reflexive cycle of length greater than three admits a near-unanimity function.) We provide two bounds on the arity of the NUF . One bound is based on the *leafage* of the graph, and the second is based on the clique-size and is obtained through the study of *tree obstructions*. We remark that the two approaches, leafage and tree obstructions, often allow us to compute the minimum k for which a graph admits a NUF_k . Typically the leafage is used to demonstrate the existence of a NUF_k , and tree obstructions are used to prove the nonexistence of a NUF_{k-1} .

3.1. Arity bounds based on leafage. Let T be a tree. A *subtree* of T is a connected subgraph of T . A *rooted subtree* R of T is a subtree of T with a distinguished vertex called the *root* of R , denoted $r(R)$.

It is well known that a graph H is chordal if and only if it is the *intersection graph* of a family $\mathcal{F}$ of subtrees of a tree; that is, there is a one-to-one correspondence between $V(H)$ and $\mathcal{F}$ such that two vertices of $V(H)$ are adjacent in H if and only if the corresponding subtrees of $\mathcal{F}$ have at least one vertex in common; see [14]. The family $\mathcal{F}$, together with the underlying tree, is called an *intersection representation* of H by subtrees. The *leafage* $l(H)$ of a chordal graph H is the minimum number of leaves of a tree in which H has an intersection representation; cf. [26].

In the following H is a chordal graph. We use $\mathcal{F}$ to denote a (fixed) family of subtrees of a tree T which gives an intersection representation of H . Further, let $\mathcal{T}$

be the set of all rooted subtrees R of T such that $R - r(R)$ is the union of some (zero, one, or more) components of $T - r(R)$, and if $R - r(R)$ is the union of zero components of $T - r(R)$, then $r(R)$ is a leaf of T . For a rooted subtree R , we denote by l_R the number of leaves of T contained in R ; note that these definitions ensure that $l_R \geq 1$ for any rooted subtree R .

Given a collection S of $k + 1$ (not necessarily distinct) subtrees in $\mathcal{F}$, we say that a rooted subtree $R \in \mathcal{T}$ is *critical with respect to S* if it satisfies the following two properties:

1. there are at least $k - l_R + 1$ subtrees of S , each of which contains a vertex of R ;
2. for each $R' \in \mathcal{T}$ contained in $R - r(R)$, there are at most $k - l_{R'}$ subtrees of S , each of which contains a vertex in R' .

The two conditions above are referred to as Properties 1 and 2.

LEMMA 3.1. *Let S be a family of $k + 1$ (not necessarily distinct) members of $\mathcal{F}$. Then*

- (a) *every rooted subtree in $\mathcal{T}$ satisfying Property 1 contains a rooted subtree which is critical with respect to S ; and*
- (b) *there are pairwise vertex disjoint critical rooted subtrees $R_1, \dots, R_p$ such that $T - \bigcup_{i=1}^p R_i$ does not contain any critical rooted subtrees.*

Proof. Let the rooted subtree X satisfy Property 1. If X also satisfies Property 2, then X is critical with respect to S , and we are done. Otherwise, $X - r(X)$ contains another rooted subtree X' which satisfies Property 1. Again, if X' also satisfies Property 2, then X' is critical; otherwise, $X' - r(X')$ contains a third rooted subtree X'' which satisfies Property 1. Continuing this way, we will find a critical subtree with respect to S . A subtree consisting of a single vertex which satisfies Property 1 trivially satisfies Property 2.

To see statement (b), note that the entire tree T (with an arbitrary root) satisfies Property 1. In addition, a rooted tree R which does not satisfy Property 1 cannot contain a subtree R' which satisfies Property 1, since $k - l_R + 1 \leq k - l_{R'} + 1$ in the case that R' is a subtree of R . $\square$

THEOREM 3.2. *Every chordal graph H of leafage k admits a NUF_{k+1} .*

Proof. Let $\mathcal{F}$ be an intersection representation of H by subtrees of a tree T with $k \geq 2$ leaves. We shall show that the intersection graph H of $\mathcal{F}$ admits a NUF_{k+1} . By Proposition 2.1, we may assume that H is connected. Further, we assume that every vertex of T belongs to a subtree in $\mathcal{F}$. For convenience, we shall not distinguish between the vertices of H and the subtrees of $\mathcal{F}$.

Let S be a collection of $k + 1$ subtrees of $\mathcal{F}$. Although some of the $k + 1$ subtrees of S may be the same, we treat them as distinct in the counting below.

By Lemma 3.1, there exist $R_1, R_2, \dots, R_p$ pairwise vertex-disjoint critical rooted subtrees with respect to S such that $T - \bigcup_{i=1}^p R_i$ does not contain any critical rooted subtrees. We claim that there is a subtree in $\mathcal{F}$ containing all the roots $r(R_i)$, $i = 1, 2, \dots, p$. When $p = 1$, this is clearly true as by assumption every vertex of T is in a subtree of $\mathcal{F}$. So assume that $p \geq 2$. In this case, we prove the stronger statement that in fact there is a subtree in S which contains all the roots. Suppose to the contrary that none of the $k + 1$ subtrees of S contains all the roots. Then each subtree of S has vertices in at most $p - 1$ critical rooted subtrees R_i . (Observe that any subtree of S that contains vertices in any two rooted subtrees must in fact contain both roots of the subtrees.) Denote by c_i ($i = 1, 2, \dots, p$) the number of subtrees of S , each of which has a vertex in R_i . Then we have $(p - 1)|S| = (p - 1)(k + 1) \geq \sum_{i=1}^p c_i$.

By Property 1, $c_i \geq k - l_{R_i} + 1$. Thus we have

$$\begin{aligned} (p-1)(k+1) &\geq \sum_{i=1}^p c_i \\ &\geq \sum_{i=1}^p (k - l_{R_i} + 1) \\ &= p(k+1) - \sum_{i=1}^p l_{R_i}, \end{aligned}$$

which gives $\sum_{i=1}^p l_{R_i} \geq (k+1)$. This implies that some leaf of T must be contained in at least two critical rooted subtrees, contradicting the assumption that the critical rooted subtrees are pairwise vertex-disjoint.

Ultimately the goal is to define a NUF on these families of subtrees. Suppose at least k subtrees of S are the same subtree X . We claim X contains all the roots. In fact, a rooted subtree $R \in \mathcal{T}$ is critical with respect to S if and only if $V(R) \cap V(X) = \{r(R)\}$. To see this, observe that R satisfying Property 1 requires $V(X) \cap V(R) \neq \emptyset$ and R satisfying Property 2 ensures $V(X) \cap V(R - r(R)) = \emptyset$. On the other hand, $V(X) \cap V(R) = \{r(R)\}$ implies that R contains vertices from at least $k \geq k - l_R + 1$ elements of S , since $l_R \geq 1$. Also, $R - r(R)$ contains no vertices from the k copies of X , which implies that $R - r(R)$ contains vertices from at most $k - l_{R-r(R)}$ subtrees of S , as $l_{R-r(R)} \geq 1$. Consequently, X contains all the roots.

It remains to define the near-unanimity function g on H . Given $k+1$ vertices of H , consider the corresponding family S of subtrees in $\mathcal{F}$. Decompose T into p critical rooted subtrees $\{R_i\}_{i=1}^p$ as described in Lemma 3.1. Define $g(S)$ as follows: When at least k subtrees of S are the same subtree X , let $g(S) = X$; otherwise, let $g(S)$ be any subtree of $\mathcal{F}$ which contains all the roots $r(R_i)$. It remains to verify that g is a homomorphism from H^{k+1} to H . Thus consider two adjacent vertices in H^{k+1} . That is, let $S = \{U_1, U_2, \dots, U_{k+1}\}$ and $S' = \{V_1, V_2, \dots, V_{k+1}\}$ be two collections of $k+1$ subtrees (from $\mathcal{F}$) such that U_j intersects with V_j for each $j = 1, 2, \dots, k+1$. Again, let $R_1, R_2, \dots, R_p$ be the decomposition of T into critical subtrees *with respect to* S . Suppose to the contrary that $g(S)$ and $g(S')$ do not intersect. Let $P : z_0 z_1 \dots z_d$ be the shortest path from $g(S)$ to $g(S')$ where $z_0 \in V(g(S))$ and $z_d \in V(g(S'))$. Let C be the component of $T - z_d$ containing z_0 and C' be the component of $T - z_0$ containing z_d . Let A be the rooted subtree consisting of C with $r(A) = z_{d-1}$. Then A is a rooted subtree containing $g(S)$ but no vertex from $g(S')$. Similarly, $B = C'$ with $r(B) = z_1$ is a rooted subtree containing $g(S')$ but no vertex from $g(S)$. Since each leaf of T is either in A or in B , we must have $l_A + l_B \geq k$. Since $g(S)$ does not intersect with B , B contains none of the roots of $R_1, R_2, \dots, R_p$. Thus, B cannot satisfy Property 1 with respect to S . This means that S contains at most $k - l_B$ subtrees such that each of them has a vertex in B . In other words, S contains at least $k+1 - (k - l_B) = l_B + 1$ subtrees, none of which has a vertex in B . Thus, these $l_B + 1$ subtrees must all be contained in A . Each of these subtrees intersects a member of S' . Hence S' must contain at least $l_B + 1$ subtrees, each of which contains a vertex in A . On the other hand, a similar argument shows that S' contains at most $k - l_A$ subtrees, each of which has a vertex in A . So we must have $l_B + 1 \leq k - l_A$, i.e., $l_A + l_B \leq k - 1$, in contradiction to the fact that $l_A + l_B \geq k$. $\square$

Lin, McKee, and West [26] proved that for every chordal graph H with n vertices, the leafage is at most $n - \lg n - \frac{1}{2} \lg \lg n + O(1)$. Hence, each chordal graph with n

vertices admits a NUF of arity $n - \lg n - \frac{1}{2} \lg \lg n + O(1)$. We improve this bound in the next section.

The upper bound on the arity in Theorem 3.2 is sharp in the sense that there are chordal graphs of leafage l which do not admit a NUF_l . In section 4.1, we construct families of graphs useful for showing lower bounds, including the one just mentioned.

3.2. Arity bounds based of tree certificates. We recall the problem EXTH from the introduction. An instance of EXTH consists of a graph G together with lists $L(v) \subseteq V(H)$, where each $L(v)$ is either a singleton set or all of $V(H)$. Let $X \subseteq V(G)$ be the set of vertices x for which $L(x)$ is a singleton. Then we may view the function $p : X \rightarrow V(H)$ defined by $p(x) = h \in L(x)$ as a preassignment of images (in H) to the vertices in X . The vertices in X are called *preassigned* vertices. The EXTH problem asks if there exists a homomorphism $f : G \rightarrow H$ which extends the preassignment p (i.e., satisfies $f(x) = p(x)$ for $x \in X$). In the case of a yes instance, we say that p is *extendible (in H)*. In the following we shall use the language of lists, or of extending preassignments, as is convenient.

Before entering the technical details of our work, we outline some key ideas used in the development. (We use the standard notation for $X \subseteq V(G)$, and $G[X]$ denotes the induced subgraph of G with vertex set X .) First, it is clear that, given a preassignment $p : X \rightarrow V(H)$, if p is not a homomorphism of $G[X]$ to H , then p is not extendible. On the other hand, if p is a homomorphism $G[X] \rightarrow H$, then a natural algorithmic idea is to successively extend p by one vertex, i.e., select $v \in V(G) \setminus X$, and define $f(v) = h \in L(v)$ (and set $f(u) = p(u)$ for all $u \in X$) such that $f : G[X \cup \{v\}] \rightarrow H$ is a homomorphism. Clearly the condition we must verify is that $f(u)f(v) \in E(H)$ whenever $uv \in E(G)$ for each $u \in X$. Such a value h is an *allowed image* for v . On the other hand, in searching $L(v)$ for an allowed image for v , we may remove from $L(v)$ any value h' such that $p(u)h' \notin E(H)$ for some $u \in X$, where $uv \in E(G)$. Below we will talk about u *causing* h' to be removed from $L(v)$. Finally, if some $L(v)$ becomes empty by removing nonallowed images, then $p : X \rightarrow V(H)$ cannot extend to a homomorphism of G to H . (The image of v under any homomorphism $\phi : G \rightarrow H$ extending p will always be an allowed image, and thus $\phi(v)$ is never removed from $L(v)$.) This process of removing nonallowed images is known as a *consistency check*; see, for example, [17].

Testing all possible extensions of p to all of G is an exponential process; however, for certain graphs H , as identified below, the extendability of p can be determined by considering only a polynomial number of possible extensions. In particular, for such a graph H and a given instance (G, p) of EXTH, the preassignment $p : X \rightarrow V(H)$ is not extendible to H if and only if there exists a *certificate of nonextendability* in the form of a tree with at most k preassigned leaves (where k is a constant depending on H). The existence of such algorithmically well-behaved NO-certificates yields that EXTH is polynomial.

We begin with the concept of a conflict.

DEFINITION 3.3. Let G and H be graphs. A set $X \subseteq V(G)$ with a preassignment $p : X \rightarrow V(H)$ which is not extendible, such that the restriction of p to any proper subset X' of X is extendible, is called a *conflict in G with respect to H* . The size of the conflict is $|X|$. A graph H has *strict extension-width k* if every conflict (in any graph G with respect to H) has size at most k .

Feder and Vardi [13] give several equivalent descriptions of graphs that admit a NUF_k . In particular, the following connection between strict extension-width and near-unanimity is presented.

THEOREM 3.4 (Feder and Vardi [13]). *A graph H admits a NUF_k if and only if it has strict extension-width $k - 1$.*

A graph has *bounded strict extension-width* if it has strict extension-width k for some integer k . Thus the theorem implies that NUF is precisely the class of graphs with bounded strict extension-width. See [17] for more details on this connection.

Reflexive graphs with a NUF_3 have been most carefully investigated. (Irreflexive graphs with a NUF_3 are characterized in [2].) The class of reflexive graphs with a NUF_3 is known to be the smallest variety containing all reflexive paths [28]. It is also known to be precisely the class of all reflexive graphs H such that H is a retract of any G of which it is an isometric subgraph [19, 28]. What this means in the language of extensions and conflicts is the following. Given a graph G containing H as a subgraph, let $p : X \rightarrow V(H)$ be a preassignment where $X = V(H)$ and p is the identity function. Note that in this context (G, p) is naturally viewed as an instance of the retraction problem for H . Either p is extendible (in the case that H is isometric) or there is a conflict of size two (otherwise). Such a conflict yields a *certificate of nonextendability*. In particular, the certificate is a path P with end vertices u and v preassigned as $p'(u) = a, p'(v) = b$, where a, b are vertices of $X = V(H)$. By taking P so that $d_P(u, v) = d_G(a, b) < d_H(a, b)$, we see that p' is extendible in G but $p \circ p'$ is not extendible in H . Hence, p is not extendible to a homomorphism of G to H ; i.e., H is not a retract of G . (For a description of NUF_k , $k \geq 3$, as a variety generated by some starting set of building blocks, see [12]. See also [19] and [4].)

We have just observed that graphs in NUF_3 have certificates of nonextendability in the form of a path whose end points have been preassigned. We now extend this concept to larger certificates.

DEFINITION 3.5. *A reflexive graph H has extension-width one if for any G and any preassignment $p : X \rightarrow V(H)$, where $X \subseteq V(G)$, either p is extendible or there exist a tree T and a set of vertices $X' \subseteq V(T)$ preassigned by a mapping $p' : X' \rightarrow X$ such that p' is extendible in G but $p \circ p'$ is not extendible in H . The tree T together with the preassignment p' is called a *tree-certificate*. Furthermore, the tree T is minimal if the preassigned vertices are precisely the leaves of T and they form a conflict (in H with respect to T).*

Thus, by the comments above the reflexive graphs with a NUF_3 are precisely the reflexive graphs of extension-width one, where the (minimal) tree-certificates can be chosen to be paths.

In [18] the concept of width one is defined as *tree duality*. A graph H has tree duality if for all G either $G \rightarrow H$ or there exists a tree T such that $T \rightarrow G$ and $T \not\rightarrow H$. The notion of extension-width defined here differs slightly from that in [13]; it has been adapted to extensions (and implicitly to retractions). Thus our condition is that either a preassignment p extends to a homomorphism of G to H or there is a tree T and a preassignment p' from T to the preassigned vertices of G (under p) such that p' extends to a homomorphism of T to G but $p \circ p'$ does not extend to a homomorphism of T to H . Clearly, the existence of a tree-certificate demonstrates that p is not extendible. The proposition below shows that each tree-certificate contains a minimal tree-certificate.

PROPOSITION 3.6. *Let (G, p) be an instance of EXT_H where $X \subseteq V(G)$ is the set of preassigned vertices. Suppose T is a tree and $p' : X' \rightarrow X$ is a preassignment, where $X' \subseteq V(T)$, such that p' is extendible (in G) but $p \circ p'$ is not extendible (in H). Then T contains, as a subtree, a minimal tree-certificate.*

Proof. Let T' be a minimal subtree of T with respect to nonextendibility in H , and let $X'' \subseteq X'$ be the set of preassigned vertices of in T' . We first claim that each vertex in X'' is a leaf. Suppose to the contrary that some vertex $v \in X''$ is not a leaf. Let $T_1, \dots, T_k$ be the subtrees of $T' - v$. Consider $T_i + v$. By the minimality of T' , $T_i + v$ is extendible in H . Moreover, since v is preassigned the same value in $T_i + v$ for all i , we obtained that the preassignment X'' of T' is extendible in H , which is a contradiction.

We claim that all the leaves in T' are preassigned. Suppose to the contrary that some leaf $v \in T'$ is not preassigned. By minimality, the preassignment from $T' - v$ to $V(H)$ is extendible. In particular, the parent of v receives an image. Since H is reflexive, v can receive the same image. Thus this preassignment of X'' to H is extendible, which is a contradiction.

Therefore, T' is a tree whose leaves are preassigned and the leaves form a conflict. That is, T' together with the preassignment is a minimal tree-certificate. $\square$

Tree-certificates are used implicitly in [9] but are first formally defined and studied in [27]. Finally, we remark that K_2 is an example of a graph with width two. That is, for the $\text{HOM}K_2$ problem, i.e., testing if a graph is bipartite, the NO-certificates are odd cycles, i.e., partial two-trees.

A special type of tree-certificate is introduced in [19]. A *hole* in a reflexive graph H is a set Z of vertices with a mapping $\delta : Z \rightarrow \{0, 1, 2, \dots\}$ such that no h in H has the distances $d_H(h, x) \leq \delta(x)$ for all $x \in Z$, but for each proper subset Z' of Z , there is an h satisfying these inequalities for all x in Z' . A *k-hole* is a hole in which the set Z has exactly k vertices. A *unit hole* is a hole in which δ is the constant mapping with range $\{1\}$. If $Z = \{x_1, x_2, \dots, x_k\}$, we may refer to the hole as a $(\delta(x_1), \delta(x_2), \dots, \delta(x_k))$ *k-hole*.

Equivalently, a hole in H is a tree T with exactly one branch vertex (vertex of degree greater than two), together with a precoloring $p : X \rightarrow V(H)$, where X is the set of leaves of T . The length of the path from the branch vertex to the leaf x is precisely $\delta(x)$. Further, (X, p) is a conflict in T with respect to H ; i.e., p is not extendible in H , but any proper subtree of T , with the same precoloring (restricted to the subtree), is extendible in H .

From the above discussion, we obtain the following corollaries of Theorem 3.4.

COROLLARY 3.7. *A reflexive graph that admits a minimal tree-certificate with k preassigned leaves cannot have a NUF_k .*

Proof. The minimal tree-certificate contains a conflict with k vertices, showing the strict extension-width is at least k . $\square$

COROLLARY 3.8. *A reflexive graph with a k -hole cannot have a NUF_k .*

It turns out that the reflexive graphs with a NUF_3 are also characterized as the class of reflexive graphs which do not have holes [19, 28]. Thus a reflexive chordal graph which is not in NUF_3 must have a hole. We shall elaborate on this fact in section 4. Specifically, such graphs must have a unit hole of a particular kind.

We now return our attention to improving the bounds on the strict extension-width of reflexive chordal graphs. In Figure 1 is an algorithm from [9]. This algorithm solves the *connected list homomorphism problem*, denoted by the $\text{CLIST-HOM}H$ problem, where H is a reflexive chordal graph and the lists of any instance induce connected subgraphs of H . The proof of correctness for the algorithm follows from properties of the perfect elimination ordering of H ; see [9]. We present the algorithm with the addition that we explicitly construct a tree-certificate for G in H when G is a NO-instance of the problem. A digraph D is used in the algorithm to retain information

CLIST-HOMH Algorithm [9]

Input: A graph G with lists $L : V(G) \rightarrow \mathcal{P}(V(H))$ such that $L(v)$ induces a connected subgraph of H .

Output: A homomorphism $f : G \rightarrow H$ such that $f(v) \in L(v)$ for all v , or a tree-certificate T proving $G \not\rightarrow H$.

1. Let $h_1, h_2, \dots, h_n$ be a perfect elimination ordering of H (an ordering such that any two neighbors of h_i among $h_{i+1}, h_{i+2}, \dots, h_n$ are adjacent).
2. Set $V(D) = \{t_g : g \in V(G)\}$; $E(D) = \emptyset$.
3. For $i = 1$ to n
 - 3.1 Remove h_i from all lists $L(g)$ in which it is not the only member.
 - 3.2 For those g which have $L(g) = \{h_i\}$:
 - 3.2.1 assign $f(g) = h_i$;
 - 3.2.2 for each g' adjacent to g , remove from $L(g')$ all vertices that are not adjacent to h_i ; and add the arc $t_{g'}t_g$ to D if some vertex is removed from $L(g')$;
 - 3.2.3 delete g from G .
 - 3.3 If some list $L(x) = \emptyset$, then let T be the subgraph of D consisting of descendants of t_x . Answer NO; return T together with the lists L_T where $L_T(t_g)$ equals the original $L(g)$ (provided as input).
4. Answer YES; return f .

FIG. 1. An algorithm for CLIST-HOMH, where H is a reflexive chordal graph, and each list induces a connected subgraph of H .

about which vertices cause the removal of elements from lists (of other vertices). In the case of a NO-instance we prove below, the digraph D contains a tree-certificate (T, p') ; thus, H has extension-width one. Also, we provide bounds on the size of the conflict contained in T , thus providing bounds on the strict extension-width of H . That is, we prove the following.

THEOREM 3.9. *Each reflexive chordal graph H has extension-width one and bounded strict extension-width.*

The proof of the theorem appears below after the development of some preliminary results.

LEMMA 3.10. *At any step of the CLIST-HOMH Algorithm in Figure 1, the digraph D is acyclic.*

Proof. If $t_{g'}t_g$ is an arc of D , then g is removed from G before g' is removed from G . $\square$

LEMMA 3.11. *Suppose the CLIST-HOMH Algorithm in Figure 1 answers NO and returns (T, L_T) for some instance (G, L) of CLIST-HOMH. Then for each vertex $t_g \in V(T)$ other than the root, the corresponding vertex $g \in V(G)$ is assigned to some h_i , i.e., $f(g) = h_i$, by the algorithm. Furthermore, this assignment is injective.*

Proof. Suppose t_{g_1} and t_{g_2} are both children of some vertex t_g in T . The arcs $t_g t_{g_1}$ and $t_g t_{g_2}$ in T are created in D after each t_{g_1} and t_{g_2} are assigned an image at step 3.2.2 of the algorithm. Thus, assume that the assignments $f(g_1) = h_i$ and $f(g_2) = h_j$ are made by the algorithm. Without loss of generality we may assume that h_i precedes h_j in the perfect elimination ordering, or $h_i = h_j$. The arc in T from t_g to t_{g_2} was added to T when the assignment $f(g_2) = h_j$ caused some vertex, say, h , to be removed from $L(g)$. In particular, $h_j h \notin E(H)$ and h appears later in the perfect elimination ordering than h_j . Also, $h_i h \in E(H)$; otherwise, h_i causes h to be removed from $L(g)$ before t_{g_2} is assigned h_j , and thus $t_g t_{g_2}$ would not appear in D .

In addition, note that t_g cannot be assigned h_j . Since h_j caused h to be removed from the list, there must be a vertex in the elimination ordering after both h_j and h that will remain in the list for g (after the current round). Recall that the lists are connected. In particular, adjacent vertices receive unique images.

We claim that h_i and h_j are nonadjacent in H . This is immediate from properties of the elimination ordering and the observations $h_i h \in E(H)$, $h_j h \notin E(H)$. The same argument shows there is no path $h_i = u_1, u_2, \dots, u_m = h_j$ in H where each u_ℓ precedes $u_{\ell+1}$ in the perfect elimination ordering. Indeed, for such a path, $h_j h \notin E(H)$ implies $u_{m-1} h \notin E(H)$, which in turn implies $u_{m-2} h \notin E(H)$. Continuing back to h_i , we conclude $h_i h \notin E(H)$, which is a contradiction.

Finally, there is no vertex h' in the perfect elimination ordering with paths $h = u_1, u_2, \dots, u_{m_1} = h_i$ and $h = w_1, w_2, \dots, w_{m_2} = h_j$ such that each u_ℓ precedes $u_{\ell+1}$ and each w_ℓ precedes $w_{\ell+1}$ in the perfect elimination ordering. Suppose to the contrary that such paths exist. Then u_2 and w_2 are common neighbors of h' and must be adjacent. Without loss of generality u_2 precedes w_2 in the elimination ordering. Thus the paths $u_2, u_3, \dots, u_{m_1} = h_i$ and $u_2, w_2, w_3, \dots, w_{m_2} = h_j$ are paths with the property above, and the first path has been shortened by one vertex. We thus repeatedly shorten the paths until one has length zero, say, the first path, at which point we have a path from h_i to h_j , which is the case analyzed in the previous paragraph.

This last result about two paths ending at h_i and h_j , respectively, shows that the descendants of t_{g_1} and t_{g_2} , respectively, receive images that are disjoint sets of vertices in H , completing the proof that the function f defined in the algorithm is indeed injective. $\square$

We now establish that in the case of a NO-instance, the CLIST-HOMH algorithm does indeed return a tree. Recall that for a list homomorphism problem vertices with lists of size one are called preassigned.

LEMMA 3.12. *If the CLIST-HOMH Algorithm in Figure 1 answers NO and returns (T, L_T) , then T is a tree rooted at t_x , where x is the vertex whose list becomes empty in step 3.3. Moreover, the preassigned vertices are leaves in T , and the preassigned images form an independent set in H .*

Proof. Every vertex in T other than the root t_x has in-degree one. Suppose to the contrary that some vertex t has predecessors t_1 and t_2 . Using the proof of Lemma 3.11, one can easily show that the descendants of t_1 and t_2 are disjoint sets and thus derive a contradiction.

To see that preassigned vertices in T are leaves, observe that internal vertices of T are assigned a value from their list and have something removed from their list by a child. Thus, the list of an internal vertex cannot be a singleton.

Suppose t_{g_1} and t_{g_2} are leaves of T , neither of which is the root. Then they have a common ancestor, say, t_g , such that the (t_g, t_{g_1}) -path and the (t_g, t_{g_2}) -path in T are internally disjoint. Let $t_{g'_1}$ and $t_{g'_2}$ be the first vertex after t_g on the two paths, respectively, with $t_{g_i} = t_{g'_i}$ possible. Without loss of generality, $f(g_1)$ precedes $f(g_2)$ in the perfect elimination ordering. If $f(g_1)f(g_2) \in E(H)$, then there are paths in H from $f(g_1)$ to $f(g'_1)$ and $f(g_1)$ to $f(g'_2)$, contrary to the proof of Lemma 3.11.

Finally, suppose that the root t' is a preassigned vertex. Then the unique child of the root, say, t_g , causes the single element from $L(t')$, say, h' , to be removed when the algorithm makes the assignment $f(g) = h$. Consider some other leaf, say, t_{g_1} , in the tree, including the possibility that t_g is a leaf. Suppose further that the assignment $f(g_1) = h_1$ is made by the algorithm. Then there is a path from h_1 to h in H whose vertices are in increasing order with respect to the perfect elimination ordering. Each vertex on this path is nonadjacent to h' . In particular, $h_1 h' \notin E(H)$. Therefore, the set of all leaves is independent. $\square$

COROLLARY 3.13. *Suppose (G, L) is an instance of EXTH where $X \subseteq V(G)$ is the set of preassigned vertices. Further, suppose the CLIST-HOMH Algorithm in Figure 1 answers NO and returns (T, L_T) . Then T contains as a subtree a minimal tree-certificate.*

Proof. Suppose (T, L_T) is returned by the CLIST-HOMH Algorithm. Because the lists L_T come from L , each is either a singleton or the entire set $V(H)$. Hence (T, L_T) is also an instance of EXTH, in particular, a NO-instance. Note the mapping $t_g \mapsto g$ is an embedding of T in G . Thus, T is extendible in G but not in H . By Proposition 3.6, T contains as a subtree a minimal tree-certificate. $\square$

Proof of Theorem 3.9. Let (G, L) be a NO-instance of EXTH. We will establish that there is a tree-certificate for G with respect to H . Moreover, this tree contains a conflict of size at most $|V(H)|$.

If H is not connected, we consider two cases. First, if some component of G has preassigned vertices in different components of H , then a path between two such vertices is a tree-certificate containing a conflict of size two. On the other hand, if for each component of G its preassigned vertices appear in the same component of H , then (G, L) is a NO-instance if and only if some component of G is a NO-instance for some component of H . Thus, we may restrict our attention to connected graphs G and H . Hence, any instance of EXTH can be viewed as an instance of CLIST-HOMH where each list is a singleton or $V(H)$, and since H is connected, we can apply the algorithm in Figure 1. Given that (G, L) is a NO-instance, the algorithm returns a tree-certificate by Corollary 3.13. Furthermore, this tree has at most $n = |V(H)|$ leaves by Lemma 3.11. The former statement shows that H has extension-width one; the latter statement shows that H has strict extension-width n . $\square$

Theorem 3.9 shows that each reflexive chordal graph with n vertices belongs to NUF_{n+1} . We now improve the bound.

THEOREM 3.14. *Let H be a reflexive chordal graph with n vertices and maximum clique-size $\omega \geq 3$. Then H admits a NUF_k with $k \leq \min\{n - \omega + 1, n - n/(\omega - 1) + 1\}$. In particular, $k \leq n - \sqrt{n} + 1$.*

Proof. First observe that $n - \sqrt{n} \geq \min\{n - n/(\omega - 1), n - \omega\}$. Hence it suffices to prove that the strict extension-width of H is less than or equal to both $n - n/(\omega - 1)$ and $n - \omega$.

Suppose G is a NO-instance of EXTH. By Corollary 3.13, there is a tree-certificate for G . In particular, there is a tree T with lists L from $V(H)$ that is nonextendible in H . Moreover, the preassigned vertices of T are precisely the leaves of T , and they form a conflict.

Suppose s vertices of T are preassigned values in H . These s vertices can only be leaves of T , including possibly the root t of T , if the root has only one child. Also, by Lemma 3.12, the preassigned values form an independent set in H . Thus, a clique with ω vertices in H contains at most one of the s preassigned images; therefore, the remaining vertices in the clique are not among the s vertices, showing that the strict extension-width is at most $s \leq n - \omega + 1$. This establishes that H admits a NUF_k for $k \leq n - \omega + 2$. In the case that $s = n - \omega + 1$, we note that H must be a split graph with s independent vertices and a clique of size $n - s$. However, such a chordal graph has an intersection representation with $s - 1$ leaves. Thus the leafage of H is at most $s - 1$, and we conclude from Theorem 3.2 that H admits a NUF_k where $k \leq n - \omega + 1$.

Let r be the number of children of t , the root of T . The r children eliminate at least r values from the list for t . If $r \neq 1$, then each of these r vertices must remove at least one value from $L(t)$ which is not among the preassigned images for the leaves

of T . Suppose to the contrary that this is not the case. Let t_g be a child of t that removes only values from $L(t)$ that come from the set of preassigned images. Then we can remove the subtree rooted at t_g from T and obtain a smaller tree-certificate, contrary to our assumption that T is minimal. Indeed, each preassigned image, say, h_i , will be removed from the list of all vertices in T at iteration i of the for-loop in the algorithm. Hence, the values that t_g removes from $L(v)$ will all be removed by the algorithm anyway, without the subtree rooted at t_g being present. (Observe that all of the preassigned vertices form an independent set, but the lists are connected. Hence $L(v)$ must contain a vertex that appears later in the perfect elimination ordering than all of the preassigned vertices. Thus v will not be assigned a preassigned image, and they will indeed all be removed from $L(v)$.)

Let i be the number of internal vertices of T (other than the root). Each of these internal vertices is assigned an image which again is not among the preassigned images for the leaves of T . There are l remaining vertices in H that are not one of the $s + i$ assigned images nor one of the r images removed from $L(t)$. Define $p = r + l + i$, unless $r = 1$, in which case $p = l + i$. Thus, $n - s = p$. We will establish the result by showing that $p \geq n/(\omega - 1)$.

Let d be the maximum number of children of a vertex in T other than the root. The number of vertices in H assigned as images to vertices in T other than the root is at most $r + id$, since the root has r children and each of the i internal vertices has at most d children. The r eliminated values from the list for the root t , and the l additional elements counted above, are the only other vertices of H , giving $n \leq 2r + l + id$.

Note finally that $\omega \geq d + 1$. If $d = 1$, the claim is trivial. Suppose $d \geq 2$. Consider the d neighbors of an internal vertex t_g (other than the root) of T . Call these neighbors $t_{g_1}, t_{g_2}, \dots, t_{g_d}$. Further, suppose the algorithm makes the assignments $f(g) = h$ and $f(g_i) = h_i$ for $i = 1, 2, \dots, d$, and suppose these assignments occur in the order $i = 1, 2, \dots, d$. The d children remove d vertices from $L(t_g)$. In particular, say that the assignment $f(g_i) = h_i$ causes x_i to be removed from $L(t_g)$. Each of $x_2, x_3, \dots, x_d$ must appear later in the perfect elimination ordering (by the construction of T) than h_1 . Also, each x_i for $2 \leq i \leq d$ must be adjacent to h_1 ; otherwise, h_1 would have caused x_i to be removed from $L(t_g)$. Hence h_1 is adjacent to h and each of x_i for $2 \leq i \leq d$. Since h_1 precedes the other vertices in the perfect elimination ordering, $h_1, h, x_2, \dots, x_d$ form a clique of size $d + 1$.

We claim $p(\omega - 1) = (p - i)(\omega - 1) + i(\omega - 1) \geq 2r + l + id \geq n$. To see this note that $\omega - 1 \geq d$, and by hypothesis $\omega - 1 \geq 2$. If $r \neq 1$, then $p - i = r + l$, and the claim follows. When $r = 1$ we have $p - i = l$, giving $l(\omega - 1) \geq 2 + l$ (and the claim), except for $l = 1$ and $\omega = 3$, or $l = 0$. First, note that $l \geq 1$; otherwise, the root t and its only child in T would map to the last two vertices in the perfect elimination ordering, which are adjacent.

Hence, assume that $l = 1$ and $\omega = 3$. If some internal vertex of T other than the root has $d' \leq d$ children, then we claim that $n \leq 2r + l + id - (d - d')$. In this final case the difference of one is accounted for unless $d = d' = 2$. This is not possible since the child of the root t would have two children that remove two values from its list, giving $l \geq 2$, contrary to $l = 1$. Therefore, $p(\omega - 1) \geq n$, or $p \geq n/(\omega - 1)$, completing the proof. $\square$

4. Reflexive graphs without NUF.

4.1. Graphs without NUF_k for some k . Having established in the previous sections that each reflexive chordal graph belongs to NUF_k for some k , we now turn our attention to providing lower bounds on the arity of the NUF.

We first show that reflexive chordal graphs without NUF_3 must have a unit hole of a particularly simple kind.

DEFINITION 4.1. *The reflexive chordal graph R_k is formed by constructing a clique on vertices $y_1, \dots, y_k$ and adding independent vertices $x_1, \dots, x_k$ and edges $x_i y_j$ for $i \neq j$.*

THEOREM 4.2. *Let H be a reflexive chordal graph which is not in NUF_3 . There exists $k \geq 3$ such that R_k is an induced subgraph of H . Furthermore, the set $X = \{x_1, x_2, \dots, x_k\} \subseteq V(R_k)$ is the vertex set of a $(1, 1, \dots, 1)$ k -hole in H .*

Proof. It suffices to show that H does indeed contain an induced copy of R_k with no vertex adjacent to all the x_i 's. Clearly, such a copy of R_k is a $(1, 1, \dots, 1)$ k -hole in H .

Since $H \notin \text{NUF}_3$, there exist a graph G and a preassignment $p : X \rightarrow V(H)$, where $X \subseteq V(G)$, $|X| \geq 3$, and X is a conflict. As described in the proof of Theorem 3.9, we may assume that G and H are connected. (A conflict must belong to a single component of G , and conflicts receiving their images in two components of H have size two.) Hence, by Corollary 3.13, there is a tree-certificate (T, p') for G with respect to H . In particular, $p' : X' \rightarrow X$, where X' is the set of leaves of T . Furthermore, $|X'| = |X|$. Suppose to the contrary that $|X'| < |X|$. Since X is a conflict, $p'(X')$ is extendible in H (as it is a proper subset of X). This contradicts the fact that (T, p') is a tree-certificate.

Hence, T is a tree, with at least three leaves which form a conflict with respect to H under the preassignment $p \circ p'$. Without loss of generality, assume T is the smallest tree whose leaves form a conflict of size $|X'|$.

We consider two cases, depending on the degree in T of the root t .

Case 1. $\deg_T(t) = r \geq 3$. Let the children of t be $t_1, t_2, \dots, t_r$. In the execution of the CLIST-HOMH Algorithm in Figure 1 these are mapped to vertices $x_1, x_2, \dots, x_r$ of H , respectively. By (the proof of) Lemma 3.12, $X = \{x_1, x_2, \dots, x_r\}$ is an independent set. Since t is the root of T (meaning that it arose from the first vertex of G whose list became empty), the vertices $x_1, x_2, \dots, x_r$ have no common neighbor in H . By the minimality of T , for $j = 1, 2, \dots, r$ the vertices in $X - \{x_j\}$ have a common neighbor y_j in H .

By the minimality of T , for $j \neq i$ the vertex y_i occurs after x_j in the perfect elimination ordering of $V(H)$. Suppose to the contrary that some child of the root, say, x_i , removes a set of images, say, Y_i , from the list of the root, all of which occur before x_j in the perfect elimination ordering. Then we may remove x_i and all of its descendants from T . The algorithm will remove all elements of Y_i from the list of the roots in Step 3.1 before the assignment $f(t_j) = x_j$ is made. Hence, there is no need to include x_i in the tree, contrary to our assumption of minimality. In particular (since $r \geq 3$), any two of $y_1, y_2, \dots, y_r$ have a common neighbor x_j occurring before them in the perfect elimination ordering. Therefore, $\{y_1, y_2, \dots, y_r\}$ induces a clique in H , and $\{x_1, x_2, \dots, x_r\} \cup \{y_1, y_2, \dots, y_r\}$ induces a copy of R_r , with no vertex in H adjacent to all the x_i 's.

Case 2. $\deg_T(t) = r \leq 2$. Since T has $s \geq 3$ leaves, at least one child of the root, say, t_{g_1} , is the root of a subtree of T consisting of a path from t_{g_1} to a vertex t' with at least two children (where $t_{g_1} = t'$ is possible). Let the children of t' be $t_2, \dots, t_k$ ($k \geq 3$), and assume these vertices have been mapped to $x_2, \dots, x_k$, respectively, by the algorithm.

If $r = 2$, then let t_{g_2} be the other child of t . Further, let h be the image assigned to t_{g_2} in the algorithm, and define v to be the vertex $v = t_{g_2}$. On the other hand,

if $r = 1$, then t is a leaf, and by Corollary 3.13 the original list for t , $L(t)$, must be a singleton, say, $\{h\}$. In this case, define the vertex v to be $v = t$.

Consider the $k - 1$ assigned images $x_2, \dots, x_k$. A vertex x_1 will be defined below. Let y_1 be the value in H assigned to t' by the algorithm. As before, for each i , the $k - 2$ vertices x_j , $j \neq i$, have a common neighbor y_i following y_1 in the perfect elimination ordering, with y_i not adjacent to x_i . Hence, for each $i = 2, \dots, k$, y_1 and y_i have a common neighbor, x_j , $j \neq i$, preceding them in the perfect elimination ordering. Thus, the y_i 's, $i = 1, 2, \dots, k$, form a clique.

Let $d = d_H(y_1, h)$ be the distance in H from y_1 to h . Since T does not map to H , $d_T(t', v) < d_H(y_1, h)$. Let $d' = d_T(t', v)$. Furthermore, for $i \neq 1$, the distance $d_i = d_H(y_i, h)$ must satisfy the inequality $d - 1 \leq d_i \leq d'$. To see the inequality, note that y_1 is adjacent to y_i ; hence, $d - 1 \leq d_i$. Also, by the minimality of T , $T - \{x_i\}$ maps to H with t' mapping to y_i . Hence, $d_i \leq d'$.

Consider a tree T' (whose leaves are vertices in H) consisting of a vertex y^* adjacent to each y_i for $i \geq 2$ together with a path of length $d - 2$ from y^* to h . If T' does not map to H , then, since T' has distances between its leaves at least their distance in H , the tree T' is another example of a tree like T , but smaller, which is a contradiction. Hence, T' maps to H , and, in particular, y^* maps to a vertex x_1 . By the distance inequalities, x_1 is not adjacent to y_1 , but it is adjacent to all y_i with $i \neq 1$.

Finally, consider x_i with $i \neq 1$. The vertex x_i is distance at least d from h in H . This follows from the algorithm and properties of the perfect elimination ordering. If h comes after y_1 in the perfect elimination ordering, then a path from x_i to h in H of length at most $d - 1$ easily yields a path of length at most $d - 1$ from y_1 to h in H , which is a contradiction. Similarly, h preceding y_1 also leads to a contradiction. Since x_1 is at distance $d - 2$ from h in H , it follows that x_1 is not adjacent to any x_i with $i \neq 1$. Furthermore, no vertex w in H is adjacent to all x_i , since otherwise we could map T to H by mapping t' to w at distance $d - 1 = d'$ from h in H . This completes the proof. $\square$

We continue our examination of holes from page 947, where we defined holes and provided a correspondence to trees with one branch point. The above proof shows that for chordal graphs H without NUF_3 , there is an instance of $\text{EXT}H$ whose tree-certificate is a $(1, 1, \dots, 1)$ hole. Historically, holes have been studied in the (equivalent) language of retractions. We adopt the retraction language for this next result. Given a graph G with a subgraph H , we say a hole in H is *filled* if G contains a vertex g whose distances to the vertices of the hole are within the distance given in the definition of the hole, i.e., $d_G(g, x) \leq \delta(x)$ for each x in the hole. Thus, a filled hole corresponds to a tree-certificate where the tree is a subdivision of a star. On the other hand, we now provide a graph G that does not retract to a reflexive chordal graph H , yet G does not fill a hole in H . In essence, the tree-certificate for G is necessarily more complex than a subdivision of a star.

THEOREM 4.3. *There is a reflexive chordal graph H with a NUF_5 (but not a NUF_4) and an instance G of RETH that does not retract to H , yet no vertex of G fills a hole in H .*

Proof. The graph H has ten vertices a_i, b_i, c_i, d_i, e_i for $i = 1, 2$ and edges $a_i c_i, a_i d_i, b_i c_i, b_i e_i, c_i d_i, c_i e_i, d_i e_i, d_1 d_2, e_1 e_2, d_1 e_2$, and $e_1 d_2$. The instance G is H together with three vertices x_1, x_2, y and edges $x_i a_i, x_i b_i, x_i y$. None of x_1, x_2, y fills a hole in H , and yet G does not retract to H . To see that no hole is filled, observe that $d(x_1, v) \geq d(c_1, v)$ for all $v \in V(H)$. Hence x_1 cannot fill a hole in H . A similar

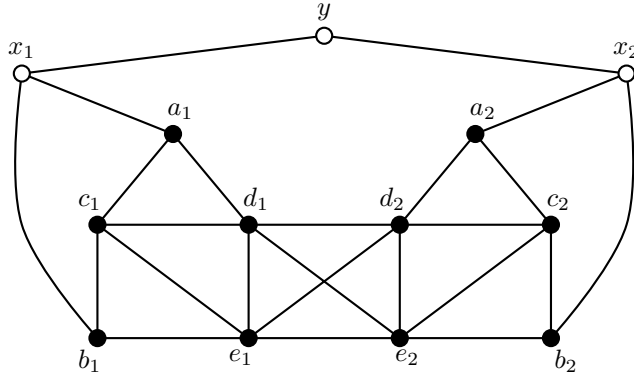


FIG. 2. A G containing a subgraph H induced by the black vertices. The graph H is in $\text{NUF}_5 - \text{NUF}_4$. The graph G is a NO -instance of RETH yet fills no holes in H .

argument shows that neither x_2 nor y fills a hole. Figure 2 contains a drawing of H (with black vertices) and G (consisting of H plus the three white vertices).

To establish that H admits a NUF_5 and not a NUF_4 , we first observe that it is easy to construct an intersection representation for H with leafage 4. Hence, by Theorem 3.2, H admits a NUF_5 . Suppose to the contrary that g is a NUF_4 for H . Then $g(a_1, a_1, a_1, c_1) = a_1$ and $g(c_1, b_1, b_1, b_1) = b_1$. Since (a_1, a_1, a_1, c_1) and (c_1, b_1, b_1, b_1) are both adjacent to (d_1, c_1, c_1, e_1) , we must have $g(d_1, c_1, c_1, e_1)$ adjacent to both a_1 and b_1 . Namely, $g(d_1, c_1, c_1, e_1) = c_1$. However, (d_1, c_1, c_1, e_1) is adjacent to (d_2, d_1, e_1, e_2) . Thus, $g(d_2, d_1, e_1, e_2)$ is adjacent to c_1 . A similar argument shows that $g(d_2, d_1, e_1, e_2)$ is adjacent to c_2 . In particular (a_2, a_2, c_2, a_2) and (b_2, c_2, b_2, b_2) are both adjacent to (c_2, d_2, e_2, c_2) . Hence, $g(c_2, d_2, e_2, c_2) = c_2$ and $g(d_2, d_1, e_1, e_2)$ is adjacent to c_2 (as well as c_1), which is a contradiction. $\square$

We complete this section by establishing that the bounds in Theorems 3.2 and 3.14 are tight.

THEOREM 4.4. *For each $i \geq 0$ and $\omega \geq 3$, there exists a reflexive chordal graph $LB(\omega, i)$ with $n = 2\omega + i(\omega - 1)$, strict extension-width $s = \omega + i(\omega - 2)$, and leafage s .*

Proof. Fix $\omega \geq 3$.

For the case $i = 0$, $LB(\omega, 0)$ is the chordal graph with vertices $x_1, \dots, x_\omega, y_1, \dots, y_\omega$, a clique on the y_l , all edges $x_k y_l$ with $k \neq l$, and all loops, i.e., a copy of R_ω as defined above. To see that $LB(\omega, 0)$ has strict extension-width at least $s = \omega$, consider an instance T (a tree defined below) of $\text{EXTLB}(\omega, 0)$. The tree T has root t adjacent to ω leaves, each assigned to a unique x_k in $LB(\omega, 0)$. Clearly, the preassignment does not extend to a homomorphism $T \rightarrow H$, but if one of the ω leaves in T is deleted, say, the leaf mapped to x_1 , then the root t can be mapped to y_1 . Thus the x_k 's form a conflict of size $s = \omega$. On the other hand, there is an intersection representation of $LB(\omega, 0)$ with s leaves. Hence, by Theorem 3.2, $LB(\omega, 0)$ belongs to NUF_{s+1} . We have $n = 2\omega$, strict extension-width $s = \omega$, and leafage s .

For the case $i = 1$, consider the element x_1 in the preceding construction, with $\omega - 1$ neighbors y_j with $j \neq 1$. To construct $LB(\omega, 1)$, add vertices $z_2, z_3, \dots, z_\omega$ adjacent to x_1 , plus edges $z_k y_l$ for $k \neq l$ and $l \neq 1$. Similarly, replace the leaf in T that is preassigned to x_1 with a vertex t' having the $\omega - 1$ children each preassigned to a unique z_k . Thus n has increased by $\omega - 1$, and the fact that T does not map to $LB(\omega, 1)$ depends on $\omega - 1$ new leaves in T but no longer depends on t' (which is no longer preassigned to x_1). Hence the size of the conflict has increased by at least $\omega - 2$;

however, it is easy to see that the leafage has increased by at most $\omega - 2$. Therefore, s has increased by $\omega - 2$.

Note that the new vertices z_k have $\omega - 1$ neighbors in $LB(\omega, 1)$ as before for the x_k in $LB(\omega, 0)$. We may repeat the process of enlarging each $LB(\omega, i)$ and changing T as follows. In $LB(\omega, i)$ add $\omega - 1$ new vertices, the $z_2, \dots, z_\omega$ above, and join each to some vertex and $\omega - 2$ of its neighbors, as for x_1 above, ensuring that no two of the new vertices have the same neighborhood. In T , select a leaf, say, u , as before, for the preimage of x_1 ; add $\omega - 1$ leaves each adjacent to u , and map each (injectively) to the new $\omega - 1$ vertices, say, $z_2, \dots, z_\omega$, as above. Thus each time we increase n by $\omega - 1$ and s by $\omega - 2$. This completes the construction. $\square$

The bound in Theorem 3.2 is tight.

COROLLARY 4.5. *For each $l \geq 3$, there is a reflexive chordal graph H with $\omega = 3$, and leafage l , such that $H \in \text{NUF}_{l+1} - \text{NUF}_l$.*

Proof. The graph $LB(3, l - 3)$ is such a graph. $\square$

The bounds in Theorem 3.14 are tight.

COROLLARY 4.6. *For each $\omega \geq 2$ there is a reflexive chordal graph $H \in \text{NUF}_k - \text{NUF}_{k-1}$, where $k = n - \omega + 1$.*

Proof. The graph $LB(\omega, 0)$ is such a graph. $\square$

COROLLARY 4.7. *For each $i \geq 0$, there is a reflexive chordal graph $H \in \text{NUF}_k - \text{NUF}_{k-1}$ with $\omega = 3$, where $k = n - n/(\omega - 1) + 1$.*

Proof. The graph $LB(3, i)$ is such a graph. $\square$

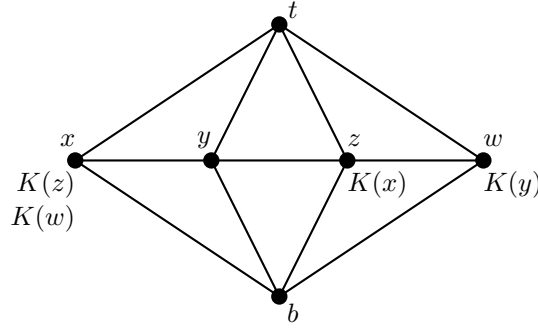
4.2. Graphs without NUF_k for any k . We now construct classes of reflexive graphs that do not admit a near-unanimity function of any arity. Given a vertex v and a subgraph S of a graph H , we use the notation $d(v, H)$ to denote the minimum distance from the vertex v to a vertex in the subgraph H , where the minimum is taken over all vertices of H . We are working with reflexive graphs; however, $N(v)$ denotes the vertices other than v that are adjacent to v , i.e., $v \notin N(v)$.

LEMMA 4.8. *Let H be a connected reflexive graph and t, b be two nonadjacent vertices of H . Suppose that for each $x \in N(t)$ there is a complete subgraph $K(x)$, which is either a maximal clique or a single vertex, with the property that $d(x, K(x)) > \max\{d(t, K(x)), d(b, K(x))\}$. Then H does not admit a NUF_k for any $k \geq 3$.*

(The hypotheses are illustrated in Figure 3, where t is the top vertex, b is the bottom vertex, and for each $x \in N(t)$, $K(x)$ is one of the middle four vertices which is not adjacent to x .)

Proof. Suppose to the contrary that g is a NUF_k on H . For each $j = 1, 2, \dots, k$, let $\bar{x}_j$ be a vertex, say, $(x_1, x_2, \dots, x_k)$, of H^k with the property that $x_i = t$ for all $i < j$, $x_i = b$ for all $i > j$, and x_j is an arbitrary vertex of H . Thus $\bar{x}_j = (t, \dots, t, x_j, b, \dots, b)$. Consider any vertex $x \in N(t)$. By hypothesis there exists a complete subgraph $K(x)$ with $d(x, K(x)) > \max\{d(t, K(x)), d(b, K(x))\}$. Thus there exist $y, z \in V(K(x))$ with $d(x, K(x)) > d(t, y)$ and $d(x, K(x)) > d(b, z)$. Hence in H^k the distance between $\bar{x}_j$, and $(y, \dots, y, x_j, z, \dots, z)$ is less than $d(x, K(x))$. For each $v \in K(x)$ the vertex $(y, \dots, y, x_j, z, \dots, z)$ is adjacent to $(v, \dots, v, x_j, v, \dots, v)$. Hence $g(y, \dots, y, x_j, z, \dots, z)$ must be adjacent to v , and as $K(x)$ is either a single vertex or a maximal complete subgraph of H , $g(y, \dots, y, x_j, z, \dots, z)$ must be a vertex of $K(x)$. We have

$$\begin{aligned} d(x, K(x)) &> d(\bar{x}_j, (y, \dots, y, x_j, z, \dots, z)) \\ &\geq d(g(\bar{x}_j), g(y, \dots, y, x_j, z, \dots, z)) \end{aligned}$$

FIG. 3. A dismantable reflexive graph without a NUF_k .

and, consequently, $g(\bar{x}_j) \neq x$. Since x was an arbitrary member of $N(t)$, we conclude that $g(\bar{x}_j) \notin N(t)$.

Since H is a connected graph, there is a (t, b) -path $(t =)v_1, v_2, \dots, v_l(= b)$. Thus, in H^k , we have the following path, say, P :

$$\begin{aligned} (t, b, b, b, \dots, b, b) &\sim (t, v_{l-1}, b, b, \dots, b, b) \sim (t, v_{l-2}, b, b, \dots, b, b) \sim \dots \\ &\sim (t, v_2, b, b, \dots, b, b) \sim (t, t, b, b, \dots, b, b) \sim (t, t, v_{l-1}, b, \dots, b, b) \sim \dots \\ &\sim (t, t, t, b, \dots, b, b) \sim (t, t, t, v_{l-1}, \dots, b, b) \dots \sim (t, t, t, \dots, t, b). \end{aligned}$$

Since $g(t, b, b, b, \dots, b) = b$ and $g(t, t, \dots, t, t, b) = t$, we must have that $g(P)$ is a (b, t) -walk in H . Since t and b are nonadjacent, some interior vertex of P must map to a neighbor of t , contrary to our claim above. $\square$

COROLLARY 4.9. *No reflexive cycle of length at least 4 admits a NUF_k for any $k \geq 3$.*

Proof. Let $C : v_1 v_2 \dots v_n v_1$ ($n \geq 4$) be a cycle. We show that C satisfies the assumption of Lemma 4.8: Set $t = v_1$, $b = v_3$, $K_1 = \langle \{v_{\lceil \frac{n-1}{2} \rceil}, v_{\lceil \frac{n}{2} \rceil}\} \rangle$, and $K_2 = \langle \{v_{\lceil \frac{n+3}{2} \rceil}, v_{\lceil \frac{n+4}{2} \rceil}\} \rangle$. Then $N(t) = \{v_2, v_n\}$ and

$$\begin{aligned} d(v_n, K_1) &= \left\lfloor \frac{n}{2} \right\rfloor > \left\lfloor \frac{n}{2} \right\rfloor - 1 = \max\{d(t, K_1), d(b, K_1)\}, \\ d(v_2, K_2) &= \left\lfloor \frac{n}{2} \right\rfloor > \left\lfloor \frac{n}{2} \right\rfloor - 1 = \max\{d(t, K_2), d(b, K_2)\}. \quad \square \end{aligned}$$

We have established that all chordal graphs belong to NUF . Recall that it is proved in [27] that the variety generated by chordal graphs is properly contained in NUF . The corollary below shows that some dismantable graphs do not admit a NUF . It has subsequently been established in [25] that NUF is properly contained in the variety generated by dismantable graphs.

COROLLARY 4.10. *The dismantable reflexive graph in Figure 3 does not admit a NUF_k for any $k \geq 3$.*

COROLLARY 4.11. *There is a reflexive graph H that does not admit a NUF of any arity such that every nonidentity retract of H admits a NUF_3 .*

We complete this section by observing for irreflexive graphs that if H admits a NUF_k , then H must be bipartite. The proof is analogous to Corollary 4.9.

PROPOSITION 4.12. *Let H be an irreflexive graph. If H is not bipartite, then H does not admit a NUF_k for any $k \geq 3$.*

5. Conservative NUF. We complete the paper by examining a special type of near-unanimity functions called conservative functions. Recall that $g \in \text{NUF}_k$ is a *conservative near-unanimity function* if $g(x_1, x_2, \dots, x_k) \in \{x_1, x_2, \dots, x_k\}$ for all $(x_1, \dots, x_k)$. Conservative near-unanimity functions have also been called *choice* or *projective* and are of particular interest since if H admits a conservative NUF, then the *list-homomorphism problem* for H is polynomial time solvable.

The computational complexity of LIST-HOMH has been determined for any fixed (general) graph H . In [11] a new class of intersection graphs called *bi-arc* graphs is introduced. The LIST-HOMH problem is polynomial time solvable if H is a bi-arc graph and is NP-complete if H is not a bi-arc graph. This result generalizes the work of [9, 10] since a reflexive graph is a bi-arc graph if and only if it is an interval graph, and an irreflexive graph is a bi-arc graph if and only if it is a complement of a circular arc graph of clique covering number two.

The polynomial time algorithm of [11] is established by showing that every bi-arc graph admits a conservative NUF_3 . Under the assumption that $P \neq \text{NP}$, the bi-arc graphs are precisely the graphs with a conservative NUF_3 . We now establish this result without the assumption $P \neq \text{NP}$.

Let C be a circle with two specified points p and q . A *bi-arc* is an ordered pair of arcs (N, S) on C such that N contains p but not q , and S contains q but not p . A graph G is a *bi-arc graph* if there is a family of bi-arcs $\{(N_x, S_x) : x \in V(G)\}$ such that, for any not necessarily distinct vertices $x, y \in V(G)$, the following hold:

- (i) if x and y are adjacent, neither N_x intersects S_y nor N_y intersects S_x ;
- (ii) if x and y are not adjacent, both N_x intersects S_y and N_y intersects S_x .

(Note that there cannot be bi-arcs $(N, S), (N', S')$ such that N intersects S' but S does not intersect N' or vice versa.)

We begin by showing that the irreflexive graphs that admit a conservative NUF_3 are precisely the complements of circular arc graphs of clique covering number two.

An *edge-asteroid* in a bipartite graph with the bipartition (X, Y) is a set of $2k+1$ edges $u_0v_0, u_1v_1, \dots, u_{2k}v_{2k}$ ($k \geq 1$ and each $u_i \in X$ and $v_i \in Y$) and $2k+1$ paths, $P_{0,1}, P_{1,2}, \dots, P_{2k,0}$, where each $P_{i,i+1}$ joins u_i to u_{i+1} , such that for each $i = 0, 1, \dots, 2k$ there is no edge between $\{u_i, v_i\}$ and $\{v_{i+k}, v_{i+k+1}\} \cup V(P_{i+k, i+k+1})$ (subscripts are modulo $2k+1$). We refer to the (odd) integer $2k+1$ as the *order* of the edge-asteroid. An edge-asteroid which has no edge between $\{u_0, v_0\}$ and $\{v_1, v_2, \dots, v_{2k}\} \cup V(P_{1,2}) \cup V(P_{2,3}) \cup \dots \cup V(P_{2k-1, 2k})$ is called a *special edge-asteroid*.

We will use the following structural characterization.

THEOREM 5.1 (see [10]). *An irreflexive bipartite graph H is the complement of a circular arc graph if and only if H is chordal bipartite and contains no special edge-asteroids.*

THEOREM 5.2. *Let H be an irreflexive graph. The following are equivalent:*

- (a) H is bipartite and the complement of a circular arc graph;
- (b) H admits a conservative NUF_3 ;
- (c) H admits a conservative NUF_k for some $k \geq 3$.

Proof. In [11], it is shown that every complement of a circular arc graph of clique covering number two has a conservative NUF_3 . Clearly, if H has a conservative NUF_3 , then H has a conservative NUF_k for some k .

To complete the proof, assume to the contrary that H is not the complement of a circular arc graph of clique covering number two, but H does admit a conservative NUF_k , say, g . If H is not bipartite, then by Proposition 4.12 H does not admit a NUF_k for any k . Therefore, H is bipartite.

Suppose H contains an induced (even) cycle, say, C , of length at least six with vertex set $0, 1, 2, \dots, 2n - 1$. In light of Proposition 2.5, g is a conservative NUF_k on the cycle C . Consider the following path in C^k :

$$\begin{aligned} (0, 0, \dots, 0, 0, 2) &\sim (2n - 1, \dots, 2n - 1, 1, 3) \sim (0, \dots, 0, 2, 2) \\ &\sim (2n - 1, \dots, 2n - 1, 1, 3, 3) \sim (0, \dots, 0, 2, 2, 2) \\ &\sim \dots \sim (2n - 1, 1, 3, \dots, 3, 3) \sim (0, 2, 2, \dots, 2, 2). \end{aligned}$$

By near unanimity, the first vertex of the path must map to 0. The vertices of the form $(2n - 1, \dots, 2n - 1, 1, 3, \dots, 3)$ cannot map to 1, since such vertices are distance $n - 2$ from $(n + 1, \dots, n + 1, n - 1, n + 1, \dots, n + 1)$ and 1 is distance n from $n + 1$. Thus the second vertex of the path must map to $2n - 1$, the third to 0, the fourth to $2n - 1$, etc. By parity, the final vertex maps to 0, which is a contradiction.

Hence H is chordal bipartite. We conclude that H contains a special edge-asteroid. Let the path $P_{0,1}$ be $u_0, p_1, p_2, \dots, p_t, u_1$. Consider the following path in H^k :

$$\begin{aligned} (u_0, \dots, u_0, u_0, u_{k+1}) &\sim (v_0, \dots, v_0, p_1, v_{k+1}) \\ &\sim (u_0, \dots, u_0, p_2, u_{k+1}) \sim (v_0, \dots, v_0, p_3, v_{k+1}) \\ &\sim \dots \sim (u_0, \dots, u_0, u_1, u_{k+1}). \end{aligned}$$

Notice that, in this path, all coordinates with the exception of coordinate $k - 1$ alternate between u_i and v_i . We will describe this situation by saying that *coordinate $k - 1$ traverses that path $P_{0,1}$ while all other coordinates alternate*.

We claim $g(u_0, \dots, u_0, u_1, u_{k+1}) = u_0$. Clearly the first vertex of the path is mapped to u_0 . Since g is a conservative function and there is no edge between $\{u_{k+1}, v_{k+1}\}$ and $\{v_0, v_1\} \cup P_{0,1}$, we know that no vertex in the above path is mapped to $\{u_{k+1}, v_{k+1}\}$. If $g(u_0, \dots, u_0, u_1, u_{k+1}) = u_1$, then we look at the path from $(u_0, \dots, u_0, u_1, u_{k+1})$ to $(u_0, \dots, u_0, u_1, u_{k+2})$ where the last coordinate traverses the path $P_{k+1,k+2}$ while all other coordinates alternate. Since there is no edge from $\{u_1, v_1\}$ to $\{v_{k+1}, v_{k+2}\} \cup P_{k+1,k+2}$, the image of the entire path must alternate between u_1 and v_1 , and thus $g(u_0, \dots, u_0, u_1, u_{k+2}) = u_1$. Similarly, we consider the path from $(u_0, \dots, u_0, u_1, u_{k+2})$ to $(u_0, \dots, u_0, u_2, u_{k+2})$ where coordinate $k - 1$ traverses $P_{1,2}$ while all other coordinates alternate. Again there is no edge between $\{u_{k+2}, v_{k+2}\} \cup \{u_0, v_0\}$ and $\{v_1, v_2\} \cup P_{1,2}$. Thus $g(u_0, \dots, u_0, u_2, u_{k+2}) = u_2$. Continuing in this manner, we can move coordinate k down to u_0 and coordinate $k - 1$ up to u_k . We find that $g(u_0, \dots, u_0, u_k, u_0) = u_k$, which is a contradiction. This proves the claim.

Consider the path from $(u_0, \dots, u_0, u_1, u_{k+1})$ to $(u_0, \dots, u_0, u_{k+1}, u_{k+1})$, where the $(k - 1)$ st coordinate traverses all the paths $P_{1,2}, P_{2,3}, \dots, P_{k,k+1}$ while the other coordinates alternate. Since there is no edge from $\{u_0, v_0\}$ to any of these paths $P_{i,i+1}$, the image of the entire path must alternate between u_0 and v_0 , and thus, $g(u_0, \dots, u_0, u_{k+1}, u_{k+1}) = u_0$. Continuing in this manner, we can show that $g(u_0, \dots, u_0, u_{k+1}, \dots, u_{k+1}) = u_0$ and thus $g(u_0, u_{k+1}, \dots, u_{k+1}) = u_0$, which is a contradiction. $\square$

We conclude by characterizing all graphs that admit a conservative near-unanimity function.

THEOREM 5.3. *Let H be a graph. The following are equivalent.*

- (a) H is a bi-arc graph;
- (b) H admits a conservative NUF_3 ;

(c) H admits a conservative NUF.

Proof. Assume H is a bi-arc graph. In [11], it is shown that H admits a conservative NUF_3 . Clearly, if H admits a conservative majority function, then it admits a conservative NUF.

Thus assume H admits a conservative NUF. Let $H^* = H \times K_2$. Note that H^* is irreflexive since K_2 is irreflexive. We have established how to construct a NUF for the product of two graphs given that both factors admit a NUF (see Proposition 2.7). This construction preserves the conservative property. Consequently, H^* is bipartite and the complement of a circular arc graph by Theorem 5.2. Using Proposition 3.1 of [11], we have H is a bi-arc graph. $\square$

Note that by Corollary 2.3 and Theorem 5.3, H admits a conservative NUF_3 if and only if H admits a conservative NUF_k for all $k \geq 3$.

It is shown in [11] that the reflexive bi-arc graphs are precisely the interval graphs and the irreflexive bi-arc graphs are precisely the complement of circular arc graphs of clique covering number two. Thus we have the following.

COROLLARY 5.4. *A reflexive graph admits a conservative NUF if and only if it is an interval graph and an irreflexive graph admits a conservative NUF if and only if it is the complement of a circular arc graph of clique covering number two.*

It is easy to construct a conservative NUF for a reflexive (resp., irreflexive) bi-arc graph using an interval (resp., a circular arc) representation of the graph. In [11], a direct construction of a conservative NUF is given for bi-arc trees. However, a direct construction for general bi-arc graphs still remains elusive.

Acknowledgment. The authors wish to thank the referees for a very careful reading of the manuscript and many helpful suggestions.

REFERENCES

- [1] K. A. BAKER AND A. F. PIXLEY, *Polynomial interpolation and the Chinese remainder theorem for algebraic systems*, Math. Z., 143 (1975), pp. 165–174.
- [2] H. BANDELT, *Graphs with edge-preserving majority functions*, Discrete Math., 103 (1992), pp. 1–5.
- [3] H. J. BANDELT, M. FARBER, AND P. HELL, *Absolute reflexive retracts*, Discrete Appl. Math., 44 (1993), pp. 9–20.
- [4] R. C. BREWSTER AND G. MACGILLIVRAY, *Building blocks for the variety of absolute retracts*, Discrete Math., 306 (2006), pp. 1758–1764.
- [5] A. A. BULATOV, *Tractable conservative constraint satisfaction problems*, in Proceedings of the 18th Annual IEEE Symposium on Logic in Computer Science (Ottawa, Canada, 2003), IEEE Computer Society, Silver Spring, MD, pp. 321–330.
- [6] V. DALMAU AND A. KROKHIN, *Majority constraints have bounded pathwidth duality*, European J. Combin., to appear.
- [7] P. DUKES, H. EMERSON, AND G. MACGILLIVRAY, *Undecidable generalized colouring problems*, J. Combin. Math. Combin. Comput., 26 (1998), pp. 97–112.
- [8] T. FEDER, *Classification of homomorphisms to oriented cycles and of k -partite satisfiability*, SIAM J. Discrete Math., 14 (2001), pp. 471–480.
- [9] T. FEDER AND P. HELL, *List homomorphisms to reflexive graphs*, J. Combin. Theory Ser. B, 72 (1998), pp. 236–250.
- [10] T. FEDER, P. HELL, AND J. HUANG, *List homomorphisms and circular arc graphs*, Combinatorica, 19 (1999), pp. 487–505.
- [11] T. FEDER, P. HELL, AND J. HUANG, *Bi-arc graphs and the complexity of list homomorphisms*, J. Graph Theory, 42 (2003), pp. 61–80.
- [12] T. FEDER, P. HELL, B. LAROSE, C. LOTEN, AND C. TARDIF, *Near Unanimity Functions and Tree Duals*, manuscript, 2008.
- [13] T. FEDER AND M. Y. VARDI, *The computational structure of monotone monadic SNP and constraint satisfaction: A study through datalog and group theory*, SIAM J. Comput., 28 (1998), pp. 57–104.

- [14] M. C. GOLUMBIC, *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, New York, 1980.
- [15] P. HELL, *Retractions de graphes*, Ph.D. thesis, Université de Montreal, Montreal, Canada, 1972.
- [16] P. HELL AND J. NEŠETŘIL, *On the complexity of H -colouring*, J. Combin. Theory Ser. B, 48 (1990), pp. 92–110.
- [17] P. HELL AND J. NEŠETŘIL, *Graphs and Homomorphisms*, Oxford Lecture Series in Mathematics and Its Applications 28, Oxford University Press, Oxford, UK, 2004.
- [18] P. HELL, J. NEŠETŘIL, AND X. ZHU, *Duality and polynomial testing of tree homomorphisms*, Trans. Amer. Math. Soc., 348 (1996), pp. 1281–1297.
- [19] P. HELL AND I. RIVAL, *Absolute retracts and varieties of reflexive graphs*, Canad. J. Math., 39 (1987), pp. 544–567.
- [20] A. P. HUHN, *Schwach distributive Verbände. I*, Acta Sci. Math. (Szeged), 33 (1972), pp. 297–305.
- [21] E. M. JAWHARI, D. MISANE, AND M. POUZET, *Retracts: Graphs and ordered sets from the metric point of view*, in Combinatorics and Ordered Sets, Contemp. Math. 57, AMS, Providence, RI, 1986, pp. 175–226.
- [22] H. W. KIN AND V. DALMAU, *Generalized majority-minority operations are tractable*, in Proceedings of the Twentieth Annual IEEE Symposium on Logic in Computer Science, IEEE Computer Society, Washington, DC, 2005, pp. 438–447.
- [23] G. KUN, B. LAROSE, AND A. MEMARTOLUI, *Private communication*, 2004.
- [24] B. LAROSE, C. LOTEN, AND C. TARDIF, *A characterisation of first-order constraint satisfaction problems*, in Proceedings of the 21st IEEE Symposium on Logic in Computer Science, IEEE Computer Society, Washington, DC, 2006, pp. 201–210.
- [25] B. LAROSE, C. LOTEN, AND L. ZÁDORI, *A polynomial time algorithm for near-unanimity graphs*, J. Algorithms, 55 (2005), pp. 177–191.
- [26] I. LIN, T. MCKEE, AND D. WEST, *The leafage of a chordal graph*, Discuss. Math. Graph Theory, 18 (1998), pp. 23–48.
- [27] C. LOTEN, *Retractions of Chordal and Related Graphs*, Ph.D. thesis, Simon Fraser University, Burnaby, BC, Canada, 2003.
- [28] R. NOWAKOWSKI AND I. RIVAL, *The smallest graph variety containing all paths*, Discrete Math., 43 (1983), pp. 223–234.
- [29] R. NOWAKOWSKI AND P. WINKLER, *Vertex-to-vertex pursuit in a graph*, Discrete Math., 43 (1983), pp. 235–239.
- [30] E. PESCH, *Retracts of Graphs*, Mathematical Systems in Economics 110, Athenäum, Frankfurt am Main, 1988.
- [31] R. W. QUACKENBUSH, I. RIVAL, AND I. ROSENBERG, *Clones, order varieties, near unanimity functions and holes*, Order, 7 (1990), pp. 239–247.

COMBINATORICS OF PEDIGREES I: COUNTEREXAMPLES TO A RECONSTRUCTION QUESTION*

BHALCHANDRA D. THATTE†

Abstract. A pedigree is a finite directed acyclic graph in which all vertices other than the founder vertices have two outgoing arcs each (to their parents), while the founder vertices have no outgoing arcs. The primary motivation to study pedigrees comes from population biology. The main result in this paper is a construction of an infinite family of counterexamples to a reconstruction problem on pedigrees, thus negatively answering a question of Steel and Hein.

Key words. pedigrees, reconstruction

AMS subject classification. 05C60

DOI. 10.1137/060675964

1. Introduction. Informally speaking, a *pedigree* is a structure that represents ancestral relationships between individuals in a population. A motivation to study pedigrees comes from population biology, where one ambitious goal is to reconstruct pedigrees of populations as far into the past as possible. The purpose of this paper is to study a reconstruction question on pedigrees. It is hoped that the main result in this paper—the nonreconstructibility of pedigrees from the pedigrees of subpopulations—will also be of interest to combinatorialists interested in the well-known reconstruction conjectures.

We now introduce the definitions and the terminology and formulate the problem.

A *general pedigree* $\mathcal{T}(X_0)$ on a set X_0 is a finite directed acyclic graph on a vertex set $V \supseteq X_0$ that satisfies the following conditions:

1. each vertex has out-degree 0 or 2; and
2. each vertex in X_0 has in-degree 0.

The vertices with out-degree 0 are called the *founders*. The vertices in X_0 are referred to as the *extant* vertices. The cardinality of X_0 is called the *order* of the pedigree. The vertices with no descendants in X_0 are called the *redundant* vertices.

A *discrete generation pedigree* on X_0 is a pedigree on vertex set $V = \cup_{i=0}^d X_i$, where X_i are disjoint sets, X_d is the set of founders, and every vertex u in $X_i, i < d$, has outgoing arcs uv and uw to vertices v and w , respectively, in X_{i+1} . In this case, d is the *depth* of the pedigree.

If there is an arc from a vertex u to a vertex v , then v is called a parent of u , and u is called a child of v . If there is a directed path from a vertex u to a vertex v in a pedigree, then v is said to be an *ancestor* of u , and u is said to be a *descendent* of v . Trivially, each vertex is its own ancestor as well as its own descendent but not its own parent or child.

A pedigree $\mathcal{P}(X_0)$ with vertex set U is said to be *isomorphic* to a pedigree $\mathcal{Q}(Y_0)$ with vertex set V if there is a one-one map $f : U \rightarrow V$ such that $u_1 - u_2$ is an arc in $\mathcal{P}(X_0)$ if and only if $f(u_1) - f(u_2)$ is an arc in $\mathcal{Q}(Y_0)$. Although this is a

*Received by the editors November 27, 2006; accepted for publication (in revised form) January 18, 2008; published electronically May 21, 2008. This research was supported by the Allan Wilson Centre for Molecular Ecology and Evolution, New Zealand.

<http://www.siam.org/journals/sidma/22-3/67596.html>

†Biomathematics Research Centre, Mathematics and Computer Science Building, University of Canterbury, Private Bag 4800, Christchurch, New Zealand (bdthatte@gmail.com).

standard definition of graph isomorphism, we will be interested in pedigrees in which the extant vertices are labeled. Therefore, if $X_0 = Y_0 = \{x_i, 1 \leq i \leq n\}$, then we will be interested only in isomorphisms π for which $\pi(x_i) = x_i$ for all $1 \leq i \leq n$.

Steel and Hein [3] posed and partially solved reconstruction and enumeration questions about pedigrees. It is well known in phylogenetics that a leaf labeled tree can be constructed up to isomorphism from the collection of pairwise distances between leaves. See, for example, [4]. This fact is a basis of many phylogenetic tree reconstruction algorithms. Motivated by such results in phylogenetics, it is natural to ask the following question: Is a pedigree determined up to isomorphism from the pairwise distances between extant vertices? A pair of extant vertices x and y in a pedigree may have several common ancestors or several paths to a common ancestor. Therefore, it is assumed that all possible distances (in the undirected sense) between all pairs of extant vertices are given. Such a question is not expected to have a positive answer, as demonstrated by a counterexample in [3]. Despite the counterexample, variations of this question are definitely significant in evolutionary biology. Steel and Hein considered the following weaker question.

Let $\mathcal{P}(X_0)$ be a pedigree without redundant vertices. A *subpedigree* $\mathcal{P}(Y)$ of $\mathcal{P}(X_0)$ is obtained by deleting every vertex in $\mathcal{P}(X_0)$ that has no descendent in Y . If the subpedigrees on all two-element subsets of X_0 are given up to isomorphism, can we construct the pedigree on X_0 up to isomorphism? Steel and Hein demonstrated that in general the collection of subpedigrees of order 2 is not sufficient to reconstruct a pedigree. They posed the following problem.

PROBLEM 1. *Is there an integer $r > 2$ such that every pedigree of order $n > r$ that has no redundant vertices is determined up to isomorphism if all its subpedigrees of order r are given up to isomorphism?*

The reason for considering only those pedigrees that have no redundant vertices is that redundant vertices do not appear in subpedigrees, so no structural information about them can be reconstructed.

It must be pointed out that the extant vertices in a pedigree are labeled, and the extant vertices in each subpedigree retain their original label. The following definitions are introduced to make this assumption precise.

Let $n > r > 2$ be positive integers. Let $\mathcal{T}(X_0)$ and $\mathcal{U}(X_0)$ be two pedigrees of order n that have no redundant vertices. The two pedigrees are said to be *r -hypomorphic* to each other if for every $Y \subset X_0, |Y| = r$, there is an isomorphism π_Y from the subpedigree $\mathcal{T}(Y)$ of $\mathcal{T}(X_0)$ to the subpedigree $\mathcal{U}(Y)$ of $\mathcal{U}(X_0)$ such that $\pi_Y(x) = x$ for all $x \in X_0$. A pedigree $\mathcal{T}(X_0)$ is said to be *r -reconstructible* if, for every pedigree $\mathcal{U}(X_0)$ that is r -hypomorphic to $\mathcal{T}(X_0)$, there is an isomorphism π from $\mathcal{T}(X_0)$ to $\mathcal{U}(X_0)$ such that $\pi(x) = x$ for all $x \in X_0$.

PROBLEM 2. *Is there an integer $r > 2$ such that all pedigrees of order $n > r$ are r -reconstructible?*

Since the extant vertices are labeled, many more vertices in a pedigree effectively receive a unique label. Therefore, a reconstruction question on pedigrees is expected to be simpler than the well-known reconstruction conjectures in graph theory.

In section 2, we construct an infinite family of nonreconstructible pedigrees. In particular, we prove that, for every $n \geq 3$, there are pedigrees of order n that are not even $(n - 1)$ -reconstructible. The problem of classification of nonreconstructible pedigrees remains open, and we suspect that it might have an algebraic structure similar to the Nash-Williams lemma in edge reconstruction theory; see [2]. We also present a few positive results on constant population size pedigrees.

2. Counterexamples to pedigree reconstruction. In this section we construct an infinite family of pairs of nonisomorphic pedigrees that have correspondingly isomorphic subpedigrees. That is, we prove the following theorem.

THEOREM 1. *For every $n > 2$, there are nonisomorphic pedigrees $\mathcal{T}(X_0)$ and $\mathcal{U}(X_0)$ of order n that are $(n - 1)$ -hypomorphic.*

Proof. The proof is divided into two cases. The case $n = 3$ gives the basic idea, which is then generalized to arbitrary values of n .

Case $n = 3$. Consider the nonisomorphic graphs $K_{1,3}$ and K_3 . Let the edges of both graphs be arbitrarily labeled e_1, e_2, e_3 , where, following the standard graph theoretic convention, an edge is a set of two vertices. It is clear that $K_{1,3} - e_i \cong K_3 - e_i$ for all i , where $-e_i$ denotes deletion of the edge e_i and the resulting isolated vertices. Now suppose that the end vertices of each edge e_i are parents of the vertex $x_i \in X_0$ in each of the pedigrees $\mathcal{T}(X_0)$ and $\mathcal{U}(X_0)$. Then the pedigrees $\mathcal{T}(\{x_i, x_j\})$ and $\mathcal{U}(\{x_i, x_j\})$ are isomorphic for all i, j , but the pedigrees $\mathcal{T}(X_0)$ and $\mathcal{U}(X_0)$ are not isomorphic. This example proves the theorem for $n = 3$. The pedigrees $\mathcal{T}(X_0)$, $\mathcal{U}(X_0)$ and their subpedigrees are shown in Figure 1.

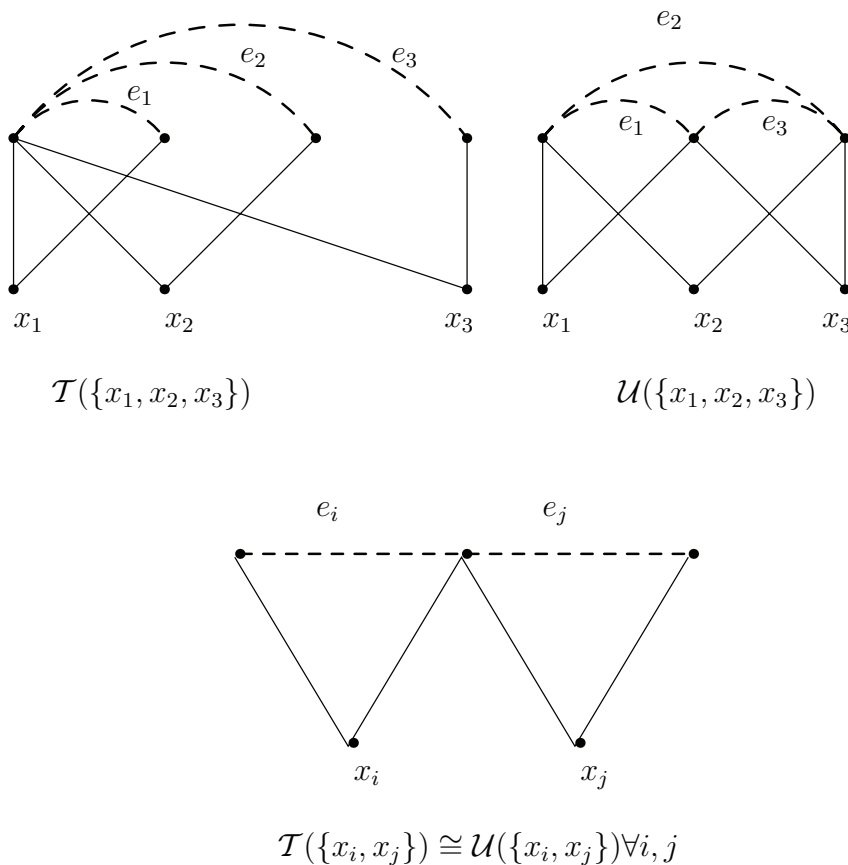


FIG. 1. Pedigrees based on $K_{1,3}$ and K_3 .

Case $n > 3$. We have to construct hypergraphs that play the role that $K_{1,3}$ and K_3 play above. We construct a hypergraph G with edge set $\{g_i, 1 \leq i \leq n\}$ and a hypergraph H with edge set $\{h_i, 1 \leq i \leq n\}$ such that the following conditions are satisfied.

1. $G \not\cong H$.
2. For each $i, 1 \leq i \leq n$, $G - g_i \cong H - h_i$; moreover, there is an isomorphism between $G - g_i$ and $H - h_i$ that *preserves the edge order*; that is, vertices in an edge g_j in $G - g_i$ are mapped to vertices in h_j under such an isomorphism for each $j \neq i$.

Once such hypergraphs are constructed, we treat each edge in each hypergraph as a founder set. We construct pedigrees T_i on founder sets g_i and pedigrees U_i on founder sets h_i such that

1. pedigrees $T_i, 1 \leq i \leq n$, are vertex-disjoint except possibly for their founder sets g_i ;
2. pedigrees $U_i, 1 \leq i \leq n$, are vertex-disjoint except possibly for their founder sets h_i ;
3. pedigrees T_i and U_i are correspondingly isomorphic; moreover, for all $i, j, i \neq j$, an isomorphism between $G - g_j$ and $H - h_j$ that preserves the edge order extends to an isomorphism between T_i and U_i ; and
4. each of the pedigrees T_i and U_i contains exactly one extant vertex x_i .

The resulting pedigrees $\cup_{i=1}^n T_i$ and $\cup_{i=1}^n U_i$ are nonisomorphic (since the hypergraphs G and H are nonisomorphic), but their subpedigrees are correspondingly isomorphic.

Construction of hypergraphs G and H . The required hypergraphs are constructed by a simple application of linear algebra.

Let each integer in $\{0, 2^n - 1\}$ be written in base 2 as an n -digit number by padding sufficiently many zeros on the left. We count its digits from the right. The set of n -digit binary numbers is denoted by $[2^n]$. The i th digit of a number k is denoted by $k(i)$, and the number obtained by setting the i th digit of k to 0 (or 1) is denoted by $k(i \leftarrow 0)$ (or, respectively, $k(i \leftarrow 1)$). The number of 1's and the number of 0's in k are denoted by $\#1(k)$ and $\#0(k)$, respectively.

The isomorphism class of a hypergraph G with edge set $\{g_i, 1 \leq i \leq n\}$ may be represented by a list of integers $a(k), k \in [2^n]$, where $a(k)$ is the number of vertices in $\cap_{i=1}^n f_i$, where $f_i = g_i$ if $k(i) = 1$, and $f_i = \bar{g}_i$ (that is, the complement of g_i) if $k(i) = 0$. In other words, we have to specify only the number of vertices in each region of the Venn diagram of $g_i, 1 \leq i \leq n$. Let the list of integers $b(k), k \in [2^n]$, similarly denote the isomorphism class of H .

The condition $G - g_i \cong H - h_i$ for $1 \leq i \leq n$, with an isomorphism between them that preserves the edge order, may be expressed as

$$(1) \quad a(k(i \leftarrow 0)) + a(k(i \leftarrow 1)) = b(k(i \leftarrow 0)) + b(k(i \leftarrow 1)), \quad k \in [2^n].$$

Since we are interested in nonisomorphic hypergraphs G and H , we must find solutions to the above equations so that $a(k) \neq b(k)$ for some $k \in [2^n]$.

We verify that

$$\begin{aligned} a(k) &= 1 \text{ and } b(k) = 0 \text{ when } k \text{ has even number of 1's,} \\ a(k) &= 0 \text{ and } b(k) = 1 \text{ when } k \text{ has odd number of 1's} \end{aligned}$$

satisfy (1).

It can be easily verified that $K_{1,3}$ and $K_3 \cup K_1$ do in fact satisfy the above solutions, where we include an isolated vertex in one of the graphs purely for algebraic convenience.

From now on we write $[2^n] = [2^n]_e \cup [2^n]_o$, where $[2^n]_e$ is the set of integers having an even number of 1's in their binary representation, and $[2^n]_o$ is the set of integers having an odd number of 1's in their binary representation. In this notation, the hypergraphs G and H are described as follows: the set $[2^n]_e$ is the set of vertices of G , and a vertex $k \in [2^n]_e$ is in g_i if and only if $k(i) = 1$. Similarly, the set $[2^n]_o$ is the vertex set of H , and a vertex $k \in [2^n]_o$ is in edge h_i if and only if $k(i) = 1$.

The vertex $k = 0$ is in G but is an isolated vertex and is included at this stage only for algebraic convenience. It may be deleted after completing the construction of nonreconstructible pedigrees.

It is clear that G and H are nonisomorphic, since each of them has 2^{n-1} vertices, but G has the isolated vertex 0, while H has no isolated vertex. What is an isomorphism between $G - g_i$ and $H - h_i$? An edge order preserving isomorphism from $G - g_i$ to $H - h_i$ must map vertices in a region of the Venn diagram of $\cup_{j|j \neq i} g_j$ to the corresponding region of the Venn diagram of $\cup_{j|j \neq i} h_j$. Consider any $k \in [2^n]$. The vertex $k(i \leftarrow 0)$ is in g_j for some $j \neq i$ if and only if the vertex $k(i \leftarrow 1)$ is in h_j , because the two vertices differ only in their i th digit. Therefore, if $k(i \leftarrow 0)$ is in G , then an edge order preserving isomorphism between $G - g_i$ and $H - h_i$ must map the vertex $k(i \leftarrow 0)$ to the vertex $k(i \leftarrow 1)$. Similarly, if the vertex $k(i \leftarrow 1)$ is in G , then an edge order preserving isomorphism between $G - g_i$ and $H - h_i$ must map the vertex $k(i \leftarrow 1)$ to the vertex $k(i \leftarrow 0)$. Moreover, this isomorphism is unique. On the standard hypercube on $[2^n]$, each vertex in $G - g_i$ is mapped to its neighbor along the i th axis, which is in $H - h_i$.

Example 1. Let $n = 4$, and let the hypergraphs G and H be defined on vertex sets $[2^4]_e$ and $[2^4]_o$ as follows:

$$\begin{aligned} g_1 &= \{0011, 0101, 1001, 1111\}, \\ g_2 &= \{0011, 0110, 1010, 1111\}, \\ g_3 &= \{0101, 0110, 1100, 1111\}, \\ g_4 &= \{1001, 1010, 1100, 1111\}, \\ h_1 &= \{0001, 0111, 1011, 1101\}, \\ h_2 &= \{0010, 0111, 1011, 1110\}, \\ h_3 &= \{0100, 0111, 1101, 1110\}, \\ h_4 &= \{1000, 1011, 1101, 1110\}, \end{aligned}$$

where g_i are the edges of G , and h_i are the edges of H . The isomorphism π_1 from $G - g_1$ to $H - h_1$ that preserves the edge order is given by $\pi_1(0000) = 0001$, $\pi_1(0011) = 0010$, $\pi_1(0101) = 0100$, $\pi_1(1001) = 1000$, $\pi_1(0110) = 0111$, $\pi_1(1010) = 1011$, $\pi_1(1100) = 1101$, $\pi_1(1111) = 1110$. Observe that $\pi_1(g_2) = h_2$, $\pi_1(g_3) = h_3$, and $\pi_1(g_4) = h_4$ under this map.

Construction of T_i and U_i . As stated earlier, for each i , pedigrees T_i and U_i must be so constructed that the edge order preserving isomorphism between $G - g_j$ and $H - h_j$ extends to an isomorphism between T_i and U_i for all $j \neq i$.

Let a balanced binary tree T_i be defined so that x_i is its root and g_i is its set of leaves. By convention, the root x_i is the lowest vertex (at depth 0) in T_i , and the leaves are the highest vertices (at depth $n-2$) in T_i . For a vertex t in T_i , let $T_i(t)$ be the subtree of T_i induced by t and all vertices in T_i that are above t . Let t_0 and t_1 be the parents of t . The subtree $T_i(t)$ is a union of subtrees $L(t)$ and $R(t)$, where $L(t)$ is induced by vertices t, t_0 , and all vertices above t_0 , and the subtree $R(t)$ is induced by vertices t, t_1 , and all vertices above t_1 . We call $L(t)$ the *left subtree* at t and $R(t)$ the *right subtree* at t .

Let $i_1, i_2, \dots, i_{n-1}$ be the integers $1 \leq j \leq n, j \neq i$, in arbitrary order. The vertices in g_i are grouped in such a way that for each vertex t at depth $k, 0 \leq k \leq n-3$, if a vertex $p \in g_i$ is a leaf of $L(t)$, then $p(i_{k+1}) = 0$, and if a vertex $p \in g_i$ is a leaf of $R(t)$, then $p(i_{k+1}) = 1$.

The vertices in h_i are partitioned, and a binary tree U_i is constructed analogously for the same ordering $i_j, 1 \leq j \leq n-1$.

For $n = 5$ and $i = 5$ and the ordering $i_1 = 2, i_2 = 3, i_3 = 1, i_4 = 4$, the trees T_5 and U_5 are shown in Figure 2.

We show that for every $j \neq i$, the unique isomorphism between $G - g_j$ and $H - h_j$ extends to an isomorphism between T_i and U_i .

Let $\bar{b} = (b_1, \dots, b_j), 0 \leq j \leq n-2$, be a j -tuple of 0's and 1's. Extending notation introduced earlier, let $g_i(\bar{b})$ denote the set $\{k \in g_i | k(i_1) = b_1, k(i_2) = b_2, \dots, k(i_j) = b_j\}$, which is the set of leaves of a binary subtree of T_i rooted at the vertex $t(\bar{b})$ at depth j . For example, when $\bar{b} = (0)$, $g_i(\bar{b})$ is the set of leaves above the left parent of x_i , and $t(\bar{b})$ is the left parent of x_i . The set $h_i(\bar{b})$ and the vertex $u(\bar{b})$ are analogously defined for U_i . By convention, an empty tuple $\bar{b}$ defines the sets g_i and h_i , and the trees T_i and U_i , rooted at x_i ; and a tuple $\bar{b}$ of length $n-2$ defines singleton subsets $\{t(\bar{b})\}$ of g_i and $\{u(\bar{b})\}$ of h_i . A tuple $\bar{b}$ of length $n-2$ also uniquely determines the digits $t(\bar{b})(i_{n-1})$ and $u(\bar{b})(i_{n-1})$, since we know that $\#1(t(\bar{b}))$ is even and $\#1(u(\bar{b}))$ is odd. Also, if $t(\bar{b})(i_{n-1}) = 1$, then $u(\bar{b})(i_{n-1}) = 0$, and if $t(\bar{b})(i_{n-1}) = 0$, then $u(\bar{b})(i_{n-1}) = 1$. Therefore, the map $t(\bar{b}) \rightarrow u(\bar{b})$ for all tuples of length at most $n-2$ extends the isomorphism between $G - g_{i_{n-1}}$ and $H - h_{i_{n-1}}$.

Let $\bar{b}$ be a tuple as above. Extending the notation $k(i \leftarrow 0)$ to $\bar{b}$, we define $\bar{b}(i \leftarrow 0)$ to be the tuple obtained by setting $b_i = 0$ in $\bar{b}$, and $\bar{b}(i \leftarrow 1)$ to be the tuple obtained by setting $b_i = 1$ in $\bar{b}$.

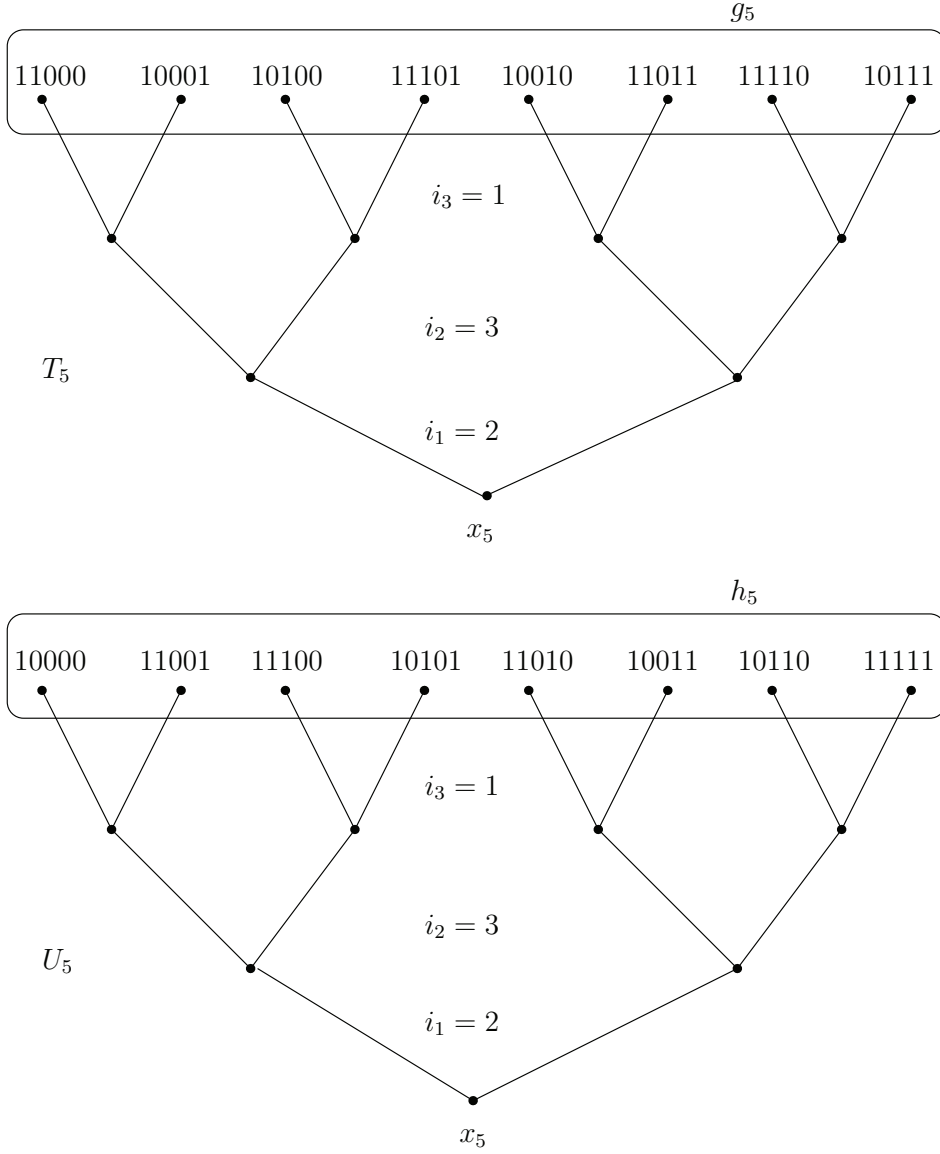
Let $\bar{b}$ be a tuple of length $n-2$ and $j \leq n-2$. By an argument as in the above paragraph, we have

1. if $t(\bar{b}(j \leftarrow 0))(i_{n-1}) = 1$, then $u(\bar{b}(j \leftarrow 1))(i_{n-1}) = 1$;
2. if $t(\bar{b}(j \leftarrow 0))(i_{n-1}) = 0$, then $u(\bar{b}(j \leftarrow 1))(i_{n-1}) = 0$;
3. if $t(\bar{b}(j \leftarrow 1))(i_{n-1}) = 1$, then $u(\bar{b}(j \leftarrow 0))(i_{n-1}) = 1$; and
4. if $t(\bar{b}(j \leftarrow 1))(i_{n-1}) = 0$, then $u(\bar{b}(j \leftarrow 0))(i_{n-1}) = 0$.

Therefore, for each $j, 1 \leq j \leq n-2$, the map defined by

1. $t(\bar{b}) \rightarrow u(\bar{b})$ for all tuples of length at most $j-1$,
2. $t(\bar{b}(j \leftarrow 0)) \rightarrow u(\bar{b}(j \leftarrow 1))$ for all tuples of length at least j , and
3. $t(\bar{b}(j \leftarrow 1)) \rightarrow u(\bar{b}(j \leftarrow 0))$ for all tuples of length at least j

extends the isomorphism between $G - g_j$ and $H - h_j$. Observe that this map sends the vertices in the left subtree $L(t(\bar{b}))$ in T_i to the vertices in the right subtree $R(u(\bar{b}))$ in U_i , and the vertices in the right subtree $R(t(\bar{b}))$ in T_i to the vertices in the left subtree $L(u(\bar{b}))$ in U_i , for each tuple $\bar{b}$ of length j . Since the trees T_i and T_j (and trees U_i and U_j) are disjoint except for their founders for all $i \neq j$, the isomorphism between

FIG. 2. Binary pedigrees T_5 and U_5 for $n = 5$.

$G - g_j$ and $H - h_j$ extends to an isomorphism between pedigrees $\mathcal{T}(X_0 \setminus \{x_j\})$ and $\mathcal{U}(X_0 \setminus \{x_j\})$ for all j .

An isomorphism between $G - g_{i_2}$ and $H - h_{i_2}$ that extends to an isomorphism between T_i and U_i , $i \neq i_2$, is schematically shown in Figure 3. $\square$

Remark 1. The pedigrees constructed above do not admit a valid gender labeling. That is, we cannot assign labels m (male) and f (female) to all vertices so that each vertex (except founders) has one male parent and one female parent. For example, in the $n = 3$ case, K_3 is not a bipartite graph, so a valid gender labeling is impossible.

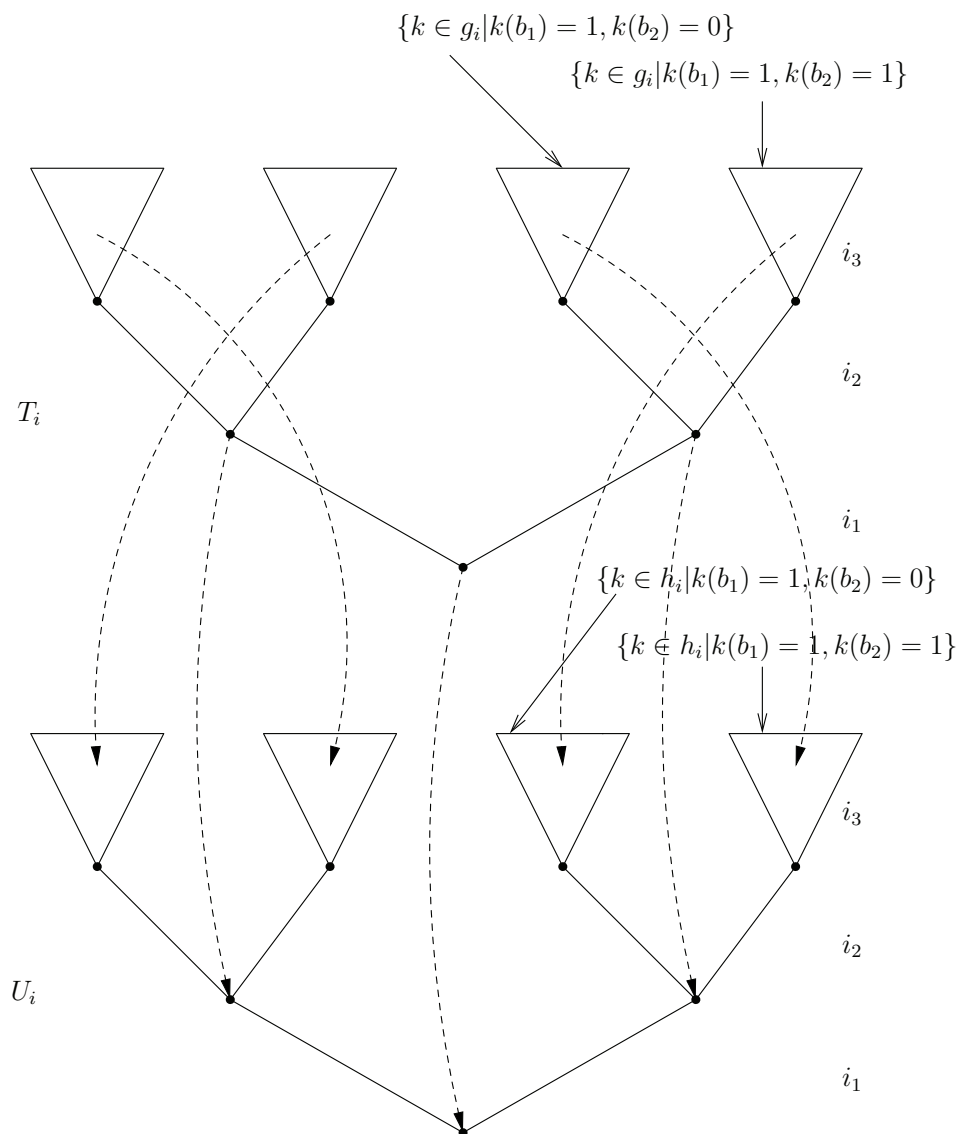


FIG. 3. Isomorphism between $G - g_{i_2}$ and $H - h_{i_2}$ extends to an isomorphism between T_i and U_i .

But the examples can be easily modified to create nonreconstructible pedigrees that also admit valid gender labels. Each vertex in a pedigree constructed above may be duplicated, and one vertex may be treated male and the other female, as shown in Figure 4. At the bottom of the tree T_i (or U_i), the vertex x_i is duplicated as vertices x_i^m and x_i^f , and the new vertex x_i is a child of x_i^m and x_i^f .

Remark 2. Let $k, k' \in [2^n]$ be any two adjacent vertices on the hypercube. From (1), if $a(k) - b(k) = p > 0$ for some p , then $b(k') - a(k') = p$, regardless of which digit k and k' differ at. In fact, by connectivity of the hypercube, we have $a(r) - b(r) = p$ for all vertices $r \in [2^n]$ that are at even distance from k on the hypercube and $b(r) - a(r) = p$

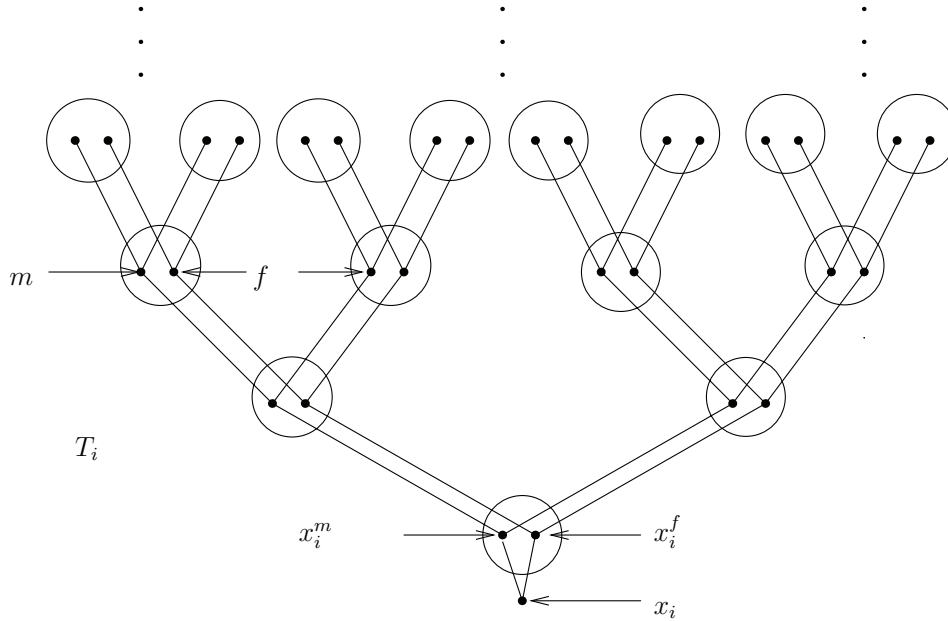


FIG. 4. Construction of a pedigree with a valid gender labeling.

for all vertices $r \in [2^n]$ that are at odd distance from k on the hypercube. This further implies that the hypergraphs G and H constructed in the above counterexample have a special structure: for each $i, 1 \leq i \leq n$, $|g_i| = |h_i| \geq 2^{n-2}$. Let $G(d)$ be the hypergraph with edge set $\{g_i(d), 1 \leq i \leq n\}$, where $g_i(d)$ is the set of ancestors of x_i at depth d in the pedigree $\mathcal{T}(X_0)$, and let $H(d)$ and $h_i(d)$ be similarly defined for the pedigree $\mathcal{U}(X_0)$; then the hypergraphs $G(d)$ and $H(d)$ must be isomorphic whenever $d < n-2$.

We end this section with a conjecture motivated by the observations made in Remark 2.

CONJECTURE 1. *The counterexample constructed in Theorem 1 is minimal. In other words, if a pedigree $\mathcal{T}(X_0)$ of order n is not $(n-1)$ -reconstructible, then it has depth at least $n-2$, and there are at least 2^{n-1} ancestors at depth $n-2$.*

Remark 3. Let G and H be simple graphs with edge sets $E(G) = \{g_i, 1 \leq i \leq m\}$ and $E(H) = \{h_i, 1 \leq i \leq m\}$, respectively, such that $G - g_i \cong H - h_i$ for all $1 \leq i \leq m$. Then the edge reconstruction conjecture states that $G \cong H$ provided $m > 3$. The condition $m > 3$ is required since $K_{1,3}$ and K_3 —the graphs used as the base case of our construction of nonreconstructible pedigrees—are not *edge reconstructible*. Although no counterexamples are yet known, Nash-Williams [2] proved a characterization of (hypothetical) counterexamples to edge reconstruction. His characterization was based on a generalization of ideas earlier introduced by Lovász [1]. Without going into detail, we note that the counterexamples presented here have certain similarities to the characterization by Nash-Williams. It may be possible to exploit such similarities to prove Conjecture 1.

3. A positive reconstruction result. Let $\mathcal{T}(X_0)$ be a discrete generation pedigree on X_0 , where $|X_0| = n > 2$. Let $S_{n-1}(\mathcal{T}) = \{\mathcal{T}(Y) | Y \subset X_0, |Y| = n-1\}$. Consider the edge labeled (multi-) graph G_1 whose vertex set is X_1 (that is, the set

of vertices at depth 1), and vertices $x, y \in X_1$ are joined by an edge e_i if they are the parents of x_i .

LEMMA 1. *If there are vertices x_i and x_j in X_0 that have the same parents, then $\mathcal{T}(X_0)$ is uniquely determined by $S_{n-1}(\mathcal{T})$.*

Proof. The situation in the lemma is recognized by looking at $\mathcal{T}(X_0 \setminus x_k)$, where $x_k \notin \{x_i, x_j\}$. Now $\mathcal{T}(X_0)$ is uniquely obtained from $\mathcal{T}(X_0 \setminus x_i)$ by joining x_i to the parents of x_j . $\square$

LEMMA 2. *If $n > 3$ and if G_1 contains a cycle, then $\mathcal{T}(X_0)$ is uniquely determined by $S_{n-1}(\mathcal{T})$.*

Proof. Let e_i be an edge in a cycle in G_1 . The end vertices of e_i are the two parents of x_i . Since the set of half brothers of x_i is known from the collection S_{n-1} , the parents of x_i are uniquely recognized in $\mathcal{T}(X \setminus x_i)$. Note that we need the condition $n > 3$ because otherwise we would get a counterexample based on $G_1 \cong K_3$ or $G_1 \cong K_{1,3}$. $\square$

COROLLARY 1. *If $|X_1| \leq n$ and $n > 3$, then $S_{n-1}(\mathcal{T})$ determines $\mathcal{T}(X_0)$ up to congruence.*

Proof. If no two vertices in X_0 have the same two parents, then G_1 has n simple edges (that is, no two edges are parallel edges), and there is a cycle in G_1 . $\square$

We end this section with another conjecture.

CONJECTURE 2. *Discrete generation pedigrees of order n that have the same number of vertices in each generation are r -reconstructible for $r > \log n$.*

This conjecture is true if Conjecture 1 is true. For suppose that Conjecture 1 is true but Conjecture 2 is not true, and that there is a constant population size pedigree of order n that is not r -reconstructible for some $r > \log n$. Therefore, for some $r > \log n$, there is a subpedigree of order $r + 1$ that is not r -reconstructible. Such a subpedigree must have depth at least $r - 1$ and must have at least 2^r vertices at depth $r - 1$, implying that $r \leq \log n$. Thus, if $r > \log n$, then we have a contradiction; therefore, all subpedigrees of order $r + 1$ are r -reconstructible when $r > \log n$, and we can complete the reconstruction inductively.

Acknowledgments. I would like to thank Mike Steel for many useful conversations and for hosting me at the University of Canterbury. Oliver Will and Mike Steel read the manuscript carefully and made many helpful comments.

REFERENCES

- [1] L. LOVÁSZ, *A note on the line reconstruction problem*, J. Combin. Theory Ser. B, 13 (1972), pp. 309–310.
- [2] C. ST. J. A. NASH-WILLIAMS, *The reconstruction problem*, in Selected Topics in Graph Theory, L. W. Beineke and R. J. Wilson, eds., Academic Press, London, 1978, pp. 205–236.
- [3] M. STEEL AND J. HEIN, *Reconstructing pedigrees: A combinatorial perspective*, J. Theoret. Biol., 240 (2006), pp. 360–367.
- [4] K. A. ZARECKIĬ, *Constructing a tree on the basis of a set of distances between the hanging vertices*, Uspekhi Mat. Nauk, 20 (1965), pp. 90–92.

RESONANCE GRAPHS AND A BINARY CODING FOR THE 1-FACTORS OF BENZENOID SYSTEMS*

HEPING ZHANG[†], PETER CHE BOR LAM[‡], AND WAI CHEE SHIU[‡]

Abstract. Applying the recently obtained distributive lattice structure on the set of 1-factors, we show that the resonance graphs of any benzenoid systems G , as well as of general plane (weakly) elementary bipartite graphs, are median graphs and thus extend greatly Klavžar et al.'s result. The n -dimensional vectors of nonnegative integers as a labelling for the 1-factors of G with n inner faces are described. The labelling preserves the partial ordering of the above-mentioned lattice and can be transformed into a binary coding for the 1-factors. A simple criterion for such a labelling being binary is given. In particular, Klavžar et al.'s algorithm is modified to generate this binary coding for the 1-factors of a cata-condensed benzenoid system.

Key words. 1-factor, benzenoid system, distributive lattice, resonance graph, Z-transformation graph, binary coding, median graph

AMS subject classifications. 05C70, 05C12, 06D05, 92E10

DOI. 10.1137/070699287

1. Introduction. A hexagonal system is a finite connected plane graph with no cut-vertices in which every interior region is bounded by a regular hexagon of side length 1. A hexagonal system with a 1-factor is regarded as a molecular graph (carbon-skeleton) of a benzenoid hydrocarbon. Hence these kinds of graphs are called *benzenoid systems* and have been extensively investigated; cf. a detailed review [23] due to Randić with the references therein. A benzenoid system is classified into two types: peri-condensed if it has an interior vertex and cata-condensed if it has none. For example, a well-known coronene is peri-condensed, while Dibenzo[a,c]anthracene is cata-condensed; both are illustrated in Figure 1.

A 1-factor or perfect matching M of a graph G is a set of pairwise nonadjacent edges (i.e., no two edges share a common endvertex) such that every vertex of G is incident with an edge in M . This graph-theoretical concept coincides with a Kekulé structure of chemical molecules. One research approach is to compute the number of Kekulé structures of a benzenoid system [5]. Another approach is to study individual Kekulé structures and relationships between them. For example, Gutman [10] and Gutman, Teodorović, and Kolaković [11] found different contributions of a leaf and a nonleaf in the tree structure on the set of 1-factors to the resonance energy of a benzenoid hydrocarbon and obtained a more accurate estimate for the resonance energy.

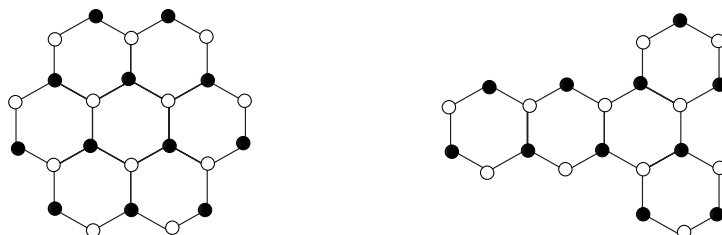
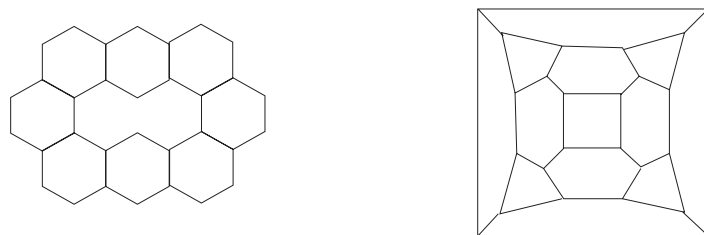
A graph was established on the set of 1-factors of a benzenoid system G : Every 1-factor of G is for a vertex, and a pair of 1-factors is for an edge if they can be transformed into each other by interchanging double and single bonds in an alternating hexagon. This new graph associated with a benzenoid system is called the

*Received by the editors August 5, 2007; accepted for publication (in revised form) February 11, 2008; published electronically May 30, 2008. This work was supported by FRG, Hong Kong Baptist University, and NSFC grant 10471058.

<http://www.siam.org/journals/sidma/22-3/69928.html>

[†]School of Mathematics and Statistics, Lanzhou University, Lanzhou, Gansu 730000, People's Republic of China (zhanghp@lzu.edu.cn).

[‡]Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, People's Republic of China (cblam@hkbu.edu.hk, weshiu@hkbu.edu.hk).

FIG. 1. *Coronene (left) and Dibenzo[a,c]anthracene (right).*FIG. 2. *A coronoid system (left) and the $B_{12}N_{12}$ fullerene (right).*

Z-transformation graph in mathematics by Zhang, Guo, and Chen [28] and the *resonance graph* in chemistry by Randić [22]. Chemical explanation for this concept can be found in [16, 22, 23]. In 1997, Randić [22] showed that the leading eigenvalues (λ) of the resonance graphs correlate with the resonance energy of benzenoids and gave a quite satisfactory regression formula.

For cata-condensed benzenoid systems with n hexagons, Chen and Zhang [2] discovered that the resonance graphs contain Hamilton paths, which can be generated by an algorithm [15]; Randić [22, 23] and Randić et al. [24] showed that some of the resonance graphs can be constructed by proposing a procedure and embedded in n -dimensional Cartesian coordinate grids; Klavžar, Žigert, and Brinkmann [17] proved that the resonance graphs are median graphs. As applications of the latter result, the resonance graphs of cata-condensed benzenoid systems can be isometrically embedded in hypercubes; an algorithm was designed [16] to give each 1-factor a sequence of ones and zeros with length n as a binary coding; a simple method was used to determine the Clar number [18].

Benzenoid systems are a kind of weakly-elementary plane bipartite graphs. In addition, the molecular graphs of most alternate compounds are also plane bipartite graphs. For example, a *coronoid* [4] is a connected subgraph of a hexagonal system such that every edge belongs to at least one hexagon and it contains at least one nonhexagonal interior face (called a corona hole); a boron-nitride (BN) fullerene [3, 8] is a cubic graph embedded on the surface of a sphere so that all faces are (exactly 6) squares and hexagons; see Figure 2. As a unified framework, the resonance graph has already been extended [30, 31, 32] in a natural way on the set of 1-factors of a plane bipartite graph. It was independently shown [7, 33] that the resonance graph of a plane bipartite graph G is connected if and only if G is weakly elementary.

The acyclic orientation of the resonance graph not only enables one to prove [31] that the block graph of the resonance graph of a plane (weakly) elementary bipartite graph G is a path, but also implies a partial ordering on the set of 1-factors of G . Lam and Zhang [20], two of the present authors, showed that this poset associated

with G is a finite distributive lattice by embedding it into an n -dimensional Cartesian coordinate grid with the usual partial ordering, where n is the number of inner faces of G , which implies that each 1-factor is assigned a sequence of nonnegative integers with length n as its labelling.

As applications of the above distributive lattice on the set of 1-factors, in this paper we mainly prove that all connected resonance graphs are median graphs and thus give a significant extension to Klavžar et al.'s result (see Klavžar, Vesel, and Žigert [15], Klavžar et al. [16], Klavžar, Žigert, and Brinkmann [17], Klavžar, Žigert, and Gutman [18]), and we further investigate the labelling for the 1-factors that preserves partial ordering. In the next section, we will introduce the distributive lattice with labelling and give a procedure for generating them. In section 3, by the established distributive lattice structure and the developed techniques, we show that the resonance graph of a plane weakly-elementary bipartite graph is a median graph. In section 4, we give a criterion that such a labelling is a binary coding for all 1-factors of a benzenoid system: it has no coronene as its nice subgraph. In general, the labelling can be transformed into a binary coding so that its length is the diameter of the resonance graph. In section 5, a simple algorithm is designed for generating the binary coding for the 1-factors of a cata-condensed benzenoid system, without producing 1-factors. We conclude this paper with two open problems.

2. Preliminaries and distributive lattices. We now introduce some concepts and notations. A *bipartite* graph means a graph for which every vertex is colored by white or black so that two adjacent vertices receive different colors; such a 2-coloring (white-black) is always given. A bipartite graph G is called *elementary* if G is connected and every edge belongs to a 1-factor of G . For example, coronene, cata-condensed benzenoid systems and boron-nitride fullerenes are plane elementary bipartite graphs.

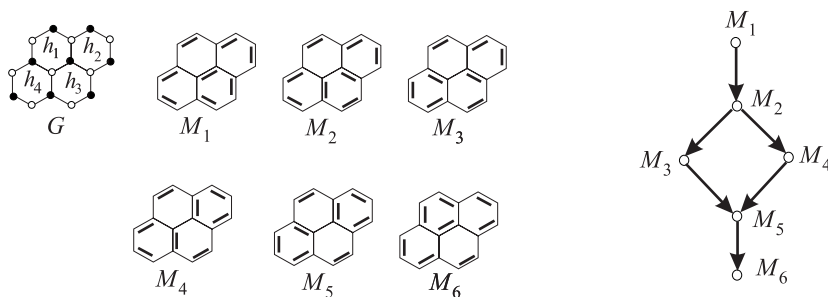
Let G be a plane bipartite graph with a 1-factor M . A cycle C is *M -alternating* if the edges of C appear alternately in and off the M . An M -alternating cycle C of G is said to be *proper* (*improper*) [30] if every edge of C belonging to M goes from white (black) endvertex to black (white) endvertex by the clockwise orientation of C .

THEOREM 2.1 (see [32]). *A nontrivial plane bipartite graph is elementary if and only if every face is resonant, i.e., every facial boundary is an alternating cycle with some 1-factor.*

Elementary components of G mean components of the subgraph obtained from G by removing all fixed single edges (those edges not contained in any 1-factors). G is called *weakly elementary* [19, 25, 32] if every inner face of every elementary component of G is still a face of the original G . For instance, benzenoid systems are these kinds of graphs.

The symmetry difference of finite sets A and B is defined as $A \oplus B := (A \cup B) \setminus (A \cap B)$. This operation can be defined among many finite sets in a natural way and is associative and commutative. If C is an M -alternating cycle of G , then $M \oplus C$ is also a 1-factor, where C may be regarded as its edge-set. It is obvious that a cycle C is proper M -alternating if and only if it is improper $M \oplus C$ -alternating. For two 1-factors M and M' of G , $M \oplus M'$ consists of pairwise disjoint M and M' -alternating cycles of G .

A partially ordered set (poset) is a set P with a binary relation $\preceq$ satisfying the reflexivity, antisymmetry, and transitivity. If $x, y \in P$, then we say y covers x if $x \prec y$ and no element z satisfies $x \prec z \prec y$. A poset $(P, \preceq)$ is determined by its covering relation and often denoted by the Hasse diagram. A one-to-one mapping ϕ from P_1 to

FIG. 3. Pyrene G , 1-factors M_1 to M_6 , and Z-transformation digraph $\vec{Z}(G)$.

P_2 is an embedding of lattice $(P_1; \vee, \wedge)$ into lattice $(P_2; \vee, \wedge)$ if $\phi(x \vee y) = \phi(x) \vee \phi(y)$ and $\phi(x \wedge y) = \phi(x) \wedge \phi(y)$ for any x and y in P_1 . Other terms on posets and lattices used in this paper can be found in [1, 9, 26].

Let G be a plane bipartite graph with a 1-factor. Let $\mathcal{M}(G)$ denote the set of all 1-factors of G and $\mathcal{F}$ the set of all inner faces of G . The *Z-transformation digraph* or *resonance digraph*, denoted by $\vec{Z}(G)$, of 1-factors of G is defined as a digraph on $\mathcal{M}(G)$ such that there exists an arc from M_1 to M_2 provided that $M_1 \oplus M_2$ is a proper M_1 - (thus improper M_2 -) alternating cycle that is an inner facial boundary of G . Ignoring all directions of arcs of $\vec{Z}(G)$, we get the usual Z-transformation graph or resonance graph, denoted by $Z(G)$.

For example, the Z-transformation digraph of pyrene is constructed in Figure 3. Note that a benzenoid system considered is always drawn so that an edge-direction is vertical and the peaks are colored black (see Figure 1). Accordingly, a proper and an improper alternating hexagon with respect to a 1-factor are the usual proper and improper *sextet*, respectively.

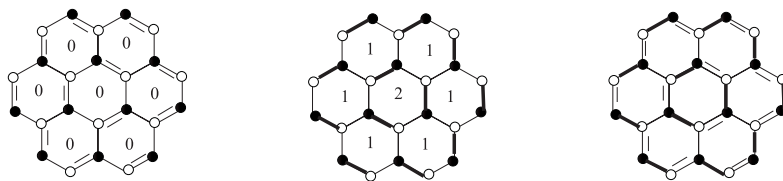
The property that $\vec{Z}(G)$ has no directed cycle implies a partial ordering: $M_1 \preceq M_2$, $M_1, M_2 \in \mathcal{M}(G)$, if and only if $\vec{Z}(G)$ has a directed path from M_2 to M_1 . Then $(\mathcal{M}(G), \preceq)$ is a poset. From now on we simply denote it by $\mathcal{M}(G)$.

THEOREM 2.2 (see [20, 29]). *For a plane bipartite graph G , $\vec{Z}(G)$ is the Hasse diagram of $\mathcal{M}(G)$.*

The plane bipartite graph G has a unique 1-factor M_0 such that G has no proper M_0 -alternating cycles [30]. Then M_0 is a minimal element of the poset $\mathcal{M}(G)$ and called the *root 1-factor* of G . Likely, G has a unique 1-factor M_1 such that G has no improper M_1 -alternating cycles. For each $M \in \mathcal{M}(G)$, we define a *function* ϕ_M on $\mathcal{F}$ as follows: For any $f \in \mathcal{F}$, $\phi_M(f)$ is the number of cycles in $M \oplus M_0$ with f in their interiors. In fact, such cycles are all proper M -alternating and pairwise disjoint if they exist. In particular, ϕ_{M_0} is constantly zero, i.e., every value on the inner faces is 0.

For example, we consider 1-factors M_0 (left) and M_1 (middle) of the coronene in Figure 4. Since $M_1 \oplus M_0$ consists of two proper M_1 -alternating cycles—the boundary and the central hexagon of the coronene (see the rightmost diagram in Figure 4), ϕ_{M_1} is easily computed and represented by the integers indicated in hexagons. As for the pyrene (see Figure 3), the functions ϕ_M of 1-factors $M_1, \dots, M_6$ are denoted, respectively, by 1111, 1101, 1100, 1001, 1000, 0000 in such an order of hexagons: h_1, h_2, h_3, h_4 .

$\mathbb{Z}^+ = \{0, 1, 2, \dots\}$ is a linear order with the usual order. $(\mathbb{Z}^+)^{\mathcal{F}}$ is the Cartesian product of $|\mathcal{F}|$ copies of $\mathbb{Z}^+$; each one is indexed by an inner face in $\mathcal{F}$. That is, $(\mathbb{Z}^+)^{\mathcal{F}} = \{(a_1, a_2, \dots, a_{|\mathcal{F}|}) : \text{each } a_i \in \mathbb{Z}^+ \text{ is indexed by the } i\text{th inner face in } \mathcal{F}\}$. It


 FIG. 4. Examples of function ϕ_M for the coronene.

is well known that $(\mathbb{Z}^+)^{\mathcal{F}}$ is an infinite distributive lattice under the usual partial ordering (i.e., for $a, b \in (\mathbb{Z}^+)^{\mathcal{F}}$, $a \leq b$ if and only if $(a)_i \leq (b)_i$ for each $i \in \mathcal{F}$).

In the following, we are given a plane (weakly) elementary bipartite graph G . Each function ϕ_M can be regarded as a vector $\phi_M(\mathcal{F})$ in $(\mathbb{Z}^+)^{\mathcal{F}}$. It was shown in [20] that $\phi : M \mapsto \phi_M$ defines an isomorphism between the poset $\mathcal{M}(G)$ and $\phi(\mathcal{M}(G))$, and the latter is a sublattice of $(\mathbb{Z}^+)^{\mathcal{F}}$.

THEOREM 2.3 (see [20]). *Let G be a plane (weakly) elementary bipartite graph. Then $\mathcal{M}(G)$ is a finite distributive lattice.*

The embedding ϕ of $\mathcal{M}(G)$ into $(\mathbb{Z}^+)^{\mathcal{F}}$ not only enables one to prove the above result, but also gives a labelling for all 1-factors of G . We now turn to discussing a generation of $\mathcal{M}(G)$ with such a labelling.

In obtaining Theorem 2.3, *unit region* and *unit decomposition* of a cycle system $M \oplus M'$ each play a crucial role. For convenience, we now introduce an equivalent form. For $M, M' \in \mathcal{M}(G)$, $\mathcal{C} := M \oplus M'$ consists of pairwise disjoint M and M' -alternating cycles. Let $\psi_{MM'}(f)$ be the number of proper M -alternating cycles in $\mathcal{C}$ with f in their interiors minus the number of improper M -alternating cycles in $\mathcal{C}$ with f in their interiors.

In a sense that $\psi_{MM'}(f)$ here may be allowed to be negative, it is slightly different from the original definition. Accordingly, by Lemmas 4.1 and 4.7 in [20], a relation among the functions $\phi_M, \phi_{M'}$, and $\psi_{MM'}$ can be expressed as follows.

LEMMA 2.4. *For $M, M' \in \mathcal{M}(G)$, $\phi_M - \phi_{M'} = \psi_{MM'}$.*

From the lemma, as a special case we immediately arrive at the following result, which will play an important role in the rest of this paper.

LEMMA 2.5. *For $M, M' \in \mathcal{M}(G)$, M covers M' if and only if $\phi_M(f) - \phi_{M'}(f) = 1$ for $f = f_0$, where f_0 is an inner face bounded by the cycle $M \oplus M'$, and 0 for the other faces in $\mathcal{F}$.*

Based on Theorems 2.2 and 2.3, we can give a generation procedure of lattice $\mathcal{M}(G)$. The main principle is the well-known Jordan–Dedekind theorem in finite distributive lattices: all maximal chains between a pair of elements have the same length. That is, all directed paths between any pair of vertices $\vec{Z}(G)$ are of the same length whenever they exist. Further, Lemma 2.5 also enables one to give the embedding ϕ of $\mathcal{M}(G)$ into $(\mathbb{Z}^+)^{\mathcal{F}}$.

We now describe an outline. First, construct the root 1-factor M_0 and let $\phi_{M_0} \equiv 0$. The inductive procedure is as follows: Suppose that a 1-factor M and its labelling ϕ_M have been already given. For each improper M -alternating face f_0 , $M \oplus f_0$ is another 1-factor covering M , while $\phi_{M \oplus f_0}(f_0) = \phi_M(f_0) + 1$ and the others remain unchanged. Note in the $M \oplus f_0$ operation, f_0 is always regarded as the set of edges bounding the face. A detailed procedure is given as follows.

ALGORITHM 2.6. Input: *A benzenoid system or plane weakly-elementary bipartite graph G .*

Output: *Lattice $\mathcal{M}$, and labelling list L for all 1-factors of G .*

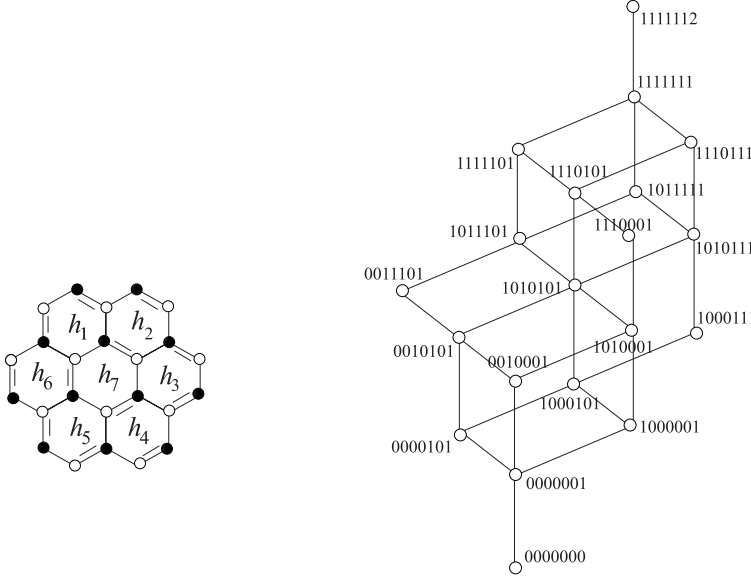


FIG. 5. The coronene and the distributive lattice on the 1-factors with labelling.

Step 0. Set $\mathcal{M} := \emptyset, L := \emptyset$, and $i = 0$; construct the root 1-factor M of G and set $S_i := \{M\}$ and $\phi_M \equiv 0$.

Step 1. Set $\mathcal{M} := \mathcal{M} \cup S_i, L := L \cup \{\phi_M : M \in S_i\}$.

Step 2. If S_i has exactly one element M and there is no improper M -alternating face, stop.

Step 3. Set $S_{i+1} := \{M \oplus f_0 : M \in S_i, f_0 \text{ is an improper } M\text{-alternating face of } G\}$. For each $M \oplus f_0$ in S_{i+1} ($M \oplus f_0$ covers M), compute $\phi_{M \oplus f_0} : \phi_{M \oplus f_0}(f_0) := \phi_M(f_0) + 1$ and $\phi_{M \oplus f_0}(f) := \phi_M(f)$ for the other faces f in $\mathcal{F}$.

Step 4. Set $i := i + 1$, then go to step 1.

When this algorithm stops, it also gives the height of $\mathcal{M}(G)$ (it is equal to the diameter of the resonance graph of G ; see Corollary 3.5). In fact, S_i consists of all 1-factors with height i . As for step 0, the root 1-factor can be obtained in the following way: First, find any 1-factor of G with the Hungarian algorithm, then repeatedly twist proper alternating inner faces. Finally, we arrive at the root 1-factor. This procedure can be completed in $O(n^2)$ steps, where n is the number of inner faces of G . This is because [29] the height is no more than $\lceil \frac{n(n+2)}{4} \rceil$.

For example, we now implement Algorithm 2.6 to the coronene. The enumeration $h_1, \dots, h_7$ of its hexagons and the root 1-factor are referred to in Figure 5 (left). When we accomplish the algorithm, the distributive lattice associated with the coronene is produced in Figure 5 (right). For convenience, all of the 1-factors are replaced with their labels. For example, 0000000 corresponds to the root 1-factor, while 1111112 represents the maximum element $M_{\hat{1}}$. In general, the following result shows that such a label can determine uniquely a 1-factor.

PROPOSITION 2.7. For any given ϕ_M , let $\phi_M \cdot \mathcal{F} := \bigoplus_{f \in \mathcal{F}} \phi_M(f)f$. Then

$$(1) \quad M = M_{\hat{0}} \oplus (\phi_M \cdot \mathcal{F}).$$

Proof. $\mathcal{M}(G)$ has a sequence $M_t (= M) \cdots M_0 (= M_{\hat{0}})$ such that each M_i covers

M_{i-1} for $1 \leq i \leq t$. By Lemma 2.5, $M_i \oplus M_{i-1}$ is a proper M_i -alternating face f_i and $\phi_{M_i}(f) - \phi_{M_{i-1}}(f) = 1$ if $f = f_i$, 0, otherwise. That is, $M_i \oplus M_{i-1} = \bigoplus_{f \in \mathcal{F}} (\phi_{M_i}(f) - \phi_{M_{i-1}}(f))f = (\phi_{M_i} - \phi_{M_{i-1}}) \cdot \mathcal{F}$. Hence $M \oplus M_0 = \bigoplus_{i=1}^t (M_i \oplus M_{i-1}) = \bigoplus_{i=1}^t (\phi_{M_i} - \phi_{M_{i-1}}) \cdot \mathcal{F} = \sum_{i=1}^t (\phi_{M_i} - \phi_{M_{i-1}}) \cdot \mathcal{F} = \phi_M \cdot \mathcal{F}$. Here the symmetry difference agrees with addition module 2 and every face $f \in \mathcal{F}$ can be always regarded as the set of edges bounding it. $\square$

Given two labels 1111112 and 1111111, for example, we shall determine their corresponding 1-factors of coronene. For the former, $M_0 \oplus (h_1 \oplus h_2 \oplus h_3 \oplus h_4 \oplus h_5 \oplus h_6 \oplus 2h_7) = M_0 \oplus (h_1 \oplus h_2 \oplus h_3 \oplus h_4 \oplus h_5 \oplus h_6) = M_0 \oplus (C \oplus h_7) = M_1$, where C denotes the boundary of the coronene. For the latter, $M_0 \oplus (h_1 \oplus h_2 \oplus h_3 \oplus h_4 \oplus h_5 \oplus h_6 \oplus h_7) = M_1 \oplus h_7$.

3. Median graph. Klavžar, Žigert, and Brinkmann [17] showed that the resonance graphs of cata-condensed benzenoid systems, or a subclass of bipartite outplane graphs, are median graphs. By the established distributive lattices and the developed techniques, we now show that the above results hold generally for benzenoid systems, or more general, plane weakly-elementary bipartite graphs G .

Given a connected graph H with vertices u and v , the *interval* $I_H(u, v)$ (or simply $I(u, v)$) between u and v consists of all vertices on shortest paths between u and v . A *median* of vertices u, v , and w is a vertex that lies in $I(u, v) \cap I(u, w) \cap I(v, w)$. A graph is called a *median graph* if every triple of its vertices has a unique median (cf. [12]). We now give the main result of this section as follows.

THEOREM 3.1. *The resonance graph of a plane weakly-elementary bipartite graph G is a median graph.*

To prove the theorem we first give three explicit formulae on distance $d_{Z(G)}(M, M')$, the length of a shortest path between a pair of vertices M and M' , in the resonance graph $Z(G)$.

THEOREM 3.2 (distance formula 1). $d_{Z(G)}(M, M') = \sum_{f \in \mathcal{F}} |\phi_M(f) - \phi_{M'}(f)|$.

Proof. Let $P = M_0 (= M)M_1 \cdots M_t (= M')$ be a path in $Z(G)$ between M and M' . Then the length

$$\begin{aligned} l(P) &= t = \sum_{i=0}^{t-1} \sum_{f \in \mathcal{F}} |\phi_{M_{i+1}}(f) - \phi_{M_i}(f)| \quad (\text{by Lemma 2.5}) \\ &= \sum_{f \in \mathcal{F}} \sum_{i=0}^{t-1} |\phi_{M_{i+1}}(f) - \phi_{M_i}(f)| \\ (2) \quad &\geq \sum_{f \in \mathcal{F}} |\sum_{i=0}^{t-1} (\phi_{M_{i+1}}(f) - \phi_{M_i}(f))| \\ &= \sum_{f \in \mathcal{F}} |\phi_M(f) - \phi_{M'}(f)|. \end{aligned}$$

We now choose a special path $P_* := N_s (= M) \cdots N_1 N_0 N_{-1} \cdots N_{-s'} (= M')$ through the *meet* $N_0 := M \wedge M'$ such that N_i covers N_{i-1} and N_{-j} covers $N_{-(j-1)}$ for all $s \geq i \geq 1$ and $s' \geq j \geq 1$. Recall that $\phi : \mathcal{M}(G) \rightarrow \phi(\mathcal{M}(G))$ is an isomorphism and $\phi(\mathcal{M}(G))$ is a sublattice of $(\mathbb{Z}^+)^{\mathcal{F}}$. Then $\phi_{N_0} = \phi_{M \wedge M'} = \phi_M \wedge \phi_{M'}$, and $\phi_{N_0}(f) = \min(\phi_M(f), \phi_{M'}(f))$ for each $f \in \mathcal{F}$. Hence by Lemma 2.5 and similar arguments as above (2) we have that

$$\begin{aligned} l(P_*) &= s + s' = \sum_{f \in \mathcal{F}} (\phi_M(f) - \phi_{N_0}(f)) + \sum_{f \in \mathcal{F}} (\phi_{M'}(f) - \phi_{N_0}(f)) \\ &= \sum_{f \in \mathcal{F}} (\phi_M(f) + \phi_{M'}(f) - 2\phi_{N_0}(f)) \\ (3) \quad &= \sum_{f \in \mathcal{F}} \{\phi_M(f) + \phi_{M'}(f) - 2\min(\phi_M(f), \phi_{M'}(f))\} \\ &= \sum_{f \in \mathcal{F}} |\phi_M(f) - \phi_{M'}(f)| \\ &\leq l(P). \end{aligned}$$

This implies that P_* is a shortest path of $Z(G)$ and therefore the theorem holds. $\square$

The above distance formula shows that the embedding $\phi : \mathcal{M}(G) \rightarrow (\mathbb{Z}^+)^{\mathcal{F}}$ is also distance-preserving.

COROLLARY 3.3 (distance formula 2). $d_{Z(G)}(M, M') = d_{Z(G)}(M \vee M', M \wedge M')$.

Proof. By Theorem 3.2 we have that

$$\begin{aligned}
 & d_{Z(G)}(M \vee M', M \wedge M') \\
 &= \sum_{f \in \mathcal{F}} \{ \phi_{M \vee M'}(f) - \phi_{M \wedge M'}(f) \} \\
 (4) \quad &= \sum_{f \in \mathcal{F}} \{ \max(\phi_M(f), \phi_{M'}(f)) - \min(\phi_M(f), \phi_{M'}(f)) \} \\
 &= \sum_{f \in \mathcal{F}} |\phi_M(f) - \phi_{M'}(f)| = d_{Z(G)}(M, M'). \quad \square
 \end{aligned}$$

Combining Lemma 2.4 with Theorem 3.2, we have the following third formula.

COROLLARY 3.4 (distance formula 3). $d_{Z(G)}(M, M') = \sum_{f \in \mathcal{F}} |\psi_{MM'}(f)|$.

The *height* (or *rank*) of $\mathcal{M}(G)$ is the length of a maximal chain between the greatest and least elements, that is, the length of a directed path of $\vec{Z}(G)$ from the source to sink. Corollary 3.3 shows that $d_{Z(G)}(M_0, M_1)$ is the diameter (i.e., the largest distance between all pairs of vertices) of $Z(G)$. So we have the following result.

COROLLARY 3.5. *The diameter of $Z(G)$ is equal to the height of $\mathcal{M}(G)$.*

LEMMA 3.6. *For any $M_1, M_2 \in \mathcal{M}(G)$, $I_{Z(G)}(M_1, M_2) = I_{\mathcal{M}(G)}(M_1 \vee M_2, M_1 \wedge M_2)$.*

Proof. For any $M \in \mathcal{M}(G)$, by Theorem 3.2 we have that

$$\begin{aligned}
 d_{Z(G)}(M_1, M_2) &= \sum_{f \in \mathcal{F}} |\phi_{M_1}(f) - \phi_{M_2}(f)| \\
 &= \sum_{f \in \mathcal{F}} |(\phi_{M_1}(f) - \phi_M(f)) + (\phi_M(f) - \phi_{M_2}(f))| \\
 (5) \quad &\leq \sum_{f \in \mathcal{F}} \{ |\phi_{M_1}(f) - \phi_M(f)| + |\phi_M(f) - \phi_{M_2}(f)| \} \\
 &= \sum_{f \in \mathcal{F}} |\phi_{M_1}(f) - \phi_M(f)| + \sum_{f \in \mathcal{F}} |\phi_M(f) - \phi_{M_2}(f)| \\
 &= d_{Z(G)}(M_1, M) + d_{Z(G)}(M, M_2).
 \end{aligned}$$

Then $M \in I_{Z(G)}(M_1, M_2) \iff$ for all f , $\phi_{M_1}(f) - \phi_M(f)$ and $\phi_M(f) - \phi_{M_2}(f)$ have the same sign $\iff$ for all f , $\min(\phi_{M_1}(f), \phi_{M_2}(f)) \leq \phi_M(f) \leq \max(\phi_{M_1}(f), \phi_{M_2}(f)) \iff \phi_{M_1 \wedge M_2} \leq \phi_M \leq \phi_{M_1 \vee M_2} \iff M_1 \wedge M_2 \preceq M \preceq M_1 \vee M_2$. $\square$

Proof of Theorem 3.1. It is sufficient to show that any triple of 1-factors M_1, M_2 , and M_3 in $\mathcal{M}(G)$ have a unique median. If M_1, M_2 , and M_3 have a median M_0 , then by Lemma 3.6 we have that $M_1 \wedge M_2 \preceq M_0 \preceq M_1 \vee M_2$, $M_2 \wedge M_3 \preceq M_0 \preceq M_2 \vee M_3$ and $M_1 \wedge M_3 \preceq M_0 \preceq M_1 \vee M_3$, which imply that

$$(6) \quad (M_1 \wedge M_2) \vee (M_2 \wedge M_3) \vee (M_1 \wedge M_3) \preceq M_0 \preceq (M_1 \vee M_2) \wedge (M_2 \vee M_3) \wedge (M_1 \vee M_3).$$

Since $\mathcal{M}(G)$ is a distributive lattice (Theorem 2.3), by [1, Theorem 8, Chapter 2] both equalities in (6) hold. This also implies that $(M_1 \wedge M_2) \vee (M_2 \wedge M_3) \vee (M_1 \wedge M_3)$ is a unique median of M_1, M_2 , and M_3 . $\square$

4. Binary coding for the 1-factors. Theorem 3.1 shows that any connected resonance graph is a median graph. Hence it can be isometrically embedded into hypercubes [12]. That is, every 1-factor can be assigned a 0,1-sequence (binary coding) of a given length so that the distance between two 1-factors is equal to the number

of positions for which both of the corresponding sequences differ. In this section, we first give a criterion when ϕ_M itself is such a binary coding for all 1-factors M of a benzenoid system, or a plane weakly-elementary bipartite graph. Then, in a general situation we show how ϕ can be transformed into a binary coding.

Let G be a plane bipartite graph with a 1-factor. A subgraph H of G is said to be *nice* if $G - V(H)$ has a 1-factor. A set S of pairwise disjoint inner faces of G is called a *resonant pattern* of G if G has a 1-factor M such that the boundary of each face in S is an M -alternating cycle. Let $c(G)$ and $k(G)$ denote the numbers of resonant patterns and 1-factors of G , respectively.

LEMMA 4.1 (see [13, 25]). *Let G be a plane bipartite graph with a 1-factor. Then $c(G) \leq k(G)$, and equality holds if and only if G is weakly elementary and for any pair of disjoint cycles that form a nice subgraph, their interiors are disjoint.*

LEMMA 4.2 (see [27]). *For a benzenoid system G , $c(G) = k(G)$ if and only if G has no coronene as its nice subgraph.*

THEOREM 4.3. *Let G be a plane weakly-elementary bipartite graph. Then ϕ is an embedding of $\mathcal{M}(G)$ into $\{0, 1\}^{\mathcal{F}}$ if and only if $c(G) = k(G)$; equivalently, for any pair of disjoint cycles that form a nice subgraph their interiors are disjoint.*

Proof. If $\phi_M(f) \geq 2$ for some $M \in \mathcal{M}(G)$ and $f \in \mathcal{F}$, then there exist two disjoint proper M -alternating cycles C_1 and C_2 such that f lies in both interiors of C_1 and C_2 . Conversely, suppose that G has a 1-factor M and a pair of disjoint proper M -alternating cycles C_1 and C_2 such that the interior of C_1 is contained in the interior of C_2 . Then $M' := M \oplus C_1 \oplus C_2 \in \mathcal{M}(G)$. Since $M \oplus M' = C_1 \cup C_2$, and both C_1 and C_2 are proper M and improper M' -alternating cycles, $\psi_{MM'}(f) = 2$ for any inner face f of G contained in the interior of C_1 . By Lemma 2.4, $\phi_M(f) - \phi_{M'}(f) = \psi_{MM'}(f) = 2$. This implies that $\phi_M(f) \geq 2$. Further, the other equivalency is implied by Lemma 4.1. $\square$

The following result is an immediate consequence of Lemma 4.2 and Theorem 4.3.

COROLLARY 4.4. *Let G be a benzenoid system. Then ϕ is an embedding of $\mathcal{M}(G)$ into $\{0, 1\}^{\mathcal{F}}$ if and only if G has no coronene as its nice subgraph.*

COROLLARY 4.5. *Let G be a benzenoid system. If G has no coronene as its nice subgraph, then the height of $\mathcal{M}(G)$ is equal to the number of resonant hexagons of G .*

Proof. Recall that M_1 is the maximum element of $\mathcal{M}(G)$. Then by distance formula 1 the height $d = \sum_{f \in \mathcal{F}} \phi_{M_1}(f)$. Since G has no coronene as its nice subgraph, by Corollary 4.4 $\phi_{M_1}(f) = 0$ or 1 for each $f \in \mathcal{F}$. For a resonant hexagon h , there exists a 1-factor M such that h is proper M -alternating. Further, by Lemma 2.5 we have that $\phi_M(h) - \phi_{M'}(h) = 1$, where $M' := M \oplus h$, which implies that $1 \leq \phi_M(h) \leq \phi_{M_1}(h) \leq 1$. Hence $\phi_{M_1}(h) = 1$. But for a nonresonant hexagon h , the $\phi_M(h)$ remain unchanged (zero). Hence d is equal to the number of resonant hexagons of G . $\square$

We have already given a criterion that every ϕ_M is a binary coding of M . In a general situation we show how ϕ can be transformed into an embedding $\bar{\phi}$ of $\mathcal{M}(G)$ into Boolean algebra $B_d = \{0, 1\}^d$, where $d = \sum_{f \in \mathcal{F}} \phi_{M_1}(f)$. In fact such an embedding gives a binary coding of length d for all 1-factors of G .

For each $M \in \mathcal{M}(G)$, the function ϕ_M on $\mathcal{F}$ is naturally transformed into the function $\bar{\phi}_M$ on $\mathcal{F}$ as follows: For each $f \in \mathcal{F}$, $\bar{\phi}_M(f)$ is a sequence of length $\phi_{M_1}(f)$ such that the first $\phi_M(f)$ positions from the left side are all placed 1 and the others 0. Since $0 \leq \phi_M(f) \leq \phi_{M_1}(f)$, $\bar{\phi}_M(f)$ is well defined. For the coronene, for example, $\phi_{M_1}(h_7) = 2$ and $\phi_{M_1}(h_i) = 1 (1 \leq i \leq 6)$. Then $\bar{\phi}_{M_1}(h_7) = 11$ and $\bar{\phi}_{M_1}(h_i) = 1 (1 \leq i \leq 6)$ (see Figure 6); while for 1-factor M with $\phi_M = 1111111$, $\bar{\phi}_M(h_7) = 10$ and $\bar{\phi}_M(h_i) = 1 (1 \leq i \leq 6)$.

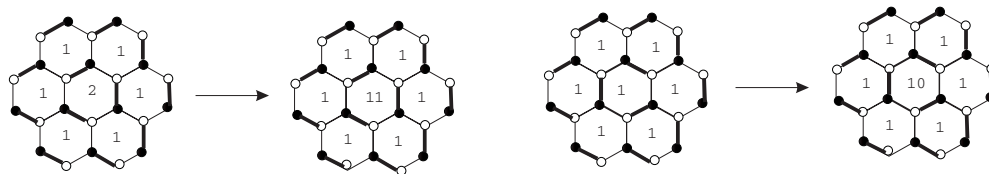
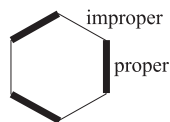
FIG. 6. The transformation of ϕ_M into a binary coding $\bar{\phi}_M$.

FIG. 7. Proper and improper positions of a hexagon.

The following result follows immediately from the property of distance-preserving embedding ϕ .

THEOREM 4.6. $\bar{\phi} : \mathcal{M}(G) \rightarrow \{0, 1\}^d$ is a distance-preserving embedding, where d is the height of $\mathcal{M}(G)$.

5. An algorithm. For a cata-condensed benzenoid system G with d hexagons, there are 2^d different ways by which the resonance graph can be isometrically embedded into a d -hypercube; for example, [15, 16] gave different binary codings. In fact such embeddings are all equivalent under the sense that they can be transformed into each other by permutating the coordinates; though, only ϕ is both order-preserving and distance-preserving embedding. In general, Algorithm 2.6 gives a generation procedure for all ϕ_M . For this special case, we now design a simpler algorithm by modifying Klavžar et al.'s method [16].

If $d \geq 2$, then G has a hexagon that shares exactly one edge with the remainder of G when removing the hexagon. This hexagon corresponds to a monovalency vertex of its inner dual graph (tree). Hence G has a sequence of hexagons $h_1, h_2, \dots, h_d$, so that each h_j shares one edge (say e_j) with exactly one hexagon (say $h_{a(j)}$) of $h_1, \dots, h_{j-1}$. We call it a *normal* sequence of hexagons of G .

For a given hexagon, the three thick edges in Figure 7 are said to be in *proper* positions, the other thin edges in *improper* positions. So the common edge of two adjacent hexagons is in a proper position of one hexagon and in an improper position of the other one.

ALGORITHM 5.1. Input: A cata-condensed benzenoid system G with a normal sequence of hexagons $h_1, h_2, \dots, h_d$ ($d \geq 1$).

Output: A binary coding ϕ for all 1-factors of G : L_d .

Step 0. $i := 1, L_i := \{0, 1\}$.

Step 1. If $i = d$, stop.

Step 2. Let e_{i+1} be the common edge of h_{i+1} with one (say $h_{a(i+1)}$) of $h_1, \dots, h_i$. If e_{i+1} lies in a proper position of $h_{a(i+1)}$, then set $L_{i+1} := \{x0 : x \in L_i\} \cup \{x1 : x \in L_i \text{ and } (x)_{a(i+1)} = 1\}$; otherwise, set $L_{i+1} := \{x1 : x \in L_i\} \cup \{x0 : x \in L_i \text{ and } (x)_{a(i+1)} = 0\}$.

Step 3. $i := i + 1$, go to step 1.

As an example, we implement the algorithm to Dibenz[a,c]anthracene. For a given normal sequence of hexagons $h_1, \dots, h_5$ (see Figure 8), e_2 and e_3 are in proper

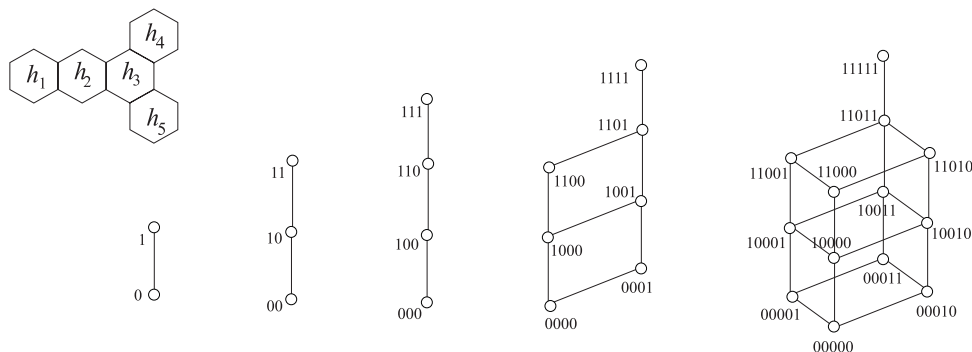


FIG. 8. An example for implementing Algorithm 5.1.

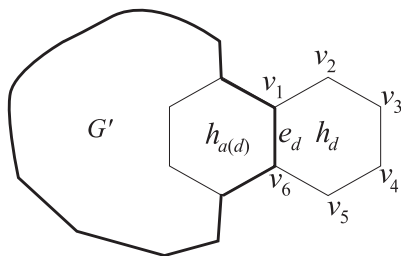


FIG. 9. Illustration for the proof of Theorem 5.2.

positions of h_1 and h_2 respectively; both e_4 and e_5 are in improper positions of h_3 . Figure 8 illustrates a process of implementing Algorithm 5.1. The binary coding for all of the 1-factors of Dibenz[a,c]anthracene with lattice structure is denoted in the last one. The following theorem gives the correctness of this algorithm.

THEOREM 5.2. *Algorithm 5.1 determines the binary coding ϕ_M for all 1-factors M of a cata-condensed benzenoid system G .*

Proof. We proceed by induction on the number d of hexagons of G . If $d = 1$, it is trivial because 1 and 0 correspond to two 1-factors entirely lying in the proper and improper positions of h_1 , respectively (see Figure 7).

From now on we suppose that $d \geq 2$. Recall that e_d is the common edge of h_d and some hexagon $h_{a(d)}$. Then $P := h_d - e_d$ is a path of length 5, denoted by $v_1 v_2 \cdots v_6$. Let G' be the graph obtained from G by removing the hexagon h_d , i.e., $G' = G - \{v_2, v_3, v_4, v_5\}$; see Figure 9. Then G' has a normal sequence $h_1, \dots, h_{d-1}$ of hexagons. We claim that the restrictions of all ϕ_M on the hexagons $h_1, \dots, h_{d-1}$ become the coding $\phi_{\widetilde{M}}$ for all 1-factors $\widetilde{M}$ of G' .

Let $\mathcal{M}_0 := \{M \in \mathcal{M}(G) : \{v_2 v_3, v_4 v_5\} \subseteq M\}$. For each $M \in \mathcal{M}_0$, $\widetilde{M} := M \setminus \{v_2 v_3, v_4 v_5\}$ is a 1-factor of G' . This naturally produces a one-to-one correspondence from the 1-factors in $\mathcal{M}_0$ to the 1-factors of G' . For any $M \in \mathcal{M}(G) \setminus \mathcal{M}_0$, $\{v_1 v_2, v_3 v_4, v_5 v_6\} \subseteq M$ and $M \oplus h_d \in \mathcal{M}_0$; by Lemma 2.5 $\phi_M(h_i) = \phi_{M \oplus h_d}(h_i)$ for all $1 \leq i \leq d-1$. Hence the above claim means that

$$(7) \quad \phi_M(h_i) = \phi_{\widetilde{M}}(h_i) \text{ for all } 1 \leq i \leq d-1 \text{ and for all } M \in \mathcal{M}_0.$$

Let $\widetilde{M}_0$ denote the minimum element of $\mathcal{M}(G')$. It follows that $\phi_{\widetilde{M}_0 \cup \{v_2v_3, v_4v_5\}}(h_i) = \phi_{\widetilde{M}_0}(h_i) = 0$ for all $1 \leq i \leq d-1$. If $M_0 = \widetilde{M}_0 \cup \{v_2v_3, v_4v_5\}$, it is trivial. Otherwise, every cycle in $M_0 \oplus (\widetilde{M}_0 \cup \{v_2v_3, v_4v_5\})$ is improper M_0 -alternating and thus proper $(\widetilde{M}_0 \cup \{v_2v_3, v_4v_5\})$ -alternating. Therefore such a cycle must pass through P ; otherwise it is a proper $\widetilde{M}_0$ -alternating cycle in G' , a contradiction. Hence such a cycle is unique and denoted by C . In fact $C = h_d$; otherwise, $C + e_d$ contains a proper $\widetilde{M}_0$ -alternating cycle of G' , a contradiction. Hence $\widetilde{M}_0 \cup \{v_2v_3, v_4v_5\}$ covers M_0 and $M_0 \oplus h_d = \widetilde{M}_0 \cup \{v_2v_3, v_4v_5\}$. So Lemma 2.5 implies the required result.

For any $M \in \mathcal{M}_0$, by Lemma 2.4 we have

$$\phi_M - \phi_{\widetilde{M}_0 \cup \{v_2v_3, v_4v_5\}} = \psi_{M(\widetilde{M}_0 \cup \{v_2v_3, v_4v_5\})},$$

and

$$\phi_M(h_i) = \psi_{M(\widetilde{M}_0 \cup \{v_2v_3, v_4v_5\})}(h_i) = \psi_{\widetilde{M}\widetilde{M}_0}(h_i) = \phi_{\widetilde{M}}(h_i) - \phi_{\widetilde{M}_0}(h_i) = \phi_{\widetilde{M}}(h_i)$$

for all $1 \leq i \leq d-1$. Claim (7) is verified.

Suppose that Algorithm 5.1 has already determined the binary coding $\phi_{\widetilde{M}}(h_i)$, $1 \leq i \leq d-1$, for all 1-factors $\widetilde{M}$ of G' . Hence for each $x \in L_{d-1}$, the corresponding $\widetilde{M} \in \mathcal{M}(G')$ has $\phi_{\widetilde{M}}(h_1, \dots, h_{d-1}) = x$. For each $M := \widetilde{M} \cup \{v_2v_3, v_4v_5\} \in \mathcal{M}_0$, we have $\phi_M(h_1, \dots, h_{d-1}) = x$ by the above claim verified.

We now consider the case that e_d is in a proper position of $h_{a(d)}$. Then both v_2v_3 and v_4v_5 are in improper positions of h_d . Therefore $\{v_2v_3, v_4v_5\} \subseteq M_0$; otherwise, h_d is a proper M_0 -alternating, a contradiction. Hence h_d lies outside all cycles in $M \oplus M_0$. This implies that $\phi_M(h_d) = 0$ and $\phi_M = x0$. Since e_d is always contained in an improper $\widetilde{M}_0$ -alternating cycle, $e_d \notin \widetilde{M}_0$. This implies that $e_d \in \widetilde{M}$ if and only if e_d is contained in a cycle C in $\widetilde{M}_0 \oplus \widetilde{M}$; that is, $h_{a(d)}$ is contained in the interior of C . If $\phi_{\widetilde{M}}(h_{a(d)}) = 1$, then, equivalently, $e_d \in M$, h_d is an improper M -alternating hexagon and $\phi_{M \oplus h_d} = x1$.

For the other case that e_d is in an improper position of $h_{a(d)}$, analogous arguments can be made. We have that $\phi_M(h_d) = 1$ and $\phi_M = x1$. If $\phi_{\widetilde{M}}(h_{a(d)}) = 0$, then, equivalently, $e_d \in M$, h_d is a proper M -alternating hexagon and $\phi_{M \oplus h_d} = x0$. These conform to step 2 in Algorithm 5.1. Therefore Algorithm 5.1 determines the binary coding ϕ_M for all 1-factors M of G . $\square$

Theorem 5.2 shows that the binary coding ϕ_M for the 1-factors M of a cata-condensed benzenoid system can be accomplished by Algorithm 5.1. Hence such a binary coding produced by Algorithm 5.1 is independent of its constructive process (i.e., choices for a normal sequence of hexagons) in a sense that each bit corresponds to a specific hexagon. This algorithm is still valid for outplane bipartite graphs.

6. Conclusion. If a plane bipartite graph G is weakly elementary, then from the distributive lattice structure on the 1-factors we know that the resonance graph is a median graph. If G is nonweakly elementary, then we can see that each component of the poset $\mathcal{M}(G)$ is a distributive lattice [29] and each component of the resonance graph is a median graph.

For the weakly-elementary case, the mapping $\phi : \mathcal{M}(G) \rightarrow (\mathbb{Z}^+)^{\mathcal{F}}$ is both order-preserving and distance-preserving embedding. When ϕ is transformed into an embedding $\bar{\phi}$ of $\mathcal{M}(G)$ into the Boolean algebra B_d , $\bar{\phi}$ gives an isometric embedding of the resonance graph $Z(G)$ into a d -hypercube—the covering graph of the Boolean

algebra B_d of rank d . Here d is equal to both the height of $\mathcal{M}(G)$ and the diameter of resonance graph $Z(G)$. We conclude by proposing two open problems.

Open problem 1. Determine the smallest integer m such that $\mathcal{M}(G)$ can be embedded into $(\mathbb{Z}^+)^m$. We may conjecture that such smallest integer m is equal to the *resonant number* or *Clar number* in benzenoid systems (i.e., the maximum size of resonant patterns). In Figure 5 we can see that the conjecture is true for the coronene since both smallest m and Clar number are equal to 3, much less than the number 7 of hexagons.

A similar problem was considered recently [6, 14]: If a graph can be isometrically embedded into the m -dimensional integer lattice $\mathbb{Z}^m$ (an infinite graph), then we find such minimum possible dimension m , called the *lattice dimension* of the graph. The lattice dimension of benzenoid systems was determined by Klavžar and Kovše [14] in terms of fundamental trees. Substantially, open problem 1 is to determine the lattice dimension of the resonance graph.

Both embeddings ϕ and $\bar{\phi}$ can be regarded as labellings for the 1-factors, which can be generated by Algorithm 2.6. For the special case of cata-condensed benzenoid system, such two labellings become the same binary coding for all 1-factors, which can be generated by Algorithm 5.1 without involving the generation of all 1-factors; its encoding gives the corresponding 1-factors (see (1)).

Open problem 2. For a general benzenoid system G , how can one design an algorithm to generate the binary coding for all 1-factors, avoiding the generation of $\mathcal{M}(G)$?

Acknowledgment. The authors are grateful to the referee for giving valuable suggestions to the improvement of this manuscript.

REFERENCES

- [1] G. BIRKHOFF, *Lattice Theory*, 3rd ed., AMS, Providence, RI, 1979.
- [2] R. CHEN AND F. ZHANG, *Hamilton paths in Z-transformation graphs of perfect matchings of hexagonal systems*, Discrete Appl. Math., 74 (1997), pp. 191–196.
- [3] N. G. CHOPRA, R. J. LUYKEN, K. CHERREY, V. H. CRESPI, M. L. COHEN, S. G. LOUIE, AND A. ZETTL, *Boron-nitride nanotubes*, Science, 269 (1995), pp. 966–967.
- [4] S. J. CYVIN, J. BRUNVOLL, AND B. N. CYVIN, *Theory of Coronoid Hydrocarbons*, Lecture Notes in Chemistry 54, Springer, Berlin, 1991.
- [5] S. J. CYVIN AND I. GUTMAN, *Kekulé Structures in Benzenoid Hydrocarbons*, Springer, Berlin, 1988.
- [6] D. EPPSTEIN, *The lattice dimension of a graph*, European J. Combin., 26 (2005), pp. 585–592.
- [7] J. C. FOURNIER, *Combinatorics of perfect matchings in plane bipartite graphs and application to tilings*, Theoret. Comput. Sci., 303 (2003), pp. 333–351.
- [8] P. W. FOWLER, T. HEINE, D. MITCHELL, R. SCHMIDT, AND G. SEIFERT, *Boron-nitrogen analogues of the fullerenes: The isolated-square rule*, J. Chem. Soc. Faraday Trans., 92 (1996), pp. 2197–2201.
- [9] G. GRÄTZER, *Lattice Theory, First Concepts and Distributive Lattice*, W. H. Freeman, San Francisco, 1971.
- [10] I. GUTMAN, *Topological properties of benzenoid systems*, in Advances in the Theory of Benzenoid Hydrocarbons II, Topics in Current Chem. 162, Springer, Berlin, 1992, pp. 1–28.
- [11] I. GUTMAN, A. V. TEODOROVIĆ, AND N. KOLAKOVIĆ, *An application of corals*, J. Serb. Chem. Soc., 55 (1990), pp. 363–368.
- [12] W. IMRICH AND S. KLAVŽAR, *Product Graphs: Structure and Recognition*, John Wiley & Sons, New York, 2000.
- [13] P. E. JOHN, H. SACHS, AND M. ZHENG, *Kekulé patterns and Clar patterns in bipartite plane graphs*, J. Chem. Inf. Comput. Sci., 35 (1995), pp. 1019–1021.
- [14] S. KLAVŽAR AND M. KOVŠE, *The lattice dimension of benzenoid systems*, MATCH Commun. Math. Comput. Chem., 56 (2006), pp. 637–648.

- [15] S. KLAVŽAR, A. VESEL, AND P. ŽIGERT, *On resonance graphs of catacondensed hexagonal graphs: Structure, coding, and Hamiltonian path algorithm*, MATCH Commun. Math. Comput. Chem., 49 (2003), pp. 99–116.
- [16] S. KLAVŽAR, A. VESEL, P. ŽIGERT, AND I. GUTMAN, *Binary coding of Kekulé structures of catacondensed benzenoid hydrocarbons*, Comput. Chem., 25 (2001), pp. 569–575.
- [17] S. KLAVŽAR, P. ŽIGERT, AND G. BRINKMANN, *Resonance graphs of catacondensed even ring systems are median*, Discrete Math., 253 (2002), pp. 35–43.
- [18] S. KLAVŽAR, P. ŽIGERT, AND I. GUTMAN, *Clar number of catacondensed benzenoid hydrocarbons*, J. Mol. Struct. (Theochem), 586 (2002), pp. 235–240.
- [19] P. C. B. LAM, W. C. SHIU, AND H. ZHANG, *Elementary blocks of plane bipartite graphs*, MATCH Commun. Math. Comput. Chem., 49 (2003), pp. 127–137.
- [20] P. C. B. LAM AND H. ZHANG, *A distributive lattice on the set of perfect matchings of a plane bipartite graph*, Order, 20 (2003), pp. 13–29.
- [21] L. LOVÁSZ AND M. D. PLUMMER, *Matching Theory*, Annals of Discrete Math. 29, North-Holland, Amsterdam, 1986.
- [22] M. RANDIĆ, *Resonance in catacondensed benzenoid hydrocarbons*, Int. J. Quantum Chem., 63 (1997), pp. 585–600.
- [23] M. RANDIĆ, *Aromaticity of polycyclic conjugated hydrocarbons*, Chem. Rev., 103 (2003), pp. 3449–3605.
- [24] M. RANDIĆ, D. J. KLEIN, S. EL-BASIL, AND P. CALKINS, *Resonance in large benzenoid hydrocarbons*, Croat. Chem. Acta, 69 (1996), pp. 1639–1660.
- [25] W. C. SHIU, P. C. B. LAM, F. ZHANG, AND H. ZHANG, *Normal components, Kekulé patterns, and Clar patterns in plane bipartite graphs*, J. Math. Chem., 31 (2002), pp. 405–420.
- [26] P. R. STANLEY, *Enumerative Combinatorics*, Vol. 1, Wadsworth, Belmont, CA, 1986.
- [27] F. ZHANG AND R. CHEN, *A theorem concerning polyhex graphs*, MATCH Commun. Math. Comput. Chem., 19 (1986), pp. 179–188.
- [28] F. ZHANG, X. GUO, AND R. CHEN, *Z-transformation graphs of perfect matchings of hexagonal systems*, Discrete Math., 72 (1988), pp. 405–415.
- [29] H. ZHANG, *Direct sum of distributive lattices on the set of perfect matchings of a plane bipartite graph*, preprint.
- [30] H. ZHANG AND F. ZHANG, *The rotation graphs of perfect matchings of plane bipartite graphs*, Discrete Appl. Math., 73 (1997), pp. 5–12.
- [31] H. ZHANG AND F. ZHANG, *Block graphs of Z-transformation graphs of perfect matchings of plane elementary bipartite graphs*, Ars Combin., 53 (1999), pp. 309–314.
- [32] H. ZHANG AND F. ZHANG, *Plane elementary bipartite graphs*, Discrete Appl. Math., 105 (2000), pp. 291–311.
- [33] H. ZHANG, F. ZHANG, AND H. YAO, *Z-transformation graphs of perfect matchings of plane bipartite graphs*, Discrete Math., 276 (2004), pp. 393–404.

LARGEST NUMBER OF SUBTREES OF TREES WITH A GIVEN MAXIMUM DEGREE*

RUSSELL KIRK[†] AND HUA WANG[‡]

Abstract. We characterize trees, given a size and maximum vertex degree, which have the greatest number of subtrees. This generalizes the recent result of Szekely and Wang. These trees coincide with those which were shown by Fischermann et al., and independently Jelen and Triesch, to minimize the Wiener index.

Key words. tree, subtree, Wiener index, maximum degree

AMS subject classifications. 05C05, 05C90

DOI. 10.1137/070687736

1. Terminology. All graphs in this paper will be finite, simple, and undirected. A *tree* $T = (V, E)$ is a connected, acyclic graph. $V(T)$ denotes the vertex set of a tree T . We refer to vertices of degree 1 of T as *leaves*. The unique path connecting two vertices v, u in T will be denoted by $P_T(v, u)$. For a tree T and two vertices v, u of T , the *distance* $d_T(v, u)$ between them is the number of edges on the path $P_T(v, u)$. For a vertex v of T , define the *distance of the vertex* as $g_T(v) = \sum_{u \in V(T)} d_T(v, u)$. Then $\sigma(T) = \frac{1}{2} \sum_{v \in V(T)} g_T(v)$ denotes the *Wiener index* of T .

We call a tree (T, r) *rooted at the vertex* r (or just T if it is clear what the root is) by specifying a vertex $r \in V(T)$. For any two different vertices u, v in a rooted tree (T, r) , we say that v is a *successor* of u if $P_T(r, u) \subset P_T(r, v)$. Furthermore, if u and v are adjacent to each other and $d_T(r, u) = d_T(r, v) - 1$, we say that u is a *parent* of v and v is a *child* of u . A subtree T' of a tree T will often be described by its vertex set, $T \setminus T'$ denotes the tree resulting from removing the edges of T' (and the resulting isolated vertices) from T .

If v is any vertex of a rooted tree (T, r) , let $T(v)$, *the subtree induced by* v , denote the rooted subtree of T that is induced by v and all of its successors in T and is rooted at v .

The *height* of a vertex v of a rooted tree T with root r is $h_T(v) = d_T(r, v)$, and the *height* of a rooted tree T is $h(T) = \max_{v \in T} h_T(v)$, the maximum height of vertices.

(T, r, k) denotes a tree T rooted at r , where the degree of any vertex is less than or equal to k . For a rooted tree T , $C_T(v)$ denotes the set of children of a vertex v .

For a tree T and vertices $v_1, v_2, \dots, v_{n-1}, v_n$ of T , let $f_T(v_1, v_2, \dots, v_{n-1}, v_n)$ denote the number of subtrees of T that contain the vertices $v_1, v_2, \dots, v_{n-1}$ and v_n , and let $F(T)$ denote the number of nonempty subtrees (connected subgraphs with at least one vertex) of T .

For a tree (T, r, k) , label the vertices in $C_T(r)$ as $r_1, r_2, \dots, r_m$ such that

$$f_{T(r_1)}(r_1) \leq f_{T(r_2)}(r_2) \leq \dots \leq f_{T(r_m)}(r_m).$$

*Received by the editors April 9, 2007; accepted for publication (in revised form) February 21, 2008; published electronically May 30, 2008.

<http://www.siam.org/journals/sidma/22-3/68773.html>

[†]Department of Mathematics, University of Florida, Gainesville, FL 32611 (russcore@ufl.edu).

[‡]Corresponding author. Department of Mathematics, University of Florida, Gainesville, FL 32611 (hua@math.ufl.edu).

Similarly, for $T(r_j)$, label the vertices in $C_T(r_j)$ as $r_{j1}, r_{j2}, \dots, r_{jl}$ according to the same rule (Figure 1). Hence for any vertex $r_{i_1 i_2 \dots i_s}$, we always have

$$(1) \quad f_{T(r_{i_1 \dots i_{s1}})}(r_{i_1 \dots i_{s1}}) \leq f_{T(r_{i_1 \dots i_{s2}})}(r_{i_1 \dots i_{s2}}) \leq \dots \leq f_{T(r_{i_1 \dots i_{st}})}(r_{i_1 \dots i_{st}}).$$

For convenience, we will use $i^{[n]}$ to denote a character i repeated n times in the subscripts, i.e., $\underbrace{i \dots i}_n$. Also, we use $T_{i_1 i_2 \dots i_n}$ to denote the tree $T(r_{i_1 i_2 \dots i_n})$ induced by $r_{i_1 i_2 \dots i_n}$.

A tree (T, r, k) is *complete* if and only if all of its vertices of height less than $h(T)$ are of maximum degree k (by abuse of the terminology, we also refer to an induced subtree as complete if all nonleaf vertices except the root are of maximum degree).

2. Introduction. To present our main result, we give more definitions.

DEFINITION 2.1. We are given a tree (T, r, k) with the following properties:

- (1) $\deg(r) = c$.
- (2) The height of any two leaves of T differs by at most 1.
- (3) For any v in T , let $v_1, v_2, \dots \in C_T(v)$, at most one $T(v_i)$ is incomplete, in which case $T(v)$ is incomplete. Then for $c = k, k-1$, and $k-2$, T is called k -good, k -rgood, and semi- k -good, respectively.

Remark. An unrooted tree T is k -good if there exists $r \in V(T)$ such that when T is considered as rooted at r , the above conditions are satisfied.

Intuitively, we get the graphs described above by simply filling the distance levels as long as there are still vertices available. In particular, a single vertex or a star rooted at the center is also considered as a “good” tree. For a *complete k -rgood (semi- k -good) tree* T , the vertices of height less than $h(T)$ are of maximum degree except the root.

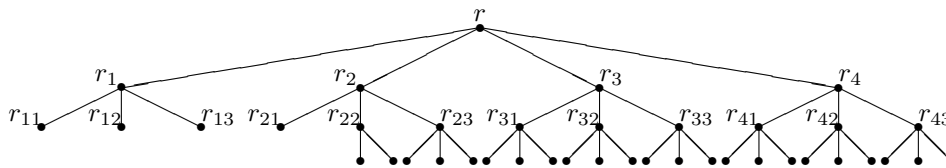


FIG. 1. A 4-good tree on 40 vertices with labels.

Note that k -good, k -rgood, and semi- k -good trees are unique in the following sense: if we have two k -good (k -rgood, semi- k -good) trees with same number of vertices, then we can label their vertices such that they are isomorphic to each other.

The Wiener index is one of the mostly wide used chemical indices in biochemistry. Since most applications under consideration have acyclic molecular structures, the Wiener index of trees has received great attention in the recent years; see [1] for a survey of related research. Although stated differently due to different terminology, Fischermann et al. [2], and independently Jelen and Triesch [4], proved the following theorem.

THEOREM 2.2. Among trees of the same size with maximum vertex degree k , precisely the k -good tree minimizes the Wiener index.

Theorem 2.2 considers trees with a general maximum vertex degree. For binary trees (with vertex degree 1 or 3) with a given number of vertices, the extremal tree that maximizes (minimizes) the number of subtrees is studied in [9]. In a related paper [7], formulas are given to calculate the number of subtrees of these extremal binary

trees. These formulas use a new representation of integers as a sum of powers of 2. Number theorists have already started investigating this new binary representation [3]. Also, the sequence of the number of subtrees of these extremal binary trees (with 2^n leaves, $n = 1, 2, \dots$) appears to be new [6].

An interesting fact is that among binary trees of the same size, the extremal one that minimizes the Wiener index is exactly the one that maximizes the number of subtrees, and vice versa. Recent work of Wagner [10] also implies a strong relation between the Wiener index and the number of subtrees, through a generating function approach. It is then natural to conjecture that for binary trees with a given number of vertices, the larger the Wiener index, the fewer the subtrees. In [7], this phenomenon is further studied, and a counterexample is presented showing that there is no such relation between the Wiener index and the number of subtrees. Yet we lack a better understanding of this relationship.

In [8], a similar question is studied for the number of subtrees that contain at least one leaf of the original tree. It turns out that the same kind of extremal trees, the ones that minimize the Wiener index and maximize the number of subtrees among binary trees of the same size, also maximize the number of leaf-containing subtrees. This was first investigated by Knudsen [5] to provide an upper bound for the time complexity of his multiple parsimony alignment with affine gap cost using a phylogenetic tree. This result on the number of leaf-containing subtrees [8] can be used to provide an upper bound for the number of acceptable residue configurations (the leaf-containing subtrees, in our terms) and helps to decide the time complexity of Knudsen's algorithm; for details see [9].

In this paper, we extend the result in [9] to general trees with a given maximum degree. We will prove the following theorem.

THEOREM 2.3. *Among trees with maximum vertex degree k and a given size, precisely the k -good tree maximizes the number of subtrees.*

Combined with Theorem 2.2, we see that the yet-to-understand relation between the Wiener index and the number of subtrees also holds for general trees, given the maximum vertex degree.

In section 3, some basic lemmas related to the number of subtrees are introduced. Then the properties of k -good, k -r-good, and semi- k -good trees are explored in section 4. We will further characterize the structure of the extremal trees that maximize the number of subtrees in section 5; these lemmas may be interesting on their own. Section 6 completes the proof of Theorem 2.3.

3. Lemmas about arbitrary trees. Two key lemmas are proved in [9]; they can be easily generalized to the following:

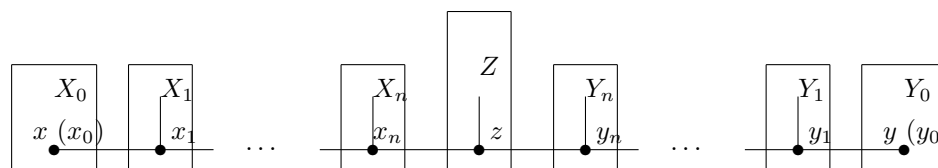


FIG. 2. Path $P_W(x, y)$ connecting vertices x and y .

Consider the tree W in Figure 2, with vertices x and y , and

$$P_W(x, y) = x_0(x)x_1 \dots x_n z y_n \dots y_1 y_0(y)(x_0(x)x_1 \dots x_n y_n \dots y_1 y_0(y))$$

if $d_W(x, y)$ is even (odd) for any $n \geq 0$.

After the deletion of all the edges of $P_W(x, y)$ from W , some connected components will remain. Let X_i (X_0) denote the component that contains x_i ($x_0 = x$), let Y_i (Y_0) denote the component that contains y_i ($y_0 = y$), for $i = 1, 2, \dots, n$, and let Z denote the component that contains z .

LEMMA 3.1. *In the above situation, if $f_{X_i}(x_i) \geq f_{Y_i}(y_i)$ for $i = 0, 1, \dots, n$, then $f_W(x) \geq f_W(y)$. Furthermore, $f_W(x) = f_W(y)$ if and only if $f_{X_i}(x_i) = f_{Y_i}(y_i)$ for all i .*

Idea of the proof. We can write $f_W(x)$ ($f_W(y)$) in terms of $f_{X_i}(x_i)$ and $f_{Y_i}(y_i)$; then the result follows by simple algebra.

A similar approach also leads to the next lemma.

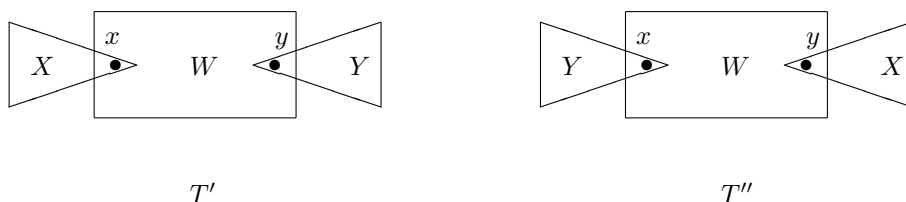


FIG. 3. Switching subtrees rooted at x and y .

If we have a tree W with vertices x and y , and two rooted trees X and Y , then we can build two new trees: first T' by identifying the root of X with x (later will simply be referred to as “attaching X to x ” or “adding $X \setminus x$ to x ” when there is no confusion) and the root of Y with y ; and second T'' by identifying the root of X with y and the root of Y with x (Figure 3). Under the circumstances below we can tell which composite tree has more subtrees.

LEMMA 3.2. *In the above situation, if $f_W(x) > f_W(y)$ and $f_X(x) < f_Y(y)$, then we have $F(T'') > F(T')$.*

Remark. When x, y are leaves, Lemmas 3.1 and 3.2 become the original lemmas in [9].

We also list the following obvious observation for our convenience.

LEMMA 3.3. *Given T rooted at r , if $v \in V(T)$ and $v \neq r$, then $f_T(r) > f_{T(v)}(v)$.*

4. Properties of the “good” trees. A few nice properties follow from the definitions of k -good, k -rgood, and semi- k -good trees; we leave the proofs of them to the reader.

LEMMA 4.1. *If T is a k -good or k -rgood tree, then for any vertex $v \in V(T)$ that is not the root, $T(v)$ is k -rgood.*

LEMMA 4.2. *If T is a k -rgood tree, $v \in V(T)$ is not a leaf, and $v' \in C_T(v)$, then $T(v) \setminus T(v')$ is semi- k -good.*

LEMMA 4.3. *For any two k -rgood or semi- k -good trees T and T' with roots r and r' , respectively, we have*

$$(2) \quad h(T) > h(T') \Rightarrow |V(T)| > |V(T')|,$$

$$(3) \quad |V(T)| \geq |V(T')| \Rightarrow h(T) \geq h(T'),$$

$$(4) \quad f_T(r) > f_{T'}(r') \Leftrightarrow |V(T)| > |V(T')|; f_T(r) = f_{T'}(r') \Leftrightarrow |V(T)| = |V(T')|.$$

We introduce the next lemma for technical reasons.

LEMMA 4.4. *Let two k -good trees T' and T'' be rooted at v' and v'' , respectively, such that $k \leq |V(T')| \leq |V(T'')|$. If T'' is not complete, then*

$$f_{T' \setminus T'_{(k-1)}}(v') < f_{T'' \setminus T''_1}(v'').$$

Proof. Consider $|V(T'_{(k-1)})|$ and $|V(T''_1)|$.

(i) If $|V(T'_{(k-1)})| > |V(T''_1)|$, then $|V(T')| - |V(T'_{(k-1)})| < |V(T'')| - |V(T''_1)|$.

(ii) If $|V(T'_{(k-1)})| \leq |V(T''_1)|$ (note that $|V(T''_1)| < |V(T''_{(k-1)})|$ as T'' is incomplete), then

$$\begin{aligned} |V(T')| - |V(T'_{(k-1)})| &= 1 + \sum_{i=1}^{k-2} |V(T'_i)| \leq 1 + (k-2)|V(T'_{(k-1)})| \\ &\leq 1 + (k-2)|V(T''_1)| < \sum_{i=2}^{k-1} |V(T''_i)| = |V(T'')| - |V(T''_1)|. \end{aligned}$$

Either way, we have

$$|V(T' \setminus T'_{(k-1)})| = |V(T')| - |V(T'_{(k-1)})| < |V(T'')| - |V(T''_1)| = |V(T'' \setminus T''_1)|.$$

The conclusion follows from Lemma 4.2 and (4). $\square$

The next two lemmas follow from the definitions; they provide a better understanding of the structures of “good” trees.

LEMMA 4.5. *Given (T, r, k) , if $\deg(r) = k-1$ and $T(v_i)$ is k -good for any $v_i \in C_T(r)$ ($1 \leq i \leq k-1$), then T is k -good if and only if one of the following conditions holds:*

(1) *For $i = 1, 2, \dots, k-2$, $h(T(v_i)) = h(T(v_{i+1}))$ and $T(v_{i+1})$ is complete.*

(2) *There exists $l \in \{1, \dots, k-1\}$ such that*

for $i = 1, 2, \dots, l-1$, $h(T(v_i)) = h(T) - 2$ and $T(v_i)$ is complete;

for $j = l, \dots, k-1$, $h(T(v_j)) = h(T(v_{j+1})) = h(T) - 1$ and $T(v_{j+1})$ is complete.

LEMMA 4.6. *Let T' and T'' be two k -good trees with $h(T'') - 1 \leq h(T') \leq h(T'')$, and let T be the tree resulting from connecting (with an edge) the roots of T' and T'' .*

If $h(T') = h(T'')$, then T is k -good if and only if one of T' and T'' is complete.

If $h(T') = h(T'') - 1$, then T is k -good if and only if T' is complete.

5. Properties of optimal trees. For convenience, we will call a tree *optimal* if it is a tree of a given size with maximum vertex degree k and it maximizes the number of subtrees. First we show that almost all vertices in an optimal tree have the maximum degree.

LEMMA 5.1. *If (T, r, k) is an optimal tree, then there is at most one nonleaf vertex that has degree not equal to k . Furthermore, if such a vertex exists, all but one of its neighbors must be leaves.*

Proof. First, suppose (for contradiction) that v_1 and v_2 are two nonleaf vertices whose degrees are less than k . Without loss of generality, assume $f_T(v_1) \geq f_T(v_2)$.

Let v_L be a neighbor of v_2 that is not on $P_T(v_1, v_2)$, and T_{v_L} is the component that contains v_L after the removal of the edge $v_2 v_L$ in T . Let T' be the tree resulting from removing a T_{v_L} from T . Notice that $f_T(v_2, v_L) > f_T(v_1, v_2, v_L) = f_T(v_1, v_L)$ as every subtree counted in $f_T(v_1, v_2, v_L)$ is counted in $f_T(v_2, v_L)$, while the subtree induced by the edge $v_2 v_L$ is counted only in $f_T(v_2, v_L)$.

Now we have

$$f_{T'}(v_1) = f_T(v_1) - f_T(v_1, v_L) > f_T(v_2) - f_T(v_L, v_2) = f_{T'}(v_2).$$

Note that, for T' , adding T_{v_L} to v_2 yields T and adding T_{v_L} to v_1 yields a new tree T'' ; by Lemma 3.2, we have $F(T) < F(T'')$ —contradicting the optimality of T .

Thus there is at most one nonleaf vertex that has degree less than k , say v ; suppose (for contradiction) that v_1 and v_2 are neighbors of v that are not leaves. Let T_1 (T_2) denote the component that contains v_1 (v_2) after the removal of the edge vv_1 (vv_2) from T .

Without loss of generality, assume $f_{T_1}(v_1) \geq f_{T_2}(v_2)$. Let T' denote the tree resulting from removing a T_u from T , where u is a neighbor of v_2 that is not on $P_T(v_1, v_2)$, and T_u is the component that contains u after the removal of the edge $v_2 u$ from T . Notice that $f_{T' \setminus T_2}(v) = f_{T \setminus T_2}(v) > f_{T_1}(v_1) \geq f_{T_2}(v_2) > f_{T_2 \setminus T_u}(v_2)$; applying Lemma 3.1 to vv_2 in T' yields that

$$f_{T'}(v) > f_{T'}(v_2).$$

Note that, for T' , adding T_u to v_2 yields T and adding T_u to v yields a new tree T'' ; by Lemma 3.2, we have $F(T) < F(T'')$ —contradicting the optimality of T . $\square$

The reader can refer to Figure 4 for the remaining lemmas in this section.

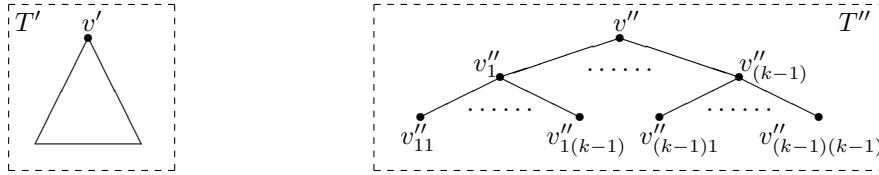


FIG. 4. T' and T'' after the removal of the edge $v'v''$ or the path $P_T(v', v'')$.

We introduce the next two lemmas for technical reasons.

LEMMA 5.2. *Let (T, r, k) be an optimal tree that is divided into two trees (T', v') and (T'', v'') by the removal of the edge $v'v''$. If $T''(v)$ is k -rgood for any $v \in C_{T''}(v'')$ and $f_{T'}(v') \geq f_{T''_{(k-1)1}}(v''_{(k-1)1})$, then, for any $n \geq 1$ such that $v''_{1[n]}$ and $v''_{(k-1)[n+1]}$ exist, we have*

$$(5) \quad |V(T''_{1[n]})| \geq |V(T''_{(k-1)[n+1]})|.$$

Proof. By induction on n , for the base case $n = 1$, we prove the following claim.

CLAIM 5.1. *Under the condition of Lemma 5.2, we must have*

$$(6) \quad |V(T''_i)| \geq |V(T''_{(k-1)j})|$$

for any $i = 1, 2, \dots, k-2$ and $j = 2, 3, \dots, k-1$.

Proof of Claim 5.1. Take the $2(k-2)$ subtrees T''_i for $i = 1, 2, \dots, k-2$ and $T''_{(k-1)j}$ for $j = 2, 3, \dots, k-1$. Put them in a sequence in the descending order of the

number of vertices; we will show that T''_i ($i = 1, 2, \dots, k-2$) must be the first half of this sequence of subtrees (which implies (6)).

Otherwise, suppose $T''_1, T''_2, \dots, T''_s$ are in the second half of the sequence; then we must have $T''_{(k-1)(k-1)}, \dots, T''_{(k-1)(k-s)}$ in the first half. Then

$$|V(T''_i)| > |V(T''_{(k-1)(k-i)})|$$

for $i = s+1, s+2, \dots, k-2$ and

$$|V(T''_j)| < |V(T''_{(k-1)(k-j)})|$$

for $j = 1, 2, \dots, s$.

By the assumption of k -rgoodness and Lemmas 4.2 and 4.3, we have

$$(7) \quad f_{T''_i}(v''_i) > f_{T''_{(k-1)(k-i)}}(v''_{(k-1)(k-i)}) \quad \text{for } i = s+1, s+2, \dots, k-2,$$

$$(8) \quad f_{T''_j}(v''_j) < f_{T''_{(k-1)(k-j)}}(v''_{(k-1)(k-j)}) \quad \text{for } j = 1, 2, \dots, s.$$

We intend to show that this contradicts the optimality of T by applying Lemmas 3.1 and 3.2 in the following setting:

$X = \bigcup_{i=1}^s T''_i \cup \{v''\}; Y = \bigcup_{i=1}^s T''_{(k-1)(k-i)} \cup \{v''_{(k-1)}\}; x = x_0 = v''; y = y_0 = v''_{(k-1)}$.

Consequently we have $W = T \setminus (X \cup Y) \cup \{x, y\}$.

Note that now $X_0 = \bigcup_{i=s+1}^{k-2} T''_i \cup T' \cup \{v''\}$ is a result of joining $T', T''_{s+1}, \dots, T''_{(k-2)}$ with v'' , and $Y_0 = \bigcup_{i=s+1}^{k-2} T''_{(k-1)(k-i)} \cup T''_{(k-1)1} \cup \{v''_{(k-1)}\}$ is a result of joining $T(v''_{(k-1)1}), \dots, T(v''_{(k-1)(k-s-1)})$ with $v''_{(k-1)}$.

By (7) and the assumption that $f_{T'}(v') \geq f_{T''_{(k-1)1}}(v''_{(k-1)1})$, we have

$$f_{X_0}(x_0) = (1 + f_{T'}(v')) \prod_{i=s+1}^{k-2} (1 + f_{T''_i}(v''_i)) > \prod_{i=1}^{k-s-1} (1 + f_{T''_{(k-1)i}}(v''_{(k-1)i})) = f_{Y_0}(y_0).$$

Similarly, note that X is a result of joining $T''_1, \dots, T''_s$ with v'' , and Y is the result of joining $T(v''_{(k-1)(k-s)}), \dots, T(v''_{(k-1)(k-1)})$ with $v''_{(k-1)}$. Equation (8) implies that

$$f_X(x) = \prod_{i=1}^s (1 + f_{T''_i}(v''_i)) < \prod_{i=k-s}^{k-1} (1 + f_{T''_{(k-1)i}}(v''_{(k-1)i})) = f_Y(y).$$

Thus, Lemma 3.1 implies that $f_W(x) > f_W(y)$, but we have $f_X(x) < f_Y(y)$. Lemma 3.2 yields that switching X and Y will increase $F(T)$, in contradiction to the optimality of T .

Hence we must have (6), in particular, $|V(T''_1)| \geq |V(T''_{(k-1)(k-1)})|$.

This completes the proof of Claim 5.1. $\square$

For the inductive step, suppose (5) holds for all values up to n . In particular, $|V(T''_{1[n]})| \geq |V(T''_{(k-1)[n+1]})|$ implies that

$$(9) \quad \sum_{j=1}^{k-1} |V(T''_{1[n]j})| = |V(T''_{1[n]})| - 1 \geq |V(T''_{(k-1)[n+1]})| - 1 = \sum_{j=1}^{k-1} |V(T''_{(k-1)[n+1]j})|.$$

For the sake of contradiction, suppose that the hypothesis does not hold for $n + 1$, yielding

$$(10) \quad |V(T''_{1[n+1]})| < |V(T''_{(k-1)[n+2]})|.$$

Now we have (we superscript some of the inequalities for reference to their proofs)

$$(11) \quad \sum_{i=2}^{k-1} |V(T''_{1[n]_i})| \stackrel{(9)-(10)}{>} \sum_{i=1}^{k-2} |V(T''_{(k-1)[n+1]_i})|.$$

We intend to show that this contradicts the optimality of T by applying Lemma 3.1 and Lemma 3.2 in the following setting:

$$W = T \setminus (X \cup Y) \cup \{x, y\}; X = T(v''_{1[n+1]}); Y = T(v''_{(k-1)[n+2]});$$

$$x = v''_{1[n+1]}; y = v''_{(k-1)[n+2]};$$

$$x_l = v''_{1[n-l+1]}; X_l = T(v''_{1[n-l+1]}) \setminus T(v''_{1[n-l+2]}) \text{ for } l = 1, 2, \dots, n,$$

$$X_{n+1} = T \setminus (T''_{(k-1)} \cup T''_1),$$

$$y_l = v''_{(k-1)[n-l+2]}; Y_l = T(v''_{(k-1)[n-l+2]}) \setminus T(v''_{(k-1)[n-l+3]}) \text{ for } l = 1, 2, \dots, n+1.$$

Note that each X_l and Y_l for $l < n + 1$ is semi- k -good by the assumption of k -rgoodness and Lemma 4.2. Also, we have

$$|V(X_l)| = |V(T''_{1[n-l+1]})| - |V(T''_{1[n-l+2]})| \geq |V(T''_{(k-1)[n-l+2]})| - |V(T''_{(k-1)[n-l+3]})| = |V(Y_l)|$$

for $l = 1, 2, \dots, n$ by induction hypothesis and (1); in addition, strict inequality holds for $l = 1$ by (11).

Hence we have $f_{X_l}(x_l) \geq f_{Y_l}(y_l)$ for $l = 2, \dots, n$ and $f_{X_1}(x_1) > f_{Y_1}(y_1)$ by Lemma 4.3.

Notice that $X_{n+1} = T \setminus (T''_{(k-1)} \cup T''_1)$ is a result of joining $T', T''_2, \dots, T''_{(k-2)}$ with v'' , while $Y_{n+1} = T(v''_{(k-1)}) \setminus T(v''_{(k-1)(k-1)})$ is the result of joining $T(v''_{(k-1)1}), \dots, T(v''_{(k-1)(k-2)})$ with $v''_{(k-1)}$. Compare these “branches”:

$$(12) |V(T''_{(k-1)j})| \stackrel{(1)}{\leq} |V(T''_{(k-1)(k-1)})| \stackrel{(5)}{\leq} |V(T''_1)| \stackrel{(1)}{\leq} |V(T''_j)| \text{ for } j = 2, \dots, k-2.$$

Equation (12) implies, with Lemma 4.3, that $f_{T''_j}(v''_j) \geq f_{T''_{(k-1)j}}(v''_{(k-1)j})$ for $j = 2, \dots, k-2$.

With the assumption that $f_{T'}(v') \geq f_{T''_{(k-1)1}}(v''_{(k-1)1})$, we have

$$f_{X_{n+1}}(x_{n+1}) = (1 + f_{T'}(v')) \prod_{i=2}^{k-2} (1 + f_{T''_i}(v''_i)) \geq \prod_{i=1}^{k-2} (1 + f_{T''_{(k-1)i}}(v''_{(k-1)i})) = f_{Y_{n+1}}(y_{n+1}).$$

Thus we have $f_{X_l}(x_l) \geq f_{Y_l}(y_l)$ for $l = 2, 3, \dots, n+1$ and $f_{X_1}(x_1) > f_{Y_1}(y_1)$, which implies that $f_W(x) > f_W(y)$ by Lemma 3.1. We also have that $f_X(x) < f_Y(y)$ by assumption of k -rgoodness, Lemma 4.3, and (10).

Note that T is the tree resulting from attaching X to x and Y to y , and let T^* be the tree resulting from attaching X to y and Y to x . Now applying Lemma 3.2 to X, Y , and W , we have $F(T^*) > F(T)$, which contradicts the optimality of T , thus proving Lemma 5.2 by induction. $\square$

Let $v'a_1 \dots a_c(z)b_c \dots b_1v''$ denote the path that connects the roots of two disjoint k -rgood trees T' and T'' in an optimal tree T . After the deletion of the edges of $P_T(v', v'')$, some connected components will remain. Let A_i denote the component

that contains a_i and B_i denote the one that contains b_i (and Z the one that contains z).

LEMMA 5.3. *As described above, if $|V(T')| \leq |V(T'')|$, $h(T') = h(T'') = m$, and $f_{A_i}(a_i) \leq f_{B_i}(b_i)$ for $i = 1, 2, \dots, c$, then T'' is complete.*

Proof. Suppose (for contradiction) that T'' is not complete.

CLAIM 5.2. $|V(T'_{(k-1)[n]})| \leq |V(T''_{1[n]})|$ for $n = 1, 2, \dots, m-1$.

Proof of Claim 5.2. By contradiction, let s be the smallest integer such that

$$(13) \quad |V(T'_{(k-1)[s]})| > |V(T''_{1[s]})|.$$

We intend to show that this contradicts the optimality of T by applying Lemmas 3.1 and 3.2 in the following setting:

$$X = T'(v'_{(k-1)[s]}); x = v'_{(k-1)[s]}; Y = T''(v''_{1[s]}); y = v''_{1[s]}.$$

We have $f_X(x) > f_Y(y)$ by (13) and Lemma 4.3.

For $l = 1, 2, \dots, s$,

$$x_l = v'_{(k-1)[s-l]}; X_l = T'_{(k-1)[s-l]} \setminus T'_{(k-1)[s+1-l]}; y_l = v''_{1[s-l]}; Y_l = T''_{1[s-l]} \setminus T''_{1[s+1-l]}.$$

$|V(T'_{(k-1)[s-l]})| \leq |V(T''_{1[s-l]})|$ by hypothesis; then $f_{X_l}(x_l) < f_{Y_l}(y_l)$ by Lemma 4.4.

For $l = s+1, \dots, s+c$,

$$x_l = a_{l-s}; X_l = A_{l-s}; y_l = b_{l-s}; Y_l = B_{l-s}.$$

We have $f_{X_l}(x_l) \leq f_{Y_l}(y_l)$ by assumption.

Let $W = T \setminus (T' \cup T'')$; then $f_W(x) < f_W(y)$ by Lemma 3.1, and thus we can switch X, Y and increase the number of subtrees by Lemma 3.2. This is a contradiction.

This completes the proof of Claim 5.2. $\square$

Now by Claim 5.2, we have

$$(14) \quad |V(T'_{(k-1)[m-1]})| \leq |V(T''_{1[m-1]})|.$$

Notice that $1 \geq h(T''_{1[m-1]}) \geq^{(14)} h(T'_{(k-1)[m-1]}) = h(T') - (m-1) = 1$. By Lemma 5.1, at least one of $v'_{(k-1)[m-1]}$ and $v''_{1[m-1]}$ is of degree k . Then $v''_{1[m-1]}$ must be of degree k by (14); hence T'' is complete. $\square$

Now we show that, after “dividing” an optimal tree into two rooted trees, if one of them is sufficiently “large,” then the other one has to be “good.”

LEMMA 5.4. *For an optimal tree (T, r, k) , assume that T is divided into two rooted subtrees (T', v') and (T'', v'') by the removal of the edge $v'v''$. Then, if for $n \geq 2$,*

$$(15) \quad f_{T'}(v') > f_{T''_{a_1 a_2 \dots a_n}}(v''_{a_1 a_2 \dots a_n}) \quad (1 \leq a_i \leq k-1 \text{ for } i = 1, 2, \dots, n)$$

hold as long as the vertex $v''_{a_1 a_2 \dots a_n}$ exists, then T'' is k -rgood.

Proof. First notice that, by Lemma 5.1, v'' must be of degree k or all of its children are leaves (in which case T'' is k -rgood).

Now we prove the lemma; by induction on $|V(T'')|$, the base case, $|V(T'')| = 1$, is trivial by definition. Now suppose Lemma 5.4 is true for any induced subtree in place of T'' with fewer vertices.

CLAIM 5.3. T''_i is k -rgood for $i = 1, 2, \dots, k-1$.

Proof of Claim 5.3. Consider T as being divided into T''' and T^* , where $T''' = (T''_i, v''_i)$ and $T^* = (T \setminus T''_i, v'')$. Now, for any $1 \leq a_i \leq k-1, i = 1, 2, \dots, n \geq 2$, by using Lemma 3.3, we have

$$f_{T^*}(v'') > f_{T''_i}(v''_i) > f_{T''_{i a_2 \dots a_n}}(v''_{i a_2 \dots a_n}) = f_{(T''''_{a_1 a_2 \dots a_n})}(v''_{i a_1 a_2 \dots a_n});$$

thus (15) holds for T^* and T''' , and then T''_i is k -rgood by hypothesis.

This completes the proof of Claim 5.3. $\square$

Now all induced subtrees of T'' are k -rgood by Lemma 4.1; consider two cases.

Case I. $h(T''_i) = h(T''_{i+1})$ for $i = 1, 2, \dots, k-2$.

In this case, we have, for $i = 1, 2, \dots, k-2$, that

$$h(T''_i) = h(T''_{i+1}) \quad \text{and} \quad |V(T''_i)| \leq^{(1)} |V(T''_{i+1})|.$$

Applying Lemma 5.3 to T''_i and T''_{i+1} yields that T''_{i+1} is complete for $i = 1, \dots, k-2$.

Since each T''_i is of the same height, we have by Lemma 4.5 that T'' is k -rgood.

Case II. There exists c , $1 \leq c \leq k-2$, such that $h(T''_c) < h(T''_{c+1})$.

Then, on the one hand, we must have $h(T''_1) < h(T''_{k-1})$.

On the other hand, by Lemma 5.2 we have

$$|V(T''_1)| \geq |V(T''_{(k-1)(k-1)})| \Rightarrow^{(3)} h(T''_1) \geq h(T''_{(k-1)(k-1)}) = h(T''_{k-1}) - 1.$$

Thus, $h(T''_1) = \dots = h(T''_c) = h(T''_{c+1}) - 1 = \dots = h(T''_{k-1}) - 1$.

Let $m = h(T''_1)$. Lemma 5.2 implies that $|V(T''_{1[m-1]})| \geq |V(T''_{(k-1)[m]})|$, coupled with Lemma 5.1, yields that $v''_{1[m-1]}$ is of degree k , which implies that T''_1 is complete.

Since each T''_i is of the same height, T''_i is complete for $i = 1, 2, \dots, c$ by (1).

Also, for $i = c+1, \dots, k-2$, T''_{i+1} is complete by applying Lemma 5.3 to T''_i and T''_{i+1} .

Thus each T''_i except T''_{c+1} is complete. T'' is k -rgood by Lemma 4.5. $\square$

Next we show that an optimal tree can be “divided” into two trees such that each of them is “large” enough to make the other one “good.”

LEMMA 5.5. *Divide an optimal tree (T, r, k) into two rooted trees (T', v') and (T'', v'') by removing an edge $v'v''$. If $|h(T'') - h(T')| \leq 1$, then both T' and T'' are k -rgood.*

Proof. Note that if we choose a longest path P and choose (v', v'') as the closest to middle edge on P , we obtain such T' and T'' .

Without loss of generality, assume $f_{T''}(v'') \geq f_{T'}(v')$; then

$$f_{T''}(v'') \geq f_{T'}(v') >^{Lemma\ 3.3} f_{T'_{a_1 a_2 \dots a_n}}(v'_{a_1 a_2 \dots a_n}).$$

So (15) holds for T'' and T' , and T' is k -rgood by Lemma 5.4.

Now, since T' is k -rgood, it contains a copy of a complete k -rgood tree (denote by T^* , $h(T^*) = h(T') - 1$) rooted at v' . Note that

$$h(T^*) = h(T') - 1 \geq h(T'') - 2 \geq h(T''_{a_1 a_2 \dots a_n})$$

for any $1 \leq a_i \leq k-1$ for $i = 1, 2, \dots, n \geq 2$. Then

$$f_{T'}(v') >^{(4)} f_{T^*}(v') \geq^{Lemma\ 3.3} f_{T''_{a_1 a_2 \dots a_n}}(v''_{a_1 a_2 \dots a_n}).$$

Thus (15) holds for T' and T'' , and T'' is k -rgood by Lemma 5.4. $\square$

6. Proof of Theorem 2.3.

Proof. Let (T, r, k) be an optimal tree. By Lemma 5.5, we can divide T into two k -rgood trees (T', v') and (T'', v'') with

$$h(T'') - 1 \leq h(T') \leq h(T'').$$

Further assume that $|V(T')| \leq |V(T'')|$. If T' is complete, T is k -good by Lemma 4.6. Now suppose T' is not complete, and consider two cases.

Case I. $h(T') = h(T'')$.

Apply Lemma 5.3 to T' and T'' ; then T'' is complete, and T is k -good by Lemma 4.6.

Case II. $h(T') = h(T'') - 1 = m$.

Note that $h(T''_{(k-1)}) = h(T') = m$, and applying Lemma 5.3 to T' and $T''_{(k-1)}$, together with the assumption that T' is not complete, we have that $T''_{(k-1)}$ must be complete.

Now considering T as being divided by the removal of the edge $v''v''_{(k-1)}$, we have $(T''', v''_{(k-1)}) = T''_{(k-1)}$ of height m being complete and $(T^*, v'') = T \setminus T''_{(k-1)}$ of height $m + 1$. Then T^* is k -good by Lemma 5.5, and T is k -good by Lemma 4.6. $\square$

Acknowledgments. The authors want to thank the referees for carefully reading such a tedious proof and providing many valuable suggestions that greatly improved the quality of this paper. The authors also want to thank the editor/typesetter for their help.

REFERENCES

- [1] A. A. DOBRYNIN, R. ENTRINGER, AND I. GUTMAN, *Wiener index of trees: Theory and applications*, Acta Appl. Math., 66 (2001), pp. 211–249.
- [2] M. FISCHERMANN, A. HOFFMANN, D. RAUTENBACH, L. A. SZÉKELY, AND L. VOLKMANN, *Wiener index versus maximum degree in trees*, Discrete Appl. Math., 122 (2002), pp. 127–137.
- [3] C. HEUBERGER AND H. PRODINGER, *On α -greedy expansions of numbers*, Adv. in Appl. Math., 38 (2007), pp. 505–525.
- [4] F. JELEN AND E. TRIESCH, *Superdominance order and distance of trees with bounded maximum degree*, Discrete Appl. Math., 125 (2003), pp. 225–233.
- [5] B. KNUDSEN, *Optimal multiple parsimony alignment with affine gap cost using a phylogenetic tree*, in Algorithms in Bioinformatics, Third International Workshop, Lect. Notes in Bioinform. 2812, Springer, Berlin, 2003, pp. 433–446.
- [6] THE ON-LINE ENCYCLOPEDIA OF INTEGER SEQUENCES, A092781, <http://www.research.att.com/~njas/sequences>.
- [7] L. A. SZÉKELY AND H. WANG, *On subtrees of trees*, Adv. in Appl. Math., 34 (2005), pp. 138–155.
- [8] L. A. SZÉKELY AND H. WANG, *Binary trees with the largest number of subtrees with at least one leaf*, Congr. Numer., 177 (2005), pp. 147–169.
- [9] L. A. SZÉKELY AND H. WANG, *Binary trees with the largest number of subtrees*, Discrete Appl. Math., 155 (2007), pp. 374–385.
- [10] S. G. WAGNER, *Correlation of graph-theoretical indices*, SIAM J. Discrete Math., 21 (2007), pp. 33–46.

REAL NUMBER CHANNEL ASSIGNMENTS FOR LATTICES*

JERROLD R. GRIGGS[†] AND XIAOHUA TERESA JIN[†]

Abstract. We consider the problem of assigning a numerical channel to each transmitter in a large regular array such that multiple levels of interference, which depend on the distance between transmitters, are avoided by sufficiently separating the channels. The goal is to find assignments that minimize the span of the labels used. A previous paper of the authors introduced a model for this problem using real number labelings of (possibly infinite) graphs G . Given reals $k_1, k_2, \dots, k_p \geq 0$, one denotes by $\lambda(G; k_1, k_2, \dots, k_p)$ the infimum of the spans of the labelings f of the vertices v of G , such that for any two vertices v and w , the difference in their labels is at least k_i , where i is the distance between v and w in G . When $p = 2$, it is enough to determine $\lambda(G; k, 1)$ for reals $k \geq 0$; for G of bounded maximum degree, this will be a continuous, piecewise linear function of k . Here we consider this function for infinite regular lattices that model large planar networks, building on earlier efforts by other researchers. For the triangular lattice, we determine the function for $k \geq 1$, which had previously been found for rational $k \geq 3$ by Calamoneri. We also give bounds for $0 \leq k \leq 1$. For the square lattice and the hexagonal lattice, we completely determine the function for $k \geq 0$, which had been given for rational $k \geq 3$ and $k \geq 2$, respectively, by Calamoneri. Portions of it have been obtained by other researchers for infinite regular lattices that model large planar networks. Here we present the complete function $\lambda(G; k, 1)$ for $k \geq 1$ when G is the triangular, square, or hexagonal lattice.

Key words. channel assignment, generalized graph coloring, planar lattice graphs

AMS subject classifications. 05C78, 05C15, 90B18

DOI. 10.1137/060650982

1. Introduction. Efficient channel assignment algorithms in wireless networks are increasingly important. There is usually a large network of transmitters in the plane, and a numerical channel must be assigned to each transmitter, where channels for nearby vertices must be assigned so as to avoid interference. The goal is to minimize the portion of the frequency spectrum that must be allocated to the problem, so it is desired to minimize the span of a feasible labeling.

In 1980 Hale [22] formulated such channel assignment problems in network engineering as graph labeling problems: each transmitter is represented by a vertex, and any pair of vertices that may interfere is represented by an edge in the graph. All labels are integers.

In 1988 Lanfear proposed to Roberts [33] a new 2-level channel assignment problem of interest to NATO, in which integer labels are assigned to transmitters in the plane, with two levels of interference, depending on the distance between transmitters; e.g., say labels differ by at least two (resp., one) when the transmitters are within some fixed distance A (resp., $2A$). Griggs introduced the analogous “lambda-labeling” problem for graphs and made the initial investigation of this graph theory problem

*Received by the editors January 26, 2006; accepted for publication (in revised form) January 2, 2008; published electronically June 11, 2008. Research was supported in part by NSF grants DMS-0072187 and DMS-0302307. The principal results here were announced in an extended abstract for the International Workshop on Wireless, Mobile, and Ad Hoc Networks in *Proceedings of the 2005 IEEE International Parallel and Distributed Computing Symposium* [16]. This research was also described in part in the second author’s dissertation [23].

<http://www.siam.org/journals/sidma/22-3/65098.html>

[†]Department of Mathematics, University of South Carolina, Columbia, SC 29208 (griggs@math.sc.edu, jin2@math.sc.edu). The research of the first author was supported in part by NSF grant DMS-0072187.

with Yeh [21]. They extended the problem in a natural way by specifying separations $k_1, \dots, k_p$ for vertices at distances $1, \dots, p$, defining an $L(k_1, k_2, \dots, k_p)$ -labeling of a graph G to be an assignment of nonnegative numbers $f(v)$ to the vertices v of G , such that $|f(u) - f(v)| \geq k_i$ if u and v are at distance i in G . The labeling f is said to belong to the set $L(k_1, k_2, \dots, k_p)(G)$. One denotes by $\lambda(G; k_1, k_2, \dots, k_p)$ the minimum span over such f , where the span is the difference between the largest and smallest labels $f(v)$. Griggs and Yeh concentrated on the fundamental case of $L(2, 1)$ -labelings, and many authors have subsequently contributed to the literature on these $L(2, 1)$ and more general labelings. For an online survey and annotated bibliography of work with conditions at distance $p = 2$, see [7]. For an overview of the recent progress and state of the general theory, refer to the recent survey [20].

In applications the frequency channel separations k_i for two transmitters are often taken to be inversely proportional to the physical distance i between them [3]. Most articles assume that the separations are nonincreasing, $k_1 \geq k_2 \geq \dots \geq k_p$. But this is not required in the theory. Since in applications one can in principle use any frequencies (channels) in the available continuous frequency spectrum, not only from a discrete set, Griggs and Jin [17] proposed extending integer graph labelings to allow the labels and separations k_i to be nonnegative *real* numbers. They used the same notation as before, $L(k_1, \dots, k_p)(G)$ and $\lambda(G; k_1, \dots, k_p)$, but now the span of a real labeling is the difference between the supremum and the infimum of the labels used, and λ is the infimum of the spans of such labelings. The authors [19] first explored this concept for simple graphs, such as paths and cycles, and then began to study optimal labelings for the lattices considered in this paper. Their early results led to the discovery of properties for general graphs, which were included in the first, foundational paper [17]. The new insights and tools developed in that project are applied to lattices in this paper. The methods described here for lattices in turn are potentially applicable to other classes of graphs.

For graphs of bounded maximum degree, the authors proved the existence of an optimal labeling of a nice form, in which all labels belong to a discrete set, denoted by $D(k_1, k_2, \dots, k_p)$, of linear combinations $\sum_i a_i k_i$, with nonnegative integer coefficients a_i .

THEOREM 1.1 (the D -Set Theorem [17]). *Let G be a graph, possibly infinite, with finite maximum degree. Let p be a positive integer, and let k_i , $1 \leq i \leq p$, be real numbers ≥ 0 . Then there exists a finite optimal $L(k_1, k_2, \dots, k_p)$ -labeling $f^* : V(G) \rightarrow [0, \infty)$ in which the smallest label is 0 and all labels belong to the set $D(k_1, k_2, \dots, k_p)$. Hence, $\lambda(G; k_1, k_2, \dots, k_p)$ belongs to $D(k_1, k_2, \dots, k_p)$.*

Due to the D -Set Theorem, previous optimal integer labeling results are compatible with new optimal real number labeling results. Some natural properties of distance-constrained labelings become more evident in the setting of real number labelings. In particular, the authors made the following observation, which was evident already in earlier work of Georges and Mauro on integer labelings.

LEMMA 1.2 (Scaling). *For real numbers $d, k_i \geq 0$, $i = 1, 2, \dots, p$,*

$$\lambda(G; d \cdot k_1, d \cdot k_2, \dots, d \cdot k_p) = d \cdot \lambda(G; k_1, k_2, \dots, k_p).$$

More subtle is the result of the authors [17, 23] that $\lambda(G; k_1, k_2, \dots, k_p)$ is a continuous function of the separations k_i for any graph G (possibly infinite) with finite maximum degree. Hence, results about the minimum spans $\lambda(G; k_1, k_2, \dots, k_p)$ for k_i being rational numbers can often be extended into the results for k_i being real numbers. Indeed, by the Scaling Lemma, it is usually enough to obtain results for

integer k_i . The analysis is often more clear when considering real number labelings. For any fixed p and any graph G with finite maximum degree, the authors conjectured [17] that $\lambda(G; k_1, k_2, \dots, k_p)$ is a piecewise linear function of real numbers k_i , where the pieces have nonnegative integer coefficients and where there are only finitely many pieces. The authors proved this if G is finite or if $p = 2$ [17]. Subsequently, Král' proved the full conjecture [25].

For the $p = 2$ case, the Scaling Lemma implies that for $k_2 > 0$, $\lambda(G, k_1, k_2) = k_2 \lambda(G; k, 1)$, where $k = k_1/k_2$. This reduces the two-parameter function to a one-parameter function, $\lambda(G; k, 1)$, $k \geq 0$. As just discussed, it is a continuous, nondecreasing, piecewise linear function with finitely many pieces. Further, each piece has the form $ak + b$ for some nonnegative integers $a, b \geq 0$. This paper concerns the function $\lambda(G; k, 1)$, $k \geq 0$, for the most natural infinite regular planar lattices (also called grid graphs), which are the triangular (6-regular), square (4-regular), and hexagonal (3-regular) lattices. The optimal span function is completely determined in the range of natural application, $k \geq 1$. It is solved as well for $0 \leq k \leq 1$ for the square and hexagonal lattice. For the triangular lattice, the problem appears to be much tougher for $0 \leq k \leq 1$, but here portions of it are solved and bounds are given in the remaining intervals. (See the last section for an update on this work.) Despite their nice properties, the optimal span function turns out to be surprisingly complicated for these three regular lattices.

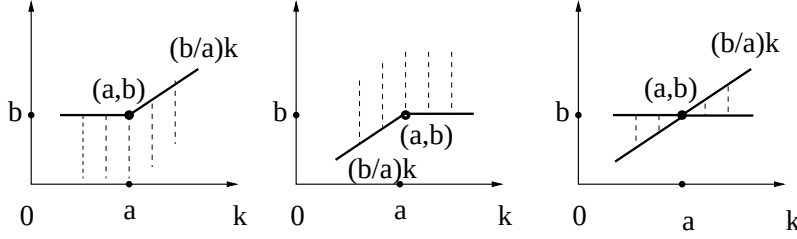
Section 2 introduces some of the general methods used to obtain optimal lattice labelings. It also reviews some of the known results for labeling infinite trees with conditions at distance two, which are closely related to the lattice results. Sections 3, 4, and 5 contain the results for the triangular, square, and hexagonal lattices, respectively. The detailed proofs, which make up most of the paper, are presented in sections 6, 7, and 8, respectively. The paper concludes with ideas for future research directions.

2. Methods. Upper bounds on $\lambda(G; k_1, k_2, \dots, k_p)$ are generally achieved by constructing an efficient labeling, sometimes discovered by computer search. Typically this can be achieved by coordinating the vertices of the lattice, giving an explicit labeling for a small piece, and repeating the pattern, tiling the whole lattice with congruent pieces. Lower bound proofs are generally more difficult. For these we identify crucial particular values of k where we need to prove a lower bound on $\lambda(G; k, 1)$. Such k are rational, say $k = a/b$ for some integers $a, b > 0$. By scaling, it is then equivalent to bound $\lambda(G; a, b)$ below, which has the advantage that we need only consider integer $L(a, b)$ -labelings, which have integer spans. We then seek to prove an integer bound, say $\lambda(G; a, b) \geq c$, by contradiction: If it is not true, then $\lambda(G; a, b) \leq c - 1$, and there must exist a labeling f of G using labels from the set $\{0, 1, \dots, c - 1\}$. We restrict f to an appropriate finite induced subgraph of G , and argue that some label, call it L , must be avoided by f . We continue to eliminate possible labels, until there remains a set of labels for which it can be shown that in fact no feasible labeling exists. In some cases we had to write a computer program to check all possible labelings from a specified label set of a particular induced subgraph.

We begin by recording a simple way to expand the set of avoided labels by using symmetry (which was introduced for the triangular lattice in [4, 14]).

LEMMA 2.1 (Symmetry). *Let S , L , and $k_1, k_2, \dots, k_p$ be nonnegative integers, and let G be a graph. If every $L(k_1, k_2, \dots, k_p)(G)$ -labeling f into $\{0, \dots, S\}$ avoids (resp., uses) label L , then every such labeling f avoids (resp., uses) label $S - L$.*

We next describe a simple method for general graphs G that is surprisingly useful,


 FIG. 1. The bound on $\lambda(G; k, 1)$.

for it permits us to extend a bound at some particular value to general values of k .

LEMMA 2.2. Let a, b be reals with $a > 0$.

If $\lambda(G; a, 1) \leq b$, then

$$\lambda(G; k, 1) \leq \begin{cases} b & \text{if } 0 \leq k \leq a, \\ \frac{b}{a}k & \text{if } k \geq a. \end{cases}$$

If $\lambda(G; a, 1) \geq b$, then

$$\lambda(G; k, 1) \geq \begin{cases} \frac{b}{a}k & \text{if } 0 \leq k \leq a, \\ b & \text{if } k \geq a. \end{cases}$$

In particular, if $\lambda(G; a, 1) = b$, then

- for $0 \leq k \leq a$, $\frac{b}{a}k \leq \lambda(G; k, 1) \leq b$;
- for $k \geq a$, $b \leq \lambda(G; k, 1) \leq \frac{b}{a}k$.

Proof. If $\lambda(G; a, 1) \leq b$, we have the following:

- For $0 \leq k \leq a$, the result follows from the fact that $\lambda(G; k, 1)$ is nondecreasing.
- For $k \geq a$, we also use the scaling lemma to obtain $\lambda(G; k, 1) \leq \lambda(G; k, \frac{k}{a}) = \frac{k}{a} \lambda(G; a, 1) \leq \frac{b}{a}k$.

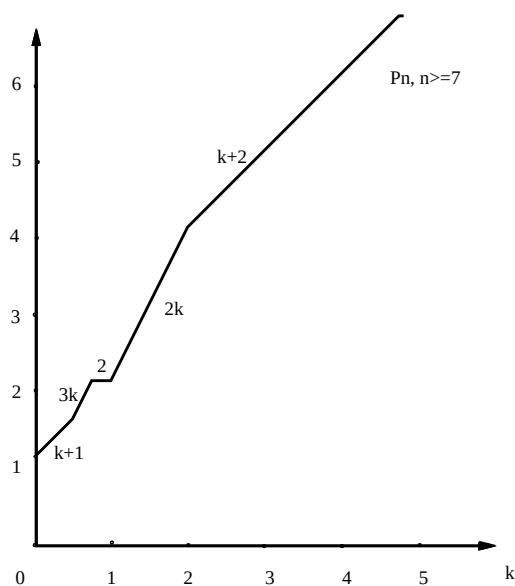
The proof is similar if $\lambda(G; a, 1) \geq b$. $\square$

It is interesting to compare the lattice problems to those for infinite trees. For integer $d > 0$, let T_d denote the tree that is regular of degree d . Note that T_d is infinite for $d \geq 2$ and T_2 is an infinite path. For the path P_n on n vertices, $n \geq 7$, the authors [19] determined the minimum span $\lambda(P_n; k, 1)$, $n \geq 7$ (see Figure 2).

Georges and Mauro [12] obtained the values of $\lambda(T_d; k_1, k_2)$ for integers $k_1 \geq k_2 \geq 0$. In a subsequent paper (with the same title!) Calamoneri, Pelc, and Petreschi [8] gave the values for integers $0 \leq k_1 \leq k_2$. By continuity and scaling, these can be restated in terms of $\lambda(T_d; k, 1)$ for reals $k \geq 0$, which is neater, so we use this format here. For $k \geq 1$ the functions get increasingly complicated as d grows, so we state formulas only for the values required here, $d = 3, 4$.

THEOREM 2.3 (see [12]). For real $k \geq 1$ we have

$$\lambda(T_3; k, 1) = \begin{cases} 3k & \text{if } 1 \leq k \leq \frac{3}{2}, \\ k + 3 & \text{if } \frac{3}{2} < k \leq 2, \\ 2k + 1 & \text{if } 2 \leq k \leq 3, \\ k + 4 & \text{if } k \geq 3. \end{cases}$$

FIG. 2. The minimum span $\lambda(P_n; k, 1)$ for path $P_n, n \geq 7$.

THEOREM 2.4 (see [12]). For real $k \geq 1$, we have

$$\lambda(T_4; k, 1) = \begin{cases} 4k & \text{if } 1 \leq k \leq \frac{4}{3}, \\ k+4 & \text{if } \frac{4}{3} < k \leq \frac{3}{2}, \\ 3k+1 & \text{if } \frac{3}{2} \leq k \leq \frac{5}{3}, \\ 6 & \text{if } \frac{5}{3} \leq k \leq 2, \\ 3k & \text{if } 2 \leq k \leq \frac{5}{2}, \\ k+5 & \text{if } \frac{5}{2} \leq k \leq 3, \\ 2k+2 & \text{if } 3 \leq k \leq 4, \\ k+6 & \text{if } k \geq 4. \end{cases}$$

THEOREM 2.5 (see [8]). For real k , $0 \leq k \leq 1$, and integer $d \geq 2$, we have

$$\lambda(T_d; k, 1) = \begin{cases} k + (d-1) & \text{if } 0 \leq k \leq \frac{1}{2}, \\ (2d-1)k & \text{if } \frac{1}{2} < k \leq \frac{d}{2d-1}, \\ d & \text{if } \frac{d}{2d-1} \leq k \leq 1. \end{cases}$$

Next we give two results that relate the optimal spans of regular trees T_d to that of general d -regular graphs G .

THEOREM 2.6 (see [13]). Let G be a regular graph of degree $d \geq 2$. Then for all real $k \geq 1$, we have $\lambda(G; k, 1) \geq \lambda(T_d; k, 1)$.

The idea of the proof is that T_d is a universal cover of G , so that we can easily define a graph homomorphism h from T_d to G . Then if f is an optimal $L(k, 1)$ -labeling of G , $f \circ h$ is an $L(k, 1)$ -labeling of T_d , so that

$$\lambda(T_d; k, 1) \leq \text{span}(f) \leq \text{span}(f') = \lambda(G; k, 1).$$

The condition $k \geq 1$ above is certainly necessary, since it could be for vertices s and t at distance two that $h(s)$ and $h(t)$ are adjacent, and we would only be certain

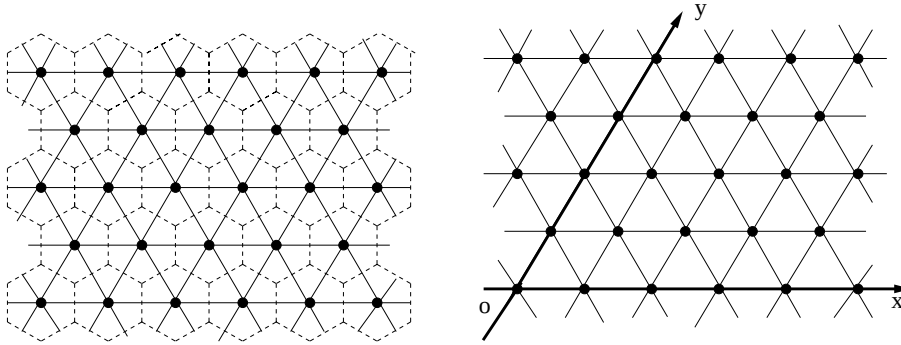


FIG. 3. The hexagonal cell covering and the triangular lattice Γ_Δ .

that $|f(s) - f(t)| \geq k$, which is not strong enough if $k < 1$. For instance, let $k < 1$. If $d = 2$, then T_d is an infinite path, and we may consider the 2-regular graph $G = C_3$. It is easily seen (by examining the two neighbors of a vertex with label 0) that $\lambda(T_2; k, 1) \geq 1 + k$, which exceeds $\lambda(C_3; k, 1) = 2k$.

However, if G is triangle-free, then it cannot be that $h(s)$ and $h(t)$ are adjacent in the problematic case above. We obtain the following result.

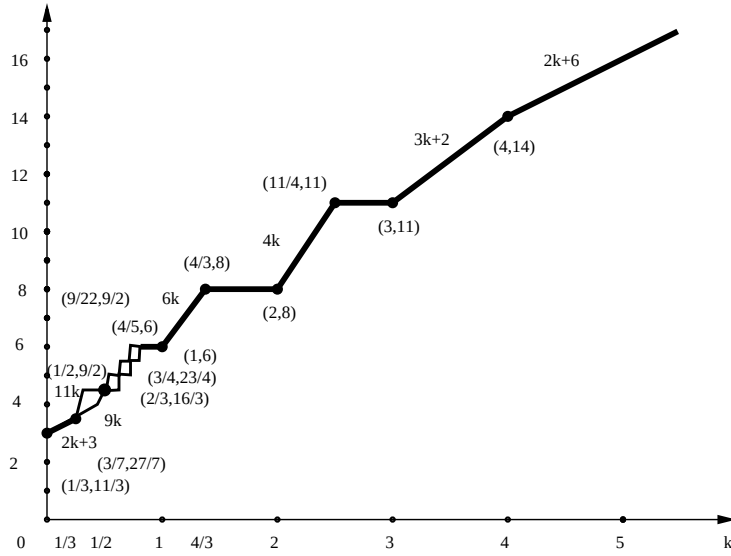
THEOREM 2.7. *Let G be a triangle-free regular graph of degree $d \geq 2$. Then for all real $k \geq 0$, we have $\lambda(G; k, 1) \geq \lambda(T_d; k, 1)$.*

3. The triangular lattice. In a radio mobile network, the large service areas are often covered by a network of nearly congruent polygonal cells, with each transmitter at the center of a cell that it covers. A honeycomb of hexagonal cells provides the most economic covering of the whole plane [11] (i.e., covers the plane with smallest possible transmitter density), where the transmitters are placed in the triangular lattice Γ_Δ (see Figure 3). We fix a point to be the original point o and impose an xy coordinate system so that we can name each point by its xy coordinate.

This problem has some history, owing to the fundamental nature of the triangular lattice for channel assignment problems. Griggs [15] formulated an integer $L(k, 1)$ -labeling problem on the triangular lattice Γ_Δ for the 2000 International Math Contest in Modeling (MCM). Among the 271 teams which worked on this problem for four days and wrote papers, five teams [4, 10, 14, 29, 34] won the contest and got their papers published. As pointed out in the last section, two student teams observed what became the Symmetry Lemma, Lemma 2.1. All winners found $\lambda(\Gamma_\Delta; k, 1)$ for $k = 2, 3$, and some gave labelings for $k = 1$ or for integers $k \geq 4$ that turn out to be optimal, but without proving the lower bound. The team of Goodwin, Johnston, and Marcus [14] proved the optimality for integers $k \geq 4$ (quite an achievement in such a short time) and considered the more general problem of $\lambda(\Gamma_\Delta; k_1, k_2)$ for integers k_1, k_2 . Subsequently, researchers Jin and Yeh [24] and Zhu and Shi [36] took over, each solving some special cases for integers $k_1 \geq k_2$. Calamoneri [6] gave the minimum span for integers $k_1 \geq 3k_2$, and she gave bounds for $k_2 \leq k_1 \leq 3k_2$.

Here we describe the solution of the $L(k, 1)$ -labeling problem for the triangular lattice for real numbers $k \geq 1$, and we give bounds for $0 \leq k \leq 1$ (see Figure 4), where considerable effort has not yet led to a full solution.

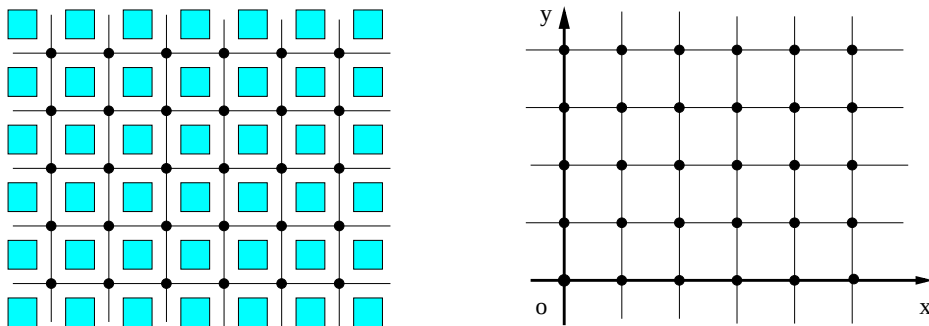
THEOREM 3.1. *For $k \geq 0$ the minimum span of any $L(k, 1)$ -labeling of the*

FIG. 4. $\lambda(\Gamma_{\Delta}; k, 1)$ for $k \geq 0$.

triangular lattice is given by

$$\lambda(\Gamma_{\Delta}; k, 1) \left\{ \begin{array}{ll} = 2k + 3 & \text{if } 0 \leq k \leq \frac{1}{3}, \\ \in [2k + 3, 11k] & \text{if } \frac{1}{3} \leq k \leq \frac{9}{22}, \\ \in [2k + 3, \frac{9}{2}] & \text{if } \frac{9}{22} \leq k \leq \frac{3}{7}, \\ \in [9k, \frac{9}{2}] & \text{if } \frac{3}{7} \leq k \leq \frac{1}{2}, \\ \in [\frac{9}{2}, \frac{16}{3}] & \text{if } \frac{1}{2} \leq k \leq \frac{2}{3}, \\ \in [\frac{16}{3}, \frac{23}{4}] & \text{if } \frac{2}{3} \leq k \leq \frac{3}{4}, \\ \in [\frac{23}{4}, 6] & \text{if } \frac{3}{4} \leq k \leq \frac{4}{5}, \\ = 6 & \text{if } \frac{4}{5} \leq k \leq 1, \\ = 6k & \text{if } 1 \leq k \leq \frac{4}{3}, \\ = 8 & \text{if } \frac{4}{3} \leq k \leq 2, \\ = 4k & \text{if } 2 \leq k \leq \frac{11}{4}, \\ = 11 & \text{if } \frac{11}{4} \leq k \leq 3, \\ = 3k + 2 & \text{if } 3 \leq k \leq 4, \\ = 2k + 6 & \text{if } k \geq 4. \end{array} \right.$$

For the proof of this theorem, go to section 6. We can use Lemma 2.2 to give a slight improvement to the stated bounds in the interval that is not yet resolved, $\frac{1}{3} \leq k \leq \frac{4}{5}$: having the exact values of lambda at $k = \frac{2}{3}, \frac{3}{4}, \frac{4}{5}$ means that there is a linear lower bound for k just below these values, of $8k$ if $k \in [\frac{9}{16}, \frac{2}{3}]$; of $\frac{23k}{3}$ if $k \in [\frac{16}{23}, \frac{3}{4}]$; and of $\frac{15k}{2}$ if $k \in [\frac{23}{30}, \frac{4}{5}]$. Similarly, there is a linear upper bound for k just above these values, of $9k$ if $k \in [\frac{1}{2}, \frac{16}{27}]$; of $8k$ if $k \in [\frac{2}{3}, \frac{23}{32}]$; and of $\frac{23k}{3}$ if $k \in [\frac{3}{4}, \frac{18}{23}]$. We conjecture that the upper bound on $\lambda(\Gamma_{\Delta}; k, 1)$ is the actual value for $\frac{1}{3} \leq k \leq \frac{1}{2}$. For $\frac{1}{2} \leq k \leq \frac{4}{5}$, we conjecture that $\lambda(\Gamma_{\Delta}; k, 1) = 5k + 2$, a formula which works already in this interval at $k = \frac{1}{2}, \frac{2}{3}, \frac{3}{4}$, and $\frac{4}{5}$. Incidentally, we compared the formulas for the triangular lattice (which is 6-regular) to that of the regular infinite tree, T_6 , and found they are quite different and not worth stating explicitly here.


 FIG. 5. A Manhattan network and the square lattice $\Gamma_{\square}$.

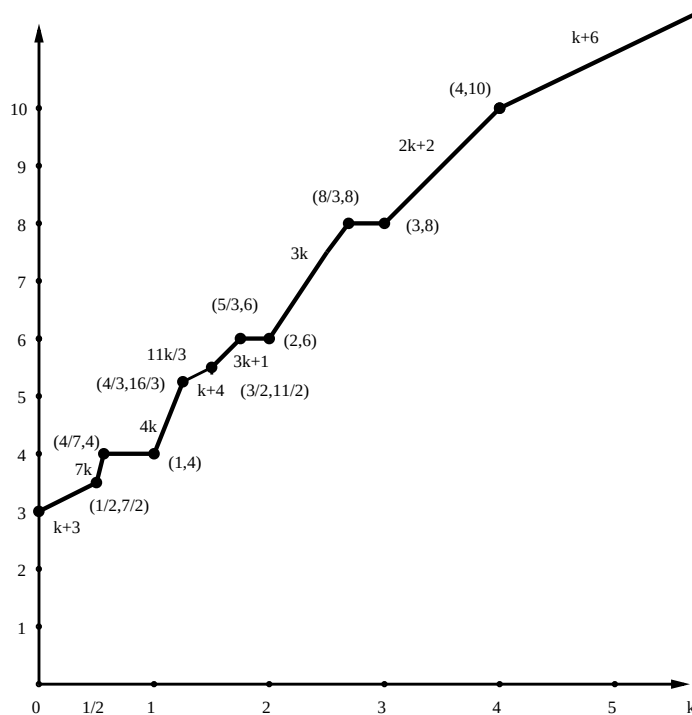
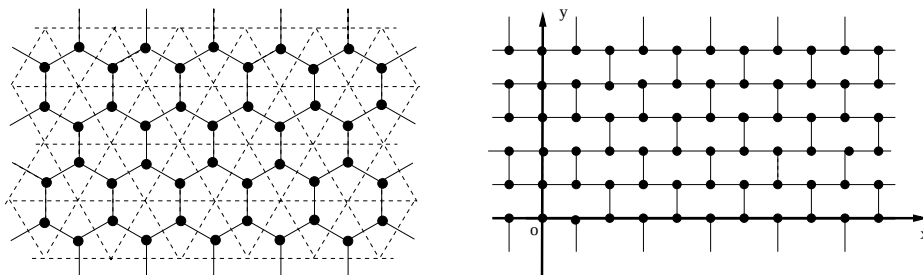
4. The square lattice. In cities, tall buildings can be obstacles in the signal path and limit the range of a cell phone. A Manhattan cellular system [3] can be used that is modeled by the square lattice $\Gamma_{\square}$ (see Figure 5). Many graphs corresponding to cellular systems are the induced subgraphs of the square lattice and the triangular lattice.

Theorem 4.1 is the complete determination of $\lambda(\Gamma_{\square}; k, 1)$ for real numbers $k \geq 0$ (see Figure 6). Previously, Calamoneri [6] independently gave the minimum (integer) span $\lambda(\Gamma_{\square}; k_1, k_2)$ for integers $k_1 \geq 3k_2$, as well as bounds when $k_2 \leq k_1 \leq 3k_2$. (It should be noted that the stated bounds in the earlier extended abstract [5] are not entirely correct, such as the claim that $\lambda(\Gamma_{\square}; 3, 2) = 12$, which is contradicted by the $L(3, 2)$ -labeling of span only 11 from [35]. However, the bounds in the subsequent preprint [6] appear to be correct.)

THEOREM 4.1. *For $k \geq 0$ the minimum span of any $L(k, 1)$ -labeling of the square lattice is given by*

$$\lambda(\Gamma_{\square}; k, 1) = \begin{cases} k + 3 & \text{if } 0 \leq k \leq \frac{1}{2}, \\ 7k & \text{if } \frac{1}{2} < k \leq \frac{4}{7}, \\ 4 & \text{if } \frac{4}{7} \leq k < 1, \\ 4k & \text{if } 1 \leq k \leq \frac{4}{3}, \\ k + 4 & \text{if } \frac{4}{3} < k \leq \frac{3}{2}, \\ 3k + 1 & \text{if } \frac{3}{2} < k \leq \frac{5}{3}, \\ 6 & \text{if } \frac{5}{3} \leq k \leq 2, \\ 3k & \text{if } 2 < k \leq \frac{8}{3}, \\ 8 & \text{if } \frac{8}{3} \leq k \leq 3, \\ 2k + 2 & \text{if } 3 \leq k \leq 4, \\ k + 6 & \text{if } k \geq 4. \end{cases}$$

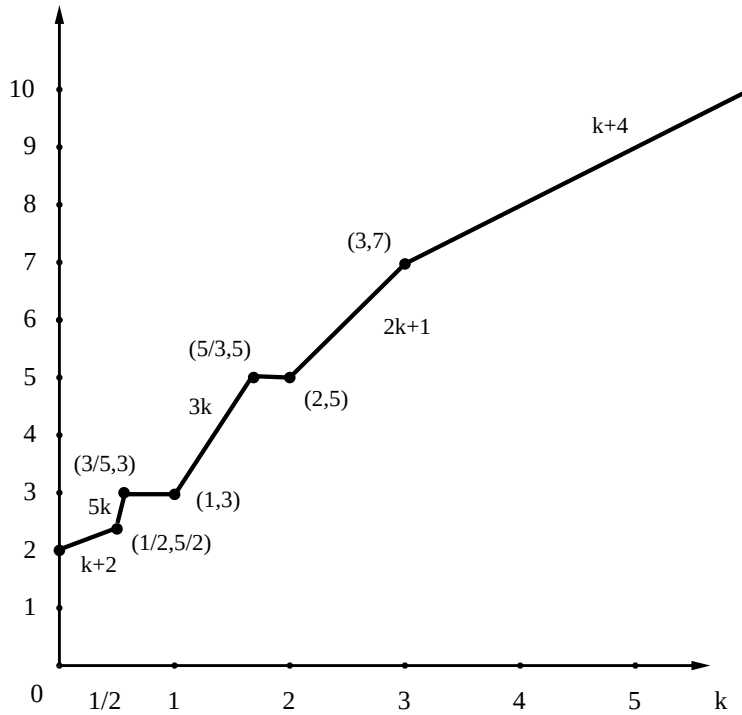
This theorem, which is proven in section 7, allows us to answer a question posed by Georges and Mauro (private communication): does $\lambda(\Gamma_{\square}; k, 1)$ agree with $\lambda(T_4; k, 1)$ for all $k \geq 0$, as stated in Theorems 2.4 and 2.5? Since $\Gamma_{\square}$ is a triangle-free regular graph of degree 4, Theorem 2.7 is applicable and tells us that $\lambda(\Gamma_{\square}; k, 1) \geq \lambda(T_4; k, 1)$ for all $k \geq 0$. Indeed, they almost always agree. However, there is one interval in which they differ: when $\frac{5}{2} < k < 3$, $\lambda(\Gamma_{\square}; k, 1) = \min\{3k, 8\}$ is strictly larger than $\lambda(T_4; k, 1) = k + 5$.

FIG. 6. The minimum span $\lambda(\Gamma_{\square}; k, 1)$.FIG. 7. The equilateral triangle cell covering and the hexagonal lattice Γ_H .

5. The hexagonal lattice. Another interesting fundamental planar array is the hexagonal lattice Γ_H (see Figure 7), which is the dual of the triangular lattice. We are not aware of its being used in real life for wireless networks, but it is mentioned in the engineering literature.

We designate a point o to be the origin, and we impose an xoy coordinate system so that we can name each point by its xoy coordinate, where (i, j) are vertices (see Figure 7). Vertices (i, j) and $(i + 1, j)$ are always adjacent, while vertices (i, j) and $(i, j + 1)$ are adjacent if and only if $i \equiv j \pmod{2}$. Calamoneri [6] gives the minimum span for the hexagonal lattice for integers $k_1 \geq 2k_2$ and bounds for $k_2 \leq k_1 \leq 2k_2$. Here we resolve the remaining open cases and completely determine the span $\lambda(\Gamma_H; k, 1)$ for real numbers $k \geq 0$ (see Figure 8).

THEOREM 5.1. For $k \geq 0$ the minimum span of any $L(k, 1)$ -labeling of the


 FIG. 8. The minimum span $\lambda(\Gamma_H; k, 1)$ for $k \geq 0$.

hexagonal lattice is given by

$$\lambda(\Gamma_H; k, 1) = \begin{cases} k+2 & \text{if } 0 \leq k \leq \frac{1}{2}, \\ 5k & \text{if } \frac{1}{2} \leq k \leq \frac{3}{5}, \\ 3 & \text{if } \frac{3}{5} \leq k \leq 1, \\ 3k & \text{if } 1 \leq k \leq \frac{5}{3}, \\ 5 & \text{if } \frac{5}{3} \leq k \leq 2, \\ 2k+1 & \text{if } 2 \leq k \leq 3, \\ k+4 & \text{if } k \geq 3. \end{cases}$$

The proof is given in section 8. We may now compare the spans of the hexagonal lattice and the regular tree of the same degree, T_3 . As before, the fact that Γ_H is triangle-free allows us to apply Theorem 2.7 to see that $\lambda(\Gamma_H; k, 1) \geq \lambda(T_3; k, 1)$ for all $k \geq 0$. Comparing the formula above for Γ_H to those from Theorems 2.3 and 2.5, we see that Γ_H agrees with T_3 except in the range $\frac{3}{2} < k < 2$, where $\lambda(\Gamma_H; k, 1) = \min\{3k, 5\}$ is strictly larger than $\lambda(T_3; k, 1) = k + 3$.

6. The proof of the Triangular Lattice Theorem (Theorem 3.1). Generally, we get upper bounds by constructing feasible labelings. For lower bounds we derive contradictions on induced subgraphs for labelings of smaller span. Especially, we denote by B_7 (resp., B_{19} , B_{37}) the induced subgraph of Γ_Δ on all 7 (resp., 19, 37) vertices at distance at most one (resp., two, three) from a fixed vertex. Lemma 2.2 is particularly useful for obtaining lower and upper bounds.

To find an upper bound on $\lambda(\Gamma_\Delta; k, 1)$, one construction method is to tile the whole lattice by a labeled parallelogram described by a matrix of labels. We de-

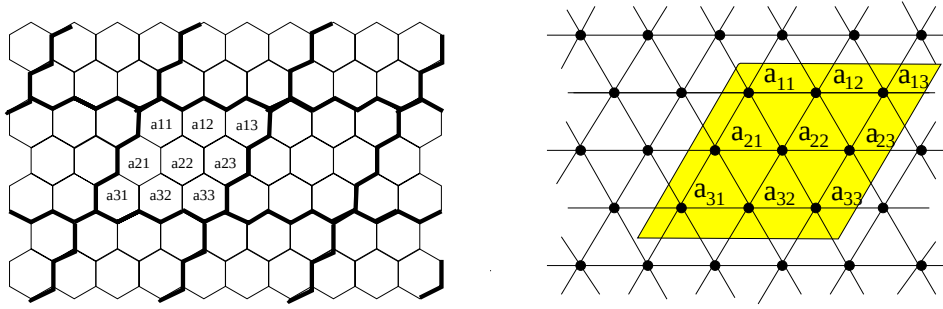


FIG. 9. The matrix labeling.

fine a doubly periodic labeling of the triangular lattice by an $m \times n$ labeling matrix $A := [a_{i,j}]$, such that we label point (i, j) by $a_{m-(j \bmod m), (i \bmod n)+1}$, where i, j are integers. For example, the following labeling (see Figure 9) is defined by the labeling matrix A :

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Then Figure 9 shows how the labels are assigned, where $a_{3,1}$ is at the vertex with coordinates $(0, 0)$ in the triangular lattice. The whole lattice is tiled with copies of the 3×3 tile as shown.

A special case of matrix labeling is defined simply by “arithmetic progressions”: For positive integers k_1, k_2 , we construct a labeling $f \in L(k_1, k_2)$ by taking $f(i, j) = (ai + bj) \bmod l$ for positive integers a, b, l . When such f is feasible, we obtain $\lambda(\Gamma_\Delta; k_1, k_2) \leq l - 1$. Some labelings of this kind were given for the triangular and square lattices in [35]. We found some new arithmetic progression labelings by computer search. We begin our constructions at $k = 0$.

PROPOSITION 6.1. *For $0 \leq k \leq \frac{1}{3}$, we have $\lambda(\Gamma_\Delta; k, 1) \leq 2k + 3$. For $\frac{1}{3} \leq k \leq \frac{9}{22}$, we have $\lambda(\Gamma_\Delta; k, 1) \leq 11k$.*

Proof. We get the upper bound $\lambda(\Gamma_\Delta; k, 1) \leq 2k + 3$ for $0 \leq k \leq \frac{1}{3}$ by defining the labeling matrix

$$A = \begin{bmatrix} k+1 & 2k & 0 & k+3 & 2k+2 & 2 \\ 2k+1 & 1 & k & 2k+3 & 3 & k+2 \\ 0 & k+3 & 2k+2 & 2 & k+1 & 2k \\ k & 2k+3 & 3 & k+2 & 2k+1 & 1 \\ 2k+2 & 2 & k+1 & 2k & 0 & k+3 \\ 3 & k+2 & 2k+1 & 1 & k & 2k+3 \end{bmatrix}.$$

In particular, $\lambda(\Gamma_\Delta; \frac{1}{3}, 1) \leq \frac{11}{3}$, and Lemma 2.2 implies that $\lambda(\Gamma_\Delta; k, 1) \leq 11k$ for $k \geq \frac{1}{3}$. $\square$

Next, we can improve upon the $11k$ upper bound for k between $\frac{9}{22}$ and $\frac{1}{2}$.

PROPOSITION 6.2. *For $\frac{9}{22} \leq k \leq \frac{1}{2}$, we have $\lambda(\Gamma_\Delta; k, 1) \leq \frac{9}{2}$.*

Proof. The upper bound $\lambda(1, 2) \leq 9$ is given in [24] by an arithmetic progression labeling in $L(1, 2)$: Label point (i, j) by $(i + 4j) \bmod 10$. By scaling, this gives the bound on $\lambda(\Gamma_\Delta; \frac{1}{2}, 1)$, which then extends to $k \leq \frac{1}{2}$ by Lemma 2.2. $\square$

The upper bound for k between $\frac{1}{2}$ and $\frac{3}{4}$ follows from the bounds at $k = \frac{2}{3}$ and $\frac{3}{4}$ by the fact that $\lambda(\Gamma_\Delta; k, 1)$ is nondecreasing (Lemma 2.2).

PROPOSITION 6.3. 1. We have $\lambda(\Gamma_\Delta; 2, 3) \leq 16$. Hence, $\lambda(\Gamma_\Delta; \frac{2}{3}, 1) \leq \frac{16}{3}$.

2. We have $\lambda(\Gamma_\Delta; 3, 4) \leq 23$. Hence, $\lambda(\Gamma_\Delta; \frac{3}{4}, 1) \leq \frac{23}{4}$.

Proof. By computer search of arithmetic progression labelings, we discovered $f_1 \in L(2, 3)(\Gamma_\Delta)$ given by $f_1(i, j) = (2i + 7j) \bmod 17$ and $f_2 \in L(3, 4)(\Gamma_\Delta)$ given by $f_2(i, j) = (3i + 10j) \bmod 24$. Hence, $\lambda(\Gamma_\Delta; 2, 3) \leq 16$ and $\lambda(\Gamma_\Delta; 3, 4) \leq 23$. $\square$

Next we extend the upper bound out to $k = \frac{4}{3}$ by applying Lemma 2.2 with the upper bound on $\lambda(\Gamma_\Delta; 1, 1)$. Note that the upper bounds we are giving here for $k = \frac{1}{2}, \frac{2}{3}, \frac{3}{4}$, and 1 are matched by the lower bounds, and so give the correct values of $\lambda(\Gamma_\Delta; k, 1)$ for these k .

PROPOSITION 6.4 (see [4, 14]). We have $\lambda(\Gamma_\Delta; 1, 1) = 6$. Hence,

$$\lambda(G; k, 1) \leq \begin{cases} 6 & \text{if } \frac{3}{4} \leq k \leq 1, \\ 6k & \text{if } 1 \leq k \leq \frac{4}{3}. \end{cases}$$

Proof. We get the upper bound, $\lambda(B_7; 1, 1) \leq 6$, from the arithmetic progression labeling $f(i, j) = (i + 3j) \bmod 7$. The rest follows from Lemma 2.2. $\square$

We now use the construction of numerous MCM teams at $k = 2$ to extend our upper bound out to $k = \frac{11}{4}$, and another construction out to $k = 4$.

PROPOSITION 6.5 (see [4, 10, 14, 29, 34]). We have $\lambda(\Gamma_\Delta; 2, 1) \leq 8$. Hence,

$$\lambda(G; k, 1) \leq \begin{cases} 8 & \text{if } \frac{4}{3} \leq k \leq 2, \\ 4k & \text{if } 2 \leq k \leq \frac{11}{4}. \end{cases}$$

Proof. Label point (i, j) by $(2i + 5j) \bmod 9$. $\square$

PROPOSITION 6.6 (see [14, 29]). For $3 \leq k \leq 4$, we have $\lambda(\Gamma_\Delta; k, 1) \leq 3k + 2$.

Proof. We get the bound by defining the labeling matrix

$$A = \begin{bmatrix} 3k & 0 & k & 2k \\ 1 & k+1 & 2k+1 & 3k+1 \\ k+2 & 2k+2 & 3k+2 & 2 \end{bmatrix}. \quad \square$$

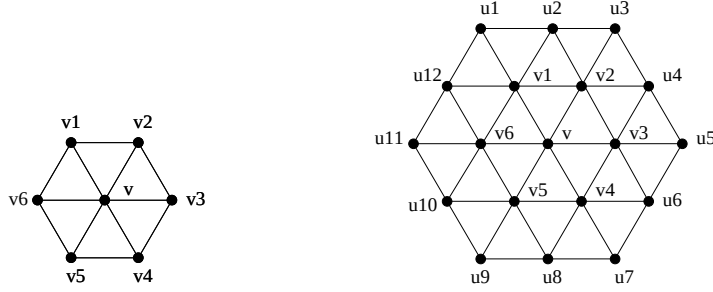
Using $\lambda(\Gamma_\Delta; 3, 1) \leq 11$, this bound of 11 extends down to $k = \frac{11}{4}$ by Lemma 2.2. A construction from the winning MCM papers takes care of all large k .

PROPOSITION 6.7 (see [4, 10, 14, 29, 34]). For $k \geq 4$, we have $\lambda(\Gamma_\Delta; k, 1) \leq 2k + 6$.

Proof. We get the labeling from the matrix

$$A = \begin{bmatrix} 2k+5 & 0 & k+4 \\ 1 & k+2 & 2k+6 \\ k+3 & 2k+4 & 2 \end{bmatrix}. \quad \square$$

We verify the lower bounds using proofs by contradiction (which can be rather complicated) and Lemma 2.2. We shall postpone the small values, $k \leq \frac{3}{4}$. We demonstrate two main methods of proof. The first method, for integers k_1, k_2 , involves the successive elimination of possible labels, until a contradiction is reached. This method was used in the contest paper of Goodwin, Johnston, and Marcus [14] to handle the case of integers $k \geq 4$ (see our comments before Proposition 6.11). We also drew ideas from [36] for the proof of the following important case.

FIG. 10. The subgraphs B_7 and B_{19} of the triangular lattice.

PROPOSITION 6.8. We have $\lambda(\Gamma_\Delta; 4, 3) \geq 24$. Hence,

$$\lambda(\Gamma_\Delta; k, 1) = \begin{cases} 6k & \text{if } \frac{3}{4} \leq k \leq \frac{4}{3}, \\ 8 & \text{if } \frac{4}{3} \leq k \leq 2. \end{cases}$$

Proof. The first statement implies the second by Lemma 2.2. It suffices to prove that $\lambda(\Gamma_\Delta; 4, 3) \geq 24$. Assume to the contrary that there exists a labeling $f \in L(4, 3)(\Gamma_\Delta)$ with its labels in $\{0, 1, \dots, 23\}$. The series of claims that follows restricts the labels f one can use until we find that no such f can exist at all, proving the proposition.

CLAIM 1. The labeling f cannot use label 3 or 20.

Proof. Assume f uses label 3 at v . By the separation conditions, the six labels around v belong to $\{7, 8, \dots, 23\}$, and the difference between any pair of them is at least 3.

Among all 49 possible labelings of B_7 with central label 0 by symmetry, we found by computer that there are just five feasible labelings of subgraph B_{19} that use 3 at the center (B_7 , B_{19} are shown in Figure 10), and none of these can be extended to B_{37} . Full details are in [23]. By the Symmetry Lemma, Lemma 2.1, f is also excluded from using the complementary label $23 - 3 = 20$, and the claim follows.

CLAIM 2. The labeling f cannot use label 7 or 16.

Proof. Assume f uses label 7 at $v \in V(\Gamma_\Delta)$. Denote the six labels around v by $x_1 < x_2 < \dots < x_6$. By the separation conditions, $x_{i+1} \geq x_i + 3$ for $i = 1, 2, \dots, 5$, and each $x_i \in \{0, 1, 2, 11, 12, \dots, 19, 21, 22, 23\}$ (recall we cannot use 3 or 20). Then even if $x_1 \leq 2$, we must have $x_2 \geq 11$, $x_3 \geq 14$, $x_4 \geq 17$, $x_5 \geq 21$, $x_6 \geq 24$, a contradiction proving the claim.

Now f has no label 3, 7, 16, 20. The proofs of Claims 3, 4, and 5 are similar to the proof of Claim 2, so we omit the details.

CLAIM 3. The labeling f cannot use label 6 or 17.

CLAIM 4. The labeling f cannot use label 10 or 13.

CLAIM 5. The labeling f cannot use label 11 or 12.

Now the set of all possible labels is $\{0, 1, 2, 4, 5, 8, 9, 14, 15, 18, 19, 21, 22, 23\}$. We cannot find seven distinct labels, such that the difference between any two of them is at least 3. So we cannot label B_7 , which is a contradiction. Thus, $\lambda(\Gamma_\Delta; 4, 3) \geq 24$. $\square$

By similar proofs, we have the following bounds (see [23] for full details).

PROPOSITION 6.9. 1. We have $\lambda(\Gamma_\Delta; 11, 4) \geq 44$. Hence,

$$\lambda(\Gamma_\Delta; k, 1) \geq \begin{cases} 4k & \text{if } 2 \leq k \leq \frac{11}{4}, \\ 11 & \text{if } \frac{11}{4} \leq k \leq 3. \end{cases}$$

2. We have $\lambda(\Gamma_\Delta; 1, 2) \geq 9$. Hence,

$$\lambda(\Gamma_\Delta; k, 1) \geq \begin{cases} 9k & \text{if } \frac{3}{7} \leq k \leq \frac{1}{2}, \\ \frac{9}{2} & \text{if } k \geq \frac{1}{2}. \end{cases}$$

3. We have $\lambda(\Gamma_\Delta; 2, 3) \geq 16$. Hence, $\lambda(\Gamma_\Delta; k, 1) \geq \frac{16}{3}$ for $k \geq \frac{2}{3}$.

4. We have $\lambda(\Gamma_\Delta; 3, 4) \geq 23$. Hence, $\lambda(\Gamma_\Delta; x, 1) \geq \frac{21}{4}$ for $k \geq \frac{3}{4}$.

5. We have $\lambda(\Gamma_\Delta; 4, 5) \geq 30$. Hence, $\lambda(\Gamma_\Delta; x, 1) \geq 6$ for $k \geq \frac{4}{5}$.

The next result takes care of all k in the interval $(3, 4)$. It can be derived by continuity and scaling from the corresponding result by Calamoneri [6] for integer labelings that give $\lambda(\Gamma_\Delta; k_1, k_2)$ for integers k_1, k_2 with $3k_2 \leq k_1 \leq 4k_2$. Her lower bound method involves looking at a small induced subgraph of the lattice and checking cases according to the numerical order of the labels. This is similar to the method devised independently by Georges and Mauro for labeling trees [12]. We discovered the result independently (but waited on the rest of this project before writing it up here). Because our proof illustrates a different method with some potential for future value, we include it here. It involves the successive removal of intervals of possible labels until there is a contradiction.

PROPOSITION 6.10. For $3 < k < 4$, we have $\lambda(\Gamma_\Delta; k, 1) = 3k + 2$.

Proof. The upper bound comes from Proposition 6.6. We prove the lower bound by contradiction: Assume $\lambda(\Gamma_\Delta; k, 1) = l < 3k + 2$. By the D -Set Theorem, there is an optimal labeling $f \in L(k, 1)(\Gamma_\Delta)$ with $\text{span}(f) = l < 3k + 2$ and $f(u) = 0$ for some $u \in V(\Gamma_\Delta)$.

CLAIM 1. The labeling f cannot use labels in $[k - 1, k) \cup (l - k, l - k + 1]$.

Proof. Assume $f(v) \in [k - 1, k)$ for some $v \in V(\Gamma_\Delta)$. The neighbors of v induce a C_6 subgraph, and their labels are all $\geq f(v) + k$. Hence, $\lambda(\Gamma_\Delta; k, 1) \geq f(v) + k + \lambda(C_6; k, 1) \geq (k - 1) + k + (k + 3) = 3k + 2$ (because $\lambda(C_6; k, 1) = k + 3$ for $k \geq 3$; see [23]). This gives a contradiction. Thus $f(v) \notin [k - 1, k)$ for all $v \in V(\Gamma_\Delta)$. By symmetry of the labels, $f(v) \notin (l - k, l - k + 1]$ for all $v \in V(\Gamma_\Delta)$, which proves the claim.

Now, $f(v) \in I_1 \cup I_2 \cup I_3$ for all $v \in V(\Gamma_\Delta)$, where $I_1 = [0, k - 1)$, $I_2 = [k, l - k]$, $I_3 = (l - k + 1, l]$. Then $|I_1| = k - 1 < k$, $|I_2| = l - 2k < k + 2$, $|I_3| = k - 1 < k$.

CLAIM 2. The labeling f cannot use labels in $[k, k + 1) \cup (l - k - 1, l - k]$.

Proof. Assume $f(v) \in [k, k + 1)$ for some $v \in V(\Gamma_\Delta)$. Among the six distinct labels around v , at most one label is in $I_1 = [0, k - 1)$ (because this label is $\leq f(v) - k < 1$), at most two labels are in $I_2 = [k, l - k]$ (because these two labels are $\geq f(v) + k \geq 2k$ and $||2k, l - k|| = l - 3k < 2$), and at most three labels are in $I_3 = (l - k + 1, l]$ (because $|I_3| < k$, these labels cannot be adjacent). Thus, one label is in I_1 , two labels are in I_2 , and three labels are in I_3 . The three labels in I_3 are for vertices that aren't adjacent. The smallest of the three labels then must be next to at least one of the labels in I_2 . This smallest label in I_3 is $\leq l - 2 < 3k$. But the two labels in I_2 are $\geq f(v) + k \geq 2k$, and the smallest label in I_3 , being next to one of these, must then be at least $3k$, a contradiction.

Thus, $f(v) \notin [k, k + 1)$ for all $v \in V(\Gamma_\Delta)$. By symmetry of the labels, $f(v) \notin (l - k - 1, l - k]$, proving the claim.

Now, $f(v) \in I_1 \cup I_2' \cup I_3 = [0, k - 1) \cup [k + 1, l - k - 1] \cup (l - k + 1, l]$, where $I_2' = [k + 1, l - k - 1]$. Then $|I_2'| \leq l - 2k - 2 < k$.

CLAIM 3. The labeling f cannot use labels in $[k + 1, k + 2) \cup (l - k - 2, l - k - 1]$.

Proof. Assume $f(v) \in [k + 1, k + 2)$ for some $v \in V(\Gamma_\Delta)$. Among the six labels around v , at most two labels are in $I_1 = [0, k - 1)$ (because these two labels are

$\leq f(v) - k < 2$), no label is in $I'_2 = [k + 1, l - k - 1]$ (because if it exists, it would be $\geq f(v) + k \geq 2k + 1 > l - k - 1$, a contradiction), and at most three labels are in $I_3 = (l - 2, l]$. We cannot label all six vertices.

Thus $f(v) \notin [k + 1, k + 2]$ for all $v \in V(\Gamma_\Delta)$. By symmetry of the labels, $f(v) \notin (l - k - 2, l - k - 1]$, proving the claim.

Now, $f(v) \in I_1 \cup I''_2 \cup I_3 = [0, k - 1] \cup [k + 2, l - k - 2] \cup (l - k + 1, l]$ for all $v \in V(\Gamma_\Delta)$, where $I''_2 = [k + 2, l - k - 2]$. Then $|I''_2| = l - 2k - 4 < k - 2 < 2$ for $k < 4$.

Since $f(u) = 0$, among the six distinct labels around u (the difference between any pair of them is at least 1), no label is in I_1 , at most two labels are in I''_2 (because $|I''_2| < 2$), and at most three labels are in $I_3 = (l - 2, l]$ (because $|I_3| < k$ means no two of its labels are for adjacent vertices). So we cannot label all six neighbors of u , a contradiction. $\square$

We next address $k \geq 4$. The modeling team of Goodwin, Johnston, and Marcus [14] obtained the correct values for the integer cases, that is, for integers $k \geq 4$. It is a pity that, due to space limitations, the elegant proof in their contest paper was omitted from the published version! It is the same method we used to prove Proposition 6.8 above.

Moreover, Goodwin et al. gave what is equivalent to the correct formula, $\lambda(\Gamma_\Delta; k_1, k_2) = 2k_1 + 6k_2$, for arbitrary integers k_1, k_2 with $k_1 > 6k_2 + 1$. By scaling and continuity, this implies $\lambda(\Gamma_\Delta; k, 1) = 2k + 6$ for all real $k \geq 6$. There appear to be some technical errors in their lower bound proof (quite understandable, since they had just four days to produce their entire paper from scratch!). However, we discovered that if one uses the D -Set Theorem, some small changes will fix their proof. We present below our own verification of the lower bound, which we need more generally for all real $k \geq 4$. We follow this with the much shorter proof for $k \geq 6$ based on the method of Goodwin et al. that does not depend on the structure of the triangular lattice, so that it can be used on other graphs for sufficiently large real k , provided that there is a linear bound for all large integers k .

PROPOSITION 6.11. *For $k \geq 4$ we have $\lambda(\Gamma_\Delta; k, 1) \geq 2k + 6$.*

Proof. Assume for contradiction that $\lambda(\Gamma_\Delta; k, 1) = l < 2k + 6$ for some $k \geq 4$. By the D -Set Theorem, there is an optimal labeling $f \in L(k, 1)(\Gamma_\Delta)$ with span and largest label l and smallest label 0.

CLAIM 1. *The labeling f cannot use labels in $[3, k]$.*

Proof. If some $f(v) \in [3, k]$, then the labels on the vertices of the C_6 neighboring v are all at least $f(v) + k$. The largest of these labels is then at least $f(v) + k + \lambda(C_6; k, 1) \geq 3 + k + (k + 3) = 2k + 6 > l$, a contradiction since $\lambda(C_6; k, 1) = k + 3$ [23, 19]), proving the claim.

By symmetry, none of the labels in f belongs to $(l - k, l - 3]$. So all labels belong to the union $I_1 \cup I_2 \cup I_3$, where $I_1 = [0, 3]$, $I_2 = [k, l - k]$, and $I_3 = (l - 3, l]$.

CLAIM 2. *The labeling f cannot use labels in $[k, k + 1]$.*

Proof. Assume some label $f(v) \in [k, k + 1]$. At most one of the six vertices next to v has a label in I_1 because any such label is $\leq f(v) - k < 1$. At most three of the six vertices have labels in I_3 as any two must be at least one apart.

First suppose that three of these labels are in I_3 . They cannot be at adjacent vertices, so suppose they are at vertices v_1, v_3 , and v_5 , with reference to the graph B_7 in Figure 10. Two of the other labels next to v must belong to $[f(v) + k, l - k]$, so the larger of the two, say it is at v_2 , must be at least $f(v) + k + 1 \geq 2k + 1$. Then both $f(v_1)$ and $f(v_3)$ are at least $f(v_2) + k$, and the larger of the two is at least $f(v_2) + k + 1 \geq 3k + 2 \geq 2k + 6 > l$, a contradiction with $k \geq 4$.

Next suppose just two of these labels next to v lie in I_3 . The two vertices are not adjacent. There must be at least three labels next to v in $[f(v) + k, l - k]$, and, because this interval has length $< k$, no two of the three are adjacent—say they are at v_1, v_3, v_5 . The largest of the three labels is at least $f(v) + k + 2$, and its neighbor with label in I_3 has label at least $f(v) + k + 2 + k \geq 3k + 2$, which is again a contradiction.

Finally, suppose at most one label next to v lies in I_3 . Then at least four labels next to v are in $[f(v) + k, l - k]$, so some two are adjacent—but this is impossible since they must differ by at least k (as $(l - k) - (f(v) + k) \leq l - 3k < 2 < k$). This proves the claim.

Hence, f has no labels in $[k, k + 1)$ or, by symmetry, in $(l - k - 1, l - k]$. So all of its labels belong to $I_1 \cup I'_2 \cup I_3$, where here $I'_2 = [k + 1, l - k - 1]$.

CLAIM 3. *The labeling f cannot use labels in $[k + 1, k + 2)$.*

Proof. Suppose some $f(v) \in [k + 1, k + 2)$. Then labels used next to v in I_1 are at most $f(v) - k < 2$, so there can be at most two such labels. On the other hand, at most three labels next to v can come from I_3 . Then some label used next to v lies in I'_2 . But such a label must be at most $l - k - 1$ and at least $f(v) + k \geq 2k + 1 \geq k + 5 > l - k - 1$, a contradiction proving the claim.

By symmetry, no label of f belongs to $(l - k - 2, l - k - 1]$. Then all of its labels belong to $I_1 \cup I''_2 \cup I_3$, where $I''_2 = [k + 2, l - k - 2]$. Let u be a vertex with $f(u) = 0$. Then its six neighbors all have labels in $I''_2 \cup I_3$. But I_3 can contain at most three of the labels, as they must be at least one apart from each other. So some three of the labels are in I''_2 . However, $(l - k - 2) - (k + 2) = l - 2k - 4 < 2$, so I''_2 can contain at most two of the labels, a contradiction, so no such f exists. $\square$

Here is the shorter proof of the restriction of the proposition above to $k \geq 6$.

PROPOSITION 6.12. *For $k \geq 6$ we have $\lambda(\Gamma_\Delta; k, 1) \geq 2k + 6$.*

Proof. Let us assume the result of Goodwin et al. that $\lambda(\Gamma_\Delta; k, 1) \geq 2k + 6$ for integers $k \geq 4$. Now consider any noninteger $k > 6$. Let $m = \lceil k \rceil - k$, so that $m \in (0, 1)$ and $k + m = \lceil k \rceil \geq 7$. Hence, $\lambda(\Gamma_\Delta; k + m, 1) \geq 2k + 2m + 6$.

Assume for contradiction that $\lambda(\Gamma_\Delta; k, 1) < 2k + 6$. Let f be an optimal labeling in $L(k, 1)(\Gamma_\Delta)$ as in the D -Set Theorem, with minimum value 0 at some vertex u and maximum value $\text{span}(f)$ at some vertex w . Define a labeling f_1 by $f_1(v) = f(v) + m \lfloor f(v)/k \rfloor$. We can check that $f_1 \in L(k + m, 1)(\Gamma_\Delta)$. Further, the minimum value of f_1 is 0, which occurs at u , and its maximum occurs at w , which thus has value $f_1(w) = \text{span}(f_1) < (2k + 6) + m \lfloor (2k + 6)/k \rfloor = 2k + 6 + 2m$ (since $k > 6$ by assumption). This contradicts the lower bound in the previous paragraph. $\square$

We cannot see how to extend the argument in the last proof to work for k between 4 and 6. It remains to do the lower bound for small k . Similar to the proof of Proposition 6.10, we can show the following (see [23]).

PROPOSITION 6.13. *For $0 < k \leq \frac{1}{2}$, we have $\lambda(\Gamma_\Delta; k, 1) \geq 2k + 3$. Hence $\lambda(\Gamma_\Delta; k, 1) = 2k + 3$ for $0 \leq k \leq \frac{3}{7}$.*

This completes the proof of Theorem 3.1.

7. The proof of the Square Lattice Theorem (Theorem 4.1). We begin by establishing the claimed upper bounds on $\lambda(\Gamma_\square; k, 1)$ for reals $k \geq 0$. In many cases, there is an explicit construction based on a modular construction, in which a particular matrix of labels is used for a rectangle of lattice points and then repeated over and over. We define a doubly periodic labeling of the square lattice by an $m \times n$ labeling matrix $A := [a_{i,j}]$, such that we label point (i, j) by $a_{m-(j \bmod m), (i \bmod n)+1}$, where i, j are integers.

PROPOSITION 7.1. *For $0 \leq k \leq \frac{1}{2}$, we have $\lambda(\Gamma_\square; k, 1) \leq k + 3$.*

Proof. Starting from an optimal $L(0, 1)$ -labeling and shifting up some labels by k , in order to satisfy the $L(k, 1)$ conditions, we constructed with the following labeling matrix that attains the upper bound:

$$A = \begin{bmatrix} 0 & k & 1 & k+1 \\ k+3 & 2 & k+2 & 3 \\ 1 & k+1 & 0 & k \\ k+2 & 3 & k+3 & 2 \end{bmatrix}. \quad \square$$

Next, applying the result above at $k = \frac{1}{2}$, Lemma 2.2 yields this upper bound for larger k .

PROPOSITION 7.2. *For $\frac{1}{2} \leq k \leq \frac{4}{7}$, we have $\lambda(\Gamma_{\square}; k, 1) \leq 7k$.*

Van den Heuvel, Leese, and Shepherd [35] gave a circular integer labeling result which is helpful for our real number labelings, as it suggests some arithmetic progression labelings that turn out to be optimal for our problem.

PROPOSITION 7.3. *We have*

$$\begin{aligned} \lambda(\Gamma_{\square}; 1, 1) &= 4, \\ \lambda(\Gamma_{\square}; 2, 1) &\leq 6, \\ \lambda(\Gamma_{\square}; 3, 1) &\leq 8, \\ \lambda(\Gamma_{\square}; 3, 2) &\leq 11. \end{aligned}$$

Proof. From [35], we have these labelings:

$\lambda(\Gamma_{\square}; 1, 1) \leq 4$ by labeling f with $f(i, j) = (i + 2j) \bmod 5$.

$\lambda(\Gamma_{\square}; 2, 1) \leq 6$ by labeling f with $f(i, j) = (2i + 3j) \bmod 7$.

$\lambda(\Gamma_{\square}; 3, 1) \leq 8$ by labeling f with $f(i, j) = (3i + 4j) \bmod 9$.

$\lambda(\Gamma_{\square}; 3, 2) \leq 11$ by labeling f with $f(i, j) = (3i + 5j) \bmod 12$.

It is easy to show $\lambda(\Gamma_{\square}; 1, 1) \geq 4$. $\square$

Applying the preceding two propositions and Lemma 2.2, we have the following upper bounds.

PROPOSITION 7.4. *We have*

$$\lambda(\Gamma_{\square}; k, 1) \leq \begin{cases} 4 & \text{if } \frac{4}{7} \leq k \leq 1, \\ 4k & \text{if } 1 \leq k \leq \frac{5}{3}, \\ 6 & \text{if } \frac{5}{3} \leq k \leq 2, \\ 3k & \text{if } 2 \leq k \leq \frac{8}{3}, \\ 8 & \text{if } \frac{8}{3} \leq k \leq 3. \end{cases}$$

The upper bounds in the proposition above are weaker than the known lower bounds for $\frac{4}{3} < k < \frac{5}{3}$. Let us consider one value in this gap, $k = \frac{11}{8}$. By Proposition 7.4, we get the upper bound $\lambda(\Gamma_{\square}; \frac{11}{8}, 1) \leq \frac{11}{2}$, so by the Scaling Lemma, $\lambda(\Gamma_{\square}; 11, 8) \leq 44$. To determine whether this is the best possible, we searched for a better labeling: We managed to construct an $L(11, 8)$ -labeling based on a matrix A in which the entries are elements of the D -set in $[0, 43]$. Since 43 can be expressed in terms of 11 and 8 in just one way, $43 = 11 + 4 \times 8$, we were able to extend this matrix labeling to cases in the range $\frac{4}{3} \leq k \leq \frac{3}{2}$, as given in the following proposition. We next took the resulting labeling at $k = \frac{3}{2}$ and found a way to extend it to the range $\frac{3}{2} \leq k \leq \frac{5}{3}$ in a way that maintains the order of the labels, while expanding

their pairwise differences, to maintain feasibility as k grows, giving Proposition 7.6. Notice that these formulas for $\lambda(\Gamma_{\square}; k, 1)$ around $k = 11/8$ are not of the simple form ck for some c , so we could not simply apply Lemma 2.2.

PROPOSITION 7.5. *For $\frac{4}{3} \leq k \leq \frac{3}{2}$, we have $\lambda(\Gamma_{\square}; k, 1) \leq k + 4$.*

Proof. The upper bound is attained by the following labeling matrix:

$$\begin{bmatrix} 5 & k & 4 & 0 & 3 & k+4 & 2 & k+3 & 1 & k+2 & 0 & k+1 \\ 0 & 3 & k+4 & 2 & k+3 & 1 & k+2 & 0 & k+1 & 5 & k & 4 \\ 2 & k+3 & 1 & k+2 & 0 & k+1 & 5 & k & 4 & 0 & 3 & k+4 \\ k+2 & 0 & k+1 & 5 & k & 4 & 0 & 3 & k+4 & 2 & k+3 & 1 \end{bmatrix}. \quad \square$$

PROPOSITION 7.6. *For $\frac{3}{2} \leq k \leq \frac{5}{3}$, we have $\lambda(\Gamma_{\square}; k, 1) \leq 3k + 1$.*

Proof. The upper bound is attained by the following labeling matrix:

$$\begin{bmatrix} 2k+2 & k & 2k+1 & 0 & 2k & 3k+1 & 2 & 3k & 1 & k+2 & 0 & k+1 \\ 0 & 2k & 3k+1 & 2 & 3k & 1 & k+2 & 0 & k+1 & 2k+2 & k & 2k+1 \\ 2 & 3k & 1 & k+2 & 0 & k+1 & 2k+2 & k & 2k+1 & 0 & 2k & 3k+1 \\ k+2 & 0 & k+1 & 2k+2 & k & 2k+1 & 0 & 2k & 3k+1 & 2 & 3k & 1 \end{bmatrix}. \quad \square$$

For larger k , we first adapt the construction given by Calamoneri for integers k_1, k_2 with $3k_2 \leq k_1 \leq 4k_2$. We then present a simple matrix $L(k, 1)$ -labeling that is optimal for all $k \geq 4$.

PROPOSITION 7.7. *For $3 \leq k \leq 4$ we have $\lambda(\Gamma_{\square}; k, 1) \leq 2k + 2$.*

Proof. Adapting the construction in [6], the upper bound is attained by the $L(k, 1)$ -labeling matrix:

$$A = \begin{bmatrix} 2k+2 & k & 2k+1 & 2 & 2k & 1 & k+2 & 0 & k+1 \\ k+2 & 0 & k+1 & 2k+2 & k & 2k+1 & 2 & 2k & 1 \\ 2 & 2k & 1 & k+2 & 0 & k+1 & 2k+2 & k & 2k+1 \end{bmatrix}. \quad \square$$

PROPOSITION 7.8. *For $k \geq 0$ we have $\lambda(\Gamma_{\square}; k, 1) \leq k + 6$.*

Proof. The upper bound is attained by the following labeling matrix:

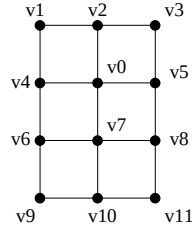
$$A = \begin{bmatrix} 0 & k+3 & 1 & k+4 \\ k+6 & 2 & k+5 & 3 \\ 1 & k+4 & 0 & k+3 \\ k+5 & 3 & k+6 & 2 \end{bmatrix}. \quad \square$$

We now consider the lower bounds to complete the proof of the formulas. It is helpful to compare our graph to T_4 , the regular infinite tree of degree 4 discussed in section 2. By Theorem 2.7 we get that for all $k \geq 0$, $\lambda(\Gamma_{\square}; k, 1) \geq \lambda(T_4; k, 1)$. From the values of $\lambda(T_4; k, 1)$ presented in Theorems 2.4 and 2.5, we obtain the claimed values of $\lambda(\Gamma_{\square}; k, 1)$ for all k outside the interval $[\frac{5}{2}, 3]$. In this remaining interval, we must improve the lower bound on $\lambda(\Gamma_{\square}; k, 1)$. In view of Lemma 2.2, all that remains to prove the theorem is to establish the lower bound at $k = \frac{8}{3}$.

PROPOSITION 7.9. *We have $\lambda(\Gamma_{\square}; 8, 3) \geq 24$. Consequently, for $2 \leq k \leq 3$, we have*

$$\lambda(\Gamma_{\square}; k, 1) \geq \begin{cases} 3k & \text{if } 2 \leq k \leq \frac{8}{3}, \\ 8 & \text{if } \frac{8}{3} \leq k \leq 3. \end{cases}$$

Proof. The second statement follows from the first by Lemma 2.2.

FIG. 11. The subgraph B_{12} of the square lattice.

Assume for contradiction that the first statement fails. Then there exists a labeling $f \in L(8, 3)(\Gamma_{\square})$ with all labels in $\{0, \dots, 23\}$. The series of claims that follows restricts the labels f one can use until we find that no such f can exist at all, proving the proposition.

Let $v_0 = (i_0, j_0) \in V(\Gamma_{\square})$. Let B_{12} be the induced subgraph as in Figure 11.

CLAIM 1. *The labeling f cannot use label 7 or 16.*

Proof. Assume $f(v_0) = 16$. Since no label can exceed 23, the four distinct labels around v_0 are each $\leq f(v_0) - 8 = 8$, which is impossible since any two must be at least 3 apart. By the Symmetry Lemma, labeling f does not use the complementary label $23 - 16 = 7$, and the claim follows.

CLAIM 2. *The labeling f cannot use label 8 or 15.*

Proof. Assume some $f(v_0) = 8$. The four labels around v_0 are each $\geq f(v_0) + 8 = 16$ or $\leq f(v_0) - 8 = 0$, and hence are 0 or ≥ 17 (because by Claim 1, f cannot use 16). Suppose they are labels $x < y < z < w$. Since the difference between any pair of the four labels is ≥ 3 , it must be that $x = 0, y = 17, z = 20, w = 23$. Suppose without loss of generality that $f(v_7) = y = 17$. Since $f(v_7) + 8 = 25$ is too large, it must be that the neighboring labels $f(v_6), f(v_8), f(v_{10})$ are all $\leq f(v_7) - 8 = 9$, and hence all $\leq f(v_0) - 3 = 8 - 3 = 5$. But this is impossible since the difference between any pair of the three labels must be at least 3. By symmetry, we can also exclude 15. This proves the claim.

Now f has no label 7, 8, 15, 16. The proofs of Claims 3 and 4 are similar to the proof of Claim 2, so we omit the details.

CLAIM 3. *The labeling f cannot use label 9 or 14.*

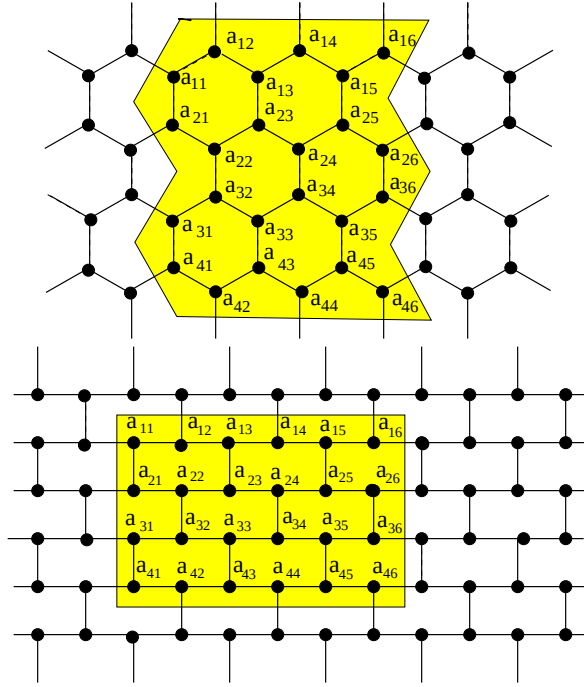
CLAIM 4. *The labeling f cannot use label 11 or 12.*

By the D -Set Theorem, there exists optimal labeling $f^* \in L(8, 3)(\Gamma_{\square})$ with smallest label 0 and all labels in $D_{8,3} \cap [0, 23] = \{0, 3, 6, 8, 9, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23\}$. Applying the claims above to f^* , we find that $f^*(v) \in \{0, 3, 6, 17, 18, 19, 20, 21, 22, 23\}$ for all $v \in V(\Gamma_{\square})$.

Let $f(v_0) = 0$. The four labels around v_0 are each $\geq f(v_0) + 8 = 8$. Their labels belong to $\{17, 18, 19, 20, 21, 22, 23\}$, a contradiction since the difference between any pair of them is ≥ 3 . Thus, it must be that $\lambda(\Gamma_{\square}; 8, 3) \geq 24$. $\square$

This completes the proof of the formulas for the square lattice, Theorem 4.1.

8. The proof of the Hexagonal Lattice Theorem (Theorem 5.1). We will find the upper bound on $\lambda(\Gamma_H; k, 1)$, $k \geq 0$, by constructions and Lemma 2.2. One construction method is to tile the whole lattice by a labeled parallelogram described by a matrix of labels. We define a doubly periodic labeling of the hexagonal lattice by an $m \times n$ labeling matrix $A := [a_{i,j}]$ for m, n even, such that we label point (i, j) by $a_{m-(j \bmod m), (i \bmod n)+1}$, where i, j are integers. For example, the following labeling


 FIG. 12. The doubly periodic labeling by matrix A .

(see Figure 12) is defined by the labeling matrix A , where

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{bmatrix}.$$

Then Figure 12 shows how the labels are assigned, where $a_{4,1}$ is at the vertex with coordinates $(0, 0)$ in the hexagonal lattice. The whole lattice is tiled with copies of the 4×6 tile as shown.

PROPOSITION 8.1. For $0 \leq k \leq \frac{1}{2}$, we have $\lambda(\Gamma_H; k, 1) \leq k + 2$.

Proof. We use the labeling matrix below, also shown in Figure 13, with the values a, b, c taken to be $k, k + 1, k + 2$, respectively:

$$A = \begin{bmatrix} 0 & a & 1 & b & 2 & c \\ b & 2 & c & 0 & a & 1 \end{bmatrix}.$$

Incidentally, this labeling was obtained by doing a first-fit labeling on one row, then on the next row, and so on. $\square$

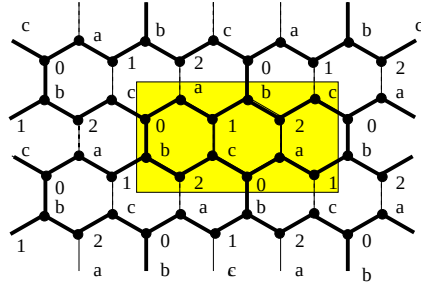
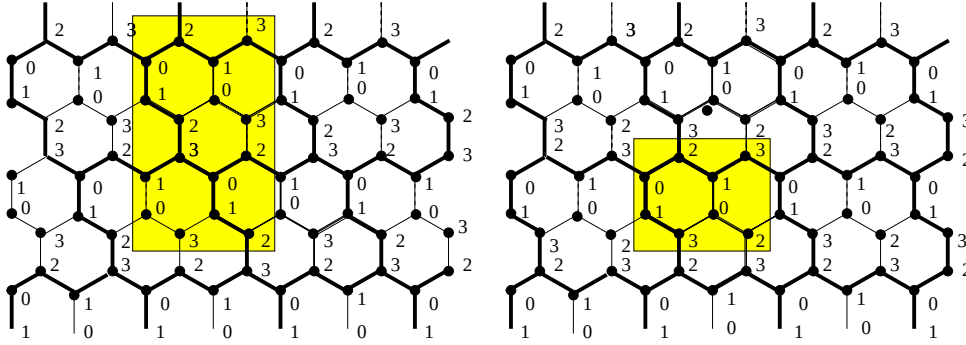
We have $\lambda(\Gamma_H; \frac{1}{2}, 1) \leq \frac{5}{2}$. By Lemma 2.2, the following proposition holds.

PROPOSITION 8.2. For $\frac{1}{2} \leq k \leq \frac{3}{5}$, we have $\lambda(\Gamma_H; k, 1) \leq 5k$.

Next we consider $k = 1$.

PROPOSITION 8.3. We have $\lambda(\Gamma_H; 1, 1) \leq 3$. Hence,

$$\lambda(\Gamma_H; k, 1) \leq \begin{cases} 3 & \text{if } \frac{3}{5} \leq k \leq 1, \\ 3k & \text{if } 1 \leq k \leq \frac{5}{3}. \end{cases}$$

FIG. 13. Optimal $L(k, 1)$ -labeling of Γ_H for $0 \leq k \leq \frac{1}{2}$ or $k \geq 3$.FIG. 14. Optimal $L(1, 1)$ -labeling of Γ_H .

Proof. Because of Lemma 2.2, it is enough to prove the upper bound at $k = 1$.

We will prove $\lambda(\Gamma_H; 1, 1) \leq 3$ by using either of the following labeling matrices. Each was obtained by a first-fit labeling process, doing one row at a time (see Figure 14).

$$A = \begin{bmatrix} 0 & 2 & 1 & 3 \\ 1 & 2 & 0 & 3 \\ 1 & 3 & 0 & 2 \\ 0 & 3 & 1 & 2 \end{bmatrix} \quad \text{or} \quad A = \begin{bmatrix} 0 & 2 & 1 & 3 \\ 1 & 3 & 0 & 2 \end{bmatrix}. \quad \square$$

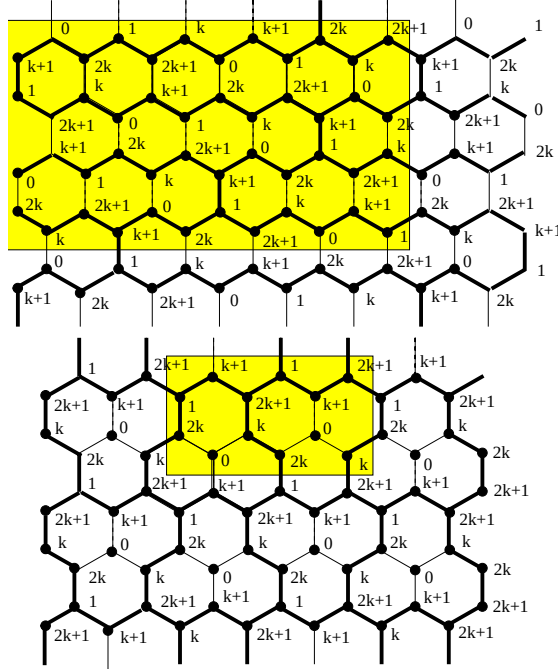
PROPOSITION 8.4. For $2 \leq k \leq 3$, we have $\lambda(\Gamma_H; k, 1) \leq 2k + 1$. For $\frac{5}{3} \leq k \leq 2$, we have $\lambda(\Gamma_H; k, 1) \leq 5$.

Proof. The second statement follows immediately from the first at $k = 2$. For $2 \leq k \leq 3$, one can prove $\lambda(\Gamma_H; k, 1) \leq 2k + 1$ by the matrix labeling with entries shown in Figure 15 (top). A simpler construction can be obtained by adapting a construction of Calamoneri [6], originally given for the corresponding integer labeling. This gives the following matrix labeling, shown in Figure 15 (bottom):

$$A = \begin{bmatrix} 1 & k+1 & 2k+1 & 1 & k+1 & 2k+1 \\ 2k & 0 & k & 2k & 0 & k \end{bmatrix}. \quad \square$$

Next we treat large k .

PROPOSITION 8.5. For $k \geq 3$, we have $\lambda(\Gamma_H; k, 1) \leq k + 4$.


 FIG. 15. Optimal $L(k, 1)$ -labelings of Γ_H for $2 \leq k \leq 3$.

Proof. Following the construction in [6] for the corresponding integer labeling, we again have the matrix labeling as in Figure 13, where this time $a = k + 4, b = k + 3, c = k + 2$:

$$A = \begin{bmatrix} 0 & a & 1 & b & 2 & c \\ b & 2 & c & 0 & a & 1 \end{bmatrix}. \quad \square$$

We next verify the lower bounds. By Theorem 2.7 we get that for all $k \geq 0$, $\lambda(\Gamma_{\square}; k, 1) \geq \lambda(T_3; k, 1)$. From the values of $\lambda(T_3; k, 1)$ presented in Theorems 2.3 and 2.5, we obtain the claimed values of $\lambda(\Gamma_{\square}; k, 1)$ for all k outside the interval $(\frac{3}{2}, 2)$. In this remaining interval, we must improve the lower bound on $\lambda(\Gamma_{\square}; k, 1)$. In view of Lemma 2.2, all that is needed to complete the proof is to establish the lower bound at $k = 5/3$.

PROPOSITION 8.6. *We have $\lambda(\Gamma_H; 5, 3) \geq 15$. Hence,*

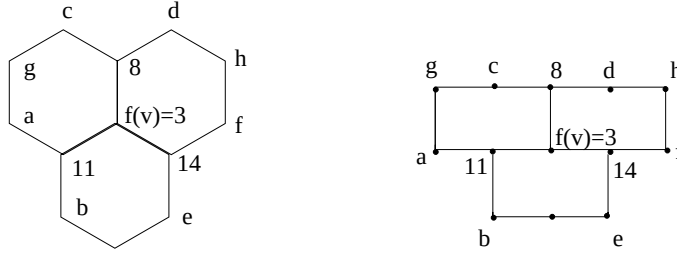
$$\lambda(\Gamma_H; k, 1) \leq \begin{cases} 3k & \text{if } 1 \leq k \leq \frac{5}{3}, \\ 5 & \text{if } \frac{5}{3} \leq k \leq 2. \end{cases}$$

Proof. It suffices to prove the first statement, due to Lemma 2.2. We will show $\lambda(\Gamma_H; 5, 3) \geq 15$.

Assume otherwise, $\lambda(\Gamma_H; 5, 3) < 15$. Then there exists an $L(5, 3)$ -labeling f with all labels in the set $\{0, \dots, 14\}$.

CLAIM 1. *The labeling f cannot use labels 4 or 10.*

Proof. Assume $f(v) = 4$ for some $v \in V(\Gamma_H)$. The three distinct labels around v are $\geq f(v) + 5 = 9$. Suppose they are labels $x_1 < x_2 < x_3$. Since any pair of the three labels differs by at least 3 (because they are at distance two from each other),

FIG. 16. The $L(5, 3)$ -labeling of a subgraph of Γ_H .

one of them is ≥ 15 , a contradiction. By the Symmetry Lemma, f cannot use label $14 - 4 = 10$. This proves the claim.

CLAIM 2. *The labeling f cannot use labels 5 or 9.*

Proof. Assume $f(v) = 5$ for some $v \in V(\Gamma_H)$. The three labels around v are $\leq f(v) - 5 = 0$ or $\geq f(v) + 5 = 10$. But 10 is excluded by the previous claim. Since any pair of the three labels differs by at least 3 it must be that the three labels used are 0, 11, 14. Then the three neighbors of the label 11 are each $\leq 11 - 5 = 6$ and any two are at least 3 apart, so they need to be 0, 3, and 6. But this is a contradiction since one of them is $f(v) = 5$. By symmetry, we must also exclude label 9, and the claim follows.

Now f has no label 4, 5, 9, 10. The proofs of Claims 3 and 4 are similar to the proof of Claim 2, so we omit the details.

CLAIM 3. *The labeling f cannot use label 7.*

CLAIM 4. *The labeling f cannot use labels 1, 2, 12, or 13.*

Now all labels of f belong to $\{0, 3, 6, 8, 11, 14\}$; call this set L .

CLAIM 5. *The labeling f cannot use labels 3 or 11.*

Proof. Assume $f(v) = 3$ for some $v \in V(\Gamma_H)$. The three labels around v are $\geq f(v) + 5 = 3 + 5 = 8$. They are 8, 11, 14 as in Figure 16. The three neighbors of the label 11 are ≤ 6 , with one of them $f(v) = 3$ and the others 0, 6. By the separation conditions and set L , we have $c, d \in \{0, 14\}$. We have two cases.

Case 1. $a = 0, b = 6$. Since $a = 0$, then $c = 14$. We cannot find a feasible label g in L , a contradiction.

Case 2. $a = 6, b = 0$. Since $b = 0$, then $e \in \{6, 8\}, f = 0$, so that $d = 14$. We cannot find a feasible label h in L , a contradiction proving the claim.

Now all labels of f belong to $\{0, 6, 8, 14\}$. We cannot label the induced subgraph $K_{1,3}$, a contradiction. $\square$

This completes the proof of the span formulas for Γ_H .

9. Concluding remarks. Despite considerable effort, we failed to complete the determination of all values $\lambda(\Gamma_\Delta; k, 1)$ for $\frac{1}{2} < k < \frac{4}{5}$. We believed that our upper bounds would be the correct values for $k \in [\frac{1}{3}, \frac{1}{2}]$, while the line $5k+2$ would be correct for $k \in [\frac{1}{2}, \frac{4}{5}]$. However, while preparing our revision of this manuscript, Král' and Škoda (building on our earlier version in circulation) found that this is not entirely the case [26]. They completed the determination of $\lambda(\Gamma_\Delta; k, 1)$. Altogether, the graph has eleven linear pieces in the interval $[0, 1]$.

As with the graphs of $\lambda(\Gamma_\square; k, 1)$ and $\lambda(\Gamma_H; k, 1)$, the graph of $\lambda(\Gamma_\Delta; k, 1)$ has the property that successive linear pieces alternately turn up and down; that is, though continuous and nondecreasing in k , the concavity keeps alternating between up and down. We continue to be puzzled why this is so.

It is natural now to extend this investigation to conditions at distance three, where very little is known. One definite result in this direction is given by Bertossi, Pinotti, and Tan [3], who determined that $\lambda(\Gamma_\Delta; 2, 1, 1) = 11$. Their construction that achieves the optimal value, 11, can be described by a labeling matrix:

$$A = \begin{bmatrix} 0 & 10 & 2 & 6 & 1 & 9 \\ 4 & 8 & 5 & 11 & 3 & 7 \\ 2 & 6 & 1 & 9 & 0 & 10 \\ 5 & 11 & 3 & 7 & 4 & 8 \\ 1 & 9 & 0 & 10 & 2 & 6 \\ 3 & 7 & 4 & 8 & 5 & 11 \end{bmatrix}.$$

On the other hand, it is easy to find 12 vertices in Γ_Δ such that the maximum distance is three, and so $\lambda(\Gamma_\Delta; 1, 1, 1) \geq 11$. It follows that $\lambda(\Gamma_\Delta; k, 1, 1) = 11$ for $1 \leq k \leq 2$.

Several papers in engineering consider labelings with conditions at distance at most p with $k_1 = k \geq k_2 = k_3 = \dots = k_p = 1$. Bertossi, Pinotti, and Tan [3] and then Panda, Kumar, and Das [30] independently obtained a lower bound for the square lattice, $\lambda(\Gamma_\square; 1, 1, \dots, 1) \geq \lfloor \frac{p^2+2p}{2} \rfloor$. This problem corresponds to integer channel assignments such that identical channels must be at least distance $p + 1$ apart in the lattice. Investigating graph models of wireless networks, Dubhashi et al. [9] presented bounds on the minimum span for $L(2, 1, 1, \dots, 1)$ -labeling of the p -dimensional square lattice (grid), in which $V(G) = \mathbb{Z}^p$, and two vertices, say $(x_1, x_2, \dots, x_p)$ and $(y_1, y_2, \dots, y_p)$, are joined by an edge whenever $\sum_{i=1}^p |x_i - y_i| = 1$.

We are continuing this project by seeking to describe *all* optimal $L(k, 1)$ -labelings of the three regular lattices, and by searching for optimal labelings with nice symmetry properties, such as being periodic or doubly periodic.

For further progress on the general theory of real number graph labelings, the reader is recommended to check the paper of Král' [25]. For an overview of the recent progress and state of the theory, refer to the recent survey [20].

Acknowledgments. The authors need to acknowledge the encouragement and sharing of ideas by their international research colleagues during this long-running project. Particular thanks go to Tiziana Calamoneri, Gerard Chang, John Georges, Dan Král', Jan Kratochvíl, Daphne Liu, David Mauro, Fred Roberts, and Roger Yeh. We are indebted to the dedicated and diligent referees for their many thoughtful editorial suggestions.

REFERENCES

- [1] A. A. BERTOSSI AND M. A. BONUCCELLI, *Code assignment for hidden terminal interference avoidance in multihop packet radio networks*, IEEE/ACM Trans. Networking, 3 (1995), pp. 441–449.
- [2] A. A. BERTOSSI, C. M. PINOTTI, R. RIZZI, AND A. M. SHENDE, *Channel assignment in honeycomb networks*, in Proceedings of the 8th Italian Conference on Theoretical Computer Science (ICTCS2003), Lecture Notes in Comput. Sci. 2841, Springer-Verlag, Berlin, 2003, pp. 150–162.
- [3] A. A. BERTOSSI, C. M. PINOTTI, AND R. B. TAN, *Channel assignment with separation for interference avoidance in wireless networks*, IEEE Trans. Parallel Distrib. Sys., 14 (2003), pp. 222–235.
- [4] R. E. BROADHURST, W. J. SHANAHAN, AND M. D. STEFFEN, “We’re sorry, you’re outside the coverage area,” UMAP J., 21 (2000), pp. 327–342.

- [5] T. CALAMONERI, *Exact solution of a class of frequency assignment problems in cellular networks and other regular grids (extended abstract)*, in Proceedings of the 8th Italian Conference on Theoretical Computer Science (ICTCS2003), Lecture Notes in Comput. Sci. 2841, Springer-Verlag, Berlin, 2003, pp. 163–173.
- [6] T. CALAMONERI, *Exact Solution of a Class of Frequency Assignment Problems in Cellular Networks and Other Regular Grids*, preprint, 2004.
- [7] T. CALAMONERI, *The $L(h, k)$ -labelling Problem: An Annotated Bibliography*, online survey updated and posted at <http://www.dsi.uniroma1.it/~calamo/PDF-FILES/survey.pdf>, 2006.
- [8] T. CALAMONERI, A. PELC, AND R. PETRESCHI, *Labeling trees with a condition at distance two*, Discrete Math., 306 (2006), pp. 1534–1539.
- [9] A. DUBHASHI, M. V. S. SASHANKA, A. PATI, S. RAMAPRASAD, AND A. M. SHENDE, *Channel assignment for wireless networks modelled as d -dimensional square grids*, in Proceedings of the 4th International Workshop on Distributed Computing, Lecture Notes in Comput. Sci. 2571, Springer-Verlag, Berlin, 2002, pp. 130–141.
- [10] D. J. DURAND, J. M. KLINE, AND K. M. WOODS, *Groovin' with the big band(width)*, UMAP J., 21 (2000), pp. 357–368.
- [11] A. GAMST, *Homogeneous distribution of frequencies in a regular hexagonal cell system*, IEEE Trans. Vehic. Tech., 31 (1982), pp. 132–144.
- [12] J. P. GEORGES AND D. W. MAURO, *Labeling trees with a condition at distance two*, Discrete Math., 269 (2003), pp. 127–148.
- [13] J. P. GEORGES AND D. W. MAURO, *On regular graphs optimally labeled with a condition at distance two*, SIAM J. Discrete Math., 17 (2003), pp. 320–331.
- [14] J. GOODWIN, D. JOHNSTON, AND A. MARCUS, *Radio channel assignments*, UMAP J., 21 (2000), pp. 369–378.
- [15] J. R. GRIGGS, *Author/Judge's commentary: The outstanding channel assignment papers*, UMAP J., 21 (2000), pp. 379–386.
- [16] J. R. GRIGGS AND X. T. JIN, *Optimal channel assignments for lattices with conditions at distance two (extended abstract)*, in Proceedings of the 5th IEEE International Parallel and Distributed Processing Symposium (International Workshop on Wireless, Mobile, and Ad Hoc Networks), extended abst., 2005.
- [17] J. R. GRIGGS AND X. T. JIN, *Real number graph labellings with distance conditions*, SIAM J. Discrete Math., 20 (2006), pp. 302–327.
- [18] J. R. GRIGGS AND X. T. JIN, *Recent progress in mathematics and engineering on optimal graph labellings with distance conditions*, J. Comb. Optim., 14 (2007), pp. 249–257.
- [19] J. R. GRIGGS AND X. T. JIN, *Real number labellings for paths and cycles*, Internet Math., to appear.
- [20] J. R. GRIGGS AND D. KRÁL', *Graph labellings with variable weights, a survey*, Discrete Appl. Math., to appear.
- [21] J. R. GRIGGS AND R. K. YEH, *Labelling graphs with a condition at distance 2*, SIAM J. Discrete Math., 5 (1992), pp. 586–595.
- [22] W. K. HALE, *Frequency assignment: Theory and applications*, Proc. IEEE, 68 (1980), pp. 1497–1514.
- [23] X. T. JIN, *Real Number Graph Labeling with Distance Conditions*, Ph.D. Dissertation, University of South Carolina, Columbia, SC, 2005.
- [24] X. T. JIN AND R. K. YEH, *Graph distance-dependent labeling related to code assignment in computer networks*, Naval Res. Logis., 52 (2005), pp. 159–164.
- [25] D. KRÁL', *The channel assignment problem with variable weights*, SIAM J. Discrete Math., 20 (2006), pp. 690–704.
- [26] D. KRÁL' AND P. ŠKODA, *Bounds for the real number graph labellings and application to labellings of the triangular lattice*, SIAM J. Discrete Math., submitted.
- [27] A. KRAMLING, M. SCHEIBENBOGEN, AND B. WALKE, *Dynamic channel allocation in wireless ATM networks*, Wireless Netw., 2 (2000), pp. 381–398.
- [28] R. A. LEESE, *A unified approach to the assignment of radio channels on a regular hexagonal grid*, IEEE Trans. Vehic. Tech., 46 (1997), pp. 968–979.
- [29] J. MINTZ, A. NEWCOMER, AND J. C. PRICE, *Channel assignment model: The span without a face*, UMAP J., 21 (2000), pp. 311–326.
- [30] B. S. PANDA, M. KUMAR, AND S. K. DAS, *Optimal schemes for channel assignment problem in wireless networks modeled as 2-dimensional square grids*, in Proceedings of the 6th International Workshop on Distributed Computing, Lecture Notes in Comput. Sci. 3326, Springer-Verlag, Berlin, 2004, pp. 424–434.
- [31] S. RAMANATHAN, *A unified framework and algorithm for channel assignment in wireless networks*, Wireless Netw., 5 (1999), pp. 81–94.

- [32] F. S. ROBERTS, *T-Colorings of graphs: Recent results and open problems*, Discrete Math., 93 (1991), pp. 229–245.
- [33] F. S. ROBERTS, *Working Group Agenda, DIMACS/DIMATIA/Renyi Working Group on Graph Colorings and Their Generalizations*, posted at <http://dimacs.rutgers.edu/Workshops/GraphColor/main.html>, 2003.
- [34] C. RUI, X. BAOXIN, AND Z. RUIDI, *Utilize the limited frequency resources efficiently*, UMAP J., 21 (2000), pp. 343–356.
- [35] J. VAN DEN HEUVEL, R. A. LEESE, AND M. A. SHEPHERD, *Graph labeling and radio channel assignment*, J. Graph Theory, 29 (1998), pp. 263–283.
- [36] D. ZHU AND A. SHI, *Optimal Channel Assignments*, preprint, 2001.

COMPUTING COMMON INTERVALS OF K PERMUTATIONS, WITH APPLICATIONS TO MODULAR DECOMPOSITION OF GRAPHS*

ANNE BERGERON[†], CEDRIC CHAUVE[‡], FABIEN DE MONTGOLFIER[§], AND
MATHIEU RAFFINOT[§]

Abstract. We introduce a new approach to compute the common intervals of K permutations based on a very simple and general notion of generators of common intervals. This formalism leads to simple and efficient algorithms to compute the set of all common intervals of K permutations that can contain a quadratic number of intervals, as well as a linear space basis of this set of common intervals. Finally, we show how our results on permutations can be used for computing the modular decomposition of graphs.

Key words. common interval, permutation, PQ -tree, modular decomposition

AMS subject classifications. 05C05, 05C62, 68R99

DOI. 10.1137/060651331

1. Introduction. The notion of *common interval* was introduced by Uno and Yagiura [19] in order to model the fact that, when comparing genomes, a group of genes can be rearranged but still remain connected. They proposed a first algorithm that computes the set of common intervals of a permutation P with the identity permutation in time $O(n + N)$, where n is the length of P , and N is the number of common intervals. However, N can be of size $O(n^2)$, thus the algorithm of Uno and Yagiura has an $O(n^2)$ time complexity. Heber and Stoye [12] defined a subset of size $O(n)$ of the common intervals of K permutations, called *irreducible intervals*, that forms a basis of the set of all common intervals: every common interval is a chain overlapping irreducible intervals. They proposed an $O(Kn)$ time algorithm to compute the set of irreducible intervals of K permutations, based on Uno and Yagiura's pioneering work.

One of the drawbacks of these algorithms is that properties of Uno and Yagiura's algorithm are not obvious [5]. Even the authors describe their $O(n + N)$ algorithm as “quite complicated,” and, in practice, simpler $O(n^2)$ algorithms run faster on randomly generated permutations [19]. On the other hand, Heber and Stoye's algorithms rely on a complex data structure that mimics what is known, in the theory of modular decomposition of graphs, as the PQ -trees of *strong intervals*. An incentive to revisit this problem is the central role that these PQ -trees seem to play in the field of comparative genomics. Strong intervals can be used to identify significant groups of genes that are conserved between genomes [13] or as guides to reconstruct evolution scenarios [1, 10].

*Received by the editors February 1, 2006; accepted for publication (in revised form) February 17, 2008; published electronically June 11, 2008.

<http://www.siam.org/journals/sidma/22-3/65133.html>

[†]Université du Québec à Montréal, Montréal H3C 3P8, QC, Canada (bergeron.anne@uqam.ca).

[‡]Simon Fraser University, Burnaby V5A 1S6, BC, Canada (cedric.chauve@sfu.ca).

[§]LIAFA, Université Denis Diderot - Case 7014, 2 place Jussieu, F-75251 Paris Cedex 05, France (fm@liafa.jussieu.fr, raffinot@liafa.jussieu.fr). Most of the work was done when the fourth author was at Poncelet Laboratory (CNRS UMI-2615), Independent University of Moscow, 11 street Bolchoï Vlassievski, 119 002 Moscow, Russia.

In order to design alternative efficient algorithms to compute common intervals, we propose a theoretical framework for common intervals based on generating families of intervals. For two permutations, these families can be computed by straightforward $O(n)$ algorithms that use only tables and stacks as data structures and that upgrade trivially to the case of K permutations. Using these families, we compute common intervals with simple $O(n + N)$ and $O(n)$ algorithms whose properties can be readily verified. We also propose a new canonical representation of the family of common intervals that is simpler than the PQ -trees. We then link this work to previous studies on common intervals and show how our new representation can be transformed in linear time into classical ones, namely PQ -trees and irreducible intervals. Conversely, generating families can be linearly built from these representations.

Finally, we extend our approach to the classical graph problem of *modular decomposition* that aims to efficiently compute a compact representation of the modules of a graph. The first linear time algorithms that were developed [8, 15] are rather complex, and many efforts have been put into the design of decomposition algorithms that are efficient in practice, even if they do not run in linear time but in quasi-linear time [9, 16].

The article is structured as follows. In section 2, we describe the notion of generators of common intervals and how to compute generators of K permutations of size n in $O(Kn)$ time. Section 3 explains how to generate the set of all N common intervals in $O(n + N)$ using a generator. Section 4 describes a new linear space basis of common intervals, called the canonical generator. Section 5 links this new representation to classical ones, namely strong intervals, irreducible intervals, and PQ -trees. Finally, in section 6, we extend our results to the modular decomposition of graphs. An extended abstract of this article appeared in [2].

2. Common intervals and generators. A *permutation* P on n elements is a complete linear order on the set of integers $\{1, 2, \dots, n\}$. We denote Id_n the identity permutation $(1, 2, \dots, n)$. An *interval* of a permutation $P = (p_1, p_2, \dots, p_n)$ is a set of consecutive elements of permutation P . An interval of the identity permutation will be denoted by giving the indices of its left and right bounds $(i..j)$.

DEFINITION 2.1. Let $\mathcal{P} = \{P_1, P_2, \dots, P_K\}$ be a set of K permutations on n elements. A *common interval* of $\mathcal{P}$ is a set of integers that is an interval in each permutation of $\mathcal{P}$.

The set $\{1, 2, \dots, n\}$ and all singletons are always common intervals of any non-empty set of permutations; they are called *trivial* intervals. In what follows, we assume, without loss of generality, that the set $\mathcal{P}$ contains the identity permutation Id_n . A common interval of $\mathcal{P}$ can thus be denoted as an interval $(i..j)$ of the identity permutation.

Algorithms that identify common intervals such as those in [12, 19] use a bottom-up approach that constructs the common intervals by extending small intervals until they form a common interval and then form longer common intervals by taking unions of common intervals that share elements. Here we adopt a dual approach of constructing common intervals as intersections of intervals that are not necessarily common. This had two major impacts on the rest of the theory. First, it allowed us to exploit the symmetric properties of minima and maxima of intervals, and it provided the natural notion of two families of intervals of size n . Moreover, by removing the need to maintain common intervals, the design of the algorithms was much easier, and the data structures became trivial.

DEFINITION 2.2. Let $\mathcal{P} = \{\text{Id}_n, P_2, \dots, P_K\}$ be a set of K permutations on n

elements. A generator for the common intervals of $\mathcal{P}$ is a pair (R, L) of vectors of size n such that

1. $R[i] \geq i$ and $L[j] \leq j$ for all $i, j \in \{1, 2, \dots, n\}$,
2. $(i..j)$ is a common interval of $\mathcal{P}$ if and only if $(i..j) = (i..R[i]) \cap (L[j]..j)$.

Definition 2.2 has the following easy property that will turn out very useful in proving results on the structure of generators.

PROPOSITION 2.3. *For any generator (R, L) , $(i..j) = (i..R[i]) \cap (L[j]..j)$ if and only if $L[i] \leq i \leq j \leq R[i]$.*

Proof. If $(i..j) = (i..R[i]) \cap (L[j]..j)$, then $(i..j) \subseteq (i..R[i])$ and $(i..j) \subseteq (L[j]..j)$, which yields $L[i] \leq i \leq j \leq R[i]$. The converse is immediate. $\square$

The following proposition shows how to construct a generator for the common intervals of a union of sets of permutations, given generators for the common intervals of each set. If X and Y are two vectors, we denote by $\min(X, Y)$ the vector $\min(X[1], Y[1]), \dots, \min(X[n], Y[n])$.

PROPOSITION 2.4. *Let (R_1, L_1) and (R_2, L_2) be generators for the common intervals of two sets $\mathcal{P}_1$ and $\mathcal{P}_2$ of permutations, both containing the identity permutation. The pair $(\min(R_1, R_2), \max(L_1, L_2))$ is a generator for the common intervals of $\mathcal{P}_1 \cup \mathcal{P}_2$.*

Proof. Interval $(i..j)$ is a common interval of $\mathcal{P}_1 \cup \mathcal{P}_2$ if and only if it is a common interval of both $\mathcal{P}_1$ and $\mathcal{P}_2$, which is equivalent, by Proposition 2.3, to $L_1[j] \leq i \leq j \leq R_1[i]$ and $L_2[j] \leq i \leq j \leq R_2[i]$ and finally to $\max(L_1[j], L_2[j]) \leq i \leq j \leq \min(R_1[i], R_2[i])$. $\square$

Proposition 2.4 implies that, given an $O(n)$ algorithm for computing generators for the common intervals of two permutations, we can easily deduce an $O(Kn)$ algorithm for computing a generator for the common intervals of K permutations.

Generators are far from unique, but some are easier to compute than others. Identifying good generators is a crucial step in the design of efficient algorithms to compute common intervals. The remainder of this section focuses on particular classes of generators that turn out to have interesting properties with respect to computations.

DEFINITION 2.5. *Let $P = (p_1, \dots, p_n)$ be a permutation on n elements. For each element p_i , we define two intervals containing p_i :*

$\text{IMax}[p_i]$ *is the largest interval of P containing p_i and whose elements are all $\geq p_i$;*

$\text{IMin}[p_i]$ *is the largest interval of P containing p_i and whose elements are all $\leq p_i$. And we define the following two integers:*

$\text{Sup}[p_i]$ *is the largest integer such that $(p_i.. \text{Sup}[p_i]) \subseteq \text{IMax}[p_i]$;*

$\text{Inf}[p_i]$ *is the smallest integer such that $(\text{Inf}[p_i]..p_i) \subseteq \text{IMin}[p_i]$.*

Remark that $(p_i.. \text{Sup}[p_i])$ and $(\text{Inf}[p_i]..p_i)$ are intervals of the identity permutation but not necessarily intervals of permutation P . For example, if $P = (1, 4, 7, 5, 9, 6, 2, 3, 8)$, we have $\text{IMax}[5] = (7, 5, 9, 6)$ and $\text{Sup}[5] = 7$, and $\text{IMin}[8] = (6, 2, 3, 8)$ and $\text{Inf}[8] = 8$.

PROPOSITION 2.6. *The pair of vectors (Sup, Inf) is a generator for the common intervals of P and Id_n .*

Proof. Suppose that $(i..j)$ is a common interval of P and Id_n ; then $\text{Sup}[i] \geq j$ and $\text{Inf}[j] \leq i$ since all elements in the set $(i..j)$ are consecutive in permutation P . Thus $(i..j) = (i.. \text{Sup}[i]) \cap (\text{Inf}[j]..j)$. On the other hand, suppose that $\text{Sup}[i] \geq j$ and $\text{Inf}[j] \leq i$; then $\text{IMax}[i]$ contains j and $\text{IMin}[j]$ contains i . Since both $\text{IMax}[i]$ and $\text{IMin}[j]$ are intervals of P , their intersection is an interval and is equal to $(i..j)$. $\square$

Example. Let $\mathcal{P} = \{\text{Id}_8, P_2\}$ and $\mathcal{Q} = \{\text{Id}_8, P_3\}$ with

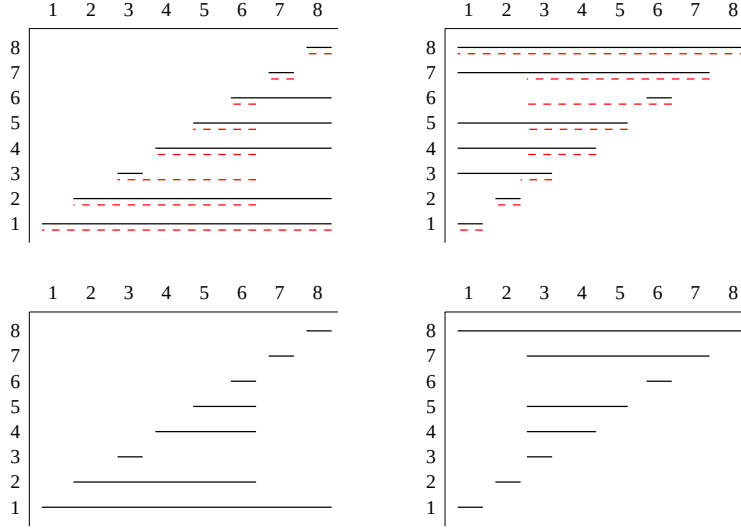


FIG. 1. The top two diagrams show the generators (Sup, Inf) of the common intervals of the set $\mathcal{P}$ in solid lines and of set $\mathcal{Q}$ in dashed lines. A line in row i of the left diagram extends from column i to column $\text{Sup}(i)$, and a line in row i of the right diagram extends from column $\text{Inf}(i)$ to column i . The bottom diagrams show a generator for the common intervals of $\mathcal{P} \cup \mathcal{Q}$ constructed using Proposition 2.4.

$$\text{Id}_8 = (1, 2, 3, 4, 5, 6, 7, 8), \quad P_2 = (1, 3, 2, 4, 5, 7, 6, 8), \quad P_3 = (2, 8, 3, 4, 5, 6, 1, 7).$$

The generators (Sup, Inf) for the common intervals of $\mathcal{P}$ and $\mathcal{Q}$ are shown in Figure 1. This figure also shows a generator for $\mathcal{P} \cup \mathcal{Q}$.

ALGORITHM 1. COMPUTING THE GENERATOR (Sup, Inf).

```

Inf[1] ← 1, Sup[n] ← n.
For k from 1 to n, m[k] ← k, M[k] ← k.
For k from 2 to n
  While m[k] - 1 is in IMin[k], m[k] ← m[m[k] - 1]
  Inf[k] ← m[k]
For k from n - 1 to 1
  While M[k] + 1 is in IMax[k], M[k] ← M[M[k] + 1]
  Sup[k] ← M[k]

```

PROPOSITION 2.7. Let P be a permutation on n elements. If the bounds of intervals $\text{IMax}[k]$ and $\text{IMin}[k]$ are known for all k , then Algorithm 1 computes (Sup, Inf) in $O(n)$ time.

Proof. We first show that Algorithm 1 is correct. Suppose that, at the beginning of the k th iteration of the second For loop, $\text{Inf}[k'] = m[k']$ for all $k' < k$, and $m[k] \in \text{IMin}[k]$. This is the case at the beginning of iteration $k = 2$, since $\text{Inf}[1] = 1$. By definition, $\text{Inf}[k] \leq k$, thus before entering the While loop, we have $\text{Inf}[k] \leq m[k]$. If the test $m[k] - 1 \in \text{IMin}[k]$ of the While loop is true, then $\text{Inf}[k] \leq m[k] - 1$, implying that $\text{Inf}[k] \leq \text{Inf}[m[k] - 1]$. Since $\text{Inf}[m[k] - 1] = m[m[k] - 1]$ by hypothesis, the instruction in the While loop preserves the invariant $\text{Inf}[k] \leq m[k]$. When the test of the While loop becomes false, then $\text{Inf}[k]$ is greater than $m[k] - 1$, thus $\text{Inf}[k] = m[k]$. The proof of correctness for Sup is similar.

Using the inverse of permutation P , the tests in the While loops can be done in

constant time. The total time complexity follows from the fact that the instruction within the *While* loop is executed exactly $n - 1$ times. Indeed, consider, at any point of the execution of the algorithm, the collection I of intervals $(m[k]..k)$ of the identity permutation that are not contained in any other interval of this type. After the initialization loop, we have n such intervals, that is, $I = \{(1..1), (2..2), \dots, (n..n)\}$, and at the completion of the algorithm, there is only one, namely $(m[n]..n) = (1..n)$, since $\text{Inf}[n] = 1$. The instruction in the *While* loop merges two consecutive intervals $(m[m[k] - 1]..m[k] - 1)$ and $(m[k]..k)$ of I into one, since the current value of $m[k]$ becomes $m[m[k] - 1]$. Since there were n intervals to start with, there can be at most $n - 1$ of these merges. $\square$

The computation of the bounds of intervals $\text{IMax}[k]$ and $\text{IMin}[k]$, as well as the computation of the inverse of permutation P , are quite straightforward. As an example, Algorithm 2 shows how to compute the left bound of $\text{IMax}[p_i]$.

PROPOSITION 2.8. *Let $P = (p_1, \dots, p_n)$ be a permutation on n elements. Algorithm 2 computes the left bound of all intervals $\text{IMax}[p_i]$ in $O(n)$ time.*

Proof. The time complexity of Algorithm 2 is immediate since each position is stacked once. Its correctness relies on the fact that, at the beginning of the i th iteration, the position j of the nearest left element such that $p_j < p_i$ must be in the stack. If it was not the case, then an element smaller than p_j was found between the positions j and i , contradicting the definition of position j . $\square$

ALGORITHM 2. COMPUTING THE LEFT BOUND OF $\text{IMax}[p_i]$ FOR ALL p_i .

S is a stack of positions; s denotes the top of S .
 Push 0 on S
 $p_0 \leftarrow 0$
 For i from 1 to n
 While $p_i < p_s$ Pop the top of S
 left bound of $\text{IMax}[p_i] \leftarrow s + 1$
 Push i on S

To summarize the results of this section, we have the following theorem.

THEOREM 2.9. *Let $\mathcal{P} = \{\text{Id}_n, P_2, \dots, P_K\}$ be a set of K permutations on n elements. A generator for the common intervals of $\mathcal{P}$ can be computed in $O(Kn)$ time.*

3. Common intervals of K permutations in optimal time. We now turn to the problem of generating all common intervals of K permutations in $O(N)$ time, where N is the number of such common intervals, given a generator satisfying the following property, based on the notion of *commuting* sets. Note that commuting families are also known as *laminar* families [18] in the field of combinatorial optimization.

DEFINITION 3.1. *Two sets A and B commute if either $A \subseteq B$, or $B \subseteq A$, or A and B are disjoint, and otherwise they overlap. A collection $\mathcal{C}$ of sets is commuting if, for any pair of sets A and B in $\mathcal{C}$, A and B commute. A generator (R, L) for the common intervals of $\mathcal{P} = \{\text{Id}_n, P_2, \dots, P_K\}$ is commuting if both the collections $\{(i..R[i])\}_{i \in (1..n)}$ and $\{(L[i]..i)\}_{i \in (1..n)}$ are commuting. If (R, L) is a commuting generator, we define $\text{Support}[i]$, for $i > 1$, to be the greatest integer $j < i$ such that $R[i] \leq R[j]$. (Similar values can be defined for L , but we will not refer to them explicitly.)*

It turns out that generators defined in section 2 are commuting. Indeed, generators defined in Proposition 2.4 are commuting if they are constructed with generators (R_1, L_1) and (R_2, L_2) that are commuting. This is a consequence of the fact that if

$a < b$ and $a' < b'$, then $\min(a, a') < \min(b, b')$ and $\max(b, b') > \max(a, a')$. For the generator (Sup, Inf) , we have the following proposition.

PROPOSITION 3.2. *The generator (Sup, Inf) for the common intervals of permutations $\mathcal{P}$ and Id_n is commuting.*

Proof. Suppose that $(k.. \text{Sup}[k])$ contains k' ; we will show that it must also contain $\text{Sup}[k']$. If k' is in $(k.. \text{Sup}[k])$, then k' is in $\text{IMax}[k]$, and $k' > k$; therefore, $\text{IMax}[k'] \subseteq \text{IMax}[k]$, and the interval $(k'.. \text{Sup}[k'])$ is included in $\text{IMax}[k]$, thus in $(k.. \text{Sup}[k])$ also. Since $\text{Sup}[k]$ is maximal, we must have $\text{Sup}[k] \geq \text{Sup}[k']$. A similar argument holds for Inf . $\square$

PROPOSITION 3.3. *Given a commuting generator (R, L) , Algorithm 3 computes the values $\text{Support}[i]$, for all $i > 1$, in linear time.*

Proof. The time complexity of Algorithm 3 is immediate. At iteration i the stack contains the left bounds of all intervals of $(j.. R[j])$ such that $R[j] \geq i$ and $j < i$, sorted in decreasing size order. It is then easy to see when equality holds. Note that $\text{Support}[1]$ is undefined and should not be used by subsequent algorithms. $\square$

THEOREM 3.4. *Given a commuting generator (R, L) , Algorithm 4 outputs all common intervals of a set $\mathcal{P}$ of K permutations on n elements in $O(n + N)$ time, where N is the number of common intervals of the set $\mathcal{P}$.*

Proof. The time complexity of Algorithm 4 is immediate. Suppose that interval $(i..j)$ is identified by the algorithm. At the start of the j th iteration of the *For* loop, $i = j$, thus $j \leq R[i]$. If the test of the *While* loop is true, then $i \geq L[j]$, and $(i..j)$ is a common interval. If $i' = \text{Support}[i]$, then $R[i'] \geq R[i]$, thus $j \leq R[i']$ at the end of the *While* loop.

On the other hand, if $(i..j)$ is a common interval of $\mathcal{P}$, with $i < j$, then $\text{Support}[j] \geq i$, since $R[i] \geq R[j]$. Let i' be the smallest integer such that $i < i'$ and $(i'..j)$ is identified by Algorithm 4 as a common interval. Such an interval exists, since $(j..j)$ is a common interval. Finally, $\text{Support}[i'] = i$ since $\text{Support}[i']$ must be greater than or equal to i . If it is greater, then $(\text{Support}[i']..j)$ is a common interval, contradicting the definition of i' . $\square$

ALGORITHM 3. COMPUTING $\text{Support}[i]$ FOR A COMMUTING GENERATOR (R, L) .

S is an empty stack; s denotes the top of S .
 Push 1 on S
 For i from 2 to n
 While $R[s] < i$ Pop the top of S
 $\text{Support}[i] \leftarrow s$
 Push i on S

ALGORITHM 4. COMMON INTERVALS OF A SET $\mathcal{P}$ GIVEN A GENERATOR (R, L) .

For j from n to 1
 $i \leftarrow j$
 While $i \geq L[j]$
 Output $(i..j)$ (* Interval $(i..j)$ is a common interval of the set $\mathcal{P}$ *)
 $i \leftarrow \text{Support}[i]$

4. A new canonical representation of closed families. The set of common intervals of a set of permutations is an example of more general families of intervals, the *closed* families. Closed families can have a quadratic number of elements, but a classical result establishes a bijection between PQ -trees with n leaves and closed families of Id_n , thus allowing a representation of size $O(n)$ ([3]—see also Proposition 5.3). In this section, we develop a new canonical representation for such families, based

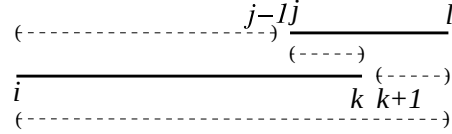


FIG. 2. Two overlapping common intervals determine four more common intervals.

on the generators of the previous section. We will discuss the relations between this representation and PQ -trees in section 5.

DEFINITION 4.1. A closed family $\mathcal{F}$ of intervals of a permutation σ on n elements is a family that contains all singletons and the interval $(1..n)$ and that has the following property: if $(i..k)$ and $(j..l)$ are in $\mathcal{F}$, and $i < j \leq k < l$, then $(i..j-1)$, $(j..k)$, $(k+1..l)$, and $(i..l)$ belong to $\mathcal{F}$.

In what follows, we will suppose, for simplicity, that the permutation σ is the identity permutation, thus every interval of the family $\mathcal{F}$ is an interval of the form $(i..j)$. The closure properties of Definition 4.1 are well-known properties of the set of common intervals of a set of permutations and are illustrated in Figure 2. Conversely, it is easy to extend the definition of generators (Definition 2.2) to the more general case of closed families, and we will refer to a *generator of a closed family* $\mathcal{F}$ as a pair (R, L) that satisfies the conditions of Definition 2.2 with respect to the members of $\mathcal{F}$. Among all possible generators, the following one provides a representation of size $O(n)$ for any closed family.

DEFINITION 4.2. A generator (R, L) for a closed family $\mathcal{F}$ is canonical if, for all $i \in (1..n)$, intervals $(i..R[i])$ and $(L[i]..i)$ belong to $\mathcal{F}$.

PROPOSITION 4.3. Let $\mathcal{F}$ be a closed family. The canonical generator of $\mathcal{F}$ always exists, and it is unique and commuting.

Proof. Let $\mathcal{F}$ be a closed family. For $1 \leq i \leq n$, define $R[i]$ as the largest integer such that $(i..R[i]) \in \mathcal{F}$, and define $L[i]$ as the smallest integer such that $(L[i]..i) \in \mathcal{F}$.

If an interval $(i..j) \in \mathcal{F}$, then $(i..j) \subseteq (i..R[i])$, and $(i..j) \subseteq (L[j]..j)$, thus (R, L) is a generator. It is canonical since we picked elements of $\mathcal{F}$. Suppose that there exists a second canonical generator (R', L') , with $R \neq R'$; then there exists $1 \leq i \leq n$ such that $R'[i] < R[i]$. Since $(i..R[i])$ is in $\mathcal{F}$, it should be generated by (R', L') , but $(i..R'[i]) \cap (L'[R[i]]..R[i])$ does not contain $R[i]$. A similar argument holds if $L \neq L'$. Finally, suppose that two intervals $(i..R[i])$ and $(j..R[j])$ overlap with $i < j < R[i] < R[j]$. Then $(i..R[j])$ is in $\mathcal{F}$, which contradicts the maximality of $(i..R[i])$. $\square$

The following elementary property of canonical generators states that if two intervals of a generator overlap, they always do it correctly, in the sense that an $(L[j]..j)$ interval is always at the left of an $(i..R[i])$ interval.

PROPOSITION 4.4. For a canonical generator (R, L) , if $(i..R[i])$ and $(L[j]..j)$ overlap, then $L[j] \leq i \leq j \leq R[i]$.

Proof. Suppose that $i \leq L[j] \leq R[i] \leq j$; then $(i..j)$ is in $\mathcal{F}$ since both $(i..R[i])$ and $(L[j]..j)$ are. Then, by definition, $(i..j) = (i..R[i]) \cap (L[j]..j)$, thus $(i..R[i]) \cup (L[j]..j) = (i..R[i]) \cap (L[j]..j)$, implying $(i..R[i]) = (L[j]..j)$, contrary to the assumption that these intervals overlap. $\square$

THEOREM 4.5. Given a commuting generator (R', L') , Algorithm 5 computes the canonical generator (R, L) of a closed family $\mathcal{F}$ in $O(n)$ time.

Proof. The time complexity of Algorithm 5 follows from the fact that testing if an interval belongs to $\mathcal{F}$ can be computed in $O(1)$ time with the generator (R', L') . Its correctness relies on the following observation: if $R[k] \neq k$, then there exists an

integer $k' > k$ such that

$$(4.1) \quad \text{Support}[k'] = k \quad \text{and} \quad R[k'] = R[k],$$

where $\text{Support}[k']$ is defined (Definition 3.1) as the greatest integer smaller than k' such that $R'[\text{Support}[k']] \geq R'[k']$. Statement (4.1) implies that, since the values of $R[k]$ are computed in decreasing order, the value of $R[k]$ will be known at the start of iteration k .

In order to prove statement (4.1), let k' be the smallest integer such that $R[k'] = R[k]$. The hypothesis $R[k] \neq k$, and the fact that (R, L) is a commuting generator, imply that $k' > k$. We must show that $\text{Support}[k'] = k$. Since (R', L') is a commuting generator, and (R, L) is the canonical generator, we must have $R'[k] \geq R'[k']$. Thus, $\text{Support}[k'] \geq k$, implying that $R[\text{Support}[k']] \leq R[k]$. Since $R[k] = R[k']$, this would imply $R[\text{Support}[k']] = R[k]$, contradicting the definition of k' . $\square$

ALGORITHM 5. CANONICAL GENERATOR (R, L) GIVEN A COMMUTING GENERATOR (R', L') .

The vector *Support* is obtained from R' using Algorithm 3
 $R[1] \leftarrow n$
 For k from 2 to n
 $R[k] \leftarrow k$
 For k from n to 2
 If $(\text{Support}[k], R[k]) \in \mathcal{F}$
 $R[\text{Support}[k]] \leftarrow \max(R[k], R[\text{Support}[k]])$
 (* Computation of L is similar, by defining the vector *Support* with respect to L' *)

Example (continued). Let $\mathcal{R} = \{\text{Id}_8, P_2, P_3\}$ with

$$\text{Id}_8 = (1, 2, 3, 4, 5, 6, 7, 8), \quad P_2 = (1, 3, 2, 4, 5, 7, 6, 8), \quad P_3 = (2, 8, 3, 4, 5, 6, 1, 7).$$

Two generators for the common interval of $\mathcal{R}$ are shown in Figure 3. The second one is canonical.

5. Transformations of canonical representations. There exist many representations of closed families, and this variety is useful. Indeed, some are better suited for algorithmic purposes [12], and others yield nice graphical representations of the organization and nesting properties of a set of common intervals [3, 14]. In this section, we present algorithms that allow the conversion between the canonical generators of the preceding section and the classic representation using PQ -trees.

5.1. Canonical representations of closed families. Let T be a tree whose n leaves are labeled with n different labels. The *frontier* of a node is the set of labels of the leaves of the subtree rooted at this node.

DEFINITION 5.1. Let $\mathcal{F}$ be a closed family. A strong interval of $\mathcal{F}$ is an interval of $\mathcal{F}$ that commutes with each interval of $\mathcal{F}$.

A proper commuting family $\mathcal{F}$ of subsets of a set V is a commuting family such that $V \in \mathcal{F}$, all the singletons are in $\mathcal{F}$, and $\emptyset \notin \mathcal{F}$. In particular, the strong common intervals of K permutations form a proper commuting family. A proper commuting family $\mathcal{F}$ can be represented by its *inclusion tree* in which the frontiers of the nodes are in bijection with the members of the family. Its root is V and its leaves are the singletons of $\mathcal{F}$. Conversely, the frontiers of a tree with V as leaf labels define a

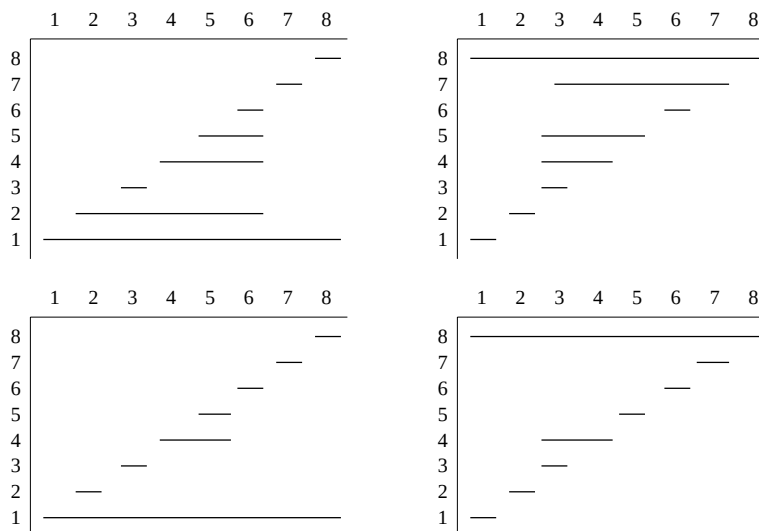


FIG. 3. The top two diagrams show a generator for the common intervals of $\mathcal{R}$ constructed using Proposition 2.4 (see Figure 1). The bottom diagrams shows the canonical generator constructed by Algorithm 5.

proper commuting family. The inclusion trees on V are thus in bijection with the proper commuting families on V .

For closed families, PQ -trees play the same role as the inclusion trees for proper commuting families.

DEFINITION 5.2. A PQ -tree on finite set V is a tree whose leaves are labeled from 1 to $|V|$ and whose internal nodes are labeled P -nodes or Q -nodes. A P -node must have at least two children, and a Q -node must have at least three children. The children of a P -node are unordered, and the children of a Q -node are totally ordered.

The reversal of a Q node consists in reversing the total order of its children. An extended frontier of a PQ -tree is either the frontier of a P node, the union of frontiers of consecutive children of a Q node, or a singleton.

PROPOSITION 5.3 (see [3, 7]). Given a closed family $\mathcal{F}$, there exists a PQ -tree such that the intervals of the family are exactly the extended frontiers. The strong intervals of the family are exactly the frontiers of this tree. Furthermore, the PQ -tree is unique up to Q node reversals.

Example. Let $\mathcal{P}_2 = \{\text{Id}_9, P_4, P_5, P_6\}$ with

$$\begin{aligned} \text{Id}_9 &= (1, 2, 3, 4, 5, 6, 7, 8, 9), & P_5 &= (6, 5, 7, 8, 9, 1, 2, 3, 4), \\ P_4 &= (9, 8, 7, 5, 6, 4, 3, 2, 1), & P_6 &= (1, 3, 2, 4, 5, 6, 9, 8, 7). \end{aligned}$$

Figure 4 shows the canonical generator of the set $\mathcal{P}_2$ and the corresponding PQ -tree.

Compared to PQ -trees, the canonical generator of a closed family $\mathcal{F}$ is much simpler since it uses only two arrays. Moreover, some operations, for example testing whether an interval $(i..j)$ belongs to the family $\mathcal{F}$, are also simpler using this representation. However, PQ -trees have the advantage of being recursive structures.

Another canonical representation, the family of *irreducible intervals*, was introduced in [12]. The links between PQ -trees and irreducible intervals have been studied in [13] that present linear-time algorithms for the conversion.

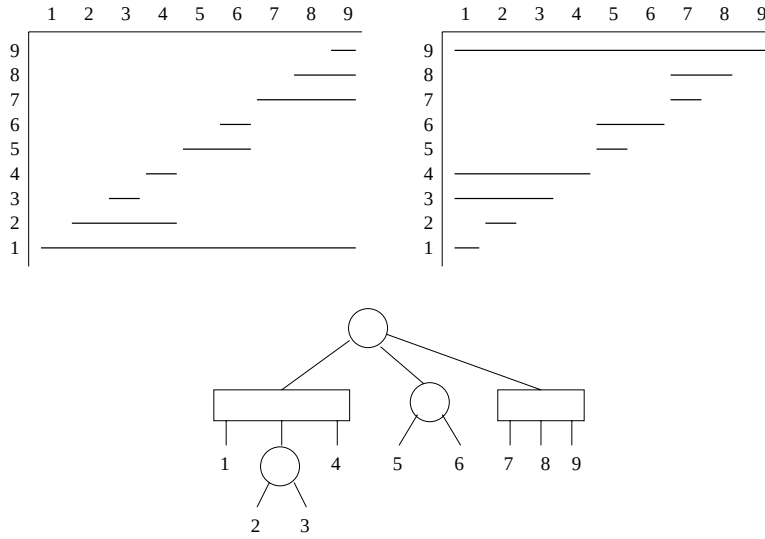


FIG. 4. The canonical generator and the associated PQ-tree for the set of permutations $\mathcal{P}_2 = \{\text{Id}_9, P_4, P_5, P_6\}$.

5.2. From canonical generators to strong intervals. Strong intervals of a closed family $\mathcal{F}$ are associated with *overlap classes*: let us consider the canonical generator (R, L) of $\mathcal{F}$. The transitive closure of the overlap relation on all intervals of $R \cup L$ is an equivalence relation, and its equivalence classes are henceforth called *overlap classes*. We will prove, in Theorem 5.8, that once the overlap classes are identified, the task of identifying the strong intervals is almost complete.

A *trivial* overlap class contains only a single interval $(i..R[i])$ or $(L[j]..j)$. We have the following lemma.

LEMMA 5.4. *A trivial overlap class is a strong interval of $\mathcal{F}$.*

Proof. Consider a trivial overlap class $\{(i..j)\}$. Either $j = R[i]$ or $L[j] = i$. Suppose that $j = R[i]$ and that $(i..j)$ is not strong. Then it is overlapped by an interval $(i'..j')$ belonging to $\mathcal{F}$ but not to the generator. If $i \leq i' \leq j \leq j'$, then the interval $(i..j')$ is in $\mathcal{F}$ but larger than $(i..R[i])$ in violation with the fact that $j' \leq R[i]$ (see Proposition 2.3). Let us then suppose $i' \leq i \leq j' \leq j$. Since $(L[j']..j')$ does not overlap $(i..R[i])$, we have $i \leq L[j'] \leq j'$. Since $i' \leq i \leq L[j']$, $(i'..j')$ is larger than $(L[j']..j')$ and thus does not belong to $\mathcal{F}$. Interval $(i..j)$ is therefore strong. The other case is similar. $\square$

For nontrivial overlap classes, we need to explain the relations between the various intervals that belong to the classes. Note that any nontrivial overlap class must contain intervals of R and L since intervals of R (resp., L) do not overlap each other. Figure 5 shows a typical nontrivial class: the number of left and right intervals is the same, left and right bounds are nicely aligned, and the left and right intervals are nested. The following lemmas establish these properties formally and can be skipped at the first reading.

LEMMA 5.5. *Let (R, L) be the canonical generator of a closed family $\mathcal{F}$. We have the following: (1) if $(i..R[i])$ overlaps $(L[j]..j)$ and $(L[j']..j')$, then $L[j] = L[j']$; and (2) if $(L[j]..j)$ overlaps $(i..R[i])$ and $(i'..R[i'])$, then $R[i] = R[i']$.*

Proof. Note that if $(i..R[i])$ overlaps $(L[j]..j)$, then $(i..j)$ belongs to $\mathcal{F}$ and, by Proposition 4.4, $i \geq L[j]$. The hypothesis that $(i..R[i])$ overlaps $(L[j]..j)$ implies fur-

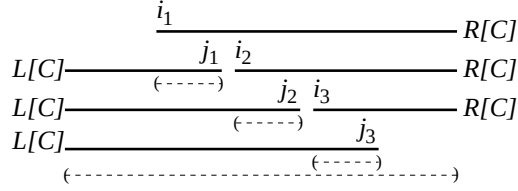


FIG. 5. Structure of a nontrivial overlap class $\mathcal{C}$ with $k = 3$. Dashed lines represent the strong intervals corresponding to the class.

ther that $i < L[j]$; otherwise we would have $(L[j]..j) \subseteq (i..R[i])$. A similar argument shows that $(i..R[i])$ overlapping $(L[j']..j')$ implies that $j' < R[i]$.

Now suppose $L[j] \neq L[j']$ and $j' > j$; then $L[j'] < L[j]$, since L is commuting. This implies that $L[j'] < L[j] < i \leq j < j' < R[i]$. Interval $(L[j']..i-1)$ belongs to $\mathcal{F}$ since it is $(L[j']..j') \setminus (i..R[i])$. Thus $(L[j']..j)$ is also in $\mathcal{F}$ since it is the union of overlapping $(L[j']..i-1)$ and $(L[j]..j)$. Applying Proposition 4.4 to interval $(L[j']..j)$ implies that $L[j'] \geq L[j]$, which yields a contradiction. Therefore $L[j] = L[j']$.

Statement (2) is proved in a similar way. $\square$

Lemma 5.5 implies that, in a nontrivial class $\mathcal{C}$, all intervals of the form $(i..R[i])$ share the same right bound, and all intervals of the form $(L[j]..j)$ share the same left bound. We will henceforth denote $L[\mathcal{C}]$ and $R[\mathcal{C}]$ these common bounds. Moreover, the interval $(L[\mathcal{C}]..R[\mathcal{C}])$ is necessarily strong.

LEMMA 5.6. *Let $\mathcal{C}$ be a nontrivial overlap class of a closed family $\mathcal{F}$. Then $(L[\mathcal{C}]..R[\mathcal{C}])$ is a strong interval of $\mathcal{F}$.*

Proof. Since $\mathcal{C}$ is a nontrivial class, it contains at least two overlapping intervals $(L[\mathcal{C}]..j)$ and $(i..R[\mathcal{C}])$. The interval $(L[\mathcal{C}]..R[\mathcal{C}])$ is an interval of $\mathcal{F}$ since it is the union of those two overlapping intervals. Suppose that $(i..j) \in \mathcal{F}$ overlaps $(L[\mathcal{C}]..R[\mathcal{C}])$ and $i < L[\mathcal{C}] \leq j < R[\mathcal{C}]$; then $(L[j]..j)$ contains $(i..j)$ and overlaps $(L[\mathcal{C}]..R[\mathcal{C}])$, thus it overlaps at least one member of $\mathcal{C}$, and therefore it belongs to $\mathcal{C}$ and $L[j] = L[\mathcal{C}]$. On the other hand, if $L[\mathcal{C}] < i \leq R[\mathcal{C}] < j$, then $(i..R[i])$ overlaps $(L[\mathcal{C}]..R[\mathcal{C}])$, thus it overlaps at least one member of $\mathcal{C}$. Thus no interval overlaps $(L[\mathcal{C}]..R[\mathcal{C}])$, and therefore it is a strong interval. $\square$

LEMMA 5.7. *Let (R, L) be the canonical generator of a closed family $\mathcal{F}$, and let $\mathcal{C}$ be a nontrivial overlap class containing $(i_1..R[\mathcal{C}]), \dots, (i_k..R[\mathcal{C}])$ and $(L[\mathcal{C}]..j_1), \dots, (L[\mathcal{C}]..j_l)$, with $i_1 < \dots < i_k$ and $j_1 < \dots < j_l$. Then $k = l$, and for all $a \in (1..k)$, $(i_a..j_a)$ is a strong interval of $\mathcal{F}$.*

Proof. We first show the result for $k = l = 1$. In this case, there are only two intervals $(i_1..R[\mathcal{C}])$ and $(L[\mathcal{C}]..j_1)$, and they overlap to form interval $(i_1..j_1)$. To show that this interval is strong, suppose that it contains more than one element and that $(i..j)$ overlaps it with $i < i_1 \leq j < j_1$. Then $(L[j]..j)$ also overlaps $(i_1..R[\mathcal{C}])$, and the class must have more than two intervals. If $(i..j)$ overlaps $(i_1..j_1)$ with $i_1 < i \leq j_1 < j$, then we obtain a similar contradiction using interval $(i..R[i])$.

Now suppose that, up to index $a \geq 1$, $(i_a..j_a)$ is a strong interval of $\mathcal{F}$, $k > a$ if and only if $l > a$, and all intervals are consecutive; that is, $i_{b+1} = i_b + 1$ when $b < a$. We will show that these statements are true for $a + 1$.

First, note that the interval $(j_a + 1..R[\mathcal{C}])$ always belongs to $\mathcal{F}$, being the difference $(i_a..R[\mathcal{C}]) \setminus (L[\mathcal{C}]..j_a)$, and since the generator is commuting, this implies that $R[j_a + 1] = R[\mathcal{C}]$. If $k > a$, then $l > a$ since $(i_{a+1}..R[\mathcal{C}])$ does not overlap any interval already processed. If $l > a$, then $(L[\mathcal{C}]..j_{a+1})$ overlaps $(j_a + 1..R[\mathcal{C}])$ and both are in $\mathcal{C}$, implying that $k > a$ and $i_{a+1} = j_a + 1$, since the hypothesis that $(i_a..j_a)$ is strong

implies that $i_{a+1} > j_a$.

To finish the proof, we must show that $(i_{a+1}..j_{a+1})$ is strong. Suppose that $(i..j) \in \mathcal{F}$ overlaps $i_{a+1}..j_{a+1}$ and $i < i_{a+1} \leq j < j_{a+1}$; then $(L[j]..j)$ overlaps $(i_{a+1}..R[\mathcal{C}])$, thus it overlaps at least one member of $\mathcal{C}$, and therefore it belongs to $\mathcal{C}$. But $i_{a+1} \leq j < j_{a+1}$ and $i_{a+1} = j_a + 1$ implies that $j_a < j < j_{a+1}$, which contradicts the fact that the list of j indices is totally ordered.

On the other hand, if $i_{a+1} < i \leq j_{a+1} < j$, then $(i..R[i])$ belongs to $\mathcal{C}$, implying that $(L[\mathcal{C}]..i-1)$ is in $\mathcal{F}$, thus $(L[i-1]..i-1)$ is in $\mathcal{C}$. Since we have that $i_{a+1} - 1 < i - 1 < j_{a+1}$, thus $j_a < i - 1 < j_{a+1}$, $i - 1$ is again out of order in the list of j indices. $\square$

Let (R, L) be the canonical generator of a closed family $\mathcal{F}$, and define the set of *traces* of overlap classes as the set $\mathcal{T}$ of intervals containing the following:

- For each nontrivial overlap class $\mathcal{C} = \{(i_1..R[\mathcal{C}]), \dots, (i_k..R[\mathcal{C}]), (L[\mathcal{C}]..j_1), \dots, (L[\mathcal{C}]..j_k)\}$
 - $(L[\mathcal{C}]..R[\mathcal{C}])$,
 - $(i_a..j_a)$ for each $1 \leq a \leq k$.
- $(i..j)$ if $\mathcal{C} = \{(i..j)\}$ is trivial.

THEOREM 5.8. *The set $\mathcal{T}$ of traces of overlap classes of the canonical generator of a closed family $\mathcal{F}$ is equal to the family of strong intervals of $\mathcal{F}$.*

Proof. The fact that all intervals of $\mathcal{T}$ are strong is proved in Lemmas 5.4, 5.6, and 5.7.

Conversely, let $S = (i..j)$ be a strong interval. If either $R[i] = j$ or $L[j] = i$, then S is in the generator and forms a trivial overlap class. Otherwise $(i..R[i])$ and $(L[j]..j)$ overlap and thus belong to an overlap class $\mathcal{C}$. Using the same notation as above we have $i = i_a$ and $j = j_b$ for some a and b . We showed in the proof of Lemma 5.7 that we must have $a \leq b$. If $a < b$, S is overlapped by $(i_b..R[\mathcal{C}])$ and thus is not strong. Therefore $a = b$. $\square$

Theorem 5.8 implies that computing the strong intervals of $\mathcal{F}$ amounts to computing $\mathcal{T}$. This can be done by a classical parenthesis-matching algorithm using a stack since we proved that the left and right bounds of the strong intervals defined by an overlap class form a balanced expression. Namely, for nontrivial classes, the corresponding expression is

$$(\dots (\dots) (\dots) \dots (\dots) \dots)$$

and for trivial classes is $(\dots)$.

More formally, let (R, L) be the canonical generator. Consider the $4n$ bounds of intervals of the families $(i..R[i])$ and $(L[j]..j)$ for $i, j \in (1..n)$. Let $(a_1, \dots, a_{4n})$ be the list of these $4n$ bounds sorted in increasing order, with the left bounds placed before the right bounds when they are equal. For the example of Figure 3 this list is

$$(1, 1, 1, \bar{1}, 2, 2, \bar{2}, \bar{2}, 3, 3, 3, \bar{3}, \bar{3}, 4, \bar{4}, 5, 5, \bar{5}, \bar{5}, \bar{5}, 6, 6, \bar{6}, \bar{6}, 7, 7, \bar{7}, \bar{7}, 8, \bar{8}, \bar{8}),$$

where $\bar{i}$ denotes a right bound. Such a list can be constructed easily by scanning the two vectors R and L and by noting that each $i \in (1..n)$ is a left bound at least once and a right bound at least once. The following algorithm takes this sorted list as input and outputs the strong intervals sorted by increasing left bounds.

ALGORITHM 6. COMPUTATION OF THE STRONG INTERVALS.

S is a stack of bounds; *s* denotes the top of *S*.
 For *i* from 1 to $4n$
 If a_i is a left bound
 Push a_i on *S*
 Else
 Output $(s..a_i)$ (* Interval $(s..a_i)$ is strong *)
 Pop the top of *S*

PROPOSITION 5.9. *Given the ordered list $(a_1, \dots, a_{4n})$ of the $4n$ bounds of a canonical generator (R, L) , Algorithm 6 computes $\mathcal{T}$ in $O(n)$ time.*

Proof. Algorithm 6 will correctly match left and right bounds of a balanced expression. Since the left and right bounds of the strong intervals defined by each overlap class are balanced, and two overlap classes are either nested or disjoint, the list $(a_1, \dots, a_{4n})$ is balanced, and a correct matching of corresponding bounds will generate all the strong intervals. Since $2n$ intervals are identified by Algorithm 6, and the number of strong intervals is between $n + 1$ and $2n - 1$, some of them may be output several times. A suitable sorting algorithm, bucket-sort for example, will allow the identification of duplicates in $O(n)$ time. $\square$

5.3. From strong intervals to PQ -trees. By Proposition 5.3, the PQ -tree of a family $\mathcal{F}$ can be generated from its canonical generator by first computing the set of strong intervals of $\mathcal{F}$, ordering them into a tree structure, and finally labeling them as P or Q nodes.

Consider the example of Figure 4; Algorithm 6 produces the stack

$(1, 1, 1, 1, 1, \bar{1}, 2, 2, \bar{2}, 3, \bar{3}, \bar{3}, 4, \bar{4}, \bar{4}, 5, 5, 5, \bar{5}, 6, \bar{6}, \bar{6}, 7, 7, 7, \bar{7}, 8, \bar{8}, 9, \bar{9}, \bar{9}, 9, \bar{9})$

and outputs the set of strong intervals

$(1..1), (2..2), (3..3), (2..3), (4..4), (1..4), (1..4), (5..5), (6..6),$
 $(5..6), (5..6), (7..7), (8..8), (9..9), (7..9), (7..9), (1..9), (1..9).$

Each of these intervals corresponds to a node of the PQ -tree.

Given a proper commuting family of $n \leq m < 2n$ intervals, the following well-known algorithm computes its inclusion tree. To upgrade an inclusion tree to a PQ -tree, it is necessary to label the nodes as P or Q and to order the children of Q -nodes. Theorem 5.10 explains how to do these last tasks.

ALGORITHM 7. BUILDING THE INCLUSION TREE OF A PROPER COMMUTING FAMILY $\mathcal{F}$.

Bucket-sort in decreasing order the intervals of $\mathcal{F}$ according to their right bound
 Bucket-sort in increasing order the intervals of $\mathcal{F}$ according to their left bound
 Let $I_1..I_m$ be the list of sorted intervals
 $F \leftarrow I_1$ (* $I_1 = V$ is the root *)
 $k \leftarrow 2$
 While $k \leq m$
 If $I_k \subset F$
 $\text{Parent}(I_k) \leftarrow F$
 $F \leftarrow I_k$
 $k \leftarrow k + 1$
 else
 $F \leftarrow \text{Parent}(F)$

Let $I_1 = (l_1..r_1), \dots, I_k = (l_k..r_k)$ be the children of a Q node. For every i , $I_i \cup I_{i+1}$ is an interval of the family $\mathcal{F}$. Intervals I_i and I_{i+1} are disjoint, thus $r_i + 1 = l_{i+1}$ or $r_{i+1} = l_i + 1$. The *canonical PQ-tree* is the one where all Q nodes are sorted in increasing order, i.e., $r_i + 1 = l_{i+1}$.

THEOREM 5.10. *Let (R, L) be the canonical generator of a closed family $\mathcal{F}$. The canonical PQ-tree of $\mathcal{F}$ can be computed from (R, L) in $O(n)$ time.*

Proof. Algorithm 7 computes the inclusion tree in $O(n)$ time. The bucket sort is a stable linear-time sorting algorithm, thus intervals with the same left bound are sorted by their right bounds. The overall time complexity is obviously $O(n)$.

Given the inclusion tree of the strong intervals of a closed family $\mathcal{F}$, a PQ -tree of $\mathcal{F}$ can be built by labeling as P or Q the internal nodes of the inclusion tree and by ordering the children of the Q nodes. Fortunately, Algorithm 7 directly orders the children of every node (including the upcoming Q nodes) by increasing values of the first bounds of their frontiers. The resulting PQ -tree is therefore the canonical one. Thus, it is necessary only to label the nodes.

Each internal node of the inclusion tree has at least two children, since a node and its only child would have the same frontier. Every node with two children is labeled P . To test whether a node with at least three children is a P or Q node, it suffices to probe its two first children: if their union is an interval of the family $\mathcal{F}$, the node is labeled Q ; otherwise it is labeled P . This can be done in $O(1)$ time per node, using the generator. $\square$

5.4. From PQ-trees to canonical generators. Given a PQ -tree T , $\sigma(T)$ is the permutation of the leaves of T obtained by a left-to-right traversal of the tree. We may always assume that $\sigma(T) = \text{Id}_n$ by renaming the leaves of T . In this section we explain how to compute the canonical generator of the closed family represented by T .

Let N be a node with children $I_1, \dots, I_k$. Let l_i and r_i be the indices of the left and right bounds of I_i . Let $\text{imin}(N)$ and $\text{imax}(N)$ be the indices, respectively, of the minimum element of N and of the maximal element of N . Computing imin and imax can be done in $O(n)$ by a simple bottom-up traversal of T .

ALGORITHM 8. COMPUTING THE CANONICAL GENERATOR FROM A PQ -TREE.

For each internal node N of T taken bottom-up

 If N is a Q node

 For i from 1 to k

$L[\text{imax}(I_i)] \leftarrow \text{imin}(N)$

$R[\text{imin}(I_i)] \leftarrow \text{imax}(N)$

 Else

$R[\text{imin}(I_1)] \leftarrow \text{imax}(N)$

$L[\text{imax}(I_1)] \leftarrow \text{imin}(N)$

$R[\text{imin}(I_k)] \leftarrow \text{imax}(N)$

$L[\text{imax}(I_k)] \leftarrow \text{imin}(N)$

 For i from 2 to $k - 1$

$L[\text{imax}(I_i)] \leftarrow \text{imin}(I_i)$

$R[\text{imin}(I_i)] \leftarrow \text{imax}(I_i)$

PROPOSITION 5.11. *Algorithm 8 computes the canonical generator of a closed family in $O(n)$ time.*

Proof. The time complexity of the algorithm is obvious because the PQ -tree T has $O(n)$ nodes. Let us prove that the algorithm computes the canonical generator.

First, for all i , the value of $R[i]$ and $L[i]$ will be determined because, for the node N that bears the leaf i as j th node, $i = \text{imin}(I_j)$ and $i = \text{imax}(I_j)$.

When $R[\text{imin}(I_i)]$ is determined, $(\text{imin}(I_i)..R[\text{imin}(I_i)])$ is an interval of $\mathcal{F}$ because if N is a P node, then $(\text{imin}(I_i)..R[\text{imin}(I_i)]) = I_i$ and if N is a Q node, then $(\text{imin}(I_i)..R[\text{imin}(I_i)])$ is the union of all consecutive children of Q from i to k . The other cases are similar.

Let $R_t[1], \dots, R_t[i]$ be the values taken by the variable $R[i]$ during the execution of the algorithm. Since the nodes are visited bottom-up, the intervals $(i..R_t[i])$ are a nested sequence of intervals. We shall now prove that this sequence ends with the largest interval starting at i (resp., ending at j). Let us suppose that $(i..R_t[i]) \neq (i..R[i])$. Then $R_t[i] < R[i]$ since $(i..R_t[i])$ belongs to $\mathcal{F}$. Let us suppose $(i..R_t[i])$ is not strong. Then it is the union of children $I_a, \dots, I_b$ of some Q -node N . The algorithm sets $R_t[i]$ to the left bound of the last child of N . Since $(i..R[i])$ is larger, it overlaps the strong interval N , which is impossible. Now let us suppose that $(i..R_t[i])$ is strong. It corresponds to a node I of the PQ -tree whose parent is N . Either $(i..R[i])$ contains N , or N is a Q -node and $(i..R[i])$ is a union of children of N . In the first case, $i = \text{imin}(N)$ and thus $R_t[i] \geq \text{imax}(N)$, a contradiction. In the second case, $R_t[i] = \text{imax}(N)$, which also contradicts $R_t[i] < R[i] = \text{imax}(N)$.

We proved that the sequence of $R_t[i]$ converges toward $R[i]$ for all i . The case for $L_t[j]$ is similar. $\square$

6. Modular decomposition. Let $G = (V, E)$ be a directed, finite, loopless graph, with $|V| = n$ and $|E| = m$. Undirected graphs may be seen as symmetrical directed graphs in this context. A *module* is a subset M of V that behaves like a single vertex: for $x \notin M$ either there are $|M|$ arcs that join x to all vertices of M , or no arc joins x to M , and conversely either there are $|M|$ arcs that join all vertices of M to x , or no arc joins M to x . A *strong* module does not overlap any other module. There may be up to 2^n modules in a graph (in the complete graph for instance), but there are at most $O(n)$ strong modules, and the *modular decomposition tree* based on the strong modules' inclusion tree is sufficient to represent all modules [17]. The modular decomposition tree is indeed the PQ -tree of the family of modules.

Modular decomposition is the first step in many graph algorithms such as graph recognition (e.g., cographs, interval graphs, permutation graphs, and other classes of perfect graphs; see [4] for a survey) and transitive orientation computation [15].

Linear-time decomposition algorithms have been discovered [8, 15] but remain rather complex. Simpler algorithms work in two steps: computing a factorizing permutation and then building a tree representation on it. The first step was published in [11]. In this paper, we simplify the second step.

A *factorizing permutation* of a graph [6] is a permutation of the vertices of the graph in which every strong module of the graph is a *factor*, that is, an interval of the permutation. Since the strong modules are a commuting family, every graph admits a factorizing permutation. A factorizing permutation of a graph can be computed in linear time [11]. In the following we assume, without loss of generality, that the vertex-set V is the set $\{1, \dots, n\}$ and that the identity permutation is a factorizing permutation of the graph.

Given an interval $(u..v)$ of the factorizing permutation, a vertex $x \notin (u..v)$ is a *splitter* of the interval if there are between 1 and $v - u$ arcs going from x to $(u..v)$, or if there are between 1 and $v - u$ arcs going from $(u..v)$ to x . A *right-module* is an interval $(u..v)$ with no splitters greater than v . A *left-module* is an interval $(u..v)$

with no splitters smaller than u . An *interval-module* is an interval $(u..v)$ with no splitters. Clearly, interval-modules are modules. However, some modules are not interval-modules, but, according to the definition of a factorizing permutation, the strong modules of the graph are interval-modules. It is well known that modules behave like intervals: unions, intersections, or differences of two overlapping modules are modules. Thus we have the following proposition.

PROPOSITION 6.1 (see [17]). *The interval-modules of a factorizing permutation of a graph G are a closed family. The strong intervals of this family are exactly the strong modules of the graph G .*

DEFINITION 6.2. *For a vertex v let $R[v]$ be the greatest integer such that $(v..R[v])$ is a left-module and $L[v]$ the smallest integer such that $(L[v]..v)$ is a right-module.*

It can be proved that, for every $w \in (L[v]..v)$, $(w..v)$ is a right-module and, for every $w < L[v]$, $(w..v)$ is not a right-module. For this reason $(L[v]..v)$ is called the maximal right-module ending at v . In a similar way, we can define the maximal left-module beginning at v . We have the following proposition.

PROPOSITION 6.3. *The pair (R, L) is a commuting generator of the interval-modules' family.*

Proof. Interval $(u..v)$ is an interval-module if and only if $R[u] \geq v$ and $L[v] \leq u$, thus (R, L) is a generator. The family defined by R is commuting because if $(u..R[u])$ overlaps $(v..R[v])$, and if, without loss of generality, $u < v$, then $(u..R[v])$ is a left-module starting at u greater than the maximal left-module $(u..R[u])$, which is a contradiction. A similar argument shows that L also is commuting. $\square$

In order to compute the maximal right-strong modules, we use a simplified version of an algorithm due to Capelle and Habib [6]. The algorithm to compute the maximal left-modules is similar.

Let us consider the maximal right-module $(L[v]..v)$ ending at v . If $L[v] > 1$, then there exists an $x > v$ that splits $(L[v] - 1..v)$; otherwise this right-module would not be maximal, and x therefore splits $(L[v] - 1..L[v])$ but does not split $(y - 1, y)$ for all $L[v] < y \leq v$. Based on this observation, Capelle and Habib's algorithm proceeds in two steps. First, for every vertex v the *rightmost splitter* $s[v]$ is computed. It is the greatest vertex, if any, that splits the pair $(v - 1..v)$. Then a loop for v from n to 2 computes all the maximal right-modules $(L[x]..x)$ such that $v = L[x]$. Computing $s[v]$ can be done by a simultaneous scan of the adjacency lists of v and $v - 1$: the greatest element occurring in only one adjacency list is kept. This can be done in time proportional to the size of the adjacency lists. The computation of $s[v]$ for all v can therefore be done in $O(n + m)$ time, that is, linear in the size of the graph. The second step is Algorithm 9. It clearly runs in $O(n)$ time, and its correctness relies on the following invariant.

INVARIANT. *At step v , for all vertices x in the stack, $(v..x)$ is a right-module, and for all $x > v$ not in the stack, $L[v] > v$.*

Proof. The invariant is initially true. Every step maintains it: if $s[v]$ does not exist, then for all x in the stack $(v - 1..x)$ is a right-module, and $(v - 1..v)$ is also a right-module. And if $s[v]$ exists, (v) is the maximal right-module ending at v . For all $x < s[v]$, $(v - 1..x)$ is not a right-module, and $(v..x)$ is therefore the maximal right-module ending at x . For all $x \geq s[v]$, $(v - 1..x)$ is still a right-module, because $s[v]$ is the greatest of the splitters of $(v - 1, v)$. $\square$

We thus have the following theorem.

THEOREM 6.4. *Given a graph G and a factorizing permutation of G , it is possible to compute the modular decomposition tree of G in time $O(n + m)$ and in a simple way.*

ALGORITHM 9. COMPUTING ALL MAXIMAL RIGHT-MODULES GIVEN $s[v]$.

```

S is a stack of vertices; t denotes the top of S.
for v from n to 2
  if  $s[v]$  exists
     $L[v] \leftarrow v$ 
    While  $t < s[v]$ 
       $L[t] \leftarrow v$ 
      Pop the top of S
  else
    Push v on S

```

7. Conclusion. In the present work, we formalized two concepts about common intervals, namely generators and canonical representation, that proved to have important algorithmic implications. Indeed, the combinatorial properties of these objects, and in particular the different links between them, are central in the design and the analysis of the simple optimal algorithms for computing the common intervals of permutations we presented. It is important to highlight that our algorithms are really “optimal” since they are based on very elementary manipulations of stacks and arrays. This is, we believe, a significant improvement over the existing algorithms that are based on intricate data structures, both in terms of ease of implementation and time efficiency and in terms of understanding the underlying concepts [12, 19].

Moreover, we showed how, transposed in the more general context of modular decomposition of graphs, our results have a similar impact and lead to a significant simplification of some existing algorithms. Indeed, modular decomposition algorithms are quite complex algorithms, but using the factorizing permutation algorithm of [11] and the right-modules identification algorithm of section 6, a generator of the interval-modules can easily be computed in linear time; tools from section 5.2 can then be used to compute the strong interval-modules, that are also the strong modules, and the *PQ*-tree, called *modular decomposition tree* in this context.

REFERENCES

- [1] S. BÉRARD, A. BERGERON, AND C. CHAUVE, *Conserved structures in evolution scenarios*, in Comparative Genomics, RECOMB 2004 International Workshop, Lect. Notes in Bioinform. 3388, Springer-Verlag, Berlin, 2004, pp. 1–15.
- [2] A. BERGERON, C. CHAUVE, F. DE MONTGOLFIER, AND M. RAFFINOT, *Computing common intervals of K permutations, with applications to modular decomposition of graphs*, in Proceedings of the 13th Annual European Symposium on Algorithms (ESA), Lecture Notes in Comput. Sci. 3669, Springer-Verlag, Berlin, 2005, pp. 779–790.
- [3] S. BOOTH AND G. LUEKER, *Testing for the consecutive ones property, interval graphs, and graph planarity using PQ -trees algorithms*, J. Comput. System Sci., 13 (1976), pp. 335–379.
- [4] A. BRANDSTÄDT, V. B. LE, AND J. P. SPINRAD, *Graph Classes: A Survey*, SIAM Monogr. Discrete Math. Appl. 3, SIAM, Philadelphia, 1999.
- [5] B. M. BUI XUAN, M. HABIB, AND C. PAUL, *Revisiting T. Uno and M. Yagiura’s algorithm*, in Proceedings of the 16th International Symposium on Algorithms and Computation (ISAAC), Lecture Notes in Comput. Sci. 3827, Springer-Verlag, Berlin, 2005, pp. 146–155.
- [6] C. CAPELLE AND M. HABIB, *Graph decompositions and factorizing permutations*, in Proceedings of the Fifth Israel Symposium on Theory of Computing and Systems (ISTCS), IEEE Computer Society, New York, 1997, pp. 132–143.
- [7] M. CHEIN, M. HABIB, AND M. C. MAURER, *Partitive hypergraphs*, Discrete Math., 37 (1981), pp. 35–50.
- [8] A. Cournier and M. Habib, *A new linear algorithm for modular decomposition*, in Proceedings of the 19th International Colloquium of Trees in Algebra and Programming (CAAP), Lecture Notes in Comput. Sci. 787, Springer-Verlag, Berlin, 1994, pp. 68–84.

- [9] E. DAHLHAUS, J. GUSTEDT, AND R. M. MCCONNELL, *Efficient and practical algorithms for sequential modular decomposition*, J. Algorithms, 41 (2001), pp. 360–387.
- [10] M. FIGEAC AND J.-S. VARRÉ, *Sorting by reversals with common intervals*, in Proceedings of the 4th International Workshop of Algorithms in Bioinformatics (WABI), Lecture Notes in Comput. Sci. 3240, Springer-Verlag, Berlin, 2004, pp. 26–37.
- [11] M. HABIB, F. DE MONTGOLFIER, AND C. PAUL, *A simple linear-time modular decomposition algorithm for graphs, using order extension*, in Proceedings of the 9th Scandinavian Workshop on Algorithm Theory (SWAT), Lecture Notes in Comput. Sci. 3111, Springer-Verlag, Berlin, 2004, pp. 187–198.
- [12] S. HEBER AND J. STOYE, *Finding all common intervals of k permutations*, in Proceedings of the 12th Annual Symposium of Combinatorial Pattern Matching (CPM), Lecture Notes in Comput. Sci. 2089, Springer-Verlag, Berlin, 2001, pp. 207–218.
- [13] G. M. LANDAU, L. PARIDA, AND O. WEIMANN, *Gene proximity analysis across whole genomes via PQ trees*, J. Comput. Biol., 12 (2005), pp. 1289–1306.
- [14] R. M. MCCONNELL AND F. DE MONTGOLFIER, *Algebraic operations on PQ-trees and modular decomposition trees*, in Proceedings of the 31st International Workshop on Graph-Theoretic Concepts in Computer Science (WG), Lecture Notes in Comput. Sci. 3787, Springer-Verlag, Berlin, 2005, pp. 421–432.
- [15] R. M. MCCONNELL AND J. SPINRAD, *Linear-time modular decomposition and efficient transitive orientation of comparability graphs*, in Proceedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 1994, pp. 536–545.
- [16] R. M. MCCONNELL AND J. SPINRAD, *Ordered vertex partitioning*, Discrete Math. Theor. Comput. Sci., 4 (2000), pp. 45–60.
- [17] R. H. MÖHRING AND F. J. RADERMACHER, *Substitution decomposition for discrete structures and connections with combinatorial optimization*, Ann. Discrete Math., 19 (1984), pp. 257–356.
- [18] A. SCHRIJVER, *Combinatorial Optimization*, Springer-Verlag, Berlin, 2003.
- [19] T. UNO AND M. YAGIURA, *Fast algorithms to enumerate all common intervals of two permutations*, Algorithmica, 26 (2000), pp. 290–309.

STRUCTURE AND COLORING OF GRAPHS WITH ONLY SMALL ODD CYCLES*

SUSAN S. WANG[†]

Abstract. In this paper we investigate the role that odd cycles play in graph structure and in graph coloring. We focus on graphs with only small odd cycles and examine how the absence of large odd cycles affects graph structure. For a number of classes of such graphs, we provide full structural characterizations. Based on this structural analysis, we develop efficient and optimal graph coloring algorithms. For example, for graphs whose only odd cycles are 3-cycles and 5-cycles, we show how to color them optimally using 3, 4, 5, or 6 colors in $O(E)$ time.

Key words. structural characterization of graphs, odd cycles, graph coloring, chromatic number, algorithms

AMS subject classifications. 05C15, 05C38, 05C75, 05C85, 68Q25, 68R10, 68W40

DOI. 10.1137/S0895480197323883

1. Introduction. In this paper we investigate the role that odd cycles play in graph structure and in graph coloring. We focus on graphs with only small odd cycles—graphs with one odd cycle size k and graphs with two odd cycle sizes, 3 and 5. A theme of the paper is *structural analysis*. For certain classes of graphs, we are able to characterize the structure of graphs directly. For others, we first use breadth-first search to decompose the graphs into components and then analyze the structure of the graphs in terms of the components. We provide full structural characterizations for those biconnected graphs with minimum degree at least three, with at least seven vertices, and with only 3-cycles and 5-cycles in odd cycles; that is, 3-cycles and 5-cycles are the only odd cycles that are present in the graphs. Characterizing graph structures enables us to identify important graph properties and to solve many problems, including graph recognition and graph coloring.

To *color* a graph means to assign a color to each vertex of the graph so that no two adjacent vertices have the same color. The graph coloring problem is to determine the minimum number of colors—denoted by the chromatic number $\chi(G)$ —needed to color a given graph G . This problem is considered hard not only because it is hard to arrive at an exact solution, even when the graph under consideration is restricted, but also because it is hard to arrive at an approximately optimal solution. Graph coloring, in its general form, in which one asks whether a graph G is k -colorable, has been shown to be NP-complete by Karp [10]. It remains NP-complete in the special case when $k = 3$ [6], even if the graph is restricted to be 4-regular planar [2]. As regards nonapproximability, Lund and Yannakakis [12] proved that there exists a fixed $c > 0$ such that approximating the chromatic number to within a factor of n^c , where n is the number of vertices, is NP-hard. Feige and Kilian [5] improved this nonapproximability result by showing that the chromatic number is hard to approximate to within a factor

*Received by the editors July 2, 1997; accepted for publication (in revised form) February 19, 2008; published electronically June 11, 2008. This work was supported in part by ARO fellowship grant DAAG29-83-G-0110, an IBM postdoctoral fellowship, and a Mills faculty research grant.

<http://www.siam.org/journals/sidma/22-3/32388.html>

[†]Department of Mathematics and Computer Science, Mills College, Oakland, CA 94613 (wang@mills.edu).

of $n^{1-\epsilon}$, for any positive ϵ , under the assumption that $\text{NP} \neq \text{ZPP}$.¹ Khanna, Linial, and Safra [11] showed that coloring 3-colorable graphs with 4 colors is NP-hard.

Graph coloring is solvable in polynomial time for bipartite graphs ($k = 2$) but is NP-complete for nonbipartite graphs ($k \geq 3$). The property that distinguishes nonbipartite from bipartite graphs is the presence of odd cycles. In this paper we investigate the role that odd cycles play in graph coloring by focusing on a special class of graphs—graphs without large odd cycles. The author has also investigated the problem for graphs with only large odd cycles; see [17].

TABLE 1

Graphs with various structure, their chromatic numbers, and the number of colors used (optimal in all cases) via algorithms developed in this paper.

G is biconnected, nonbipartite, with:	$\chi(G)$	Number of colors used
One odd cycle size, 3, and		
No K_4	3	3
With K_4	4	4
One odd cycle size, k , $k \geq 5$	3	3
Two odd cycle sizes, 3 and 5, and		
No K_4 and no W_5	3	3
No K_4 and with W_5	4	4
With K_4 but no K_5	4	4
With K_5 but no K_6	5	5
With K_6	6	6

In [3], Bollobás and Erdős posed the following conjecture: If a graph G has s different odd cycle sizes, then $\chi(G) \leq 2s + 2$, with equality if and only if G contains a complete subgraph on $2s + 2$ vertices. Gyárfás proved this conjecture in [8] and a stronger theorem showing that every biconnected graph G with s odd cycle sizes is either K_{2s+2} or contains a vertex of degree at most $2s$, implying that $\chi(G)$ is at most $2s + 2$. In this paper, we shall provide structural characterizations of biconnected graphs for the case of $s = 1$, and for the case of $s = 2$ where the odd cycle sizes are 3 and 5. In particular, we shall provide full structural characterizations for graphs that are biconnected and with one odd cycle size, 3, and for graphs that are biconnected, with minimum degree at least three, with at least seven vertices, and with two odd cycle sizes, 3 and 5. Based on the structural analysis, we shall present linear-time algorithms that optimally color all such graphs. Table 1 summarizes the graph coloring results of this paper. All the algorithms are optimal, in terms of the number of colors used, and efficient— $O(E)$ in complexity. Some of the results of this paper constituted a part of author's doctoral dissertation in [17] and were briefly summarized in [13, p. 149] and in [14, p. 28].

The paper is organized as follows. In section 2, we provide some background and then derive some graph-theoretic bounds on cycle sizes and the chromatic number for graphs with only small odd cycles. In section 3, we focus on graphs with only one odd cycle size, say k , treating first the case $k = 3$ and then the case $k \geq 5$. In section 4, we focus on graphs with two odd cycle sizes, 3 and 5. In section 5, we pose some open problems.

2. Odd cycles, even cycles, and the chromatic number. In this paper $G = (V, E)$ denotes a simple, connected, finite, and undirected graph with $|V|$ vertices and $|E|$ edges. The *size* of a graph is the number of vertices in the graph. The *degree*

¹ZPP denotes the class of languages decidable by a random expected polynomial-time algorithm that makes no errors. See [9] for a precise definition.

of a vertex is the number of edges incident with it. The *minimum degree* of a graph is the minimum degree of vertices of the graph. We denote a *path* by the edges on the path: $v_1 - v_2 - \cdots - v_i$, where $v_1 - v_2$ represents an edge (a path of length one) from v_1 to v_2 . The *length* of a path P , $|P|$, is the number of edges on path P . If P is a path in G , and if v_1, v_2 are two vertices on the path, then $v_1 \xrightarrow{P} v_2$ denotes that part of path on P that goes from v_1 to v_2 . A complete graph on n vertices is denoted by K_n . An n -wheel, W_n , is an n -cycle together with another vertex that is adjacent to all vertices of the n -cycle. If H_1, H_2 are two subgraphs of G , each with at least one edge, then $H_1 \Delta H_2$ denotes the *symmetric difference* of H_1 and H_2 , that is, $H_1 \Delta H_2 = (H_1 \cup H_2) - (H_1 \cap H_2)$.

Breadth-first search is a procedure of visiting the vertices of a graph in a specific order. We begin the search procedure at a vertex r called the *root* (often the root is chosen arbitrarily). First, we search r , that is, we visit all neighbors of r , then we search all neighbors of r , and then we search all of neighbors of neighbors of r , and so on. Each vertex in G is searched exactly once, and the breadth-first search procedure halts after all vertices have been searched. It is clear that the breadth-first search procedure takes time $O(E)$, when the graph is given in terms of its adjacency list.

Let $BFS(G_r)$ denote the breadth-first search graph of G rooted at r ; that is, $BFS(G_r)$ is a leveled graph isomorphic to G . The *level* $l_r(v)$ of a vertex v of a graph is defined as follows: the level of the root of $BFS(G_r)$ is 0 ($l_r(r) = 0$); the level of all other vertices is defined inductively as follows: if vertex v is first visited through the edge (u, v) , then $l_r(v) = l_r(u) + 1$. A vertex u is a *parent* of a vertex v in $BFS(G_r)$ if $(u, v) \in E$ and $l_r(u) = l_r(v) - 1$. If u is a parent of v , then v is a *child* of u . A vertex u is an *ancestor* of a vertex v if either u is a parent of v , or u is an ancestor of a parent of v . A vertex w is a *common ancestor* of vertices u and v if w is an ancestor of both u and v . A vertex w is a *nearest-common ancestor* of vertices u and v if w is a common ancestor of u and v , and the distance between w and u ($l_r(u) - l_r(w)$) is no greater than that of any other common ancestor of u and v .

Here is a useful characterization of *biconnected* graphs.

THEOREM 1 (see [1]). *The following properties are equivalent in a graph G of size at least three:*

- (A) G is biconnected.
- (B) Any two vertices of G lie on a common simple cycle.
- (C) Every vertex and edge of G lies on a common simple cycle.
- (D) Any two edges of G lie on a common simple cycle.

In studying graphs with only small odd cycles, we are interested in knowing whether the size of a largest odd cycle determines or bounds the chromatic number. There are some interesting and rather curious results concerning this question. On the one hand, a $2p$ -chromatic graph—as exemplified by K_{2p} , the complete graph on $2p$ vertices—need not contain odd cycles of size $2p + 1$ or longer. On the other hand, Erdős and Hajnal [4] proved that every $(2p + 1)$ -chromatic graph must contain an odd cycle of size at least $2p + 1$.

THEOREM 2. *If G does not contain odd cycles of size more than $2p - 1$ ($p \geq 1$), then G is $2p$ -colorable, that is, $\chi(G) \leq 2p$.*

Hence, the size plus one of a largest odd cycle of a graph provides an upper bound for the chromatic number of the graph. The Erdős–Hajnal result is tight. Unfortunately their proof is nonconstructive.

One may wonder whether limiting the size of odd cycles affects the size of even cycles of a graph. It has been observed that if G is biconnected, nonbipartite, and does

not contain large odd cycles, then G does not contain large even cycles. Theorem 5 below describes this result more precisely. Before proving Theorem 5, we first establish two lemmas.

LEMMA 3. *Let G be a biconnected graph, and let D be a simple even cycle that edge-overlaps some simple odd cycle. Then there exists a simple odd cycle C such that $C\Delta D$ is a simple odd cycle.*

Proof. Consider the set of all simple odd cycles that edge-overlap D (this set is nonempty by assumption). In this set, choose a cycle C such that $|C - D|$ is minimal (a cardinality argument shows that this is possible). Now we claim that $C\Delta D$ is a simple odd cycle. First, $C\Delta D$ is odd, because $|C\Delta D| = |C| + |D| - 2|C \cap D|$. We use an argument by contradiction to show that $C\Delta D$ is a simple cycle. If $C\Delta D$ is not simple, then it is the disjoint union of simple cycles. Since $C\Delta D$ is odd, at least one of the simple cycles, say C' , is odd. Because $C' \subset C\Delta D$ and C is simple, C' edge-overlaps D . Since both C and C' edge-overlap D , and since $C' \subset C\Delta D$, it follows that $|C' - D| < |C - D|$, which contradicts the minimality of C . $\square$

LEMMA 4. *In a biconnected, nonbipartite graph G , every edge of G lies on a simple odd cycle.*

Proof. Since G is biconnected, every edge in G lies on a simple cycle (Theorem 1). Let e_1 be one of the edges under consideration, and suppose that e_1 lies on the simple cycle C . If C is odd, then we are done. Suppose C is even. Since G is nonbipartite, there exists a simple odd cycle D in G . Let e_2 be an edge on D . Since G is biconnected, there exists a simple cycle E that contains edges e_1 and e_2 (Theorem 1). If E is odd, then again we are done. Suppose E is even. Since E and D edge-overlap (in e_2) and D is odd, there exists a simple odd cycle F such that $F\Delta E$ is a simple odd cycle, by Lemma 3. Either e_1 lies on F or on $F\Delta E$. $\square$

THEOREM 5.²

(A) *In a biconnected, nonbipartite graph G , if there exists a simple even cycle of size $2m$, then there exists a simple odd cycle of size at least $m + 1$. Equivalently, the following holds:*

(B) *In every biconnected, nonbipartite graph G , we have*

$$l_o \leq l \leq 2(l_o - 1),$$

where l_o is the size of a maximum odd cycle and l is the size of a maximum cycle.

Proof. Let D be a simple even cycle of size $2m$ in G . Then D edge-overlaps some simple odd cycle, by Lemma 4. Hence, by Lemma 3, there is a simple odd cycle C such that $C\Delta D$ is a simple odd cycle. If $|C| \geq m + 1$, then we are done. Suppose $|C| \leq m$. Since D is a simple cycle, C is not contained in D . Hence, $|C \cap D| \leq m - 1$. It follows that the size of $C\Delta D$ is at least $m + 1$, because $|C\Delta D| = |C| + |D| - 2|C \cap D| \geq |D| - |C \cap D| \geq 2m - (m - 1) = m + 1$. This proves (A).

To establish the equivalence of (A) and (B), we first form the contrapositive of (A); it says that in every biconnected, nonbipartite graph, if $l_o < m + 1$, then $l_e < 2m$, where l_e is the size of a maximum even cycle. To show that (A) implies (B), set $m = l_o$. Then the hypothesis of the contrapositive is satisfied, so $l_e < 2l_o$. Since l_e is even, this means $l_e \leq 2(l_o - 1)$ ($l_o \geq 3$). Therefore, $l_o \leq l \leq 2(l_o - 1)$. To show that (B) implies (A), assume that (B) holds and that $l_o < m + 1$. By simple arithmetic, $2(l_o - 1) < 2m$, so $l < 2m$ by (B). Therefore, $l_e < 2m$. $\square$

Let G_i denote the graph obtained from G by iteratively deleting vertices of degree less than i . Thus, the minimum degree of G_i is at least i .

²Theorem 5 was found by the author and published in [17]. Subsequently a proof was published in [16].

LEMMA 6. *If we can k -color G_i in time T , then we can k -color G in time $T + O(E)$ for all $k \geq i$.*

*Proof.*³ Once G_i is k -colored, we can k -color G by adding the deleted vertices back to G_i , one at a time, in reverse order. Each time we add a vertex to G_i , its degree is less than i (due to the way G_i is created); therefore it can be (legally) colored with one of the k colors, since $k \geq i$. The above condition holds for all vertices added back to G_i . Hence, we can k -color G .

To create G_i , we use a queue to keep track of vertices of degree less than i . As we delete vertices of degree less than i , we keep them on a stack so that later we can color them in reverse order. In the process of coloring G , we go through the adjacency list of G a constant number of times: first to determine the initial degrees of the vertices; then to update the degrees as vertices are deleted; and finally to color the vertices when they are added back to G_i . Hence, the total running time of the algorithm is $T + O(E)$, where T is the amount of time it takes to color G_i . $\square$

3. Odd cycles = k -cycles. In studying graphs with only small odd cycles, we first focus on graphs with only one odd cycle size k , i.e., graphs with no odd cycles other than k -cycles. We show that the restriction of having only one odd cycle size in the context of biconnected graphs gives rise to tightly structured graphs. Being able to characterize the structure of graphs enables us to recognize the graphs and color them optimally.

3.1. $k = 3$.

LEMMA 7. *If a biconnected graph G contains no odd cycles other than 3-cycles, then G does not contain two edge-disjoint 3-cycles.*

Proof. Suppose that G contains two edge-disjoint 3-cycles, say $C_1 = v_1 - v_2 - v_3 - v_1$ and $C_2 = v_a - v_b - v_c - v_a$. Since G is biconnected, the edges (v_1, v_2) and (v_a, v_b) lie on a common simple cycle (Theorem 1), say C_3 . The size of C_3 is either three or four by Theorem 5. If the size of C_3 is three, then cycles C_1 and C_2 have a common vertex. Without loss of generality, assume that $v_2 = v_b$. The edge on C_3 that connects C_1 and C_2 must then be (v_1, v_a) , and therefore the cycle $v_1 - v_a - v_c - v_b - v_3 - v_1$ is a 5-cycle, which is a contradiction. Hence, the size of C_3 must be four. But then the cycle $C_1 \Delta C_3$ is a 5-cycle, which again is a contradiction. $\square$

THEOREM 8. *If G is a biconnected graph that contains no odd cycles other than 3-cycles, then G is either K_4 or belongs to the family of graphs of Figure 1. This family can be recognized in $O(E)$ time.*

Proof. First, we show that if G contains a 4-clique $K_4 = (v_1, v_2, v_3, v_4)$, then G coincides with K_4 . Argue by contradiction: suppose v_a is a vertex in G that is not in K_4 , and let (v_a, v_b) be an edge incident with v_a for some vertex v_b in G . From Lemma 4, we know that every edge in G lies on a 3-cycle. In particular, (v_a, v_b) lies on some 3-cycle, say $C = v_a - v_b - v_c - v_a$. According to Lemma 7, C shares an edge with every 3-cycle in K_4 . But this is not possible for any combination of v_b and v_c . Thus, no vertex $v_a \in G - K_4$ can exist; i.e., $G = K_4$.

Next we show that if G is not a 4-clique, then it must belong to the family of graphs illustrated in Figure 1. First, we know that G contains a 3-cycle, by assumption. Second, every edge of G lies on a 3-cycle, by Lemma 4. Third, since G does not contain a 4-clique, all 3-cycles in G must share the same edge, by Lemma 7. Thus, every vertex not on the shared edge is on some 3-cycle containing the shared edge. Fourth, if we delete the two vertices associated with the shared edge from the graph, then all the

³The algorithm presented in the proof is a well-known heuristic for graph coloring.



FIG. 1. A family of graphs containing only 3-cycles in odd cycles and no K_4 .

vertices remaining in the graph are independent. Indeed, if an edge exists in the remaining graph, then that edge must lie on a 3-cycle, and the 3-cycle must contain the shared edge, which is not possible (since the two edges have four distinct vertices).

Due to the restricted structure of graphs satisfying the hypotheses of Theorem 8, we can recognize such graphs in $O(E)$ time (using breadth-first search, for example). $\square$

THEOREM 9. *If G is a biconnected graph that contains no odd cycles other than 3-cycles, then G can be optimally colored in $O(E)$ time.*

Proof. We have established in Theorem 8 that G is either K_4 or a member of the family of graphs shown in Figure 1. In the first case, the chromatic number of G is 4; in the second case, it is 3. Observe that all vertices except the two on the shared edge in Figure 1 have degree two. This observation implies that a biconnected graph with minimum degree at least three and containing no odd cycles other than 3-cycles must coincide with K_4 . Thus, one way of optimally coloring G is as follows. First, iteratively delete vertices of degree less than three from G to obtain G_3 . If G_3 is nonempty, then G must be K_4 and we can 4-color it in $O(E)$ time. If G_3 is empty, then G is 3-colorable in $O(E)$ time, by Lemma 6. $\square$

The proof of Theorem 9 shows that if a graph G contains no odd cycles other than 3-cycles, then G is perfect;⁴ that is, the chromatic number is the same as the clique number for G and every induced subgraph of G .

3.2. $k \geq 5$. All graphs that contain only k -cycles in odd cycles for some fixed $k \geq 5$, can be 3-colored in time $O(E)$. To prove this assertion, we first recall the following theorem.

THEOREM 10 (see [16]). *Every biconnected, nonbipartite graph of size at least seven and with minimum degree at least three contains an odd cycle with at least two diagonals.*

THEOREM 11. *If G is a biconnected graph that contains only k -cycles in odd cycles for some fixed $k \geq 5$, then G can be 3-colored in $O(E)$ time.*

Proof. We iteratively delete vertices of degree at most two from G to obtain biconnected components each with minimum degree at least three. All of these biconnected components G' must be bipartite. Indeed, for $|V(G')| < 7$, this can be established without difficulty, by a case argument, using the fact that any odd cycle has to be a 5-cycle by our assumption and that G' has minimum degree three. For $|V(G')| \geq 7$,

⁴Graph coloring for perfect graphs is solvable in polynomial time using the ellipsoid method for linear programming [7].

we argue by contradiction. If G' were nonbipartite, then by Theorem 10, it would contain an odd cycle with at least two diagonals. This would force G' , and hence also G , to have odd cycles of different sizes, which contradicts the hypotheses of the theorem. All of the biconnected bipartite components G' can be colored in $O(E')$ time. Combining Lemma 6 with the fact that biconnected components of a graph can be found in $O(E)$ time [15], we see that G can be 3-colored in $O(E)$ time. $\square$

The preceding argument actually proves that biconnected graphs containing only k -cycles in odd cycles, for some fixed $k \geq 5$, with minimum degree at least three do not exist.

The author has recently obtained a complete characterization of graphs that are biconnected and contain only k -cycles in odd cycles for all $k \geq 5$. This result will be published elsewhere.

4. Odd cycles = 3-cycles + 5-cycles. We now move on to consider graphs with two odd cycle sizes, more specifically, to graphs with only 3-cycles and 5-cycles in odd cycles. We focus on graphs that are biconnected with minimum degree at least three and with at least seven vertices. We analyze their structures and present full structural characterizations of these graphs in Theorems 13, 15, and 16 and an $O(E)$ -time graph recognition algorithm in Algorithm 1. With regards to coloring, graphs with only 3-cycles and 5-cycles in odd cycles may contain a K_6 —a 6-clique—as a subgraph and hence may require at least six colors. However, as Theorem 2 indicates, they are 6-colorable. Based on the structural analysis, we develop an algorithm, Algorithm 2, that optimally colors all graphs with only 3-cycles and 5-cycles in odd cycles in $O(E)$ time.

From the graph structure perspective, it is of interest to focus on the quintessential aspects of the structure that exhibit the characteristics of the graph. For that reason we focus on biconnected graphs, as every edge of a biconnected graph lies on an odd cycle (Lemma 4). In addition, all graphs can be decomposed into biconnected components in $O(E)$ time [15]. In connection with graph coloring, since all nonbipartite graphs require at least three colors and in consideration of Lemma 6, we chose to examine graphs with minimum degree at least three.

LEMMA 12. *If G is a biconnected graph with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices, then G contains no K_5 .*

Proof. Suppose for a contradiction that G contains a K_5 . Since G is of size at least seven, $G - K_5$ is of size at least two. If $G - K_5$ contains an edge, say e_1 , then since G is biconnected, there exists a simple cycle C that contains the edge e_1 and an edge of K_5 (Theorem 1). Because K_5 is a complete graph on five vertices, it is not hard to construct an odd cycle of size at least seven from C and K_5 . Thus, $G - K_5$ must be an independent set. Let v_1 and v_2 be two vertices in $G - K_5$. Because the degrees of v_1 and v_2 are both at least three, each of v_1, v_2 must be adjacent to three vertices of K_5 . It is not hard to construct a 7-cycle from v_1, v_2 , and K_5 . $\square$

Since the graphs that satisfy the hypotheses of Lemma 12 do not contain a K_5 , we may partition them into two classes: those that contain a K_4 and those that do not. A full structural characterization of the first class is given in Theorem 13, and a full structural characterization of the second class is given in Theorems 15 and 16.

THEOREM 13. *Assume G is a biconnected graph with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices. If G contains a K_4 , then G belongs to one of the two families of graphs shown in Figure 2.*

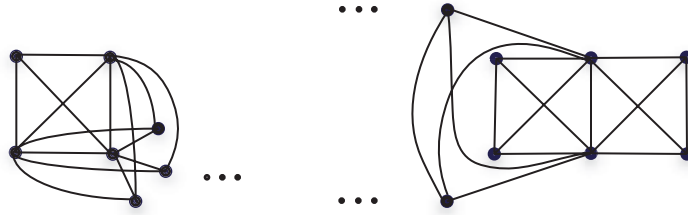


FIG. 2. Two families of biconnected graphs with two odd cycle sizes, 3 and 5, with minimum degree at least three, with at least seven vertices, and containing a K_4 .

Proof. Suppose that we have a 4-clique in G , say $K_4 = (v_1, v_2, v_3, v_4)$. First, we argue that all vertices in $G - K_4$ are distance one from K_4 . Suppose not. Without loss of generality, let v_a be a vertex that is (shortest) distance two or more from K_4 . Since G is biconnected, there exists a simple cycle C that contains v_a and an edge of K_4 (Theorem 1). Because v_a is distance two or more from K_4 , the total length of paths on C that connect v_a with K_4 is at least four. Thus, from C and K_4 we can easily construct a simple odd cycle of size at least seven, which is a contradiction.

Second, we argue that $G - K_4$ does not contain any paths of length two. Suppose, to the contrary, that it does; say $v_a - v_b - v_c$ is a path of length two in $G - K_4$. Both v_a and v_c must be adjacent to exactly one and the same vertex of K_4 . If not, then, in view of the first observation, v_a and v_c are adjacent to distinct vertices of K_4 , say v_1 and v_2 , respectively. Then $v_a - v_b - v_c - v_2 - v_3 - v_4 - v_1 - v_a$ is a 7-cycle, which is not possible. Hence, we can and do assume that v_a and v_c are adjacent to the same vertex v_1 . Since the degree of v_c is at least three, at least one of the following is true: (a) v_c is adjacent to v_a ; or (b) v_c is adjacent to some vertex v_d such that v_d is not in K_4 (and $v_d \neq v_a$). In either case, we claim that v_b is adjacent to v_1 but not to v_2, v_3 , or v_4 . According to our first observation, v_b is adjacent to a vertex of K_4 . In case (a), if v_b were adjacent to v_2, v_3 , or v_4 , then it would be easy to construct a 7-cycle. (For example, if v_b were adjacent to v_2 , then we would obtain the 7-cycle $v_a - v_b - v_2 - v_3 - v_4 - v_1 - v_c - v_a$.) In case (b), using similar arguments as those used to show that v_a and v_c are adjacent to exactly one and the same vertex (v_1) of K_4 , we conclude that the same is true for v_a and v_d , and for v_b and v_d . Hence, we have established our claim. The above argument shows that, for any path of length two or more outside of K_4 , the vertices on the path are all adjacent to exactly one and the same vertex of K_4 . Hence, any such path must enter K_4 (in either direction) through the same vertex of K_4 . But G is biconnected, and we have a contradiction.

From the first observation we know that every vertex in $G - K_4$ is adjacent to at least one vertex of K_4 . From the second observation we know that each connected component of $G - K_4$ is either a vertex or a path of length one. Note that each vertex in $G - K_4$ is adjacent to at most three vertices of K_4 , because G does not contain a K_5 (Lemma 12). Since the minimum degree of G is at least three, an independent vertex in $G - K_4$ is adjacent to exactly three vertices of K_4 , and a vertex on a path of length one in $G - K_4$ is adjacent to either two or three vertices of K_4 . Hence,

$G - K_4$ cannot contain an independent vertex and a path of length one; otherwise it would be easy to construct a 7-cycle. Thus, the connected components of $G - K_4$ are either all independent vertices or all paths of length one. A similar argument shows that if $G - K_4$ contains all independent vertices, then all the independent vertices are adjacent to the same three vertices of K_4 (recall we have assumed that G has at least seven vertices); and if $G - K_4$ contains paths of length one, then all vertices in $G - K_4$ are adjacent to the same two vertices of K_4 (otherwise we could construct a 7-cycle). This completely characterizes the structure of G : it belongs to one of the two families of graphs shown in Figure 2. $\square$

Next we shall consider graphs that do not contain K_4 . We use breadth-first search to decompose the graphs and obtain additional structural results.

LEMMA 14. *Assume G is a biconnected graph with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices. If G contains no K_4 , then a breadth-first-search graph of G — $BFS(G)$ —has the following properties:*

- (A) $BFS(G)$ contains at most four levels.
- (B) If level 1 is not independent, then $BFS(G)$ contains at most three levels.
- (C) Levels 3 and 4 of $BFS(G)$ are independent.
- (D) Levels 1 and 2 of $BFS(G)$ are bipartite; moreover, level 2 is a forest.

Proof. We shall prove each part individually.

- (A) $BFS(G)$ contains at most four levels.

Suppose, to the contrary, that $BFS(G_r)$ has more than four levels, where r is the root. Let u be a vertex at level 5. Since G is biconnected, u and r lie on a common simple cycle (Theorem 1), say C . The size of C is at least ten because u is at level 5 and r is at level 0. The size of C cannot be odd, by our assumption about odd cycle sizes, and it cannot be even (of size ten or more), by Theorem 5. Thus, we have a contradiction.

- (B) If level 1 is not independent, then $BFS(G)$ contains at most three levels.

Assume, to the contrary, that $BFS(G_r)$ has more than three levels, where r is a vertex on a 3-cycle. Let u be a vertex at level 4 and (v_a, v_b) a cross edge at level 1 forming a 3-cycle with r . Since G is biconnected, u and (v_a, v_b) lie on a common simple cycle (Theorem 1), say C . The size of C is at least seven because u is at level 4 and (v_a, v_b) is at level 1. The size of C cannot be odd, by our assumption on odd cycle sizes, and it cannot be even (of size ten or more), by Theorem 5. Thus, the size of C must be eight. Since C contains the edge (v_a, v_b) and is of size eight, C does not contain the root r . Because edge (v_a, v_b) forms a 3-cycle with the root, we can construct a 9-cycle from C by replacing edge (v_a, v_b) by edges (v_a, r) and (r, v_b) . This contradicts the assumption about odd cycle sizes.

- (C) Levels 3 and 4 of $BFS(G)$ are independent.

We first show that level 4 is independent, using an argument by contradiction. Suppose (v_1, v_2) is a cross edge at level 4 and r is the root. Since G is biconnected, r and (v_1, v_2) lie on a common simple cycle (Theorem 1), say C . The size of C is at least nine. The size of C cannot be odd, by our assumption about odd cycle sizes, and it cannot be even (of size ten or more), by Theorem 5. The desired contradiction has arrived.

To show that level 3 is independent, we consider two cases: (i) root r is on a 3-cycle, and (ii) root r is not on a 3-cycle.

Case (i). Root r is on a 3-cycle; that is, there are cross edges at level 1.

Suppose, for a contradiction, that level 3 is not independent. Let (v_1, v_2) be a cross edge at level 3, and let (v_a, v_b) be a cross edge at level 1 that forms a 3-cycle

with r . Since G is biconnected, (v_a, v_b) and (v_1, v_2) lie on a common simple cycle (Theorem 1), say C . The size of C is at least six, which means that it must be even, by our assumption about odd cycle sizes. If C does not contain the root, then we construct an odd cycle of size at least seven from C by replacing edge (v_a, v_b) by edges (v_a, r) and (r, v_b) . If C contains the root, then the size of C must be eight, by Theorem 5. Thus, C contains exactly two cross edges, including (v_a, v_b) , and either (r, v_a) or (r, v_b) . Without loss of generality, assume that C contains (r, v_a) . Then we obtain a 7-cycle from C by replacing edges (r, v_a) and (v_a, v_b) by edge (r, v_b) . This contradicts the assumption about odd cycle sizes.

Case (ii). Root r is not on a 3-cycle; that is, level 1 is independent.

Suppose the contrary, and let (v_1, v_2) be a cross edge at level 3. Since G is biconnected, r and (v_1, v_2) lie on a common simple cycle (Theorem 1), say C . The size of C is at least seven. The size of C cannot be odd, by our assumption about odd cycle sizes, and it cannot be even of size ten or more, by Theorem 5. Thus, the size of C must be eight, which means that C contains two cross edges. Since level 1 is independent by assumption and level 4 is independent as shown above, the second cross edge of C lies either at level 3 or at level 2. The argument now breaks down into several subcases.

Suppose, for a contradiction, that the second cross edge of C lies at level 3; without loss of generality, assume that it is (v_2, v_3) . Then cycle C is of the form $r - g_1 - p_1 - v_1 - v_2 - v_3 - p_3 - g_3 - r$, where g_1, g_3 and p_1, p_3 are vertices at levels 1 and 2, respectively. If p_1 is a parent of v_2 , then we arrive at a 7-cycle, namely $r - g_1 - p_1 - v_2 - v_3 - p_3 - g_3 - r$. Similarly, if p_3 is a parent of v_2 , then we have the 7-cycle $r - g_1 - p_1 - v_1 - v_2 - p_3 - g_3 - r$. Thus, v_2 must have a parent other than p_1, p_3 ; call it p_2 . Exploring the possible parents of p_2 , we find a similar situation: If g_1 is a parent of p_2 , we get the 7-cycle $r - g_1 - p_2 - v_2 - v_3 - p_3 - g_3 - r$; if g_3 is a parent of p_2 , we obtain the 7-cycle $r - g_1 - p_1 - v_1 - v_2 - p_2 - g_3 - r$. Thus, p_2 must have a parent other than g_1, g_3 ; call it g_2 . But now we have the 7-cycle $r - g_1 - p_1 - v_1 - v_2 - p_2 - g_2 - r$, which yields the desired contradiction.

Next suppose the second cross edge of C lies at level 2. Without loss of generality, assume that (p_2, p_3) is the second cross edge and cycle C is of the form $r - g_1 - p_1 - v_1 - v_2 - p_2 - p_3 - g_3 - r$, where g_1, g_3 and p_1, p_2, p_3 are vertices at levels 1 and 2, respectively. Note that g_1 is the unique grandparent of v_1 and v_2 (if not, then it would be easy to construct a 7-cycle that contains edge (v_1, v_2) and root r). Thus, g_1 is the unique parent of p_1 and p_2 . For the rest of the analysis, we shall focus on vertices v_1, v_2 , and p_1 . We shall use the fact that all of them have degree at least three to analyze possible structures that can exist at each vertex and show that no combination of these structures is possible.

Focusing on v_1 . Since the degree of v_1 is at least three, at least one of the following is true: (a) v_1 has another parent besides p_1 ; (b) v_1 is incident with another cross edge besides (v_1, v_2) ; (c) v_1 has a child. Suppose (a) holds. Observe that v_1 does not have p_2 as a parent; otherwise we would obtain the 7-cycle $r - g_1 - p_1 - v_1 - p_2 - p_3 - g_3 - r$. Nor does v_1 have p_3 as a parent, because v_1 has a unique grandparent. It is possible that v_1 has p_4 as a parent for some $p_4 \neq p_1, p_2, p_3$. Since g_1 is the unique grandparent of v_1 , it is the unique parent of p_4 . We call this Structure (1) (see Figure 3).

We now take up the case when (b) holds; that is, suppose v_1 is incident with cross edge (v_1, v_3) , $v_3 \neq v_2$. Note that v_3 does not have p_3 as a parent; otherwise we would get the 7-cycle $r - g_1 - p_1 - v_1 - v_3 - p_3 - g_3 - r$. Nor does v_3 have a parent, say p , that is different from p_2 and p_3 ; otherwise it would be easy to construct an odd cycle

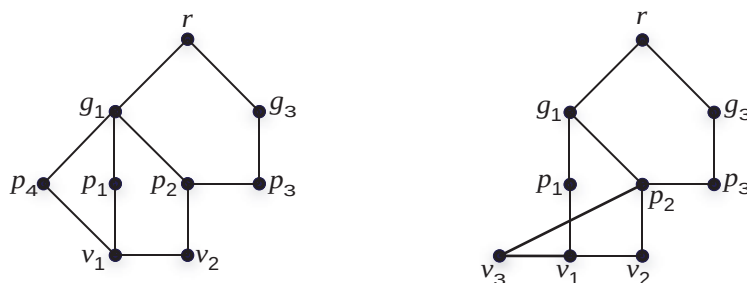


FIG. 3. Structure (1).

Structure (2).

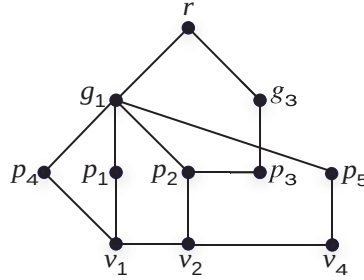
of size at least seven from path $p - v_3 - v_1 - v_2 - p_2 - p_3$ and two paths from p and p_3 to their nearest-common ancestor in $BFS(G_r)$. It is possible that v_3 has p_2 as a parent. We call this Structure (2) (see Figure 3).

We now turn to the case when (c) holds; that is, suppose v_1 has a child, say u . Since G is biconnected, (r, g_1) and (v_1, u) lie on a common simple cycle (Theorem 1), say C_2 . In view of our assumption about odd cycle sizes and Theorem 5, the size of C_2 is eight, because r is at level 0 and u is at level 4. Thus, C_2 does not contain cross edges. Combining this observation with the fact that g_1 is the unique grandparent of v_1 , we see that C_2 must contain a path—call it P_1 —of length two from v_1 to g_1 . This implies that C_2 does not contain p_2 or v_2 ; if it did, then the fact that g_1 is the unique parent of p_2 and the unique grandparent of v_2 would imply that cycle C_2 is closed at g_1 instead of at r . We can replace P_1 by $g_1 - p_2 - v_2 - v_1$ in C_2 and arrive at a 9-cycle, which is a contradiction.

Thus, if we focus on v_1 , Structures (1) and (2) are possible.

Focusing on v_2 . Since the degree of v_2 is at least three, at least one of the following is true: (a) v_2 has another parent besides p_2 ; (b) v_2 is incident with another cross edge besides (v_1, v_2) ; (c) v_2 has a child. Suppose (a) holds. Note that v_2 does not have p_3 as a parent, since v_2 has a unique grandparent. Nor does v_2 have a parent p_i different from p_2 and p_3 , since if it did, then p_i would have g_1 as its parent (because v_2 has a unique grandparent), and we would have the 7-cycle $r - g_1 - p_i - v_2 - p_2 - p_3 - g_3 - r$. Thus, v_2 does not have any other parent besides p_2 .

Next suppose (b) holds; that is, suppose v_2 is incident with cross edge (v_2, v_4) for some $v_4 \neq v_1$. Observe that v_4 does not have p_2 as a parent; otherwise we would obtain the 9-cycle $r - g_1 - p_1 - v_1 - v_2 - v_4 - p_2 - p_3 - g_3 - r$. It follows that g_1 is the unique grandparent of v_4 ; if not, then it would be easy to construct a 7-cycle that contains edge (v_2, v_4) and the root r . This implies that v_4 does not have p_3 as a parent. Also, v_4 does not have p_1 as a parent; if it did, then in combination with Structure (1), we would have the 7-cycle $g_1 - p_4 - v_1 - p_1 - v_4 - v_2 - p_2 - g_1$; and in combination with Structure (2), we would have the 7-cycle $g_1 - p_1 - v_4 - v_2 - v_1 - v_3 - p_2 - g_1$. The only possibility left is that v_4 has a parent p_5 that is different from p_1, p_2 , and p_3 . Note that g_1 is the unique parent of p_5 . Having p_5 as a parent for v_4 is not compatible with Structure (2); if it were compatible, we would arrive at the 7-cycle

FIG. 4. *Structure (3).*

$g_1 - p_2 - v_3 - v_1 - v_2 - v_4 - p_5 - g_1$. It is possible for v_4 to have p_5 as a parent, in combination with Structure (1), and this leads to Structure (3) (see Figure 4).

We now turn to the case when (c) holds; that is, suppose v_2 has a child. Arguments analogous to those which showed that v_1 does not have any children now show that v_2 does not have any children.

Hence, if we focus on both v_1 and v_2 , we see that Structure (2) is not possible, but Structure (3), which includes Structure (1) as a subgraph, is possible.

Focusing on p_1 . We shall now base our discussions on Structure (3). Since the degree of p_1 is at least three, at least one of the following is true: (a) p_1 has another parent besides g_1 ; (b) p_1 is incident with a cross edge; (c) p_1 has another child besides v_1 . Possibility (a) cannot hold because p_1 has a unique parent, namely g_1 . Nor can (b) hold. Indeed, suppose, to the contrary, that p_1 is incident with a cross edge. If p_1 is adjacent to p_2 , we have the 7-cycle $g_1 - p_2 - p_1 - v_1 - v_2 - v_4 - p_5 - g_1$. If p_1 is adjacent to p_3 , we get the 7-cycle $g_1 - p_1 - p_3 - p_2 - v_2 - v_4 - p_5 - g_1$. If p_1 is adjacent to some p_j different from p_2 and p_3 via a cross edge, then it is easy to construct an odd cycle of size at least seven from path $p_j - p_1 - v_1 - v_2 - p_2 - p_3$ and two paths from p_j and p_3 to their nearest-common ancestor in $BFS(G_r)$. Thus, possibility (b) leads to a contradiction in every case.

We now turn to possibility (c). Suppose p_1 has a child, say v_i , that is different from v_1 . Note that $v_i \neq v_2$, since v_2 does not have another parent besides p_2 . Also $v_i \neq v_4$; otherwise we would obtain the 7-cycle $g_1 - p_4 - v_1 - p_1 - v_4 - v_2 - p_2 - g_1$. Since the degree of v_i is at least three, v_i is adjacent to at least two other vertices. We consider three possible adjacency structures: (i) v_i has another parent; (ii) v_i is incident with a cross edge; (iii) v_i has a child. We proceed to show that none of the possibilities can actually occur. As regards (i), observe that v_i does not have p_2 as a parent; otherwise we would have the 7-cycle $r - g_1 - p_1 - v_i - p_2 - p_3 - g_3 - r$. Nor does v_i have p_k as a parent for some $p_k \neq p_1, p_2$, as it would be easy to construct an odd cycle of size at least seven from the path $p_k - v_i - p_1 - v_1 - v_2 - p_2$ and the two paths from p_k and p_2 to their nearest-common ancestor in $BFS(G_r)$. Thus, v_i has a unique parent, namely p_1 .

To see that (ii) cannot hold; suppose v_i is incident with a cross edge (v_i, v_j) . Since G is biconnected, r and (v_i, v_j) lie on a common simple cycle (Theorem 1), say C_3 .

Using similar arguments as those presented in earlier paragraphs, we see that C_3 is of size eight and contains two cross edges with the second cross edge lying at level 2. This implies that v_i and v_j have distinct parents and a unique grandparent g_1 . Since v_i has p_1 as its unique parent, v_j must have a parent other than p_1 , say p_m . This leads to the 7-cycle $g_1 - p_4 - v_1 - p_1 - v_i - v_j - p_m - g_1$, which is impossible.

To dispose of possibility (iii), assume v_i has a child, say u . Since G is biconnected, r and (v_i, u) lie on a common simple cycle (Theorem 1), say C_4 . Again using arguments similar to those presented in earlier paragraphs, we see that C_4 is of size eight. Hence, C_4 does not contain cross edges. Observe that C_4 contains the path $r - g_1 - p_1 - v_i - u$, because p_1 is the unique parent of v_i , and g_1 is the unique parent of p_1 . In addition, C_4 does not contain vertices v_1 , v_2 , or p_2 (if it did and since C_4 does not contain cross edges, C_4 would have to contain a path from v_1 , v_2 , or p_2 to g_1 , closing the cycle C_4 at g_1 instead of at r). We can replace edge $g_1 - p_1$ on C_4 by path $g_1 - p_2 - v_2 - v_1 - p_1$ to obtain an 11-cycle, which is not possible.

The assumption that (c) holds led to three subcases (i), (ii), (iii), each of which has been shown to be impossible. Thus, (c) cannot hold. Therefore, Structure (3) is not possible.

In summary, in consideration of the degree requirements of v_1, v_2 , and p_1 , none of the structures in support of the degree requirements is possible. Therefore, level 3 is independent.

(D) Levels 1 and 2 of $BFS(G)$ are bipartite; moreover, level 2 is a forest.

To show that levels 1 and 2 are bipartite, we break up the proof into two parts, showing, respectively, that levels 1 and 2 do not contain 5-cycles and 3-cycles. We then proceed to show that level 2 is a forest (i.e., a collection of trees).

Claim D.1. Levels 1 and 2 do not contain 5-cycles.

First, we show that if there exists a 5-cycle at level 2, then the vertices on the 5-cycle have the same unique parent. Label the vertices on the 5-cycle v_1 through v_5 , and assume, to the contrary, that two vertices, say v_1 and v_3 , have distinct parents. There then exist two disjoint paths from v_1 and v_3 to their common ancestor r , which is two levels above them. The cycle that consists of path $v_1 - v_5 - v_4 - v_3$ plus the two disjoint paths from v_1 and v_3 to r is a 7-cycle, which is not possible.

Second, we show that G cannot contain a W_5 as a subgraph. (Recall that W_5 is a 5-cycle together with another vertex that is adjacent to all vertices of the 5-cycle.) Suppose, to the contrary, that there is such a W_5 . Since G contains at least seven vertices, there is a vertex u in G that is not part of W_5 . Because G is biconnected, u and an edge of W_5 lie on a common simple cycle (Theorem 1), say C . Due to the structure and symmetry of W_5 , there exist paths of length two, three, four, and five between any two vertices of W_5 . Consequently, it is not hard to construct an odd cycle of size at least seven from C and W_5 , which is not possible.

The above two observations immediately imply Claim D.1.

Claim D.2. Levels 1 and 2 do not contain 3-cycles.

Since G does not contain K_4 as a subgraph, level 1 does not contain 3-cycles. Suppose, for a contradiction, that level 2 contains a 3-cycle $C_1 = v_1 - v_2 - v_3 - v_1$. The three vertices on the cycle cannot have the same parent; otherwise we would obtain a K_4 . Thus, two of the vertices, say v_1 and v_2 , have different parents, say p_1 and p_2 , respectively.

Since the degree of r is at least three, r must have a third child, say p_3 . Since G is biconnected, (r, p_3) and (v_1, v_2) lie on a common simple cycle (Theorem 1), say C_2 . Let P_1 be the path on C_2 containing edge (r, p_3) with two endpoints v_a and v_b on C_1

such that if we traverse C_2 from edge (r, p_3) in one direction, we enter C_1 for the first time at v_a ; and if we traverse C_2 from edge (r, p_3) in the other direction, we enter C_1 for the first time at v_b . Since P_1 contains (r, p_3) and two vertices on C_1 , its length must be four or more. In fact, its length must be exactly four, because if it were five or more, then we could easily construct an odd cycle of size at least seven from P_1 and C_1 . This means that p_3 must be adjacent to a vertex on C_1 .

We claim that p_3 is not incident with a cross edge. Indeed, suppose, to the contrary, that p_3 is adjacent to p_i via a cross edge. Note that $p_i \neq p_1$; otherwise we would have the 7-cycle $r - p_3 - p_1 - v_1 - v_3 - v_2 - p_2 - r$. An analogous argument shows that $p_i \neq p_2$. Recall that p_3 is adjacent to a vertex on C_1 (i.e., either v_1, v_2 , or v_3). If p_3 is adjacent to v_1 , then we obtain the 7-cycle $r - p_i - p_3 - v_1 - v_3 - v_2 - p_2 - r$. Symmetrically, if p_3 is adjacent to v_2 , then we arrive at another 7-cycle. If p_3 is adjacent to v_3 , then we get the 7-cycle $r - p_i - p_3 - v_3 - v_2 - v_1 - p_1 - r$. Thus, all possibilities lead to contradictions.

Since the degree of p_3 is at least three and p_3 is not incident with a cross edge, p_3 must have at least two children. From the above discussion, it follows that at least one child of p_3 belongs to the set $\{v_1, v_2, v_3\}$. We claim that exactly one child of p_3 belongs to this set. Suppose, to the contrary, that two children of p_3 belong to the set $\{v_1, v_2, v_3\}$. If v_1 and v_3 are children of p_3 , then we get the 7-cycle $r - p_1 - v_1 - p_3 - v_3 - v_2 - p_2 - r$. Similarly, if v_2 and v_3 are children of p_3 , we again obtain a 7-cycle. The third possibility is that v_1 and v_2 are children of p_3 . In this case, if p_1 is a parent of v_3 , we have the 7-cycle $r - p_1 - v_3 - v_1 - p_3 - v_2 - p_2 - r$; and if p_j is a parent of v_3 for some $p_j \neq p_1, p_3$, we arrive at the 7-cycle $r - p_1 - v_1 - p_3 - v_2 - v_3 - p_j - r$. In all cases, we arrive at a contradiction.

Because p_3 has at least two children, and exactly one of them belongs to the set $\{v_1, v_2, v_3\}$, it follows that at least one child of p_3 is distinct from v_1, v_2, v_3 . Let v_4 be such a child. Since the degree of v_4 is at least three, v_4 must be adjacent to at least two other vertices. Assume w is one of them (thus $w \neq p_3$). Certainly w is different from each of v_1, v_2, v_3, p_1, p_2 ; otherwise it would be easy to construct a 7-cycle. Also note that w is not adjacent to vertices v_1, v_2, v_3 ; otherwise it would again be easy to construct a 7-cycle (for example, if w were adjacent to v_2 , then we would have the 7-cycle $r - p_3 - v_4 - w - v_2 - v_1 - p_1 - r$). Since G is biconnected, (v_4, w) and (v_1, v_2) lie on a common simple cycle (Theorem 1), say C_3 . Let P_2 be the path on C_3 containing edge (v_4, w) with two endpoints v_c and v_d on C_1 such that if we traverse C_3 from edge (v_4, w) in one direction, we enter C_1 for the first time at v_c ; and if we traverse C_3 from edge (v_4, w) in the other direction, we enter C_1 for the first time at v_d . The length of P_2 cannot be less than five; this would imply that at least one of v_4, w is adjacent to v_1, v_2 , or v_3 . The length of P_2 cannot be five or more; then we could easily construct an odd cycle of size at least seven from P_2 and C_1 (for example, if the length of P_2 were five and $v_c = v_1, v_d = v_2$, then we would have the 7-cycle $v_1 \xrightarrow{P_2} v_2 - v_3 - v_1$). The desired contradiction has arrived. We conclude that level 2 does not contain 3-cycles.

Claim D.3. Level 2 is a forest.

Assume, to the contrary, that level 2 is not a forest; then it contains even cycles of size four, six, or eight (Theorem 5). Suppose it contains a 4-cycle, say $v_1 - v_2 - v_3 - v_4 - v_1$. If two adjacent vertices v_i and v_{i+1} , where $v_{4+1} = v_1$, have different parents, then we obtain a 7-cycle; for example, if v_2 and v_3 have different parents p_2 and p_3 , respectively, then we get the 7-cycle $v_3 - v_4 - v_1 - v_2 - p_2 - r - p_3 - v_3$. Thus, all vertices v_1 through v_4 have exactly one and the same parent, say p_1 . Since G is

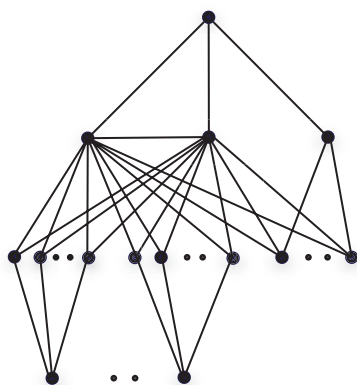


FIG. 5. A family of biconnected graphs with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices, where root r is on a 3-cycle, and $BFS(G_r)$ has three levels.

biconnected, (r, p_1) and (v_1, v_2) lie on a common simple cycle (Theorem 1), say C . It is easy to construct a 7-cycle from C and vertices v_1 through v_4 and vertex p_1 , and this leads to a contradiction. Using similar arguments, we can also show that level 2 does not contain 6-cycles or 8-cycles. Therefore, level 2 is a forest. $\square$

THEOREM 15. Assume G is a biconnected graph with two odd cycle sizes, 3 and 5, with minimum degree at least three, with at least seven vertices, and with no K_4 . If a breadth-first-search graph of G from a vertex on a 3-cycle contains three levels, then G belongs to the family of graphs shown in Figure 5.

Proof. Let r be a vertex on a 3-cycle such that $BFS(G_r)$ has three levels. Let (v_1, v_2) be a cross edge at level 1 of $BFS(G_r)$ forming a 3-cycle with r . Let u_1 be a vertex at level 3 of $BFS(G_r)$. Since the degree of u_1 is at least three and since $BFS(G_r)$ has at most three levels, with the third level being independent (properties (B) and (C) of Lemma 14), the vertex u_1 has at least three parents. We prove that $BFS(G_r)$ has the following properties, which reflect the structure shown in Figure 5:

- (A) Level 1 contains exactly one cross edge.
- (B) Level 2 is independent.
- (C) All vertices at level 2 have degree exactly three.
- (D) Level 1 contains exactly three vertices.

We begin with the first property.

- (A) Level 1 contains exactly one cross edge.

This property is proved by establishing two claims.

Claim A.1. There exists a 5-cycle containing u_1 and (v_1, v_2) .

Since G is biconnected, u_1 and (v_1, v_2) lie on a common simple cycle (Theorem 1), say C . By Theorem 5 and our assumptions that (v_1, v_2) is a cross edge at level 1 and u_1 a vertex at level 3, the size of C is at least five and at most eight. We first argue by contradiction that the size of C cannot be eight. Suppose, to the contrary, that it is. Then r must lie on C ; otherwise we could replace edge (v_1, v_2) by edges

(v_1, r) and (r, v_2) to obtain a 9-cycle. Also, C contains neither (r, v_1) nor (r, v_2) ; if C contained, say (r, v_1) , then we could replace edges (r, v_1) and (v_1, v_2) by edge (r, v_2) and arrive at a 7-cycle. In order for C to be of size eight, to contain u_1, r , and (v_1, v_2) , and not to contain (r, v_1) or (r, v_2) , it must contain a path from cross edge (v_1, v_2) to another cross edge and then to r . Without loss of generality, assume that C contains cross edge (v_2, v_3) . We then get a 7-cycle by replacing edges (v_2, v_3) and (v_3, r) by edge (v_2, r) in C . We conclude that the size of C is less than eight. This size cannot be seven, by our assumption on odd cycle sizes. Nor can it be six; otherwise we could replace edge (v_1, v_2) by edges (v_1, r) and (r, v_2) to get a 7-cycle. Thus, the size of C must be five.

Claim A.2. v_1 and v_2 are the only grandparents of u_1 .

From Claim A.1 we see that v_1 and v_2 are two grandparents of u_1 . Suppose, for a contradiction, that v_3 is a third grandparent of u_1 through child p_1 (i.e., p_1 is a child of v_3 and a parent of u_1). Since there exists a 5-cycle containing u_1 and (v_1, v_2) (Claim A.1), there exists a parent of u_1 , say p_2 , distinct from p_1 , such that p_2 is a child of v_1 or v_2 . If p_2 is a child of v_1 , then we get the 7-cycle $u_1 - p_1 - v_3 - r - v_2 - v_1 - p_2 - u_1$; a similar contradiction arises if p_2 is a child of v_2 . Thus, no third grandparent of u_1 can exist.

Now we are ready to show that level 1 contains exactly one cross edge. Suppose that (v_a, v_b) is any cross edge at level 1. Using analogous arguments as those in Claims A.1 and A.2, we conclude that v_a and v_b are the only grandparents of u_1 . Thus, the edge (v_a, v_b) coincides with the edge (v_1, v_2) .

(B) Level 2 is independent.

Consider a vertex p at level 2; there are three possibilities:

- (1) p is a child of v_1 or v_2 and a parent of a vertex at level 3;
- (2) p is a child of v_1 or v_2 but not a parent of a vertex at level 3;
- (3) p is neither a child of v_1 nor of v_2 .

We shall prove that, in each case, p is not incident with a cross edge.

Claim B.1. If a vertex at level 2 is a child of v_1 or v_2 and a parent of a vertex at level 3, then it is not incident with a cross edge.

Suppose the contrary. Without loss of generality, let p_1 be a parent of u_1 and a child of v_1 such that p_1 is incident with a cross edge (p_1, p_a) . Since there exists a 5-cycle containing u_1 and (v_1, v_2) (Claim A.1), there exists a vertex p_2 , distinct from p_1 , such that p_2 is a parent of u_1 and a child of v_2 . If p_a has a parent other than v_2 , say v_p , then we have the 7-cycle $u_1 - p_1 - p_a - v_p - r - v_2 - p_2 - u_1$. Hence, the only parent of p_a is v_2 . Since u_1 has at least three parents, let p_3 be a third parent of u_1 . Note that v_1 is not a parent of p_3 ; otherwise we would get the 7-cycle $u_1 - p_1 - p_a - v_2 - r - v_1 - p_3 - u_1$. Hence, the only parent of p_3 is v_2 by Claim A.2. Similarly, the only parent of p_2 is v_2 . It is not possible that $p_a = p_2$ (or p_3); otherwise we would arrive at the 7-cycle $u_1 - p_2 - p_1 - v_1 - r - v_2 - p_3 - u_1$.

We now focus on possible structures surrounding p_2 . Since the degree of p_2 is at least three (note p_2 is adjacent to v_2 and u_1), p_2 must either (i) be incident with a cross edge or (ii) have another child besides u_1 . Suppose, for a contradiction, that (i) holds; say p_2 is incident with cross edge (p_2, p_b) . Observe that $p_b \neq p_a$; otherwise we would have the 7-cycle $u_1 - p_2 - p_a - v_2 - r - v_1 - p_1 - u_1$. Either v_1 is a parent of p_b or it is not. If it is, then we obtain the 7-cycle $u_1 - p_2 - p_b - v_1 - v_2 - p_a - p_1 - u_1$. If it is not, then p_b has a parent v_3 different from v_1 , and we get the 7-cycle $u_1 - p_2 - p_b - v_3 - r - v_1 - p_1 - u_1$. Thus, both possibilities lead to a contradiction, so (i) cannot hold. Suppose now that (ii) holds. Let u_2 , different from u_1 , be a child of p_2 . Since u_2 is a vertex at level 3,

it has the same structure as u_1 ; that is, u_2 has at least three parents, and v_1 and v_2 are the only grandparents of u_2 . Without loss of generality, assume p_c is a parent of u_2 and a child of v_1 . Observe that $p_c \neq p_a$, since p_a has exactly one parent, namely v_2 . If $p_c = p_1$, then we arrive at the 7-cycle $u_1 - p_2 - u_2 - p_1 - v_1 - v_2 - p_3 - u_1$; if $p_c \neq p_1$, we get the 7-cycle $u_2 - p_c - v_1 - p_1 - p_a - v_2 - p_2 - u_2$. Both possibilities lead to contradictions, so (ii) cannot hold either. Thus, p_2 cannot have degree at least three, and therefore, by our assumption, it cannot exist. It follows that p_1 cannot be incident with a cross edge.

Claim B.2. If a vertex at level 2 is a child of v_1 or v_2 but not a parent of a vertex at level 3, then it is not incident with a cross edge.

Suppose, to the contrary, that p_a is a child of v_2 , has no children, and is incident with cross edge (p_a, p_b) . Since there exists a 5-cycle containing u_1 and (v_1, v_2) (Claim A.1), there exist vertices p_1, p_2 such that they are parents of u_1 and children of v_1, v_2 , respectively. Let v_i be a parent of p_b . If $v_i = v_1$, we get the 7-cycle $u_1 - p_2 - v_2 - p_a - p_b - v_1 - p_1 - u_1$. If $v_i \neq v_2$, then since $v_i \neq v_1$, we obtain the 9-cycle $u_1 - p_2 - v_2 - p_a - p_b - v_i - r - v_1 - p_1 - u_1$. Thus, $v_i = v_2$; in other words, p_b has v_2 as its only parent. The same arguments apply to p_a ; that is, p_a has v_2 as its only parent. Having v_2 as a parent, and being incident with a cross edge, p_b has no children by Claim B.1. Since p_a and p_b do not have children, they must be incident with other cross edges, because the degrees of both vertices are at least three. Assume that p_a is incident with cross edge (p_a, p_c) and p_b incident with cross edge (p_b, p_d) . Using the same arguments as before, we conclude that v_2 is the unique parent of p_c and p_d , that p_c and p_d do not have any children, and that p_c and p_d are incident with other cross edges. Repeat this argument to conclude that all paths from p_a to r must include v_2 , which contradicts the assumption that G is biconnected.

Claim B.3. All vertices at level 2 are children of v_1 or v_2 .

Suppose, for a contradiction, that there exists a vertex p at level 2 such that p is a child of neither v_1 nor v_2 . Since G is biconnected, u_1 and p lie on a common simple cycle (Theorem 1), say C . Since the parents of u_1 are not incident with cross edges (Claim B.1) and since the grandchildren of v_1 and v_2 do not have a third grandparent (Claim A.2), C must contain a path from u_1 to v_1 and another path from u_1 to v_2 . Thus, C consists of a path from u_1 to v_1 plus a path from v_1 to p plus a path from p to v_2 plus the path v_2 to u_1 . Since neither v_1 nor v_2 is a parent of p , since (v_1, v_2) is the only cross edge at level 1 (property (A)), and since a child of v_1 or v_2 is not incident with a cross edge (Claim B.2), the length of the path from v_1 (v_2) to p in C is at least three. Hence, the size of C is at least ten, which contradicts the fact the size of C is at most eight (Theorem 5).

It follows from Claim B.3 that case (3)— p is neither a child of v_1 nor v_2 —does not exist. Combine Claims B.1 and B.2 to conclude that level 2 is independent.

(C) All vertices at level 2 have degree exactly three.

We consider two cases: those vertices at level 2 with children and those without.

Claim C.1. If a vertex at level 2 has children, then it has exactly one child and exactly two parents, namely v_1 and v_2 .

Suppose, for a contradiction, that a vertex at level 2 has more than one child. Without loss of generality, let p_1 be a vertex at level 2 such that p_1 is a parent of u_1 and a child of v_1 (Claim A.2), and let u_2 be a second child of p_1 . Because there exists a 5-cycle containing u_1 and (v_1, v_2) (Claim A.1), there exists a vertex p_2 , distinct from p_1 , such that p_2 is a parent of u_1 and a child of v_2 . Since u_2 is a vertex at level 3, u_2 has the same structure as u_1 . In particular, u_2 has at least

three parents, so one of them is different from p_1 and p_2 ; say it is p_3 . Note that p_3 has v_1 or v_2 as a parent, by Claim A.2. If v_1 were a parent of p_3 , then we would obtain the 7-cycle $u_2 - p_3 - v_1 - v_2 - p_2 - u_1 - p_1 - u_2$. Hence, v_2 is the only parent of p_3 . A symmetric argument shows that v_2 is the only parent of p_2 . Since the degree of p_2 is at least three and p_2 is not incident with a cross edge (Claim B.1), p_2 must have another child, say u_3 . Observe that $u_3 \neq u_2$; otherwise we would arrive at the 7-cycle $u_2 - p_2 - u_1 - p_1 - v_1 - v_2 - p_3 - u_2$. Since u_3 is a vertex at level 3, u_3 has the same structure as u_1 . In particular, u_3 has at least three parents, so one of them is different from p_1 and p_2 ; say it is p_4 . Note that $p_4 \neq p_3$; otherwise we would get the 7-cycle $u_3 - p_3 - u_2 - p_1 - v_1 - v_2 - p_2 - u_3$. By Claim A.2, p_4 has v_1 or v_2 as a parent. If v_1 were a parent of p_4 , then we would get the 9-cycle $u_3 - p_4 - v_1 - v_2 - p_3 - u_2 - p_1 - u_1 - p_2 - u_3$; if v_2 were a parent of p_4 , then we would obtain the 7-cycle $u_3 - p_4 - v_2 - v_1 - p_1 - u_1 - p_2 - u_3$. These contradictions imply that the degree of p_2 cannot be at least three. Consequently, the vertex p_1 cannot have more than one child. In other words, if a vertex p at level 2 has children, then it has just one child, say u . Since p has degree at least three and is not incident with a cross edge (Claim B.1), it must have at least two parents, and these parents are the grandparents of u . Apply Claim A.2 (with “ u_1 ” replaced by “ u ”) to conclude that v_1 and v_2 are the only parents of p .

Claim C.2. If a vertex at level 2 has no children, then it has exactly three parents, including v_1 and v_2 .

Let p_1 be a vertex at level 2 such that p_1 has no children. Since p_1 has degree at least three, is not incident with a cross edges (Claim B.2), and has no children, it must have at least three parents, and therefore one of them, say v_3 , is different from v_1 and v_2 . Since the degree of v_3 is at least three and (v_1, v_2) is the only cross edge at level 1 (property (A)), v_3 must have another child, say p_2 . From Claim B.3 we know that v_1 or v_2 is a parent of p_2 . Without loss of generality, suppose it is v_1 . It is not possible for p_1 to have parent v_4 that is different from v_1, v_2 , and v_3 ; that would yield the 7-cycle $r - v_4 - p_1 - v_3 - p_2 - v_1 - v_2 - r$. Therefore, the parents of p_1 are exactly v_1, v_2 , and v_3 .

(D) Level 1 contains exactly three vertices.

Since r has degree at least three, it has at least three children: v_1, v_2 , and, say, v_3 . From Claim C.1, we know that v_3 has no grandchildren. We claim that r must have exactly three children. If r had a fourth child, say v_4 , then the structure of v_4 and its children must be similar to that of v_3 and its children (see Claim C.2); namely the children of v_4 have exactly three parents: v_1, v_2 , and v_4 . This would lead to the 7-cycle $r - v_3 - p_1 - v_1 - v_2 - p_2 - v_4 - r$, where p_1, p_2 are children of v_3, v_4 , respectively.

We are now in a position to show that $BFS(G_r)$ belongs to the family of graphs shown in Figure 5. Level 1 contains exactly three vertices (property (D))— v_1, v_2 , and, say, v_3 . Since the degree of v_3 is at least three and (v_1, v_2) is the only cross edge at level 1 (property (A)), v_3 has at least two children, each of which has degree exactly three with v_1, v_2 , and v_3 as its parents (Claim C.2). All the other vertices on level 2 have degree exactly three—one child and the two parents v_1 and v_2 (Claim C.1). We have completely characterized the structure of G when a BFS graph of G from a vertex on a 3-cycle contains three levels: it belongs to the family of graphs shown in Figure 5. $\square$

THEOREM 16. Assume G is a biconnected graph with two odd cycle sizes, 3 and 5, with minimum degree at least three, with at least seven vertices, and with no K_4 . If

a breadth-first-search graph of G from a vertex on a 3-cycle contains two levels, then G belongs to one of the families of graphs shown in Figures 6–10.

Proof. Let r be a vertex on a 3-cycle such that $BFS(G_r)$ has two levels. We shall prove that $BFS(G_r)$ has the following properties:

- (A) Level 2 does not contain paths of length two or more.
- (B) If level 2 contains paths of length one, then the graph belongs to the family of graphs shown in Figure 6.
- (C) Level 1 does not contain paths of length three or more.
- (D) If level 1 contains paths of length two and level 2 is independent, then the graph belongs to one of the families of graphs shown in Figures 7 and 8.
- (E) If level 1 contains paths of length one only and level 2 is independent, then the graph belongs to one of the families of graphs shown in Figures 9 and 10.

We begin with the first property.

- (A) Level 2 does not contain paths of length two or more.

We proceed by eliminating the possibilities of paths of length five or more and then length four, length three, and finally length two.

Claim A.1. Level 2 does not contain paths of length five or more.

Let P_1 be a maximum-length path at level 2 and m be its length. Since the degree of every vertex is at least three, since $BFS(G_r)$ contains 2 levels, and since level 2 contains no cycles (property (D) of Lemma 14), each of the endpoints of P_1 must have at least two parents. Thus, there exists a cycle in $BFS(G_r)$ that contains the path P_1 and the root r and has length $m + 4$. From Theorem 5 we know that maximum cycle length in G is eight. Therefore, m is at most four.

Claim A.2. Level 2 does not contain paths of length four.

Suppose, for a contradiction, that level 2 contains a path of length four, say $v_1 - v_2 - v_3 - v_4 - v_5$. Argue as in the proof of Claim A.1 and use Claim A.1 to conclude that v_1 and v_5 must have at least two parents. Let p_1 be a parent of v_1 , and let p_2 , distinct from p_1 , be a parent of v_5 . If v_2 had a parent different from p_2 , say p_3 , we would get a 7-cycle $r - p_3 - v_2 - v_3 - v_4 - v_5 - p_2 - r$. Thus, the only parent of v_2 is p_2 . Similarly, the only parent of v_4 is p_1 . The second parent of v_1 must be p_2 ; otherwise if p_4 were a parent of v_1 , then we would have the 7-cycle $r - p_4 - v_1 - p_1 - v_4 - v_5 - p_2 - r$. Similarly, the second parent of v_5 must be p_1 . But then we arrive at the 7-cycle $p_1 - v_5 - v_4 - v_3 - v_2 - p_2 - v_1 - p_1$, which is a contradiction.

Claim A.3. Level 2 does not contain paths of length three.

Suppose, for a contradiction, that level 2 contains a path of length three, say $v_1 - v_2 - v_3 - v_4$. As before, and in view of Claim A.2, v_1 and v_4 must have at least two parents. If p_1 is a parent of v_1 , and p_2 , distinct from p_1 , is a parent of v_4 , then $r - p_1 - v_1 - v_2 - v_3 - v_4 - p_2 - r$ is a 7-cycle, which is not possible.

Claim A.4. Level 2 does not contain paths of length two.

Suppose, for a contradiction, that level 2 contains a path of length two, say $v_1 - v_2 - v_3$. As before, and in view of Claim A.3, v_1 and v_3 must have at least two parents. We now consider three cases: (i) v_1 and v_3 share two or more parents, (ii) v_1 and v_3 share exactly one parent, and (iii) v_1 and v_3 share no parents, and we show that in each case we have a contradiction.

Case (i). v_1 and v_3 share two or more parents.

Let p_1 and p_2 be two parents that v_1 and v_3 share. If v_2 had a parent other than p_1, p_2 , say p_3 , then we would arrive at the 7-cycle $r - p_3 - v_2 - v_1 - p_2 - v_3 - p_1 - r$. Thus, v_2 must have p_1 and/or p_2 as parents. Since the degree of r is at least three, r

has a third child, say p_4 . Because G is biconnected, there exists a simple cycle C that contains (r, p_4) and p_1 (Theorem 1). Due to the multiple number of edges that exist among vertices p_1, p_2, v_1, v_2, v_3 , it is not hard to use these edges and the cycle C to build an odd cycle of size seven or greater, leading to a contradiction.

Case (ii). v_1 and v_3 share exactly one parent.

Let p_1 be the parent that v_1 and v_3 share, and let p_2 and p_3 be the second parents of v_1 and v_3 , respectively. We now focus on possible parents of v_2 and show that in each case we reach a contradiction. If p_1 is a parent of v_2 , we get the 7-cycle $r - p_2 - v_1 - p_1 - v_2 - v_3 - p_3 - r$. If p_2 is a parent of v_2 , we arrive at the 7-cycle $r - p_2 - v_2 - v_1 - p_1 - v_3 - p_3 - r$. If some vertex p_4 , different from p_1 and p_2 , is a parent of v_2 , then we obtain the 7-cycle $r - p_4 - v_2 - v_3 - p_1 - v_1 - p_2 - r$.

Case (iii). v_1 and v_3 share no parents.

Since v_1 and v_3 each has at least two parents, let p_1 and p_2 be parents of v_1 , and let p_3 and p_4 be parents of v_3 . Note that none of p_1, p_2, p_3 , and p_4 can be a parent of v_2 ; otherwise we would get a 7-cycle (for instance, if p_1 were a parent of v_2 , then we would have the 7-cycle $r - p_2 - v_1 - p_1 - v_2 - v_3 - p_3 - r$). Thus, a parent of v_2 , say p_5 , is different from p_1, p_2, p_3, p_4 . Since the degree of p_1 is at least three, p_1 is either incident with a cross edge or has another child. If p_1 were incident with cross edge (p_1, p_6) , we would arrive at the 7-cycle $r - p_6 - p_1 - v_1 - v_2 - v_3 - p_3$ (or p_4) $- r$. Thus, p_1 must have another child, say v_4 . Observe that $v_4 \neq v_2$ because v_2 does not have p_1 as a parent, and $v_4 \neq v_3$ because v_1 and v_3 share no parents. In fact, v_4 has exactly one parent, namely p_1 ; if v_4 had another parent, say p_7 , then we would have the 7-cycle $r - p_7 - v_4 - p_1 - v_1 - v_2 - p_5 - r$ (in the case where $p_7 = p_5$ we would get the 7-cycle $r - p_1 - v_4 - p_5 - v_2 - v_3 - p_3 - r$). Since the degree of v_4 is at least three and since v_4 has exactly one parent, v_4 must be incident with two cross edges, say (v_4, v_5) and (v_4, v_6) . Now $v_5 - v_4 - v_6$ is a path of length two on level 2, so the argument at the beginning of Claim A.4 shows that v_5 must have at least two parents. Let p_8 be a parent of v_5 such that $p_8 \neq p_1$. If $p_8 = p_2$, we arrive at the 9-cycle $r - p_2 - v_5 - v_4 - p_1 - v_1 - v_2 - v_3 - p_3 - r$; if not, we obtain the 7-cycle $r - p_8 - v_5 - v_4 - p_1 - v_1 - p_2 - r$. Both of these situations are impossible.

(B) If level 2 contains paths of length one, then the graph belongs to the family of graphs shown in Figure 6.

Suppose that $v_1 - v_2$ is a path of length one at level 2. The usual argument and property (A) show that v_1 and v_2 must each have at least two parents. We now consider three cases: (i) v_1 and v_2 share no parents, (ii) v_1 and v_2 share exactly one parent, and (iii) v_1 and v_2 share two or more parents, and we show that cases (i) and (ii) are not possible, and case (iii) leads to the family of graphs shown in Figure 6.

Case (i). v_1 and v_2 share no parents.

Let p_1, p_2 be parents of v_1 , and let p_3, p_4 be parents of v_2 . Since the degree of p_1 is at least three, p_1 must either have another child or be incident with a cross edge. Suppose p_1 has a child v_3 ; clearly, $v_3 \neq v_2$ because v_1 and v_2 share no parents. Since the degree of v_3 is at least three and level 2 does not contain paths of length two (property (A)), v_3 must have a second parent; call it p_5 . But this leads to the 7-cycle $r - p_5 - v_3 - p_1 - v_1 - v_2 - p_3$ (or p_4) $- r$. Thus, p_1 cannot have another child beside v_1 , so it must be incident with a cross edge, say (p_1, p_6) . A similar argument shows that all parents of v_1 and v_2 have exactly one child and are incident with cross edges. In particular, this is true of p_3 ; say (p_3, p_7) is the cross edge. If p_6 is a parent of v_2 , we have the 7-cycle $r - p_2 - v_1 - p_1 - p_6 - v_2 - p_4 - r$; if not, we arrive at the 7-cycle

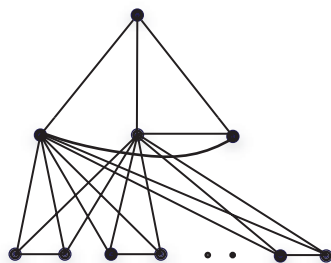


FIG. 6. A family of biconnected graphs with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices, where root r is on a 3-cycle, $BFS(G_r)$ has two levels, and level 2 is not independent.

$r - p_6 - p_1 - v_1 - v_2 - p_3 - p_7 - r$. We conclude that case (i) leads to a contradiction, so it is not possible.

Case (ii). v_1 and v_2 share exactly one parent.

Let p_1 be the parent that v_1 and v_2 share, and let p_2 and p_3 be second parents of v_1 and v_2 , respectively. Since the degree of p_2 is at least three, p_2 must either have another child or be incident with a cross edge. Suppose p_2 has a child v_3 ; clearly, $v_3 \neq v_2$ because v_1 and v_2 share exactly one parent. Since the degree of v_3 is at least three and level 2 does not contain paths of length two (property (A)), v_3 must have a second parent; call it p_4 . But this leads to the 7-cycle $r - p_4 - v_3 - p_2 - v_1 - v_2 - p_1$ (or p_3) $- r$. Thus, p_2 cannot have another child beside v_1 , so it must be incident with a cross edge, say (p_2, p_5) . For the same reasons, p_3 has exactly one child, namely v_2 , and is incident with a cross edge, say (p_3, p_6) . If $p_5 \neq p_1, p_3$, we obtain the 7-cycle $r - p_5 - p_2 - v_1 - p_1 - v_2 - p_3 - r$. If $p_5 = p_1$ and $p_6 = p_1$, then we have a W_5 , which does not exist in G , as shown in property (D) of Lemma 14. The only possibility left is $p_5 = p_3$ and hence $p_6 = p_2$ (i.e., we have the cross edge (p_2, p_3)). Observe that now v_1 and v_2 have exactly two parents. If p_7 were, say, a third parent of v_1 , then we would arrive at the 7-cycle $r - p_7 - v_1 - p_2 - p_3 - v_2 - p_1 - r$. In particular, the degrees of the vertices v_1, v_2, p_2, p_3 are exactly three.

Let us focus on a seventh vertex of G , which could be a third child of p_1 (recall that p_2 and p_3 have exactly one child each) or a fourth child of r . Suppose p_1 has a third child v_4 . Since the degree of v_4 is at least three and level 2 does not contain paths of length two (property (A)), v_4 must have a second parent. This second parent cannot be p_2 or p_3 , so it must be another child of r ; call it p_8 . This leads to the 7-cycle $r - p_8 - v_4 - p_1 - v_1 - v_2 - p_3 - r$. Thus, p_1 cannot have a third child, so r must have a fourth child; call it p_9 . Since G is biconnected, there exists a simple cycle C containing p_9 and p_1 . Let us focus on the path from p_9 to p_1 on C without going through r ; call it P . The second to last vertex on P (just before reaching p_1) must be another child of r (call it p_{10} (p_{10} could be p_9)), since p_1 does not have a third child and p_2, p_3, v_1, v_2 have degrees exactly three. This leads to the 7-cycle $r - p_{10} - p_1 - v_1 - v_2 - p_3 - p_2 - r$, which is not possible. We conclude that case (ii) leads to a contradiction and is therefore not possible.

Case (iii). v_1 and v_2 share two or more parents.

Let p_1 and p_2 be two parents that v_1 and v_2 share. Recall that G has no K_4 ; thus, p_1 is not adjacent to p_2 via a cross edge. Since the degree of r is at least three and since r is on a 3-cycle, r must have a third child, say p_3 , such that p_3 is on a 3-cycle with r . Since G is biconnected, there exists a path P from p_3 to the subgraph induced by the vertices v_1, v_2, p_1, p_2 that does not pass through r . If P had length two or more, then we could use P and the edges among vertices $r, v_1, v_2, p_1, p_2, p_3$ to build an odd cycle of size seven or more. Consequently, P has length one; i.e., all paths from p_3 to the subgraph induced by the vertices v_1, v_2, p_1, p_2 not going through r are of length one (or edges). Note that p_3 is not incident with a cross edge (p_3, p_i) for some $p_i \neq p_1, p_2$; otherwise it would be easy to construct a 7-cycle from the path $r - p_i - p_3$ and a path from p_3 back to r of length five involving intermediate vertices v_1, v_2, p_1, p_2 . In addition, p_3 does not have children that are different from v_1 and v_2 . Suppose, to the contrary, that v_i is a child of p_3 , different from v_1 and v_2 . Since the degree of v_i is at least three and since level 2 does not contain paths of length two (property (A)), v_i must have a second parent; call it p_j . If $p_j = p_1$, we have the 7-cycle $r - p_3 - v_i - p_1 - v_1 - v_2 - p_2 - r$; if $p_j = p_2$, we obtain a similar 7-cycle. If $p_j \neq p_1, p_2$, we can build a 7-cycle from the path $r - p_j - v_i - p_3$ and a path from p_3 back to r of length four involving three intermediate vertices from the set $\{v_1, v_2, p_1, p_2\}$. We conclude that if p_3 is adjacent to a vertex different from r , that vertex must be in the set $\{v_1, v_2, p_1, p_2\}$. Since p_3 has degree at least three, it must be adjacent to at least two vertices in this set.

The graph G has at least seven vertices; a seventh vertex could be a fourth child of r , or a child of p_1 or p_2 . Suppose r has a fourth child; call it p_4 . Using arguments analogous to those above, we see that the structure for p_4 is the same as that for p_3 ; that is, p_4 is not incident with a cross edge (p_4, p_k) for some $p_k \neq p_1, p_2$; p_4 has no children different from v_1 and v_2 ; and p_4 is adjacent to at least two of the vertices in the set $\{v_1, v_2, p_1, p_2\}$. Recall that p_3 is on a 3-cycle with r . Hence, p_3 is adjacent to at least one of p_1, p_2 . Without loss of generality, assume p_3 is adjacent to p_1 . Note that there exist paths of length three from p_1 to each of v_1, v_2, p_2 . It is easy now to construct a 7-cycle: take the path $r - p_3 - p_1$, plus a path from p_1 to p_4 of length four involving intermediate vertices v_1, v_2, p_2 , plus the path $p_4 - r$. We conclude that r cannot have a fourth child. In other words, the only vertices on level 1 are p_1, p_2 , and p_3 .

A seventh vertex—call it v_3 —must therefore be a child of p_1 or p_2 , and say it is a child of p_1 . Since v_3 has degree at least three and level 2 has no paths of length two (property (A)), v_3 must have another parent besides p_1 . That parent cannot be p_3 , because p_3 has no children different from v_1 and v_2 , so it must be p_2 . Since level 1 has exactly three vertices, the third adjacency of v_3 must be a cross edge. There are no paths of length two at level 2, so that cross edge must be to a new vertex on level 2, say v_4 . The vertex v_4 has degree at least three, and level 2 contains no paths of length two, so v_4 must have two parents. Neither of these parents can be p_3 (which has no children different from v_1 and v_2), so p_1 and p_2 must be the parents of v_4 as well. The existence of v_3 and v_4 imply that p_3 cannot be adjacent to v_1 or v_2 ; that would lead to a 7-cycle (for instance, if p_3 were adjacent to p_1 and v_1 , then we would get the 7-cycle $r - p_3 - v_1 - v_2 - p_1 - v_3 - p_2 - r$). Thus, p_3 must be incident with both p_1 and p_2 via cross edges. Note that the structures of v_3 and v_4 are identical to those of v_1 and v_2 . Similarly, if p_1 has additional children (they must come in pairs), then their structures must be the same as those of v_1 and v_2 . Thus, we arrive at the family of graphs shown in Figure 6.

(C) Level 1 does not contain paths of length three or more.

Claim C.1. Level 1 does not contain paths of length four or more.

Since each vertex on level 1 is adjacent to r , level 1 does not contain paths of length five or more; otherwise it would be easy to build a 7-cycle using a length-five path together with r . Now suppose, for a contradiction, that level 1 contains a path of length four, say $v_1 - v_2 - v_3 - v_4 - v_5$. Since G contains at least seven vertices, there exists a seventh vertex (call it u) that is either a child of r or a child of a vertex on level 1. In either case, since G is biconnected, there exist two paths from u to the subgraph induced by the vertices $r, v_1, v_2, v_3, v_4, v_5$. Due to the multiple number of edges that exist among vertices $r, v_1, v_2, v_3, v_4, v_5$, it is not hard to construct an odd cycle of size seven or more from those two paths and the edges among vertices $r, v_1, v_2, v_3, v_4, v_5$, which gives the desired contradiction.

Claim C.2. Level 1 does not contain paths of length three.

Suppose, for a contradiction, that level 1 contains a path of length three, say $v_1 - v_2 - v_3 - v_4$. Note that level 2 must be independent, because we have shown in property (A) that level 2 does not contain paths of length two or more, and the only family of graphs where level 2 contains paths of length one is shown in Figure 6. Also note that the family of graphs shown in Figure 6 does not contain paths of length three on level 1. We now consider two cases: (i) v_1 and v_4 have no children, and (ii) v_1 and v_4 have children, and we show that they are not possible.

Case (i). v_1 and v_4 have no children.

Because the degrees of v_1 and v_4 are at least three and v_1 and v_4 have no children, v_1 and v_4 must each be incident with at least two cross edges. Since level 1 is bipartite (property (D) of Lemma 14) and since level 1 does not contain paths of length four or more (Claim C.1), v_1 must be adjacent to v_4 (i.e., (v_1, v_4) must be the second cross edge). Since G has at least seven vertices, a sixth vertex can either be a child of r or a child of a vertex on level 1. Suppose, for a contradiction, that r has a fifth child; call it v_5 . Since level 1 does not contain paths of length four or more (Claim C.1), v_5 is not adjacent to any of v_1, v_2, v_3, v_4 . Because G is biconnected, there must exist a path of length two or more from v_5 to the subgraph induced by the vertices v_1, v_2, v_3, v_4 that does not pass through r . Due to the multiple number of edges that exist among vertices r, v_1, v_2, v_3, v_4 , it is not hard to construct an odd cycle of size seven or more using that path and the edges among vertices $r, v_1, v_2, v_3, v_4, v_5$. Thus, r has only four children. Next suppose that, say, v_2 has a child; call it w . Since the degree of w is at least three and level 2 is independent, w must have three parents. But we have just shown that r has only four children, and v_1 and v_4 have no children by assumption. Thus, it is not possible for w to have three parents, which implies that a sixth vertex cannot exist.

Case (ii). v_1 and v_4 have children.

As usual, we argue by contradiction to show that this cannot occur. First, consider the case where v_1 and v_4 have at least two children between them; say w_1 is a child of v_1 , and w_2 is a child of v_4 , and $w_1 \neq w_2$. Since G is biconnected, (v_1, w_1) and (v_4, w_2) lie on a common simple cycle (Theorem 1); call it C . Due to the multiple number of edges that exist among vertices r, v_1, v_2, v_3, v_4 , it is not hard to construct an odd cycle of size seven or more from C and the edges among vertices r, v_1, v_2, v_3, v_4 , leading to a contradiction.

Now consider the case where v_1 and v_4 have exactly one child between them; call it w . Since G has at least seven vertices, a seventh vertex, say u , can either be a child of r or a child of a vertex on level 1. Suppose, for a contradiction, that u is a fifth child

of r . Because G is biconnected, there must exist a path of length one or more from u to the subgraph induced by the vertices v_1, v_2, v_3, v_4, w that does not pass through r . Due to the multiple number of edges that exist among vertices r, v_1, v_2, v_3, v_4, w , it is not hard to construct an odd cycle of size seven or more from that path and the edges among vertices $r, v_1, v_2, v_3, v_4, w, u$. Thus, r has only four children. It follows that u must be a child of a vertex on level 1. Since the degree of u is at least three and since level 2 is independent, u must have three parents. But r has only four children, and v_1 and v_4 have exactly one child (w) between them, so it is not possible for u to have three parents.

(D) If level 1 contains paths of length two and level 2 is independent, then the graph belongs to one of the families of graphs shown in Figure 7 or 8.

Suppose that level 1 contains a path of length two, say $v_1 - v_2 - v_3$. Since the degree of v_1 is at least three, since level 1 is bipartite (property (D) of Lemma 14), and since level 1 does not contain paths of length three (property (C)), v_1 must have at least one child. A similar argument and conclusion apply to v_3 . We consider two cases: (i) v_1 and v_3 together have two or more children, and (ii) v_1 and v_3 have exactly one and the same child.

Case (i). v_1 and v_3 together have two or more children.

Without loss of generality, let w_1 be a child of v_1 , and w_2 a child of v_3 , such that $w_1 \neq w_2$. Since the degree of w_1 is at least three and since level 2 is independent, w_1 must have at least two other parents besides v_1 ; and similarly for w_2 . We consider now two subcases: (a) w_1 and w_2 have v_1, v_2, v_3 only as parents, (b) at least one of w_1 and w_2 has a parent outside of v_1, v_2, v_3 .

Subcase (a). w_1 and w_2 have v_1, v_2, v_3 only as parents.

Since G has at least seven vertices and is biconnected, there must exist two paths from the seventh vertex to the subgraph induced by the vertices $r, v_1, v_2, v_3, w_1, w_2$. It is then easy to build an odd cycle of size at least seven from those two paths and the subgraph induced by $r, v_1, v_2, v_3, w_1, w_2$. Hence, this subcase cannot happen.

Subcase (b). At least one of w_1 and w_2 has a parent outside of v_1, v_2, v_3 .

In this subcase, we shall explore the structure of vertices and work toward building up the family of graphs shown in Figure 7. Without loss of generality, let v_4 be a parent of w_1 . We focus on potential (second and third) parents of w_2 . First, we show that v_1, v_2, v_4 cannot be parents of w_2 . If v_1 were a parent of w_2 , we would get the 7-cycle $r - v_4 - w_1 - v_1 - w_2 - v_3 - v_2 - r$. If v_2 were a parent of w_2 , we would have the 7-cycle $r - v_4 - w_1 - v_1 - v_2 - w_2 - v_3 - r$. If v_4 were a parent of w_2 , we would obtain the 7-cycle $r - v_1 - w_1 - v_4 - w_2 - v_3 - v_2 - r$. Thus, w_2 must have a new vertex (different from v_1, v_2, v_4) as a second parent. Similarly, the third parent of w_2 must also be a new vertex, and the same is true for the third parent of w_1 .

We now explore the structure of the parents of w_1 (or w_2), such as v_4 . Since the degree of v_4 is at least three, v_4 must have another child or be incident with a cross edge. Suppose, for a contradiction, that v_4 has another child; call it w_3 . Note that $w_3 \neq w_2$, since v_4 cannot be a parent of w_2 . If w_3 has a parent other than v_1, v_2 , say v_i , then we obtain the 7-cycle $r - v_i - w_3 - v_4 - w_1 - v_1 - v_2 - r$. If not, then since the degree of w_3 is at least three and level 2 is independent, w_3 must have v_1 and v_2 as parents. This yields the 7-cycle $r - v_4 - w_1 - v_1 - w_3 - v_2 - v_3 - r$. It follows from this observation that v_4 must be incident with a cross edge, say (v_4, v_5) . Note that $v_5 \neq v_1, v_3$, since level 1 does not contain paths of length three (property (C)). If v_5 were different from v_2 , we would get the 7-cycle $r - v_5 - v_4 - w_1 - v_1 - v_2 - v_3 - r$. Consequently, $v_5 = v_2$ (i.e., v_4 is adjacent to v_2). The same arguments

can be applied to all parents of w_1 and w_2 , including v_1 and v_3 , to conclude that all parents of w_1 and w_2 have degree exactly three, are adjacent to v_2 , and have only one child.

With regard to the structure of v_2 , observe that v_2 does not have any children. Indeed, if v_2 has a child w_j , then $w_j \neq w_1, w_2$, because (as noted earlier) v_2 cannot be a parent of w_1 or w_2 . Since the degree of w_j is at least three and level 2 is independent, w_j must have another parent (call it v_k), different from v_1 and v_3 , since v_1 and v_3 have exactly one child. This leads to the 7-cycle $r - v_k - w_j - v_2 - v_1 - w_1 - v_4 - r$.

Next we explore the structure of the children of r that are not parents of w_1 or w_2 . Let v_m be such a child. Let G' represent the subgraph induced by the vertices v_2, w_1, w_2 , and the parents of w_1 and w_2 . So v_m is outside of G' . Because G is bi-connected, there must exist a path P from v_m to G' that does not pass through r . Since the parents of w_1 and w_2 have degree exactly three and since w_1 and w_2 are not incident with cross edges, the only way to enter G' from outside not through r is via a cross edge to v_2 . Thus, the last two vertices on P are v_n and v_2 , where v_n is a child of r (v_n could be v_m). Observe that the length of the path from v_m to v_n on P is either zero (that is, $v_m = v_n$) or two, because there exist paths from v_m to v_n through r of length three ($v_m - r - v_2 - v_n$), of length four ($v_m - r - v_1 - v_2 - v_n$), and of length six ($v_m - r - v_1 - w_1 - v_4 - v_2 - v_n$) (so if there existed paths of lengths one, three, or more from v_m to v_n on P , we would have an odd cycle of size seven or more). Since the degree of v_n is at least three, v_n must either be incident with a second cross edge or have a child. It cannot be incident with a second cross edge, because level 1 is bipartite (property (D) of Lemma 14) and level 1 does not contain paths of length three (property (C)). It follows that v_n must have a child, say w_3 . Observe that $v_1 - v_2 - v_n$ is a path of length two on level 1, v_1 and v_n have two children (w_1 and w_3), and these two children have parents outside of v_1, v_2, v_3 . Thus, the parents of w_3 , such as v_n , must have the structure described earlier; namely they have degree exactly three, are adjacent to v_2 , and have only one child. If $v_m \neq v_n$, then v_m is also a parent of w_3 , because v_n has exactly one child and the path from v_m to v_n is of length two (via child w_3). Therefore, the structure of v_m (a child of r but not a parent of w_1 or w_2) is the same as that of v_n ; that is, v_m has degree exactly three, is adjacent to v_2 , and has only one child. Thus, we arrive at the family of graphs shown in Figure 7.

Case (ii). v_1 and v_3 have exactly one and the same child.

Notice that in Case (i) Subcase (b), $v_1 - v_2 - v_4$ is a path of length two on level 1, and v_1 and v_4 have exactly one and the same child. Hence, some of the graphs of Case (ii) are subgraphs of the family of graphs shown in Figure 7. For the discussion that follows, we shall focus on graphs that do not belong in Case (i), that is, graphs G such that $BFS(G_r)$ contains paths of length two on level 1 and all endpoints of the paths of length two have exactly one and the same child. Again, we shall explore the structure of vertices and work toward building up the family of graphs shown in Figure 8.

Let w_1 be the unique child of v_1 and v_3 . We make several observations.

First, v_2 does not have a child, except possibly for w_1 . Indeed, suppose, to the contrary, that w_2 is a child of v_2 different from w_1 . Since the degree of w_2 is at least three and level 2 is independent, w_2 must have at least two other parents besides v_2 . Let v_p be a parent of w_2 ; certainly $v_p \neq v_1, v_3$ since v_1 and v_3 have exactly one child, namely w_1 . Therefore, we obtain the 7-cycle $r - v_1 - w_1 - v_3 - v_2 - w_2 - v_p - r$.

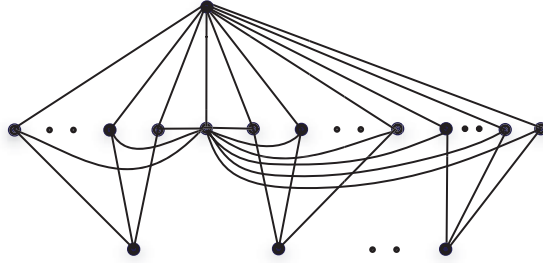


FIG. 7. A family of biconnected graphs with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices, where root r is on a 3-cycle, $BFS(G_r)$ has two levels, level 2 is independent, and level 1 contains paths of length two.

Second, all parents of w_1 have only one child, namely w_1 itself. Suppose, to the contrary, that v_q is a parent of w_1 and v_q has a child w_3 different from w_1 . Since the degree of w_3 is at least three and level 2 is independent, w_3 must have at least two other parents besides v_q . Let v_r be a parent of w_3 ; then $v_r \neq v_1, v_3$ since v_1 and v_3 have w_1 as their only child, and $v_r \neq v_2$ by the first observation. This leads to the 7-cycle $r - v_r - w_3 - v_q - w_1 - v_1 - v_2 - r$.

Third, all parents of w_1 are adjacent to v_2 but not incident with other cross edges. Suppose, indeed, that v_s is a parent of w_1 such that $v_s \neq v_1, v_3$, and v_s is adjacent to a vertex v_t on level 1 with $v_t \neq v_2$. Then we arrive at the 7-cycle $r - v_t - v_s - w_1 - v_1 - v_2 - v_3 - r$, which is a contradiction. Thus, v_s is not adjacent to a level 1 vertex different from v_2 . Since v_s has degree at least three and has just one child, by the second observation, it follows that v_s must be adjacent to v_2 . In particular, the degree of v_s is exactly three.

Fourth, due to our assumption that endpoints of paths of length two on level 1 have one and the same child, if v_2 is incident with a cross edge, say with a vertex v_t (so now $v_1 - v_2 - v_t$ is a path of length two), then v_t has exactly one child, namely w_1 .

Since G contains at least seven vertices, since w_1 is the only child of v_1, v_3 , and since w_1 is the only possible child of v_2 , a sixth vertex of G must be a child of r ; call it v_4 . We now explore the structure of v_4 . Let G' represent the subgraph induced by the vertices v_1, v_2, v_3, w_1 . Because G is biconnected, there exists a path P from v_4 to G' that does not pass through r . The ways for P to enter G' from outside, not going through r , are via v_2 through a cross edge or via w_1 through a parent-child edge. Let us focus on the second to last vertex on P ; call it u . Because the degree of u is at least three, observations two through four above imply that u must be adjacent to v_2 via a cross edge, must be a parent of w_1 , and must have degree exactly three (so $u = v_4$).

Using analogous arguments as above, we see that a seventh vertex must be a child of r with degree exactly three and adjacent to v_2 and w_1 . If G has more than seven vertices, they must also have the same structure as the sixth and the seventh vertices. Hence, the graph belongs to the family of graphs in Figure 8, with the edge (v_2, w_1) being optional.

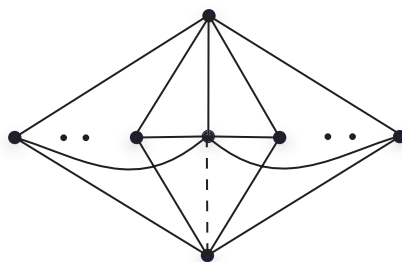


FIG. 8. A second family of biconnected graphs with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices, where root r is on a 3-cycle, $BFS(G_r)$ has two levels, level 2 is independent, and level 1 contains paths of length two.

We have completely characterized the structure of G when level 1 of $BFS(G_r)$ contains paths of length two, where r is a vertex on a 3-cycle. If there is a path of length two such that the endpoints of the path have two or more children, then G belongs to the family of graphs shown in Figure 7; otherwise it belongs to the family of graphs shown in Figure 8.

(E) If level 1 contains paths of length one only, and if level 2 is independent, then the graph belongs to one of the families of graphs in Figure 9 or 10.

We consider two cases: (i) level 1 contains two or more paths of length one, and (ii) level 1 contains exactly one path of length one.

Case (i). Level 1 contains two or more paths of length one.

Let $v_1 - v_2$ and $v_3 - v_4$ be two paths of length one on level 1. Since the degree of v_1 is at least three and since level 1 does not contain paths of length two or more by assumption, v_1 must have at least one child. A similar conclusion holds for vertices v_2, v_3 , and v_4 . We now argue that v_1, v_2, v_3 , and v_4 together cannot have two or more children. Suppose, to the contrary and without loss of generality, that v_1 and v_3 have distinct children; say, w_1 is a child of v_1 , and w_2 is a child of v_3 . Since G is biconnected, there exists a simple cycle that contains the edges (v_1, w_1) and (r, v_3) (Theorem 1), which means there exists a path from w_1 to v_3 without going through r ; call it P . Since the maximum cycle length in G is eight (Theorem 5) and since there is a path from w_1 to v_3 through r of length three ($w_1 - v_1 - r - v_3$), the length of P is at most five. We shall examine various possible paths of P according to length and show that no such path can exist. Keep in mind that since level 2 is independent and since the degrees of w_1 and w_2 are at least three, w_1 and w_2 must both have at least three parents.

Suppose the length of P is one, that is, $P = w_1 - v_3$. If v_1 and v_2 are both parents of w_2 , then we have the 7-cycle $r - v_2 - w_2 - v_1 - w_1 - v_3 - v_4 - r$; if not, we obtain the 7-cycle $r - v_2 - v_1 - w_1 - v_3 - w_2 - v_i - r$, where v_i is a parent of w_2 different from v_1 and v_2 . Suppose, next, that the length of P is two. Since level 2 is independent and since level 1 contains paths of length one only by assumption, P must be $w_1 - v_4 - v_3$. If v_1 and v_4 are both parents of w_2 , then we get the 7-cycle $r - v_2 - v_1 - w_1 - v_4 - w_2 - v_3 - r$; if not, then we arrive at the 7-cycle $r - v_1 - w_1 - v_4 - v_3 - w_2 - v_j - r$, where v_j is a parent of w_2 different from v_1 and v_4 .

If the length of P is three, then either P contains v_2 , in which case it does not contain v_4 and we get the 7-cycle $r - v_1 - w_1 \xrightarrow{P} v_3 - v_4 - r$; or P does not contain v_2 and we have the 7-cycle $r - v_2 - v_1 - w_1 \xrightarrow{P} v_3 - r$. If the length of P is four, then we arrive at the 7-cycle $r - v_1 - w_1 \xrightarrow{P} v_3 - r$. Finally, suppose the length of P is five. If P does not contain v_4 , then we obtain the 9-cycle $r - v_1 - w_1 \xrightarrow{P} v_3 - v_4 - r$. If P contains v_4 , let P_1 represent the path in P from w_1 to v_4 . The length of P_1 is four or less. If the length of P_1 is four, we get the 7-cycle $r - v_1 - w_1 \xrightarrow{P_1} v_4 - r$. If the length of P_1 is three or less, then the path $P_1 - v_3$ is a path from w_1 to v_3 of length four or less, which we have already shown does not exist.

The preceding argument shows that it is not possible for v_1, v_2, v_3 , and v_4 together to have two or more children. They must therefore all have exactly one and the same child, say w . Note that in this case v_1, v_2, v_3 , and v_4 have degree exactly three. Since G contains at least seven vertices, a seventh vertex of G must be a child of r ; call it v_5 . Because G is biconnected, there exists a simple cycle that contains the edges (v_1, w) and (r, v_5) (Theorem 1), which means there exists a path from w to v_5 that does not go through r . Since level 2 is independent, the first vertex after w on the path from w to v_5 must be a parent of w ; call it v_p (it is possible that $v_5 = v_p$). Note that v_p does not have another child besides w ; if it did, say w_k , then we would obtain the 7-cycle $r - v_1 - v_2 - w - v_p - w_k - v_s - r$, where v_s is a parent of w_k . Thus, in order for v_p to have degree at least three, v_p must be incident with a cross edge, say (v_p, v_q) . By taking v_1, v_2, v_p, v_q instead of v_1, v_2, v_3, v_4 and using the argument of the preceding paragraphs, it follows that v_q must have w as its only child. Both v_p and v_q , like v_1 through v_4 , have degrees exactly three and have identical structure. Since all paths on level 1 have length one, it is the case that either v_p or v_q is v_5 .

If G has additional vertices, using analogous arguments we see that those vertices must have the same structure as v_5 ; that is, they have degree exactly three, they have w as a child, and they are incident with a cross edge. We now arrive at the family of graphs shown in Figure 9.

Case (ii). Level 1 contains exactly one path of length one.

Let $v_1 - v_2$ be a path of length one on level 1. We shall mostly use the degree requirement to analyze the rest of the graph. Since the degree of v_1 is at least three and since level 1 does not contain paths of length two or more by assumption, v_1 must have at least one child, say w_1 . Because the degree of w_1 is at least three and level 2 is independent, w_1 must have a parent on level 1 that is different from v_1 and v_2 ; call it v_3 . Since the degree of v_3 is at least three and since level 1 contains exactly one path of length one, v_3 must have another child (besides w_1); call it w_2 . Again, because the degree of w_2 is at least three and level 2 is independent, w_2 must have at least three parents. Note that if w_2 has a parent other than v_1, v_2, v_3 , say v_p , then we obtain a 7-cycle $r - v_2 - v_1 - w_1 - v_3 - w_2 - v_p - r$. Hence, w_2 must have v_1, v_2 , and v_3 as its parents. Similarly, w_1 must have v_2 as its third parent; if w_1 has a parent v_q different from v_1, v_2, v_3 , then we get the 7-cycle $r - v_2 - v_1 - w_2 - v_3 - w_1 - v_q - r$. This argument shows that w_1 and w_2 have exactly three parents, namely v_1, v_2, v_3 . The same structure applies to all the children of v_3 : they all have exactly three parents, namely v_1, v_2, v_3 . If r had additional children, say v_4 , then using analogous arguments as above, we could show that v_4 has at least two children, and all of v_4 's children have exactly three parents, namely v_1, v_2, v_4 . But we would arrive at the 7-cycle $r - v_4 - w_3 - v_1 - v_2 - w_1 - v_3 - r$, where w_3 is a child of v_4 different from w_1 . We conclude that r does not have a fourth child; in other words, r has

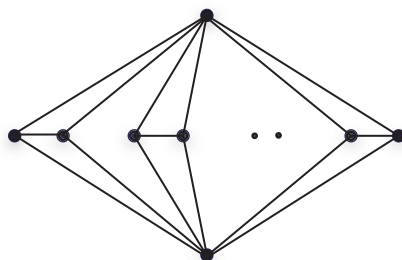


FIG. 9. A family of biconnected graphs with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices, where root r is on a 3-cycle, $BFS(G_r)$ has two levels, level 2 is independent, and level 1 contains paths of length one only.

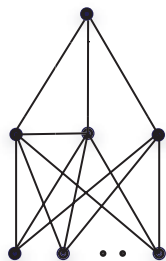


FIG. 10. A second family of biconnected graphs with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices, where root r is on a 3-cycle, $BFS(G_r)$ has two levels, level 2 is independent, and level 1 contains paths of length one only.

exactly three children. Since G contains at least seven vertices, a seventh vertex and any additional vertices of G must be children of v_1, v_2, v_3 . Thus, G belongs to the family of graphs shown in Figure 10. $\square$

Theorems 13, 15, and 16 provide a complete structural characterization of biconnected graphs with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices: the graph belongs to one of the family of graphs shown in Figures 2, 5–10. Observe that some of these families of graphs are *equivalent*, in the sense that they contain the same graphs but are drawn differently. In fact, families of graphs in Figures 5 and 7 are equivalent, as are those in Figures 6 and 9, and in Figures 8 and 10.

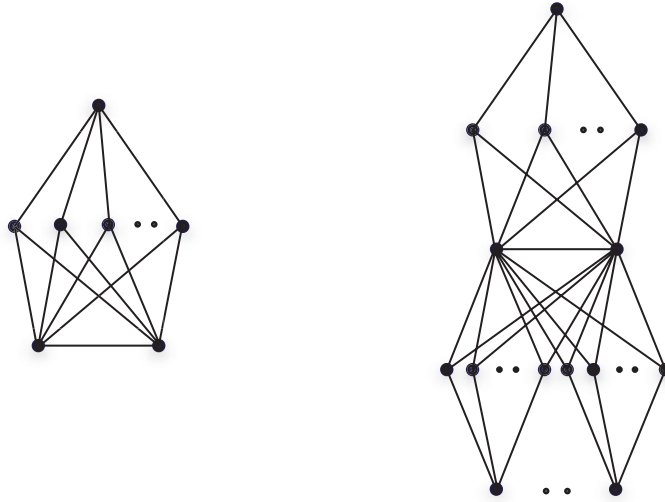


FIG. 11. Two families of biconnected graphs with two odd cycle sizes, 3 and 5, with minimum degree at least three, with at least seven vertices, and in which level 1 of $BFS(G)$ is independent.

Also notice that every vertex of the family of graphs shown in Figures 2, 6, and 9 is on a 3-cycle, while some vertices of the families of graphs shown in Figures 5, 7, 8, and 10 lie on 5-cycles only. After doing breadth-first searches on those vertices that lie on 5-cycles only, we arrive at the two families of graphs shown in Figure 11, with the first family equivalent to the families in Figures 8 and 10 and the second family equivalent to the families in Figures 5 and 7. Observe that the first family is a subgraph of the second family. The graphs in Figure 11 have an independent level 1 with structures consistent with the properties in Lemma 14; that is, $BFS(G)$ contains at most four levels; levels 3 and 4 of $BFS(G)$ are independent; and level 2 of $BFS(G)$ is bipartite. We state this structural characterization as Corollary 17.

COROLLARY 17. *Assume G is a biconnected graph with two odd cycle sizes, 3 and 5, with minimum degree at least three, with at least seven vertices, and with no K_4 . If level 1 of a breadth-first-search graph of G is independent, then G belongs to one of the family of graphs shown in Figure 11.*

Another important point regarding the families of graphs in Figures 2, 5–10 is that a vertex of maximum degree is always on a 3-cycle. In addition, if we do breadth-first searches on maximum degree vertices only, then the structures we get are those in Figure 2 (modified by breadth-first search) and in Figures 7, 8, and 9. This observation simplifies the number of cases we need to consider when approaching problems such as graph recognition and graph coloring.

We now present a simple and efficient graph recognition algorithm (Algorithm 1) for biconnected graphs with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices.

Input. A biconnected graph G with minimum degree at least three and with at least seven vertices.

Output. *Accept* if G contains only 3-cycles and 5-cycles in odd cycles; *Reject* otherwise.

1. Let r be a vertex of maximum degree in G .
2. Build the breadth-first-search graph of G rooted at r — $BFS(G_r)$.
3. If $BFS(G_r)$ has one level, then do
 4. If G belongs to the families of graphs shown in Figure 2, then accept.
5. Else if $BFS(G_r)$ has two levels, then do
 6. If there are two or more independent vertices at level 2 of $BFS(G_r)$, then do
 7. If G belongs to the family of graphs shown in Figure 7, then accept.
 8. Else if there is a single vertex at level 2 of $BFS(G_r)$, then do
 9. If level 1 of $BFS(G_r)$ has paths of length two, then do
 10. If G belongs to the family of graphs shown in Figure 8, then accept
 11. Else if level 1 of $BFS(G_r)$ has paths of length one only, then do
 12. If G belongs to the family of graphs shown in Figure 9, then accept.
13. Reject.

ALGORITHM 1. An $O(E)$ -time graph recognition algorithm for biconnected graphs with two odd cycle sizes, 3 and 5, with minimum degree at least three, and with at least seven vertices.

THEOREM 18. If G is a biconnected graph with minimum degree at least three, with at least seven vertices, and with two odd cycle sizes, 3 and 5, then Algorithm 1 recognizes G in $O(E)$ time.

Proof. Graph biconnectivity can be checked in $O(E)$ time [15]. In the following analysis, we assume that G is a biconnected graph with minimum degree at least three and at least seven vertices. As observed earlier, if G contains two odd cycle sizes, 3 and 5, and we perform breadth-first search of G on a vertex of maximum degree, then we get one of the structures shown in Figure 2 (modified by breadth-first search) and in Figures 7, 8, and 9. Thus, the tasks associated with verifying that G contains only 3-cycles and 5-cycles in odd cycles are the following: first, find a maximum degree vertex r ; second, build $BFS(G_r)$; and, third, check to see if $BFS(G_r)$ belongs to one of the families of graphs in Figure 2 (modified by breadth-first search) and in Figures 7, 8, and 9. This is basically Algorithm 1.

Note that when breadth-first search is performed on graphs in Figure 2 with root r being a maximum degree vertex, the breadth-first-search graph has only one level. Thus, verifying that G belongs to the first family of graphs in Figure 2 involves the following: check if level 1 contains exactly two vertices also of the maximum degree, and check if all other vertices on level 1 have degree exactly three and are adjacent to the three maximum degree vertices. Verifying that G belongs to other families of graphs involves similar checks and is graph specific. For instance, to check if G belongs to the family of graphs in Figure 7, we first check if level 1 contains a vertex that is adjacent to all other vertices on level 1 but has no children, and then we check whether each of the other vertices on level 1 has degree exactly three and has exactly one child. Verification for each of the families of graphs can be done in $O(E)$ time. Breadth-first search takes $O(E)$ time. Therefore, Algorithm 1 runs in $O(E)$ time. $\square$

Using the same framework as Algorithm 1, we present, in Algorithm 2, a simple

and efficient algorithm that colors G optimally in time $O(E)$.

Input. A biconnected graph G , with two odd cycle sizes, 3 and 5, and with minimum degree at least three.

Output. A p -coloring of G (see Table 1 for the value of p).

1. If G contains no more than six vertices, then color G using $\chi(G)$ colors.
2. Else do (G contains no K_5)
 3. Let r be a vertex of maximum degree in G .
 4. Build the breadth-first-search graph of G rooted at r — $BFS(G_r)$.
 5. Color root vertex r with color A.
 6. If $BFS(G_r)$ has one level, then do
 7. If level 1 of $BFS(G_r)$ contains two vertices of degree greater than three, then color them with colors B and C, and color the other vertices on level 1 with color D.
 8. Else color the vertex that has degree greater than three at level 1 of $BFS(G_r)$ with color B, and color the other vertices on level 1 with colors C and D.
 9. Else ($BFS(G_r)$ has two levels) do
 10. Color vertices at level 1 of $BFS(G_r)$ with colors B and C.
 11. Color vertices at level 2 of $BFS(G_r)$ with color A.

ALGORITHM 2. An optimal p -coloring algorithm (see Table 1 for the value of p) for graphs with two odd cycle sizes, 3 and 5.

THEOREM 19. If G is a biconnected graph with two odd cycle sizes, 3 and 5, then Algorithm 2 optimally colors G in $O(E)$ time.

Proof. In the following analysis, we assume that G is a graph with minimum degree at least three and then use Lemma 6 to relax the degree restriction. We also assume that the size of G is at least seven (if G contains no more than six vertices, we optimally color G by exhaustive search in step 1). Based on these assumptions, G does not contain a K_5 (Lemma 12). As noted in the proof of Theorem 18, $BFS(G_r)$ belongs to one of the families of graphs in Figure 2 (modified by breadth-first search) and in Figures 7, 8, and 9, where r is a maximum degree vertex in G .

Since the graphs in Figure 2 contain a K_4 and $BFS(G_r)$ has only one level, three colors are needed for the vertices on level 1 (steps 6–8). The graphs in Figures 7, 8, and 9 contain two levels and are straightforward to color: the root r gets one color; level 1 is bipartite, so it needs two colors; and level 2 is independent, so one color suffices (steps 5, 9–11). Breadth-first search takes $O(E)$ time. Therefore, Algorithm 2 runs in $O(E)$ time. $\square$

5. Summary and open problems. In this paper we have shown that biconnected graphs containing only small odd cycles possess many special structures. We have used breadth-first search to identify various structures and to show that certain structures do not exist. We have completely characterized the structure of graphs that are biconnected, with minimum degree at least three, with at least seven vertices, and with two odd cycle sizes, 3 and 5. Knowing the structure of these graphs enables us, among other things, to recognize them and to color them. We have developed graph coloring algorithms that are simple, efficient, and optimal in the number of colors used.

One direction of future work is to focus on graphs with different combinations of two odd cycle sizes, for example, 3 and 7, 3 and 9, 5 and 13, etc. Another direction is to consider graphs with three different odd cycle sizes (such as 3, 5, and 7), four different odd cycle sizes, and so on.

We have used case analysis to study the structure of graphs with only small odd cycles. As the number s of different odd cycle sizes increases, the complexity of the structure and the number of cases to be considered increase. For large s , case analysis becomes infeasible. It would be very desirable to find a more uniform and general method of analyzing and identifying the global structure of graphs.

Acknowledgments. I thank my advisor Andrea LaPaugh for helpful discussions and comments, for carefully reading earlier drafts of this paper, and for her advice and support. I thank an anonymous referee for helpful, detailed comments and for suggestions on how to improve the paper.

REFERENCES

- [1] C. BERGE, *Graphs and Hypergraphs*, North-Holland, Amsterdam, London, 1973.
- [2] D. P. DAILEY, *Uniqueness of colorability and colorability of planar 4-regular graphs are NP-complete*, Discrete Math., 30 (1980), pp. 289–293.
- [3] P. ERDŐS, *Some of my favourite unsolved problems*, in A Tribute to Paul Erdős, A. Baker, B. Bollobás, and A. Hajnal, eds., Cambridge University Press, Cambridge, UK, 1990, pp. 467–478.
- [4] P. ERDŐS AND A. HAJNAL, *On chromatic numbers of graphs and set systems*, Acta Math. Acad. Sci. Hungar., 17 (1966), pp. 61–99.
- [5] U. FEIGE AND J. KILIAN, *Zero knowledge and chromatic number*, J. Comput System Sci., 57 (1998), pp. 187–199.
- [6] M. R. GAREY, D. S. JOHNSON, AND L. J. STOCKMEYER, *Some simplified NP-complete graph problems*, Theoret. Comput. Sci., 1 (1976), pp. 237–267.
- [7] M. GRÖTSCHEL, L. LOVÁSZ, AND A. SCHRIJVER, *The ellipsoid method and its consequence in combinatorial optimization*, Combinatorica, 1 (1981), pp. 169–197.
- [8] A. GYÁRFÁS, *Graphs with k odd cycle lengths*, Discrete Math., 103 (1992), pp. 41–48.
- [9] D. S. JOHNSON, *A catalog of complexity classes*, in Handbook of Theoretical Computer Science, J. Van Leeuwen, ed., Elsevier, Amsterdam, 1990, pp. 67–161.
- [10] R. M. KARP, *Reducibility among combinatorial problems*, in Complexity of Computer Computations, R. E. Miller and J. W. Thatcher, eds., Plenum Press, New York, 1972, pp. 85–103.
- [11] S. KHANNA, N. LINIAL, AND S. SAFRA, *On the hardness of approximating the chromatic number*, Combinatorica, 20 (2000), pp. 393–415.
- [12] C. LUND AND M. YANNAKAKIS, *On the hardness of approximating minimization problems*, J. Assoc. Comput. Mach., 41 (1994), pp. 960–981.
- [13] P. MIHÓK AND I. SCHIERMEYER, *Cycle lengths and chromatic number of graphs*, Discrete Math., 286 (2004), pp. 147–149.
- [14] B. RANDERATH AND I. SCHIERMEYER, *Vertex colouring and forbidden subgraphs—a survey*, Graphs Combin., 20 (2004), pp. 1–40.
- [15] R. TARJAN, *Depth-first search and linear graph algorithms*, SIAM J. Comput., 1 (1972), pp. 146–160.
- [16] H.-J. VOSS, *Cycles and Bridges in Graphs*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.
- [17] S. S. WANG YEH, *Odd Cycles, Bipartite Subgraphs, and Approximate Graph Coloring*, Ph.D. thesis, Princeton University, Princeton, NJ, 1989 (also Princeton University CS Dept. Technical Report 208, June 1989).

ON THE GRAPH BISECTION CUT POLYTOPE*

MICHAEL ARMBRUSTER[†], CHRISTOPH HELMBERG[†], MARZENA FÜGENSCHUH[‡],
AND ALEXANDER MARTIN[‡]

Abstract. Given a graph $G = (V, E)$ with node weights $\varphi_v \in \mathbb{N} \cup \{0\}$, $v \in V$, and some number $F \in \mathbb{N} \cup \{0\}$, the convex hull of the incidence vectors of all cuts $\delta(S)$, $S \subseteq V$, with $\varphi(S) \leq F$ and $\varphi(V \setminus S) \leq F$ is called the *bisection cut polytope*. We study the facial structure of this polytope which shows up in many graph partitioning problems with applications in VLSI design or frequency assignment. We give necessary and in some cases sufficient conditions for the knapsack tree inequalities introduced in [C. E. Ferreira et al., *Math. Programming*, 74 (1996), pp. 247–267] to be facet-defining. We extend these inequalities to a richer class by exploiting the fact that each cut intersects each cycle in an even number of edges. Finally, we present a new class of inequalities that are based on nonconnected substructures yielding nonlinear right-hand sides. We show that the supporting hyperplanes of the convex envelope of this nonlinear function correspond to the faces of the so-called *cluster weight polytope*, for which we give a complete description under certain conditions.

Key words. graph partitioning, clustering, integer polyhedra

AMS subject classifications. 90C57, 90C10

DOI. 10.1137/060675253

1. Introduction. Let $G = (V, E)$ be an undirected graph with $V = \{1, \dots, n\}$ and $E \subseteq \{\{i, j\} : i, j \in V, i < j\}$. For given vertex weights $\varphi_v \in \mathbb{N} \cup \{0\}$ for all $v \in V$ and edge costs $w_{\{i, j\}} \in \mathbb{R}$ for all $\{i, j\} \in E$, a partition of the vertex set V into two disjoint clusters S and $V \setminus S$ with sizes $\varphi(S) \leq F$ and $\varphi(V \setminus S) \leq F$ for fixed $F \in [\frac{1}{2}\varphi(V), \varphi(V)] \cap (\mathbb{N} \cup \{0\})$ is called a *bisection*. Finding a bisection such that the total cost of edges in the cut $\delta(S) := \{\{i, j\} \in E : i \in S, j \in V \setminus S\}$ is minimal is the *minimum bisection problem* (MB). Its decision form is known to be NP-complete [10].

In this paper we investigate the *bisection cut polytope* P_B associated with MB. To define P_B note that a cut $\delta(S)$ can be described by its incidence vector $\chi^{\delta(S)}$ with respect to the edge set E . Then

$$P_B := \text{conv}\{y \in \mathbb{R}^{|E|} : y = \chi^{\delta(S)}, S \subseteq V, \varphi(S) \leq F, \varphi(V \setminus S) \leq F\}.$$

MB as well as P_B are related to other problems and polytopes in the literature. Obviously, the bisection cut polytope is contained in the *cut polytope* [2, 6]

$$(1) \quad P_C := \text{conv}\{y \in \mathbb{R}^{|E|} : y = \chi^{\delta(S)}, S \subseteq V\}.$$

If $F = \varphi(V)$, then MB is equivalent to the maximum cut problem (using the negative cost function) and $P_B = P_C$. For $F = \lceil \frac{1}{2}\varphi(V) \rceil$ MB is equivalent to the *equipartition problem* [5] and the bisection cut polytope equals the *equipartition polytope* [3, 4, 14]

$$P_E := \text{conv}\{y \in \mathbb{R}^{|E|} : y = \chi^{\delta(S)}, S \subseteq V, |\varphi(S) - \varphi(V \setminus S)| \leq 1\}.$$

*Received by the editors November 17, 2006; accepted for publication (in revised form) March 11, 2008; published electronically July 2, 2008. This work was supported by the German Research Foundation (DFG) under grants HE 3524/2-1/2 and MA 1324/2-1/2.

<http://www.siam.org/journals/sidma/22-3/67525.html>

[†]Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany (michael.armbruster@mathematik.tu-chemnitz.de, helmberg@mathematik.tu-chemnitz.de).

[‡]FB Mathematik, Technische Universität Darmstadt, Schlossgartenstr. 7, D-64289 Darmstadt, Germany (mfuegenschuh@mathematik.tu-darmstadt.de, martin@mathematik.tu-darmstadt.de).

Furthermore, MB is a special case of the minimum node capacitated graph partitioning problem (MNCGP) [8], where two or more clusters are available for the partition of the node set and each cluster has a limited capacity. The objective in MNCGP is the same as in MB, i.e., to minimize the total cost of edges having endpoints in distinct clusters. Finally, we mention the *knapsack polytope* [17]

$$(2) \quad P_K := \text{conv} \left\{ x \in \{0, 1\}^{|V|} : \sum_{v \in V} \varphi_v x_v \leq F \right\}.$$

P_K plays a fundamental role in the inequalities which we derive for the bisection cut polytope.

Graph partitioning problems in general have numerous applications, for instance, in numerics [12], VLSI design [15], compiler design [13], and frequency assignment [7].

The main contribution of this paper is threefold. First, in [8] the so-called knapsack tree inequalities have been introduced. These inequalities relate the knapsack conditions on the nodes with the edge variables defining the cuts and turn out to be computationally very effective. However, no theoretical justification has been found so far for this behavior. In this paper, we give necessary conditions for the knapsack tree inequality to be facet-defining, which turn out to be also sufficient in certain cases. Second, we can generalize the knapsack tree inequalities in the case of bisections by exploiting the well-known fact that any cut intersects a cycle an even number of times. This new class of inequalities, called *bisection knapsack walk inequalities*, subsume the knapsack tree inequalities and yield computationally more flexibility in finding strong inequalities. The third class of inequalities, called *capacity reduced bisection knapsack walk inequalities*, extend both classes of inequalities to nonconnected substructures. The idea is to exploit the weights of the nodes that are not end-nodes of walks to reduce the capacity of the corresponding knapsack inequality, yielding in this way stronger right-hand sides for the knapsack tree and bisection knapsack walk inequalities. These stronger conditions result in nonlinear right-hand sides. We consider the convex envelope of this nonlinear function and show that the supporting hyperplanes are in one-to-one correspondence to the faces of a certain polytope, called *cluster weight polytope*. For the case of a star without capacity restriction we are able to give a complete description of the cluster weight polytope yielding in this case the tightest strengthening possible for the capacity reduced bisection knapsack walk inequalities.

The outline of the paper is as follows. In section 2 we introduce an integer programming formulation for MB building on the formulation of MNCGP given in [8]. Section 3 treats the known knapsack tree inequalities valid for both MB and MNCGP, while section 5 introduces the new bisection knapsack walk inequalities which are only valid for MB and which subsume the knapsack tree inequalities. Section 4 shows a strengthening applicable only to knapsack tree inequalities. Furthermore, we state necessary and sufficient conditions for knapsack tree inequalities to define facets of P_B . Finally, section 6 introduces a strengthening of the bisection knapsack walk inequalities. For this purpose we investigate the facial structure of the cluster weight polytope on stars.

We frequently denote an edge $\{i, j\}$ by ij . Let A and B be discrete sets such that $A \subseteq B$. The incidence vector of A with respect to B is a vector $\chi^A \in \{0, 1\}^{|B|}$ with

$$\chi_a^A = \begin{cases} 1 & \text{if } a \in A, \\ 0 & \text{if } a \in B \setminus A. \end{cases}$$

For a vector $x \in \mathbb{R}^{|B|}$ we define $x(A) = \sum_{a \in A} x_a$. The vector $0^{|A|}$ is the zero vector of dimension $|A|$ and e_a is the unit vector of dimension $|A|$, which is indexed by the elements of A and has entry 1 in coordinate $a \in A$. For a graph $G = (V, E)$ the edge set of the subgraph induced by $\bar{V} \subseteq V$ will be denoted by $E(\bar{V})$ and the node set of the subgraph induced by $\bar{E} \subseteq E$ by $V(\bar{E})$. The convex hull of a set $A \subseteq \mathbb{R}^n$ will be denoted by $\text{conv}\{A\}$.

2. An integer programming formulation of MB. The integer programming formulation for MB given below is based on the formulation for MNCGP presented in [8]. We introduce variables z_i^k for each node $i \in V$ and each cluster $k = 1, 2$ and edge variables y_{ij} for each edge $ij \in E$. z_i^k is set to 1 if node i is in cluster k and 0 otherwise. Variable y_{ij} is set to 1 if edge ij is in the cut, i.e., i and j are not in the same cluster. Then MB can be written as

$$\begin{aligned}
 \text{(MB)} \quad & \min \sum_{e \in E} w_e y_e \\
 \text{s.t.} \quad & z_i^1 + z_i^2 = 1 \quad \forall i \in V, \\
 & \sum_{i \in V} \varphi_i z_i^k \leq F, \quad k = 1, 2, \\
 & y_{ij} \geq z_i^1 - z_j^1 \quad \forall ij \in E, \\
 & y_{ij} \geq z_j^1 - z_i^1 \quad \forall ij \in E, \\
 & y_{ij} \leq 2 - z_i^1 - z_j^1 \quad \forall ij \in E, \\
 & y_{ij} \leq 2 - z_i^2 - z_j^2 \quad \forall ij \in E, \\
 & y_{ij} \in \{0, 1\} \quad \forall ij \in E, \\
 & z_i^k \in \{0, 1\} \quad \forall i \in V, k = 1, 2.
 \end{aligned}$$

The first constraints assure that each node i is packed into exactly one cluster k . The second constraints enforce the capacity restriction on each cluster k . Together with the first set of equations, the next four constraints transmit for each edge $ij \in E$ the values of variables z_i^1 and z_j^1 to the edge variable y_{ij} in the sense that $y_{ij} = 1$ if and only if $z_i^1 \neq z_j^1$ and $y_{ij} = 0$ otherwise. The last two constraints are the binary restrictions on the variables.

Noting that the variables z_i^k do not appear in the objective function, we can consider model

$$\begin{aligned}
 \min \quad & \sum_{e \in E} w_e y_e \\
 \text{s.t.} \quad & y \in Y_{\text{MB}},
 \end{aligned}$$

where $Y_{\text{MB}} \subseteq \mathbb{R}^{|E|}$ is the projection onto the y -space of the feasible region of model (MB). It can be worked out that $P_B = \text{conv}(Y_{\text{MB}})$.

3. Known valid inequalities for MNCGP and MB. A large variety of valid inequalities for the polytope associated to MNCGP is known and, since MB is a special case of MNCGP, all those inequalities are also valid for P_B : cycle inequalities of the cut polytope [2], tree inequalities [3], star inequalities [3], cycle inequalities of capacitated graph partitioning [4], cycle with tails inequalities [8], suspended tree inequalities [14], path block cycle inequalities [14], cycle with ear inequalities [8], strengthened cycle

with ear inequalities [8], knapsack tree inequalities [8], and strengthened knapsack tree inequalities [8].

In the remainder of the paper we specialize and improve the knapsack tree inequality for MB. First we recall its definition for MNCGP from [8].

DEFINITION 1 (knapsack tree inequality [8]). *Let $\sum_{v \in V} a_v x_v \leq a_0$ be a valid inequality for the knapsack polytope P_K with $a_v \geq 0$ for all $v \in V$. For a fixed node $r \in V$ and a subtree (T, E_T) of G rooted at r we define the knapsack tree inequality*

$$(3) \quad \sum_{v \in T} a_v \left(1 - \sum_{e \in P_{rv}} y_e \right) \leq a_0,$$

where for each $v \in T$ the edge set of the path joining node v to root r in (T, E_T) is denoted by P_{rv} .

If (T, E_T) is a star rooted at r , i.e., $E_T = \{\{r, t\} : t \in T, t \neq r\}$, then we call inequality (3) the knapsack star inequality.

In general, there is an exponential number of these knapsack tree inequalities, since for each combination of a valid knapsack inequality with a choice of a rooted tree there is one knapsack tree inequality.

PROPOSITION 2 (see [8]). *The knapsack tree inequality (3) is valid for the polytope P_B .*

The above statement follows from the fact that MB is a special case of MNCGP.

It will be useful to write inequality (3) in the form

$$(4) \quad \sum_{e \in E_T} \left(\sum_{v: e \in P_{rv}} a_v \right) y_e \geq \sum_{v \in T} a_v - a_0.$$

The term on the right-hand side may be interpreted as the excess if all vertices $v \in T$ are packed into the cluster containing the root node r , while we are only allowed to pack a total weight of a_0 . The left-hand side has to compensate for this; i.e., it has to force some edges into the cut so that not all vertices are placed into the same cluster as the current root. We use this reformulation to apply a folklore approach to strengthen coefficients in general binary programs and obtain

$$(5) \quad \sum_{e \in E_T} \min \left\{ \sum_{v: e \in P_{rv}} a_v, \sum_{v \in T} a_v - a_0 \right\} y_e \geq \sum_{v \in T} a_v - a_0.$$

We call this inequality the *truncated knapsack tree inequality*.

Remark 3. Note that this strengthening was already proposed in Proposition 3.12 in [8] applied to the knapsack tree inequality for MNCGP. For MNCGP those authors also proposed a second case of strengthening, namely (in our notation) to reduce α_e to a_0 . However, for MB we always have $\alpha_0 \leq a_0$ due to the following reason. In the bisection case if $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)^T \in P_K \cap \{0, 1\}^n$, then also $\hat{x} = (1 - \bar{x}_1, \dots, 1 - \bar{x}_n)^T$ lies in $P_K \cap \{0, 1\}^n$. This follows from the fact that the total node weight of each of the clusters $\{v \in V : \bar{x}_v = 1\}$ and $\{v \in V : \bar{x}_v = 0\}$ cannot exceed F . If $\sum_{i \in V} a_v \bar{x}_v \leq a_0$, then also $\sum_{v \in V} a_v \hat{x}_v = \sum_{v \in V} a_v (1 - \bar{x}_v) \leq a_0$. Summing up these two inequalities yields $\sum_{v \in V} a_v \leq 2a_0$ and thus $\alpha_0 = \sum_{v \in V} a_v - a_0 \leq a_0$.

4. Minimum root strengthening of knapsack tree inequalities. Given a knapsack inequality $\sum_{v \in V} a_v x_v \leq a_0$ with $a_v \geq 0$, $v \in V$, let a corresponding knapsack tree inequality be defined on a tree (T, E_T) rooted at r . If we replace r by another node from T , the paths change. The corresponding change of the coefficients

of the inequality will be exploited in the strengthening of the truncated knapsack tree inequality presented in this section. Our strengthening aims at reducing the coefficients of the left-hand side while keeping the value of the right-hand side. We are going to show that the strongest or in some cases even facet-defining inequality is achieved if r enforces a sort of balance with respect to the cumulated node weights on the paths to r . To emphasize that the coefficients in (5) depend on the root node r we introduce the notation

$$(6) \quad \alpha_0 := \sum_{v \in T} a_v - a_0, \quad \alpha_e^r := \min \left\{ \sum_{v: e \in P_{rv}} a_v, \alpha_0 \right\}, \quad e \in E_T,$$

and consider (5) in the form

$$(7) \quad \sum_{e \in E_T} \alpha_e^r y_e \geq \alpha_0.$$

Note that if we change the root of (T, E_T) , the right-hand side of (7) remains the same, since by this operation we do not eliminate nodes of (T, E_T) .

At first we derive some relations based on the definition of the coefficients α_e^r , $r \in T$, $e \in E_T$, which we will exploit in the proofs of the results presented in this section. The following lemma states that along a path from the root to some node the coefficients cannot increase.

LEMMA 4. *Let (T, E_T) be a tree in G rooted at node r and let e and f be two edges on a path to r such that e is closer to r than f with respect to the number of edges. Then*

$$(8) \quad \alpha_e^r \geq \alpha_f^r.$$

Proof. It suffices to consider incident edges e and f . Setting $e := ij$ and $f := jk$ we obtain

$$\sum_{v: e \in P_{rv}} a_v = \sum_{v: f \in P_{rv}} a_v + a_j + \sum_{\bar{e} \in \bar{E}} \left(\sum_{v: \bar{e} \in P_{rv}} a_v \right) \geq \sum_{v: f \in P_{rv}} a_v,$$

where $\bar{E}$ contains edges incident to j except e and f . Hence if $\alpha_f^r = \alpha_0$, then also $\alpha_e^r = \alpha_0$; otherwise $\alpha_f^r \leq \min\{\sum_{v: e \in P_{rv}} a_v, \alpha_0\} = \alpha_e^r$. $\square$

In the next lemma we investigate the change of coefficients if the root is moved from a node r to an adjacent node s .

LEMMA 5. *Let (T, E_T) be a tree in G and let $r, s \in T$ be two adjacent nodes with $\bar{e} = rs$. We have*

- (a) $\alpha_e^r = \alpha_e^s$ for all $e \in E_T$ such that $e \neq rs$;
- (b) $\alpha_{\bar{e}}^r = \min\{a_s + \sum_{e \in \delta(\{s\}) \setminus \{\bar{e}\}} \alpha_e^s, \alpha_0\}$;
- (c) if $\alpha_{\bar{e}}^r \leq \alpha_{\bar{e}}^s$, then $\alpha_e^r \leq \alpha_e^s$ for all $\bar{v} \in V_{\bar{e}}^r := \{v \in T : \bar{e} \in P_{rv}\}$ and all $e \in E_T$.

Proof. (a) For $e \neq rs$ we have $\{v : e \in P_{rv}\} = \{v : e \in P_{sv}\}$ and thus $\sum_{v: e \in P_{rv}} a_v = \sum_{v: e \in P_{sv}} a_v$.

(b) Using the notation from (6) we have

$$\sum_{v: \bar{e} \in P_{rv}} a_v = a_s + \sum_{e \in \delta(\{s\}) \setminus \{\bar{e}\}} \left(\sum_{v: e \in P_{sv}} a_v \right).$$

(c) Consider a $\bar{v} \in T$ with $\bar{e}$ on the path $\Pi = P_{r\bar{v}}$ with $V_\Pi = \{v_1, \dots, v_p\}$, $p \geq 2$, where $v_1 = r$, $v_2 = s$, $v_p = \bar{v}$, and v_k, v_{k+1} , $1 \leq k \leq p-1$, are adjacent in T . Applying (a) recursively to nodes v_i, v_{i+1} , $i = 1, \dots, p-1$, we obtain

$$(9) \quad \alpha_e^r = \alpha_e^{\bar{v}} \quad \forall e \in E_T \setminus E_\Pi.$$

As $\bar{e}$ is outside the path from s to $\bar{v}$, the same argument for root s and the assumption yield

$$(10) \quad \alpha_{\bar{e}}^{\bar{v}} = \alpha_{\bar{e}}^s \geq \alpha_{\bar{e}}^r.$$

By Lemma 4, coefficients cannot increase along paths from the root, so

$$(11) \quad \begin{aligned} \alpha_{\bar{v}v_{p-1}}^{\bar{v}} &\geq \alpha_{v_{p-1}v_{p-2}}^{\bar{v}} \geq \dots \geq \alpha_{v_3v_2}^{\bar{v}} \geq \alpha_{v_2r}^{\bar{v}} = \alpha_{\bar{e}}^{\bar{v}}, \\ \alpha_{\bar{e}}^r &= \alpha_{rv_2}^r \geq \alpha_{v_2v_3}^r \geq \dots \geq \alpha_{v_{p-2}v_{p-1}}^r \geq \alpha_{v_{p-1}\bar{v}}^r. \end{aligned}$$

Thus, putting (9)–(11) together, the claim is proved. $\square$

This allows us to characterize exactly the set of roots giving the best coefficients.

LEMMA 6. *Let (T, E_T) be a tree in G . The set of minimal roots*

$$\mathcal{R} := \{r \in T : \alpha_e^r \leq \alpha_e^v \quad \forall v \in V_T \text{ and } e \in E_T\}$$

is nonempty and induces a connected subtree in T . A node $r \in T$ satisfies $r \in \mathcal{R}$ if and only if $\alpha_e^r \leq \alpha_e^s$ for all $e = rs \in E_T$.

Proof. To see that $\mathcal{R}$ is nonempty, orient the edges $e = uv \in T$ with $\alpha_e^u < \alpha_e^v$ towards v ; do not orient edges with $\alpha_e^u = \alpha_e^v$. As T contains no cycle, there must be a node $r \in T$ so that all incident edges are either not oriented or point towards r , i.e., $\alpha_e^r \leq \alpha_e^s$ for all $e = rs \in E_T$. By Lemma 5(c) and using the notation defined there, this r is in $\mathcal{R}$, because $V = \bigcup_{rs \in E_T} V_{rs}^r \cup \{r\}$.

Next, we show connectedness for $r, s \in \mathcal{R}$. These satisfy $\alpha_e^r = \alpha_e^s$ for all $e \in E_T$. Assume there is some inner vertex $\bar{v} \in T$ on the path between r and s . Apply Lemma 5(a) on the edges outside the paths $P_{r\bar{v}}$ and $P_{s\bar{v}}$ to see that $\alpha_e^r = \alpha_e^{\bar{v}}$ for all $e \in E_T$ and so $\bar{v} \in \mathcal{R}$.

The characterization of the elements of $\mathcal{R}$ is obtained via Lemma 5(c) directly. $\square$

For each choice of a roots out of R we obtain the same coefficient for each edge. Thus, the strongest truncated knapsack tree inequality is independent of the choice of $r \in R$.

THEOREM 7. *Let (T, E_T) be a tree in G and $\mathcal{R}$ as defined in Lemma 6. The strongest truncated knapsack tree inequality, with respect to the knapsack inequality $\sum_{v \in V} a_v x_v \leq a_0$, $a_v \geq 0$, $v \in V$, defined on (T, E_T) is obtained for a root $r \in \mathcal{R}$; i.e., if $r \in \mathcal{R}$, then*

$$(12) \quad \sum_{e \in E_T} \alpha_e^s y_e \geq \sum_{e \in E_T} \alpha_e^r y_e \geq \alpha_0$$

holds for all $s \in T$ and all $y \in P_B$. In particular,

$$(13) \quad \sum_{e \in E_T} \alpha_e^r y_e = \sum_{e \in E_T} \alpha_e^s y_e$$

holds for all $r, s \in \mathcal{R}$ and all $y \in P_B$.

Proof. The proof is obtained directly by Lemma 6. $\square$

In what follows the elements of the set $\mathcal{R}$ will be called *minimal roots* of a given tree (T, E_T) . In order to obtain the strongest truncated knapsack tree inequality it is sufficient, by Theorem 7, to identify any minimal root. Given a tree (T, E_T) rooted at some node r one can find a minimal root along the lines of the proof of Lemma 6 by proceeding iteratively as follows. Select a node $s \in T$ adjacent to r such that $\alpha_{rs}^r = \max\{\alpha_{rv}^r : rv \in E_T\}$. If $\alpha_{rs}^r > \alpha_{rs}^s$, then also $\sum_{e \in E_T} \alpha_e^r > \sum_{e \in E_T} \alpha_e^s$ by Lemma 5(c). Hence r can be discarded and s is marked as the next root of (T, E_T) . Otherwise $\sum_{e \in E_T} \alpha_e^r y_e \geq \alpha_0$ is the strongest truncated knapsack tree inequality with respect to all possible choices of roots in (T, E_T) .

In the remainder of this section we show that the assumption on r to be a minimal root is not only a necessary condition for a truncated knapsack tree inequality to be facet-defining for the polytope P_B , which follows from Theorem 7, but in some cases is also sufficient.

For this purpose we assume that $G = (T, E_T)$ is a tree and $\varphi_v = 1$ for all $v \in T$. Then the knapsack polytope P_K is defined by the inequality $\sum_{v \in T} x_v \leq F$, and the corresponding knapsack tree inequality (3) defined on (T, E_T) takes the form

$$\sum_{v \in T} \left(1 - \sum_{e \in P_{rv}} y_e \right) \leq F.$$

Applying the strengthening (5) and notation (6) we obtain $\alpha_0 = |T| - F$ and $\alpha_e^r = \min\{|V_e^r|, |T| - F\}$, for all $e \in E$, where V_e^r is the set of nodes whose path to $r \in T$ contains the edge e ; see, e.g., Figure 3. To emphasize the special case that we treat in what follows we set $\kappa_e^r := \alpha_e^r$, $\bar{F} := \alpha_0$ and consider the inequality $\sum_{e \in E_T} \kappa_e^r y_e \geq \bar{F}$, or $(\kappa^r)^T y \geq \bar{F}$ for short. For ease of exposition we call κ_e^r the *knapsack weight* of $e \in E$ with respect to the root r of (T, E_T) . If $\kappa_e^r = \bar{F}$ and $\bar{F} < |V_e^r|$, we say that e has the *reduced knapsack weight*. Furthermore, we introduce the term *branchless path*, which is a path in (T, E_T) whose inner nodes are all of degree 2. We consider a path consisting of an edge and both its end-nodes as a trivial case of a branchless path. We call an edge a *leaf* if one of its endpoints is of degree 1.

THEOREM 8. *Assume that $G = (T, E_T)$ is a tree rooted at a node $r \in T$, $\varphi_v = 1$ for all $v \in T$ and $\frac{|T|}{2} + 1 \leq F < |T|$. The truncated knapsack tree inequality $(\kappa^r)^T y \geq \bar{F}$ is facet-defining for P_B if and only if one of the following conditions is satisfied:*

- (a) r is a minimal root and (T, E_T) satisfies the following branchless path condition: each branchless path with F nodes has one end-edge that is a leaf in (T, E_T) ;
- (b) $F = |T| - 1$.

Remark 9. Note the following:

- (1) Given a graph $G = (V, E)$, P_B is full-dimensional under the assumptions that $\varphi_v = 1$ for all $v \in V$ and $F \geq \frac{|V|}{2} + 1$; see [8]. If $F = \frac{|V|}{2}$, the bisection cut polytope is not full-dimensional, and hence this case needs a special treatment which can be found in [9].
- (2) In case $F = |V|$ the knapsack inequality $\sum_{v \in V} x_v \leq F$ is redundant for P_K and thus the corresponding truncated knapsack tree inequalities are redundant for P_B . Therefore we assume that $F < |V|$, in particular, $\bar{F} > 0$.

Due to the complexity of the proof of Theorem 8, we complete it in several steps. First we outline the general idea of the sufficiency part. Let $\mathcal{F}$ be a face of P_B induced by $(\kappa^r)^T y \geq \bar{F}$ and let $\mathcal{F}_b$ be the facet of P_B defined by the inequality $b^T y \geq b_0$ such that $\mathcal{F} \subseteq \mathcal{F}_b$. To show that $(\kappa^r)^T y \geq \bar{F}$ is a facet-defining inequality for P_B , we prove

that $\mathcal{F} = \mathcal{F}_b$; i.e., there exists $\gamma \in \mathbb{R} \setminus \{0\}$ such that

$$(14) \quad \begin{aligned} b_e &= \gamma \kappa_e^r \quad \forall e \in E, \\ b_0 &= \gamma \bar{F} \end{aligned}$$

hold.

We introduce now further definitions and lemmas required to prove the above relations. Given a partition of the node set T , we denote by V_r the cluster containing r ; see, e.g., Figure 3. We say that two edges $e, f \in E$ are *related* if there exists a path to the root containing both e and f . An edge e is related to itself. For an edge e we set $B_e = \{f \in E : f \text{ is related to } e\}$ and call the graph $(\bigcup_{f \in B_e} f, B_e)$ a *branch* (induced by e); it is a subtree of (T, E_T) and is said to be incident on r if e and r are incident. If any two edges e and f are incident and related and such that e is closer to the root than f (with respect to the number of edges), then f is a *child* of e . We denote the set of children of e by

$$S_e = \{f \in E_T : f \text{ is a child of } e\}.$$

With this, the recursive construction rule of Lemma 5(b) specialized to κ_e^r reads

$$(15) \quad \kappa_e^r = \min \left\{ 1 + \sum_{\bar{e} \in S_e} \kappa_{\bar{e}}^r, \bar{F} \right\}.$$

We say that a bisection cut $\delta(V_r)$ is *tight* for $(\kappa^r)^T y \geq \bar{F}$ if it satisfies $(\kappa^r)^T \chi^{\delta(V_r)} = \bar{F}$. As we will show soon, $|V_r| = F$ holds if $\delta(V_r)$ is tight for $(\kappa^r)^T y \geq \bar{F}$ and all $e \in \delta(V_r)$ satisfy $\kappa_e^r = |V_e^r| \leq \bar{F}$, i.e., all edges in the cut do not have reduced knapsack weights. In this case we will call the cut $\delta(V_r)$ *double-tight* for $(\kappa^r)^T y \geq \bar{F}$.

Next, we derive some properties of bisection cuts tight for $(\kappa^r)^T y \geq \bar{F}$.

LEMMA 10. *No two edges in a bisection cut tight for $(\kappa^r)^T y \geq \bar{F}$ are related.*

Proof. Assume, for contradiction, that $\delta(V_r)$ is a bisection cut tight for $(\kappa^r)^T y \geq \bar{F}$ containing two related edges $e = v_{j-1}v_j$ and $f = v_{k-1}v_k$ on some path $P_{v_k}^r = \{v_1v_2, v_2v_3, \dots, v_{j-1}v_j, \dots, v_{k-1}v_k\}$ with $r = v_1$ and $1 < j < k$. Without loss of generality (w.l.o.g.), we may assume that $v_{j-1} \in V_r$ and no further edge $v_{i-1}v_i$ is in the cut for $j < i < k$. Then $\{v_j, \dots, v_{k-1}\} \subset T \setminus V_r$ and $v_k \in V_r$. The cut induced by $\bar{V}_r = V_r \cup \{v_j\} \setminus \{v_k\}$ is again a bisection cut with $\delta(\bar{V}_r) \subseteq (\delta(V_r) \setminus \{e, f\}) \cup S_e \cup S_f$; we know $e \notin \delta(\bar{V}_r)$ (this will not hold for f if $f \in S_e$). Note that e and f must be unreduced because $\delta(V_r)$ is tight, $(\kappa^r)^T \chi^{\delta(V_r)} = \bar{F}$. Thus, by (15),

$$(\kappa^r)^T \chi^{\delta(V_r)} - (\kappa^r)^T \chi^{\delta(\bar{V}_r)} \geq \kappa_e^r + \kappa_f^r - \sum_{\bar{e} \in S_e \cup S_f} \kappa_{\bar{e}}^r \geq 1,$$

and that yields $(\kappa^r)^T \chi^{\delta(\bar{V}_r)} + 1 \leq (\kappa^r)^T \chi^{\delta(V_r)} = \bar{F}$. This contradicts $(\kappa^r)^T \chi^{\delta(\bar{V}_r)} \geq \bar{F}$, which holds because $\delta(\bar{V}_r)$ is a bisection cut and $(\kappa^r)^T y \geq \bar{F}$ is feasible for P_B . $\square$

LEMMA 11. *Let (T, E_T) be rooted at $r \in T$. A bisection cut $\delta(V_r)$ is double-tight for $(\kappa^r)^T y \geq \bar{F}$ if and only if $|V_r| = F$ and $(V_r, E(V_r))$ is connected.*

Proof. Assume first that $\delta(V_r)$ is double-tight for $(\kappa^r)^T y \geq \bar{F}$. By Lemma 10 any two edges in $\delta(V_r)$ are not related. This implies that V_r induces a connected subgraph of (T, E_T) . Hence $T \setminus V_r = \bigcup_{e \in \delta(V_r)} V_e^r$ and $V_e^r \cap V_f^r = \emptyset$ for any $e, f \in \delta(V_r)$. Furthermore, $\kappa_e^r = |V_e^r|$ holds for each $e \in \delta(V_r)$ and we obtain $|T \setminus V_r| = \sum_{e \in \delta(V_r)} \kappa_e^r = \bar{F}$, i.e., $|V_r| = F$.

Now consider a bisection $(V_r, T \setminus V_r)$ such that $(V_r, E(V_r))$ is connected and $|V_r| = F$ (i.e., $|T \setminus V_r| = \bar{F}$). We show first that $\delta(V_r)$ contains only edges whose knapsack weights are not reduced. Assume for contradiction that $\delta(V_r)$ contains an edge f with a reduced knapsack weight. Since $\kappa_f^r = \bar{F}$, this is the only edge in $\delta(V_r)$; otherwise $\delta(V_r)$ is not tight for $(\kappa^r)^T y \geq \bar{F}$. Hence $\delta(V_r) = \{f\}$, and since f has a reduced knapsack weight, $|T \setminus V_r| = |V_f^r| > \bar{F}$ holds, contradicting the assumption that $|V_r| = F$. To show that $\delta(V_r)$ is tight for $(\kappa^r)^T y \geq \bar{F}$, we use the assumption that $(V_r, E(V_r))$ is connected. We have

$$\sum_{e \in \delta(V_r)} \kappa_e^r = \sum_{e \in \delta(V_r)} |V_e^r| = \left| \bigcup_{e \in \delta(V_r)} V_e^r \right| = \bar{F}.$$

Hence $\delta(V_r)$ is double-tight for $(\kappa^r)^T y \geq \bar{F}$. $\square$

Next, we provide some results following from the assumption that (T, E_T) is rooted at a minimal root. As we will show in the following lemmas, this ensures the existence of bisection cuts tight for $(\kappa^r)^T y \geq \bar{F}$, which we will consider in the proof of further lemmas preceding the proof of Theorem 8.

LEMMA 12. *Let $B = (V_B, E_B) \subseteq (T, E_T)$ be a branch incident on root $r \in T$. If r is a minimal root of (T, E_T) , then $|V_B \setminus \{r\}| \leq F$.*

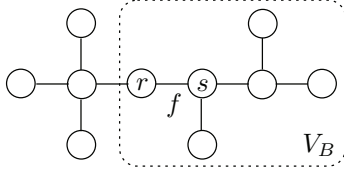


FIG. 1. Node set V_B .

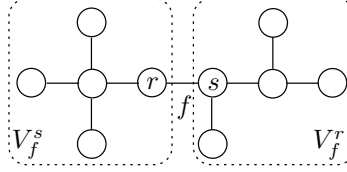


FIG. 2. Node sets V_f^s and V_f^r .

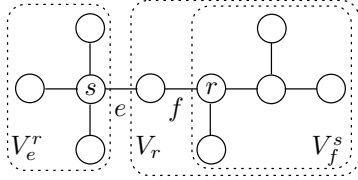
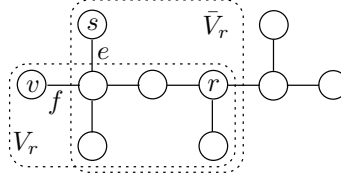
Proof. Let $B = (V_B, E_B)$ be a branch incident on r and assume that $|V_B \setminus \{r\}| > F$. We are going to show that in this case r cannot be a minimal root. Let s be the node in V_B adjacent to r , and let $f = rs \in E_B$; see Figures 1 and 2. Note that $V_f^r \cup V_f^s = T$. Since $|V_f^r| = |V_B \setminus \{r\}| > F > \bar{F}$, we have $\kappa_f^r = \bar{F}$ on the one hand. On the other hand, $\kappa_f^s = \min\{|T \setminus V_f^r|, \bar{F}\} < \bar{F}$. Hence $\kappa_f^s < \kappa_f^r$, and by Lemma 6, r is not a minimal root. $\square$

LEMMA 13. *Assume (T, E_T) , rooted at a minimal root r , has an edge $e \in E_T$ such that $\kappa_e^r = \bar{F}$. The cut $\delta(V_e^r) = \{e\}$ is a bisection cut tight for $(\kappa^r)^T y \geq \bar{F}$.*

Proof. Note that in the considered case we have $V_r = T \setminus V_e^r$. We show that the cut $\delta(V_r)$, which is obviously tight for $(\kappa^r)^T y \geq \bar{F}$, is also a bisection cut. Assume that $\delta(V_r)$ is not a bisection cut. Then either $|V_r| < \bar{F}$ or $|T \setminus V_r| < \bar{F}$. In the first case, let s be a node incident to e such that the path $\Pi_{rs} = (V_{rs}, E_{rs})$ joining r and s contains e ; see Figure 3. For all $f \in E_T \setminus E_{rs}$ there holds $\kappa_f^r = \kappa_f^s$ due to Lemma 5(a). For $f \in E_{rs}$ we obtain by Lemma 4 that $\kappa_f^r \geq \kappa_e^r = \bar{F} > |V_r| \geq |V_f^s| = \kappa_f^s$. By Lemma 6 this contradicts the assumption that r is a minimal root. One can show in a similar way that $|T \setminus V_r| < \bar{F}$ not possible either. $\square$

The following result exhibits the central structural property required for cross-connecting double-tight cuts of which one contains edge e and the next contains its children S_e .

LEMMA 14. *Given (T, E_T) with minimal root r . For each unreduced $e \in E_T$ that is not a leaf there exists a double-tight cut $\delta(V_r)$ that contains e , S_f , and S_g for two*

FIG. 3. Node sets V_r , V_e^r , and V_f^r .FIG. 4. Node sets V_r and $\bar{V}_r$ (Lemma 15(b)).

further unreduced edges f and g with $\{e, f, g\}$ unrelated if and only if (T, E_T) satisfies the branchless path condition.

Proof. First suppose that in (T, E_T) the branchless path condition does not hold, so there is a branchless path on nodes $\bar{V} = \{v_1, \dots, v_F\}$ with $\{v_i, v_{i+1}\} \in E_T$ for $1 \leq i < F$ and neither $\{v_1, v_2\}$ nor $\{v_{F-1}, v_F\}$ are leaves. By the minimality of r and Lemma 12 we must have $r \in \bar{V}$. One of the edges $\{v_1, v_2\}$ and $\{v_{F-1}, v_F\}$ must be unreduced (otherwise one of them would cut the graph into two parts, one containing more than F and one at least $\bar{F}+1$ nodes). W.l.o.g., let $e = \{v_1, v_2\}$ be unreduced and satisfy $\kappa_e^r \leq \kappa_{v_{F-1}v_F}^r$. For this e there exist no further two unrelated edges f and g satisfying the requirements, because at least one of them must be located on the path (by Lemma 11 we need $|V_r| = F$) and must thus be related either to e or to the other edge.

Now suppose that (T, E_T) satisfies the branchless path conditions. Let e be an unreduced edge that is not a leaf. Edge e is contained in some branch incident on r ; call this branch $B = (V_B, E_B)$. By Lemmas 11 and 12 we may construct a bisection $(V_r, T \setminus V_r)$ with $e \in \delta(V_r)$ and $\delta(V_r)$ double-tight by extending $V_B \setminus V_e^r$ to V_r via successively adding adjacent points from $V \setminus V_B$ so that $|V_r| = F$ and the subgraph $\bar{T} = (V_r, E_r)$ induced by V_r is a tree. Note that by Lemma 12 root r can be a leaf of $\bar{T}$ only if it is incident to e . First suppose $\bar{T}$ has two leaves so that the corresponding nodes of degree 1 are not incident to e , then take these for f and g . Otherwise $\bar{T}$ forms a path $\{v_1, \dots, v_F\}$ with v_i adjacent to v_{i+1} in $\bar{T}$ for $1 \leq i < F$ with e incident to one of its endpoints, say $v_1 \in e$. By the branchless path condition at least one of the nodes v_i , $1 \leq i < F-1$, has degree at least 3 in T (otherwise one can find a branchless path with F nodes not having one end edge a leaf); pick one and call it $\bar{v}$. Incident to this $\bar{v}$ there is an edge $f \in \delta(V_r)$ with $f \neq e$. Now observe that, by Lemma 11, $\delta(V_r \cup f \setminus \{v_F\})$ is a double-tight cut that contains e , S_f , and S_g for $g = \{v_{F-2}, v_{F-1}\}$ and e, f , and g are unrelated. $\square$

LEMMA 15. Given (T, E_T) satisfying the branchless path condition and a minimal root $r \in T$, let $\mathcal{F}, \mathcal{F}_b$ be faces of P_B defined by $(\kappa^r)^T y \geq \bar{F}$ and $b^T y \geq b_0$, respectively, such that $\mathcal{F} \subseteq \mathcal{F}_b$. For each branch $B = (V_B, E_B)$ in (T, E_T) incident on r there is a $\gamma_B \geq 0$ such that

- (a) $b_e = b_0$ holds for any edge $e \in E_T$ with $\kappa_e^r = \bar{F}$;
- (b) $b_e = b_f =: \gamma_B$ holds for any two leaves $e, f \in E_B$ of T ;
- (c) $b_e = \gamma_B \kappa_e^r$ holds for unreduced $e \in E_B$;
- (d) $\gamma_B = b_0/\bar{F}$ and $b_e = \gamma_B \kappa_e^r$ holds for $e \in E_B$.

Proof. (a) Let e be an edge in E_T with $\kappa_e^r = \bar{F}$. By Lemma 13 the cut $\delta(V_e^r) = \{e\}$ is a bisection cut tight for $(\kappa^r)^T y \geq \bar{F}$. Since $\chi^{\{e\}}$ is in $\mathcal{F}_b$, we obtain $b_e = b_0$.

(b) Let e and f be leaves in E_B and denote by s, v their respective end-nodes of degree 1. $|V_B \setminus \{s\}| \leq F$ follows from Lemma 12. Hence there exists a bisection $(V_r, T \setminus V_r)$ with V_r connected such that $V_B \setminus \{s\} \subseteq V_r$, $s \notin V_r$, and $|V_r| = F$. Then, by Lemma 11, $\bar{V}_r = V_r \cup \{s\} \setminus \{v\}$ also yields a double-tight bisection cut, so $b_0 = \sum_{\bar{e} \in \delta(V_r)} b_{\bar{e}} = \sum_{\bar{e} \in \delta(\bar{V}_r)} b_{\bar{e}}$ and therefore $b_e = b_f =: \gamma_B$; see Figure 4.

(c) We use induction by distance of the edges in E_B to the deepest leaves in their induced branches. By (b) the claim holds for the leaves. Let e be an unreduced edge in E_B that is not a leaf. By Lemma 14 there exists a double-tight cut $\delta(V_r)$ and unreduced edges f, g unrelated to e and to each other so that $e \in \delta(V_r)$, $S_f \subset \delta(V_r)$, and $S_g \subset \delta(V_r)$. Pick some $h \in S_e$. Lemma 4 implies that h is unreduced and hence (15) combined with the induction hypothesis yields

$$(16) \quad b_h - \sum_{\bar{e} \in S_h} b_{\bar{e}} = \gamma_B \left(\kappa_h^r - \sum_{\bar{e} \in S_h} \kappa_{\bar{e}}^r \right) = \gamma_B.$$

Denote by v_e (v_f, v_g, v_h) that node of e (f, g, h) whose distance to r is greater. Then by Lemma 11

$$\bar{V}_f = V_r \cup \{v_e\} \setminus \{v_f\}, \quad \bar{V}_g = V_r \cup \{v_e\} \setminus \{v_g\}, \quad \text{and} \quad \bar{V}_h = V_r \cup \{v_e, v_h\} \setminus \{v_f, v_g\}$$

(see Figure 5) also yield double-tight cuts with edge sets (for appropriately chosen $D \subset E_T$)

$$\begin{aligned} \delta(V_r) &= \{e\} \dot{\cup} S_f \dot{\cup} S_g \dot{\cup} D, \\ \delta(\bar{V}_f) &= S_e \dot{\cup} \{f\} \dot{\cup} S_g \dot{\cup} D, \\ \delta(\bar{V}_g) &= S_e \dot{\cup} S_f \dot{\cup} \{g\} \dot{\cup} D, \\ \delta(\bar{V}_h) &= (S_e \setminus \{h\}) \dot{\cup} S_h \dot{\cup} \{f\} \dot{\cup} \{g\} \dot{\cup} D. \end{aligned}$$

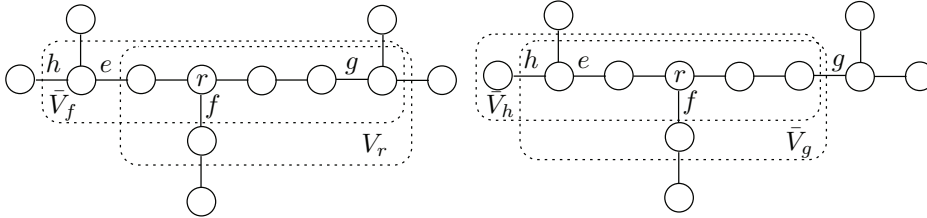


FIG. 5. Node sets V_r , $\bar{V}_f$, $\bar{V}_g$, and $\bar{V}_h$.

Exploiting

$$b_0 = \sum_{\bar{e} \in \delta(V_r)} b_{\bar{e}} = \sum_{\bar{e} \in \delta(\bar{V}_f)} b_{\bar{e}} = \sum_{\bar{e} \in \delta(\bar{V}_g)} b_{\bar{e}} = \sum_{\bar{e} \in \delta(\bar{V}_h)} b_{\bar{e}}$$

one obtains after a few rearrangements together with (16)

$$b_e - \sum_{\bar{e} \in S_e} b_{\bar{e}} = b_f - \sum_{\bar{e} \in S_f} b_{\bar{e}} = b_g - \sum_{\bar{e} \in S_g} b_{\bar{e}} = b_h - \sum_{\bar{e} \in S_h} b_{\bar{e}} = \gamma_B.$$

By induction, $b_{\bar{e}} = \gamma_B \kappa_{\bar{e}}^r$ for all $\bar{e} \in S_e$, so $b_e = \gamma_B(1 + \sum_{\bar{e} \in S_e} \kappa_{\bar{e}}^r) = \gamma_B \kappa_e^r$.

(d) We first show $\gamma_{\bar{B}} = \gamma_B$ for all branches $\bar{B}$ incident on r with respect to a special branch B incident on r . Observe that there is always a branch $B = (V_B, E_B)$ incident on r with $|V_B \setminus \{r\}| < F$ (because there exist at least two branches by minimality of r and because $|T| \leq 2F$), so let B be this branch. Let $\bar{B} = (\bar{V}, \bar{E})$ be

some other branch incident on r . Denote by $\bar{S} \subset \bar{E}$ the set of unreduced edges in $\bar{E}$ that have no unreduced predecessors, i.e.,

$$\bar{S} = \{\bar{e} \in \bar{E} : |V_e^r| = \kappa_e^r \text{ and if } \bar{e} \in S_{e'} \text{ for some } e' \in \bar{E}, \text{ then } |V_{e'}^r| > |\bar{F}|\}.$$

Next, the set $\hat{V} = V_B \cup \bar{V} \cup \bigcup_{\bar{e} \in \bar{S}} V_e^r$ induces a tree $\hat{T} = (\hat{V}, \hat{E})$ with $|\hat{V}| \leq F$ ($|V_B| \leq F$ by construction and removing any leaf $v \in \hat{V} \setminus V_B$ corresponds to cutting a reduced edge f in $\bar{B}$ so that $|\hat{V} \setminus \{v\}| \leq |T| - |V_f^r| < |T| - \bar{F} = F$). Thus, we may extend $\hat{V}$ by nodes from $V \setminus (V_B \cup \bar{V})$ to a set V_r with $|V_r| = F$ so that $\delta(V_r)$ is a double-tight cut with $\bar{S} \subseteq \delta(V_r)$ and $E_B \cap \delta(V_r) = \emptyset$. Take a leaf $e \in E_B$, an edge $f \in \bar{S}$, and denote by v_e and v_f those of the two nodes of e and f whose distance to r is greater. Then $V_e = V_r \cup \{v_f\} \setminus \{v_e\}$ yields another double-tight cut $\delta(V_e) = \delta(V_r) \cup \{e\} \cup S_f \setminus \{f\}$, and therefore $\gamma_B = b_e = b_f - \sum_{\bar{e} \in S_f} b_{\bar{e}} = \gamma_{\bar{B}}$.

As $\gamma_{\bar{B}} = \gamma_B$ for all branches $\bar{B}$ and B , any double-tight cut $\delta(V_r)$ yields with (c)

$$b_0 = \sum_{e \in \delta(V_r)} b_e = \gamma_B \sum_{e \in \delta(V_r)} \kappa_e^r = \gamma_B \bar{F}.$$

Thus, $\gamma_B = \frac{b_0}{\bar{F}}$. $\square$

Proof of Theorem 8. (a) Because of (b) we require $\bar{F} \geq 2$ within the proof of (a). For sufficiency assume that the branchless path condition is satisfied. Given a face $\mathcal{F}_b$ of P_B defined by $b^T y \geq b_0$ that contains all roots of $(\kappa^r)^T y \geq \bar{F}$, Lemma 15 shows that $b^T y \geq b_0$ is a nonnegative multiple of $(\kappa^r)^T y \geq \bar{F}$, so the latter is facet-inducing.

Necessity of the minimal root condition is a consequence of Theorem 7, so consider the case that r is a minimal root but the branchless path condition does not hold. In this case (T, E_T) contains a path on nodes $\bar{V} = \{v_1, \dots, v_F\}$ with $\{v_i, v_{i+1}\} \in E_T$ for $1 \leq i < F$ and neither $\{v_1, v_2\}$ nor $\{v_{F-1}, v_F\}$ is a leaf. By the minimality of r and Lemma 12 there must be some $k \in \{1, \dots, F\}$ with $r = v_k$. Split (T, E_T) at r into two edge disjoint subtrees (V_1, E_1) and (V_2, E_2) with $V_1 \cap V_2 = \{r\}$, $E_T = E_1 \cup E_2$ and set

$$\begin{aligned} S_1 &:= \delta(\{v_1\}) \setminus \{v_1 v_2\}, \\ S_2 &:= \delta(\{v_F\}) \setminus \{v_{F-1} v_F\}, \\ P_1 &:= \{v_i v_{i+1} : 1 \leq i < k\}, \\ P_2 &:= \{v_i v_{i+1} : k \leq i < F\}, \\ D_i &:= E_i \setminus P_i \quad \text{for } i \in \{1, 2\}, \\ n_i &:= \sum_{e \in S_i} |V_e^r| \quad \text{for } i \in \{1, 2\}. \end{aligned}$$

Note that $1 \leq n_i < \bar{F}$ for $i \in \{1, 2\}$, because otherwise the cut S_i would induce a partition requiring at least $n_i + F + 1 > \bar{F} + F = |T|$ nodes. Furthermore, $n_1 + n_2 = \bar{F}$ because $S_1 \cup S_2$ induces a double-tight cut by Lemma 11. We show that all roots of $(\kappa^r)^T y \geq \bar{F}$ also satisfy the equation

$$(17) \quad n_2 \sum_{e \in D_1} \kappa_e^r y_e + n_1 \sum_{e \in P_1} (\bar{F} - \kappa_e^r) y_e - n_2 \sum_{e \in P_2} (\bar{F} - \kappa_e^r) y_e - n_1 \sum_{e \in D_2} \kappa_e^r y_e = 0.$$

Indeed, let V_r be a node set inducing a bisection cut tight for $(\kappa^r)^T y \geq \bar{F}$ and consider the number of times the path can be cut in view of Lemma 10:

$|\delta(V_r) \cap (P_1 \cup P_2)| = 0$: This implies $\delta(V_r) = S_1 \cup S_2$ by Lemma 11 and (17) holds.
 $|\delta(V_r) \cap (P_1 \cup P_2)| = 1$: By symmetry it suffices to consider the case $\delta(V_r) \cap P_1 = \{f\}$.
 By Lemma 10 we have $\delta(V_r) \cap D_1 = \emptyset$ and as the cut is tight, $\sum_{e \in D_2} \kappa_e^r y_e = \bar{F} - \kappa_f^r$, so (17) holds.
 $|\delta(V_r) \cap (P_1 \cup P_2)| \geq 2$: By Lemma 10 it cannot be greater than two, but even two is impossible for a tight cut. Indeed, Lemma 10 ensures $\delta(V_r) \cap (P_1) = \{e_1\}$, $\delta(V_r) \cap (P_2) = \{e_2\}$, and $\delta(V_r) \cap (D_1 \cup D_2) = \emptyset$. Since $n_i < \bar{F}$ we obtain $\kappa_{e_i}^r \geq n_i + 1$ for $i \in \{1, 2\}$, so $(\kappa^r)^T \chi^{\delta(V_r)} = \kappa_{e_1}^r + \kappa_{e_2}^r \geq n_1 + n_2 + 2 > \bar{F}$ yields the desired contradiction.

As (17) is not a scalar multiple of $(\kappa^r)^T y \geq \bar{F}$, the latter cannot be a facet if $\bar{F} \geq 2$ and the branchless path condition is not fulfilled.

(b) If $\bar{F} = 1$, then $\kappa_e^r = 1$ for all $e \in E_T$ and each e alone forms a bisection cut that is tight for $(\kappa^r)^T y \geq \bar{F}$. As these are $|E_T|$ affinely independent roots, the inequality is facet-inducing. $\square$

Remark 16. The statements of Theorem 8 cannot be easily carried forward for graphs denser than trees. So far, we observed that if there are cycles in the graph, then not all trees in this graph yield a facet-defining truncated knapsack tree inequality, even if conditions (a) in Theorem 8 are satisfied. Some additional assumptions must be figured out. For instance, it can be shown that if $F = |T| - 1$ (condition (b) in Theorem 8) and the graph G contains at least one cycle, then $(\kappa^r)^T y \geq \bar{F}$ does not define a facet of P_B .

5. The bisection knapsack walk inequalities for MB. In this section we exploit the special structure of MB in order to derive an improved version of the knapsack tree inequality. Note that in the MNCGP case with $K > 2$ a walk $\{e_1 = \{v_1, v_2\}, e_2 = \{v_2, v_3\}\}$ with $y_{e_1} = y_{e_2} = 1$ does not imply any relation between nodes v_1 and v_3 , while in the MB case where $K = 2$ it follows from $y_{e_1} = y_{e_2} = 1$ that v_1 and v_3 belong to the same cluster.

More generally, whenever there is a walk (for ease of exposition we assume throughout that a walk traverses an edge at most once; for the general case, see [1]) between two nodes of the graph with an even number of edges in the cut, we know in the case of MB that the two end-nodes of the walk have to be in the same cluster. We may therefore replace the indicator term $1 - \sum_{e \in P_{rv}} y_e$ of (3) by

$$(18) \quad 1 - \sum_{e \in P_{rv} \setminus H_v} y_e - \sum_{e \in H_v} (1 - y_e),$$

where $H_v \subseteq P_{rv}$ with even cardinality. So if $y \in \{0, 1\}^{|E|}$ is a valid solution of MB and P_{rv} is a walk from r to v in G with $H_v = \{e \in P_{rv} : y_e = 1\}$ and $|H_v|$ even, then expression (18) is equal to one, indicating that r and v belong to the same cluster. If, however, $H_v \neq \{e \in P_{rv} : y_e = 1\}$, the value of (18) is less than or equal to zero.

LEMMA 17. *Given a root node $r \in V$, walks $P_{rv} \subseteq E$, and even subsets $H_v \subseteq P_{rv}$ for all $v \in V$, let $y = \chi^{\delta(S)}$ for some $S \subseteq V$ with $r \in S$ and put $z = \chi^S$. Then for all $v \in V$*

$$(19) \quad 1 - \sum_{e \in P_{rv} \setminus H_v} y_e - \sum_{e \in H_v} (1 - y_e) \leq z_v.$$

Proof. For $v \in S$ we have $z_v = 1$ and inequality (19) is satisfied, because $y_e \geq 0$ and $1 - y_e \geq 0$ for all $e \in E$. If $v \notin S$, the set $C = \{e \in P_{rv} : y_e = 1\}$ must be of odd

cardinality (otherwise r and v would be together in S_1). Since H_v is of even cardinality and both C and H_v are subsets of P_{rv} , there exists an $e \in P_{rv}$ with $e \in C \setminus H_v$ or $e \in H_v \setminus C$. If $e \in C \setminus H_v$, then $y_e = 1$ and the left-hand side of (19) is smaller or equal to $1 - y_e = 0 = z_v$. If $e \in H_v \setminus C$, then $y_e = 0$ and the left-hand side of (19) is smaller than or equal to $1 - (1 - y_e) = 0 = z_v$. $\square$

Now we are ready to sum up all the evaluation terms.

DEFINITION 18 (bisection knapsack walk inequality). *Let $\sum_{v \in V} a_v x_v \leq a_0$ be a valid inequality for the knapsack polytope P_K with $a_v \geq 0$ for all $v \in V$. For a subset $V' \subseteq V$, a fixed root node $r \in V'$, walks $P_{rv} \subseteq E$, and sets $H_v \subseteq P_{rv}$ with $|H_v|$ even, the bisection knapsack walk inequality reads*

$$(20) \quad \sum_{v \in V'} a_v \left(1 - \sum_{e \in P_{rv} \setminus H_v} y_e - \sum_{e \in H_v} (1 - y_e) \right) \leq a_0.$$

Lemma 17 directly implies the following proposition.

PROPOSITION 19. *The bisection knapsack walk inequality (20) is valid for the polytope P_B .*

Note that knapsack tree inequalities are a special case of the bisection knapsack walk inequalities where the walks P_{rv} form a tree, all nodes on these walks are contained in V' , and all $H_v = \emptyset$. Again, we may rewrite the bisection knapsack walk inequality so as to pronounce its strength in forcing cut variables to increase:

$$\sum_{e \in E} \left(\sum_{v \in V': e \in P_{rv}} a_v - \sum_{v \in V': e \in H_v} 2a_v \right) y_e \geq \sum_{v \in V'} a_v - a_0 - \sum_{v \in V'} a_v |H_v|.$$

Remark 20. For $y = \chi^{\delta(S)}$ with $r \in S \subseteq V$, $z = \chi^S$, and $U_v \subseteq P_{rv}$ with $|U_v|$ odd ($v \in V \setminus \{r\}$), one can show $\sum_{e \in P_{rv} \setminus U_v} y_e + \sum_{e \in U_v} (1 - y_e) \geq z_v$ for $v \in V \setminus \{r\}$. In case of (MB), a valid knapsack inequality $\sum_{v \in V} a_v x_v \leq a_0$ implies the validity of $\sum_{v \in V'} a_v z_v \geq a(V') - a_0$ for all $V' \subseteq V$. Thus the so-called *odd bisection knapsack walk inequality*

$$a_r + \sum_{v \in V' \setminus \{r\}} a_v \left(\sum_{e \in P_{rv} \setminus U_v} y_e + \sum_{e \in U_v} (1 - y_e) \right) \geq a(V') - a_0$$

is valid for P_B , too. Due to their close relation to the (even) bisection knapsack walk inequalities (20) we will not treat these inequalities further in this paper but refer the interested reader to [1].

Remark 21. The bisection knapsack walk inequalities are closely linked to the cycle inequalities [2], which are defined for cycles $C = (V_C, E_C)$ in G and subsets $U \subseteq E_C$ with $|U|$ odd by

$$\sum_{e \in E_C \setminus U} y_e - \sum_{e \in U} y_e \geq 1 - |U|.$$

Indeed, consider the case that in a bisection knapsack walk inequality a path P_{rv} with even subset H_v can be shortcut by the edge rv ; put $E_C = P_{rv} \cup \{rv\}$ and $U = H_v \cup \{rv\}$ and rewrite the corresponding cycle inequality in the form

$$(21) \quad 1 - \sum_{e \in P_{rv} \setminus H_v} y_e - \sum_{e \in H_v} (1 - y_e) \leq 1 - y_{rv}.$$

If $1 - y_{rv}$ is interpreted as z_v (indicating whether v is in the same set as the root r), this is just (19). Thus, whenever all cycle inequalities are enforced and a direct edge rv exists between roots r and v , then $P_{rv} = \{rv\}$ is always the best possible choice. It is, however, not true in general that in the presence of all cycle inequalities a shorter path (with respect to the number of edges) dominates longer paths; see Example 22.

Example 22. Let G be the cycle on five nodes of Figure 6. The solution $y = (y_{12}, y_{23}, y_{34}, y_{45}, y_{15})^T = (0.5, 0.5, 0, 0, 0)^T$ fulfills all cycle inequalities because it is a convex combination of the two cuts $(0, 0, 0, 0, 0)^T$ and $(1, 1, 0, 0, 0)^T$. Now look at the bisection knapsack walk inequalities with $V' = \{1, 3\}$ and $r = 1$. The shorter path P_{13}^s from root node 1 to node 3 uses the edge set $\{\{1, 2\}, \{2, 3\}\}$ with $H_3^s = \emptyset$ or $H_3^s = \{\{1, 2\}, \{2, 3\}\}$, and the longer path P_{13}^l uses the edge set $\{\{3, 4\}, \{4, 5\}, \{1, 5\}\}$ with $H_3^l = \emptyset$, $H_3^l = \{\{3, 4\}, \{4, 5\}\}$, $H_3^l = \{\{3, 4\}, \{1, 5\}\}$, or $H_3^l = \{\{4, 5\}, \{1, 5\}\}$. For the shorter path of the two possible bisection knapsack walk inequalities the left-hand side value is $a_3 \cdot 0$, whereas the best possible bisection knapsack walk inequality on the longer path uses $H_3^l = \emptyset$ and yields left-hand side value $a_3 \cdot 1$.

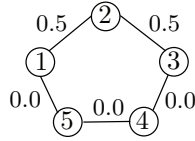


FIG. 6. Graph for the counterexample of Example 22.

6. Capacity improved bisection knapsack walk inequalities and the lower envelope for stars. To motivate another strengthening for bisection knapsack walk inequalities consider the case of a disconnected graph with two components, one of them being a single edge $\{u, v\}$, the other connected one being $V' = V \setminus \{u, v\}$. Even though one cannot include the edge $\{u, v\}$ directly in a bisection knapsack walk inequality rooted at some $r \in V'$, one can at least improve the inequality if $y_{uv} = 1$. In this case u and v belong to different clusters, and therefore the capacity F of both clusters can be reduced by $\min\{\varphi_u, \varphi_v\}$. Since F is the right-hand side of the inequality $\sum_{v \in V} \varphi_v x_v \leq F$ used to define the knapsack polytope P_K , this reduction may help to derive stronger bisection knapsack walk inequalities. For instance, one can look at a given valid inequality $\sum_{v \in V} a_v x_v \leq a_0$ for the original knapsack polytope with capacity F , and in case $y_{uv} = 1$ we are allowed to reduce the right-hand side a_0 by $\min\{a_u, a_v\}$, thus also improving the bisection knapsack walk inequality.

To generalize this idea we define for $\bar{G} \subseteq G$ with $\bar{V} \subseteq V$, $\bar{E} \subseteq E(\bar{V})$ and $a \in \mathbb{R}_+^{|\bar{V}|}$ a function $\beta_{\bar{G}} : \{0, 1\}^{|\bar{E}|} \rightarrow \mathbb{R} \cup \{\infty\}$ with

$$(22) \quad \beta_{\bar{G}}(y) = \inf \left\{ a(S), a(\bar{V} \setminus S) : S \subseteq \bar{V}, \max \{a(S), a(\bar{V} \setminus S)\} \leq a_0, y = \chi^{\delta_G(S)} \right\}.$$

Now we look at the lower convex envelope $\check{\beta}_{\bar{G}} : \mathbb{R}^{|\bar{E}|} \rightarrow \mathbb{R} \cup \{\infty\}$ of $\beta_{\bar{G}}(y)$, i.e.,

$$(23) \quad \check{\beta}_{\bar{G}}(x) = \sup \left\{ \check{\beta}(x) : \check{\beta} : \mathbb{R}^{|\bar{E}|} \rightarrow \mathbb{R}, \check{\beta} \text{ convex}, \check{\beta}(y) \leq \beta_{\bar{G}}(y), y \in \{0, 1\}^{|\bar{E}|} \right\}.$$

Notice that $\check{\beta}_{\bar{G}}$ is a piecewise linear function on its domain. We will see that given a bisection knapsack walk inequality (20) on some $V' \subseteq V$ and $\bar{V} \subseteq V \setminus V'$, subtracting

any affine minorant $c_0 + \sum_{e \in \bar{E}} c_e y_e$ of $\check{\beta}_{\bar{G}}$, i.e.,

$$(24) \quad c_0 + \sum_{e \in \bar{E}} c_e y_e \leq \check{\beta}_{\bar{G}}(y),$$

on the right-hand side of (20) yields again a valid inequality for P_B . It yields an improvement with respect to a given y if the minorant is positive for this y . For convenience, the next proposition states this for several disjoint subsets $\bar{V}$.

PROPOSITION 23. *Let $\sum_{v \in V} a_v x_v \leq a_0$ with $a_v \geq 0$ for all $v \in V$ be a valid inequality for the knapsack polytope P_K . Choose a nonempty $V' \subseteq V$ and subgraphs $(\bar{V}_l, \bar{E}_l) = \bar{G}_l \subset G$ with $\bar{V}_l \cap V' = \emptyset$, $\bar{E}_l \subseteq E(\bar{V}_l)$ for $l = 1, \dots, L$ and pairwise disjoint sets $\bar{V}_l$. Find for each l an affine minorant $c_0^l + \sum_{e \in \bar{E}_l} c_e y_e$ for the convex envelope $\check{\beta}_{\bar{G}_l}$ so that (24) holds for all y in P_B . Then the capacity reduced bisection knapsack walk inequality*

$$(25) \quad \sum_{v \in V'} a_v \left(1 - \sum_{e \in P_{rv} \setminus H_v} y_e - \sum_{e \in P_{rv} \cap H_v} (1 - y_e) \right) \leq a_0 - \sum_{l=1}^L \left(c_0^l + \sum_{e \in \bar{E}_l} c_e y_e \right)$$

is valid for P_B .

Proof. Let $y \in P_B$ such that $y = \chi^{\delta(S)}$ with $S \subseteq V$; then $\varphi(S) \leq F$ and $\varphi(V \setminus S) \leq F$. W.l.o.g., let $r \in S$ and put $z = \chi^S$. Then for all $l = 1, \dots, L$

$$c_0^l + \sum_{e \in \bar{E}_l} c_e y_e \leq \check{\beta}_{\bar{G}_l}(y) \leq \beta_{\bar{G}_l}(y) = \min \{a(\bar{V}_l \cap S), a(\bar{V}_l \setminus S)\} \leq \sum_{v \in \bar{V}_l \cap S} a_v = \sum_{v \in \bar{V}_l \cap S} a_v z_v.$$

Furthermore, by Lemma 17 we have $1 - \sum_{e \in P_{rv} \setminus H_v} y_e - \sum_{e \in H_v} (1 - y_e) \leq z_v$ for $v \in V'$. Thus

$$\begin{aligned} & \sum_{v \in V'} a_v \left(1 - \sum_{e \in P_{rv} \setminus H_v} y_e - \sum_{e \in H_v} (1 - y_e) \right) + \sum_{l=1}^L \sum_{e \in \bar{E}_l} c_e y_e \\ & \leq \sum_{v \in V'} a_v z_v + \sum_{l=1}^L \sum_{v \in \bar{V}_l} a_v z_v \leq \sum_{v \in V} a_v z_v \leq a_0. \quad \square \end{aligned}$$

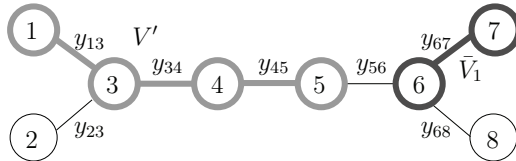


FIG. 7. Graph considered in Example 24 (1). $F = 4$, $\varphi_i = 1$ for all $i \in V$, $\sum_{i \in V} x_i \leq 4$.

Example 24. For the graph G displayed in Figure 7 with $\varphi_v = 1$ for all $v \in V$ the polytope P_B has 74 facets (computed by *polymake* [11]). Among these are 14 trivial facets, only 2 pure bisection knapsack walk facets, 19 truncated bisection knapsack walk facets, 16 capacity reduced bisection knapsack walk facets (some truncated), 4 capacity reduced odd bisection knapsack walk facets, and 19 facets for which we are not yet able to recognize a construction rule. Here we want to give a first simple example for a capacity reduced bisection knapsack walk inequality. Two more involved examples will follow at the end of this section. We use the knapsack inequality $\sum_{v \in V} x_v \leq 4$ in all three examples; thus $a_v = 1$ for all $v \in V$:

- (1) For $V' = \{1, 3, 4, 5\}$, root node $r = 3$ and $H_v = \emptyset$ for all $v \in V'$ the bisection knapsack walk inequality is $1 + (1 - y_{13}) + (1 - y_{34}) + (1 - y_{34} - y_{45}) \leq 4$. We choose $\bar{G} = (\bar{V}, \bar{E})$ with $\bar{V} = \{6, 7\}$ and $\bar{E} = \{67\}$. We will see that the unique best minorizing function for $\check{\beta}_{\bar{G}}$ is y_{67} , and thus the bisection knapsack walk inequality can be strengthened to $1 + (1 - y_{13}) + (1 - y_{34}) + (1 - y_{34} - y_{45}) \leq 4 - y_{67}$. Now rewrite this inequality to $y_{13} + 2y_{34} + y_{45} - y_{67} \geq 0$ and observe that, as in (5), the coefficient of y_{34} can be strengthened to 1 in order to find the facet $y_{13} + y_{34} + y_{45} - y_{67} \geq 0$ of P_B .

To find inequalities (24) to apply in Proposition 23 we take a closer look at the lower envelope defined in (23). In certain cases, e.g., for the case of $\bar{G} = (\bar{V}, \bar{E})$ being a star with $a(\bar{V}) \leq a_0$, we are able to give a full description of $\check{\beta}_{\bar{G}}$ by giving a complete description of the cluster weight polytope defined below. This will provide the tightest improvement possible in (25).

DEFINITION 25. Given a graph $G = (V, E)$ with node weights $a_v \geq 0$ for $v \in V$, for a set $S \subseteq V$ we define the following point in $\mathbb{R}^{|E|+1}$:

$$h_G^S = \begin{pmatrix} a(S) \\ \chi^{\delta(S)} \end{pmatrix}.$$

With respect to a given nonnegative $a_0 \in \mathbb{R}$ we define

$$P_{CW}(G) = \text{conv}\{h_G^S : S \subseteq V, a(S) \leq a_0, a(V \setminus S) \leq a_0\}$$

and call this set the cluster weight polytope.

As usual, we will drop G in h_G^S and $P_{CW}(G)$ if the graph is clear from the context. The purpose of studying $P_{CW}(\bar{G})$ is that its polyhedral description immediately yields the epigraph of $\check{\beta}_{\bar{G}}$ via $\text{epi}(\check{\beta}_{\bar{G}}) = P_{CW}(\bar{G}) + \{\lambda(1, 0^T)^T : \lambda \geq 0\}$. This is the content of the next proposition.

PROPOSITION 26. Given a subgraph $\bar{G} = (\bar{V}, \bar{E})$ of G with node weights $a_v \geq 0$ for $v \in V$, an inequality of the form $y_0 + \sum_{e \in \bar{E}} \gamma_e y_e \geq \gamma_0$ is valid for $P_{CW}(\bar{G})$ if and only if $\gamma_0 - \sum_{e \in \bar{E}} \gamma_e y_e$ is an affine minorant of $\check{\beta}_{\bar{G}}$.

Proof. $y_0 + \sum_{e \in \bar{E}} \gamma_e y_e \geq \gamma_0$ is valid for $P_{CW}(\bar{G})$ if and only if

$$y_0 + \sum_{e \in \bar{E}} \gamma_e y_e \geq \gamma_0 \text{ for } \begin{pmatrix} y_0 \\ y \end{pmatrix} \in \{h_{\bar{G}}^S : S \subseteq \bar{V}, a(S) \leq a_0, a(\bar{V} \setminus S) \leq a_0\}$$

if and only if

$$y_0 \geq \gamma_0 - \sum_{e \in \bar{E}} \gamma_e y_e \text{ for } \left\{ \begin{pmatrix} \min\{a(S), a(\bar{V} \setminus S)\} \\ \chi^{\delta_{\bar{G}}(S)} \end{pmatrix} : S \subseteq \bar{V}, \max\{a(S), a(\bar{V} \setminus S)\} \leq a_0 \right\}$$

if and only if $\beta_{\bar{G}}(y) \geq \gamma_0 - \sum_{e \in \bar{E}} \gamma_e y_e$ for $y \in \{0, 1\}^{|\bar{E}|}$ (by (22), recall that $\inf \emptyset = \infty$ by definition) if and only if $\check{\beta}_{\bar{G}}(y) \geq \gamma_0 - \sum_{e \in \bar{E}} \gamma_e y_e$ for $y \in \mathbb{R}^{|\bar{E}|}$ (see (23)). $\square$

Hence, the “lower” facets of P_{CW} are in one-to-one correspondence to the linear components of $\check{\beta}$. For a star $\bar{G} = (\bar{V}, \bar{E})$ we are able to exhibit facets of $P_{CW}(\bar{G})$, which in certain problems enables us to strengthen bisection knapsack walk inequalities of P_B to facet-defining inequalities of P_B (see Example 40 at the end of this section).

Let us first look at a symmetry of P_{CW} for general graphs $G = (V, E)$, a property which we will later use frequently to reduce our efforts in the proofs.

PROPOSITION 27. P_{CW} is symmetric to the hyperplane $\{y \in \mathbb{R}^{|E|} : 2y_0 = a(V)\}$.

Proof. Observe that for any point h^S used in the definition of P_{CW} the point $h^{V \setminus S}$ is contained in P_{CW} , too. Since $\chi^{\delta(S)} = \chi^{\delta(V \setminus S)}$, we have for all those pairs $(h^S, h^{V \setminus S})$

$$\left(\frac{\frac{1}{2}a(V)}{\chi^{\delta(S)}} \right) - h^S = h^{V \setminus S} - \left(\frac{\frac{1}{2}a(V)}{\chi^{\delta(S)}} \right). \quad \square$$

Another useful result for a star $G = (V, E)$ is the following.

LEMMA 28. Let $G = (V, E)$ be a star with center $r \in V$, $a_v \geq 0$ for all $v \in V$, and $a_{v'} = a(V \setminus \{v'\})$ for at least one $v' \in V \setminus \{r\}$. Then $a(S) = a(V \setminus S)$ for all $S \subseteq V$ with $v' \in S$ and $r \in V \setminus S$ if and only if $a_{v'} = a_r$ and $a_v = 0$ for all $v \in V \setminus \{v', r\}$.

Proof. Sufficiency is obvious. We will show necessity: Suppose $a(S) = a(V \setminus S)$ for all $S \subseteq V$ with $v' \in S$ and $r \in V \setminus S$. Then, in particular, this is true for $V \setminus S = \{r\}$, i.e., $a_r = a(V \setminus \{r\}) = a_{v'} + a(V \setminus \{v', r\}) = a(V \setminus \{v'\}) + a(V \setminus \{v', r\}) = a_r + 2a(V \setminus \{v', r\})$. Thus, $a_v = 0$ for all $v \in V \setminus \{v', r\}$ and $a_{v'} = a_r$. $\square$

In the remaining part of the section we will look into P_{CW} for stars $G = (V, E)$ with center node $r \in V$ and the constraint $\sum_{v \in V} a_v x_v \leq a_0$. At first we determine the dimension of the polytope.

PROPOSITION 29. For a star $G = (V, E)$ with center $r \in V$ and $a \geq 0$ with $a(V) \leq a_0$, the polytope P_{CW} has full dimension $|E| + 1$ for $a \neq 0^{|E|}$ and dimension $|E| = |V|$ for $a = 0^{|E|}$.

Proof. Since G is a star and by assumption $a(V) \leq a_0$, the $1 + |E|$ points $h^\emptyset$ and $h^{\{v\}}$ for all $v \in V \setminus \{r\}$ are contained in P_{CW} and affinely independent. Thus the dimension of P_{CW} is at least $|E|$. If $a \neq 0^{|E|}$, then h^V is affinely independent from all points listed previously, and thus P_{CW} is full-dimensional with dimension $|E| + 1$. For $a = 0^{|E|}$ all points lie on the hyperplane $y_0 = 0$. $\square$

For $G = (V, E)$ a star with center $r \in V$ and weights $a_v = 0$ for all $v \in V$ and $a_0 \geq 0$, it can easily be worked out that P_{CW} is completely described by the equality $y_0 = 0$ and the inequalities $0 \leq y_{rv} \leq 1$ for all $v \in V \setminus \{r\}$. So from now on we assume $a_v > 0$ for at least one $v \in V$. Let us first state trivial valid inequalities and facets of P_{CW} .

PROPOSITION 30. For a star $G = (V, E)$ with center $r \in V$, $a \geq 0$ with $a \neq 0^{|E|}$, and $a(V) \leq a_0$ the trivial inequalities

$$(26) \quad 0 \leq y_{rv} \leq 1 \quad \forall v \in V \setminus \{r\}$$

are facet-inducing for P_{CW} except for one particular case: if there is exactly one $v' \in V \setminus \{r\}$ with $a_{v'} = a_r = \frac{1}{2}a(V)$, then $y_{rv'} \leq 1$ does not induce a facet.

Proof. The validity of the inequalities $y_{rv'} \geq 0$ and $y_{rv'} \leq 1$ for all $v' \in V \setminus \{r\}$ follows from the definition of P_{CW} . In general, to prove that a valid inequality defines a facet of P_{CW} we have to find $\dim(P_{\text{CW}})$ affinely independent points of P_{CW} which fulfill it with equality. From Proposition 29 we know that $\dim(P_{\text{CW}}) = |V|$ if $a \neq 0^{|E|}$. For $y_{rv'} \geq 0$ we choose the $|V|$ points $h^\emptyset$, h^V , and $h^{\{v\}}$ for all $v \in V \setminus \{r, v'\}$. For $y_{rv'} \leq 1$ the accumulation of affinely independent points on the inequality is a bit more involved: If $a_{v'} \neq a(V \setminus \{v'\})$, we can choose the $|V|$ points $h^{\{v'\}}$, $h^{V \setminus \{v'\}}$, and $h^{\{v', v\}}$ for all $v \in V \setminus \{r\}$ with $v \neq v'$. If $a_{v'} = a(V \setminus \{v'\})$, we look at two cases:

1. $a_r \neq a_{v'}$: Then there is a $\tilde{v} \in V \setminus \{r, v'\}$ with $a_{\tilde{v}} > 0$. Furthermore, since $a_{v'} = a(V \setminus \{v'\})$, we have $a_{v'} = \frac{1}{2}a(V)$. Together with $a_{\tilde{v}} > 0$ this implies $a(\{v', \tilde{v}\}) \neq a(V \setminus \{v', \tilde{v}\})$, i.e., $h^{\{v', \tilde{v}\}} \neq h^{V \setminus \{v', \tilde{v}\}}$. Thus we can choose the $|V|$ points $h^{\{v'\}}$, $h^{\{v', v\}}$ for all $v \in V \setminus \{r, v'\}$ and $h^{V \setminus \{v', \tilde{v}\}}$.

2. $a_r = a_{v'}$: The set of points contained in the definition of P_{CW} which fulfill $y_{rv'} = 1$ is $\{h^S, h^{V \setminus S} : S \subseteq V, v' \in S, r \in V \setminus S\}$. Lemma 28 implies for every pair $(h^S, h^{V \setminus S})$ in this set that $a(S) = a(V \setminus S)$. Since $a(S) + a(V \setminus S) = a(V)$, we get $a(S) = \frac{1}{2}a(V)$ for all S with $y = \chi^{\delta(S)}$ and $y_{rv'} = 1$. Thus all vertices of P_{CW} fulfilling $y_{rv'} = 1$ live in the hyperplane $\{y \in \mathbb{R}^{|E|+1} : y_0 = \frac{1}{2}a(V)\}$. Therefore, $y_{rv'} \leq 1$ cannot induce a facet of P_{CW} . $\square$

In the following two propositions we look into nontrivial facets of P_{CW} . Proposition 31 deals with the case $a(V \setminus \{r\}) > a_r$ and Proposition 32 with the case $a(V \setminus \{r\}) \leq a_r$.

PROPOSITION 31. *Given a star $G = (V, E)$ with center $r \in V$, $a \geq 0$ with $a \neq 0^{|E|}$, $a(V) \leq a_0$, and $a(V \setminus \{r\}) > a_r$, we call a triple $(V_p, \bar{v}, V_n)$ feasible if it fulfills $V = \{r, \bar{v}\} \dot{\cup} V_p \dot{\cup} V_n$ and $a(V_p) \leq \frac{1}{2}a(V) < a(V_p) + a_{\bar{v}}$. For all feasible triples $(V_p, \bar{v}, V_n)$ the inequalities*

$$(27) \quad y_0 + \sum_{v \in V_p} a_v y_{rv} + (a(V) - 2a(V_p) - a_{\bar{v}}) y_{r\bar{v}} - \sum_{v \in V_n} a_v y_{rv} \leq a(V),$$

$$(28) \quad y_0 - \sum_{v \in V_p} a_v y_{rv} - (a(V) - 2a(V_p) - a_{\bar{v}}) y_{r\bar{v}} + \sum_{v \in V_n} a_v y_{rv} \geq 0$$

are facet-inducing for P_{CW} .

Note that it is possible that either V_p or V_n of feasible triples $(V_p, \bar{v}, V_n)$ might be empty, but for $a(V \setminus \{r\}) > a_r$ there always is the special element $\bar{v}$.

Proof of Proposition 31. To reduce our efforts in this proof and the ones to follow, observe that for each feasible triple $(V_p, \bar{v}, V_n)$ the corresponding pair of inequalities (27) and (28) is symmetric to the hyperplane $\{y \in \mathbb{R}^{|E|} : 2y_0 = a(V)\}$. To see this, subtract the equation $2y_0 = a(V)$ from (27) to obtain (28). Thus, by Proposition 27, it suffices to show that (27) is valid and facet-defining. Furthermore, to show the validity of (27) it is sufficient to look only at the “upper” points defining P_{CW} , i.e., if w.l.o.g. $S \subseteq V$ such that $a(S) \geq a(V \setminus S)$, then we only need to check validity of (27) for $h^S = (a(S), \chi^{\delta(S)})^T$.

Consider an arbitrary $S \subseteq V$ such that $a(S) \geq a(V \setminus S)$. Let $V^1 = \{v \in V : rv \in \delta(S)\}$. We discern the following four cases:

1. $\bar{v} \in V^1 = S$: For $\begin{pmatrix} a(S) \\ \chi^{\delta(S)} \end{pmatrix}$ the left-hand side of (27) equals

$$\begin{aligned} a(V^1) + a(V_p \cap V^1) + a(V) - 2a(V_p) - a_{\bar{v}} - a(V_n \cap V^1) \\ = 2a(V_p \cap V^1) + a(V) - 2a(V_p) \\ = a(V) - 2a(V_p \setminus V^1) \leq a(V), \end{aligned}$$

where the first equality uses $a(V^1) = a(V_p \cap V^1) + a_{\bar{v}} + a(V_n \cap V^1)$ and the inequality is due to $a(V_p \setminus V^1) \geq 0$.

2. $\bar{v} \notin V^1 = S$: For $\begin{pmatrix} a(S) \\ \chi^{\delta(S)} \end{pmatrix}$ the left-hand side of (27) equals

$$a(V^1) + a(V_p \cap V^1) - a(V_n \cap V^1) = 2a(V_p \cap V^1) \leq 2a(V_p) \leq a(V),$$

where the equality uses $a(V^1) = a(V_p \cap V^1) + a(V_n \cap V^1)$ and the last inequality is due to $a(V_p) \leq \frac{1}{2}a(V)$ by the definition of V_p .

3. $\bar{v} \in V^1 = V \setminus S$: For $\binom{a(S)}{\chi^{\delta(S)}}$ the left-hand side of (27) equals

$$\begin{aligned} a(V) - a(V^1) + a(V_p \cap V^1) + a(V) - 2a(V_p) - a_{\bar{v}} - a(V_n \cap V^1) \\ = 2a(V) - 2a(V_p) - 2a_{\bar{v}} - 2a(V_n \cap V^1) \\ < 2a(V) - a(V) - 2a(V_n \cap V^1) \leq a(V), \end{aligned}$$

where the first equality uses $a(V^1) = a(V_p \cap V^1) + a_{\bar{v}} + a(V_n \cap V^1)$, the strict inequality is due to $a(V_p) + a_{\bar{v}} > \frac{1}{2}a(V)$ by the definition of V_p and $\bar{v}$, and the inequality holds since $a(V_n \cap V^1) \geq 0$.

4. $\bar{v} \notin V^1 = V \setminus S$: For $\binom{a(S)}{\chi^{\delta(S)}}$ the left-hand side of (27) equals

$$a(V) - a(V^1) + a(V_p \cap V^1) - a(V_n \cap V^1) = a(V) - 2a(V_n \cap V^1) \leq a(V),$$

where the first equality uses $a(V^1) = a(V_p \cap V^1) + a(V_n \cap V^1)$ and the inequality is due $a(V_n \cap V^1) \geq 0$.

In order to show that (27) is also facet-defining, let $V_p = \{v_1^p, \dots, v_{|V_p|}^p\}$ and $V_n = \{v_1^n, \dots, v_{|V_n|}^n\}$. Then the $|V|$ points

$$\begin{aligned} h^V, h^{V \setminus \{v_1^p\}}, \dots, h^{V \setminus \{v_1^p, \dots, v_{|V_p|}^p\}}, h^{\{v_1^p, \dots, v_{|V_p|}^p, \bar{v}\}}, \\ h^{\{v_1^p, \dots, v_{|V_p|}^p, \bar{v}, v_1^n\}}, \dots, h^{\{v_1^p, \dots, v_{|V_p|}^p, \bar{v}, v_1^n, \dots, v_{|V_n|}^n\}} \end{aligned}$$

fulfill inequality (27) with equality and are affinely independent, and thus (27) is a facet-inducing inequality. $\square$

In the case of $a(V \setminus \{r\}) \leq a_r$ the set V_n is empty, there is no $\bar{v}$, and the inequalities (27) and (28) take the following form.

PROPOSITION 32. *For a star $G = (V, E)$ with root $r \in V$ and $a \geq 0$ with $a \neq 0^{|E|}$, $a(V) \leq a_0$, and $a(V \setminus \{r\}) \leq a_r$ the inequalities*

$$(29) \quad y_0 + \sum_{v \in V \setminus \{r\}} a_v y_{e_v} \leq a(V),$$

$$(30) \quad y_0 - \sum_{v \in V \setminus \{r\}} a_v y_{e_v} \geq 0$$

are facet-inducing for P_{CW} .

Proof. We start again by observing the symmetry of inequalities (29) and (30) to the hyperplane $\{y \in \mathbb{R}^{|E|} : 2y_0 = a(V)\}$. To see this, subtract the equation $2y_0 = a(V)$ from inequality (29) to obtain inequality (30). Thus, by Proposition 27, it suffices to prove the validity and facet-induction of (29). Take an $S \subseteq V$ with $a(S) \geq a(V \setminus S)$. Then $h^S = \binom{a(S)}{\chi^{\delta(S)}}$ is one of the points defining P_{CW} . We see that $V \setminus S = \{v \in V : rv \in \delta(S)\}$. Now plug h^S into the left-hand side of (29) to get $a(S) + a(V \setminus S) = a(V)$. The point $h^{V \setminus S} = \binom{a(V \setminus S)}{\chi^{\delta(V \setminus S)}}$ can also not violate (29) since $a(V \setminus S) \leq a(S)$, and thus (29) is valid for P_{CW} .

In order to show that (29) is facet-inducing let $v_1, \dots, v_{|V|-1}$ be an arbitrary ordering of the nodes in $V \setminus \{r\}$. Since $a(S) + a(V \setminus S) = a(V)$ holds, the $\dim(P_{\text{CW}}) = |V|$ points

$$h^V, h^{V \setminus \{v_1\}}, \dots, h^{V \setminus \{v_1, \dots, v_{|V|-1}\}}$$

fulfill inequality (29) with equality and are affinely independent. $\square$

All possible facets of P_{CW} fall into one of the following three classes:

$$(31) \quad y_0 + \sum_{v \in V \setminus \{r\}} \gamma_v y_{rv} \leq \gamma_0,$$

$$(32) \quad \sum_{v \in V \setminus \{r\}} \gamma_v y_{rv} \leq \gamma_0,$$

$$(33) \quad -y_0 + \sum_{v \in V \setminus \{r\}} \gamma_v y_{rv} \leq \gamma_0.$$

In the next two lemmas we will look closer into coefficients of facets of the form (31). The following three propositions state that we have found all facets of P_{CW} of the forms (31), (32), and (33), respectively. Finally, Theorem 38 summarizes the results. The section is accompanied by two small examples on how to apply the inequalities to derive capacity reduced bisection knapsack walk inequalities.

LEMMA 33. *For an arbitrary facet of P_{CW} of the form (31) we have for all $v \in V \setminus \{r\}$*

$$-a_v \leq \gamma_v \leq a_v.$$

Proof. We give the proof for the case $\gamma_{\tilde{v}} > 0$ (the case $\gamma_{\tilde{v}} < 0$ can be proved by analogous arguments). The facet has a root $(\hat{y}_0, \hat{y}^T)^T$ with $\hat{y}_{r\tilde{v}} = 0$, because otherwise all roots $\hat{y}$ would lie on the equation $\hat{y}_{r\tilde{v}} = 1$; thus (31) could not induce a facet. Let $S \subseteq V$ be the corresponding subset satisfying $\hat{y} = \chi^{\delta(S)}$ and $\hat{y}_0 = a(S) \geq a(V \setminus S)$. To bound $\gamma_{\tilde{v}}$ we look at $\bar{y} = \hat{y} + e_{r\tilde{v}}$, i.e., the cut $\delta(S) \cup \{r\tilde{v}\}$. We discern three cases concerning the location of node $\tilde{v}$ and the size of the bigger cluster:

1. $\tilde{v} \in V \setminus S$: Because $a(S) \geq a(V \setminus S)$ we obtain $a(S \cup \{\tilde{v}\}) \geq a(V \setminus (S \cup \{\tilde{v}\}))$. Set $\bar{y}_0 = a(S \cup \{\tilde{v}\})$, i.e., $(\bar{y}_0, \bar{y}^T)^T = h^{S \cup \{\tilde{v}\}} \in P_{CW}$. In order for (31) to be feasible for $(\bar{y}_0, \bar{y}^T)^T$ we need $\gamma_0 \geq \bar{y}_0 + \sum_{v \in V \setminus \{r\}} \gamma_v \bar{y}_{rv}$. As $(\hat{y}_0, \hat{y}^T)^T$ is a root of (31), we have $\gamma_0 = \hat{y}_0 + \sum_{v \in V \setminus \{r\}} \gamma_v \hat{y}_{rv}$. Thus, $\hat{y}_0 + \sum_{v \in V \setminus \{r\}} \gamma_v \hat{y}_{rv} \geq \bar{y}_0 + \sum_{v \in V \setminus \{r\}} \gamma_v \bar{y}_{rv}$, i.e., $\hat{y}_0 \geq \bar{y}_0 + \gamma_{\tilde{v}}$. Therefore, $\gamma_{\tilde{v}} \leq \hat{y}_0 - \bar{y}_0 = -a_{\tilde{v}}$. This contradicts our assumption $\gamma_{\tilde{v}} > 0$; thus the case $\tilde{v} \in V \setminus S$ is not possible.
2. $\tilde{v} \in S$ and $a(S \setminus \{\tilde{v}\}) \geq a((V \setminus S) \cup \{\tilde{v}\})$: Set $\bar{y}_0 = a(S \setminus \{\tilde{v}\})$, i.e., $(\bar{y}_0, \bar{y}^T)^T = h^{S \setminus \{\tilde{v}\}} \in P_{CW}$. As $(\bar{y}_0, \bar{y}^T)^T$ is feasible for (31) we derive, as in the previous case, $\hat{y}_0 \geq \bar{y}_0 + \gamma_{\tilde{v}}$, and hence $\gamma_{\tilde{v}} \leq \hat{y}_0 - \bar{y}_0 = a(S) - a(S \setminus \{\tilde{v}\}) = a_{\tilde{v}}$.
3. $\tilde{v} \in S$ and $a(S \setminus \{\tilde{v}\}) < a((V \setminus S) \cup \{\tilde{v}\})$: This implies $a(S \setminus \{\tilde{v}\}) < \frac{1}{2}a(V)$. Set $\bar{y}_0 = a((V \setminus S) \cup \{\tilde{v}\})$, i.e., $(\bar{y}_0, \bar{y}^T)^T = h^{(V \setminus S) \cup \{\tilde{v}\}} \in P_{CW}$. From the feasibility of (31) we conclude $\hat{y}_0 \geq \bar{y}_0 + \gamma_{\tilde{v}}$. Therefore, $\gamma_{\tilde{v}} \leq \hat{y}_0 - \bar{y}_0 = a(S) - a((V \setminus S) \cup \{\tilde{v}\}) = a_{\tilde{v}} + 2a(S \setminus \{\tilde{v}\}) - a(V) < a_{\tilde{v}}$, where the last inequality uses $a(S \setminus \{\tilde{v}\}) - \frac{1}{2}a(V) < 0$. $\square$

LEMMA 34. *For an arbitrary facet of P_{CW} of the form (31) we have $\gamma_0 = a(V)$ and $\sum_{v \in V \setminus \{r\}} \gamma_v \leq a_r$.*

Proof. In order for (31) to be valid for $h^V = (a(V), (\chi^{\delta(V)})^T)^T \in P_{CW}$ we get $\gamma_0 \geq a(V)$. We discern two cases regarding the weight of the root node r .

$a_r < a(V \setminus \{r\})$: (31) has to be valid for $(a(V \setminus \{r\}), (\chi^{\delta(V \setminus \{r\})})^T)^T = h^{V \setminus \{r\}} \in P_{CW}$, and thus $\sum_{v \in V \setminus \{r\}} \gamma_v \leq \gamma_0 - a(V \setminus \{r\})$.

$a_r \geq a(V \setminus \{r\})$: (31) has to be valid for $(a_r, (\chi^{\delta(\{r\})})^T)^T = h^{\{r\}} \in P_{\text{CW}}$, and thus $\sum_{v \in V \setminus \{r\}} \gamma_v \leq \gamma_0 - a_r \leq \gamma_0 - a(V \setminus \{r\})$.

Thus in any case we have

$$(34) \quad \sum_{v \in V \setminus \{r\}} (a_v + \gamma_v) \leq \gamma_0.$$

Now use $a_v + \gamma_v \geq 0$ (by Lemma 33) and $y_{rv} \in [0, 1]$ for all $(y_0, y^T)^T \in P_{\text{CW}}$ to conclude that

$$(35) \quad \sum_{v \in V \setminus \{r\}} (a_v + \gamma_v) y_{rv} \leq \gamma_0$$

is a valid inequality for P_{CW} . Additionally, $a_r = a(V) - a(V \setminus \{r\})$; thus it is sufficient to show that $\gamma_0 = a(V)$ if (31) induces a facet of P_{CW} , because then (34) implies $\sum_{v \in V \setminus \{r\}} \gamma_v \leq a_r$.

So assume, for contradiction, that $\gamma_0 > a(V)$ with (31) facet-defining. In search for roots of (31) let $S \subseteq V$ be such that $\tilde{y}_0 = a(S) \geq a(V \setminus S)$, put $\tilde{y} = \chi^{\delta(S)}$, and consider the following two cases:

$r \in S$: Then $\tilde{y}_0 + \sum_{v \in V \setminus \{r\}} \gamma_v \tilde{y}_{rv} = a(S) + \sum_{v \in V \setminus S} \gamma_v \leq a(V) < \gamma_0$, where the $\leq$ -inequality is due to $\gamma_v \leq a_v$ by Lemma 33. Therefore, $(\tilde{y}_0, \tilde{y}^T)^T$ cannot lie on the facet.

$r \in V \setminus S$: We show that all such roots also satisfy (35) with equality and so (31) cannot be a facet. Indeed, if it is a root, $(\tilde{y}_0, \tilde{y}^T)^T$ satisfies $\tilde{y}_0 + \sum_{v \in V \setminus \{r\}} \gamma_v \tilde{y}_{rv} = \gamma_0$. Since $\tilde{y}_0 = \sum_{v \in V \setminus \{r\}} a_v \tilde{y}_{rv}$, we obtain $\gamma_0 = \sum_{v \in V \setminus \{r\}} (a_v + \gamma_v) \tilde{y}_{rv}$.

Hence, any facet inducing inequality (31) has $\gamma_0 = a(V)$. $\square$

PROPOSITION 35. *For a star $G = (V, E)$ with root $r \in V$, $a \geq 0$ with $a \neq 0^{|E|}$, and $a(V) \leq a_0$ all facets of the form (31) for P_{CW} are defined by (27) if $a(V \setminus \{r\}) > a_r$ and by (29) if $a(V \setminus \{r\}) \leq a_r$.*

Proof. We have shown in Lemma 33 that each coefficient γ_v for all $v \in V \setminus \{r\}$ in all facets of P_{CW} of the form (31) fulfills

$$(36) \quad -a_v \leq \gamma_v \leq a_v.$$

Lemma 34 tells us that for each individual facet of P_{CW} of the form (31) the coefficients fulfill

$$(37) \quad \sum_{v \in V \setminus \{r\}} \gamma_v \leq a_r$$

and

$$(38) \quad \gamma_0 = a(V).$$

For any given $y \in [0, 1]^{|E|}$ we will now determine the best γ_0 and γ subject to the constraints (36), (37), and (38), so that the right-hand side of $y_0 \leq \gamma_0 - \sum_{v \in V \setminus \{r\}} \gamma_v y_{rv}$ is as small as possible. If we can always exhibit an optimal solution γ_0^*, γ^* that corresponds to the coefficients of (27) if $a(V \setminus \{r\}) > a_r$ or of (29) if $a(V \setminus \{r\}) \leq a_r$,

then the proof is complete. At first note that (38) directly fixes γ_0 to $a(V)$, which corresponds to the right-hand sides of (27) and (29). Now look at the problem

$$(39) \quad \begin{aligned} \min \quad & a(V) - \sum_{v \in V \setminus \{r\}} y_{rv} \gamma_v \\ \text{s.t.} \quad & \sum_{v \in V \setminus \{r\}} \gamma_v \leq a_r, \\ & -a_v \leq \gamma_v \leq a_v \quad \forall v \in V \setminus \{r\}. \end{aligned}$$

In case $a(V \setminus \{r\}) \leq a_r$, the optimal solution sets $\gamma_v = a_v$ for all $v \in V \setminus \{r\}$ and we have determined inequality (29). So assume $a(V \setminus \{r\}) > a_r$. Using the variable transformation $\tilde{\gamma}_v = \gamma_v + a_v$, problem (39) is equivalent to

$$(40) \quad \begin{aligned} \max \quad & \sum_{v \in V \setminus \{r\}} y_{rv} \tilde{\gamma}_v - \sum_{v \in V \setminus \{r\}} y_{rv} a_v \\ \text{s.t.} \quad & \sum_{v \in V \setminus \{r\}} \tilde{\gamma}_v \leq a(V), \\ & 0 \leq \tilde{\gamma}_v \leq 2a_v \quad \forall v \in V \setminus \{r\}. \end{aligned}$$

Problem (40) is the canonical continuous bounded knapsack problem (see sections 3.2 and 3.3.1 in [16]) with continuous variables γ_v , profits y_{rv} , weights 1, and upper bound $2a_v$ for all items $v \in V \setminus \{r\}$ and knapsack capacity $a(V)$. An optimal solution can be found by sorting the items v with respect to nonincreasing profit-to-weight ratios $y_{rv}/1$ —w.l.o.g., let this ordering be $1, 2, \dots, |V| - 1$ —and by using this ordering to pack the knapsack in the following way: $\tilde{\gamma}_v = 2a_v$ for all $v = 1, \dots, \bar{v} - 1$ with $\bar{v}$ so that $2a(\{1, \dots, \bar{v} - 1\}) \leq a(V)$ and $2a(\{1, \dots, \bar{v} - 1\}) + 2a_{\bar{v}} > a(V)$, $\tilde{\gamma}_{\bar{v}} = a(V) - 2a(\{1, \dots, \bar{v} - 1\})$, and $\tilde{\gamma}_v = 0$ for all $v = \bar{v} + 1, \dots, |V| - 1$. The item $\bar{v}$ is called the critical item. Note that if one $\bar{v}$ can be chosen as the critical item, then so can all $v \neq \bar{v}$ with $y_{rv} = y_{r\bar{v}}$.

Backsubstitution of $\tilde{\gamma}_v = \gamma_v + a_v$ yields the optimal solution of problem (39): $\gamma_v = a_v$ for all $v = 1, \dots, \bar{v} - 1$ with $a(\{1, \dots, \bar{v} - 1\}) \leq \frac{1}{2}a(V)$ and $a(\{1, \dots, \bar{v} - 1\}) + a_{\bar{v}} > \frac{1}{2}a(V)$, $\gamma_{\bar{v}} = a(V) - 2a(\{1, \dots, \bar{v} - 1\}) - a_{\bar{v}}$, and $\gamma_v = -a_v$ for all $v = \bar{v} + 1, \dots, |V| - 1$. Finally we observe that we have determined a feasible triple $(V_p = \{1, \dots, \bar{v} - 1\}, \bar{v}, V_n = \{\bar{v} + 1, \dots, |V| - 1\})$, i.e., we have found an inequality of (27). $\square$

PROPOSITION 36. *For a star $G = (V, E)$ with root $r \in V$, $a \geq 0$ with $a \neq 0^{|E|}$, and $a(V) \leq a_0$ all facets of the form (33) for P_{CW} are defined by (28) if $a(V \setminus \{r\}) > a_r$ and by (30) if $a(V \setminus \{r\}) \leq a_r$.*

Proof. Use the symmetry of P_{CW} , of pairs (31) and (33) with the same γ_v and γ_0 , of pairs (27) and (28) and of pairs (29) and (30) to the hyperplane $\{y \in \mathbb{R}^{|E|} : 2y_0 = a(V)\}$, and apply Proposition 35. $\square$

PROPOSITION 37. *For a star $G = (V, E)$ with root $r \in V$, $a \geq 0$ with $a \neq 0^{|E|}$, and $a(V) \leq a_0$ all facets of the form (32) for P_{CW} are defined by (26).*

Proof. It is trivial to show that facets of a polytope with coefficient zero for a fixed variable are also facets of the projection of this polytope if one projects out this variable. Since the hyperplanes defined by inequalities of the form (32) have coefficient zero for variable y_0 , we have to look at the projection of P_{CW} onto the space $\mathbb{R}^{|E|}$ and show that this projection only has facets of the form (26). A point $(a(S), (\chi^{\delta(S)})^T)^T \in \mathbb{R}^{|E|+1}$ used to define P_{CW} is projected to $\chi^{\delta(S)} \in \mathbb{R}^{|E|}$, and since

$a(V) \leq a_0$ the polytope P_{CW} contains the points $(a(S), (\chi^{\delta(S)})^T)^T \in \mathbb{R}^{|E|+1}$ for all $S \subseteq V$; thus its projection contains all possible points $\{0, 1\}^{|E|}$. Furthermore, the projection of any other point of P_{CW} can be written as the convex combination of points $\{0, 1\}^{|E|}$. Thus the projection of P_{CW} is exactly the $|E|$ -dimensional hypercube. To finish the proof we note that the $|E|$ -dimensional hypercube is completely described by the projection of the inequalities (26). $\square$

THEOREM 38. For a star $G = (V, E)$ with root $r \in V$, $a \geq 0$ with $a \neq 0^{|E|}$, and $a(V) \leq a_0$ we have

$$P_{\text{CW}} = \{y \in \mathbb{R}^{|E|+1} : y \text{ fulfills (26), (27), and (28)}\} =: Y \text{ if } a(V \setminus \{r\}) > a_r, \text{ and} \\ P_{\text{CW}} = \{y \in \mathbb{R}^{|E|+1} : y \text{ fulfills (26), (29), and (30)}\} =: Y^r \text{ if } a(V \setminus \{r\}) \leq a_r.$$

Proof. If $a(V \setminus \{r\}) > a_r$, Propositions 30 and 31 show that $Y \supseteq P_{\text{CW}}$, and to show $Y \subseteq P_{\text{CW}}$ we can use Propositions 35, 36, and 37. If $a(V \setminus \{r\}) \leq a_r$, Propositions 30 and 32 show that $Y^r \supseteq P_{\text{CW}}$, and to prove $Y^r \subseteq P_{\text{CW}}$ we can use again Propositions 35, 36, and 37. $\square$

Remark 39. Note that in all assertions of this section we have assumed $a(V) \leq a_0$. This assumption guarantees that every $S \subseteq V$ contributes its point h^S to P_{CW} . If we reduce a_0 below $a(V)$, the facial structure of P_{CW} becomes much more complicated, because suddenly the whole complexity of the knapsack polytope P_K comes into play. So far a complete description of P_{CW} with $a(V) > a_0$ seems out of reach for nontrivial graphs, even if we assume $a_v = 1$ for all $v \in V$.

Example 40. We continue Example 24. For the choice of the subgraphs $\bar{G}_l$ compare Figures 8 and 9.

- (2) The bisection knapsack walk inequality on $V' = \{1, 2, 3\}$ with root node $r = 3$ and $H_v = \emptyset$ for all $v \in V'$ is $1 + (1 - y_{13}) + (1 - y_{23}) \leq 4$. With $\bar{G}_1$ and $\bar{G}_2$ such that $\bar{V}_1 = \{4, 5\}$, $\bar{V}_2 = \{6, 7\}$, $\bar{E}_1 = \{45\}$, and $\bar{E}_2 = \{67\}$ the capacity reduced bisection knapsack walk inequality reads $1 + (1 - y_{13}) + (1 - y_{23}) \leq 4 - y_{45} - y_{67}$ and is a facet of P_B .

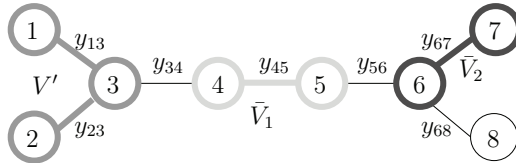


FIG. 8. Graph for Example 40 (2). $F = 4$, $\varphi_i = 1$ for all $i \in V$, $\sum_{i \in V} x_i \leq 4$.

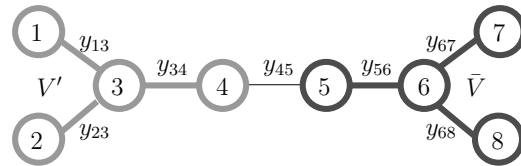


FIG. 9. Graph for Example 40 (3). $F = 4$, $\varphi_i = 1$ for all $i \in V$, $\sum_{i \in V} x_i \leq 4$.

- (3) For $V' = \{1, 2, 3, 4\}$, $r = 3$, and $H_v = \emptyset$ for all $v \in V'$ the bisection knapsack walk inequality is $1 + (1 - y_{13}) + (1 - y_{23}) + (1 - y_{34}) \leq 4$. Proposition 31 establishes that for $\bar{G}$ with $\bar{V} = \{5, 6, 7, 8\}$ and $\bar{E} = \{56, 67, 68\}$ one of the best minorizing functions for $\beta_{\bar{G}}$ is $y_{56} + y_{67} - y_{68}$. Thus the resulting capacity

reduced bisection knapsack walk inequality reads $1 + (1 - y_{13}) + (1 - y_{23}) + (1 - y_{34}) \leq 4 - y_{56} - y_{67} + y_{68}$. It is a facet of P_B .

7. Conclusions. We investigated the bisection cut polytope P_B associated with the minimum graph bisection problem (MB). In particular, we exploited the knapsack condition ($f(S) \leq F$ and $f(V \setminus S) \leq F$, $S \subseteq V$) in the formulation of the problem, which makes it NP-hard. As one would expect, inequalities basing on this knapsack constraint define high-dimensional faces of P_B . In the first part of the paper we showed that in case the underlying graph G is a tree the knapsack tree inequalities define facets of P_B . The situation becomes more complicated if one considers a denser graph. We suppose that also in this case there are facet-defining knapsack tree inequalities for P_B . However, so far we have not been able to identify sufficient conditions which must be fulfilled by the tree supporting the inequality. Here there is certainly room for further research. In the second part of the paper we worked out a version of knapsack related inequalities—the bisection knapsack walk inequalities—which exploit the special bisection case. We took a closer look at their strengthening resulting in capacity reduced bisection knapsack walk inequalities. The right-hand sides of these inequalities may be reduced by exploiting weights of nodes that are not endpoints of walks within the respective inequality. The best possible reduction of the right-hand side is achieved by applying facets of the newly introduced cluster weight polytope P_{CW} . As one would expect, the facial structure of P_{CW} is not trivial due to its relation to the knapsack polytope. We gave the full description of P_{CW} for the case that the complement of the walk in G is a star, all nodes of which fit into the knapsack. Even though this simple case was already challenging we encourage the investigation of P_{CW} on more complex graphs than a star and on graphs with an active capacity restriction on the node weight. The practical value of the strengthenings will be investigated in a follow-up paper, which also features extensive results on a comparison of semidefinite versus pure linear branch and cut approaches for bisection problems on large sparse graphs.

REFERENCES

- [1] M. ARMBRUSTER, *Branch-and-Cut for a Semidefinite Relaxation of Large-scale Minimum Bisection Problems*, Ph.D. thesis, Technische Universität Chemnitz, Chemnitz, Germany, 2007.
- [2] F. BARAHONA AND A. R. MAHJOUR, *On the cut polytope*, Math. Programming, 36 (1986), pp. 157–173.
- [3] M. CONFORTI, M. R. RAO, AND A. SASSANO, *The equipartition polytope I*, Math. Programming, 49 (1990), pp. 49–70.
- [4] M. CONFORTI, M. R. RAO, AND A. SASSANO, *The equipartition polytope II*, Math. Programming, 49 (1990), pp. 71–90.
- [5] C. C. DE SOUZA, *The Graph Equipartition Problem: Optimal Solutions, Extensions and Applications*, Ph.D. thesis, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, 1993.
- [6] M. DEZA AND M. LAURENT, *Geometry of Cuts and Metrics*, Algorithms Combin. 15, Springer, Berlin, 1997.
- [7] A. EISENBLÄTTER, *Frequency Assignment in GSM Networks*, Ph.D. thesis, Technische Universität Berlin, Berlin, 2001.
- [8] C. E. FERREIRA, A. MARTIN, C. C. DE SOUZA, R. WEISMANTEL, AND L. A. WOLSEY, *Formulations and valid inequalities for the node capacitated graph partitioning problem*, Math. Programming, 74 (1996), pp. 247–267.
- [9] M. FÜGENSCHUH, *Relaxations and Solutions for the Minimum Graph Bisection Problem*, Ph.D. thesis, Darmstadt University of Technology, Darmstadt, Germany, 2007.
- [10] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability*, W. H. Freeman, New York, 1979.

- [11] E. GAWRILOW AND M. JOSWIG, *polymake: A framework for analyzing convex polytopes*, in Polytopes—Combinatorics and Computation, G. Kalai and G. M. Ziegler, eds., Birkhäuser, Basel, 2000, pp. 43–74.
- [12] J. R. GILBERT AND R. E. TARJAN, *The analysis of a nested dissection algorithm*, Numer. Math., 50 (1987), pp. 377–404.
- [13] E. JOHNSON, A. MEHROTRA, AND G. L. NEMHAUSER, *Min-cut clustering*, Math. Programming, 62 (1993), pp. 133–152.
- [14] M. LAURENT AND C. C. DE SOUZA, *Some new classes of facets for the equicut polytope*, Discrete Appl. Math., 62 (1995), pp. 167–191.
- [15] T. LENGAUER, *Combinatorial Algorithms for Integrated Circuit Layout*, John Wiley, Chichester, UK, 1990.
- [16] S. MARTELLO AND P. TOTH, *Knapsack Problems*, John Wiley, Chichester, UK, 1990.
- [17] R. WEISMANTEL, *On the 0/1 knapsack polytope*, Math. Programming, 77 (1997), pp. 49–68.

LABELED PARTITIONS AND THE q -DERANGEMENT NUMBERS*

WILLIAM Y. C. CHEN[†] AND DEHENG XU[†]

Abstract. Inspired by MacMahon's original proof of his celebrated theorem on the distribution of the major index over permutations, we give a reformulation of his argument in terms of labeled partitions. In this framework, we establish a decomposition theorem for labeled partitions which is analogous to the decomposition of a permutation into derangement points and fixed points. This decomposition implies a reformulation of Wachs's formula concerning the derangement parts and the major index on permutations which was derived in order to present a bijective treatment of Gessel's formula on the q -derangement numbers.

Key words. q -derangement number, major index, bijection, partitions, labeled partitions

AMS subject classifications. 05A30, 05A19, 05A15

DOI. 10.1137/06066326X

1. Introduction. We will follow the terminology and notation on permutations and partitions and q -series in Andrews [2] and Stanley [10]. The set of permutations on $\{1, 2, \dots, n\}$ is denoted by S_n . For any permutation $\pi = \pi_1\pi_2\cdots\pi_n \in S_n$, an index i with $1 \leq i \leq n-1$ is called a *descent* of π if $\pi_i > \pi_{i+1}$. The major index $\text{maj}(\pi)$ of π , introduced by MacMahon [9], is defined as the sum of all descents of π . The following formula is well known:

$$(1.1) \quad \sum_{\pi \in S_n} q^{\text{maj}(\pi)} = [n]! = 1 \cdot (1+q) \cdot (1+q+q^2) \cdots (1+q+\cdots+q^{n-1}).$$

The underlying idea of MacMahon's proof goes as follows. It is easier to consider sequences and partitions than solely permutations for the purpose of studying the major index. MacMahon established (1.1) by proving an equivalent formula,

$$(1.2) \quad \frac{1}{(q)_n} \sum_{\pi \in S_n} q^{\text{maj}(\pi)} = \frac{1}{(1-q)^n},$$

where $(q)_n = (1-q) \cdots (1-q^n)$, and $(q)_n^{-1}$ is the generating function for partitions with n parts, where zero parts are allowed. We will give a reformulation of MacMahon's proof in section 2 by using the notion of standard labeled partitions.

The main objective of this paper is to employ MacMahon's method to deal with the major index of derangements. An integer i with $1 \leq i \leq n$ is said to be a *fixed point* of $\pi \in S_n$ if $\pi_i = i$, and a *derangement point* otherwise. *Derangements* are permutations with no fixed points. Let D_n be the set of all derangements in S_n . The q -derangement numbers are defined by $d_0(q) = 1$ and for $n \geq 1$

$$d_n(q) = \sum_{\pi \in D_n} q^{\text{maj}(\pi)}.$$

*Received by the editors June 19, 2006; accepted for publication (in revised form) March 19, 2008; published electronically July 2, 2008. This work was supported by the 973 Project, the PCSIRT Project of the Ministry of Education, the Ministry of Science and Technology, and the National Science Foundation of China.

<http://www.siam.org/journals/sidma/22-3/66326.html>

[†]Center for Combinatorics, LPMC-TJKLC, Nankai University, Tianjin 300071, People's Republic of China (chen@nankai.edu.cn, xudeheng@eyou.com).

The following elegant formula was first derived by Gessel and published in [6] as a consequence of the quasi-symmetric generating function encoding the descents and the cycle structure of permutations:

$$(1.3) \quad d_n(q) = [n]! \sum_{k=0}^n \frac{(-1)^k}{[k]!} q^{\binom{k}{2}}.$$

A combinatorial proof of (1.3) has been obtained by Wachs [12]. Let us review the combinatorial settings of Wachs. Suppose that the derangement points of π are $p_1, p_2, \dots, p_k$. The reduction of π to its *derangement part*, denoted by $dp(\pi)$, is defined as a permutation on $\{1, 2, \dots, k\}$ induced by the relative order of $\pi_{p_1}, \pi_{p_2}, \dots, \pi_{p_k}$. For example, the derangement points of $\pi = (1, 5, 3, 7, 6, 2, 9, 8, 4)$ are 2, 4, 5, 6, 7, 9, and $\pi_2\pi_4\pi_5\pi_6\pi_7\pi_9 = (5, 7, 6, 2, 9, 4)$. Then $dp(\pi) = (3, 5, 4, 1, 6, 2)$. Clearly $dp(\pi) \in D_k$ if π has k derangement points. On the other hand, we can insert a fixed point j with $1 \leq j \leq k+1$ into $\pi \in S_k$ to obtain a permutation

$$(1.4) \quad \bar{\pi} = \pi'_1 \pi'_2 \cdots \pi'_{j-1} j \pi'_j \cdots \pi'_k \in S_{k+1},$$

where $\pi'_i = \pi_i$ if $\pi_i < j$ and $\pi'_i = \pi_i + 1$ if $\pi_i \geq j$. Such an insertion operation produces one extra fixed point and will be used in the proof of Theorem 2.2. Wachs [12] has established the following relation.

THEOREM 1.1. *Let $0 \leq k \leq n$ and $\sigma \in D_k$. Then we have*

$$(1.5) \quad \sum_{\substack{dp(\pi)=\sigma \\ \pi \in S_n}} q^{\text{maj}(\pi)} = q^{\text{maj}(\sigma)} \begin{bmatrix} n \\ k \end{bmatrix},$$

where $\begin{bmatrix} n \\ k \end{bmatrix} = \frac{[n]!}{[k]![n-k]!}$ is the q -binomial coefficient.

Summing over all derangements $\sigma \in D_k$ and $0 \leq k \leq n$, and applying (1.1), we can deduce from (1.5) that

$$[n]! = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} d_k(q).$$

Thus (1.3) follows from the q -binomial inversion [1, Corollary 3.38].

In order to justify the relation (1.5), Wachs found a bijection on S_n by rearranging a permutation π according to *excedant* (where $\pi_i > i$), fixed point, and *subcedant* (where $\pi_i < i$). She showed that this bijection preserves the major index. Then a result of Garsia and Gessel [4, Theorem 3.1] on shuffles of permutations is applied to establish Theorem 1.1.

Inspired by MacMahon's original proof of (1.1), we present an alternative approach to Wachs's formula (1.5) based on the following reformulation:

$$(1.6) \quad \frac{1}{(q)_n} \sum_{\substack{dp(\pi)=\sigma \\ \pi \in S_n}} q^{\text{maj}(\pi)} = \frac{1}{(q)_k(q)_{n-k}} q^{\text{maj}(\sigma)}.$$

We will use the terminology of labeled partitions and will introduce the notion of standard labeled partitions. In this framework, MacMahon's proof can be easily stated. Moreover, a combinatorial justification of (1.6) reduces to a decomposition theorem which is analogous to the decomposition of a permutation by separating the derangement points and the fixed points.

2. Labeled partitions. Let $\lambda = (\lambda_1, \dots, \lambda_n)$ be a partition, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$. We say that λ is a partition with n parts, or of length n , and we write $|\lambda| = \lambda_1 + \dots + \lambda_n$. A *labeled partition* of length n is defined as a pair (λ, π) (or $\binom{\lambda}{\pi}$), where λ is a partition with n parts and π is a permutation $\pi = \pi_1\pi_2 \cdots \pi_n$ on $[n] = \{1, 2, \dots, n\}$. A labeled partition is also represented in the following two-row form as in Andrews [2, p. 43]:

$$\begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \pi_1 & \pi_2 & \cdots & \pi_n \end{pmatrix}.$$

A labeled partition (λ, π) is said to be *standard* if $\pi_i > \pi_{i+1}$ implies $\lambda_i > \lambda_{i+1}$. For example, the labeled partition in (2.1) is standard.

The following Lemma 2.1 is straightforward to verify, which is MacMahon's original approach for studying the major index with the aid of partitions; see MacMahon [9], Andrews [2, Theorem 3.7], Knuth [8, p. 18] or [7]. This method was further extended by Stanley [11]. For other applications, see [4].

LEMMA 2.1. *Given $\pi \in S_n$, there is a bijection $\psi_\pi: \lambda \mapsto \mu$ between the set of partitions λ with n parts and the set of partitions μ with n parts such that (μ, π) is a standard labeled partition and $|\lambda| + \text{maj}(\pi) = |\mu|$. Formally, we write $\psi(\lambda, \pi) = (\psi_\pi(\lambda), \pi)$.*

The bijection ψ_π (or simply ψ when π is understood from the context) is given as follows:

$$\mu = \psi_\pi(\lambda) = (\lambda_1 + \phi_1, \lambda_2 + \phi_2, \dots, \lambda_n + \phi_n),$$

where ϕ_i is the number of descents in $\pi_i\pi_{i+1} \cdots \pi_n$. One may also view ψ as the operation of adding 1 to $\lambda_1, \dots, \lambda_i$ whenever i is a descent of π . The inverse map ψ^{-1} can be easily described.

Below is a restatement of MacMahon's proof of (1.2) in the terminology of labeled partitions.

Proof of (1.2). Given a sequence $a_1a_2 \cdots a_n$ of nonnegative integers, we associate it with weight $q^{a_1+a_2+\cdots+a_n}$. Let us construct a two-row array:

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

Observe that a labeled partition (λ, π) is standard if $\lambda_i = \lambda_{i+1}$ implies $\pi_i < \pi_{i+1}$. Therefore, by permuting the columns of the above array, we get a unique standard labeled partition (μ, π) with $|\mu| = a_1 + a_2 + \cdots + a_n$. Applying Lemma 2.1, we obtain a partition λ with $|\lambda| + \text{maj}(\pi) = |\mu|$. Clearly, the above steps are reversible. This completes the proof. $\square$

Let $n = 9$ and $a_1a_2 \dots a_9$ be the first row in the following array:

$$\begin{pmatrix} 3 & 6 & 8 & 3 & 1 & 3 & 6 & 4 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{pmatrix}.$$

By permuting the columns we get the standard labeled partition,

$$(2.1) \quad \begin{pmatrix} \mu \\ \pi \end{pmatrix} = \begin{pmatrix} 8 & 8 & 6 & 6 & 4 & 3 & 3 & 3 & 1 \\ 3 & \underline{9} & 2 & 7 & \underline{8} & 1 & 4 & \underline{6} & 5 \end{pmatrix},$$

where we have underlined those elements π_i whenever i is a descent of π .

Applying the inverse map ψ^{-1} we obtain

$$\begin{pmatrix} \lambda \\ \pi \end{pmatrix} = \begin{pmatrix} 5 & 5 & 4 & 4 & 2 & 2 & 2 & 2 & 1 \\ 3 & \underline{9} & 2 & 7 & \underline{8} & 1 & 4 & \underline{6} & 5 \end{pmatrix},$$

which is the corresponding labeled partition.

We remark that the idea of standard labeled partitions appeared in [4, p. 292], though it was not used to prove (1.2).

We are now ready to present a decomposition theorem on standard labeled partitions with respect to the fixed points. Let σ be a given derangement in D_k and π a permutation in S_n such that $dp(\pi) = \sigma$. Assume that (μ, π) is a standard labeled partition of length n . Let $i_1 < i_2 < \cdots < i_{n-k}$ be the fixed points and $j_1 < j_2 < \cdots < j_k$ the derangement points of π . We now define the following decomposition of a standard labeled partition:

$$(2.2) \quad \Delta: (\mu, \pi) \mapsto (\beta, \gamma),$$

where $\beta = \mu_{j_1} \mu_{j_2} \cdots \mu_{j_k}$ and $\gamma = \mu_{i_1} \mu_{i_2} \cdots \mu_{i_{n-k}}$ are the partitions corresponding to the derangement points and fixed points, respectively. Evidently, μ consists of the parts from β and γ , or in common notation, $\mu = \beta \cup \gamma$.

THEOREM 2.2. *Let $\sigma \in D_k$. There is a bijection between the set of standard labeled partitions (μ, π) of length n with $dp(\pi) = \sigma$ and the set of pairs of partitions (β, γ) such that β is a partition with k parts, γ is a partition with $n-k$ parts, $\mu = \beta \cup \gamma$, and (β, σ) is a standard labeled partition.*

For example, let

$$\begin{pmatrix} \lambda \\ \pi \end{pmatrix} = \begin{pmatrix} 5 & 4 & 4 & 4 & 4 & 3 & 2 \\ \underline{5} & \underline{2} & 1 & 4 & \underline{7} & 3 & 6 \end{pmatrix}.$$

Applying ψ , we get

$$\begin{pmatrix} \mu \\ \pi \end{pmatrix} = \begin{pmatrix} 8 & \mathbf{6} & \mathbf{5} & \mathbf{5} & 5 & 3 & 2 \\ 5 & \mathbf{2} & \mathbf{1} & \mathbf{4} & 7 & 3 & 6 \end{pmatrix}.$$

The fixed points of π are 2, 4, as signified in boldface. Hence $\sigma = dp(\pi) = (3 \ 1 \ 5 \ 2 \ 4)$. Applying Δ on (μ, π) gives $(\beta, \gamma) = ((8 \ 5 \ 5 \ 3 \ 2), (6 \ 5))$. Finally, applying ψ^{-1} to

$$\begin{pmatrix} \beta \\ \sigma \end{pmatrix} = \begin{pmatrix} 8 & 5 & 5 & 3 & 2 \\ 3 & 1 & 5 & 2 & 4 \end{pmatrix},$$

we obtain

$$\begin{pmatrix} \alpha \\ \sigma \end{pmatrix} = \begin{pmatrix} 6 & 4 & 4 & 3 & 2 \\ 3 & 1 & 5 & 2 & 4 \end{pmatrix}.$$

We note that the above theorem and Lemma 2.1 lead to a combinatorial interpretation of the relation (1.6). Given a labeled partition (λ, π) , we first use the map ψ to transform it into a standard labeled partition (μ, π) . Let $\sigma = dp(\pi)$. Using the above decomposition for (μ, π) , we obtain a pair of partitions (β, γ) such that (β, σ) is a standard labeled partition and γ is a partition with $n-k$ parts. Moreover, we can find a partition α with k parts such that $\psi_\sigma(\alpha, \sigma) = (\beta, \sigma)$. Thus we are led to the relation

$$(2.3) \quad |\lambda| + \text{maj}(\pi) = |\alpha| + |\gamma| + \text{maj}(\sigma),$$

which yields (1.6).

Proof of Theorem 2.2. We first verify that (β, σ) is standard. Suppose there exists i such that $\sigma_i > \sigma_{i+1}$. Then we need to show that $\beta_i > \beta_{i+1}$. Clearly, σ_i and σ_{i+1} correspond to two elements π_j and π_k ($j < k$) in π such that $\pi_j > \pi_k$, and the points $j+1, j+2, \dots, k-1$, if any, are fixed points of π . In other words, it is necessary to show that $\mu_j > \mu_k$ since $\beta_i = \mu_j$ and $\beta_{i+1} = \mu_k$ by the decomposition. If $j = k-1$, since (μ, π) is standard, we have $\mu_j > \mu_k$. For the case $j < k-1$, we see that either $\pi_j > \pi_{j+1} = j+1$ or $\pi_{k-1} = k-1 > \pi_k$. Otherwise, we have $\pi_j < j+1 \leq k-1 < \pi_k$, a contradiction. Therefore, we have either $\mu_j > \mu_{j+1}$ or $\mu_{k-1} > \mu_k$. This implies that $\mu_j > \mu_k$. So we conclude that (β, σ) is a standard labeled partition.

We now proceed to construct the inverse map φ which corresponds to the procedure to recover π by inserting the fixed points into the derangement σ on $\{1, 2, \dots, k\}$. It turns out that the order of insertions reflects the property of standard labeled partitions.

We begin with $(\mu^0, \pi^0) = (\beta, \sigma)$ and assume that (μ^i, π^i) is obtained from (μ^{i-1}, π^{i-1}) by *inserting* γ_i , where γ_i is the i th part of γ . Let r be the first position such that the insertion of γ_i produces a partition. In other words, if μ^{i-1} already contains some parts equal to γ_i , then we insert γ_i as the first occurrence. This partition is denoted by μ^i . We need to determine the corresponding fixed point of π^i caused by the insertion of γ_i .

Clearly, $\mu_{r-1}^i > \mu_r^i = \gamma_i$. We may assume that $\mu_r^i = \dots = \mu_t^i > \mu_{t+1}^i$ for some $t \geq r$. As in the proof of (1.2), in order to get a standard labeled partition after each insertion of γ_i , we should insert s such that the subsequence

$$\pi_r^{i-1'}, \dots, \pi_{s-1}^{i-1'}, s, \pi_s^{i-1'}, \dots, \pi_{t-1}^{i-1'}$$

of the permutation obtained by inserting s into π^{i-1} is increasing. We see that the position s is uniquely determined. If $r = t$, then we set $s = r$. Otherwise we find the first position s ($r \leq s \leq t$) such that $\pi_{s-1}^{i-1} < s \leq \pi_s^{i-1}$. Strictly speaking, we have adopted the convention that $\pi_{r-1}^{i-1} = -\infty$ and $\pi_t^{i-1} = \infty$. Consequently, we can insert s into π^{i-1} as a fixed point of π^i .

It remains to show that $\pi^{n-k} = \pi$. For notational simplicity, we write π^{n-k} as $\bar{\pi}$. By removing the common fixed points, we may assume that the first fixed point f of π is different from the first fixed point $\bar{f}$ of $\bar{\pi}$. We find that f satisfies the condition $\pi_{f-1} < f \leq \pi_{f+1} - 1$. Furthermore, by the insertion procedure, $\bar{f}$ is the first position we like to locate, and hence we may assume that $\bar{f} < f$. Clearly, $\mu_f = \mu_{\bar{f}}$. Since (μ, π) and $(\mu, \bar{\pi})$ are standard labeled partitions, we have

$$\pi_{\bar{f}} < \pi_{\bar{f}+1} < \dots < \pi_f, \quad \bar{\pi}_{\bar{f}} < \bar{\pi}_{\bar{f}+1} < \dots < \bar{\pi}_f.$$

Now we see that $\pi_f = f$ and $\bar{\pi}_{\bar{f}} = \bar{f}$. But $\pi_{\bar{f}} < \pi_{\bar{f}+1} < \dots < \pi_f$ and $\pi_f = f$, so we can deduce that $\pi_{\bar{f}} \leq \bar{f}$. Since f is the first fixed point of π , we obtain $\sigma_{\bar{f}} = \pi_{\bar{f}} < \bar{f}$. From the construction of $\bar{\pi}$, it follows that $\bar{\pi}_{\bar{f}} \leq \sigma_{\bar{f}} < \bar{f}$, which is impossible given that $\bar{\pi}_{\bar{f}} = \bar{f}$. This completes the proof. $\square$

For example, given $\sigma = (3 \ 1 \ 5 \ 2 \ 4)$, $\beta = (8, 5, 5, 3, 2)$, and $\gamma = (6, 5)$, we may recover (μ, π) through the following steps:

$$\begin{pmatrix} \beta \\ \sigma \end{pmatrix} = \begin{pmatrix} 8 & 5 & 5 & 3 & 2 \\ 3 & 1 & 5 & 2 & 4 \end{pmatrix} \xrightarrow{\gamma_1=6} \begin{pmatrix} 8 & \underline{6} & 5 & 5 & 3 & 2 \\ 4 & \underline{2} & 1 & 6 & 3 & 5 \end{pmatrix} \xrightarrow{\gamma_2=5} \begin{pmatrix} 8 & 6 & 5 & \underline{5} & 5 & 3 & 2 \\ 5 & 2 & 1 & \underline{4} & 7 & 3 & 6 \end{pmatrix} = \begin{pmatrix} \mu \\ \pi \end{pmatrix}.$$

Acknowledgments. The authors would like to thank Henry Y. Gao and Guoce Xin for their valuable suggestions. They are also grateful to the referee for helpful comments leading to an improvement of an earlier version.

REFERENCES

- [1] M. AIGNER, *Combinatorial Theory*, Springer-Verlag, New York, 1979.
- [2] G. E. ANDREWS, *The Theory of Partitions*, Encyclopedia Math. Appl. 2, Addison-Wesley, Reading, MA, 1976.
- [3] W. Y. C. CHEN AND G.-C. ROTA, *q-Analogs of the inclusion-exclusion principle and permutations with restricted position*, Discrete Math., 104 (1992), pp. 7–22.
- [4] A. M. GARSIA AND I. GESSEL, *Permutation statistics and partitions*, Adv. Math., 31 (1979), pp. 288–305.
- [5] A. M. GARSIA AND J. REMMEL, *A combinatorial interpretation of q-derangement and q-Laguerre numbers*, European J. Combin., 1 (1980), pp. 47–59.
- [6] I. GESSEL AND C. REUTENAUER, *Counting permutations with given cycle structure and descent set*, J. Combin. Theory Ser. A, 64 (1993), pp. 189–215.
- [7] D. E. KNUTH, *A note on solid partitions*, Math. Comp., 24 (1970), pp. 955–962.
- [8] D. E. KNUTH, *The Art of Computer Programming, Vol. 3. Sorting and Searching*, Addison-Wesley Series in Computer Science and Information Processing, Addison-Wesley, Reading, MA, 1973.
- [9] P. A. MACMAHON, *Combinatory Analysis*, Vol. 1, Cambridge University Press, London, New York, 1915 (reprinted by Chelsea, New York, 1960).
- [10] R. P. STANLEY, *Enumerative Combinatorics*, Vol. 1, Cambridge University Press, Cambridge, UK, 1997.
- [11] R. P. STANLEY, *Ordered Structures and Partitions*, Mem. Amer. Math. Soc. 119, AMS, Providence, RI, 1972.
- [12] M. L. WACHS, *On q-derangement numbers*, Proc. Amer. Math. Soc., 106 (1989), pp. 273–278.

EULER–MAHONIAN STATISTICS ON ORDERED SET PARTITIONS*

MASAO ISHIKAWA[†], ANISSE KASRAOUI[‡], AND JIANG ZENG[‡]

Abstract. An ordered partition with k blocks of $[n] := \{1, 2, \dots, n\}$ is a sequence of k disjoint and nonempty subsets, called blocks, whose union is $[n]$. Clearly the number of such ordered partitions is $k!S(n, k)$, where $S(n, k)$ is the Stirling number of the second kind. A statistic on ordered partitions of $[n]$ with k blocks is called an *Euler–Mahonian statistic* if its generating polynomial is $[k]_q!S_q(n, k)$, which is a natural q -analogue of $k!S(n, k)$. Motivated by Steingrímsson’s conjectures dating back to 1997, we consider two different methods to produce Euler–Mahonian statistics on ordered set partitions: (a) we give a bijection between ordered partitions and weighted Motzkin paths by using a variant of Françon–Viennot’s bijection to derive many Euler–Mahonian statistics by expanding the generating function of $[k]_q!S_q(n, k)$ as an explicit continued fraction; (b) we encode ordered partitions by walks in some digraphs and then derive new Euler–Mahonian statistics by computing their generating functions using the transfer-matrix method. In particular, we prove several conjectures of Steingrímsson.

Key words. ordered partitions, Euler–Mahonian statistics, q -Stirling numbers of the second kind, transfer-matrix method

AMS subject classifications. Primary, 05A18; Secondary, 05A15, 05A30

DOI. 10.1137/060672340

1. Introduction. The systematic study of statistics on permutations and words has its origins in the work of MacMahon, at the turn of the last century. MacMahon [11] considered four different statistics for a permutation π : The number of descents ($\text{des } \pi$), the number of excedances ($\text{exc } \pi$), the number of inversions ($\text{inv } \pi$), and the major index ($\text{maj } \pi$). These are defined as follows: A descent in a permutation $\pi = a_1 a_2 \cdots a_n$ is an i such that $a_i > a_{i+1}$, an excedance is an i such that $a_i > i$, an inversion is a pair (i, j) such that $i < j$ and $a_i > a_j$, and the major index of π is the sum of the descents in π .

In fact, MacMahon studied these statistics in greater generality, namely, over the rearrangement class of an arbitrary word w with repeated letters. The rearrangement class $R(w)$ of a word $w = a_1 a_2 \cdots a_n$ is the set of all words obtained by permutating the letters of w . All of the statistics mentioned above generalize to words, and in each case, except for that of exc , the generalization is trivial.

MacMahon showed, algebraically, that exc is *equidistributed* with des , and that inv is equidistributed with maj , over $R(w)$ for any word w . That is to say,

$$\sum_{z \in R(w)} t^{\text{exc } z} = \sum_{z \in R(w)} t^{\text{des } z} \quad \text{and} \quad \sum_{z \in R(w)} t^{\text{inv } z} = \sum_{z \in R(w)} t^{\text{maj } z}.$$

Any statistic that is equidistributed with des is said to be *Eulerian*, while any statistic equidistributed with inv is said to be *Mahonian*. Foata [5] coined the name *Euler–Mahonian* statistic for a bivariate statistic $(\text{eul } \pi, \text{mah } \pi)$, where eul is Eulerian and mah is Mahonian, and carried out the first study of such pairs.

*Received by the editors October 14, 2006; accepted for publication (in revised form) March 19, 2008; published electronically July 2, 2008.

<http://www.siam.org/journals/sidma/22-3/67234.html>

[†]Faculty of Education, Tottori University, Koyama, Tottori, Japan (ishikawa@fed.tottori-u.ac.jp).

[‡]Institut Camille Jordan, Université Claude Bernard (Lyon I), F-69622, Villeurbanne Cedex, France (anisse@math.univ-lyon1.fr, zeng@math.univ-lyon1.fr).

In 1997 Steingrímsson [18] introduced the notion of *Euler–Mahonian* statistic on ordered partitions. In order to motivate this definition we recall a well-known basic identity relating Eulerian and Stirling numbers along with some basic definitions.

For a permutation $\sigma = \sigma(1)\sigma(2)\dots\sigma(n)$ of $[n] = \{1, 2, \dots, n\}$, the integer $i \in [n-1]$ is called a *descent* (resp., *ascent*) of σ if $\sigma(i) > \sigma(i+1)$ (resp., $\sigma(i) < \sigma(i+1)$). Let $\text{des } \sigma$ be the number of descents of σ and S_n be the set of permutations of $[n]$; then the *Eulerian polynomial* $A_n(x)$ is the generating polynomial of permutations in S_n with respect to the number of descents, i.e.,

$$A_n(x) = \sum_{\sigma \in S_n} x^{1+\text{des } \sigma} = \sum_{k=0}^{n-1} A(n, k) x^{1+k},$$

where the coefficients $A(n, k)$ are called the *Eulerian numbers*. The Eulerian polynomials can also be defined by the following exponential generating function:

$$(1.1) \quad 1 + \sum_{n \geq 1} A_n(x) \frac{z^n}{n!} = \frac{1-x}{1-xe^{(1-x)z}}.$$

A partition $P = B_1/B_2/\dots/B_k$ of $[n]$ is a collection of disjoint and nonempty subsets B_i 's, called *blocks*, whose union is $[n]$, where we write P in the standard way; i.e., the blocks B_i are arranged in increasing order of their minimal elements and separated by $/$. Let $\mathcal{P}_n^k$ be the set of partitions of $[n]$ with k blocks. Now, if σ is a permutation in S_k , the sequence $\pi = B_{\sigma(1)}/B_{\sigma(2)}/\dots/B_{\sigma(k)}$ is called an *ordered partition* of $[n]$ with k blocks. In other words, an ordered partition π of $[n]$ with k blocks can be viewed as a pair (P, σ) , where P is a partition in $\mathcal{P}_n^k$ and σ a permutation in S_k . We shall denote the permutation σ by $\text{perm}(\pi)$. Let $\mathcal{OP}_n^k$ be the set of ordered partitions of $[n]$ into k blocks.

A *preferential arrangement* of $[n]$ is a permutation of $[n]$, of which each ascent is *underlined or not*. A *run* in a permutation σ is an increasing sequence of consecutive terms $\sigma(i) < \sigma(i+1) < \dots < \sigma(j)$ such that $\sigma(i-1) > \sigma(i)$ and $\sigma(j) > \sigma(j+1)$, where $\sigma(0) = n+1$ and $\sigma(n+1) = 0$. Clearly each preferential arrangement can be decomposed uniquely, from left to right, into *runs* or *increasing sequences ending with an underlined ascent*. Hence a preferential arrangement of $[n]$ can be identified with an *ordered set partition* of $[n]$. For example, we have the following correspondence between a preferential arrangement σ and an ordered partition π of [9]:

$$\sigma = \underline{2}931485\underline{6}7 \longleftrightarrow \pi = 2/9/3/148/5\underline{6}/7.$$

Clearly, if k is the number of blocks in the ordered set partition π , then $k-1$ equals the number of descents and underlined ascents of σ . Since the number of ordered partitions of $[n]$ into k blocks is $k!S(n, k)$, where $S(n, k)$ denotes the *Stirling numbers of the second kind*, this correspondence implies immediately the identity:

$$(1.2) \quad k!S(n, k) = \sum_{m=1}^k \binom{n-m}{n-k} A(n, m-1).$$

Indeed, we can construct the corresponding preferential arrangements of $[n]$ by first choosing a permutation of $[n]$ with $m-1$ descents and then underline $k-1-(m-1) = k-m$ ascents.

Let $\mathcal{OP}_n$ be the set of all ordered partitions of $[n]$, and let $\underline{S}_n$ be the set of all preferential arrangements of $[n]$. Clearly the cardinality of $\mathcal{OP}_n$ is given by $|\mathcal{OP}_n| =$

$\sum_{k=0}^n k! S(n, k)$ or $|\mathcal{OP}_n| = \sum_{\sigma \in S_n} 2^{\text{des } \sigma}$, by the correspondence between $\mathcal{OP}_n$ and S_n . We then derive from (1.1) the following exponential generating function:

$$1 + \sum_{n \geq 1} |\mathcal{OP}_n| \frac{z^n}{n!} = \frac{1}{2 - e^z} = 1 + z + 3 \frac{z^2}{2!} + 13 \frac{z^3}{3!} + 75 \frac{z^4}{4!} + \cdots$$

Carlitz [1, 2] was the first to study q -analogues of *Eulerian numbers* and *Stirling numbers of the second kind*, which we recall in the following.

Define the p, q -integer $[n]_{p,q} = \frac{p^n - q^n}{p - q}$, the p, q -factorial $[n]_{p,q}! = [1]_{p,q} [2]_{p,q} \cdots [n]_{p,q}$, and the p, q -binomial coefficient

$$\begin{bmatrix} n \\ k \end{bmatrix}_{p,q} = \begin{cases} \frac{[n]_{p,q}!}{[k]_{p,q}! [n-k]_{p,q}!} & \text{if } 0 \leq k \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

If $p = 1$, we shall write $[n]_q$, $[n]_{q!}$, and $\begin{bmatrix} n \\ k \end{bmatrix}_q$ for $[n]_{1,q}$, $[n]_{1,q!}$, and $\begin{bmatrix} n \\ k \end{bmatrix}_{1,q}$, and usually called q -integer, q -factorial, and q -binomial coefficient, respectively.

The q -Eulerian numbers $A_q(n, k)$ ($n \geq k \geq 0$) are defined by

$$A_q(n, k) = q^k [n - k]_q A_q(n - 1, k - 1) + [k + 1]_q A_q(n - 1, k).$$

The first values of the q -Eulerian numbers $\langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle_q$ ($n \geq k \geq 0$) read

$n \setminus k$	0	1	2	3
1	1			
2	1	q		
3	1	$2q + 2q^2$	q^3	
4	1	$3q + 5q^2 + 3q^3$	$3q^3 + 5q^4 + 3q^5$	q^6

The *major index* of σ , noted $\text{maj } \sigma$, is the sum of its descents, i.e., $\text{maj } \sigma = \sum_i i$, where the summation is over all descents i of σ . It is well known that maj is a Mahonian statistic on S_n , i.e.,

$$\sum_{\sigma \in S_n} q^{\text{maj } \sigma} = [n]_{q!}.$$

Carlitz [2] gave the following combinatorial interpretation of q -Eulerian numbers:

$$A_q(n, k) = \sum_{\sigma} q^{\text{maj } \sigma},$$

where the summation is over all permutations of $[n]$ with k descents.

The two natural q -Stirling numbers $S_q(n, k)$ and $\tilde{S}_q(n, k)$ of the second kind are defined by

$$(1.3) \quad S_q(n, k) = q^{k-1} S_q(n - 1, k - 1) + [k]_q S_q(n - 1, k) \quad (n \geq k \geq 0)$$

and

$$(1.4) \quad \tilde{S}_q(n, k) = \tilde{S}_q(n - 1, k - 1) + [k]_q \tilde{S}_q(n - 1, k) \quad (n \geq k \geq 0),$$

where $S_q(n, k) = \tilde{S}_q(n, k) = \delta_{nk}$ if $n = 0$ or $k = 0$, and δ_{nk} is the Kronecker delta.

It is easy to see that $S_q(n, k) = q^{\binom{k}{2}} \widetilde{S}_q(n, k)$. The first values of the q -Stirling numbers $S_q(n, k)$ read

$n \setminus k$	1	2	3	4
1	1			
2	1	q		
3	1	$2q + q^2$	q^3	
4	1	$3q + 3q^2 + q^3$	$3q^3 + 2q^4 + q^5$	q^6

There has been a considerable amount of recent interest in properties and combinatorial interpretations of the q -Eulerian numbers and q -Stirling numbers and related numbers (see, e.g., [1, 2, 3, 10, 12, 13, 14, 15, 16, 18, 20, 21, 22]).

Garsia [8] found a q -analogue of the Frobenius formula relating q -Eulerian numbers and q -Stirling numbers of the second kind:

$$(1.5) \quad \sum_{k=1}^n \frac{[k]! S_q(n, k) x^k}{(x; q)_{k+1}} = \sum_{k=1}^{\infty} [k]^n x^k = \frac{\sum_{\sigma \in S_n} x^{1+\text{des } \sigma} q^{\text{maj } \sigma}}{(x; q)_{n+1}},$$

where $(x; q)_n = (1-x)(1-xq) \cdots (1-xq^{n-1})$. From (1.5) Zeng and Zhang [24] then derived the following q -analogue of (1.2):

$$(1.6) \quad [k]_q! S_q(n, k) = \sum_{m=1}^k q^{k(k-m)} \begin{bmatrix} n-m \\ n-k \end{bmatrix}_q A_q(n, m-1).$$

In his attempt to give a combinatorial proof of (1.6), Steingrímsson [18] introduced the following definition.

DEFINITION 1.1. *A statistic STAT on ordered set partitions is called Euler–Mahonian if for any $n \geq k \geq 1$ its generating function over $\mathcal{OP}_n^k$ is equal to $[k]_q! S_q(n, k)$, i.e.,*

$$\sum_{\pi \in \mathcal{OP}_n^k} q^{\text{STAT } \pi} = [k]_q! S_q(n, k).$$

An Euler–Mahonian statistic on ordered set partitions can be derived from a result of Wachs [12, Theorem 2.1] (see also [18, Theorem 4]). Steingrímsson [18] gave several Euler–Mahonian statistics and conjectured more such statistics.

The aim of this paper is to provide many more such statistics on ordered set partitions by taking two different approaches. In the first approach our starting point is an explicit continued fraction expansion of the generating function $\sum_{n,k} [k]_q! S_q(n, k) u^k z^n$. By embedding the ordered partitions into preferential arrangements we can encode the latter by weighted Motzkin paths, whose generating function is known to have an explicit continued fraction expansion. By identifying these two continued fractions we find many Euler–Mahonian statistics on ordered partitions. In the second approach we shall construct a bijection ψ between ordered partitions and walks in some digraphs (see section 3). This bijection keeps track of several statistics on ordered partitions. Then, by the transfer-matrix method, we evaluate the generating functions of these statistics on ordered partitions. Finally, by identifying the generating functions we derive several Euler–Mahonian statistics. We conclude this paper with some open problems.

2. Definitions and main results. Let $\sigma = \sigma_1\sigma_2\ldots\sigma_n \in S_n$. By convention, we assume that $\sigma_0 = n+1$ and $\sigma_{n+1} = 0$. For $1 \leq i \leq n$, we say that σ_i is a

- *valley* if $\sigma_{i-1} > \sigma_i < \sigma_{i+1}$;
- *underlined valley* if $\sigma_{i-1} > \sigma_i < \sigma_{i+1}$ and σ_i is underlined;
- *peak* if $\sigma_{i-1} < \sigma_i > \sigma_{i+1}$;
- *double ascent* if $\sigma_{i-1} < \sigma_i < \sigma_{i+1}$;
- *underlined double ascent* if $\sigma_{i-1} < \sigma_i < \sigma_{i+1}$ and σ_i is underlined;
- *double descent* if $\sigma_{i-1} > \sigma_i > \sigma_{i+1}$.

We will denote by $\text{DA}(\sigma)$, $\underline{\text{DA}}(\sigma)$, $\text{VA}(\sigma)$, $\underline{\text{VA}}(\sigma)$, $\text{DD}(\sigma)$, $\text{PE}(\sigma)$ the set of double ascents, underlined double ascents, valleys, underlined valleys, double descents, and peaks of σ , respectively. For instance, if $\sigma = \underline{2}931485\underline{6}7$, then

$$\begin{aligned} \text{DA}(\sigma) &= \{4\}, & \underline{\text{DA}}(\sigma) &= \{6\}, & \text{VA}(\sigma) &= \{1, 5\}, \\ \underline{\text{VA}}(\sigma) &= \{2\}, & \text{DD}(\sigma) &= \{3\}, & \text{PE}(\sigma) &= \{7, 8, 9\}. \end{aligned}$$

For $i \in [n]$, define $\text{lsg}_i(\sigma)$ (resp., $\text{rsg}_i(\sigma)$) as the number of runs of σ strictly to the left (resp., right) of i which contain an element smaller than i and an element greater than i . Then we define

$$\text{lsg}(\sigma) = \sum_{i=1}^n \text{lsg}_i(\sigma), \quad \text{rsg}(\sigma) = \sum_{i=1}^n \text{rsg}_i(\sigma).$$

More generally, for a sequence of statistics STAT_i ($i \geq 0$) on a set E , and for F a mapping which associates to any element in E a finite set of integers, we define the statistic $\text{STAT}(F)$ on E by

$$(2.1) \quad \text{STAT}(F)(e) = \sum_{i \in F(e)} \text{STAT}_i(e) \quad \text{for any } e \in E.$$

For instance, the statistic $\text{lsg}(\text{DD} \cup \text{VA})$ is defined for any $\sigma \in S_n$ by

$$\text{lsg}(\text{DD} \cup \text{VA})(\sigma) = \sum_{i \in \text{DD}(\sigma) \cup \text{VA}(\sigma)} \text{lsg}_i(\sigma).$$

Consider the generating polynomial of preferential arrangements:

$$\begin{aligned} \mu_n &= \sum_{\pi \in S_n} \alpha^{\#\text{DD}(\pi)} \beta^{\#\text{DA}(\pi)} \gamma^{\#\underline{\text{DA}}(\pi)} \delta^{\#\text{VA}(\pi)} \xi^{\#\underline{\text{VA}}(\pi)} \eta^{\#\text{PE}(\pi)} \\ &\quad \times a^{\text{lsg}(\text{DD})(\pi)} b^{\text{rsg}(\text{DD})(\pi)} c^{\text{lsg}(\text{DA})(\pi)} d^{\text{rsg}(\text{DA})(\pi)} f^{\text{lsg}(\underline{\text{DA}})(\pi)} g^{\text{rsg}(\underline{\text{DA}})(\pi)} \\ &\quad \times s^{\text{lsg}(\text{VA})(\pi)} t^{\text{rsg}(\text{VA})(\pi)} x^{\text{lsg}(\underline{\text{VA}})(\pi)} y^{\text{rsg}(\underline{\text{VA}})(\pi)} v^{\text{lsg}(\text{PE})(\pi)} w^{\text{rsg}(\text{PE})(\pi)}. \end{aligned}$$

Our first result is an explicit continued fraction expansion for the ordinary generating function of μ_n .

THEOREM 2.1. *We have the following continued fraction expansion:*

$$(2.2) \quad 1 + \sum_{n \geq 1} \mu_n z^n = \frac{1}{1 - b_0 z - \frac{\lambda_1 z^2}{1 - b_1 z - \frac{\lambda_2 z^2}{\ddots}}},$$

where $b_n = \alpha[n+1]_{a,b} + \beta[n]_{c,d} + \gamma[n]_{f,g}$, $\lambda_n = (\delta[n]_{s,t} + \xi[n]_{x,y}) \eta[n]_{v,w}$.

In order to derive Euler–Mahonian statistics on ordered set partitions we need the following lemma.

LEMMA 2.2. *We have*

$$(2.3) \quad \sum_{n,k} [k]_q! S_q(n, k) u^k z^n = \frac{1}{1 - b_0 z - \frac{\lambda_1 z^2}{1 - b_1 z - \frac{\lambda_2 z^2}{\ddots}}},$$

where $b_n = uq^n[n+1]_q + [n]_q(1 + uq^n)$, $\lambda_n = uq^{n-1}[n]_q^2(1 + uq^n)$.

The proofs of Theorem 2.1 and Lemma 2.2 will be given in section 3.

There are several possibilities to specialize the parameters in (2.2) to obtain (2.3). Clearly such a specialization will induce an Euler–Mahonian statistic on ordered set partitions. For example, setting $\alpha = \eta = u$, $\gamma = \xi = uq$, $\beta = \delta = c = s = 1$, $a = f = x = v = q^2$, and $b = d = g = t = y = w = q$ in Theorem 2.1, the right-hand side of (2.2) reduces to that of (2.3). On the other hand, for $\sigma \in \underline{S}_n$, let

$$s(\sigma) = (2 \operatorname{lsg} + \operatorname{rsg} - 2 \operatorname{lsg}(\operatorname{DA}) - 2 \operatorname{lsg}(\operatorname{VA}))(\sigma) + \#\operatorname{VA}(\sigma) + \#\operatorname{DA}(\sigma).$$

Then

$$(2.4) \quad \sum_{\pi \in \underline{S}_n} u^{n - \#\operatorname{DA}(\sigma) - \#\operatorname{VA}(\sigma)} q^{s(\sigma)} = \sum_{k=1}^n [k]_q! S_q(n, k) u^k.$$

As $n - \#\operatorname{DA}(\sigma) - \#\operatorname{VA}(\sigma)$ is the number of blocks of σ identified as an ordered set partition, extracting the coefficients of u^k in (2.4) yields the following result.

THEOREM 2.3. *We have*

$$(2.5) \quad \sum_{\sigma \in \underline{S}_n^k} q^{s(\sigma)} = [k]_q! S_q(n, k),$$

where $\underline{S}_n^k$ is the set of underlined permutations of $[n]$ whose number of descents and underlined ascents equals $k - 1$.

Remark 2.4. By varying the coefficients lsg and rsg in the specialization we have $2^6 = 64$ statistics satisfying (2.5), i.e., $\{a, b\} = \{f, g\} = \{x, y\} = \{v, w\} = \{q, q^2\}$ and $\{c, d\} = \{s, t\} = \{1, q\}$.

It is not difficult to translate all these statistics into the model of ordered set partitions, but we decide to omit their description because they are too involved.

Let $\pi = B_1/B_2/\cdots/B_k$ be an ordered partition in $\mathcal{OP}_n^k$. The *opener* of a block is its least element, while the *closer* is its greatest element. We will denote by $\operatorname{open}(\pi)$ and $\operatorname{clos}(\pi)$ the set of openers and closers, respectively, of π . The letters (integers) of π are divided into four classes:

- *singletons*: elements of the singleton blocks;
- *strict openers*: smallest elements of the nonsingleton blocks;
- *strict closers*: largest elements of the nonsingleton blocks;
- *transients*: all other elements, i.e., nonextremal elements of nonsingleton blocks.

The sets of strict openers, strict closers, singletons, and transients of π will be denoted by $\mathcal{O}(\pi)$, $\mathcal{C}(\pi)$, $\mathcal{S}(\pi)$, and $\mathcal{T}(\pi)$, respectively. Obviously we have

$$\operatorname{open}(\pi) = \mathcal{O}(\pi) \cup \mathcal{S}(\pi), \quad \operatorname{clos}(\pi) = \mathcal{C}(\pi) \cup \mathcal{S}(\pi), \quad \mathcal{S}(\pi) = \operatorname{open}(\pi) \cap \operatorname{clos}(\pi).$$

The 4-tuple $(\mathcal{O}(\pi), \mathcal{C}(\pi), \mathcal{S}(\pi), \mathcal{T}(\pi))$ is called the *type* of π . For instance, if $\pi = 35/246/1/78$, then

$$\mathcal{O}(\pi) = \{2, 3, 7\}, \quad \mathcal{C}(\pi) = \{5, 6, 8\}, \quad \mathcal{S}(\pi) = \{1\}, \quad \text{and} \quad \mathcal{T}(\pi) = \{4\}.$$

We define a *partial order* on blocks B_i 's as follows: $B_i > B_j$ if all the letters of B_i are greater than those of B_j —in other words, if the opener of B_i is greater than the closer of B_j .

DEFINITION 2.5. For a permutation σ of $[n]$, the pair (i, j) is an inversion if $1 \leq i < j \leq n$ and $\sigma(i) > \sigma(j)$. Let $\text{inv } \sigma$ be the number of inversions in σ and

$$\text{cinv } \sigma = \binom{n}{2} - \text{inv } \sigma.$$

By convention, for any ordered set partition π , we put $\text{Inv } \pi = \text{inv}(\text{perm}(\pi))$ and $\text{cInv } \pi = \text{cinv}(\text{perm}(\pi))$.

DEFINITION 2.6. A block inversion in π is a pair (i, j) such that $i < j$ and $B_i > B_j$. We denote by $\text{bInv } \pi$ the number of block inversions in π . We also set $\text{cbInv} = \binom{k}{2} - \text{bInv}$.

Let w_i be the index of the block (counting from left to right) containing i , namely, the integer j such that $i \in B_j$. Following Steingrímsson [18], for $1 \leq i \leq n$ we define ten coordinate statistics on $\pi \in \mathcal{OP}_n^k$:

$$\begin{aligned} \text{ros}_i(\pi) &= \#\{j \in \text{open}(\pi) \mid i > j, w_j > w_i\}, \\ \text{rob}_i(\pi) &= \#\{j \in \text{open}(\pi) \mid i < j, w_j > w_i\}, \\ \text{rcs}_i(\pi) &= \#\{j \in \text{clos}(\pi) \mid i > j, w_j > w_i\}, \\ \text{rcb}_i(\pi) &= \#\{j \in \text{clos}(\pi) \mid i < j, w_j > w_i\}, \\ \text{los}_i(\pi) &= \#\{j \in \text{open}(\pi) \mid i > j, w_j < w_i\}, \\ \text{lob}_i(\pi) &= \#\{j \in \text{open}(\pi) \mid i < j, w_j < w_i\}, \\ \text{lcs}_i(\pi) &= \#\{j \in \text{clos}(\pi) \mid i > j, w_j < w_i\}, \\ \text{lcb}_i(\pi) &= \#\{j \in \text{clos}(\pi) \mid i < j, w_j < w_i\}, \end{aligned}$$

and let $\text{rsb}_i(\pi)$ (resp., $\text{lsb}_i(\pi)$) be the number of blocks B in π to the right (resp., left) of the block containing i such that the opener of B is smaller than i and the closer of B is greater than i . Then define ros , rob , rcs , rcb , lob , los , lcs , lcb , lsb , and rsb as the sum of their coordinate statistics, e.g.,

$$\text{ros} = \sum_i \text{ros}_i.$$

Remark 2.7. Note that ros is the abbreviation of “right, opener, smaller,” while lcb is the abbreviation of “left, closer, bigger,” etc.

Given an ordered set partition π , let π^r be the ordered set partition obtained from π by reversing the order of the blocks. This turns a left (resp., right) opener into a right (resp., left) opener, and likewise for the closers. Moreover, let π^c be the ordered set partition obtained by complementing each of the letters in π , that is, by replacing the letter i by $n + 1 - i$. Then, it is easy to see that $\text{rob}(\pi^c) = \text{rcs}(\pi)$ and $\text{ros}(\pi^c) = \text{rcb}(\pi)$, and likewise for the left and closer statistics. Thus the eight statistics obtained by independently varying left/right, opener/closer, and smaller/bigger fall into only two categories when it comes to their distribution on ordered partitions. One

of these categories consists of rob, lob, rcs, and lcs, and the other contains ros, los, rcb, and lcb. Note that these results are completely false on the unordered set partitions.

For instance, we give the values of the coordinate statistics computed on the ordered partition $\pi = 6\ 8/5/1\ 4\ 7/3\ 9/2$:

$$\begin{array}{rcccccccc} \pi = & 6\ 8 & / & 5 & / & 1\ 4\ 7 & / & 3\ 9 & / & 2 \\ \text{los}_i : & 0\ 0 & / & 0 & / & 0\ 0\ 2 & / & 1\ 3 & / & 1 \\ \text{ros}_i : & 4\ 4 & / & 3 & / & 0\ 2\ 2 & / & 1\ 1 & / & 0 \\ \text{lob}_i : & 0\ 0 & / & 1 & / & 2\ 2\ 0 & / & 2\ 0 & / & 3 \\ \text{rob}_i : & 0\ 0 & / & 0 & / & 2\ 0\ 0 & / & 0\ 0 & / & 0 \\ \text{lcs}_i : & 0\ 0 & / & 0 & / & 0\ 0\ 1 & / & 0\ 3 & / & 0 \\ \text{rcs}_i : & 2\ 3 & / & 1 & / & 0\ 1\ 1 & / & 1\ 1 & / & 0 \\ \text{lcb}_i : & 0\ 0 & / & 1 & / & 2\ 2\ 1 & / & 3\ 0 & / & 4 \\ \text{rcb}_i : & 2\ 1 & / & 2 & / & 2\ 1\ 1 & / & 0\ 0 & / & 0 \\ \text{lsb}_i : & 0\ 0 & / & 0 & / & 0\ 0\ 1 & / & 1\ 0 & / & 1 \\ \text{rsb}_i : & 2\ 1 & / & 2 & / & 0\ 1\ 1 & / & 0\ 0 & / & 0 \end{array}$$

Note that there are four block inversions: $\{6, 8\} > \{5\}$, $\{6, 8\} > \{2\}$, $\{5\} > \{2\}$, and $\{3, 9\} > \{2\}$, and thus $\text{bInv } \pi = 4$. Moreover, $\text{perm}(\pi) = 54132$ and thus $\text{Inv}(\pi) = 8$.

Inspired by the statistic mak on the permutations in [6], Steingrímsson introduced its analogue on $\mathcal{OP}_n^k$ as follows:

$$(2.6) \quad \text{mak} = \text{ros} + \text{lcs},$$

$$(2.7) \quad \text{lmak} = n(k-1) - [\text{los} + \text{rcs}],$$

$$(2.8) \quad \text{mak}' = \text{lob} + \text{rcb},$$

$$(2.9) \quad \text{lmak}' = n(k-1) - [\text{lcb} + \text{rob}].$$

The following result due to Ksavlouf and Zeng [10] permit us to reduce the original conjectures in [18] almost by half. For completeness, we include a more straightforward proof.

PROPOSITION 2.8. *On $\mathcal{OP}_n^k$ the following functional identities hold:*

$$\text{mak} = \text{lmak}' \quad \text{and} \quad \text{mak}' = \text{lmak}.$$

Proof. Let $i \in [n]$. For any $\pi \in \mathcal{OP}_n^k$ the value $(\text{los}_i + \text{lob}_i)\pi$ (resp., $(\text{ros}_i + \text{rob}_i)\pi$) is equal to the number of blocks on the left (resp., right) of the block containing i , so $(\text{los}_i + \text{lob}_i + \text{ros}_i + \text{rob}_i)\pi = k-1$. Similarly, we have $(\text{lcs}_i + \text{lcb}_i + \text{rcs}_i + \text{rcb}_i)\pi = k-1$. Summing over all $i \in [n]$ yields that $\text{los} + \text{lob} + \text{ros} + \text{rob} = \text{lcs} + \text{lcb} + \text{rcs} + \text{rcb} = n(k-1)$. The proposition is then equivalent to $\text{lob} + \text{los} = \text{lcb} + \text{lcs}$, which is obvious. $\square$

Let $\mathcal{OP}^k := \bigcup_{n \geq k} \mathcal{OP}_n^k$ be the set of all ordered set partitions with k blocks. Define also the statistic cinvLSB on $\mathcal{OP}^k$ by

$$(2.10) \quad \text{cinvLSB} := \text{lsb} + \text{cbInv} + \binom{k}{2}.$$

Consider the generating function of ordered partitions with k blocks,

$$(2.11) \quad \phi_k(a; x, y, t, u) := \sum_{\pi \in \mathcal{OP}^k} x^{(\text{mak} + \text{bInv})\pi} y^{\text{cinvLSB } \pi} t^{\text{Inv } \pi} u^{\text{cInv } \pi} a^{|\pi|},$$

and its variant by replacing mak by lmak,

$$(2.12) \quad \varphi_k(a; x, y, t, u) := \sum_{\pi \in \mathcal{OP}^k} x^{(\text{lmak} + \text{bInv})\pi} y^{\text{cinvLSB } \pi} t^{\text{Inv } \pi} u^{\text{cInv } \pi} a^{|\pi|},$$

where $|\pi| = n$ if π is an ordered partition of $[n]$.

One of the main results of this paper is the following theorem, which in particular shows that the above two generating functions are identical.

THEOREM 2.9. *We have*

$$(2.13) \quad \phi_k(a; x, y, t, u) = \frac{a^k (xy)^{\binom{k}{2}} [k]_{tx,uy}!}{\prod_{i=1}^k (1 - a[i]_{x,y})},$$

$$(2.14) \quad \varphi_k(a; x, y, t, u) = \frac{a^k (xy)^{\binom{k}{2}} [k]_{tx,uy}!}{\prod_{i=1}^k (1 - a[i]_{x,y})}.$$

The proof of this theorem will be given in sections 4 and 5. We first show how to derive Euler–Mahonian statistics on ordered set partitions from the above theorem.

By (2.11) and (2.12) we have

$$\begin{aligned} \sum_{\pi \in \mathcal{OP}^k} q^{(\text{mak} + \text{bInv})\pi} a^{|\pi|} &= \phi_k(a; q, 1, 1, 1), \\ \sum_{\pi \in \mathcal{OP}^k} q^{(\text{lmak} + \text{bInv})\pi} a^{|\pi|} &= \varphi_k(a; q, 1, 1, 1), \\ \sum_{\pi \in \mathcal{OP}^k} q^{\text{cinvLSB } \pi} a^{|\pi|} &= \phi_k(a; 1, q, 1, 1) = \varphi_k(a; 1, q, 1, 1). \end{aligned}$$

Theorem 2.9 implies that the right-hand sides of the above three identities are all equal to

$$(2.15) \quad \frac{a^k q^{\binom{k}{2}} [k]_q!}{\prod_{i=1}^k (1 - a[i]_q)} = \sum_{n \geq k} [k]_q! S_q(n, k) a^n,$$

where the last equality follows directly from (1.4). Thus we have proved the following result, which was conjectured by Steingrímsson [18].

THEOREM 2.10. *The following inversion-like statistics are Euler–Mahonian:*

$$\text{mak} + \text{bInv}, \quad \text{lmak} + \text{bInv}, \quad \text{cinvLSB}.$$

In other words, the generating functions of the above statistics over $\mathcal{OP}_n^k$ are all equal to $[k]_q! S_q(n, k)$.

Similarly, we have

$$\begin{aligned} \sum_{\pi \in \mathcal{OP}^k} q^{(\text{mak} + \text{bInv} - \text{inv} + \text{cinv})\pi} a^{|\pi|} &= \phi_k(a; q, 1, 1/q, q), \\ \sum_{\pi \in \mathcal{OP}^k} q^{(\text{lmak} + \text{bInv} - \text{inv} + \text{cinv})\pi} a^{|\pi|} &= \varphi_k(a; q, 1, 1/q, q), \\ \sum_{\pi \in \mathcal{OP}^k} q^{(\text{cinvLSB} + \text{inv} - \text{cinv})\pi} a^{|\pi|} &= \phi_k(a; 1, q, q, 1/q) = \varphi_k(a; 1, q, q, 1/q), \end{aligned}$$

and the three right-hand sides are all equal to (2.15). In other words, we have the following result.

THEOREM 2.11. *The following statistics are Euler–Mahonian:*

$$\text{mak} + \text{bInv} - (\text{inv} - \text{cinv}), \quad \text{lak} + \text{bInv} - (\text{inv} - \text{cinv}), \quad \text{cinvLSB} + (\text{inv} - \text{cinv}).$$

As a further consequence of Theorem 2.9, we can give an alternative proof of the following “hard” combinatorial interpretations for q -Stirling numbers of the second kind. According to Wachs and White [20], a statistic on $\mathcal{P}_n^k$, whose generating function equals the q -Stirling number $S_q(n, k)$, is called “hard” if it is difficult to show that its generating function satisfies the recurrence relation (1.3).

COROLLARY 2.12. *We have*

$$S_q(n, k) = \sum_{\pi \in \mathcal{P}_n^k} q^{\text{mak } \pi} = \sum_{\pi \in \mathcal{P}_n^k} q^{\text{lak } \pi} = \sum_{\pi \in \mathcal{P}_n^k} q^{\text{lsb } \pi + \binom{k}{2}}.$$

Remark 2.13. The first interpretation was proved by Ksavrelof and Zeng [10], while the second is due to Wachs and White [20] and the third interpretation was first proved by Stanton (see [20]).

Proof. An unordered set partition can be identified with an ordered set partition without inversion, i.e., $\mathcal{P}^k = \{\pi \in \mathcal{OP}^k \mid \text{Inv } \pi = 0\}$. Since the statistic bInv vanishes on $\mathcal{P}^k$, we derive from (2.11) and (2.12) that

$$\begin{aligned} \sum_{\pi \in \mathcal{P}^k} q^{\text{mak } \pi} a^{|\pi|} &= \phi_k(a; q, 1, 0, 1), \\ \sum_{\pi \in \mathcal{P}^k} q^{\text{lak } \pi} a^{|\pi|} &= \varphi_k(a; q, 1, 0, 1). \end{aligned}$$

Now, by definition (2.10), the identity $\text{cinvLSB} = \text{lsb} + 2\binom{k}{2}$ holds on $\mathcal{P}^k$. It then follows from (2.11) that

$$\sum_{\pi \in \mathcal{P}^k} q^{\binom{k}{2} + \text{lsb } \pi} a^{|\pi|} = q^{-\binom{k}{2}} \phi_k(a; 1, q, 0, 1).$$

Now, applying Theorem 2.9 we see that the right-hand sides of the above three identities are equal to $\sum_{n \geq k} S_q(n, k) a^n$ in view of (2.15). $\square$

3. Ordered partitions and Motzkin paths.

3.1. Encoding ordered partitions by weighted Motzkin paths. Recall that a Motzkin path of length n is a path $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_n)$ from $\gamma_0 = (0, 0)$ to $\gamma_n = (n, 0)$ in the first quadrant with unit steps $\gamma_i - \gamma_{i-1} = (1, 0)$, $(1, 1)$, or $(1, -1)$, which are called East, Northeast, and Southeast, respectively. For convenience, we will distinguish three kinds of East steps, E , $\overline{E}$, and $\widetilde{E}$, and two kinds of Northeast steps, NE and $\overline{NE}$. The weight of a step at height h is a_h (resp., $\bar{a}_h$, c_h , b'_h , b''_h , and b_h''') if it is NE (resp., $\overline{NE}$, SE , E , $\overline{E}$, and $\widetilde{E}$).

Consider the weighted Motzkin paths with six kinds of steps and weights defined by the following array:

Step	NE	$\overline{NE}$	E	$\overline{E}$	$\widetilde{E}$	SE
Weight	$a_h = h + 1$	$\bar{a}_h = h + 1$	$b_h = h$	$\bar{b}_h = h$	$\widetilde{b}_h = h + 1$	$c_h = h$

A *Meixner diagram* of length n is a pair $h = (\omega, \xi)$, where ω is a Motzkin path of length n and $\xi = (\xi_i)_{1 \leq i \leq n}$ is a sequence of integers such that if the i th step of ω is of height h , then

$$\begin{aligned} 1 \leq \xi_i \leq h+1 & \quad \text{if the step is } NE, \overline{NE}, \text{ or } \widetilde{E}, \\ 2 \leq \xi_i \leq h+1 & \quad \text{if the step is } E, \overline{E}, \text{ or } SE. \end{aligned}$$

We denote by Γ_n the set of *Meixner diagrams* of length n .

There is a classical bijection due to Françon and Viennot [7] (see also [19, Chapter II-38]) between the symmetric group S_n and a set of weighted Motzkin paths of length n , called *Laguerre diagrams*. We shall adapt this classical bijection to give a bijection between $\underline{S}_n$ and Γ_n .

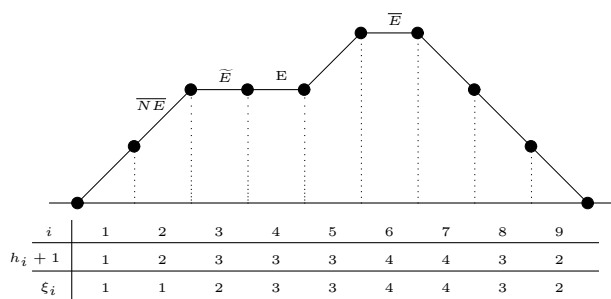


FIG. 1. A *Meixner diagram* of length 9.

Let $h = (\omega, \xi) \in \Gamma_n$. We construct a sequence of words w_i on the alphabet $\{-\} \cup [i]$, $i = 0 \dots n$, satisfying the following conditions:

- (a) for $i \in [n]$, the number of occurrences of $-$ in w_{i-1} is equal to $h_i + 1$, where h_i is the height of the i th step of ω ;
- (b) for each i , $0 \leq i \leq n$, the first letter of the word w_i is the symbol $-$;

and defined by induction as follows:

1. $w_0 = -$;
2. for $i \in [n]$, the word w_i is obtained from w_{i-1} by replacing the ξ_i th occurrences of $-$ (from left to right) in w_{i-1} by i (resp., $-i$, $-\underline{i}$, i , $\underline{i}$, $-i$) if the i th step of ω is SE (resp., NE , $\overline{NE}$, E , $\overline{E}$, $\widetilde{E}$).

It is clear that w_0 satisfies (a) and (b). Now, let $i \in [n]$ and suppose that $w_0, \dots, w_{i-1}$ are defined and that they satisfy the conditions (a) and (b). The definition of a path diagram and condition (a) allow us to construct w_i thanks to condition (2). Step 2 shows also that w_i satisfy conditions (a) and (b). Then the word w_n has only one occurrence of $-$ in the beginning. We then replace it by $n+1$. We can identify such permutations with underlined permutations on $[n]$ by deleting the first letter.

For instance, if h is the Meixner diagram in Figure 1, the step-by-step construction of $\phi(h)$ goes as follows:

$$\begin{aligned} w_0 &= -, & w_1 &= -1-, & w_2 &= -\underline{2}-1-, & w_3 &= -\underline{2}-31-, & w_4 &= -\underline{2}-314-, \\ w_5 &= -\underline{2}-314-5-, & w_6 &= -\underline{2}-314-5\underline{6}-, & w_7 &= -\underline{2}-314-5\underline{6}7, \\ w_8 &= -\underline{2}-31485\underline{6}7, & w_9 &= -\underline{2}931485\underline{6}7. \end{aligned}$$

Hence $\phi(h) = \underline{2}931485\underline{6}7$.

As in [15, 19, 3], we have the following result. The proof is left to the reader.

THEOREM 3.1. *The application ϕ is a bijection between Γ_n and $\underline{S}_n$. Moreover if $h = (\omega, \xi) \in \Gamma_n$ and $\sigma = \phi(h) \in \underline{S}_n$, then for $1 \leq i \leq n$*

- (a) *$i \in \text{VA}(\sigma)$ (resp., $\underline{\text{VA}}(\sigma)$, $\underline{\text{DA}}(\sigma)$, $\underline{\text{DA}}(\sigma)$, $\text{DD}(\sigma)$, $\text{PE}(\sigma)$) if and only if the i th step of ω is NE (resp., $\overline{NE}$, E , $\overline{E}$, $\tilde{E}$, SE);*
 (b) *$\text{rsg}_i(\sigma) = h_i + 1 - \xi_i$ and*

$$\begin{cases} \text{lsg}_i(\sigma) = \xi_i - 1 & \text{if the } i\text{th step is NE, } \overline{NE}, \text{ or } \tilde{E}; \\ \text{lsg}_i(\sigma) = \xi_i - 2 & \text{if the } i\text{th step is E, } \overline{E}, \text{ or SE.} \end{cases}$$

Theorem 2.1 then follows immediately from Theorem 3.1 by applying Flajolet's combinatorial theorem of continued fractions [4].

3.2. A continued fraction expansion. The recurrence (1.3) is equivalent to

$$\sum_{n \geq k} S_q(n, k) z^n = \frac{z^k}{(1-z)(1-[2]_q z) \cdots (1-[k]_q z)}.$$

Let $[b; n]_q! = b(b+q) \cdots (b+q+\cdots+q^{n-1})$. Then

$$(3.1) \quad \sum_{n, k} [b; k]_q! S_q(n, k) u^k z^n = \sum_{k \geq 0} \frac{[b; k]_q! u^k z^k}{(1-z)(1-[2]_q z) \cdots (1-[k]_q z)}.$$

Using the method in [23], we can derive an explicit continued fraction expansion of the last expression.

THEOREM 3.2. *We have*

$$\sum_{n, k} [b; k]_q! S_q(n, k) u^k z^n = \frac{1}{1 - b_0 z - \frac{\lambda_1 z^2}{1 - b_1 z - \frac{\lambda_2 z^2}{\ddots}}},$$

where

$$(3.2) \quad \begin{cases} b_n = uq^n(b+q+\cdots+q^n) + [n]_q(1+uq^n), & n \geq 0, \\ \lambda_n = uq^{n-1}(b+q+\cdots+q^{n-1})[n]_q(1+uq^n), & n \geq 1. \end{cases}$$

Proof. Let $f(b, z) := \sum_{n, k} [b; k]_q! S_q(n, k) u^k z^n$. Then it is easy to see that f satisfies the following functional equation:

$$(3.3) \quad f(b, z) = 1 + \frac{ubz}{1-z} f\left(\frac{b}{q} + 1, \frac{qz}{1-z}\right).$$

Suppose that

$$f(b, z) = \frac{1}{1 - \frac{c_1(b)z}{1 - \frac{c_2(b)z}{\ddots}}}.$$

By contraction from the 1st row we have

$$f(b, z) = 1 + \frac{c_1(b)z}{1 - (c_1(b) + c_2(b))z - \frac{c_2(b)c_3(b)z^2}{\ddots}}.$$

By contraction from the second row we have

$$(3.4) \quad f(b, z) = \frac{1}{1 - c_1(b)z - \frac{c_1(b)c_2(b)z^2}{1 - (c_2(b) + c_3(b))z - \frac{c_3(b)c_4(b)z^2}{\ddots}}}.$$

It follows that

$$\frac{ubz}{1-z} f\left(\frac{b}{q} + 1, \frac{qz}{1-z}\right) = \frac{ubz}{1 - (qc_1(b/q + 1) + 1)z - \frac{c_1(b/q + 1)c_2(b/q + 1)q^2 z^2}{\ddots}}.$$

Hence

$$\begin{aligned} c_1(b) &= ub, \\ c_1(b) + c_2(b) &= qc_1(b/q + 1) + 1 \implies c_2(b) = uq + 1, \\ c_2(b)c_3(b) &= c_1(b/q + 1)c_2(b/q + 1)q^2 \implies c_3(b) = u(b + q)q, \\ c_3(b) + c_4(b) &= 1 + qc_2(b/q + 1) + qc_3(b/q + 1) \implies c_4(b) = (1 + q)(1 + uq^2), \\ c_4(b)c_5(b) &= q^2 c_3(b/q + 1)c_4(b/q + 1) \implies c_5(b) = uq^2(b + q + q^2). \end{aligned}$$

This implies immediately by induction that

$$\begin{aligned} c_{2i+1}(b) &= uq^i(b + q + \cdots + q^i), \quad i \geq 0, \\ c_{2i}(b) &= [i]_q(1 + uq^i), \quad i \geq 1. \end{aligned}$$

Substituting these in (3.4) yields the desired result. $\square$

Note that when $b = 1$ Theorem 3.2 reduces to Lemma 2.2.

4. Ordered set partitions and walks in digraphs.

4.1. Encoding ordered set partitions by walk diagrams. Let $\pi = B_1/B_2/\cdots/B_k$ be an ordered partition of $[n]$ and i an integer in $[n]$. The restriction $B_j \cap [i]$ of a block B_j to $[i]$ is said to be *active* if $B_j \neq [i]$ and $B_j \cap [i] \neq \emptyset$, and *complete* if $B_j \subseteq [i]$. We define the *trace* of the ordered partition π on $[i]$ as an ordered partition $T_i(\pi)$ on $[i]$ with two kinds of blocks, active or complete, i.e.,

$$T_i(\pi) = B_1(\leq i)/B_2(\leq i)/\cdots/B_k(\leq i),$$

where $B_j(\leq i)$ is the complete or active block $B_j \cap [i]$, while empty sets are omitted. The sequence $(T_i(\pi))_{1 \leq i \leq n}$ is called the *trace* of the ordered set partition π .

DEFINITION 4.1. Let $D = (V, E)$ be the digraph on $V = \mathbb{N}^2$ with edge set E defined by

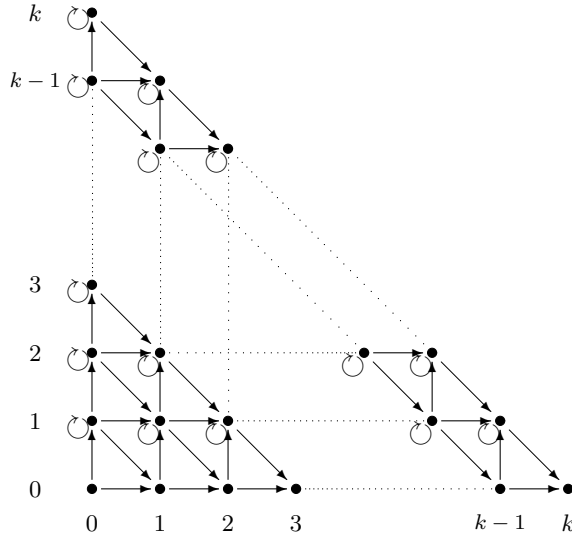
$$E = \{(u, v) \in V^2 \mid u = v = (x, y) \text{ with } y > 0 \text{ or } u - v = (0, 1), (1, 0), (1, -1)\}.$$

For any integer $k \geq 0$, let $V_k = \{(i, j) \in V \mid i + j \leq k\}$ and let D_k be the restriction of the digraph D on V_k .

An illustration of D_k is given in Figure 2.

Clearly there are four types of edges in D . For convenience, an edge (u, v) is called

- *North* if $v = u + (0, 1)$;
- *East* if $v = u + (1, 0)$;

FIG. 2. The digraph D_k .

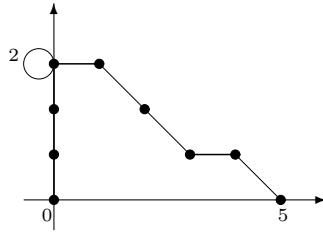
- *Southeast* if $v = u + (1, -1)$;
- *Null* if $v = u$.

DEFINITION 4.2. A walk of depth k and length n in D is a sequence $\omega = (\omega_0, \omega_1, \dots, \omega_n)$ of vertices in D such that $\omega_0 = (0, 0)$, $\omega_n = (k, 0)$ and (ω_i, ω_{i+1}) is an edge of D for $i = 0, \dots, n-1$. Moreover, the abscissa and ordinate of ω_i are called the abscissa and height of the step (ω_i, ω_{i+1}) , respectively. Let Ω_n^k be the set of walks of depth k and length n and let Ω^k be the set of walks of depth k .

We can visualize a walk ω by drawing a segment from ω_{i-1} to ω_i in the xy plane. For instance, if

$$\omega = ((0, 0), (0, 1), (0, 2), (0, 3), (0, 3), (0, 3), (1, 3), (2, 2), (3, 1), (4, 1), (5, 0)),$$

then the illustration is given in Figure 3.

FIG. 3. A walk in Ω_{10}^5 with two successive Null steps from $(0, 3)$ to $(0, 3)$.

DEFINITION 4.3. A walk diagram of depth k and length n is a pair (ω, ξ) , where $\omega = ((p_i, q_i))_{0 \leq i \leq n}$ is a walk in Ω_n^k and $\xi = (\xi_i)_{1 \leq i \leq n}$ is a sequence of integers such that

- $1 \leq \xi_i \leq q_{i-1}$ if the i th step of ω is Null or Southeast,
- $1 \leq \xi_i \leq p_{i-1} + q_{i-1} + 1$ if the i th step of ω is North or East.

Denote by Δ_n^k the set of walk diagrams of depth k and length n and by $\Delta^k = \bigcup_{n \geq 0} \Delta_n^k$ the set of walk diagrams of depth k . The following is the main result of this section.

THEOREM 4.4. *For each $n \geq k \geq 1$, there is a bijection $\psi : \Delta_n^k \rightarrow \mathcal{OP}_n^k$ such that if $\psi((\omega, \xi)) = \pi$ for $(\omega, \xi) \in \Delta_n^k$, then*

(a) *if the i th step of ω is North (resp., East), then $i \in \mathcal{O}(\pi)$ (resp., $i \in \mathcal{S}(\pi)$) and*

$$\begin{aligned} (\text{lcs} + \text{rcs})_i(\pi) &= p_{i-1}, & \text{los}_i(\pi) &= \xi_i - 1, \\ (\text{lsb} + \text{rsb})_i(\pi) &= q_{i-1}, & \text{ros}_i(\pi) &= p_{i-1} + q_{i-1} + 1 - \xi_i; \end{aligned}$$

(b) *if the i th step of ω is Southeast (resp., Null), then $i \in \mathcal{C}(\pi)$ (resp., $i \in \mathcal{T}(\pi)$) and*

$$\begin{aligned} (\text{lcs} + \text{rcs})_i(\pi) &= p_{i-1}, & \text{lsb}_i(\pi) &= \xi_i - 1, \\ (\text{lsb} + \text{rsb})_i(\pi) &= q_{i-1} - 1, & \text{rsb}_i(\pi) &= q_{i-1} - \xi_i. \end{aligned}$$

Proof. Given a walk diagram $(\omega, \xi) \in \Delta_n^k$, we define the ordered partition $\pi = \psi((\omega, \xi)) \in \mathcal{OP}_n^k$ by constructing successively its traces $T_0 := \emptyset, T_1, \dots, T_n := \pi$. For $1 \leq i \leq n$, assume that $T_{i-1} = B_{i1}/B_{i2}/\dots/B_{i\ell}$ with $\ell = p_{i-1} + q_{i-1}$. Then there are $\ell + 1$ places to insert an active block with i as an opener or a singleton block $\{i\}$: before B_{i1} , between B_{ij} and $B_{i(j+1)}$, where $1 \leq j \leq \ell - 1$, or after $B_{i\ell}$. There are q_{i-1} places to insert i as a transient or closer: in any of the q_{i-1} active blocks. We label these places from left to right. We extend T_{i-1} to T_i as follows:

- if the i th step of ω is North (resp., East), then we create an active block (resp., singleton) with i in T_{i-1} at the ξ_i th place and close it if the step is East;
- if the i th step of ω is Null (resp., Southeast), then we insert i as a transient (resp., closer) in the ξ_i th active block of T_{i-1} .

Consider the walk diagram $h = (\omega, \xi) \in \Omega_{10}^5$, where ω is the walk in Figure 3 and $\xi = (1, 2, 1, 2, 1, 1, 1, 2, 4, 1)$; then the step-by-step construction of $\psi(h)$ goes as follows:

i	ω_i	step_i	ξ_i	T_i
0	(0, 0)			$\emptyset$
1	(0, 1)	North	1	$\{1, \dots\}$
2	(0, 2)	North	2	$\{1, \dots\} - \{2, \dots\}$
3	(0, 3)	North	1	$\{3, \dots\} - \{1, \dots\} - \{2, \dots\}$
4	(0, 3)	Null	2	$\{3, \dots\} - \{1, 4, \dots\} - \{2, \dots\}$
5	(0, 3)	Null	1	$\{3, 5, \dots\} - \{1, 4, \dots\} - \{2, \dots\}$
6	(1, 3)	East	1	$\{6\} - \{3, 5, \dots\} - \{1, 4, \dots\} - \{2, \dots\}$
7	(2, 2)	Southeast	1	$\{6\} - \{3, 5, 7\} - \{1, 4, \dots\} - \{2, \dots\}$
8	(3, 1)	Southeast	2	$\{6\} - \{3, 5, 7\} - \{1, 4, \dots\} - \{2, 8\}$
9	(4, 1)	East	4	$\{6\} - \{3, 5, 7\} - \{1, 4, \dots\} - \{9\} - \{2, 8\}$
10	(5, 0)	Southeast	1	$\{6\} - \{3, 5, 7\} - \{1, 4, 10\} - \{9\} - \{2, 8\}$

Thus $\psi(h) = 6/3 \ 5 \ 7/1 \ 4 \ 10/9/2 \ 8$.

To show that ψ is bijective, we construct its inverse ϕ . Given an ordered partition π in $\mathcal{OP}_n^k$ we denote by $\text{act}_i \pi$ and $\text{com}_i \pi$ the numbers of active and complete blocks in $T_i(\pi)$, respectively. Let $\omega_0(\pi) = (0, 0)$ and $\omega_i(\pi) = (\text{com}_i \pi, \text{act}_i \pi)$ for $1 \leq i \leq n$. If i is a singleton or opener, let ξ_i be the index (from left to right) of the block containing i in $T_i(\pi)$. If i is a transient or closer, let ξ_i be the index (from left to right) of the active block which will contain i in $T_i(\pi)$ among the active blocks in the trace $T_{i-1}(\pi)$. Let $\omega = (\omega_i(\pi))_{0 \leq i \leq n}$ and $\xi = (\xi_i)_{1 \leq i \leq n}$. It is readily seen that (ω, ξ) is a walk diagram in Ω_n^k and ϕ is the inverse of ψ .

It remains to check the other properties.

- (a) If the i th step of ω is North (resp., East), by construction, the element i will be an opener (resp., singleton) of the partitions T_j for $j \geq i$. Since $(\text{lcs} + \text{rcs})_i(\pi)$ (resp., $(\text{lsb} + \text{rsb})_i(\pi)$) is equal to the number of complete (resp., active) blocks in $T_{i-1}(\pi)$, we have

$$(\text{lcs} + \text{rcs})_i(\pi) = p_{i-1} \quad \text{and} \quad (\text{lsb} + \text{rsb})_i(\pi) = q_{i-1}.$$

Since $\text{los}_i \pi$ (resp., $\text{ros}_i \pi$) is the number of blocks in $T_i(\pi)$ on the left (resp., right) of the block containing i , we get $\text{los}_i(\pi) = \xi_i - 1$ and $\text{ros}_i(\pi) = p_{i-1} + q_{i-1} + 1 - \xi_i$.

- (b) If the i th step of ω is Southeast (resp., Null), by construction, the element i is a closer (resp., transient) of the ordered partition T_j for $j \geq i$. As in (a), we have $(\text{lcs} + \text{rcs})_i(\pi) = p_{i-1}$. Note that $\text{lsb}_i(\pi)$ (resp., $\text{rsb}_i(\pi)$) is equal to the number of active blocks in $T_i(\pi)$ on the left (resp., on the right) of the block containing i ; as we insert i in the ξ_i th active block in $T_{i-1}(\pi)$ among q_{i-1} active blocks, we have $\text{lsb}_i(\pi) = \xi_i - 1$ and $\text{rsb}_i(\pi) = q_{i-1} - \xi_i$. Finally we have $(\text{lsb} + \text{rsb})_i(\pi) = (\xi_i - 1) + (q_{i-1} - \xi_i) = q_{i-1} - 1$. $\square$

4.2. Generating functions of walks. Given a walk ω of finite length in D_k , define the valuation of a step (ω_i, ω_{i+1}) of abscissa p and height q by

$$(4.1) \quad v(\omega_i, \omega_{i+1}) = \begin{cases} t_1^p t_7^q [p + q + 1]_{t_5, t_6} & \text{if the step is North or East;} \\ t_2^p [q]_{t_3, t_4} & \text{if the step is Null or Southeast.} \end{cases}$$

The valuation $v(\omega)$ of ω is the product of the weights of all its steps.

Let $\mathbf{t} = (t_1, t_2, t_3, t_4, t_5, t_6, t_7)$. For $k \geq 0$, define

$$(4.2) \quad Q_k(a; \mathbf{t}) := \sum_{w \in \Omega^k} v(w) a^{|w|},$$

where $|w|$ denotes the length of the walk w .

By Theorem 4.4 one can rewrite the generating function $Q_k(a; \mathbf{t})$ of ordered partitions as

$$(4.3) \quad Q_k(a; \mathbf{t}) = \sum_{\pi \in \mathcal{OP}^k} t_1^{(\text{lcs} + \text{rcs})(\mathcal{O} \cup \mathcal{S})\pi} t_2^{(\text{lcs} + \text{rcs})(\mathcal{T} \cup \mathcal{C})\pi} t_3^{\text{rsb}(\mathcal{T} \cup \mathcal{C})\pi} \\ \times t_4^{\text{lsb}(\mathcal{T} \cup \mathcal{C})\pi} t_5^{\text{ros}(\mathcal{O} \cup \mathcal{S})\pi} t_6^{\text{los}(\mathcal{O} \cup \mathcal{S})\pi} t_7^{(\text{lsb} + \text{rsb})(\mathcal{O} \cup \mathcal{S})\pi} a^{|\pi|}.$$

It is obvious that the number of vertices of D_k is equal to

$$\widehat{k} := 1 + 2 + \cdots + (k + 1) = \frac{(k + 1)(k + 2)}{2}.$$

We label the vertices of D_k by their ranks in the following total ordering: $(i, j) \leq (i', j')$ if and only if $i + j < i' + j'$ or $(i + j = i' + j'$ and $j \geq j')$. Thus $v_1 = (0, 0)$, $v_2 = (0, 1)$, $v_3 = (1, 0)$, $v_4 = (0, 2)$, $v_5 = (1, 1)$, $v_6 = (2, 0), \dots, v_{\widehat{k}} = (k, 0)$.

The *adjacency matrix* of D_k relative to the valuation $\mathbf{v}$ is the $\widehat{k} \times \widehat{k}$ matrix A_k defined by

$$A_k(i, j) = \begin{cases} \mathbf{v}(v_i, v_j) & \text{if } (v_i, v_j) \text{ is an edge of } D_k; \\ 0 & \text{otherwise.} \end{cases}$$

Applying the transfer-matrix method (see, e.g., [17, Theorem 4.7.2]), we derive

$$(4.4) \quad Q_k(a; \mathbf{t}) = \frac{(-1)^{1+\widehat{k}} \det(I_k - aA_k; \widehat{k}, 1)}{\det(I_k - aA_k)},$$

where $(B; i, j)$ denotes the matrix obtained by removing the i th row and j th column of B and I_k is the $\widehat{k} \times \widehat{k}$ identity matrix.

For instance, when $k = 2$, we have

$$A_2 = \left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & t_7 [2]_{t_5, t_6} & t_7 [2]_{t_5, t_6} & 0 \\ 0 & 0 & 0 & 0 & t_1 [2]_{t_5, t_6} & t_1 [2]_{t_5, t_6} \\ \hline 0 & 0 & 0 & [2]_{t_3, t_4} & [2]_{t_3, t_4} & 0 \\ 0 & 0 & 0 & 0 & t_2 & t_2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

and

$$Q_2(a; \mathbf{t}) = -\frac{\det(I_2 - aA_2; 6, 1)}{\det(I_2 - aA_2)} = \frac{a^2 [2]_{t_5, t_6} (at_2 t_7 + t_1 (1 - a[2]_{t_3, t_4}))}{(1 - a)(1 - a[2]_{t_3, t_4})(1 - at_2)}.$$

In order to prove Theorem 2.9 it is sufficient to evaluate the following special cases of $Q_k(a; \mathbf{t})$:

$$(4.5) \quad f_k(a; x, y, t, u) = Q_k(a; x, x, x, y, t, u, y),$$

$$(4.6) \quad g_k(a; x, y, t, u) = Q_k(a; 1, x, 1, xy, t, u, y).$$

Invoking (4.3) and remarking that $\text{Inv} = \text{ros}(\mathcal{O} \cup \mathcal{S})$ and $\text{cInv} = \text{los}(\mathcal{O} \cup \mathcal{S})$, we have the combinatorial interpretations

$$f_k(a; x, y, t, u) = \sum_{\pi \in \mathcal{OP}^k} x^{(\text{lcs} + \text{rcs})\pi + \text{rsb}(\mathcal{T} \cup \mathcal{C})} \pi_y^{((\text{lsb} + \text{rsb})(\mathcal{O} \cup \mathcal{S}) + \text{lsb}(\mathcal{T} \cup \mathcal{C}))} \pi_t^{\text{Inv}} \pi_u^{\text{cInv}} \pi_a^{|\pi|},$$

$$g_k(a; x, y, t, u) = \sum_{\pi \in \mathcal{OP}^k} x^{(\text{lcs} + \text{rcs} + \text{lsb})(\mathcal{T} \cup \mathcal{C})} \pi_y^{((\text{lsb} + \text{rsb})(\mathcal{O} \cup \mathcal{S}) + \text{lsb}(\mathcal{T} \cup \mathcal{C}))} \pi_t^{\text{Inv}} \pi_u^{\text{cInv}} \pi_a^{|\pi|}.$$

In the next section, we shall prove the following result.

THEOREM 4.5. *For $k \geq 1$, we have*

$$(4.7) \quad f_k(a; x, y, t, u) = \frac{a^k x^{\binom{k}{2}} [k]_{t, u}!}{\prod_{i=1}^k (1 - a[i]_{x, y})},$$

$$(4.8) \quad g_k(a; x, y, t, u) = \frac{a^k [k]_{t, u}!}{\prod_{i=1}^k (1 - ax^{k-i} [i]_{xy})}.$$

The f_k 's and g_k 's are closely related to the ϕ_k 's and φ_k 's, respectively. To explain a relation between these generating functions, we need the following lemma.

LEMMA 4.6. *The following functional identities hold on $\mathcal{OP}_n^k$:*

$$\begin{aligned} \text{mak} + \text{bInv} &= (\text{lcs} + \text{rcs}) + \text{rsb}(\mathcal{T} \cup \mathcal{C}) + \text{Inv}, \\ \text{lmak} + \text{bInv} &= n(k-1) - (\text{lcs} + \text{rcs})(\mathcal{T} \cup \mathcal{C}) - \text{lsb}(\mathcal{T} \cup \mathcal{C}) - \text{cInv}, \\ \text{cinvLSB} &= (\text{lsb} + \text{rsb})(\mathcal{O} \cup \mathcal{S}) + \text{lsb}(\mathcal{T} \cup \mathcal{C}) + \text{Inv} + 2 \text{cInv}. \end{aligned}$$

Proof. Since $\text{bInv} = \text{rcs}(\mathcal{O} \cup \mathcal{S})$ and $\text{Inv} = \text{ros}(\mathcal{O} \cup \mathcal{S})$ we can rewrite

$$\begin{aligned} \text{mak} + \text{bInv} &= \text{lcs} + \text{ros} + \text{rcs}(\mathcal{O} \cup \mathcal{S}) \\ &= (\text{lcs} + \text{rcs}) + \text{ros} + (\text{rcs}(\mathcal{O} \cup \mathcal{S}) - \text{rcs}) \\ &= (\text{lcs} + \text{rcs}) + \text{ros} - \text{rcs}(\mathcal{T} \cup \mathcal{C}) \\ &= (\text{lcs} + \text{rcs}) + (\text{ros} - \text{rcs})(\mathcal{T} \cup \mathcal{C}) + \text{ros}(\mathcal{O} \cup \mathcal{S}) \\ &= (\text{lcs} + \text{rcs}) + \text{rsb}(\mathcal{T} \cup \mathcal{C}) + \text{Inv} \end{aligned}$$

and

$$\begin{aligned} \text{cinvLSB} &= k(k-1) + \text{lsb} - \text{bInv} \\ &= 2(\text{Inv} + \text{cInv}) + \text{lsb} - \text{rcs}(\mathcal{O} \cup \mathcal{S}) \\ &= 2(\text{Inv} + \text{cInv}) + (\text{lsb} + \text{rsb})(\mathcal{O} \cup \mathcal{S}) \\ &\quad - (\text{rcs} + \text{rsb})(\mathcal{O} \cup \mathcal{S}) + \text{lsb}(\mathcal{T} \cup \mathcal{C}) \\ &= 2(\text{Inv} + \text{cInv}) + (\text{lsb} + \text{rsb})(\mathcal{O} \cup \mathcal{S}) - \text{ros}(\mathcal{O} \cup \mathcal{S}) + \text{lsb}(\mathcal{T} \cup \mathcal{C}) \\ &= (\text{lsb} + \text{rsb})(\mathcal{O} \cup \mathcal{S}) + \text{lsb}(\mathcal{T} \cup \mathcal{C}) + \text{Inv} + 2 \text{cInv}. \end{aligned}$$

Invoking (2.7) we have

$$\begin{aligned} n(k-1) - (\text{lmak} + \text{bInv}) &= (\text{los} + \text{rcs}) - \text{rcs}(\mathcal{O} \cup \mathcal{S}) \\ &= \text{los} + \text{rcs}(\mathcal{T} \cup \mathcal{C}) \\ &= \text{los}(\mathcal{O} \cup \mathcal{S}) + \text{rcs}(\mathcal{T} \cup \mathcal{C}) + \text{los}(\mathcal{T} \cup \mathcal{C}) \\ &= \text{cInv} + (\text{lcs} + \text{rcs})(\mathcal{T} \cup \mathcal{C}) + (\text{los} - \text{lcs})(\mathcal{T} \cup \mathcal{C}) \\ &= \text{cInv} + (\text{lcs} + \text{rcs})(\mathcal{T} \cup \mathcal{C}) + \text{lsb}(\mathcal{T} \cup \mathcal{C}). \end{aligned}$$

This completes the proof of Lemma 4.6. $\square$

The following lemma is an immediate consequence of Lemma 4.6.

LEMMA 4.7. *For $k \geq 1$, we have*

$$(4.9) \quad \phi_k(a; x, y, t, u) = f_k(a; x, y, txy, uy^2),$$

$$(4.10) \quad \varphi_k(a; x, y, t, u) = g_k(ax^{k-1}; 1/x, y, ty, uy^2/x).$$

Theorem 2.9 follows immediately from Theorem 4.5 and Lemma 4.7.

5. Determinantal computations. Let A'_k and A''_k be the matrices obtained from A_k by making the substitutions corresponding to (4.5) and (4.6), respectively. Now, for each $k \geq 0$, let

$$M_k = I_k - aA'_k \quad \text{and} \quad N_k = I_k - aA''_k.$$

Then we derive from (4.4), (4.5), and (4.6) that for each $k \geq 1$,

$$(5.1) \quad f_k(a; x, y, t, u) = \frac{(-1)^{1+\widehat{k}} \det(M_k; \widehat{k}, 1)}{\det M_k},$$

$$(5.2) \quad g_k(a; x, y, t, u) = \frac{(-1)^{1+\widehat{k}} \det(N_k; \widehat{k}, 1)}{\det N_k}.$$

It is not difficult to see that M_n and N_n are upper triangular matrices and to obtain then that for each $n \geq 1$

$$(5.3) \quad \det M_n = \prod_{m=1}^n \prod_{i=0}^m (1 - ax^i[m-i]_{x,y}),$$

$$(5.4) \quad \det N_n = \prod_{m=1}^n \prod_{k=0}^{n-m} (1 - ax^k[m]_{xy}).$$

The evaluation of $\det(M_n; \widehat{n}, 1)$ and $\det(N_n; \widehat{n}, 1)$ is highly nontrivial.

THEOREM 5.1. *Let $n \geq 1$ be a positive integer. Then*

$$(5.5) \quad \det(M_n; \widehat{n}, 1) = (-1)^{\binom{n}{2}} a^n x^{\binom{n}{2}} [n]_{t,u}! \prod_{m=1}^{n-1} \prod_{i=1}^m (1 - ax^i[m-i+1]_{x,y}).$$

THEOREM 5.2. *Let $n \geq 1$ be a positive integer. Then*

$$(5.6) \quad \det(N_n; \widehat{n}, 1) = (-1)^{\binom{n}{2}} a^n [n]_{t,u}! \prod_{m=1}^{n-1} \prod_{k=1}^{n-m} (1 - ax^{k-1}[m]_{xy}).$$

Note that in view of (5.1) and (5.2), the above results complete the proof of Theorem 4.5.

5.1. Proof of Theorem 5.1. By definition of the matrix M_n we have

$$M_1 = \left(\begin{array}{c|cc} 1 & -a & -a \\ \hline 0 & 1-a & -a \\ \hline 0 & 0 & 1 \end{array} \right)$$

and

$$M_2 = \left(\begin{array}{ccc|ccc} 1 & -a & -a & 0 & 0 & 0 \\ 0 & 1-a & -a & -ay(t+u) & -ay(t+u) & 0 \\ 0 & 0 & 1 & 0 & -ax(t+u) & -ax(t+u) \\ \hline 0 & 0 & 0 & 1-a(x+y) & -a(x+y) & 0 \\ 0 & 0 & 0 & 0 & 1-ax & -ax \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right).$$

Clearly the matrix M_n is of order $\widehat{n}$ and can be defined recursively by

$$(5.7) \quad M_0 = (1), \quad M_n = \left(\begin{array}{c|c} M_{n-1} & \overline{M}_{n-1} \\ \hline O_{n+1, \widehat{n-1}} & \widehat{M}_{n-1} \end{array} \right),$$

where $n \geq 1$,

$$(5.8) \quad \widehat{M}_{n-1} = (\delta_{ij} - ax^{i-1}[n+1-i]_{x,y}(\delta_{ij} + \delta_{i+1,j}))_{1 \leq i,j \leq n+1}$$

and $\overline{M}_{n-1}$ is the $\widehat{n-1} \times (n+1)$ matrix

$$\overline{M}_{n-1} = \left(\frac{O_{\widehat{n-2}, n+1}}{\check{M}_{n-1}} \right)$$

with the $n \times (n+1)$ matrix

$$(5.9) \quad \check{M}_{n-1} = (-ax^{i-1}y^{n-i}[n]_{t,u}(\delta_{ij} + \delta_{i+1,j}))_{1 \leq i \leq n, 1 \leq j \leq n+1}.$$

Here δ_{ij} stands for the Kronecker delta and $O_{m,n}$ denotes the $m \times n$ zero matrix.

Let

$$K_n = \widehat{n} - 1 = \frac{n(n+3)}{2},$$

and let $P_n = (M_n; \widehat{n}, 1)$, i.e., the $K_n \times K_n$ matrix obtained from M_n by deleting the $\widehat{n}$ th row and the first column. For instance, we have

$$P_1 = \begin{pmatrix} -a & -a \\ 1-a & -a \end{pmatrix}$$

and

$$P_2 = \left(\begin{array}{cc|ccc} -a & -a & 0 & 0 & 0 \\ 1-a & -a & -ay(t+u) & -ay(t+u) & 0 \\ \hline 0 & 1 & 0 & -ax(t+u) & -ax(t+u) \\ 0 & 0 & 1-a(x+y) & -a(x+y) & 0 \\ 0 & 0 & 0 & 1-ax & -ax \end{array} \right).$$

In general we can define P_n as follows:

$$P_n = \left(\frac{P_{n-1} \mid \overline{P}_{n-1}}{X_{n-1} \mid \widehat{P}_{n-1}} \right),$$

where $\overline{P}_{n-1}$ is a $K_{n-1} \times (n+1)$ matrix, X_{n-1} is a $(n+1) \times K_{n-1}$ matrix, and $\widehat{P}_{n-1}$ is an $(n+1) \times (n+1)$ matrix. We shall compute $\det P_n$ by the following well-known formula for any block matrix with an invertible square matrix A :

$$\det \left(\frac{A \mid B}{C \mid D} \right) = \det A \cdot \det (D - CA^{-1}B).$$

Since the entries of $CA^{-1}B$ are also written by minors, we guess these entries and prove it by induction (see Theorem 5.3). Thus, looking at P_2 as the block matrix composed of P_1 , X_1 , $\overline{P}_1$, and $\widehat{P}_1$, we have

$$\overline{P}_1 = \begin{pmatrix} 0 & 0 & 0 \\ -ay(t+u) & -ay(t+u) & 0 \end{pmatrix}$$

and

$$\hat{P}_1 = \begin{pmatrix} 0 & -ax(t+u) & -ax(t+u) \\ 1-a(x+y) & -a(x+y) & 0 \\ 0 & 1-ax & -ax \end{pmatrix}.$$

Since $\bar{P}_n$ is a $K_n \times (n+2)$ matrix, we can write

$$\bar{P}_n = \left(\frac{O_{K_{n-1}, n+2}}{U_n} \right),$$

where U_n is the $(n+1) \times (n+2)$ matrix composed of the last $(n+1)$ rows of $\bar{P}_n$. For $1 \leq k \leq n+2$, let

$$P_n^k = \left(\begin{array}{c|c} P_{n-1} & \bar{P}_{n-1} \\ \hline X_{n-1} & \hat{P}_{n-1}^k \end{array} \right)$$

denote the $K_n \times K_n$ matrix obtained from P_n by replacing the rightmost column with the k th column of $\bar{P}_n$. Here $\hat{P}_{n-1}^k$ is the $(n+1) \times (n+1)$ matrix obtained from $\hat{P}_{n-1}$ by replacing the rightmost column with the k th column of U_n . For example,

$$P_2^2 = \left(\begin{array}{cc|ccc} -a & -a & 0 & 0 & 0 \\ 1-a & -a & -ay(t+u) & -ay(t+u) & 0 \\ \hline 0 & 1 & 0 & -ax(t+u) & 0 \\ 0 & 0 & 1-a(x+y) & -a(x+y) & -ay^2(t^2+tu+t^2) \\ 0 & 0 & 0 & 1-ax & -axy(t^2+tu+t^2) \end{array} \right).$$

THEOREM 5.3. *Let $n \geq 1$ be a positive integer. Then we have*

$$(5.10) \quad \frac{\det P_n}{\det P_{n-1}} = (-1)^{n-1} ax^{n-1} [n]_{t,u} \prod_{i=1}^{n-1} (1 - ax^i [n-i]_{x,y})$$

and

$$(5.11) \quad \frac{\det P_n^k}{\det P_n} = ax^{\frac{(k-1)(k-2)}{2} - \frac{n(n-1)}{2}} y^{\frac{(n+1-k)(n+2-k)}{2}} [n+1]_{t,u} \begin{bmatrix} n+1 \\ k-1 \end{bmatrix}_{x,y}$$

for $1 \leq k \leq n$,

$$(5.12) \quad \frac{\det P_n^{n+1}}{\det P_n} = ay [n+1]_{t,u} [n]_{x,y},$$

and $\det P_n^{n+2} = 0$.

We first prove an elementary identity, which is obvious if it is written in a suitable manner.

LEMMA 5.4. *For $0 \leq m \leq n$,*

$$(5.13) \quad \begin{aligned} & \sum_{k=0}^m (-1)^{m-k} x^{\binom{k}{2}} y^{\binom{n-k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_{x,y} \prod_{i=0}^{k-1} \{1 - ax^i [n-i]_{x,y}\} \prod_{i=k}^{m-1} \{-ax^i [n-i]_{x,y}\} \\ &= x^{\binom{m}{2}} y^{\binom{n-m}{2}} \begin{bmatrix} n \\ m \end{bmatrix}_{x,y} \prod_{i=1}^m \{1 - ax^i [n-i]_{x,y}\}. \end{aligned}$$

Proof. Note that

$$[n]_{x,y} = y^{n-1} [n]_{x/y}, \quad \begin{bmatrix} n \\ k \end{bmatrix}_{x,y} = y^{kn-k^2} \begin{bmatrix} n \\ k \end{bmatrix}_{x/y}.$$

Since $\binom{n-k}{2} + \binom{k}{2} + kn - k^2 = \binom{n}{2}$, setting $c = ay^{n-1}$ and $q = x/y$, we can rewrite (5.13) as follows:

$$(5.14) \quad \begin{aligned} & \sum_{k=0}^m (-1)^{m-k} q^{\binom{k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q \prod_{i=0}^{k-1} \{1 - cq^i [n-i]_q\} \prod_{i=k}^{m-1} \{-cq^i [n-i]_q\} \\ &= q^{\binom{m}{2}} \begin{bmatrix} n \\ m \end{bmatrix}_q \prod_{i=1}^m \{1 - cq^i [n-i]_q\}. \end{aligned}$$

Setting

$$X = 1 + \frac{cq^n}{1-q}, \quad Y = \frac{c}{1-q+cq^n}, \quad Z = \frac{cq^n}{1-q},$$

then $1 - cq^i [n-i]_q = X(1 - Yq^i)$ and $-cq^i [n-i]_q = Z(1 - q^{i-n})$. Hence, in (5.14) making the substitutions

$$\begin{aligned} & \prod_{i=0}^{k-1} \{1 - cq^i [n-i]_q\} = X^k (Y; q)_k, \\ & \prod_{i=1}^m \{1 - cq^i [n-i]_q\} = X^m (Yq; q)_m, \\ & \prod_{i=k}^{m-1} \{-cq^i [n-i]_q\} = (-1)^m Z^{m-k} q^{\binom{m}{2} - mn} \frac{(q^{n-m+1}; q)_m}{(q^{-n}; q)_k} \end{aligned}$$

and writing the q -binomial coefficients as

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = (-1)^k q^{kn-k(k-1)/2} \frac{(q^{-n}; q)_k}{(q; q)_k}, \quad \begin{bmatrix} n \\ m \end{bmatrix}_q = (-1)^m q^{mn-m(m-1)/2} \frac{(q^{-n}; q)_m}{(q; q)_m},$$

we see, after simplifying, that identity (5.14) is equivalent to the special case $Y = c/(1 - q + cq^n)$ of the identity,

$$\sum_{k=0}^m \frac{(Y; q)_k}{(q; q)_k} Y^{m-k} = \frac{(Yq; q)_m}{(q; q)_m},$$

which can be easily verified by induction. $\square$

Proof of Theorem 5.3. We proceed by induction on n . When $n = 1$, by a direct computation we obtain $\det P_1 = a$, $\det P_1^1 = \det P_1^2 = a^2 y [2]_{t,u}$, and $\det P_1^3 = 0$. This shows the theorem is true when $n = 1$. Let n be an integer ≥ 2 . Assume the theorem is true for $n - 1$.

(i) We get

$$\det P_n = \det \left(\begin{array}{c|c} P_{n-1} & \overline{P}_{n-1} \\ \hline X_{n-1} & \widehat{P}_{n-1} \end{array} \right) = \det P_{n-1} \cdot \det \left(\widehat{P}_{n-1} - X_{n-1} P_{n-1}^{-1} \overline{P}_{n-1} \right)$$

and

$$\det P_n^k = \det \left(\begin{array}{c|c} P_{n-1} & \bar{P}_{n-1} \\ \hline X_{n-1} & \hat{P}_{n-1}^k \end{array} \right) = \det P_{n-1} \cdot \det \left(\hat{P}_{n-1}^k - X_{n-1} P_{n-1}^{-1} \bar{P}_{n-1} \right).$$

- (ii) By direct computation we can see that the (i, j) th entry of $X_{n-1} P_{n-1}^{-1} \bar{P}_{n-1}$ ($1 \leq i, j \leq n+1$) is equal to

$$\begin{cases} \frac{\det P_{n-1}^j}{\det P_{n-1}} & \text{if } i = 1, \\ 0 & \text{otherwise.} \end{cases}$$

By the induction hypothesis, the $(1, j)$ th entry of $X_{n-1} P_{n-1}^{-1} \bar{P}_{n-1}$ equals

$$(5.15) \quad \begin{cases} a x^{\binom{j-1}{2} - \binom{n-1}{2}} y^{\binom{n+1-j}{2}} [n]_{t,u} \begin{bmatrix} n \\ j-1 \end{bmatrix}_{x,y} & \text{if } 1 \leq j \leq n-1, \\ a y [n-1]_{x,y} [n]_{t,u} & \text{if } j = n, \\ 0 & \text{if } j = n+1. \end{cases}$$

- (iii) Put $W_{n-1}^k = \hat{P}_{n-1}^k - X_{n-1} P_{n-1}^{-1} \bar{P}_{n-1}$ and $W_{n-1} = \hat{P}_{n-1} - X_{n-1} P_{n-1}^{-1} \bar{P}_{n-1}$. Then, by (i), we have $\frac{\det P_n^k}{\det P_{n-1}} = \det W_{n-1}^k$ and $\frac{\det P_n}{\det P_{n-1}} = \det W_{n-1}$. By (5.9) and (5.15), we can see that the $(1, j)$ th entry of W_{n-1}^k is

$$-a x^{\binom{j-1}{2} - \binom{n-1}{2}} y^{\binom{n+1-j}{2}} [n]_{t,u} \begin{bmatrix} n \\ j-1 \end{bmatrix}_{x,y}$$

for $1 \leq j \leq n$, and the $(1, n+1)$ th entry is 0 (the top row does not depend on k). It is also easy to see that the $(1, j)$ th entry of W_{n-1} is

$$-a x^{\binom{j-1}{2} - \binom{n-1}{2}} y^{\binom{n+1-j}{2}} [n]_{t,u} \begin{bmatrix} n \\ j-1 \end{bmatrix}_{x,y}$$

for $1 \leq j \leq n+1$.

- (iv) We claim that

$$\det W_{n-1} = (-1)^{n-1} a x^{n-1} [n]_{t,u} \prod_{i=1}^{n-1} (1 - a x^i [n-i]_{x,y}).$$

In fact, the (i, j) th entry of W_{n-1} is

$$\begin{cases} -a y^{\frac{(n-j)(n-j+1)}{2}} x^{\frac{(j-1)(j-2)}{2} - \frac{(n-1)(n-2)}{2}} [n]_{t,u} \begin{bmatrix} n \\ j-1 \end{bmatrix}_{x,y} & \text{if } i = 1 \text{ and } 1 \leq j \leq n+1, \\ 1 - a x^{j-1} [n+1-j]_{x,y} & \text{if } i = j+1 \text{ and } 1 \leq j \leq n, \\ -a x^{j-2} [n+2-j]_{x,y} & \text{if } i = j \text{ and } 2 \leq j \leq n+1, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, if we expand $\det W_{n-1}$ along the top row, then we obtain

$$\begin{aligned} \det W_{n-1} &= -a x^{-\frac{(n-1)(n-2)}{2}} [n]_{t,u} \\ &\quad \times \sum_{j=1}^{n+1} (-1)^{j+1} y^{\frac{(n-j)(n-j+1)}{2}} x^{\frac{(j-1)(j-2)}{2}} \begin{bmatrix} n \\ j-1 \end{bmatrix}_{x,y} \det W_{n-1}(1; j). \end{aligned}$$

If we use

$$\det W_{n-1}(1; j) = \prod_{\nu=0}^{j-2} (1 - ax^\nu[n - \nu]_{x,y}) \prod_{\nu=j-1}^{n-1} (-ax^\nu[n - \nu]_{x,y}),$$

then we obtain

$$\begin{aligned} \det W_{n-1} &= -ax^{-\frac{(n-1)(n-2)}{2}} [n]_{t,u} \sum_{j=1}^{n+1} (-1)^{j+1} y^{\frac{(n-j)(n-j+1)}{2}} x^{\frac{(j-1)(j-2)}{2}} \begin{bmatrix} n \\ j-1 \end{bmatrix}_{x,y} \\ &\quad \times \prod_{\nu=0}^{j-2} (1 - ax^\nu[n - \nu]_{x,y}) \prod_{\nu=j-1}^{n-1} (-ax^\nu[n - \nu]_{x,y}) \\ &= (-1)^{n-1} ax^{n-1} [n]_{t,u} \prod_{i=1}^{n-1} (1 - ax^i[n - i]_{x,y}) \end{aligned}$$

by (5.13). Thus, by (i), we conclude that

$$(5.16) \quad \frac{\det P_n}{\det P_{n-1}} = \det W_{n-1} = (-1)^{n-1} ax^{n-1} [n]_{t,u} \prod_{i=1}^{n-1} (1 - ax^i[n - i]_{x,y}).$$

(v) We claim that

$$\frac{\det P_n^k}{\det P_n} = \frac{\det W_{n-1}^k}{\det W_{n-1}} = ax^{\frac{(k-1)(k-2)}{2} - \frac{n(n-1)}{2}} y^{\frac{(n+1-k)(n+2-k)}{2}} [n+1]_{t,u} \begin{bmatrix} n+1 \\ k-1 \end{bmatrix}_{x,y}$$

for $1 \leq k \leq n$. Because the rightmost column of $\hat{P}_{n-1}^k$ is the k th column of U_n , the $(i, n+1)$ th entry of $\hat{P}_{n-1}^k$ is

$$\begin{cases} -ay^n[n+1]_{t,u} & \text{if } i = 2, \\ 0 & \text{otherwise} \end{cases}$$

when $k = 1$,

$$\begin{cases} -ay^{n+2-k}x^{k-2}[n+1]_{t,u} & \text{if } i = k, \\ -ay^{n+1-k}x^{k-1}[n+1]_{t,u} & \text{if } i = k+1, \\ 0 & \text{otherwise} \end{cases}$$

when $2 \leq k \leq n$,

$$\begin{cases} -ayx^{n-1}[n+1]_{t,u} & \text{if } i = n+1, \\ 0 & \text{otherwise} \end{cases}$$

when $k = n+1$, and all zero when $k = n+2$. By the induction hypothesis, the $(1, n+1)$ th entry of $X_{n-1}P_{n-1}^{-1}\bar{P}_{n-1}$ is $\frac{\det P_{n-1}^{n+1}}{\det P_{n-1}} = 0$. Thus the $(n+1)$ th column of $W_{n-1}^k = \hat{P}_{n-1}^k - X_{n-1}P_{n-1}^{-1}\bar{P}_{n-1}$ equals the $(n+1)$ th column of $\hat{P}_{n-1}^k$.

- (a) When $k = 1$, we expand $\det W_{n-1}^1$ along the $(n+1)$ th column; then, by direct computation, we obtain

$$\det W_{n-1}^1 = (-1)^{n+3}(-ay^n[n+1]_{t,u}) \det W_{n-1}(2; n+1).$$

By expanding $\det W_{n-1}^1(2; n+1)$ along the top row we obtain

$$\det W_{n-1}(2; n+1) = \left(-ay^{\frac{n(n-1)}{2}} x^{-\frac{(n-1)(n-2)}{2}} [n]_{t,u}\right) \prod_{\nu=1}^{n-1} (1-ax^\nu[n-\nu]_{x,y}).$$

Thus we conclude that

$$\begin{aligned} \det W_{n-1}^1 &= (-1)^{n-1} a^2 y^{\frac{n(n+1)}{2}} x^{-\frac{(n-1)(n-2)}{2}} [n]_{t,u} \\ &\quad \times [n+1]_{t,u} \prod_{\nu=1}^{n-1} (1-ax^\nu[n-\nu]_{x,y}). \end{aligned} \quad (5.17)$$

By (5.16), this implies

$$\frac{\det W_{n-1}^1}{\det W_{n-1}} = ay^{\frac{n(n+1)}{2}} x^{-\frac{n(n-1)}{2}} [n+1]_{t,u},$$

which is the desired identity.

- (b) When $2 \leq k \leq n$, expanding $\det W_{n-1}^k$ along the $(n+1)$ th column, we obtain

$$\begin{aligned} \det W_{n-1}^k &= (-1)^{k+n+1}(-ay^{n+2-k}x^{k-2}[n+1]_{t,u}) \det W_{n-1}(k; n+1) \\ &\quad + (-1)^{k+n+2}(-ay^{n+1-k}x^{k-1}[n+1]_{t,u}) \det W_{n-1}(k+1; n+1). \end{aligned} \quad (5.18)$$

By expanding along the top row, we obtain

$$\begin{aligned} \det W_{n-1}(k; n+1) &= -ax^{-\frac{(n-1)(n-2)}{2}} [n]_{t,u} \sum_{j=1}^{k-1} (-1)^{j+1} y^{\frac{(n-j)(n-j+1)}{2}} x^{\frac{(j-1)(j-2)}{2}} \\ &\quad \times \begin{bmatrix} n \\ j-1 \end{bmatrix}_{x,y} \det W_{n-1}(1, k; j, n+1), \\ \det W_{n-1}(k+1; n+1) &= -ax^{-\frac{(n-1)(n-2)}{2}} [n]_{t,u} \sum_{j=1}^k (-1)^{j+1} y^{\frac{(n-j)(n-j+1)}{2}} x^{\frac{(j-1)(j-2)}{2}} \\ &\quad \times \begin{bmatrix} n \\ j-1 \end{bmatrix}_{x,y} \det W_{n-1}(1, k+1; j, n+1), \end{aligned}$$

where $W_{n-1}(1, k; j, n+1) = W_{n-1}(k; n+1)(1; j)$ and $W_{n-1}(1, k+1; j, n+1)$

$= W_{n-1}(k+1; n+1)(1; j)$. If we use

$$\begin{aligned} \det W_{n-1}(1, k; j, n+1) &= \prod_{\nu=0}^{j-2} (1 - ax^\nu[n-\nu]_{x,y}) \prod_{\nu=j-1}^{k-3} (-ax^\nu[n-\nu]_{x,y}) \\ &\quad \times \prod_{\nu=k-1}^{n-1} (1 - ax^\nu[n-\nu]_{x,y}), \\ \det W_{n-1}(1, k+1; j, n+1) &= \prod_{\nu=0}^{j-2} (1 - ax^\nu[n-\nu]_{x,y}) \prod_{\nu=j-1}^{k-2} (-ax^\nu[n-\nu]_{x,y}) \\ &\quad \times \prod_{\nu=k}^{n-1} (1 - ax^\nu[n-\nu]_{x,y}), \end{aligned}$$

then, by (5.13), we obtain

$$\begin{aligned} \det W_{n-1}(k; n+1) &= (-1)^{k-1} ax^{\frac{(k-2)(k-3)}{2} - \frac{(n-1)(n-2)}{2}} y^{\frac{(n-k+2)(n-k+1)}{2}} \\ &\quad \times [n]_{t,u} \left[\begin{matrix} n \\ k-2 \end{matrix} \right]_{x,y} \prod_{\nu=1}^{n-1} (1 - ax^\nu[n-\nu]_{x,y}), \\ \det W_{n-1}(k+1; n+1) &= (-1)^k ax^{\frac{(k-1)(k-2)}{2} - \frac{(n-1)(n-2)}{2}} y^{\frac{(n-k)(n-k+1)}{2}} \\ &\quad \times [n]_{t,u} \left[\begin{matrix} n \\ k-1 \end{matrix} \right]_{x,y} \prod_{\nu=1}^{n-1} (1 - ax^\nu[n-\nu]_{x,y}). \end{aligned}$$

Thus, from (5.18), we conclude that

$$\begin{aligned} \det W_{n-1}^k &= (-1)^{n-1} a^2 x^{-\frac{(n-1)(n-2)}{2} + \frac{(k-1)(k-2)}{2}} y^{\frac{(n+1-k)(n+2-k)}{2}} [n]_{t,u} [n+1]_{t,u} \\ &\quad \times \left(y^{n-k+2} \left[\begin{matrix} n \\ k-2 \end{matrix} \right]_{x,y} + x^{k-1} \left[\begin{matrix} n \\ k-1 \end{matrix} \right]_{x,y} \right) \prod_{\nu=1}^{n-1} (1 - ax^\nu[n-\nu]_{x,y}) \\ &= (-1)^{n-1} a^2 x^{-\frac{(n-1)(n-2)}{2} + \frac{(k-1)(k-2)}{2}} y^{\frac{(n+1-k)(n+2-k)}{2}} \\ &\quad \times [n]_{t,u} [n+1]_{t,u} \left[\begin{matrix} n+1 \\ k-1 \end{matrix} \right]_{x,y} \prod_{\nu=1}^{n-1} (1 - ax^\nu[n-\nu]_{x,y}). \end{aligned}$$

Using (5.16), we obtain

$$\frac{\det W_{n-1}^k}{\det W_{n-1}} = ax^{-\frac{n(n-1)}{2} + \frac{(k-1)(k-2)}{2}} y^{\frac{(n+1-k)(n+2-k)}{2}} [n+1]_{t,u} \left[\begin{matrix} n+1 \\ k-1 \end{matrix} \right]_{x,y},$$

which is the desired identity.

- (c) When $k = n+1$, we also expand $\det W_{n-1}^k$ along the $(n+1)$ th column and repeat the same argument. It is not hard to obtain

$$\frac{\det W_{n-1}^{n+1}}{\det W_{n-1}} = ay[n+1]_{t,u} [n]_{x,y}.$$

The details are left to the reader.

- (d) When $k = n+2$, $\det \widehat{P}_{n-1}^k$ vanishes since all the entries of the last column of $\det \widehat{P}_{n-1}^k$ are zero.

This proves the theorem is true for n . By induction we conclude that the theorem is true for all $n \geq 1$. This completes the proof. $\square$

Since $\det(P_1) = a$, Theorem 5.1 (5.5) follows easily from (5.10). $\square$

5.2. Proof of Theorem 5.2. We first introduce two generalizations of q -binomial coefficients. For any sequence of nonzero functions $F = \{F_n\}_{n=1}^\infty$ in finitely many variables $t_1, t_2, \dots$ we define $F_n! = \prod_{k=1}^n F_k$, and for any positive integers n and k ,

$$\begin{bmatrix} n \\ k \end{bmatrix}_F = \begin{cases} \frac{F_n!}{F_k! F_{n-k}!} & \text{if } 0 \leq k \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

Next, for any positive integers n and k , let

$$]n, k[_{x,y} = [n]_{xy} - x^{n-k} [k]_{xy}$$

and

$$]n[_{x,y} \begin{bmatrix} k \\ \end{bmatrix} = \begin{cases} \frac{\prod_{i=0}^{k-1}]n, i[_{x,y}}{[k]_{xy}!} & \text{if } 0 \leq k \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

For instance, we have $]3, 1[_{x,y} = 1 + xy + x^2y^2 - x^2$ and

$$]3[_{x,y} \begin{bmatrix} 2 \\ \end{bmatrix} = \frac{(1 + xy + x^2y^2)(1 + xy + x^2y^2 - x^2)}{(1 + xy)}.$$

Note that $]n, k[_{x,1} = [n - k]_x$ and $]n[_{x,1} \begin{bmatrix} k \\ \end{bmatrix} = \begin{bmatrix} n \\ k \end{bmatrix}_x$.

We prove Theorem 5.2 (5.6) by considering the following matrix $N_n(\lambda, a)$, which reduces to the matrix N_n by setting $\lambda = 1$ and $F_n = [n]_{t,u}$. Let $N_n(\lambda, a)$ be the matrix defined inductively as follows:

$$N_0(\lambda, a) = (\lambda)$$

and

$$(5.19) \quad N_n(\lambda, a) = \left(\begin{array}{c|c} N_{n-1}(\lambda, a) & \overline{N}_{n-1}(\lambda, a) \\ \hline O_{n+1, \widehat{n-1}} & \widehat{N}_{n-1}(\lambda, a) \end{array} \right),$$

where $\widehat{N}_{n-1}(\lambda, a)$ is the $(n+1) \times (n+1)$ matrix defined by

$$(5.20) \quad \widehat{N}_{n-1}(\lambda, a) = (\lambda \delta_{ij} - ax^{i-1} [n+1-i]_{xy} (\delta_{ij} + \delta_{i+1,j}))_{1 \leq i,j \leq n+1}$$

and $\overline{N}_{n-1}(\lambda, a)$ is the $\widehat{n-1} \times (n+1)$ matrix

$$\left(\begin{array}{c} O_{\widehat{n-2}, n+1} \\ \hline \check{N}_{n-1} \end{array} \right)$$

with the $n \times (n+1)$ matrix

$$(5.21) \quad \check{N}_{n-1} = (-ay^{n-i} F_n \cdot (\delta_{ij} + \delta_{i+1,j}))_{1 \leq i \leq n, 1 \leq j \leq n+1}.$$

For instance, we get

$$N_2(\lambda, a) = \begin{pmatrix} \lambda & -aF_1 & -aF_1 & 0 & 0 & 0 \\ 0 & \lambda - a & -a & -ayF_2 & -ayF_2 & 0 \\ 0 & 0 & \lambda & 0 & -aF_2 & -aF_2 \\ 0 & 0 & 0 & \lambda - a(1 + xy) & -a(1 + xy) & 0 \\ 0 & 0 & 0 & 0 & \lambda - ax & -ax \\ 0 & 0 & 0 & 0 & 0 & \lambda \end{pmatrix}.$$

Let $\dot{N}_n(\lambda, a)$ denote the matrix obtained from $N_n(\lambda, a)$ by deleting the $\hat{n}$ th row and the first column. Then the following theorem is sufficient to prove our result. Here our strategy is as follows. We regard $\det N_n(\lambda, a)$ as a polynomial in λ and find all linear factors. Finally we check the leading coefficient on both sides.

THEOREM 5.5. *We have*

$$(5.22) \quad \det \dot{N}_n(\lambda, a) = (-1)^{n(n-1)/2} a^n F_n! \lambda^n \prod_{m=1}^{n-1} \prod_{k=1}^{n-m} (\lambda - ax^{k-1}[m]_{xy}).$$

Then by setting $\lambda = 1$ and $F_n = [n]_{t,u}$ we obtain Theorem 5.2 (5.6). For instance, we have

$$\det \dot{N}_1(\lambda, a) = \det \begin{pmatrix} -aF_1 & -aF_1 \\ \lambda - a & -a \end{pmatrix} = a F_1 \lambda$$

and

$$\begin{aligned} \det \dot{N}_2(\lambda, a) &= \det \begin{pmatrix} -aF_1 & -aF_1 & 0 & 0 & 0 \\ \lambda - a & -a & -ayF_2 & -ayF_2 & 0 \\ 0 & \lambda & 0 & -aF_2 & -aF_2 \\ 0 & 0 & \lambda - a(1 + xy) & -a(1 + xy) & 0 \\ 0 & 0 & 0 & \lambda - ax & -ax \end{pmatrix} \\ &= -a^2 F_1 F_2 \lambda^2 (\lambda - a). \end{aligned}$$

Fix positive integers m and k . Define the row vectors $\mathbf{X}_n^{m,k}$ of degree $\hat{n}$ as follows: For $1 \leq i \leq n+1$ and $1 \leq j \leq i$, the $\left(\frac{i(i-1)}{2} + j\right)$ th entry of $\mathbf{X}_n^{m,k}$ is equal to

$$(5.23) \quad \begin{aligned} X_{i,j}^{m,k} &= (-1)^{i+m+k} x^{-(m+k-1)(i-m-k) + \binom{j-k}{2}} y^{\binom{i-m-k}{2}} \\ &\quad \times \frac{F_{i-m-k}!}{[i-m-k]_{xy}!} \left[\begin{matrix} i-1 \\ m+k-1 \end{matrix} \right]_F \frac{m}{m+k-j} \left[\begin{matrix} m \\ x, y \end{matrix} \right]. \end{aligned}$$

Here we use the convention that $F_n! = [n]_{xy}! = 1$ if $n \leq 0$. For example, if $n = 3$, $m = k = 1$, then

$$\mathbf{X}_3^{1,1} = \left(0, 1, 1, -\frac{F_2}{x}, -\frac{F_2}{x}, 0, \frac{F_2 F_3 y}{x^2 [2]_{xy}!}, \frac{F_2 F_3 y}{x^2 [2]_{xy}!}, 0, 0 \right).$$

LEMMA 5.6. Let n be a positive integer. Let m and k be positive integers such that $0 \leq m \leq n-1$ and $1 \leq k \leq n-m$. Then we have

$$(5.24) \quad \mathbf{X}_n^{m,k} N_n(\lambda, a) = (\lambda - ax^{k-1}[m]_{xy}) \mathbf{X}_n^{m,k}.$$

Before we proceed to the proof of the lemma, we give an example to illustrate (5.24). If $n = 2$ and $m = k = 1$, then we have

$$\begin{aligned} & \left(0, 1, 1, -\frac{F_2}{x}, -\frac{F_2}{x}, 0\right) \begin{pmatrix} \lambda & -F_1 & -F_1 & 0 & 0 & 0 \\ 0 & \lambda - a & -a & -ayF_2 & -ayF_2 & 0 \\ 0 & 0 & \lambda & 0 & -aF_2 & -aF_2 \\ 0 & 0 & 0 & \lambda - a(1+xy) & -a(1+xy) & 0 \\ 0 & 0 & 0 & 0 & \lambda - ax & -ax \\ 0 & 0 & 0 & 0 & 0 & \lambda \end{pmatrix} \\ &= \left(0, \lambda - a, \lambda - a, -\frac{F_2}{x}(\lambda - a), -\frac{F_2}{x}(\lambda - a), 0\right). \end{aligned}$$

Proof of Lemma 5.6. We proceed by induction on n . When $n = 0$ or $n = 1$, our claim is easy to check by direct computation. Assume (5.24) is true up to $n-1$. Then the first $\widehat{n-1}$ entries of $\mathbf{X}_n^{m,k} N_n(\lambda, a)$ agree with those of $(\lambda - x^{k-1}[m]_{xy}) \mathbf{X}_n^{m,k}$ by the induction hypothesis. So we have to check the last $n+1$ entries. In fact we verify the following three cases.

- (i) If $i = n+1$ and $j = 1$, then the $(\frac{n(n+1)}{2} + 1)$ th entry of $\mathbf{X}_n^{m,k} N_n(\lambda, a)$ is equal to

$$(-ay^{n-1} F_n) X_{n,1}^{m,k} + (\lambda - a[n]_{xy}) X_{n+1,1}^{m,k}.$$

Note that the coefficient $\left]_{m+k-1}^m \left[\begin{smallmatrix} n \\ x, y \end{smallmatrix} \right] \right]$ becomes zero by definition unless $k = 1$. Thus, by direct computation, one can easily check that this sum equals

$$\begin{aligned} & (-1)^{n+m+k+1} x^{-(m+k-1)(n-m-k+1) + \binom{1-k}{2}} y^{\binom{n-m-k+1}{2}} \\ & \times \frac{F_{n-m-k+1}!}{[n-m-k+1]_{xy}!} \left[\begin{smallmatrix} n \\ m+k-1 \end{smallmatrix} \right]_F \left]_{m+k-1}^m \left[\begin{smallmatrix} n \\ x, y \end{smallmatrix} \right] \right] (\lambda - ax^{k-1}[m]_{xy}). \end{aligned}$$

- (ii) If $i = n+1$ and $2 \leq j \leq n$, then the $(\frac{n(n+1)}{2} + j)$ th entry of $\mathbf{X}_n^{m,k} N_n(\lambda, a)$ is equal to

$$\begin{aligned} & (-ay^{n-j+1} F_n) X_{n,j-1}^{m,k} + (-ay^{n-j} F_n) X_{n,j}^{m,k} \\ & + (-ax^{j-2}[n-j+2]_{xy}) X_{n+1,j-1}^{m,k} + (\lambda - ax^{j-1}[n-j+1]_{xy}) X_{n+1,j}^{m,k}. \end{aligned}$$

By direct computation, one can easily check that this equals

$$\begin{aligned} & (-1)^{n+m+k+1} x^{-(m+k-1)(n-m-k+1) + \binom{j-k}{2}} y^{\binom{n-m-k+1}{2}} \\ & \times \frac{F_{n-m-k+1}!}{[n-m-k+1]_{xy}!} \left[\begin{smallmatrix} n \\ m+k-1 \end{smallmatrix} \right]_F \left]_{m+k-j}^m \left[\begin{smallmatrix} n \\ x, y \end{smallmatrix} \right] \right] (\lambda - ax^{k-1}[m]_{xy}). \end{aligned}$$

- (iii) If $i = j = n+1$, then the $\frac{(n+1)(n+2)}{2}$ th entry of $\mathbf{X}_n^{m,k} N_n(\lambda, a)$ is equal to

$$(-aF_n) X_{n,n}^{m,k} + (-ax^{n-1}) X_{n+1,n}^{m,k} + \lambda X_{n+1,n+1}^{m,k}.$$

One can easily check that this is always equal to zero.

Thus this completes the proof of our lemma. $\square$

COROLLARY 5.7. *Let n be a positive integer. Then there exists a polynomial $\varphi(\lambda)$ such that*

$$\det \dot{N}_n(\lambda, a) = \varphi(t)\lambda^n \prod_{m=1}^{n-1} \prod_{k=1}^{n-m} (\lambda - ax^{k-1}[m]_{xy}).$$

Proof. Let $\dot{\mathbf{X}}_n^{m,k}$ (resp., $\ddot{\mathbf{X}}_n^{m,k}$) denote the vector of degree $\hat{n} - 1$ obtained from $\mathbf{X}_n^{m,k}$ by deleting the last (resp., first) entry. Note that we have shown, in the proof of Lemma 5.6, that the last entry of $\mathbf{X}_n^{m,k}$ is always zero. Thus, by (5.24), we obtain

$$\dot{\mathbf{X}}_n^{m,k} \dot{N}_n(\lambda, a) = (\lambda - ax^{k-1}[m]_{xy}) \ddot{\mathbf{X}}_n^{m,k}.$$

By substituting $\lambda = ax^{k-1}[m]_{xy}$ into this identity we obtain

$$(5.25) \quad \dot{\mathbf{X}}_n^{m,k} \dot{N}_n(ax^{k-1}[m]_{xy}, a) = \mathbf{0}$$

for $0 \leq m \leq n-1$ and $1 \leq k \leq n-m$. If $1 \leq m \leq n-1$ and $1 \leq k \leq n-m$, then $\dot{\mathbf{X}}_n^{m,k}$ is a nonzero vector so that (5.25) implies $\dot{N}_n(ax^{k-1}[m]_{xy}, a)$ is singular, i.e., $\det \dot{N}_n(ax^{k-1}[m]_{xy}, a) = 0$. Since $\det \dot{N}_n(\lambda, a)$ is a polynomial of λ , $\det \dot{N}_n(\lambda, a)$ is divisible by $\lambda - ax^{k-1}[m]_{xy}$. If $m = 0$, then $\dot{\mathbf{X}}_n^{0,k}$, $1 \leq k \leq n$, are evidently linearly independent and (5.25) implies $\det \dot{N}_n(\lambda, a)$ is divisible by λ^n . Summarizing we immediately obtain the corollary. $\square$

Now we are in position to complete the proof of Theorem 5.5.

Proof of Theorem 5.5. To complete the proof of Theorem 5.5, we need to show that the degree of $\det \dot{N}_n(\lambda, a)$ is $\frac{n(n+1)}{2}$ as a polynomial in λ , and the leading coefficient of $\det \dot{N}_n(\lambda, a)$ is equal to $(-1)^{n(n-1)/2} a^n F_n!$. Let $K_n = \hat{n} - 1$, which is the degree of the matrix $\dot{N}_n(\lambda, a)$. Let $\dot{b}_{ij}$ denote the (i, j) th entry of $\dot{N}_n(\lambda, a)$. By the definition of determinants we have

$$\det \dot{N}_n(\lambda, a) = \sum_{\pi \in S_{K_n}} \text{sgn } \pi \dot{b}_{\pi(1)1} \dot{b}_{\pi(2)2} \cdots \dot{b}_{\pi(K_n)K_n}.$$

We use the two-line notation

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & K_n \\ \pi(1) & \pi(2) & \cdots & \pi(K_n) \end{pmatrix}$$

to express a permutation π of $[K_n]$. For each j , if $\pi(j) = j + 1$, then the entry $\dot{b}_{\pi(j)j}$ is of degree 1 as a polynomial in λ , and otherwise it is a constant. Thus $\det \dot{N}_n(\lambda, a)$ is apparently of at most $K_n - 1 = \frac{(n+1)(n+2)}{2} - 2$ degree as a polynomial in λ . For example, $\dot{N}_3(\lambda, a)$ looks as follows:

$$\begin{pmatrix} \begin{array}{cc|ccc} -aF_1 & -a\mathbf{F}_1 & 0 & 0 & 0 \\ \lambda - a & -a & -ayF_2 & -ayF_2 & 0 \\ 0 & \lambda & 0 & -aF_2 & -a\mathbf{F}_2 \end{array} & \begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \\ \hline \begin{array}{cc|ccc} 0 & 0 & \lambda - a[2]_{xy} & -a[2]_{xy} & 0 \\ 0 & 0 & 0 & \lambda - a\mathbf{x} & -ax \\ 0 & 0 & 0 & 0 & \lambda \end{array} & \begin{array}{ccc|ccc} -ay^2F_3 & -ay^2F_3 & 0 & 0 & 0 & 0 \\ 0 & -ayF_3 & -ayF_3 & 0 & 0 & 0 \\ 0 & 0 & -aF_3 & -a\mathbf{F}_3 & 0 & 0 \end{array} \\ \hline \begin{array}{cc|ccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} & \begin{array}{ccc|ccc} \lambda - a[3]_{xy} & -a[3]_{xy} & 0 & 0 & 0 & 0 \\ 0 & \lambda - a\mathbf{x}[2]_{xy} & -ax[2]_{xy} & 0 & 0 & 0 \\ 0 & 0 & \lambda - a\mathbf{x}^2 & -ax^2 & 0 & 0 \end{array} \end{pmatrix}.$$

Our first claim is that $\det \dot{N}_n(\lambda, a)$ is a polynomial of degree $\frac{n(n+1)}{2}$. For each $i = 1, \dots, n$, let Col_i denote the set of columns $j = \frac{i(i+1)}{2}, \frac{i(i+1)}{2} + 1, \dots, \frac{(i+1)(i+2)}{2} - 1$. Note that Col_i includes $i+1$ columns. We claim that $\pi(j) = j+1$ can happen at most i column indices j in each block Col_i . Otherwise $\dot{b}_{\pi(1)1} \dot{b}_{\pi(2)2} \cdots \dot{b}_{\pi(K_n)K_n}$ vanishes. In fact, assume that $\pi(j) = j+1$ for all j in a certain block Col_i . Then this must be the case for the block Col_{i+1} . There is no other choice if we assume $\dot{b}_{\pi(1)1} \dot{b}_{\pi(2)2} \cdots \dot{b}_{\pi(K_n)K_n}$ is nonzero. And this must also be the case for the block Col_{i+2} , and so on. Finally we have to take $\pi(j) = j+1$ for all j in the block Col_n , but this is impossible. Thus we reach a contradiction. We conclude that the degree of $\det \dot{N}_n(\lambda, a)$ is at most $\frac{n(n+1)}{2}$. In fact there is a permutation which realizes this degree, i.e.,

$$\pi = \left(\begin{array}{cc|cc|ccc|cccc} 1 & 2 & 3 & 4 & 5 & \dots & \dots & K_{n-1}+1 & K_{n-1}+2 & \dots & K_n-1 & K_n \\ 2 & 1 & 4 & 5 & 3 & \dots & \dots & K_{n-1}+2 & K_{n-1}+3 & \dots & K_n & K_{n-1}+1 \end{array} \right).$$

It is easy to see that this π is the only permutation with which $\dot{b}_{\pi(1)1} \dot{b}_{\pi(2)2} \cdots \dot{b}_{\pi(K_n)K_n}$ does not vanish and is of degree $\frac{n(n+1)}{2}$. Thus we conclude that the leading coefficient of $\det \dot{N}_n(\lambda, a)$ equals

$$\text{sgn } \pi \cdot (-1)^n a^n F_n!.$$

This immediately implies the resulting identity (5.22). $\square$

Remark 5.8. One may notice that M_n in (5.7) and N_n in (5.19) are in a similar form, but our methods to evaluate them are far from parallel. It seems that the first method does not work with the matrix N_n since we can't guess the entries of $CA^{-1}N$ as we did in (5.11). Meanwhile, the second method does not work with the matrix M_n at this point since even if we generalize M_n to $M_n(\lambda, a)$, we don't know the general form of the eigenvectors of $M_n(\lambda, a)$. The reader can find the general guidance about matrix evaluation in [9]. We may say that the second proof follows this general philosophy.

6. Concluding remarks. We record some open problems in this section. First of all, it is desirable to find a simpler proof of Theorem 2.10. Second, regarding Steingrímsson's original problems in [18], there are still three conjectures remaining open, which we recall as follows. For an ordered partition $\pi = B_1/B_2/\dots/B_k$ in $\mathcal{OP}_n^k$ we say that i is a *block descent* in π if $B_i > B_{i+1}$; the *block major index* of π , denoted $\text{bMaj}(\pi)$, is the sum of the block descents in π . Consider the ordered partition $\pi = 6\ 8/5/1\ 4\ 7/3\ 9/2$, since $\{6, 8\} > \{5\}$ and $\{3, 9\} > \{2\}$, so 1 and 4 are the two block descents in π and $\text{bMaj } \pi = 1 + 4 = 5$. Then Steingrímsson presented the following conjecture in [18].

CONJECTURE 6.1 (Steingrímsson). *The following statistics are Euler–Mahonian:*

$$\text{mak} + \text{bMaj}, \quad \text{lmak} + \text{bMaj}, \quad \text{cmajLSB} := \text{lsb} + \text{cbMaj} + \binom{k}{2},$$

where $\text{cbMaj} = \binom{k}{2} - \text{bMaj}$.

Consider the following two generating functions of ordered partitions with $k \geq 0$ blocks:

$$(6.1) \quad \xi_k(a; x, y) := \sum_{\pi \in \mathcal{OP}^k} x^{(\text{mak} + \text{bMaj})\pi} y^{\text{cmajLSB } \pi} a^{|\pi|},$$

$$(6.2) \quad \eta_k(a; x, y) := \sum_{\pi \in \mathcal{OP}^k} x^{(\text{lmak} + \text{bMaj})\pi} y^{\text{cmajLSB } \pi} a^{|\pi|}.$$

Then we expect the following more general conjecture would hold.

CONJECTURE 6.2. *For $k \geq 0$, the following identities hold.*

$$(6.3) \quad \xi_k(a; x, y) = \eta_k(a; x, y) = \frac{a^k(xy)^{\binom{k}{2}} [k]_{x,y}!}{\prod_{i=1}^k (1 - a[i]_{x,y})}.$$

Note added in proof. A bijective approach to the conjectures of Steingrímsson has been recently discovered by the last two authors. This approach, as well as the proof of Conjectures 6.1 and 6.2, is given in a forthcoming paper (see <http://arxiv.org/abs/0712.1755v1>).

Acknowledgment. The authors would like to thank the referee for his/her careful reading of a previous version of this paper.

REFERENCES

- [1] L. CARLITZ, *q-Bernoulli and Eulerian numbers*, Trans. Amer. Math. Soc., 76 (1954), pp. 332–350.
- [2] L. CARLITZ, *A combinatorial property of q-Eulerian numbers*, Amer. Math. Monthly, 82 (1975), pp. 51–54.
- [3] R. CLARKE, E. STEINGRÍMSSON, AND J. ZENG, *New Euler-Mahonian statistics on permutations and words*, Adv. Appl. Math., 18 (1997), pp. 237–270.
- [4] PH. FLAJOLET, *Combinatorial aspects of continued fractions*, Discrete Math., 32 (1980), pp. 125–161.
- [5] D. FOATA, *Distribution Eulériennes et Mahoniennes sur le groupe des permutations*, in Higher Combinatorics, Berlin Combinatorics Symposium, M. Aigner, ed., D. Reidel, Boston, 1976, pp. 27–49.
- [6] D. FOATA AND D. ZEILBERGER, *Denert’s permutation statistic is indeed Euler-Mahonian*, Stud. Appl. Math., 83 (1990), pp. 31–59.
- [7] J. FRANÇON AND G. VIENNOT, *Permutations selon leurs pics, creux, doubles montées et doubles descentes, nombres d’Euler et de Genocchi*, Discrete Math., 28 (1979), pp. 21–35.
- [8] A. GARSIA, *On the “maj” and “inv” q-analogue of Eulerian numbers*, J. Linear Multilinear Algebra, 8 (1979), pp. 21–34.
- [9] C. KRATTENTHALER, *Advanced determinant calculus*, Sémin. Lothar. Combin., 42 (1999), article B42q.
- [10] G. KSAVRELOF AND J. ZENG, *Nouvelles statistiques de partitions pour les q-nombres de Stirling de seconde espèce*, Discrete Math., 256 (2002), pp. 743–758.
- [11] P. A. MACMAHON, *Combinatorial Analysis*, Vols. 1 and 2, Cambridge University Press, Cambridge, UK, 1915 (reprinted by Chelsea, New York, 1955).
- [12] S. MILNE, *Restricted growth functions, rank row matching of partition lattices, and q-Stirling numbers*, Adv. Math., 43 (1982), pp. 173–196.
- [13] B. REMMEL AND M. WACHS, *Rook theory, generalized Stirling numbers and (p,q)-analogues*, Electron. J. Combin., 11 (2004), no. 1, research paper 84.
- [14] B. SAGAN, *A maj statistics for set partitions*, European J. Combin., 12 (1991), pp. 69–79.
- [15] R. SIMION AND D. STANTON, *Specializations of generalized Laguerre polynomials*, SIAM J. Math. Anal., 25 (1994), pp. 712–719.
- [16] R. SIMION AND D. STANTON, *Octabasic Laguerre polynomials and permutation statistics*, J. Comput. Appl. Math., 68 (1996), pp. 297–329.
- [17] R. P. STANLEY, *Enumerative Combinatorics*, Vol. 1, Cambridge Stud. Adv. Math. 49, Cambridge University Press, Cambridge, UK, 1997.
- [18] E. STEINGRÍMSSON, *Statistics on Ordered Partitions of Sets*, preprint, 1999; available online from <http://arxiv.org/abs/math/0605670v4>.
- [19] X. VIENNOT, *Une théorie combinatoire des polynômes orthogonaux généraux*, Notes de conférences à l’UQAM, Montréal, 1983.
- [20] M. WACHS AND D. WHITE, *p, q-Stirling numbers and set partition statistics*, J. Combin. Theory Ser. A, 56 (1991), pp. 27–46.
- [21] M. WACHS, *σ -Restricted growth functions and p, q-Stirling numbers*, J. Combin. Theory Ser. A, 68 (1994), pp. 470–480.

- [22] D. WHITE, *Interpolating set partition statistics*, J. Combin. Theory Ser. A, 68 (1994), pp. 262–295.
- [23] J. ZENG, *The q -Stirling numbers, continued fractions and the q -Charlier and q -Laguerre polynomials*, J. Comput. Appl. Math., 57 (1995), pp. 413–424.
- [24] J. ZENG AND C. G. ZHANG, *A q -analog of Newton's series, Stirling functions and Eulerian functions*, Results Math., 25 (1994), pp. 370–391.

LONG LOCAL SEARCHES FOR MAXIMAL BIPARTITE SUBGRAPHS*

HEMANSHU KAUL[†] AND DOUGLAS B. WEST[‡]

Abstract. Given a partition of the vertices of a graph into two sets, a *flip* is a move of a vertex from its own set to the other, under the condition that it has more incident edges to vertices in its own set than in the other. Every sequence of flips eventually produces a bipartite subgraph capturing more than half of the edges in the graph. Each flip gains at least one edge. For an n -vertex loopless multigraph, we show that there is always a sequence of at most $n/2$ flips that cannot be extended, and we construct a graph having a sequence of $\frac{2}{25}(n^2 + n - 31)$ flips.

Key words. large bipartite subgraph, MAX-CUT problem, local switching, hill-climbing algorithm

AMS subject classifications. 05C35, 05C85, 68R10

DOI. 10.1137/060666615

1. Introduction. Finding a bipartite subgraph with the maximum number of edges in a given graph is a classical problem in combinatorial optimization and extremal graph theory. It has been studied in computer science from an algorithmic point of view (as the MAX-CUT problem) and in combinatorics from a more structural point of view. The problem is NP-complete (Karp [11]), so the focus of research has been on heuristics, approximation algorithms, and extremal results (conditions on graph parameters that guarantee large bipartite subgraphs).

For a graph G , let $b(G)$ denote the maximum number of edges in a bipartite subgraph of G . For the graph G in which we study $b(G)$, we always let n and m denote the numbers of vertices and edges. Classical guarantees include $b(G) \geq \frac{m}{2} + \frac{1}{8}(\sqrt{8m+1} - 1)$ (Edwards [5]) and, for connected graphs, $b(G) \geq \frac{m}{2} + \frac{1}{4}(n - 1)$ (conjectured by Erdős [7], and proved by Edwards [6]). Bylka, Idzik, and Tuza [3] strengthened the Edwards–Erdős bound to $b(G) \geq \lceil \frac{m}{2} + \frac{1}{2} \lceil \frac{2t-1}{2t} (n-1) \rceil \rceil$ when G is connected and has no odd cycle of length less than $2t$. Furthermore, this is sharp (using graphs whose blocks are $(2t+1)$ -cycles), and there is a linear-time algorithm to find a bipartite subgraph this big.

Bollobás and Scott [1] gave an algorithm running in time $O(2^{ck^4} + m + n)$ that finds a bipartite subgraph with at least $\frac{m}{2} + \sqrt{\frac{m}{8}} + k$ edges, if one exists, and otherwise provides an optimal partition (they also provide a survey of work in the area, updating the survey [14]). Goemans and Williamson ([9]) gave a randomized .87-approximation algorithm for $b(G)$ based on semidefinite programming. However, Håstad [10] proved that no .942-approximation algorithm exists unless $P=NP$.

An elementary heuristic approach to finding large bipartite subgraphs is to start with an arbitrary vertex partition and then make local improvements. Bylka, Idzik, and Tuza [3] found a local switching algorithm that guarantees a cut as big as the bound of Edwards [5]. A simple type of local improvement is to move a vertex having

*Received by the editors August 1, 2006; accepted for publication (in revised form) February 11, 2008; published electronically July 3, 2008.

<http://www.siam.org/journals/sidma/22-3/66661.html>

[†]Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616 (kaul@math.iit.edu).

[‡]Department of Mathematics, University of Illinois, Urbana, IL 61801 (west@math.uiuc.edu). This author's work supported in part by NSA grant H98230-06-1-0065.

more incident edges to vertices in its own set than in the other set; moving it to the other set increases the size of the cut. We call such a change a *flip* of the vertex. Erdős [8] introduced this notion to give a short proof that every graph G has a bipartite subgraph with at least half its edges. When flips are no longer possible, for each vertex at least half of the incident edges appear in the current bipartite subgraph, and hence by summing degrees the subgraph contains at least half the edges of G .

Since the process also must terminate (in at most m flips), this simple process is a .5-approximation algorithm for $b(G)$. Furthermore, it lends itself to parallel computation, since the determination of whether a vertex can flip is independent of that determination for other vertices. Parallelizable local-improvement algorithms based on switching protocols have been studied by Bylka, Idzik, and Tuza [3], by Bylka [2], and by others.

Our purpose in this paper is to study the *length* (number of flips) of a maximal sequence of flips, where *maximal* means that at the end every vertex has at least as many neighbors in the opposite side of the cut as in its own side. At that point, the cut is a local optimum for the MAX-CUT problem.

For the complete graph K_n , every maximal flip sequence has length $\lfloor n/2 \rfloor$. One might think that there is always a linear bound on the length of flip sequences, but Cowen and West [4] showed that this is false. They posed several questions about the lengths of maximal flip sequences in a graph G (we can consider these also for loopless multigraphs).

(1) Let $l(G)$ denote the maximum length of a flip sequence in G starting from the trivial vertex partition $(\emptyset, V(G))$. What is the maximum of $l(G)$ over n -vertex graphs? Over n -vertex graphs with m edges?

(2) Let $s(G)$ denote the minimum length of a maximal flip sequence starting from the trivial vertex partition. What is the maximum of $s(G)$ over n -vertex graphs? Is it true that always $s(G) \leq n/2$, and is there a fast algorithm to produce a flip sequence of length at most $n/2$?

Since every flip increases the number of edges across the cut, $l(G) \leq m$. When n is a perfect square and $m = \frac{1}{2}(n-1)\sqrt{n}$, Cowen and West [4] constructed a graph with a flip sequence of length m , matching the edge bound. It remains open whether $\frac{1}{2}n^{3/2}$ is asymptotically the maximum number of edges in a graph with a flip sequence whose length is the number of edges (increasing the size of the cut by only 1 with each flip).

The construction with length $\frac{1}{2}n^{3/2}$ was at that time the longest known flip sequence in an n -vertex graph. Since each flip increases the size of the cut, $\frac{1}{4}n^2$ is a trivial upper bound. Our main result, Theorem 3, is construction of an n -vertex graph having a flip sequence of length at least $\frac{2}{25}(n^2 + n - 31)$. This graph has roughly $\frac{1}{5}n^2$ edges.

Flips can be found in time proportional to maximum degree, so $l(G)$ times maximum degree is a bound on the worst-case running time of this simple heuristic algorithm. See Korte and Lovász [12], Loeb [13], and Tovey [15] for further discussion of issues related to the complexity of greedy and local-search based algorithms. Our result shows that one cannot guarantee less than $O(n^2)$ for the worst-case number of steps.

It would be interesting to know the coefficient on n^2 in the largest $l(G)$. We believe our construction is near optimal. In Proposition 2, we show that the maximum length of a flip sequence in a graph G with minimum degree d is at most $b(G) - \lfloor d^2/4 \rfloor$.

We also answer most of question (2), proving in Theorem 1 that $s(G) \leq \lfloor n/2 \rfloor$ for every n -vertex loopless multigraph G . The bound is sharp, as every maximal

flip sequence in K_n has length $\lfloor n/2 \rfloor$. It is also sharp for $K_{n/2, n/2}$, where every local optimum splits the vertices in half (including the “worst” local optimum with $n/4$ vertices of each partite set on each side when 4 divides n). Our proof does not provide a fast algorithm to find such a sequence, since it assumes knowledge of an optimal partition (in fact, it reaches some optimal partition in at most $\lfloor n/2 \rfloor$ flips). Thus the complexity of computing $s(G)$ or of finding a flip sequence of length at most $\lfloor n/2 \rfloor$ remains open. Is there a polynomial-time algorithm to find a sequence of $O(n)$ flips that results in a maximal bipartite subgraph? In terms of parallel computation, is there an NC algorithm that will construct a maximal bipartite subgraph?

There remain other open questions about local switching algorithms for MAX-CUT. Bylka, Idzik, and Tuza [3] showed that no algorithm using flips and certain other local switches guarantees a cut as big as the Edwards–Erdős bound [6]. A more general algorithm might come closer to that bound. Fix a positive integer k . Given a current vertex partition, a *local k -switch* selects subsets of size at most k from the two parts and switches them; it is *improving* if the new cut is bigger than the old one. Tuza [14] asked for the largest constant $c(k)$ such that a sequence of improving local k -switches can produce a bipartite subgraph with $\frac{m}{2} + c(k)n - o(n)$ edges.

2. Upper bound on $s(G)$. We use $N(v)$ to denote the neighborhood of a vertex v (in G). For $S \subseteq V(G)$ and $v \in V(G)$, let $d_S(v)$ be the number of edges incident to v whose other endpoints are in S . Recall that $s(G)$ is the minimum length of a maximal flip sequence from the trivial partition.

THEOREM 1. *If G is an n -vertex loopless multigraph, then $s(G) \leq \lfloor n/2 \rfloor$. In particular, given a globally optimal partition, a sequence of at most $\lfloor n/2 \rfloor$ flips can be constructed that produces some globally optimal partition.*

Proof. Let (X, Y) be a globally optimal partition, with $|X| = k \leq n/2$. By starting with $(\emptyset, V(G))$ and flipping to the first part of some or all of the members of the set X , we will obtain a globally optimal partition $(\hat{X}, \hat{Y})$ with $\hat{X} \subseteq X$ and $\hat{Y} \supseteq Y$.

Let X' denote the set of vertices in X having exactly half their neighbors in X and half in Y . Let $G_1, \dots, G_r$ be the components of the induced subgraph $G[X']$. Select $x_i \in V(G_i)$, for $1 \leq i \leq r$. Let $Z = \{x_1, \dots, x_r\}$. Index the rest of X as $x_{r+1}, \dots, x_k$ in such a way that each vertex of $X' - Z$ has an earlier neighbor in X' . In particular, choose a spanning tree of each G_i , root it at x_i , and search the tree from x_i , assigning labels in increasing order as vertices are found.

Now flip the vertices $x_k, \dots, x_{r+1}$ (from the second part to the first) in decreasing order of index. We claim that these flips are all available when desired and that the resulting partition $(\hat{X}, \hat{Y})$ is globally optimal. Since (X, Y) is globally optimal, $d_X(u) \leq d_Y(u)$ whenever $u \in X$, and $d_Y(v) \leq d_X(v)$ whenever $v \in Y$. Furthermore, these inequalities are strict for all vertices of odd degree.

With the given ordering of X , let $X_i = \{x_{k+1-i}, \dots, x_k\}$ for $1 \leq i \leq k-r$, with $X_0 = \emptyset$. Let $Y_i = V(G) - X_i$ for $0 \leq i \leq k-r$. Since $X_{i-1} \subset X$ and $Y \subset Y_{i-1}$, we have $d_{X_{i-1}}(x_i) \leq d_X(x_i) < d_Y(x_i) \leq d_{Y_{i-1}}(x_i)$ when $x_i \notin X'$. For $x_i \in X' - Z$, the indexing gives x_i a neighbor in $X \cap Y_{i-1}$. Thus the inequalities become $d_{X_{i-1}}(x_i) < d_X(x_i) \leq d_Y(x_i) < d_{Y_{i-1}}(x_i)$. The two cases allow us to flip $x_k, \dots, x_{r+1}$ in order.

We have reached a partition $(\hat{X}, \hat{Y})$ with $\hat{X} = X - Z$ and $\hat{Y} = Y \cup Z$. It remains to show that $(\hat{X}, \hat{Y})$ is a global optimum, which also implies that no further flips can be made. Reaching (X, Y) from $(\hat{X}, \hat{Y})$ is accomplished by moving the vertices of Z . Because Z consists of one vertex from each component of $G[X']$, Z is an independent

set. Since also $Z \subseteq X'$, $d_{\hat{X}}(x) = d_{\hat{Y}}(x)$ for each $x \in Z$. Therefore, moving all of Z does not change the size of the cut, and $(\hat{X}, \hat{Y})$ is already optimal. $\square$

If every vertex of G has odd degree, then the first statement of Theorem 1 can be proved more simply. In this case, we do not need to know a globally optimal partition to obtain a maximal flip sequence of length at most $\lfloor n/2 \rfloor$. We can start with any locally optimal partition, which can be found in polynomial time (as discussed in the introduction), and this partition can be reached from the trivial partition simply by flipping each vertex of one part. When G has vertices of even degree, our proof uses a given globally optimal partition to prevent vertices of Y from flipping; finding such a partition is NP-hard.

3. Upper bound on $l(G)$. Henceforth we consider only simple graphs. Recall that $l(G)$ is the maximum length of a flip sequence from the trivial partition of G . Since each flip increases the number of edges across the cut, $l(G) \leq b(G)$, and it is well known that $b(G) \leq n^2/4$. The example constructed by Cowen and West [4] is bipartite, with $l(G) = b(G) = m$. Achieving length $b(G)$ requires gaining only one edge with each flip, so G must have minimum degree 1. Their construction uses about $\sqrt{n}$ vertices of degree 1. It is natural to wonder by how much the upper bound $b(G)$ can be lowered when G has minimum degree d .

The basic idea is that the first few switches must gain many edges, leaving fewer for later flips. It may be that the later flips gain only one edge each. Indeed, this is what happens in our construction, where the minimum degree is linear in n and yet the number of flips is quadratic.

PROPOSITION 2. *If a graph G has minimum degree d , then $l(G) \leq b(G) - \lfloor d^2/4 \rfloor$.*

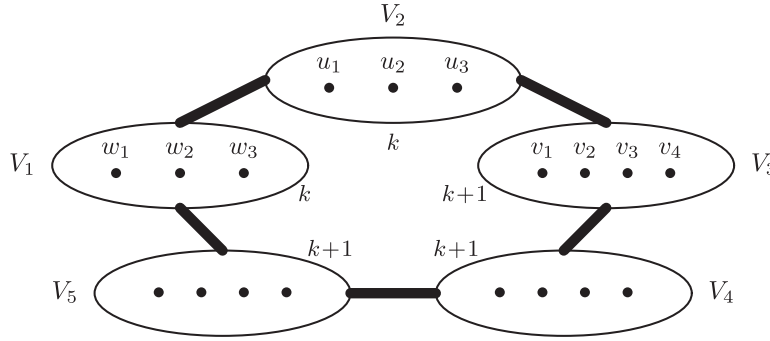
Proof. In a flip sequence on G from the initial partition with $(X, Y) = (\emptyset, V(G))$, the first move gains at least d edges. The second flip must move another vertex to X . As long as at most $1 + d/2$ vertices have been flipped, no vertex of X can have more neighbors in X than in Y , so no vertex in X can be flipped back to Y .

When $k = \lfloor d/2 \rfloor$, the first k flips thus yield a partition with $|X| = k$; at most $\binom{k}{2}$ edges are induced by X . Hence at this point the cut has at least $kd - 2\binom{k}{2}$ edges. Each subsequent flip after the first k gains at least one edge in the cut until at most $b(G)$ edges have been captured. The total number of flips is thus at most $k + b(G) - (kd - 2\binom{k}{2})$, which simplifies to $b(G) - k(d - k)$. With $k = \lfloor d/2 \rfloor$, the bound follows. $\square$

To see the effect of this bound, suppose that $4 \nmid n$ and $b(G)$ is $n^2/4$, the maximum possible value. This requires G to have the complete bipartite graph $K_{n/2, n/2}$ as a subgraph, and hence G has minimum degree at least $n/2$. Now Proposition 2 yields $l(G) \leq \frac{n^2}{4} - \frac{1}{4} \frac{n^2}{4} = \frac{3}{16} n^2$. More generally, if G is αn -regular, where $0 < \alpha \leq 1$, using $b(G) \leq m$ yields $l(G) \leq m - \frac{1}{4} \alpha^2 n^2 = m(1 - \frac{1}{2} \alpha)$.

For the graph G we construct in Theorem 3, where $l(G) > \frac{2}{25} n^2$, we have $b(G) \approx \frac{4}{25} n^2$, and the minimum degree is about $\frac{2}{5} n$. The resulting upper bound on $l(G)$ from Proposition 2 is about $\frac{3}{25} n^2$, which equals $\frac{3}{2} (\frac{2}{25} n^2)$. This suggests that it will be hard to improve Proposition 2 without knowing something about the structure of G . For example, when G is triangle-free, the gain in the first k flips in the proof of Proposition 2 is at least $kd - 2k^2/4$. This improves the upper bound in Proposition 2 to roughly $b(G) - \frac{3}{8} d^2$, and it improves the upper bound on $l(G)$ for the graph in Theorem 3 to about $\frac{5}{4} (\frac{2}{25} n^2)$.

4. Lower bound construction for $l(G)$. In this section, we construct a graph G and a flip sequence of length about $\frac{2}{25} n^2$ on it. As mentioned earlier, Cowen and

FIG. 1. Graph G in Theorem 3, with $k = 3$.

West [4] achieved a sequence of length $\frac{1}{2}n^{3/2}$ using a graph with that many edges and a sequence that increases the size of the cut by 1 with each flip. We achieve quadratic length by sacrificing equality with m and enlarging the cut rapidly at first in order to permit many more small steps later in which vertices repeatedly move back and forth. With such a sequence we come near the counting bound from Proposition 2.

THEOREM 3. *For every positive integer n , there exists an n -vertex graph G such that $l(G) \geq \frac{2}{25}(n^2 + n - 31)$.*

Proof. We first present an explicit construction for $n \equiv 3 \pmod{5}$, with $l(G) > \frac{2}{25}n^2$. Given a positive integer k , let $n = 5k + 3$. Create vertex subsets $V_1, \dots, V_5$, with $|V_1| = |V_2| = k$ and $|V_3| = |V_4| = |V_5| = k + 1$. Form G by expanding the vertices of a 5-cycle into these sets. That is, given the 5-cycle with vertices $z_1, \dots, z_5$, replace z_i with V_i for each i , with every vertex in V_i adjacent to every vertex in V_{i-1} and V_{i+1} , treating subscripts modulo 5. Label the vertices in the first three sets as follows: $V_1 = \{w_1, \dots, w_k\}$, $V_2 = \{u_1, \dots, u_k\}$, and $V_3 = \{v_1, \dots, v_{k+1}\}$. See Figure 1.

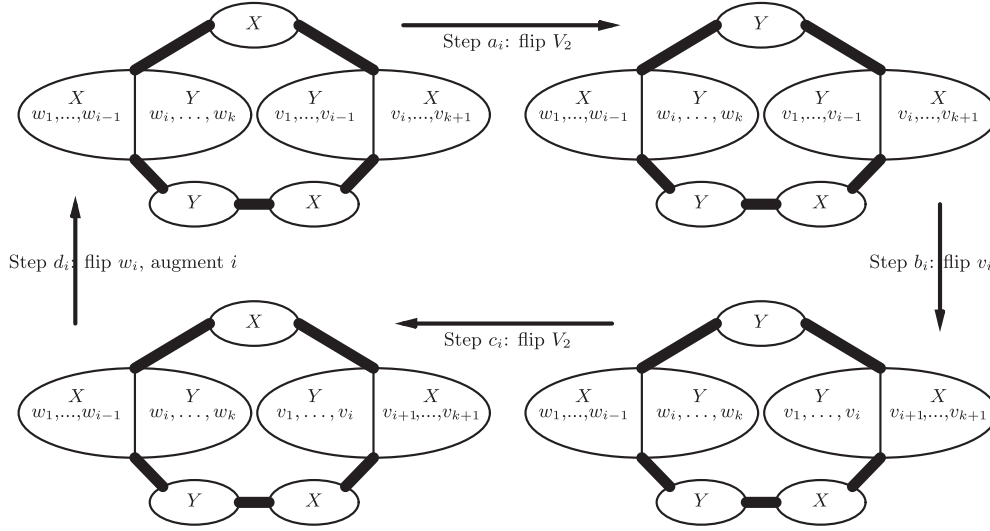
We start with the trivial partition $(V(G), \emptyset)$. During the flip sequence, we use X and Y , respectively, to denote the first and the second set in the current partition, so initially $X = V(G)$. At the end of the flip sequence, we will have $X = V_1 \cup V_2 \cup V_4$ and $Y = V_3 \cup V_5$. As explained after the proof, this in fact is an optimal cut; it has $2k(k+1) + 2(k+1)^2$ edges, including all the edges of G except the k^2 edges between V_1 and V_2 .

Our flip sequence for G begins with a “preprocessing” Phase 0 that gains many edges for each flip. Since G has a large minimum degree, the argument in Proposition 2 implies that the flip sequence must start this way.

Phase 0 consists of two steps. We first flip each vertex of V_1 from X to Y . This is valid because all the neighbors of each vertex of V_1 lie in $X - V_1$ throughout this phase. We next flip each vertex of V_5 from X to Y . This is valid because, after moving V_1 , each vertex of V_5 has $k + 1$ neighbors in X (all of V_4) and k neighbors in Y (all of V_1) throughout the remainder of this phase.

After Phase 0, we have $V_4 \subset X$ and $V_5 \subset Y$, and this remains true for the remainder of the flip sequence. We group the remainder of the flips into Phases 1 through $k + 1$.

Let $V_1(i) = \{w_1, \dots, w_{i-1}\}$ and $V'_1(i) = \{w_i, \dots, w_k\}$. Similarly, let $V_3(i) = \{v_i, \dots, v_{i-1}\}$ and $V'_3(i) = \{v_i, \dots, v_{k+1}\}$. Thus $V_1 = V'_1(1)$ and $V_3 = V'_3(1)$ and $V_1(1) = V_3(1) = \emptyset$. Hence the cut at the end of Phase 0 can be described as $X = V_1(1) \cup V_2 \cup V'_3(1) \cup V_4$ and $Y = V'_1(1) \cup V_3(1) \cup V_5$, which is the description of the starting cut for Phase 1. The four steps of Phase i are shown in Figure 2.

FIG. 2. Steps in Phase i .

Phase i . At the start of Phase i , $X = V_1(i) \cup V_2 \cup V_3'(i) \cup V_4$ and $Y = V_1'(i) \cup V_3(i) \cup V_5$.

Step a_i : Flip all of V_2 from X to Y ; this step is valid because each vertex in V_2 now has $k + 1$ neighbors in X (all of $V_1(i) \cup V_3'(i)$) and k neighbors in Y (all of $V_1'(i) \cup V_3(i)$).

Step b_i : Flip v_i in V_3 from X to Y ; this step is valid because v_i now has $k + 1$ neighbors in X (all of V_4) and k neighbors in Y (all of V_2).

Step c_i : Flip all of V_2 from Y to X ; this step is valid because each vertex in V_2 now has $k + 1$ neighbors in Y (all of $V_3(i + 1) \cup V_1'(i)$) and k neighbors in Y (all of $V_3(i) \cup V_1'(i)$).

Step d_i : If $i \leq k$, then flip w_i in V_1 from Y to X ; otherwise stop. For $i \leq k$, the flip is valid because w_i now has $k + 1$ neighbors in Y (all of V_5) and k neighbors in X (all of V_2).

At the end of Phase $k + 1$, the partition is $X = V_1 \cup V_2 \cup V_4$ and $Y = V_3 \cup V_5$. The flip sequence ends, since each vertex having any neighbors in its own side of the cut has k neighbors in its side and $k + 1$ in the opposite side.

In Phase 0, both V_1 and V_5 were flipped; altogether $2k + 1$ flips. Each subsequent Phase flips V_2 twice and flips one vertex of V_3 and one vertex of V_1 , except that the last phase omits the last flip. Thus the total number of flips is $2k + (k + 1)(2k + 2)$. Since $n = 5k + 3$, we obtain $l(G) \geq \frac{2}{25}(n^2 + 9n - 11)$.

Finally, suppose that $n \not\equiv 3 \pmod{5}$. Let n' be the largest integer congruent to 3 modulo 5 that is at most n . We use an instance of the construction above with n' vertices and leave the remaining vertices isolated. Since $n' \geq n - 4$, we have $l(G) \geq \frac{2}{25}(n'^2 + 9n' - 11) \geq \frac{2}{25}(n^2 + n - 31)$. $\square$

In fact, the final partition in the flip sequence we constructed in the proof of Theorem 3 is globally optimal, as we now show. Let G be any expansion of a 5-cycle replacing the i th vertex with an independent V_i , indexed cyclically. For any partition (X, Y) of $V(G)$ yielding a bipartite subgraph H , let $a_i = |X \cap V_i|$ and $b_i = |Y \cap V_i|$. Each vertex v in V_i has $a_{i-1} + a_{i+1}$ neighbors in X and $b_{i-1} + b_{i+1}$ neighbors in Y . Vertex v has largest degree in H if it is in the set opposite the larger of these two quan-

ties. This applies for all $v \in V_i$ (and for all i), so we achieve the largest bipartite subgraph by putting each V_i onto one side of the cut. Under this restriction, the best we can achieve is to capture all but the smallest of the five complete bipartite subgraphs expanding edges of the cycle, which is what the final partition of Theorem 3 does.

REFERENCES

- [1] B. BOLLOBÁS AND A. D. SCOTT, *Better bounds for max cut*, in Contemporary Combinatorics, Bolyai Soc. Math. Stud. 10, B. Bollobás, ed., János Bolyai Math. Soc., Budapest, 2002, pp. 185–246.
- [2] S. BYLKA, *Local improving algorithms for large cuts in graphs with maximum degree three*, Discrete Math., 235 (2001), pp. 53–67.
- [3] S. BYLKA, A. IDZIK, AND Z. TUZA, *Maximum cuts—improvements and local algorithmic analogues of the Edwards-Erdős inequality*, Discrete Math., 194 (1999), pp. 39–58.
- [4] L. J. COWEN AND D. B. WEST, *Greedy vertex partition of a graph, problem 382*, in Research Problems, Kleitman and Combinatorics: A Celebration, Discrete Math., 257 (2002), pp. 599–624.
- [5] C. S. EDWARDS, *Some extremal properties of bipartite subgraphs*, Canad. J. Math., 25 (1973), pp. 475–485.
- [6] C. S. EDWARDS, *An improved lower bound for the number of edges in a largest bipartite subgraph*, in Recent Advances in Graph Theory (Proc. Second Czechoslovak Sympos., Prague, 1974), Academia, Prague, 1975, pp. 167–181.
- [7] P. ERDŐS, *On some extremal problems in graph theory*, Israel J. Math., 3 (1965), pp. 113–116.
- [8] P. ERDŐS, *Problems and results on finite and infinite graphs*, in Recent Advances in Graph Theory (Proc. Second Czechoslovak Sympos., Prague, 1974), Academia, Prague, 1975, pp. 183–192.
- [9] M. X. GOEMANS AND D. P. WILLIAMSON, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. Assoc. Comp. Mach., 42 (1995), pp. 1115–1145.
- [10] J. HÅSTAD, *Some optimal inapproximability results*, in Proceedings of the 29th Annual ACM Symposium on Th. Comput., El Paso, 1997, J. Assoc. Comp. Mach., 48 (2001), pp. 798–859.
- [11] R. M. KARP, *Reducibility among combinatorial problems*, in Complexity of Computer Computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972), R. Miller and J. Thatcher, eds., Plenum Press, New York, 1972, pp. 85–103.
- [12] B. KORTE AND L. LOVÁSZ, *Mathematical structures underlying greedy algorithms*, in Fundamentals of Computational Theory, Lecture Notes in Comput. Sci. 117, F. Gécseg, ed., Springer, Berlin, 1981, pp. 205–209.
- [13] M. LOEBL, *Efficient maximal cubic graph cuts*, in ICALP, Lecture Notes in Comput. Sci. 510, J. Albert, B. Monien, and M. Rodrigues-Artalejo, eds., Springer, Berlin, 1991, pp. 351–362.
- [14] S. POLJAK AND Z. TUZA, *Maximum cuts and large bipartite subgraphs*, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 20 (1995), pp. 181–244.
- [15] C. TOVEY, *Local improvement on discrete structures*, in Local Search in Combinatorial Optimization, Wiley-Interscience Series, E. Aarts and J. K. Lenstra, eds., Wiley, New York, 1997, pp. 57–89.

ELEMENTARY INCIDENCE THEOREMS FOR COMPLEX NUMBERS AND QUATERNIONS*

JÓZSEF SOLYMOSI[†] AND KONRAD J. SWANEPOEL[‡]

Abstract. We present some elementary ideas to prove the following Sylvester–Gallai type theorems involving incidences between points and lines in the planes over the complex numbers and quaternions.

1. Let A and B be finite sets of at least two complex numbers each. Then there exists a line ℓ in the complex affine plane such that $|(A \times B) \cap \ell| = 2$.
2. Let S be a finite noncollinear set of points in the complex affine plane. Then there exists a line ℓ such that $2 \leq |S \cap \ell| \leq 5$.
3. Let A and B be finite sets of at least two quaternions each. Then there exists a line ℓ in the quaternionic affine plane such that $2 \leq |(A \times B) \cap \ell| \leq 5$.
4. Let S be a finite noncollinear set of points in the quaternionic affine plane. Then there exists a line ℓ such that $2 \leq |S \cap \ell| \leq 24$.

Key words. Sylvester–Gallai theorem, complex affine plane, quaternionic affine plane, incidence geometry

AMS subject classifications. Primary, 52C10; Secondary, 51A30, 52C35

DOI. 10.1137/070685117

1. Introduction. The Sylvester–Gallai theorem [12, 11] states that, for any finite noncollinear set of points in the Euclidean plane $\mathbb{R}^2$, there is a line intersecting the set in exactly two points. See [1, section 7.2] for a recent list of references regarding this theorem and its relatives. It is interesting to ask what happens if the underlying field is not the real numbers. In this paper we consider the complex numbers $\mathbb{C}$ and the quaternions $\mathbb{H}$. For more general considerations, see [8]. We consider the plane $\mathbb{K}^2$ over the skew field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ to be the two-dimensional left vector space over $\mathbb{K}$. Lines are defined to be the translates of 1-dimensional subspaces. Thus lines are either vertical, i.e., of the form $\{(x, y) \in \mathbb{K}^2 : x = c\}$ for some $c \in \mathbb{K}$, or nonvertical: $\{(x, y) \in \mathbb{K}^2 : y = xm + c\}$ for some $m, c \in \mathbb{K}$. (The equation of a nonvertical line would be the more familiar $y = mx + c$ if we took $\mathbb{K}^2$ to be a right vector space). When $\mathbb{C}^2$ is identified with $\mathbb{R}^4$, lines are special types of 2-dimensional real subspaces, and therefore incidence problems involving them cannot be reduced to the real case. A similar remark holds for the identification of $\mathbb{H}^2$ with $\mathbb{R}^8$.

As observed by Kelly [6], it is a consequence of an inequality of Hirzebruch [5] that any finite noncollinear subset of $\mathbb{C}^2$ always intersects some line in 2 or 3 points (which is best possible, as shown by the nine inflection points of a cubic curve). No elementary proof is known of either the inequality or this consequence (see [1, section 7.3, Problem 7]). However, by adapting Kelly’s proof [2] of the Sylvester–Gallai theorem that uses

*Received by the editors March 13, 2007; accepted for publication (in revised form) March 10, 2008; published electronically July 3, 2008.

<http://www.siam.org/journals/sidma/22-3/68511.html>

[†]Department of Mathematics, The University of British Columbia, Vancouver, V6T 1Z2 BC, Canada (solymosi@math.ubc.ca). This author was partially supported by NSERC and OTKA grants and by a Sloan Fellowship.

[‡]Faculty of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany (konrad.swanepoel@gmail.com). This author was partially supported by the South African National Research Foundation. Most of this paper was written while he was a member of the Department of Mathematical Sciences, University of South Africa. He also gratefully acknowledges the invitation and support of the Department of Mathematics, University of British Columbia.

the smallest perpendicular distance between points and nonincident lines, it is possible to prove the following somewhat weaker statement.

THEOREM 1. *Let S be a finite noncollinear subset of $\mathbb{C}^2$. Then there exists a line ℓ such that $2 \leq |S \cap \ell| \leq 5$.*

Using the consequence of the inequality of Hirzebruch mentioned above, Kelly [6] showed that, for any finite nonplanar subset of $\mathbb{C}^3$, there exists a line intersecting the set in exactly two points, thus answering a question of Serre [9]. Another proof can be found in [4] that avoids the use of Hirzebruch's inequality. There it is also shown that any finite subset of $\mathbb{H}^4$ not on a hyperplane has a line intersecting the set in exactly two points. The approach in our proof of Theorem 1 also gives the following.

THEOREM 2. *Let S be a finite noncollinear subset of $\mathbb{H}^2$. Then there exists a line ℓ such that $2 \leq |S \cap \ell| \leq 24$.*

Note that although the number 24 is most likely too high, no example is known that shows it has to be larger than 3.

Many applications of incidence bounds to combinatorial number theory use only the special case where the set of points is assumed to be a grid, i.e., a Cartesian product of finite subsets of the ground field (see, for example, [10] and the survey paper of Elekes [3]). For this special case we obtain the following much stronger conclusions.

THEOREM 3. *Let $A, B \subset \mathbb{C}$, with $2 \leq |A|, |B| < \infty$. Then there exists a line ℓ in $\mathbb{C}^2$ such that $|(A \times B) \cap \ell| = 2$.*

THEOREM 4. *Let $A, B \subset \mathbb{H}$, with $2 \leq |A|, |B| < \infty$. Then there exists a line ℓ in $\mathbb{H}^2$ such that $2 \leq |(A \times B) \cap \ell| \leq 5$.*

2. Proofs. We consider $\mathbb{C}^2$ as a complex inner product space in the usual way, i.e., we define the inner product of vectors (a_1, b_1) and (a_2, b_2) in $\mathbb{C}^2$ as $\langle (a_1, b_1), (a_2, b_2) \rangle = a_1 \bar{a}_2 + b_1 \bar{b}_2$, the norm of $(a, b) \in \mathbb{C}^2$ as $\|(a, b)\| = \sqrt{\langle (a, b), (a, b) \rangle}$, and the distance between points (a_1, b_1) and (a_2, b_2) in $\mathbb{C}^2$ as $\|(a_1 - a_2, b_1 - b_2)\|$. Working carefully with inner products one can then determine the shortest distance between a point and a line in $\mathbb{C}^2$, as will be done in the next proof.

Proof of Theorem 1. As in Kelly's proof of the real Sylvester–Gallai theorem, we choose three noncollinear points $p, q, r \in S$ such that the distance between p and the line qr is a minimum. We show that qr contains at most 5 points.

After a unitary transformation and a dilatation (thus preserving the minimum distance) we may assume that $p = (0, 1)$, and that the line qr is the x -axis, i.e., $qr \cap S = \{(z_1, 0), \dots, (z_k, 0)\}$. Thus the distance between p and qr equals 1. Choose any two points $(z_i, 0), (z_j, 0) \in qr \cap S$. Then the distance between $(z_j, 0)$ and the line ℓ through $(0, 1)$ and $(z_i, 0)$ must be at least 1. Considering an arbitrary point $(1 - \lambda)(z_i, 0) + \lambda(0, 1)$ on ℓ , a simple calculation gives the square of its distance to $(z_j, 0)$ to be

$$\begin{aligned} & \|(1 - \lambda)(z_i, 0) + \lambda(0, 1) - (z_j, 0)\|^2 \\ &= \overline{(z_i - z_j)}^2 - (z_i - z_j)\lambda \bar{z}_i - \lambda z_i \overline{(z_i - z_j)} + |\lambda|^2 |z_i|^2 + |\lambda|^2 \\ &= \frac{|z_i - z_j|^2}{1 + |z_i|^2} + \frac{1 + |z_i|^2}{|z_i|^2} \left| \frac{|z_i|^2}{1 + |z_i|^2} (z_i - z_j) - \lambda z_i \right|^2. \end{aligned}$$

The last expression is minimized when

$$\lambda = \frac{(z_i - z_j)\bar{z}_i}{1 + |z_i|^2},$$

which gives that the square of the distance between $(z_j, 0)$ and ℓ is

$$\frac{|z_i - z_j|^2}{1 + |z_i|^2}.$$

Since this quantity is at least 1 for all distinct i and j , we obtain, in particular, that $|z_i - z_j| > |z_i|$ and $|z_i - z_j| > |z_j|$. Identify $\mathbb{C}$ with $\mathbb{R}^2$. Then for each triangle $\triangle 0z_i z_j$ in the Euclidean plane, the side $z_i z_j$ is the unique longest side. This implies that the angle $\angle z_i 0 z_j > 60^\circ$, giving $|qr \cap S| < 6$. $\square$

Proof of Theorem 2. We consider $\mathbb{H}^2$ to be a left vector space, which can be turned into an inner product space in exactly the same way as $\mathbb{C}^2$ where we recall that the conjugate of a quaternion $\alpha = a + bi + cj + dk$ is $\bar{\alpha} = a - bi - cj - dk$. The distance between two points is defined as before, and we may follow the previous proof almost verbatim, since the commutative law is not used anywhere. The only change is that the Euclidean plane has to be replaced by the identification of $\mathbb{H}$ with $\mathbb{R}^4$. We obtain that $|qr \cap S|$ is bounded above by the number of rays from the origin in $\mathbb{R}^4$ such that the angle at the origin between each pair exceeds 60° . This is bounded above by the kissing number of a ball in $\mathbb{R}^4$, which is 24 [7]. $\square$

We in fact obtained in the above proof that the *strict kissing number* is an upper bound, i.e., the largest number of nontouching unit balls in $\mathbb{R}^4$ that all touch a single unit ball. It is conjectured that the kissing arrangement of 24 balls in $\mathbb{R}^4$ is unique [7], which would imply that the strict kissing number is at most 23.

Proof of Theorem 3. Suppose that the conclusion is false. Let a_1 and a_2 be a closest pair in A , and b_1 and b_2 a furthest pair in B . Then there exists a third point $(a_3, b_3) \in A \times B$ such that (a_1, b_1) , (a_2, b_2) , (a_3, b_3) are collinear. Let the equation of the (nonvertical) line in $\mathbb{C}^2$ on which they lie be $y = xm + c$, $m, c \in \mathbb{C}$. Again identify $\mathbb{C}$ with the Euclidean plane $\mathbb{R}^2$. Then the map $\varphi : \mathbb{C} \rightarrow \mathbb{C}$, $z \mapsto zm + c$ is an orientation-preserving similarity transformation, with $\varphi(a_i) = b_i$, $i = 1, 2, 3$. Since a closest pair in $\{a_1, a_2, a_3\}$ is mapped to a furthest pair in $\{b_1, b_2, b_3\}$ and ratios of distances are preserved, $\triangle a_1 a_2 a_3$ and $\triangle b_1 b_2 b_3$ are both equilateral triangles in the plane.

We now interchange b_1 and b_2 , i.e., we consider the line through (a_1, b_2) and (a_2, b_1) with equation $y = xm' + c'$. By assumption there is a third point (a'_3, b'_3) on this line, and exactly as before, the map $\varphi' : \mathbb{C} \rightarrow \mathbb{C}$, $z \mapsto zm' + c'$ preserves orientation and maps the equilateral triangle $\triangle a_1 a_2 a'_3$ to $\triangle b_2 b_1 b'_3$.

Suppose that $a_3 = a'_3$. Then $b_3 \neq b'_3$, since $\triangle b_1 b_2 b_3$ and $\triangle b_2 b_1 b_3$ have opposite orientations. Thus $\triangle b_1 b_2 b_3$ and $\triangle b_1 b_2 b'_3$ are distinct equilateral triangles with a common edge; hence $|b_3 - b'_3| = \sqrt{3}|b_1 - b_2| > |b_1 - b_2|$, which contradicts the choice of b_1 and b_2 .

Therefore $a_3 \neq a'_3$. We have shown that any closest pair of points in A is contained in the vertex sets of two equilateral triangles. This, however, contradicts the finiteness of A . $\square$

Note that we made essential use of the fact that the similarities preserve orientation. Without this, we would only be able to conclude that there is a line ℓ such that $2 \leq |(A \times B) \cap \ell| \leq 3$.

Proof of Theorem 4. We proceed along the same lines as in the previous proof. Let $\{a_1, a_2\}$ be a closest pair in A , and $\{b_1, b_2\}$ a furthest pair in B . Consider the line ℓ through (a_1, b_1) and (a_2, b_2) . Since $a_1 \neq a_2$, ℓ is of the form $y = xm + c$. Then the mapping $\varphi : \mathbb{H} \rightarrow \mathbb{H}$; $x \mapsto xm + c$ maps $A' := \{x \in \mathbb{H} : (x, y) \in \ell \cap (A \times B)\}$ to $B' := \{y \in \mathbb{H} : (x, y) \in \ell \cap (A \times B)\}$. Identifying $\mathbb{H}$ with $\mathbb{R}^4$ we see that φ is a

a similarity transformation of $\mathbb{R}^4$, i.e., a composition of a translation, rotation and dilation. Thus B' is similar to A' . Because the smallest distance in A' is mapped to the largest distance in B' , A' and B' are both equilateral sets in $\mathbb{R}^4$; hence $|\ell \cap (A \times B)| = |A'| = |B'| \leq 5$. $\square$

Acknowledgments. We thank the anonymous referees for their valuable remarks.

REFERENCES

- [1] P. BRASS, W. MOSER, AND J. PACH, *Research Problems in Discrete Geometry*, Springer, New York, 2005.
- [2] H. S. M. COXETER, *A problem of collinear points*, Amer. Math. Monthly, 55 (1948), pp. 26–28.
- [3] GY. ELEKES, *SUMS versus PRODUCTS in number theory, algebra and Erdős geometry*, in Paul Erdős and His Mathematics, II (Budapest, 1999), Bolyai Soc. Math. Stud. 11, L. Lovasz, M. Simonovitz, and V. T. Sós, eds., János Bolyai Math. Soc., Budapest, 2002, pp. 241–290.
- [4] N. ELKIES, L. M. PRETORIUS, AND K. J. SWANEPOEL, *Sylvester-Gallai theorems for complex numbers and quaternions*, Discrete Comput. Geom., 35 (2006), pp. 361–373.
- [5] F. HIRZEBRUCH, *Arrangements of Lines and Algebraic Surfaces, Arithmetic and Geometry, Vol. II*, Birkhäuser Boston, Cambridge, MA, 1983, pp. 113–140.
- [6] L. M. KELLY, *A resolution of the Sylvester-Gallai problem of J.-P. Serre*, Discrete Comput. Geom., 1 (1986), pp. 101–104.
- [7] O. R. MUSIN, *The Kissing Number in Four Dimensions*, preprint, <http://arxiv.org/math.MG/0309430>, 2005.
- [8] L. M. PRETORIUS AND K. J. SWANEPOEL, *The Sylvester-Gallai theorem, colourings and algebra*, Discrete Math., to appear, doi:10.1016/j.disc.2007.12.027.
- [9] J.-P. SERRE, *Problems*, Amer. Math. Monthly, 73 (1966), p. 89.
- [10] J. SOLYMOSI AND G. TARDOS, *On the number of k -rich transformations*, in Proceedings of the 23rd Annual ACM Symposium on Comp. Geom., Gyeongju, Republic of Korea, 2007, pp. 227–231.
- [11] R. STEINBERG, *Solution to problem 4065*, Amer. Math. Monthly, 51 (1944), pp. 169–171.
- [12] J. J. SYLVESTER, *Educational Times*, 46 (1893), pp. 156.

SEQUENCES WITH CHANGING DEPENDENCIES*

PAUL BALISTER[†], BÉLA BOLLOBÁS[‡], AND STEFANIE GERKE[§]

Abstract. Consider words over an alphabet with n letters. Fisher [*Amer. Math. Monthly*, 96 (1989), pp. 610–614] calculated the number of distinct words of length ℓ assuming certain pairs of letters commute. In this paper we are interested in a more general setting where the pairs of letters that commute at a certain position of a word depend on the initial segment of the word. In particular, we show that if for each word at each position any letter fails to commute with at most a constant number of other letters, then the number of distinct words of length ℓ is at most $C^{n+\ell}$ for some constant C . We use this result to obtain a lower bound on the number of diagonal flips required in the worst case to transform one n -vertex labeled triangulated planar graph into some other one. This has previously been proved in [D. D. Sleator, R. E. Tarjan, and W. P. Thurston, *SIAM J. Discrete Math.*, 5 (1992), pp. 428–450] by different methods.

Key words. sequences, triangulations, commuting sets

AMS subject classification. 05C30

DOI. 10.1137/060663611

1. Introduction. Let $\mathcal{A} = \{a_1, \dots, a_m\}$ be a finite alphabet consisting of letters $a_1, \dots, a_m$. We denote the set of all words of finite length with letters from $\mathcal{A}$ by $\mathcal{A}^*$. Given a word $w \in \mathcal{A}^*$, we write $[w]_t$ for the prefix of w of length t (that is, $[w]_t$ consists of the first t letters of w). For each word w in $\mathcal{A}^*$ we fix a set $C_w \subseteq \mathcal{A}$. We call the collection $\mathcal{C} = (C_w : w \in \mathcal{A}^*)$ the *commuting family*. Let $\sim_{\mathcal{C}}$ be the smallest equivalence relation such that

$$a_1 \dots a_{t-1} a_t a_{t+1} \dots a_{\ell} \sim_{\mathcal{C}} a_1 \dots a_{t-1} a_{t+1} a_t \dots a_{\ell}$$

whenever $a_{t+1} \in C_{a_1 \dots a_t}$. By enlarging the sets C_w , if necessary, we may assume that for all $a, b \in \mathcal{A}$ and $w \in \mathcal{A}^*$, we have $a \in C_{wb} \Leftrightarrow b \in C_{wa}$. Note that this does not change the equivalence relation. We call two words w and w' equivalent (with respect to $\mathcal{C}$) if $w \sim_{\mathcal{C}} w'$.

Fisher [4] considered the number T_{ℓ} of equivalent classes of words of length ℓ , when C_w depends only on the last letter of w (that is, we specify the pairs of letters that commute). In this case one can construct a dependency graph $G = (V, E)$ where the vertices correspond to the letters of the alphabet and two vertices are adjacent if the corresponding letters commute. Fisher showed that for $\ell > 0$, $T_{\ell} = \sum_{k=1}^{\ell} (-1)^{k+1} c_k T_{\ell-k}$ with $T_0 = 1$, where c_k is the number of complete subgraphs of size k in G . It follows that the generating function $t(x) = \sum_{k=0}^{\infty} T_k x^k$ is closely related to the well-known independence polynomial $Z_G(x) = \sum_{k=0}^{|V|} s_k x^k$, where s_k is the number of independent sets with k vertices. More precisely, $t(x) = 1/Z_{\overline{G}}(-x)$,

*Received by the editors June 23, 2006; accepted for publication (in revised form) April 6, 2008; published electronically July 3, 2008. Research was performed while the third author was visiting the University of Memphis and also while all authors were visiting the Institute for Mathematical Sciences at the National University of Singapore.

<http://www.siam.org/journals/sidma/22-3/66361.html>

[†]Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152-3240 (pbalistr@memphis.edu).

[‡]Trinity College, Cambridge, CB2 1TQ, UK, and University of Memphis, Memphis, TN 38152-3240 (B.Bollobas@dpmms.cam.ac.uk).

[§]Royal Holloway College, University of London, Egham TW20 0EX, UK (stefanie.gerke@rhul.ac.uk).

where $\overline{G}$ is the complement of G . The notation we use here, Z_G , is the one used in statistical physics, where (a generalization of) the polynomial goes under the name of the “lattice-gas partition function” associated with the graph G . For a review of various combinatorial results concerning the independence polynomial, see Levit and Mandrescu [7].

Clearly, $s_0 = 1$, $s_1 = |V|$ for every graph $G = (V, E)$, and $s_n = 0$ unless G has no edges, in which case it is 1. In particular, Z_G is nonzero in some complex disc

$$D_R = \{z \in \mathbb{C} : |z| \leq R\},$$

where $R > 0$.

Extending a result of Kotecký and Preiss [6], Dobrushin [1] (see also [2]) proved the following fundamental result in statistical physics, widely known as *Dobrushin’s theorem*.

THEOREM 1.1. *Let Z_G be the independence polynomial of a graph $G = (V, E)$. Let r_v , $v \in V$, be constants such that $0 \leq r_v < 1$, and*

$$R \leq r_v \prod_{u: uv \in E} (1 - r_u)$$

for every $v \in V$. Then Z_G is nonzero in the disc $|z| \leq R$.

In fact, if $|z| \leq R$, then

$$(1) \quad |Z_G(z)| \geq Z_G(-R) \geq \prod_{v \in V} (1 - r_v).$$

Twenty years earlier, Lovász had proved a celebrated result in combinatorics with analogous conditions (see [3]). This result, the *Lovász local lemma*, gives necessary conditions for the probability of the intersection of a family of events to be strictly positive, when these events are not far from being independent. Recently, Scott and Sokal [9], [8] showed that these two fundamental results, the Lovász local lemma and Dobrushin’s theorem, are, in fact, equivalent. The proof of this equivalence is not too easy: it is based on the result of Shearer [10] that the conditions in the Lovász local lemma are best possible.

As a consequence of Dobrushin’s theorem, if the degree of every vertex in $G = (V, E)$ is bounded by k , then one can take $R = R(k) > 0$ independently of the size of G . Since $t(x) = 1/Z_{\overline{G}}(-x)$, this implies, for example, that there are at most $C(k)^{\ell+|V|}$ different equivalence classes of words of length ℓ if there are at most k letters that do not commute with any given letter (in other words the dependency graph G has minimum degree at least $|V| - k - 1$).

Sleater, Tarjan, and Thurston [11] used a different setup to count the number of different outcomes of sequences of transformations of graphs. Each transformation is obtained by replacing a subgraph isomorphic to one of a fixed family of graphs by another graph. Whether or not two transformations commute depends on the previous transformations applied. In this sense it is very similar to our approach, but ours appears to be more general as well as simpler.

As an example where the sets C_w depend on the entire word and not only the last letter, we consider words in connection with the class of labeled triangulations on n vertices, that is, maximal planar graphs on the vertex set $\{1, \dots, n\}$. Consider a fixed triangulation that is embedded in the plane. Due to Whitney [14], this embedding is essentially unique. If one removes an edge $\{i, j\}$ of a triangulation, then there is

a face $\{i, k, j, l\}$ of size four, and if the edge $\{k, l\}$ does not already exist, then one can add $\{k, l\}$ instead of $\{i, j\}$ to obtain a new triangulation. Such an operation is called a *flip*. It is well known [13] that each triangulation can be reached from any other triangulation by a finite number of flips. We consider such *flip sequences* from a fixed given embedded triangulation and also *extended flip sequences* where we allow multiple edges but insist that each of the multiple edges has its own face, namely, the face of size four that is obtained by deleting the edge before the flip.

Let $\mathcal{A} = \{a_0, \dots, a_{3n-6}\}$ be the alphabet where for $i = 1, \dots, 3n - 6$ the flip of edge i is represented by the letter a_i and a nonflip is represented by a_0 . After a flip of edge i , the new edge created will still be numbered i . Some flip sequences may lead to the same triangulation. To capture this we define for an extended flip sequence $w = a_i \dots a_j$, the set C_w as the nonflip a_0 , and the set of all edges except for the four edges having a common face with the last edge a_j after the flips of the sequence w have taken place.

Consider the commuting family $\mathcal{C} = (C_w : w \in \mathcal{A}^*)$ and the set $\mathcal{A}^\ell$ of all flip sequences of length ℓ starting with a given embedded triangulation. Note that if two flip sequences in $\mathcal{A}^\ell$ are equivalent, then the two sequences yield the same triangulation. This can be seen by observing that it does not matter at which point an edge is flipped as long as the edges incident to a common face have not been changed, and $\mathcal{C}$ is chosen in such a way that the order of edges incident to a common face can never be changed. Hence the number of equivalence classes is at least the number of triangulations that can be reached with ℓ flips from the given triangulation. We want to find a lower bound on the number of flips ℓ needed to reach all triangulations from a fixed given triangulation. Since Tutte [12] has shown that the number of triangulations is at least $C^n n!$ for any $C < 256/27$ and sufficiently large n , it suffices to find an upper bound on the number of equivalence classes of $\mathcal{A}^\ell$, as we need more equivalence classes than triangulations to reach all triangulations from a given one. We find such an upper bound in Theorem 2.2, where we show that the number of equivalence classes is at most $\tilde{C}^{|\mathcal{A}|+\ell}$ for some constant $\tilde{C}$. Thus to reach all triangulations from a fixed triangulation we have to ensure that the number of flips ℓ satisfies $\tilde{C}^{3n-5+\ell} \geq C^n n!$. It follows that $\Omega(n \log n)$ flips are needed. This lower bound matches asymptotically the upper bound, as it is known (see [11] and [5]) that $O(n \log n)$ flips suffice. The lower bound has also been shown in [11] with a slightly longer, more complicated proof. In particular, in [11] one needs to consider the dual of a triangulation before one can apply their method.

2. Main result. To prove our main result we first show that the number of words over the alphabet $\{1, \dots, m\}$ that never decrease much can be bounded from above.

LEMMA 2.1. *Let $k \in \mathbb{N}$ be a constant. The number of words $a_1 \dots a_\ell$ of length ℓ over the alphabet $\mathcal{A} = \{1, \dots, m\}$ such that $a_{j+1} - a_j \geq -k$ for all $1 \leq j \leq \ell - 1$ is at most*

$$\binom{m + \ell k + \ell}{\ell} \leq ((k + 1)e)^\ell e^{m/(k+1)},$$

where e denotes the base of the natural logarithm.

Proof. We can view the sequence $a_1, \dots, a_\ell$ as the following walk on the integers starting at 0. We consider $a_1, \dots, a_\ell$ in turn. When considering a_i , $b_i = a_i - a_{i-1} + k$ steps of size 1 are taken up (where we set $a_0 = k$) and then one step of size k is taken down; see Figure 1. Note that $b_i \geq 0$ as $a_{j+1} - a_j \geq -k$ for all $1 \leq j \leq \ell - 1$. Given

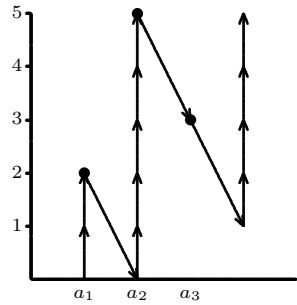


FIG. 1. Sequence $UUDUUUUUDDUUUU$ corresponds to $(a_1, a_2, a_3) = (2, 5, 3)$, where $k = 2$, $m = 5$.

such a walk we can easily reconstruct the sequence $a_1 \dots a_\ell$, as a_i is just the height of the walk before stepping down for the i th time. Hence it remains to count the number of such walks.

Note that when considering the word $a_1, \dots, a_\ell$ we can step up a total of at most $m + \ell k$ times. By adding additional steps up after the last step down, we may assume without loss of generality that there are exactly $m + \ell k$ steps up and ℓ steps down. Therefore the number of such walks is at most the number of ways we can choose the positions of the ℓ steps down out of the $m + \ell k + \ell$ steps and is thus at most

$$\begin{aligned} \binom{m + \ell k + \ell}{\ell} &\leq \left(\frac{e(m + \ell(k + 1))}{\ell} \right)^\ell \\ &\leq \left((k + 1)e \left(\frac{m}{(k + 1)\ell} + 1 \right) \right)^\ell \leq ((k + 1)e)^\ell e^{m/(k+1)}. \end{aligned}$$

(Here we used the well-known facts that $\binom{a}{b} \leq (ea/b)^b$ and $(1 + x) \leq e^x$.) $\square$

THEOREM 2.2. *Let $k \in \mathbb{N}$ be a constant, let $\mathcal{A}$ be a finite alphabet with m letters, and let $\mathcal{A}^\ell$ be a subset of all words in $\mathcal{A}^*$ of length ℓ . If every member C_w of a commuting family $\mathcal{C}$ has at least $m - k$ elements, then there exists a set $\mathcal{S}$ of size at most $((k + 1)e)^\ell e^{m/(k+1)}$ such that for each word $w \in \mathcal{A}^\ell$ there exists a word $w' \in \mathcal{S}$ with $w \sim_{\mathcal{C}} w'$.*

Proof. We shall define an injection from the set of equivalence classes to the set of sequences $b_1, \dots, b_\ell$ of length ℓ with $b_j \in \{1, \dots, m\}$ and the property that $b_{j+1} - b_j \geq -k$. The result then follows by Lemma 2.1.

Let $\mathcal{A} = \{a_1, \dots, a_m\}$. We start with an arbitrary ordering of the letters, say, $a_1 < a_2 < \dots < a_m$. This ordering might change at each step of the following procedure. We assign the following sequence $b_1, \dots, b_\ell$ to a word $w = a_{i_1}, \dots, a_{i_\ell} \in \mathcal{A}^\ell$: At each step j , $b_j \in \{1, \dots, m\}$ is equal to the position of a_{i_j} in the current ordering. Thus $b_1 = i_1$. Given the ordering when we have considered k letters of w , the next ordering is obtained by moving the letters not in $C_{[w]_k}$ that are below a_{i_k} directly below a_{i_k} so that the relative order is conserved; see Table 1. Observe that letters that are above a_{i_k} in the current ordering are not moved.

Note first that given the set $\mathcal{C}$ and the initial ordering $a_1 < a_2 < \dots < a_m$, a sequence $b_1, \dots, b_\ell$ with $b_i \in \{1, \dots, m\}$ for all $1 \leq i \leq \ell$ represents a unique word in $\mathcal{A}^*$. Thus it remains to show that we can choose a word in each equivalence class in

TABLE 1

The change of the ordering with the word *decaf* which is mapped to $(4, 5, 4, 3, 6)$. We start with the ordering a, b, c, d, e, f and consider the letters of the word *decaf* in turn. After considering d we have to move the letters not in C_d that are below d directly behind d . C_d consists of two elements b, f . Thus we have to move a, c directly behind d to obtain b, a, c, d, e, f . After considering de we have to move all the elements below e that are not in $C_{de} = \{a, d\}$ directly behind e . Thus we have to move b and c to obtain a, d, b, c, e, f . We continue and obtain the word $(4, 5, 4, 3, 6)$.

Commuting sets: $C_d = \{b, f\}$, $C_{de} = \{a, d\}$, $C_{dec} = \{b, d\}$, $C_{deca} = \{e, f\}$					
Order	Initially	After:			
		d	de	dec	$deca$
6	f	f	f	f	f
5	e	e	e	e	e
4	d	d	c	c	c
3	c	c	b	a	a
2	b	a	d	b	b
1	a	b	a	d	d

such a way that each word is mapped to a sequence $b_1, \dots, b_\ell$ with $b_{j+1} - b_j \geq -k$ for all $j \in \{1, \dots, \ell - 1\}$. For each equivalence class we choose the word that is assigned the lexicographically smallest sequence, that is, if $w = a_{i_1} \dots a_{i_\ell}$ is chosen and is mapped to $b_1, \dots, b_\ell$, then every word that is equivalent to w is assigned to a sequence that is bigger than $b_1, \dots, b_\ell$ in the lexicographic order.

We now claim that these chosen words are assigned to sequences $b_1, \dots, b_\ell$ with $b_{j+1} - b_j \geq -k$ for all $1 \leq i \leq \ell$. Assume for contradiction that a chosen word $w = a_{i_1} \dots a_{i_\ell}$ is mapped to a sequence $b_1, \dots, b_\ell$ such that $b_{t+1} - b_t < -k$ for some index $t \in \{1, \dots, \ell - 1\}$. Then by the assumption that $|C_{[w]_t}| \geq m - k$ and the choice of our ordering, $a_{i_{t+1}} \in C_{[w]_t}$. Thus w is equivalent to $w' = a_{i_1} \dots a_{i_{t-1}} a_{i_{t+1}} a_{i_t} a_{i_{t+2}} \dots a_{i_\ell}$. But it is clear that $a_{i_1} \dots a_{i_{t-1}}$ is mapped to the same subsequence in both words. Moreover, after we have considered $a_{i_1} \dots a_{i_{t-1}}$, in the current ordering we have $a_{i_{t+1}} < a_{i_t}$, since $a_{i_{t+1}} \geq a_{i_t}$ at this point would imply that $a_{i_{t+1}} \geq a_{i_t}$ in the ordering after $[w]_t$ has been considered. But this contradicts our assumption. Hence w' is mapped to a sequence that is smaller with respect to the lexicographic ordering, a contradiction to the choice of w . $\square$

As we have discussed in the introduction one can consider the set $\mathcal{A}^\ell$ of all flip sequences of length ℓ from a given triangulation T , and construct sets C_w , $w \in \mathcal{A}^*$ with $|C_w| \leq (3n - 5) - 4$ for all $C_w \in \mathcal{C}$ such that every triangulation T' that can be reached from T with at most ℓ flips is represented by at least one equivalence class. For any $C < 256/27$, there are at least $C^n n!$ triangulations of size n , and by Theorem 2.2 we can reach at most $t(\ell) = (5e)^\ell e^{(3n-5)/5}$ triangulations from T with ℓ flips. Thus we have to choose ℓ such that $t(\ell) \geq C^n n!$ in order to achieve that the number of triangulations that can be reached from T with at most ℓ flips is bigger than the number of all triangulations. But $t(\ell) \geq C^n n!$ is equivalent to $\ell \geq \alpha n \log n$ for some constant $\alpha > 0$. Hence we obtain the following result, a more complicated proof of which can be found in [11].

COROLLARY 2.3. *There are labeled triangulations T and T' on n vertices such that $\Theta(n \log n)$ flips are needed to transform T into T' .*

The alert reader might have observed that we did not need all the properties of the equivalence relation $\sim_{\mathcal{C}}$. Thus we can slightly generalize the definition of a commuting family. We consider sets $\mathcal{B} \subseteq \mathcal{A}^*$ where some of the words in $\mathcal{B}$ are *dominated* by others. More precisely, we are given a set $(C'_w \subseteq \mathcal{A} : w \in \mathcal{A}^*)$ that we call a *generalized*

commuting family, and a word $w_1 \in \mathcal{B}$ dominates a word $w_2 \in \mathcal{B}$ if and only if there exists a sequence of words $w'_1, w'_2, \dots, w'_l \in \mathcal{A}^*$ with $w_1 = w'_1$ and $w_2 = w'_l$ such that for each $j = 1, \dots, l-1$, w'_{j+1} can be obtained from w'_j by exchanging two consecutive letters at positions t_j and $t_j + 1$ and the $t_j + 1$ st letter of w'_{j+1} is contained in $C'_{[w'_j]_{t_j}}$. For example if $\mathcal{B}$ consists of all words of length ℓ and $C'_w = \mathcal{A}$ for all $w \in \mathcal{A}^*$, then every word dominates every other word. If $C'_w = \emptyset$ for all $w \in \mathcal{A}^*$, then no word dominates any other word. Finally, for a set $\mathcal{B}$ with generalized commuting family $\mathcal{C}$, a *basic set* is a subset of $\mathcal{A}^*$ such that each word of $\mathcal{B}$ dominates at least one word of the basic set. Trivial basic sets of $\mathcal{B}$ with commuting family $\mathcal{C}$ are supersets of $\mathcal{B}$. The size of a smallest basic set $\mathcal{B}$ with commuting family $\mathcal{C}$ is denoted by $|\mathcal{B}, \mathcal{C}|$. Now, as in Theorem 2.2 one can prove that $|\mathcal{B}, \mathcal{C}| \leq C(k)^{m+\ell}$ if every element of the generalized commuting family contains at least $m - k$ elements.

REFERENCES

- [1] R. L. DOBRUSHIN, *Estimates of semi-invariants for the Ising model at low temperatures*, in Topics in Statistical and Theoretical Physics, Amer. Math. Soc. Transl. Ser. 2, vol. 177, AMS, Providence, RI, 1996, pp. 59–81.
- [2] R. L. DOBRUSHIN, *Perturbation methods of the theory of Gibbsian fields*, in Lectures on Probability Theory and Statistics, P. Bernard, ed., École d'Été de Probabilités de Saint-Flour XXIV—1994, Lecture Notes in Math. 1648, Springer-Verlag, Berlin, 1996, pp. 1–66.
- [3] P. ERDŐS AND L. LOVÁSZ, *Problems and results on 3-chromatic hypergraphs and some related questions*, in Infinite and Finite Sets (Colloq., Keszthely, 1973), Vol. II, Colloq. Math. Soc. János Bolyai 10, North-Holland, Amsterdam, 1975, pp. 609–627.
- [4] D. C. FISHER, *The number of words of length n in a graph monoid*, Amer. Math. Monthly, 96 (1989), pp. 610–614.
- [5] Z. GAO, J. URRUTIA, AND J. WANG, *Diagonal flips in labelled planar triangulations*, Graphs Combin., 17 (2001), pp. 647–657.
- [6] R. KOTECKÝ AND D. PREISS, *Cluster expansion for abstract polymer models*, Comm. Math. Phys., 103 (1986), pp. 491–498.
- [7] V. E. LEVIT AND E. MANDRESCU, *The independence polynomial of a graph—a survey*, in Proceedings of the 1st International Conference on Algebraic Informatics, Aristotle Univ. Thessaloniki, Thessaloniki, 2005, pp. 233–254.
- [8] A. D. SCOTT AND A. D. SOKAL, *The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma*, J. Statist. Phys., 118 (2005), pp. 1151–1261.
- [9] A. D. SCOTT AND A. D. SOKAL, *On dependency graphs and the lattice gas*, Combin. Probab. Comput., 15 (2006), pp. 253–279.
- [10] J. B. SHEARER, *On a problem of Spencer*, Combinatorica, 5 (1985), pp. 241–245.
- [11] D. D. SLEATOR, R. E. TARJAN, AND W. P. THURSTON, *Short encodings of evolving structures*, SIAM J. Discrete Math., 5 (1992), pp. 428–450.
- [12] W. T. TUTTE, *A census of planar triangulations*, Canad. J. Math., 14 (1962), pp. 21–38.
- [13] K. WAGNER, *Bemerkungen zum Vierfarbenproblem*, Jahresber. Deutsch. Math.-Verein., 46 (1936), pp. 26–32.
- [14] H. WHITNEY, *2-isomorphic graphs*, Amer. J. Math., 55 (1933), pp. 245–254.

SPLIT HYPERGRAPHS*

ÁDÁM TIMÁR†

Abstract. Generalizing the notion of split graphs to uniform hypergraphs, we prove that the class of these hypergraphs can be characterized by a finite list of excluded induced subhypergraphs. We show that a characterization by generalized degree sequences is impossible, unlike in the well-known case of split graphs. We also give an algorithm to decide whether a given uniform hypergraph is a split hypergraph. If it is, the algorithm gives a splitting of it; the running time is $O(N \log N)$. These answer questions of Sloan, Turán, and Peled.

Key words. split hypergraph, split graph, excluded induced subgraph, degree sequence, sparse and dense graph

AMS subject classifications. Primary, 05C65; Secondary, 05C85, 68R05

DOI. 10.1137/060675812

1. Introduction, split graphs. We call a graph G *split* if there exists a partition $A \cup B$ of its vertex set so that there is no edge between any two points of A and there is an edge between every two points of B . Such a partition $A \cup B$ of $V(G)$ is called a *splitting*.

The following theorem describes split graphs in two different ways.

THEOREM 1.1. *For a graph G the following are equivalent:*

1. G is split.
2. G does not contain C_4 , C_5 , or $K_2 \cup K_2$ as an induced subgraph.
3. There is an m such that $\sum_{i=1}^m d_i - m(m-1) = \sum_{i=m+1}^n d_i$, where $d_1 \geq d_2 \geq \dots \geq d_n$ are the degrees in G .

The equivalence of 1 and 2 is first shown in [5], while the equivalence of 1 and 3 is in [6]. See also [1] or [4] for history and further references.

The generalizations of many graph theoretical questions to hypergraphs have practical significance. Most frequently, posing these problems for hypergraphs makes them rather difficult. For example, testing the existence of a perfect matching can be solved by a well-known polynomial algorithm for graphs. Generalizing it to 3-uniform hypergraphs, we get an NP-complete problem (see [7], [3]).

The class of line graphs can be characterized by a finite list of excluded induced subgraphs—on the other hand, there are infinitely many pairwise nonisomorphic graphs with the property that none of them is the line graph of any 3-uniform hypergraph, but every proper subgraph of any of them is (see [8]). Studying extremal questions for graphs turned out to be very fruitful. The generalization of this theory to hypergraphs comes up against serious difficulties.

Recent research in learning theory has led to a generalization of the notion of split graphs (see [9]). Call a k -uniform hypergraph G a *split hypergraph* if its vertices can be partitioned into two classes so that every k -tuple in one of the classes is an edge and no k -tuple in the other class is an edge in G . We call such a partition a *splitting*.

*Received by the editors November 23, 2006; accepted for publication (in revised form) February 28, 2008; published electronically July 3, 2008. This research was done when the author was at the University of Szeged [10] and was partially supported by Hungarian National Foundation for Scientific Research grant OTKA49398.

<http://www.siam.org/journals/sidma/22-3/67581.html>

†Department of Mathematics, University of British Columbia, 121-1984 Mathematics Road, Vancouver, BC V6T1Z1, Canada (timar@math.ubc.ca, <http://www.math.ubc.ca/~timar/>).

Note that our definition allows one of the classes to have size smaller than k , in which case the requirement on that class automatically holds.

When G is not a split hypergraph, we say that it is not split or nonsplit. For a hypergraph G , $V(G)$ will denote its vertex set and $E(G)$ its edge set. From now on, n denotes $|V(G)|$ and N is $|V(G)| + |E(G)|$.

The following question arose naturally: Can we state something similar to Theorem 1.1 for the case of split hypergraphs? Peled asked whether the class of k -uniform split hypergraphs can be characterized by a finite list of excluded induced subhypergraphs. We give an affirmative answer to this question in section 2. Section 3 gives examples of minimal nonsplit hypergraphs in the 3-uniform case. We did not strive for completeness of the list. Nevertheless, it may show that the number of excluded induced subhypergraphs increases considerably as we go from graphs to 3-uniform hypergraphs. In section 4 we examine if the connection between degree sequences and the split property of graphs can be generalized. We find a negative answer.

In their paper [9], where they apply split hypergraphs in learning theory, Sloan and Turán gave an algorithm that decides whether a k -uniform hypergraph is split and gives a splitting if possible. The running time of their algorithm is $O(n^{2k})$; they mention that finding a more efficient algorithm is an open question. Feder et al. [2] introduce the notion of a sparse and dense partition of a graph, which is a kind of generalization of a splitting. Their definition can be extended to hypergraphs and the algorithm they present (in Theorem 3.1) can be used to find a splitting (or all splittings) of a uniform hypergraph. The running time of their algorithm in this context is the same as the time requirement of the one given by Sloan and Turán. In section 5 we give an $O(N \log N)$ algorithm for the problem. Our method can also be used to give an algorithm for sparse and dense partitions that in certain cases is faster than the one in [2].

The restriction of an edge to an $X \subset V(G)$ is its intersection with X . Minimality of a graph with respect to a certain property will always mean that no *induced* proper subgraph of it has the property. As usual, $X \Delta Y$ will stand for the symmetric difference between sets X and Y . When $x \in V(G)$, we sometimes simply use x instead of $\{x\}$ in set operations.

2. The existence of a finite characterization. In this section we prove our main theorem. First we state two lemmas. The following lemma holds for hypergraphs with edges of sizes at most k , which is a relaxation of k -uniformity, and is convenient for our inductive proof.

LEMMA 2.1. *Let k and m be positive integers and Γ a hypergraph with edges of size at most k . Then Γ has at most k^m different minimal cover sets of at most m points. Furthermore, there is an algorithm of time requirement $O(|V(\Gamma)| + |E(\Gamma)|)$ that outputs a list of these minimal cover sets.*

Because of the short running time of our algorithm, some more details about the implementation could be added. However, the design of a data structure and the implementation of the algorithm do not match the texture of our paper, so we leave these to the interested reader.

Proof. We prove the first part of the claim by induction on m . The case $m = 1$ is obvious. Suppose that for $m < r$ the statement is true, and let Γ be any hypergraph with edges of size at most k . Let H be an arbitrary edge. Any minimal cover set of Γ with at most m points consists of a point x of H and a minimal cover set of $\Gamma|_{V(\Gamma) \setminus \{x\}}$ with at most $m - 1$ elements. Applying the induction hypothesis to the latter one and noting that x could be chosen in k different ways, we get the statement for $m = r$.

The algorithm will imitate the proof in the previous paragraph. For $i = 1, \dots, m$, pick the first edge e_i that has not been deleted yet, a vertex v_i of e_i , delete the edges that contain v_i , and let $i := i + 1$. The output is $\{v_1, \dots, v_m\}$. Since the number of minimal cover sets is bounded by a constant, we can list all of them by systematically going through all possible ways of picking v_i in the sequence of e_i , and still use only linear time. $\square$

In what follows, for the sake of simplicity, the edges of the k -uniform hypergraph G will be referred to as *type 0 edges*, and the k -tuples not contained in $E(G)$ as *type 1 edges*. (Thus we identify G with a labeling of the edges of a complete k -uniform hypergraph by numbers 0 and 1.)

Consider some minimal nonsplit hypergraph G . Minimality implies that for any $x \in G$ there is a splitting (A_0^x, A_1^x) of $G \setminus x$; i.e., A_0^x induces only edges of type 0, and A_1^x induces only edges of type 1. (A class that contains only edges of type i will sometimes be called a *type i class*.)

LEMMA 2.2. *Fix a splitting (A_0^v, A_1^v) of $G \setminus v$ for every $v \in V(G)$. If $x, y \in V(G)$ are vertices in a minimal nonsplit hypergraph G , then $|A_i^x \cap A_{1-i}^y| < k$ and $|A_i^x \Delta A_i^y| \leq 2k$ for $i = 0, 1$.*

Proof. The first inequality follows by the fact that a k -tuple cannot be contained in the type i class of a splitting (of $G \setminus x$) and a type $1 - i$ class of a splitting (of $G \setminus y$) at the same time. For the second inequality, use the facts

$$A_i^x \setminus A_i^y \subset A_i^x \cap (A_{1-i}^y \cup \{y\}) \quad \text{and} \quad |A_i^x \cap (A_{1-i}^y \cup \{y\})| \leq k$$

by the first inequality. Conclude similarly with x and y interchanged to obtain the bound on the symmetric difference. $\square$

In the proof, we shall consider sets subindexed by 0 or 1. Such a set, of subindex i , will always be a class of edges of type i , or some “preliminary set” in constructing a class only with edges of type i . This notation will provide us with the following convenience in terminology. Given some set $S \subset V(G)$ by a name subindexed by i ($i = 0, 1$), let the *bad edges* of S be the edges of type $1 - i$ on S .

THEOREM 2.3. *If a k -uniform nonsplit hypergraph has more than $4(k+1)k^{4k(k+1)+1}$ vertices, then it has a proper induced subhypergraph that is nonsplit.*

Fix k . We shall give an algorithm. Its brief description is given in the next two paragraphs. The input is $G, x, (A_0^x, A_1^x), \alpha, \beta$, where G is a minimal nonsplit hypergraph, $x \in V(G)$, (A_0^x, A_1^x) is a splitting of $G \setminus x$, $\alpha \in \{1, \dots, 4k\}$, and $\beta \in \{1, \dots, k+1\}$. The algorithm either stops without any output, or outputs a splitting of $G \setminus y$ for some $y \in V(G)$. It has a randomized part to simplify its description.

At the beginning our algorithm adds x to A_0^x , creating some type 1 edges (bad edges) in $B_0 := A_0^x \cup x$. There is another class B_1 , set to be A_1^x at the start. Then we choose a random minimal cover set of size at most $k+1$ for the bad edges and move it over from B_0 to B_1 . In general, at each step at most one of B_0 and B_1 contains any bad edges, depending on the parity of the step. The algorithm always chooses uniformly a random minimal cover set of size at most $k+1$ for these bad edges, and consisting of points that have not been moved between the two classes yet, and any such set is chosen with positive probability. (If there is no such set, then the algorithm fails.) Move this minimal cover set to the other class. When we arrive at the α th step, let the β th element of the minimal cover set be called y . (We may assume that there is some fixed ordering on any subset of the vertices, for example, the one given by the listing of the vertices in the input.) Remove y from the minimal cover set (and hence from both classes in the future) and continue the procedure. We run the algorithm for

at most $4k$ steps. With positive probability, at some step, (B_0, B_1) will be a splitting of $G \setminus y$, and this is the output.

The inputs α and β , together with the choices that the random part of the algorithm makes, determine y . Using Lemma 2.1, we shall give a bound on the number of possible choices and finally get the general upper bound in Theorem 2.3 for the number of vertices in G .

Here is a formal presentation of the algorithm.

ALGORITHM 1.

FUNCTION $\text{BadEdges}(i, B_i)$ returns the list of type $1-i$ (bad) edges of B_i .

FUNCTION $\text{MinCoverSet}(R, L)$ returns a uniformly chosen random minimum cover set of R of size at most $k+1$ that is disjoint from L . If there is no such set, then the algorithm ends and outputs ‘‘Unsuccessful.’’

Start of Algorithm

INPUT: a k -uniform hypergraph G , vertex x of G , a splitting (A_0^x, A_1^x) of $G \setminus x$, numbers $\alpha \in \{1, \dots, 4k\}$ and $\beta \in \{1, \dots, k+1\}$.

$B_0 := A_0^x \cup \{x\}$, $B_1 := A_1^x$, $L := \emptyset$, $i := 0$, $j := 1$, $\ell := \emptyset$

while $j \leq 4k$ **do** {

$R := \text{BadEdges}(i, B_i)$

if $R = \emptyset$ **then** **output** (B_0, B_1) , **end of algorithm**

$\ell := \text{MinCoverSet}(R, L)$

$B_i := B_i \setminus \ell$, $L := L \cup \ell$

if $j = \alpha$ **then** { **if** $|\ell| < \beta$ **then** **output** $(\text{‘‘Unsuccessful’’})$, **end of algorithm**,

else $\ell := \ell \setminus \{y\}$, where y is the β th element of ℓ }

$B_{1-i} := B_{1-i} \cup \ell$

$i := 1 - i$, $j := j + 1$

}

output $(\text{‘‘Unsuccessful’’})$, **end of algorithm**

LEMMA 2.4. Fix G , x , and (A_0^x, A_1^x) in the input. For any $y \in V(G)$, Algorithm 1 outputs a splitting of $G \setminus y$ with positive probability for some values of α, β .

Proof. Fix y and a splitting (A_0^y, A_1^y) of $G \setminus y$. None of the bad edges in $A_0^x \cup \{x\}$ can be contained in A_0^y , so there exists some minimal cover set of these bad edges that is contained in $A_1^y \cup \{y\}$. (Such a minimal cover set is chosen by MinCoverSet with positive probability.) This minimal cover set cannot have more than $k+1$ elements, because then a k -tuple would be present in both A_0^x and A_1^y , contradicting Lemma 2.2. Since the points that we moved are in $A_1^y \cup \{y\}$ by our choice, they get to their final class by the move, so we may assume that they are not moved in later steps of the algorithm. Hence, when the algorithm chooses a minimal cover set of the bad edges of size at most $k+1$ and moves this set from $A_0^x \cup \{x\}$ to A_1^x (call the resulting sets C_0 and C_1 , respectively), we have $|C_0 \Delta A_0^y| < |A_0^x \Delta A_0^y|$, and also $|C_1 \Delta A_1^y| \leq |A_1^x \Delta A_1^y|$. These inequalities are crucial in that we succeed after at most $4k$ repetitions. When y is in ℓ , if α is equal to the actual j and y is the β th element in ℓ , then y is removed from ℓ , and its removal ensures that the remaining graph has a splitting. (This possible removal causes the second inequality above to be not necessarily strict.)

Similarly, denote by D_0 and D_1 the sets B_0 and B_1 , respectively, at some step, and let C_0 and C_1 be B_0 and B_1 , respectively, in the next step. Let ℓ be the set that was moved, as in the algorithm; we may assume by symmetry that it was moved from D_0 to D_1 . So, $C_0 = D_0 \setminus \ell$ and $C_1 = D_1 \cup (\ell \setminus y)$. There were some bad edges in

D_0 because the algorithm did not stop. None of these bad edges can be contained in A_0^y , so there exists some minimal cover set of the bad edges that is contained in A_1^y . We may assume that the elements of this minimal cover set are points that have not been moved at any previous step (i.e., they are not in L), because we are considering the outcome of the algorithm (of positive probability), when the moved vertices must reach the type of class that contains them in (A_0^y, A_1^y) by the move. (The only vertex not in any of A_0^y or A_1^y is y , but it is removed from ℓ if α and β are suitable.) The algorithm chooses such a minimal cover set ℓ with positive probability, and then we have

$$|C_0 \Delta A_0^y| < |D_0 \Delta A_0^y| \text{ and } |C_1 \Delta A_1^y| \leq |D_1 \Delta A_1^y|.$$

We conclude that $|B_0 \Delta A_0^y| + |B_1 \Delta A_1^y|$ decreases in each iteration step by at least 1 with positive probability. At some step, y has to be contained in ℓ , since G itself is not split. If α and β are suitable, then y is removed from ℓ (and the future $B_0 \cup B_1$). Then the algorithm can stop for two reasons. The first is when it outputs “Unsuccessful” because of the random choices made by **MinCoverSet**. However, we have seen that, for a certain sequence of the random choices made by **MinCoverSet** along the iteration steps, there always exists a minimal cover set of the bad edges consisting of points not moved yet, as long as the set of the bad edges is nonempty. On the other hand, when the set of bad edges becomes empty, the actual (B_0, B_1) is a splitting. This happens either when the algorithm has already run for $4k$ steps ($j > 4k$), in which case $|B_0 \Delta A_0^y| + |B_1 \Delta A_1^y|$ has to be 0 (at the first step, $|B_0 \Delta A_0^y| + |B_1 \Delta A_1^y| = |A_0^x \Delta A_0^y| + |A_1^x \Delta A_1^y| \leq 4k$ by Lemma 2.2), and we achieved the splitting (A_0^y, A_1^y) as an output, or when there are no bad edges at some earlier step. In that case the actual (B_0, B_1) does not necessarily coincide with the fixed splitting (A_0^y, A_1^y) , but it is still a splitting of $G \setminus y$. $\square$

Proof of Theorem 2.3. Fix again G , x , and (A_0^x, A_1^x) . According to Lemma 2.4, for every $y \in V(G)$ there is some input α, β such that with positive probability Algorithm 1 gives a splitting of $G \setminus y$. The corresponding sequence of ℓ 's (along the iteration steps of the algorithm) and the α, β determine the output and y . There are at most k^{k+1} choices for ℓ in each step, by Lemma 2.1, and the rest of the algorithm is deterministic. So there are at most $k^{4k(k+1)}$ possible sequences of ℓ 's along the iteration steps of the algorithm. There are $4k(k+1)$ possible inputs α, β . Hence there are at most $k^{4k(k+1)} 4k$ possible y 's, which is our upper bound for the number of vertices in a minimal nonsplit k -uniform hypergraph. $\square$

A straightforward corollary is our main theorem.

THEOREM 2.5. *The family of k -uniform split hypergraphs can be characterized by a finite set of forbidden induced subhypergraphs.*

3. 3-uniform minimal nonsplit hypergraphs. Although Theorem 2.3 gave a very rough upper bound on the size of minimal nonsplit hypergraphs, in this section we illustrate with a few examples that the list of excluded induced subhypergraphs increases considerably compared to the case of graphs.

For $|V(G)| = 6$, G is nonsplit if and only if for any edge its complement is also an edge and if the subhypergraph induced by any four points is neither the empty nor the complete 3-uniform hypergraph. It is clear that these are minimal nonsplit hypergraphs too.

We get examples of minimal nonsplit hypergraphs for the case of $|V(G)| = 7, 8, 9$ in the following manner. Write the vertices along a circle and let the edges be exactly those 3-tuples that consist of three consecutive nodes along the circle. G is not split,

since we could put into the type 0 class at most 3 vertices, but then the rest contains a type 0 edge (and similarly if we put only 2 arbitrary points in the type 0 class). However, dropping any vertex, two suitably chosen points can cut the remaining arc into parts with at most two vertices. In the case of 8 (respectively, 9) vertices, we can erase 1 (respectively, 1, 2, 3, or 4) edges from the hypergraph just described, and similarly, we still get a minimal nonsplit hypergraph. The argument is nearly the same for these new graphs.

We have found minimal 3-uniform nonsplit hypergraphs on vertex sets of 11 and 12 elements, but the details are lengthy. The examples above may show that even in the 3-uniform case there are at least 100 (and probably even many more) minimal excluded subhypergraphs.

4. Degree sequences of hypergraphs. The question arises of whether a characterization like that in part 3 of Theorem 1.1 can be given for split hypergraphs. Since the equivalence of 3 and 1 provides us with an $O(n)$ algorithm to decide whether a graph is split or not, a characterization for split hypergraphs using degree sequences could be promising.

PROPOSITION 4.1. *There are 3-uniform hypergraphs G and G' on the vertex set $\{1, \dots, n\}$ such that G is split and G' is nonsplit; moreover, $d_i = d'_i$ and $\delta_{ij} = \delta'_{ij}$ for any $i, j \in \{1, 2, \dots, n\}$. Here d_i (d'_i) stands for the degree of i in G (G'); δ_{ij} (δ'_{ij}) is the number of edges containing both i and j in G (G').*

Proof. Call a 6-element subset of the vertices of a 3-uniform hypergraph H exchangeable if there is a list $a_1, a_2, a_3, a_{-1}, a_{-2}, a_{-3}$ of these vertices so that $\{a_1, a_2, a_3\}$, $\{a_1, a_{-2}, a_{-3}\}$, $\{a_{-1}, a_2, a_{-3}\}$, $\{a_{-1}, a_{-2}, a_3\}$ are edges of H , and the other 3-tuples not containing a_i and a_{-i} together for some i are not in H . We say that we exchange the exchangeable point set $\{a_1, a_2, a_3, a_{-1}, a_{-2}, a_{-3}\}$ when we erase the four edges described above and add the 3-tuples $\{a_{-1}, a_{-2}, a_{-3}\}$, $\{a_{-1}, a_2, a_3\}$, $\{a_1, a_{-2}, a_3\}$, and $\{a_1, a_2, a_{-3}\}$ to $E(H)$. We may think of the exchangeable 6 vertices as the nodes of an octahedron. Then “exchangeable” means that, if the nodes were labeled appropriately, the four 3-tuples in $E(H)$ are determined by four faces whose pairwise intersection is one point. “Exchanging” means that we replace these 3-tuples in $E(H)$ by the other four faces of the octahedron. One can easily see that these are well defined; the listing of the 6 vertices is essentially unique. On the other hand, the degrees and the number of edges incident to two points do not change after applying an exchange. Thus if we get a nonsplit hypergraph from a split hypergraph after serial exchanging, then the proposition follows.

Now, denote $X = \{1, 2, 3, 4\}$, $Y = \{5, 6, 7, 8\}$. Define G on the vertex set $\{1, \dots, 14\}$ as follows. Every 3-tuple of $X \cup \{9, 10, 11\}$ is red, every 3-tuple of $Y \cup \{12, 13, 14\}$ is blue; if a 3-tuple intersects both X and $\{12, 13, 14\}$, then it is blue; if a 3-tuple intersects both Y and $\{9, 10, 11\}$, then it is red. Moreover, the 3-tuples of $\{9, 10, \dots, 14\}$ that have not been defined yet get colors so that $\{9, 10, \dots, 14\}$ give an exchangeable set of points, as in the first sentence of this proof, with $(a_1, a_2, a_3, a_{-1}, a_{-2}, a_{-3}) = (9, 10, 11, 12, 13, 14)$. The colors of the remaining 3-tuples are arbitrary. G is obviously split with color classes $X \cup \{9, 10, 11\}$ as red and $Y \cup \{12, 13, 14\}$ as blue. Let G' be the hypergraph obtained from G by exchanging $\{9, \dots, 14\}$. In G' , $\{9, 10, 11\}$ has become blue, so at least one of its points must be put into the blue class. But there, this point gives a red edge with any two points of Y . Since at least two points of Y must remain in the blue class, we conclude that G' is indeed nonsplit. $\square$

For an arbitrary k , one can give a k -uniform counterexample with a construction similar to the one above by using the k -dimensional cross-polytope $\{x \in \mathbb{R}^k : \|x\|_1 \leq 1\}$ (the “octahedron” of dimension k).

5. Algorithm for splitting.

THEOREM 5.1. *For an arbitrary, but fixed, k there is an algorithm that decides if a G k -uniform hypergraph is a split hypergraph and gives a splitting when it is. The running time is $O(N \log N)$ (where we defined $N = |V(G)| + |E(G)|$).*

As in the case of Lemma 2.1, we mention that the running time is understood with an appropriately chosen data structure. Let us also point out that the constant in the $O(N \log N)$ bound is exponential in k .

The following lemma is used repeatedly to obtain splittings of larger and larger subhypergraphs of G .

LEMMA 5.2. *Let H be a k -uniform hypergraph, and let (V^1, V^2) be a partition of $V(H)$. Denote by H^1 and H^2 the subhypergraphs of H induced by V^1 and V^2 , respectively. Suppose that (A_0, A_1) is a splitting of H and (A_0^i, A_1^i) is a splitting of H^i ($i = 1, 2$). Then $|A_j \setminus (A_j^1 \cup A_j^2)| \leq 2k - 2$, $j = 0, 1$.*

Proof. If the lemma did not hold, then some k -tuple in A_j would be present in A_{1-j}^1 or A_{1-j}^2 ($j \in \{0, 1\}$), contradicting the fact that both are monochromatic of different types. $\square$

We present an algorithm whose existence proves Theorem 5.1. Call it Algorithm 2. We do not give the pidgin Pascal program this time, because the main structure of the algorithm is very simple, and the more particular elements are similar to Algorithm 1 (except that the random part is replaced by a deterministic choice, as we shall see). Let us start with an overview of the algorithm. We use notation $\langle m \rangle$ for the modulo 2 value of an integer m .

The crucial component of our algorithm is a subroutine called **Tree**. Its input is (H, X_0, X_1) , where H is a k -uniform hypergraph and (X_0, X_1) partitions $V(H)$. **Tree** will create labels for the vertices of a rooted tree T of depth $2k - 2$, where each inner node has k^{2k-2} children. For each vertex, we assume that there is some ordering on its children, so we can talk about the “ m th child” (when $m \leq k^{2k-2}$). Similarly, we assume that for any set of k -tuples of G , there is some fixed ordering on the minimal cover sets of size at most $2k - 2$ of these k -tuples, so we can talk about the “ m th minimal cover set” (when m is small enough). Note that such an ordering of the minimal cover sets can be easily implemented by an algorithm that is linear in the input size, using Lemma 2.1. Finally, fix some ordering on the vertices of T so that vertices at a smaller depth precede vertices at a larger depth.

Say that the root *has depth* 0, its children *have depth* 1, etc. The root r will have label $(X_0^r, X_1^r) := (X_0, X_1)$. If X_0^r contains some edge of type 1 in G , let $s := 0$. Otherwise let $s := 1$.

Suppose that v is a vertex in T , whose parent is u , and v is at depth g . Assume that v is the m th child of u . Let v get label (ℓ_v, X_0^v, X_1^v) , where ℓ_v is the m th minimal cover set of size at most $2k - 2$ of the bad edges in $X_{\langle g+s \rangle}^u$. Let $X_{\langle g+s \rangle}^v := X_{\langle g+s \rangle}^u \setminus \ell_v$ and $X_{\langle 1+g+s \rangle}^v = X_{\langle 1+g+s \rangle}^u \cup \ell_v$. If m is such that there is no m th minimal cover set of size at most $2k - 2$ for the actual bad edges, then delete v and all its offspring from T . If there are no bad edges in $X_{\langle g+s \rangle}^u$, then **Tree** outputs the first such (X_0^u, X_1^u) (in the ordering of the vertices), which is a splitting of the input graph H . If there are bad edges for all the leaves, then **Tree** outputs “No,” indicating H is not split.

Algorithm 2 receives input G , a k -uniform hypergraph. In the preparatory step it partitions $V(G)$ to sets $A_1, \dots, A_{\lceil n/k-1 \rceil}$, where each A_i has $k - 1$ elements, with

the possible exception of $A_{\lceil n/k-1 \rceil}$. Then we call the iteration part of the algorithm with input $((A_1, \emptyset), \dots, (A_{\lceil n/k-1 \rceil}, \emptyset))$.

In each cycle of the iteration part of Algorithm 2 there is an incoming list $((X_0^1, X_1^1), (X_0^2, X_1^2), \dots, (X_0^m, X_1^m))$, where (X_0^i, X_1^i) is a splitting of $G|_{X_0^i \cup X_1^i}$, and the $(X_0^i \cup X_1^i)_{i=1}^m$ give a partition of $V(G)$. Now form pairs $(Y_0^i, Y_1^i) := (X_0^{2i-1} \cup X_0^{2i}, X_1^{2i-1} \cup X_1^{2i})$ as $i = 1, \dots, \lfloor m/2 \rfloor$; if m is odd, then define $(Y_0^{\lceil m/2 \rceil}, Y_1^{\lceil m/2 \rceil}) := (X_0^m, X_1^m)$. Now, call **Tree** for each $i \in \{1, \dots, \lceil m/2 \rceil\}$, with input (Y_0^i, Y_1^i) . If any of these returns “No,” then Algorithm 2 ends and outputs “ G has no splitting.” Otherwise, if **Tree** returns (A_0^i, A_1^i) at the i th call, then the input for the next iteration step is $((A_0^1, A_1^1), (A_0^2, A_1^2), \dots, (A_0^{\lceil m/2 \rceil}, A_1^{\lceil m/2 \rceil}))$. The last cycle of the iteration is when $m = 2$, that is, when we call **Tree** only once. Suppose then that it returns the pair (A_0^1, A_1^1) . In this case Algorithm 2 outputs this pair, which is a splitting of G .

LEMMA 5.3. Algorithm 2 provides a splitting of G , if there is any, and answers “ G has no splitting” if G is nonsplit.

Proof. We shall prove that **Tree** indeed decides if there is a splitting of the hypergraph that it receives while Algorithm 2 is running, and outputs a splitting if there is one. (Note that the hypergraphs that **Tree** receives in Algorithm 2 have a very specific form.) Then it is clear that the iteration steps give splittings for families of bigger and bigger graphs and finally for G , if G is split. So, as in Lemma 5.2, let H be a k -uniform hypergraph, and let (V^1, V^2) be a partition of $V(H)$. Denote by H^1 and H^2 the subhypergraphs of H induced by V^1 and V^2 , respectively. Suppose that (A_0^i, A_1^i) is a splitting of H^i ($i = 1, 2$). We need to show that if the input is $(H, A_0^1 \cup A_0^2, A_1^1 \cup A_1^2)$, then **Tree** finds a splitting of H , if there is any, and answers “No” if H is not split. Note that the inputs that **Tree** can get in Algorithm 2 indeed have this special form.

Suppose that H has a splitting; fix one such (A_0, A_1) . The process of finding a splitting of H corresponds to a walk on T , starting from the root and moving towards a child in each step. Each step of the walk corresponds to finding a minimal cover set for the bad edges in the class actually being examined and putting this set into the other class (starting with the two classes from the input of **Tree**). By the same argument as in Lemma 2.4 (now using the analogue Lemma 5.2 of Lemma 2.2), we shall conclude that if there is a splitting, we get it in $\leq 2k - 2$ steps, or equivalently, at the latest when arriving at a leaf of T .

To be more detailed, suppose that $(A_0^1 \cup A_0^2, A_1^1 \cup A_1^2) =: (B_0, B_1)$ is not yet a splitting (otherwise we are done; the algorithm outputs the splitting after making the label for the root). Then there are type 1 edges in B_0 , or there are type 0 edges in B_1 (call any of these a *bad edge*). We may assume by symmetry that B_0 has bad edges. Then some minimal cover set of these bad edges is contained in A_1 , and the root has a child v such that ℓ_v is this minimal cover set. The iteration part of Algorithm 2 moves over ℓ_v from B_0 to B_1 , hence creating X_0^v and X_1^v . We have $|X_0^v \Delta A_0| + |X_1^v \Delta A_1| < |B_0 \Delta A_0| + |B_1 \Delta A_1|$, similarly to the displayed line in the proof of Lemma 2.4. Note that after the first step of the walk on T only one of X_0^v and X_1^v can have bad edges.

Proceed similarly as follows: When the walk in T is in u and the bad edges are in X_i^u ($i \in \{0, 1\}$), there is a minimal cover set for these bad edges that is contained in A_{1-i} , and such that it is disjoint from $\bigcup \ell_w$, where w ranges through the ancestors of u . (This latter assumption can be made because in each step we moved points that got to their “final class” in (A_0, A_1) by this move.) Some child v of u is such that this

minimal cover set is ℓ_v , and for the arising (X_0^v, X_1^v) we again have

$$|X_0^v \Delta A_0| + |X_1^v \Delta A_1| < |X_0^u \Delta A_0| + |X_1^u \Delta A_1|.$$

Since for the root r we have $|X_0^r \Delta A_0| + |X_1^r \Delta A_1| \leq 2k - 2$ by Lemma 5.2, in at most $2k - 2$ steps we necessarily arrive to a vertex where the two cases contain no bad edges. At this step **Tree** outputs the two classes, which is a splitting of H .

If H is not split, then it is clear that there are bad edges in some class for any partition of the vertex set, so **Tree** will eventually output “No.” $\square$

LEMMA 5.4. *The algorithm runs in $O(N \log N)$ time.*

Proof. First, **Tree** runs in linear time, because the size of the underlying tree is a constant (determined by k), and the labels consisting of minimal cover sets can be constructed in a number of steps that is linear in the input of **Tree**, using the method in Lemma 2.1.

In each iteration step, each element of $V(G) \cup E(G)$ is present in at most one of the subhypergraphs that some **Tree** subroutine receives, and then in each iteration cycle there is a linear number of extra steps when the iteration cycle creates what it returns from the outputs of the **Tree** subroutines. This is $O(N)$, and there are $O(\log N)$ iteration steps, so we get the claim. $\square$

Acknowledgments. I am grateful to Péter Hajnal for his support and suggestions. I also thank Csaba Biró for his comments on the manuscript.

REFERENCES

- [1] A. BRANDSTÄDT, V. B. LE, AND J. P. SPINRAD, *Graph Classes: A Survey*, SIAM Monogr. Discrete Math. Appl. 3, SIAM, Philadelphia, 1999.
- [2] T. FEDER, P. HELL, S. KLEIN, AND R. MOTWANI, *List partitions*, SIAM J. Discrete Math., 16 (2003), pp. 449–478.
- [3] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability. A Guide to the Theory of NP-Completeness*, A Series of Books in the Mathematical Sciences, W. H. Freeman and Co., San Francisco, CA, 1979.
- [4] M. C. GOLUBIC, *Algorithmic Graph Theory and Perfect Graphs*, Ann. Discrete Math. 57, Elsevier, Amsterdam, 2004.
- [5] A. GYÁRFÁS AND J. LEHEL, *A Helly-type problem in trees*, in Combinatorial Theory and Its Applications, II (Proc. Colloq., Balatonfüred, 1969), North-Holland, Amsterdam, 1970, pp. 571–584.
- [6] P. L. HAMMER AND B. SIMEONE, *The Splittance of a Graph*, Res. Report CORR 77-39, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON, 1977.
- [7] R. M. KARP, *Reducibility among combinatorial problems*, in Proceedings of the IBM Symposium on Complexity of Computer Computations, Thomas J. Watson Research Center (Yorktown Heights, NY, 1972), Plenum Press, New York, 1972, pp. 85–103.
- [8] R. N. NAIK, S. B. RAO, S. S. SHRIKHANDE, AND N. M. SINGHI, *Intersection graphs of k -uniform linear hypergraphs*, European J. Combin., 3 (1982), pp. 159–172.
- [9] R. H. SLOAN AND GY. TURÁN, *Learning from incomplete boundary queries using split graphs and hypergraphs*, extended abstract in Computational Learning Theory, Lecture Notes in Comput. Sci. 1208, Springer, Berlin, 1997, pp. 38–50.
- [10] Á. TIMÁR, *Split Hypergraphs*, Diploma thesis, University of Szeged, Szeged, Hungary, 2000 (in Hungarian).

DETECTING A THETA OR A PRISM*

MARIA CHUDNOVSKY[†] AND ROHAN KAPADIA[‡]

Abstract. A theta in a graph is an induced subgraph consisting of two nonadjacent vertices joined by three disjoint paths. A prism in a graph is an induced subgraph consisting of two disjoint triangles joined by three disjoint paths. This paper gives a polynomial-time algorithm to test whether a graph has an induced subgraph that is either a prism or a theta.

Key words. induced subgraph, detection algorithm, theta, prism

AMS subject classification. 05C85

DOI. 10.1137/060672613

1. Introduction. The vertex set of a graph G is denoted by $V(G)$ and the edge set by $E(G)$. All graphs in this paper are simple. A *cycle* is a connected subgraph in which all vertices have degree two. A *path* P in G is an induced connected subgraph of G such that either P is a 1-vertex graph, or two vertices of P have degree one and all others have degree two. (Please note that we are using a nonstandard definition.) The *ends* of a path are the two vertices with degree one in the path. If $X \subseteq V(G)$, the subgraph with vertex set X and edge set all edges of G with both ends in X is denoted $G[X]$ and called the subgraph *induced* on X .

A *prism* is a graph consisting of two disjoint triangles $\{a_1, a_2, a_3\}$ and $\{b_1, b_2, b_3\}$ together with three paths P_1 , P_2 , and P_3 such that the union of every two of P_1 , P_2 , and P_3 is an induced cycle. A *theta* is a graph consisting of two nonadjacent vertices s and t , and three paths P , Q , and R with ends s and t , such that the union of every two of P , Q , and R is an induced cycle. The vertices s, t are called the *ends* of the theta. Finally, a *pyramid* is a graph consisting of a triangle $\{b_1, b_2, b_3\}$, a vertex $a \notin \{b_1, b_2, b_3\}$, and, for $i = 1, 2, 3$, a path P_i between a and b_i , such that a is adjacent to at most one of b_1, b_2, b_3 and the union of every two of P_1, P_2 , and P_3 is an induced cycle. An induced subgraph H of G is a *prism*, *theta*, or *pyramid* in G if H is a prism, theta, or pyramid, respectively. We say that a graph G *contains* a prism, a theta, or a pyramid if some induced subgraph of G is a prism, a theta, or a pyramid, respectively. A *hole* in a graph is an induced cycle of length at least four. A hole is *odd* if it has an odd number of vertices, and is even otherwise. Given a graph G , the *complement* $\overline{G}$ of G is the graph with the same vertex set as G , and such that two vertices are adjacent in $\overline{G}$ if and only if they are nonadjacent in G . A graph G is called *Berge* if neither G nor $\overline{G}$ contains an odd hole.

In [1] a polynomial-time algorithm is presented to test whether a graph contains a pyramid. The main result of [1] is a polynomial-time algorithm to test if a graph is Berge. Since every pyramid contains an odd hole, no Berge graph contains a pyramid, and testing for pyramids is the first step in testing for Berge-ness. Similarly to

*Received by the editors October 17, 2006; accepted for publication (in revised form) February 11, 2008; published electronically July 16, 2008.

<http://www.siam.org/journals/sidma/22-3/67261.html>

[†]IEOR, Columbia University, 500 W. 120th Street, New York, NY 10027 (mchudnov@columbia.edu). The research of this author was conducted during the period the author served as a Clay Mathematics Institute Research Fellow at Princeton University.

[‡]Department of Mathematics, McGill University, Montreal, Quebec H3A 2K6, Canada (rkapadia@math.mcgill.ca). The research of this author was conducted at Princeton University.

pyramids, every theta and every prism contains an even hole, and therefore a graph with no even holes contains no theta or prism. In [4] Maffray and Trotignon prove that the problem of deciding if a graph contains a prism is NP-complete, and in [2] Chudnovsky and Seymour show that testing if a graph contains a theta can be done in polynomial time. However, the complexity of deciding if a graph contains a prism or a theta remains open (and this is the analogue of the pyramid problem for graphs with no even holes). Our main result here is a polynomial-time algorithm that, given a graph G , decides if G contains either a prism or a theta. The algorithm is described in section 5, and its running time is $O(|V(G)|^{35})$.

Before continuing with the outline of this algorithm, we need a few more definitions. The *interior* of a path P is the set of vertices that have degree two in P , and it is denoted by P^* . If u and v are two vertices in a path P , then u - P - v denotes the subpath of P whose ends are u and v , and if u and v are consecutive, then we may denote u - P - v by just u - v . For a subgraph X of G we say that $v \in V(G) \setminus V(X)$ has a neighbor in X if there exists a vertex in $V(X)$ adjacent to v .

For $A \subseteq V(G)$, a vertex is called *A-complete* if it is adjacent to every vertex in A . Two subsets $A, B \subseteq V(G)$ are *anticomplete* to each other if no vertex of A is adjacent to any vertex of B . For any two vertices a and b , let $N(a)$ be the set of neighbors of a in G , and let $N(a, b)$ be the set of all $\{a, b\}$ -complete vertices in G .

Let K be a prism. Label the triangles of K $\{a_1, a_2, a_3\}$ and $\{b_1, b_2, b_3\}$, and for $i = 1, 2, 3$ let P_i be the path from a_i to b_i . For $i = 1, 2, 3$ let m_i be the vertex in P_i such that for the paths a_i - P_i - m_i , denoted by S_i , and m_i - P_i - b_i , denoted by T_i , we have $|E(T_i)| \leq |E(S_i)| \leq |E(T_i)| + 1$. A *K-major* vertex, or just a *major* vertex when there is no danger of confusion, is a vertex not in $V(K)$ whose neighbors in K are not a subset of the vertex set of a 3-vertex path contained in K .

Finally, a subgraph H is *smaller* than a subgraph H' if $|V(H)| < |V(H')|$. An induced subgraph X of G that is either a theta or a prism is called *smallest* if no other theta or prism in G is smaller than X .

The idea of the algorithm is as follows. Let G be the graph input to the algorithm and suppose there is a prism K in G . A shortest path Q with the same ends as S_1 is called a *shortcut* across K . If the subgraph induced by $V(K) \cup Q^* \setminus S_1^*$ is also a prism, then Q is *good*, and if not, then Q is *bad*. Shortest paths with the same ends as one of T_1, S_2, T_2, S_3 , and T_3 are also called shortcuts across K . Shortcuts will be defined in the same way for thetas. A theta or prism is *clean* if all shortcuts across it are good. The algorithm uses a procedure called *cleaning*, first introduced by Conforti and Rao in [3]. The idea of cleaning is to identify a *cleaner*, which is a subset $X \subseteq V(G)$ such that if G contains a prism or a theta, then the graph G' induced by G on $V(G) \setminus X$ contains a clean prism or a clean theta.

In G' , it becomes easy to detect the clean theta or prism. For example, if G' contains a clean prism K , then for any shortcut Q across the path S_1 of K , the subgraph induced on $V(K) \cup Q^* \setminus S_1^*$ is also a prism, because Q is a good shortcut. The same is true for any other shortcut across K . So the algorithm checks all 9-tuples of vertices $a_1, a_2, a_3, m_1, m_2, m_3, b_1, b_2, b_3$; for each of the pairs $a_1m_1, a_2m_2, a_3m_3, m_1b_1, m_2b_2, m_3b_3$, it finds a shortest path joining the pair, then tests whether the subgraph induced by the union of these paths is a prism. If there is a clean smallest prism in G' , then it can be shown that a prism is found when considering one of the 9-tuples. A similar procedure works to find clean smallest thetas.

There are three stages of cleaning required. The first two cleaning algorithms output a list of polynomially many subsets of $V(G)$, one of which is a cleaner if G

contains a smallest prism or theta. The first outputs a cleaner when G contains a smallest theta and is given in section 2. The second outputs a cleaner when G contains a smallest prism K and is given in section 3. However, the second algorithm requires that the input graph G contain no K -major vertices. For this reason, we need the third cleaning algorithm, which outputs a list of subsets of which one contains all K -major vertices and is disjoint from $V(K)$. This algorithm is given in section 4. Finally, the algorithms for detecting clean prisms and thetas are described in section 5.

2. Theta shortcuts. Let K be a theta in G with ends a and b and paths P_1, P_2 , and P_3 between them. For $1 \leq i \leq 3$, let s_i be the neighbor of a in P_i , and let t_i be the neighbor of b in P_i . Also, for $1 \leq i \leq 3$ define $m_i \in P_i^*$ and subpaths S_i and T_i of P_i such that $S_i = s_i P_i m_i$, $T_i = m_i P_i t_i$, and $|V(T_i)| \leq |V(S_i)| \leq |V(T_i)| + 1$. We call P_1, P_2 , and P_3 the *paths* of K and call S_1, T_1, S_2, T_2, S_3 , and T_3 the *half-paths* of K . The *frame* of the theta K is the 11-tuple

$$\{a, b, s_1, s_2, s_3, m_1, m_2, m_3, t_1, t_2, t_3\},$$

and we say that K has a *tidy frame* if

- a and b have no neighbors in $V(G) \setminus V(K)$, and
- for $i = 1, 2, 3$, if m_i is adjacent to a and b , then m_i has no neighbors in $V(G) \setminus V(K)$.

If S is one of the half-paths of K , then a *shortcut* across S is a shortest path in G with the same ends as S . A shortcut across K is a shortcut across any half-path of K . A shortcut S' across S is *good* if no interior vertex of S' has neighbors in $V(K) \setminus V(S)$ and *bad* otherwise.

THEOREM 2.1. *Let K be a smallest theta in G . Assume K has a tidy frame. Then there is a vertex $c \in P_2^* \cup P_3^*$ such that c has a neighbor in the interior of every bad shortcut across S_1 or T_1 .*

Proof. Let C be the cycle induced by G on $V(P_2) \cup V(P_3)$. Let Q be a bad shortcut across S_1 . Suppose there is no edge between $V(Q)$ and $V(C)$ except s_1-a . Since Q is a bad shortcut, it follows that some vertex of Q has a neighbor in $V(T_1) \setminus \{m_1\}$, and then the graph induced by G on $V(C) \cup V(Q) \cup V(T_1)$ contains a theta smaller than K , a contradiction. Since m_1 has no neighbors in C and a is the only neighbor of s_1 in C , this proves that some vertex in Q^* has a neighbor in C . Let $q \in Q^*$ be the vertex with neighbors in C such that no other vertex in $q-Q-m_1$ has neighbors in C .

CLAIM 2.1. *The vertex q has exactly two neighbors in C , and they are adjacent to each other.*

If $m_1 = t_1$, then any shortcut between m_1 and s_1 has empty interior, and therefore is a good shortcut. Consequently, we may assume $m_1 \neq t_1$.

Suppose q has two neighbors in C that are not adjacent to each other. Let u and v be the neighbors of q in C such that a path between u and v in C contains b and no other neighbor of q . Then u and v are not adjacent to each other. Let C_u be the path of C from u to b such that $v \notin V(C_u)$ and let C_v be the path of C from v to b such that $u \notin V(C_v)$. The paths $q-u-C_u-b$, $q-v-C_v-b$ and a path between q and b with interior in $V(q-Q-m_1-T_1-b)$ form a theta, K' . Since $V(K') \subseteq (V(K) \setminus V(S_1)) \cup (V(Q) \setminus \{s_1\})$, and $|V(Q)| \leq |V(S_1)|$, we have $|V(K')| < |V(K)|$, contradicting the fact that no theta in G is smaller than K . Therefore all neighbors of q in C are pairwise adjacent.

Suppose q has exactly one neighbor in C ; call it u . If u is not adjacent to b , then the two paths between u and b in C and a path with interior in $V(u-q-Q-m_1-T_1-b)$ form a theta, K' . Since $V(K') \subseteq (V(K) \setminus V(S_1)) \cup (V(Q) \setminus \{s_1\})$, K' is smaller than K , a contradiction. Therefore u is adjacent to b . Assume u is not adjacent to a . Since

$m_1 \neq t_1$, the two paths between u and a in C and a path P' between u and a with interior in $V(u-q-Q-m_1-S_1-s_1-a)$ form a theta, K' . If q is nonadjacent to s_1 , then $|V(q-Q-m_1)| < |V(T_1)|$, and therefore $|V(K')| < |V(K)|$; and if q is adjacent to s_1 , then $V(P') = \{q, s_1, a\}$, and so $V(K') = V(C) \cup \{s_1\}$ and, again, $|V(K')| < |V(K)|$; in both cases, a contradiction. Therefore u is adjacent to a . Since u is adjacent to both a and b , it is one of m_2 or m_3 , contradicting the fact that K has a tidy frame. Therefore, q has two neighbors in C , and they are adjacent. This proves Claim 2.1.

Let R be a bad shortcut across T_1 . Suppose there is no edge between $V(R)$ and $V(C)$ except t_1-b . Since R is a bad shortcut, it follows that some vertex of R has a neighbor in $V(S_1) \setminus \{m_1\}$, and then the graph induced by G on $V(C) \cup V(R) \cup V(S_1)$ contains a theta smaller than K , a contradiction. Since m_1 has no neighbors in C , and b is the only neighbor of t_1 in C , this proves that some vertex in R^* has a neighbor in C . Let $r \in R^*$ be the vertex with neighbors in C such that no other vertex in $r-R-m_1$ has neighbors in C .

CLAIM 2.2. *The vertex r has exactly two neighbors in C , and they are adjacent to each other.*

This proof is very similar to that of Claim 2.1. If $m_1 = s_1$, then $m_1 = t_1$, and there are no shortcuts between m_1 and t_1 . Therefore we may assume $m_1 \neq s_1$.

Suppose r has two neighbors in C that are not adjacent to each other. Let u and v be the neighbors of r in C such that a path between u and v in C contains a and no other neighbor of r . Then u and v are not adjacent to each other. Let C_u be the path of C from u to a such that $v \notin V(C_u)$, and let C_v be the path of C from v to a such that $u \notin V(C_v)$. The paths $r-u-C_u-a$, $r-v-C_v-a$ and a path between r and a with interior contained in $V(r-R-m_1-S_1-a)$ form a theta, K' . Since $V(K') \subseteq (V(K) \setminus V(T_1)) \cup (V(R) \setminus \{t_1\})$, and $|V(R)| \leq |V(T_1)|$, we have $|V(K')| < |V(K)|$, contradicting the fact that no theta in G is smaller than K . Therefore all neighbors of r in C are pairwise adjacent.

Suppose r has exactly one neighbor in C ; call it u . If u is not adjacent to a , then the two paths between u and a in C and a path with interior in $V(u-r-R-m_1-S_1-a)$ form a theta, K' . Since $V(K') \subseteq (V(K) \setminus V(T_1)) \cup (V(R) \setminus \{t_1\})$, K' is smaller than K , a contradiction. Therefore u is adjacent to a . Assume u is not adjacent to b . Since $m_1 \neq s_1$, the two paths between u and b in C and a path with interior in $V(u-r-R-m_1-T_1-t_1-b)$ form a theta, K' . Since $r-R-m_1$ contains fewer vertices than R and $|V(R)| \leq |V(S_1)|$, we have $|V(K')| < |V(K)|$, a contradiction. Therefore u is adjacent to b . Since u is adjacent to both a and b , it is one of m_2 or m_3 , contradicting the fact that K has a tidy frame. Therefore, r has two neighbors in C , and they are adjacent. This proves Claim 2.2.

Let Q' be another bad shortcut across S_1 , and let $q' \in Q'^*$ be the vertex with neighbors in C such that no other vertex in $q'-Q'-m_1$ has neighbors in C . Denote by x_q and y_q the neighbors of q in C , and denote by x'_q and y'_q the neighbors of q' in C .

Let R' be another bad shortcut across T_1 , and let $r' \in R'^*$ be the vertex with neighbors in C such that no other vertex in $r'-R'-m_1$ has neighbors in C . Denote by x_r and y_r the neighbors of r in C , and denote by x'_r and y'_r the neighbors of r' in C .

CLAIM 2.3. *No two of the subsets $\{x_q, y_q\}$, $\{x'_q, y'_q\}$, $\{x_r, y_r\}$, and $\{x'_r, y'_r\}$ are disjoint.*

Assume that $\{x_q, y_q\}$ is disjoint from $\{x'_q, y'_q\}$. Since K has a tidy frame, $\{a, b\} \cap \{x_q, x'_q, y_q, y'_q\} = \emptyset$. Let Q'' be the path between q and q' with interior in $V(q-Q-m_1) \cup V(q'-Q'-m_1)$. The triangles $\{q, x_q, y_q\}$ and $\{q', x'_q, y'_q\}$, the two paths between them in

C , and the path Q'' form a prism, K' . Since

$$|V(Q'')| \leq 2|V(S_1)| - 3 < |V(S_1)| + |V(T_1)| - 1,$$

and $V(K') \subseteq V(C) \cup V(Q'')$, we deduce that K' is smaller than K , a contradiction. Therefore, $\{x_q, y_q\} \cap \{x'_q, y'_q\} \neq \emptyset$.

Assume that $\{x_r, y_r\}$ is disjoint from $\{x'_r, y'_r\}$. Since K has a tidy frame, $\{a, b\} \cap \{x_r, x'_r, y_r, y'_r\} = \emptyset$. Let R'' be the path between r and r' with interior in $V(r-R-m_1) \cup V(r'-R'-m_1)$. The triangles $\{r, x_r, y_r\}$ and $\{r', x'_r, y'_r\}$, the two paths between them in C and the path R'' form a prism, K' . Since

$$|V(R'')| \leq 2|V(T_1)| - 3 < |V(S_1)| + |V(T_1)| - 2,$$

and $V(K') \subseteq V(C) \cup V(R'')$, we deduce that K' is smaller than K , a contradiction. Therefore, $\{x_r, y_r\} \cap \{x'_r, y'_r\} \neq \emptyset$.

Suppose $\{x_q, y_q\}$ is disjoint from $\{x_r, y_r\}$. Then the triangles $\{q, x_q, y_q\}$ and $\{r, x_r, y_r\}$, the two paths in C between $\{x_q, y_q\}$ and $\{x_r, y_r\}$, and the path between q and r with interior in $V(q-Q-m_1-R-r)$ form a prism, K' . Since $V(K') \subseteq V(C) \cup V(Q) \cup V(R) \setminus \{s_1, t_1\}$, the prism K' is smaller than K , a contradiction. Therefore $\{x_q, y_q\} \cap \{x_r, y_r\} \neq \emptyset$. By symmetry, we also have that $\{x_q, y_q\} \cap \{x'_r, y'_r\} \neq \emptyset$, $\{x'_q, y'_q\} \cap \{x_r, y_r\} \neq \emptyset$, and $\{x'_q, y'_q\} \cap \{x'_r, y'_r\} \neq \emptyset$. This proves Claim 2.3.

Let U, W , and Z each be a bad shortcut across either S_1 or T_1 . Since the choices of Q, Q', R , and R' were arbitrary, it follows from Claims 2.1, 2.2, and 2.3 that there exist $u \in U^*$, $w \in W^*$, and $z \in Z^*$ such that u has two adjacent neighbors $x_u, y_u \in V(C)$, w has two adjacent neighbors $x_w, y_w \in V(C)$, z has two adjacent neighbors $x_z, y_z \in V(C)$, and no two of $\{x_u, y_u\}$, $\{x_w, y_w\}$, and $\{x_z, y_z\}$ are disjoint.

Assume that $\{x_w, y_w\} \cap \{x_z, y_z\} = \{x_w\}$ and $\{x_w, y_w\} \cap \{x_u, y_u\} = \{y_w\}$. Then, since x_w and y_w are adjacent and x_z and y_z are adjacent, $\{x_z, y_z\}$ contains x_w and the vertex adjacent to x_w in $V(C) \setminus \{y_w\}$. Similarly, since x_u and y_u are adjacent, $\{x_u, y_u\}$ contains y_w and the vertex adjacent to y_w in $V(C) \setminus \{x_w\}$. Since x_w and y_w have no common neighbors in C , we have that $\{x_z, y_z\} \cap \{x_u, y_u\} = \emptyset$, a contradiction. Therefore, by symmetry between x_w and y_w we may assume that $x_w \in \{x_z, y_z\}$ and $x_w \in \{x_u, y_u\}$. Therefore every bad shortcut across S_1 or T_1 contains an interior vertex adjacent to x_w ; taking $c = x_w$ proves Theorem 2.1. $\square$

THEOREM 2.2. *There is an algorithm with the following specifications:*

- **Input:** A graph G .
- **Output:** A sequence of subsets $X_1, \dots, X_r$ of $V(G)$ with $r \leq |V(G)|^3$, such that for every smallest theta K in G with a tidy frame and $i \in \{1, 2, 3\}$, one of $X_1, \dots, X_r$ is disjoint from $V(K)$ and contains a vertex of every bad shortcut across the half-paths S_i and T_i of K .
- **Running Time:** $O(|V(G)|^4)$.

Proof. The algorithm is as follows. Enumerate all triples of vertices (c, u, v) for which u and v are adjacent to c . For each such triple, compute the subset $N(c) \setminus \{u, v\}$. Let $X_1, \dots, X_r$ be the subsets generated. We output the list $\emptyset, X_1, \dots, X_r$. That concludes the description of the algorithm; we now prove that it works correctly. The number of subsets generated is at most $|V(G)|^3$ and the running time is $O(|V(G)|^4)$ because it takes linear time to compute $N(c)$. It remains to check that the sequence $X_1, \dots, X_r$ has the properties claimed. Let K be a smallest theta in G and assume it has a tidy frame. By Theorem 2.1, for every $i = 1, 2, 3$ there is a vertex $c' \in V(K)$ such that every bad shortcut across S_i or T_i contains an interior vertex adjacent to c' . The vertex c' has exactly two neighbors in K ; call them u' and v' . One of the triples

enumerated by the algorithm will be (c', u', v') . The subset $N(c') \setminus \{u', v'\}$ includes all neighbors of c' , except its two neighbors that are in K , and is therefore disjoint from $V(K)$. This proves Theorem 2.2. $\square$

3. Prism shortcuts. Let K be a prism in G . Label the triangles of K $\{a_1, a_2, a_3\}$ and $\{b_1, b_2, b_3\}$ and the paths P_1 , P_2 , and P_3 such that for $i = 1, 2, 3$ the path P_i has ends a_i and b_i . For $i = 1, 2, 3$ choose $m_i \in P_i^*$ and subpaths S_i and T_i such that $S_i = a_i - P_i - m_i$, $T_i = m_i - P_i - b_i$, $|V(S_i)| \leq |V(T_i)| + 1$, and $|V(T_i)| \leq |V(S_i)| + 1$. We call P_1 , P_2 , and P_3 the *paths* of K and S_1 , T_1 , S_2 , T_2 , S_3 , and T_3 the *half-paths* of K . The 9-tuple

$$\{a_1, a_2, a_3, m_1, m_2, m_3, b_1, b_2, b_3\}$$

is the *frame* of K . A prism K has a *tidy frame* if for $i = 1, 2, 3$ the following hold:

- Neither a_i nor b_i has neighbors in $V(G) \setminus V(K)$, and
- if m_i is adjacent to a_i and b_i , then m_i has no neighbors in $V(G) \setminus V(K)$.

The prism K in G is *almost clean* if it has a tidy frame and G contains no K -major vertices.

If S is one of the half-paths of K , then a *shortcut* across S is a shortest path in G with the same ends as S . A shortcut across K is a shortcut across some half-path of K . A shortcut Q across S is *good* if no interior vertex of Q has neighbors in $V(K) \setminus V(S)$, and *bad* otherwise.

Let C be the cycle induced by G on $V(P_2) \cup V(P_3)$.

THEOREM 3.1. *Let K be a smallest prism. For every bad shortcut Q across S_1 , some vertex of Q^* has a neighbor in C .*

Proof. Suppose no vertex of Q^* has a neighbor in C . Then, since Q is bad, some vertex of Q^* has a neighbor in $V(T_1) \setminus \{m_1\}$. Therefore, there exists a path P'_1 from a_1 to b_1 with $P'_1 \subset V(Q) \cup V(T_1) \setminus \{m_1\}$. So there is a prism contained in $V(C) \cup V(P'_1)$ that is smaller than K , which is a contradiction because K is a smallest prism. This proves Theorem 3.1. $\square$

In view of Theorem 3.1, for a bad shortcut Q across S_1 or T_1 , define $\beta_Q \in Q^*$ to be the vertex with neighbors in $V(C)$ such that no other vertex in $\beta_Q - Q - m_1$ has neighbors in $V(C)$.

THEOREM 3.2. *Let K be an almost clean smallest prism. For a bad shortcut Q across S_1 , either $|N(\beta_Q) \cap V(K)| = |N(\beta_Q) \cap V(C)| = 1$ or $N(\beta_Q) \cap V(K)$ is the vertex set of a three-vertex path of C .*

Proof. Since K is almost clean, β_Q is not a K -major vertex, and therefore $N(\beta_Q) \cap V(K) \subseteq N(\beta_Q) \cap V(C)$. Assume β_Q has exactly two neighbors in C , u , and v . If u and v are not adjacent, then $u - \beta_Q - v$ and the two paths in C between u and v form a theta that has fewer vertices than K , a contradiction. So u and v are adjacent. Denote $\beta_Q - Q - m_1$ by U . Since K is almost clean, $\{u, v\} \cap \{b_1, b_2, b_3\} = \emptyset$, so there is a prism K' containing the triangles $\{b_1, b_2, b_3\}$ and $\{\beta_Q, u, v\}$ with $V(K') \subseteq V(C) \cup V(T_1) \cup V(U)$ (the three paths of K' are two paths of C and the path contained in $V(U) \cup V(T_1)$). Since $|V(U)| < |V(S_1)|$ (because $a_1 \notin V(U)$), K' is smaller than K , a contradiction. Therefore, β_Q has either a unique neighbor in C or at least three neighbors in C . Since K is almost clean, β_Q is not K -major, so if it has at least three neighbors in C , then it has exactly three neighbors in C , which are the vertex set of a 3-vertex path of C . This proves Theorem 3.2. $\square$

THEOREM 3.3. *Let K be a smallest prism in G . Assume K is almost clean. Then there exist two edges e_1 and e_2 , one in $E(P_2)$ and the other in $E(P_3)$, such that either*

- there exists a bad shortcut T across S_1 with $|N(\beta_T) \cap V(C)| = 1$ and for every bad shortcut Q across S_1 , if $|N(\beta_Q) \cap V(C)| = 1$, then β_Q is adjacent to an end of e_1 and if $|N(\beta_Q) \cap V(C)| > 1$, then β_Q is adjacent to both ends of e_1 , or
- for every bad shortcut Q , $|N(\beta_Q) \cap V(C)| = 3$ and β_Q is adjacent either to both ends of e_1 or to both ends of e_2 .

Proof. We start with a few observations about bad shortcuts across S_1 .

CLAIM 3.1. *There exists an edge $e_1 \in E(C)$ such that for every bad shortcut Q across S_1 , if $|N(\beta_Q) \cap V(C)| = 1$, then β_Q is adjacent to an end of e_1 .*

We may assume that there exist two distinct bad shortcuts Q and Q' across S_1 such that β_Q has a unique neighbor x in C , $\beta_{Q'}$ has a unique neighbor y in C , and $x \neq y$. Let U be the path β_Q - Q - m_1 , let U' be the path $\beta_{Q'}$ - Q' - m_1 , and let W be a path between β_Q and $\beta_{Q'}$ with interior in $V(U) \cup V(U')$. If x and y are nonadjacent, then the two paths between them in C and the path x - β_Q - W - $\beta_{Q'}$ - y form a theta, K' . Since $|V(W)| \leq 2|V(S_1)| - 3$ and $|V(P_1)| \geq 2|V(S_1)| - 2$, K' is smaller than K , a contradiction. Therefore, x is adjacent to y .

Let Q'' be any other bad shortcut across S_1 such that $\beta_{Q''}$ has a unique neighbor z in C . By the previous argument, applied to the pairs Q, Q'' and Q', Q'' , it follows that z is equal or adjacent to x and equal or adjacent to y . Then either $z \in \{x, y\}$ or z is adjacent to both x and y . But since x and y are adjacent vertices of C , they have no common neighbor in C , so $z \in \{x, y\}$. Therefore, there exists an edge $e_1 = x$ - y such that whenever $\beta_{Q''}$ has a unique neighbor in C , $\beta_{Q''}$ is adjacent to an end of e_1 . This proves Claim 3.1.

CLAIM 3.2. *There exist edges $e_1 \in E(P_2)$ and $e_2 \in E(P_3)$ such that for every bad shortcut Q across S_1 , if $|N(\beta_Q) \cap V(C)| = 3$, then β_Q is either adjacent to both ends of e_1 or to both ends of e_2 .*

Let Q be a bad shortcut across S_1 such that $\beta_{Q'}$ has three neighbors in C . By Theorem 3.2, $N(\beta_Q) \cap V(C)$ is the vertex set of a 3-vertex path of C . Since K is almost clean, $N(\beta_Q) \cap V(C)$ is contained in either P_2^* or P_3^* .

Therefore, in order to prove Claim 3.2, it is enough to show that there exist edges $e_1 \in E(P_2)$ and $e_2 \in E(P_3)$ such that for every bad shortcut Q across S_1 , if $|N(\beta_Q) \cap V(P_2)| = 3$, then β_Q is adjacent to both ends of e_1 , and if $|N(\beta_Q) \cap V(P_3)| = 3$, then β_Q is adjacent to both ends of e_2 .

Let Q and Q' be bad shortcuts across S_1 . Suppose that each of β_Q and $\beta_{Q'}$ has three neighbors in P_2 . Then each of $N(\beta_Q) \cap V(C)$ and $N(\beta_{Q'}) \cap V(C)$ is the vertex set of a 3-vertex path of P_2 .

We claim that $|N(\beta_Q) \cap V(C) \cap N(\beta_{Q'})| \geq 2$. Assume that $|N(\beta_Q) \cap V(C) \cap N(\beta_{Q'})| < 2$. If β_Q and $\beta_{Q'}$ are adjacent to each other, then there is a path P'_2 between a_2 and b_2 whose interior is contained in $P_2^* \cup \{\beta_Q, \beta_{Q'}\}$ and is anticomplete to $V(K) \setminus P_2^*$. Then $(V(K) \setminus P_2^*) \cup P_2'^*$ contains a prism smaller than K , a contradiction. So we may assume that β_Q and $\beta_{Q'}$ are not adjacent to each other. Then there are two paths A and B between β_Q and $\beta_{Q'}$ whose interiors are anticomplete to each other and are contained in $V(C)$. The paths A , B and a path between β_Q and $\beta_{Q'}$ with interior in $V(Q) \cup V(Q')$ form a theta, K' . Since $|V(Q) \cup V(Q')| \leq 2|V(S_1)| - 3 < |V(S_1)| + |V(T_1)| - 1 = |V(P_1)|$, the theta K' is smaller than the prism K , a contradiction. This proves the claim.

Label the neighbors of β_Q in C as x, y , and z and the neighbors of $\beta_{Q'}$ in C as x', y' , and z' such that x - y - z and x' - y' - z' are paths of C . Let Q'' be another bad shortcut across S_1 such that $\beta_{Q''}$ has three neighbors $x'', y'',$ and z'' in P_2

and $x''-y''-z''$ is a path of P_2 . Then we have that $|\{x'', y'', z''\} \cap \{x, y, z\}| \geq 2$ and $|\{x'', y'', z''\} \cap \{x', y', z'\}| \geq 2$. Since $x-y-z$ is a path, either $\{x, y\} \subset \{x', y', z'\}$ or $\{y, z\} \subset \{x', y', z'\}$, and similarly, either $\{x, y\} \subset \{x'', y'', z''\}$ or $\{y, z\} \subset \{x'', y'', z''\}$. Assume that $\{x, y, z\} \cap \{x', y', z'\} = \{x, y\}$ and $\{x, y, z\} \cap \{x'', y'', z''\} = \{y, z\}$. Then since $x-y-z$ is a path and $x'-y'-z'$ is a path, $\{x', y', z'\}$ contains $\{x, y\}$ and the vertex adjacent to x in $V(P_2) \setminus \{y\}$. Similarly, since $x''-y''-z''$ is a path, $\{x'', y'', z''\}$ contains $\{y, z\}$ and the vertex adjacent to z in $V(P_2) \setminus \{y\}$. Since x and z are not adjacent and x and z have no common neighbors in $V(P_2) \setminus \{y\}$, it follows that $\{x', y', z'\} \cap \{x'', y'', z''\} = \{y\}$, a contradiction. Therefore, by symmetry between $\{x, y\}$ and $\{y, z\}$ we may assume that $\{x, y\} \subset \{x', y', z'\}$ and $\{x, y\} \subset \{x'', y'', z''\}$. This proves that for every bad shortcut R across S_1 , if $|N(\beta_R) \cap V(P_2)| = 3$, then β_R is adjacent to both ends of $e_1 = x-y$.

From the symmetry, there exists an edge $e_2 \in E(P_3)$ such that for every bad shortcut R across S_1 , if $|N(\beta_R) \cap V(P_3)| = 3$, then β_R is adjacent to both ends of e_2 . This proves Claim 3.2.

CLAIM 3.3. *If Q and Q' are bad shortcuts across S_1 with $|N(\beta_Q) \cap V(C)| = 1$ and $|N(\beta_{Q'}) \cap V(C)| = 3$, then the neighbor of β_Q in C is adjacent to $\beta_{Q'}$.*

Denote by x the neighbor of β_Q in C . By Theorem 3.2, the neighbors of $\beta_{Q'}$ in C are the vertex set of a 3-vertex path of C , say $c_1-c_2-c_3$.

Assume that $x \notin \{c_1, c_2, c_3\}$. Then the subgraph induced by G on $(V(C) \setminus \{c_2\}) \cup V(Q) \cup V(Q')$ contains a theta, K' (the three paths of K' are two paths between $\beta_{Q'}$ and x with interior in C and the path with interior in $V(\beta_{Q'}-Q'-m_1-Q-\beta_Q-x)$). Since $V(K') \subseteq V(C) \cup Q^* \cup Q'^* \cup \{m_1\}$ and $|Q^*| + |Q'^*| + 1 \leq 2|V(S_1)| - 3 < |V(S_1)| + |V(T_1)| - 1$, the theta K' is smaller than K , a contradiction. Therefore, $x \in \{c_1, c_2, c_3\}$. This proves Claim 3.3.

If for every bad shortcut Q , $|N(\beta_Q) \cap V(C)| = 3$, then the second outcome of the theorem holds by Claim 3.2. So we may assume that there exists a bad shortcut Q with $N(\beta_Q) \cap V(C) = \{x\}$. Suppose there exists a bad shortcut Q' with $|N(\beta_{Q'}) \cap V(C)| = \{y\}$ such that $x \neq y$. Then by Claims 3.1 and 3.3 the first outcome of the theorem holds with e_1 being the edge as in Claim 3.1. So we may assume that for every bad shortcut Q' with $|N(\beta_{Q'}) \cap V(C)| = 1$, $\beta_{Q'}$ is adjacent to x . Let c_1-x-c_2 be the 3-vertex path in C with x as its interior vertex. If there exist bad shortcuts Q' and Q'' with $|N(\beta_{Q'}) \cap V(C)| = 3$ and $|N(\beta_{Q''}) \cap V(C)| = 3$ such that $\beta_{Q'}$ is not adjacent to c_1 and $\beta_{Q''}$ is not adjacent to c_2 , then $N(\beta_{Q'}) \cap V(C) \cap N(\beta_{Q''}) = \{x\}$, contradicting Claim 3.2. Thus one of c_1 and c_2 , say c_1 , is complete to $\{\beta_{Q'}, \beta_{Q''}\}$, and the theorem holds with $e_1 = c_1-x$. This completes the proof of Theorem 3.3. $\square$

THEOREM 3.4. *Let K be a smallest prism in G . Assume K is almost clean. Then there exist two edges e_1 and e_2 , one in $E(P_2)$ and the other in $E(P_3)$, such that either*

- *there exists a bad shortcut T across S_1 or T_1 with $|N(\beta_T) \cap V(C)| = 1$, and for every bad shortcut Q across S_1 or T_1 , if $|N(\beta_Q) \cap V(C)| = 1$, then β_Q is adjacent to an end of e_1 , and if $|N(\beta_Q) \cap V(C)| > 1$, then β_Q is adjacent to both ends of e_1 , or*
- *for every bad shortcut Q across S_1 or T_1 , $|N(\beta_Q) \cap V(C)| = 3$ and β_Q is adjacent either to both ends of e_1 or to both ends of e_2 .*

Proof. We start with a few observations about bad shortcuts across S_1 and T_1 .

CLAIM 3.4. *Let Q be a bad shortcut across S_1 with $|N(\beta_Q) \cap V(C)| = 1$ and R a bad shortcut across T_1 with $|N(\beta_R) \cap V(C)| = 1$. Then the neighbor of β_Q in C and the neighbor of β_R in C are either equal or adjacent to each other.*

Let $\{x_q\} = N(\beta_Q) \cap V(C)$ and $\{x_r\} = N(\beta_R) \cap V(C)$. Assume that x_q and x_r are distinct and nonadjacent. There are two paths between them, A and B , each of length at least two and contained in C . There is a path P between x_q and x_r with interior in $V(\beta_Q-Q-m_1-R-\beta_R)$. The three paths A , B , and P form a theta, K' . Since $|V(P)| \leq |Q^* \cup \{m_1\} \cup R^*| < |V(P_1)|$, the theta K' is smaller than the prism K , a contradiction. This proves Claim 3.4.

CLAIM 3.5. *Let Q be a bad shortcut across S_1 with $|N(\beta_Q) \cap V(C)| = 3$ and $N(\beta_Q) \cap V(C) \subseteq P_2^*$. Let R be a bad shortcut across T_1 with $|N(\beta_R) \cap V(C)| = 3$ and $N(\beta_R) \cap V(C) \subseteq P_2^*$. Then there is an edge in $E(P_2)$ whose ends are $\{\beta_Q, \beta_R\}$ -complete.*

Assume that $|N(\beta_Q) \cap V(C) \cap N(\beta_R)| < 2$. By Theorem 3.2, each of $N(\beta_Q) \cap V(K)$ and $N(\beta_R) \cap V(K)$ is the vertex set of a 3-vertex path of C . If β_Q is adjacent to β_R , then there is a path A between a_2 and b_2 with interior in $P_2^* \cup \{\beta_Q, \beta_R\}$ such that $\{\beta_Q, \beta_R\} \subset V(A)$. Since $|N(\beta_Q) \cap V(C)| = |N(\beta_R) \cap V(C)| = 3$, it follows that $|N(\beta_Q) \cup N(\beta_R) \cap V(C)| \geq 5$. So $|V(P_2) \setminus V(A)| \geq 3$ and $|V(A) \setminus V(P_2)| = 2$, which implies that $|V(A)| < |V(P_2)|$. It follows from Theorem 3.2 that A^* is anticomplete to P_1^* and P_3^* , and so the induced subgraph with vertex set $(V(K) \setminus V(P_2)) \cup V(A)$ is a prism smaller than K , a contradiction. This proves that β_Q is not adjacent to β_R . It follows that there are two paths B and B' between β_Q and β_R whose interiors are anticomplete to each other and are contained in $V(C)$. There is also a path P between them with interior contained in $V(\beta_Q-Q-m_1-R-\beta_R)$. Since β_Q is not adjacent to β_R , P has length at least two. Since $|V(P)| \leq |Q^* \cup \{m_1\} \cup R^*| < |V(P_1)|$, the paths B , B' , and P form a theta smaller than K , a contradiction. This proves that $|N(\beta_Q) \cap V(C) \cap N(\beta_R)| \geq 2$. Since each of $N(\beta_Q) \cap V(C)$ and $N(\beta_R) \cap V(C)$ is the vertex set of a 3-vertex subpath of P_2 , Claim 3.5 follows.

CLAIM 3.6. *Let Q be a bad shortcut across S_1 with $|N(\beta_Q) \cap V(C)| = 1$ and R a bad shortcut across T_1 with $|N(\beta_R) \cap V(C)| = 3$. Then the neighbor of β_Q in C is adjacent to β_R .*

Let $\{x_q\} = N(\beta_Q) \cap V(C)$. Assume that $x_q \notin N(\beta_R)$. There are two paths A and B between x_q and β_R whose interiors are anticomplete to each other and are contained in $V(C)$. There is a path P between x_q and β_R with interior in $V(\beta_Q-Q-m_1-R-\beta_R)$. Since $|V(P)| \leq |Q^* \cup \{m_1\} \cup R^*| + 1 < |V(P_1)|$, the paths A , B , and P form a theta smaller than K , a contradiction. This proves Claim 3.6.

CLAIM 3.7. *Suppose that there exist $p \in V(P_2)$ and $s \in V(P_3)$ such that for every bad shortcut Q across S_1 or T_1 , β_Q is adjacent to p or s . Then the theorem holds.*

Since there are no K -major vertices in G , for every bad shortcut Q across S_1 or T_1 with $N(\beta_Q) \cap V(C) \subseteq P_2^*$, $N(\beta_Q) \cap V(C)$ is contained in the vertex set of the 5-vertex path $c_1-c_2-p-c_3-c_4$ contained in C . Suppose there exist two bad shortcuts P and P' across either S_1 or T_1 with $N(\beta_P) \cap V(C) = \{c_1, c_2, p\}$ and $N(\beta_{P'}) \cap V(C) = \{p, c_3, c_4\}$. Then $N(\beta_P) \cap V(C) \cap N(\beta_{P'}) = \{p\}$. If both P and P' are shortcuts across S_1 or both are shortcuts across T_1 , then this contradicts both outcomes of Theorem 3.3. If one of P and P' is a shortcut across S_1 and the other is a shortcut across T_1 , then this contradicts Claim 3.5. Therefore, there exists a 4-vertex path A of C with $p \in A^*$ such that for every bad shortcut P across S_1 or T_1 , if $N(\beta_P) \cap V(C) \subseteq P_2^*$, then $N(\beta_P) \cap V(C) \subset V(A)$. Let e_1 be the edge with both ends in A^* . It follows that for every bad shortcut P across S_1 or T_1 with $N(\beta_P) \cap V(C) \subseteq P_2^*$, if $|N(\beta_P) \cap V(C)| = 1$, then β_P is adjacent to the end p of e_1 , and if $|N(\beta_P) \cap V(C)| > 1$, then β_P is adjacent to both ends of e_1 . By symmetry between P_2 and P_3 , it also follows that there is an edge $e_2 \in E(P_3)$ such that for every bad shortcut P across S_1 or T_1 with

$N(\beta_P) \cap V(C) \subseteq P_3^*$, if $|N(\beta_P) \cap V(C)| = 1$, then β_P is adjacent to an end of e_2 , and if $|N(\beta_P) \cap V(C)| > 1$, then β_P is adjacent to both ends of e_2 . So if for all bad shortcuts Q across S_1 and T_1 , $|N(\beta_Q) \cap V(C)| > 1$, then the second outcome of the theorem holds. If not, then by Claims 3.4 and 3.6, either for all bad shortcuts Q across S_1 and T_1 , $N(\beta_Q) \cap V(C) \subseteq P_2^*$, or for all bad shortcuts Q across S_1 and T_1 , $N(\beta_Q) \cap V(C) \subseteq P_3^*$. In both cases, (with e_2 in place of e_1 in the second case), the first outcome of the theorem holds. This proves Claim 3.7.

First assume that there exist a bad shortcut Q across S_1 with $|N(\beta_Q) \cap V(C)| = 1$ and a bad shortcut R across T_1 with $|N(\beta_R) \cap V(C)| = 1$. The first outcome of Theorem 3.3 holds for both S_1 and T_1 . Let $\{x_q\} = N(\beta_Q) \cap V(C)$ and $\{x_r\} = N(\beta_R) \cap V(C)$. Assume that $x_q \neq x_r$. By Claim 3.4, x_q is adjacent to x_r . Let Q' be any other bad shortcut across S_1 and let R' be any other bad shortcut across T_1 . It follows from Claim 3.4 that if $|N(\beta_{Q'}) \cap V(C)| = 1$, then $\beta_{Q'}$ is adjacent to x_q or x_r ; it also follows that if $|N(\beta_{R'}) \cap V(C)| = 1$, then $\beta_{R'}$ is adjacent to x_r or x_q . Also, Theorem 3.3 implies that if $|N(\beta_{Q'}) \cap V(C)| > 1$, then $\beta_{Q'}$ is adjacent to x_q , and if $|N(\beta_{R'}) \cap V(C)| > 1$, then $\beta_{R'}$ is adjacent to x_r . By Claim 3.6, if $|N(\beta_{Q'}) \cap V(C)| > 1$, then $\beta_{Q'}$ is adjacent to x_r , and if $|N(\beta_{R'}) \cap V(C)| > 1$, then $\beta_{R'}$ is adjacent to x_q . Therefore, the first outcome of the theorem holds with e_1 being x_q-x_r . Now we may assume that for every bad shortcut Q across S_1 with $|N(\beta_Q) \cap V(C)| = 1$ and R across T_1 with $|N(\beta_R) \cap V(C)| = 1$, if $\{x_q\} = N(\beta_Q) \cap V(C)$ and $\{x_r\} = N(\beta_R) \cap V(C)$, then $x_q = x_r$. By the first outcome of Theorem 3.3, for any bad shortcut Q' across S_1 with $|N(\beta_{Q'}) \cap V(C)| > 1$, $\beta_{Q'}$ is adjacent to x_q , and for any bad shortcut R' across T_1 with $|N(\beta_{R'}) \cap V(C)| > 1$, $\beta_{R'}$ is adjacent to x_r , which equals x_q . By Claim 3.7 with $p = x_q$, the first outcome of the theorem holds.

Next assume that there exists a bad shortcut Q across S_1 with $|N(\beta_Q) \cap V(C)| = 1$. Then we may assume because of the previous paragraph that for every bad shortcut R across T_1 , $|N(\beta_R) \cap V(C)| > 1$. The first outcome of Theorem 3.3 holds for shortcuts across S_1 , and the second outcome of Theorem 3.3 holds for shortcuts across T_1 . Let $\{x_q\} = N(\beta_Q) \cap V(C)$. Suppose there is a bad shortcut Q' across S_1 with $N(\beta_{Q'}) \cap V(C) = \{x_{q'}\}$ such that $x_{q'} \neq x_q$. By the first outcome of Theorem 3.3, x_q is adjacent to $x_{q'}$, and for every bad shortcut Q'' across S_1 , if $|N(\beta_{Q''}) \cap V(C)| = 1$, then $\beta_{Q''}$ is adjacent to either x_q or to $x_{q'}$, and if $|N(\beta_{Q''}) \cap V(C)| > 1$, then $\beta_{Q''}$ is adjacent to both x_q and $x_{q'}$. By Claim 3.6, for every bad shortcut R' across T_1 , both x_q and $x_{q'}$ are adjacent to $\beta_{R'}$. Then the first outcome of the theorem holds with e_1 being the edge $x_q-x_{q'}$. So for every bad shortcut Q' across S_1 with $|N(\beta_{Q'}) \cap V(C)| = 1$ we may assume that $\beta_{Q'}$ is adjacent to x_q . By the first outcome of Theorem 3.3, for any bad shortcut Q' across S_1 with $|N(\beta_{Q'}) \cap V(C)| > 1$, $\beta_{Q'}$ is adjacent to x_q . By Claim 3.6, for any bad shortcut R' across T_1 , $\beta_{R'}$ is adjacent to x_q . Therefore, by Claim 3.7 with $p = x_q$, the first outcome of the theorem holds.

If there exists a bad shortcut R across T_1 with $|N(\beta_R) \cap V(C)| = 1$, then by symmetry, the argument of the previous paragraph applies. So we may finally assume that for every shortcut P across either S_1 or T_1 , $|N(\beta_P) \cap V(C)| > 1$. Then the second outcome of Theorem 3.3 holds for both S_1 and T_1 . Let Q be a bad shortcut across S_1 with $N(Q) \cap V(C) \subseteq P_2^*$. Let c_1-p-c_2 be a subpath of P_2 such that $N(Q) \cap V(C) = \{c_1, p, c_2\}$. By the second outcome of Theorem 3.3, for any bad shortcut Q' across S_1 with $N(\beta_{Q'}) \cap V(C) \subseteq P_2^*$, $\beta_{Q'}$ is adjacent to p . Also, by Claim 3.5, for any bad shortcut R across T_1 with $N(\beta_R) \cap V(C) \subseteq P_2^*$, β_R is adjacent to p . By symmetry between P_2 and P_3 , there is also a vertex $s \in P_3^*$ such that for every bad shortcut R across S_1 or T_1 with $N(\beta_R) \cap V(C) \subseteq P_3^*$, β_R is adjacent to s . Therefore, by Claim 3.7, the second case of the theorem holds. This proves Theorem 3.4. $\square$

THEOREM 3.5. *There is an algorithm with the following specifications:*

- **Input:** A graph G .
- **Output:** A sequence of subsets $X_1, \dots, X_r$ of $V(G)$ with $r \leq 2|V(G)|^4$, such that for every almost clean smallest prism K in G and $i \in \{1, 2, 3\}$, one of $X_1, \dots, X_r$ is disjoint from $V(K)$ and contains a vertex of every bad shortcut across the half-paths S_i and T_i of K .
- **Running Time:** $O(|V(G)|^5)$.

Proof. The algorithm is as follows. Enumerate all quadruples of distinct vertices (u_1, u_2, u_3, u_4) . For each, define the subset

$$Y = (N(u_2) \cup N(u_3)) \setminus \{u_1, u_2, u_3, u_4\}$$

and the subset

$$Z = N(u_1, u_2) \cup N(u_3, u_4).$$

Let $Y_1, \dots, Y_s$ and $Z_1, \dots, Z_s$ be the subsets generated. Output $\emptyset, Y_1, \dots, Y_s, Z_1, \dots, Z_s$.

That concludes the description of the algorithm; we now prove that it works correctly. It takes time $O(|V(G)|^5)$ to find the sets $Y_1, \dots, Y_s$ and $Z_1, \dots, Z_s$, and the number of subsets in the output is $\leq 2|V(G)|^4$. It remains to check that the output sequence has the properties claimed. Let K be an almost clean smallest prism in G with half-paths S_1, T_1, S_2, T_2, S_3 , and T_3 . We may assume by symmetry that $i = 1$. Apply Theorem 3.4 to S_1 and T_1 .

If the first outcome of Theorem 3.4 holds, then there is an edge in $P_2^* \cup P_3^*$, say u_2-u_3 , such that every bad shortcut across S_1 or T_1 contains an interior vertex adjacent to either u_2 or u_3 . Choose u_1 to be the neighbor of u_2 in $V(K) \setminus \{u_3\}$ and u_4 to be the neighbor of u_3 in $V(K) \setminus \{u_2\}$. Then the subset $(N(u_2) \cup N(u_3)) \setminus \{u_1, u_2, u_3, u_4\}$ is disjoint from $V(K)$ and contains an interior vertex of every bad shortcut across S_1 or T_1 .

If the second outcome of Theorem 3.4 holds, then there are two edges in $P_2^* \cup P_3^*$, say u_1-u_2 and u_3-u_4 , such that the set $N(u_1, u_2) \cup N(u_3, u_4)$ contains an interior vertex of every bad shortcut across S_1 or T_1 . Since u_1 and u_2 are adjacent, they have no common neighbors in $V(K)$, and since u_3 and u_4 are adjacent, they have no common neighbors in $V(K)$, so $N(u_1, u_2) \cup N(u_3, u_4)$ is disjoint from $V(K)$. This proves Theorem 3.5. $\square$

4. Major vertices. In this section, let K be a smallest prism in a graph G , and assume that K has a tidy frame. Let the vertices of P_1 be $f_1, f_2, \dots, f_n$, numbered in order with $f_1 = a_1$ and $f_n = b_1$. Let the vertices of P_2 be $h_1, h_2, \dots, h_m$, numbered in order with $h_1 = a_2$ and $h_m = b_2$, and let the vertices of P_3 be $g_1, g_2, \dots, g_p$, numbered in order with $g_1 = a_3$ and $g_p = b_3$.

Recall that a K -major vertex is a vertex in $V(G) \setminus V(K)$ whose neighbors in K are not contained in a 3-vertex path in K . Since K has a tidy frame, every K -major vertex is anticomplete to $\{a_1, a_2, a_3, b_1, b_2, b_3\}$. For a K -major vertex x , if x has neighbors in P_1 , let j_x be minimal such that x is adjacent to f_{j_x} and let k_x be maximal such that x is adjacent to f_{k_x} . Similarly, let s_x be minimal and t_x maximal such that x is adjacent to h_{s_x} and h_{t_x} , and let c_x be minimal and d_x maximal such that x is adjacent to g_{c_x} and g_{d_x} .

We start with a few easy but useful lemmas.

THEOREM 4.1. *A K -major vertex has at least three neighbors in K .*

Proof. Let x be a major vertex with $N(x) \cap V(K) = \{u, v\}$. We may assume by symmetry that $\{u, v\} \subset V(P_2) \cup V(P_3)$. Since x is major, u and v are not adjacent, so the subgraph induced by G on $V(P_2) \cup V(P_3) \cup \{x\}$ is a theta smaller than K , a contradiction. $\square$

THEOREM 4.2. *There is no K -major vertex with exactly two neighbors in P_2 and no neighbors in P_3 .*

Proof. Let x be a K -major vertex with $N(x) \cap V(P_2) = \{h_{s_x}, h_{t_x}\}$ and $N(x) \cap V(P_3) = \emptyset$. By Theorem 4.1, x has at least one neighbor in P_1 . If $s_x + 1 < t_x$, then the subgraph induced by G on $V(P_2) \cup V(P_3) \cup \{x\}$ is a theta smaller than K , a contradiction. So we may assume that $s_x + 1 = t_x$. Now the triangles $\{h_{s_x}, h_{t_x}, x\}$ and $\{a_1, a_2, a_3\}$ and the paths $h_{s_x}-P_2-a_2$, $h_{t_x}-P_2-b_2-b_3-P_3-a_3$, and $x-f_{j_x}-P_1-a_1$ form a prism, K' . Since $V(K') \subseteq V(K) \cup \{x\} \setminus \{f_{j_x+1}, b_1\}$, the prism K' is smaller than K unless $f_{j_x+1} = b_1$. Also, the triangles $\{x, h_{s_x}, h_{t_x}\}$ and $\{b_1, b_2, b_3\}$ and the paths $h_{s_x}-P_2-a_2-a_3-P_3-b_3$, $h_{t_x}-P_2-b_2$, and $x-f_{k_x}-P_1-b_1$ form a prism, K'' . Since $V(K'') \subseteq V(K) \cup \{x\} \setminus \{f_{k_x-1}, a_1\}$, the prism K'' is smaller than K unless $f_{k_x-1} = a_1$. Therefore, since K is a smallest prism, $f_{j_x+1} = b_1$ and $f_{k_x-1} = a_1$. It follows that $j_x = k_x$ and f_{j_x} is adjacent to both a_1 and b_1 , so $f_{j_x} = m_1$, contradicting the fact that K has a tidy frame. $\square$

THEOREM 4.3. *Let x and y be vertices, each with at least two nonadjacent neighbors in a path $P = v_1 \cdots v_n$. If there do not exist paths A and B between x and y with $A^* \subset V(P)$ and $B^* \subset V(P)$ such that A^* and B^* are anticomplete to each other, then for some $1 \leq i \leq n$, $N(x) \cap V(P) \subset \{v_1, \dots, v_{i+1}\}$ and $N(y) \cap V(P) \subset \{v_i, \dots, v_n\}$.*

Proof. Let v_s and v_t be the neighbors of x in P with s minimum and t maximum. Let v_p and v_r be the neighbors of y in P with p minimum and r maximum. Since x and y each have two nonadjacent neighbors in P , $s + 1 < t$ and $p + 1 < r$. If $p < s$ and $r > t$, then there are paths A and B between x and y with $A^* \subseteq \{v_p, \dots, v_s\}$ and $B^* \subseteq \{v_t, \dots, v_r\}$, and A^* and B^* are anticomplete to each other, a contradiction. So either $p \geq s$ or $r \leq t$. By symmetry, we may assume that $p \geq s$. If $r \leq t$, then there are paths A and B between x and y with $A^* \subseteq \{v_s, \dots, v_p\}$ and $B^* \subseteq \{v_r, \dots, v_t\}$, and A^* and B^* are anticomplete to each other, a contradiction. Therefore, $r > t$. If $p < t - 1$, then there are paths A and B between x and y with $A^* \subseteq \{v_s, \dots, v_p\}$ and $B^* \subseteq \{v_t, \dots, v_r\}$, and A^* and B^* are anticomplete to each other, a contradiction. So $p \geq t - 1$, and the theorem holds with $i = t - 1$. This proves Theorem 4.3. $\square$

THEOREM 4.4. *Every K -major vertex has neighbors in at least two of P_1 , P_2 , and P_3 .*

Proof. Let v be a K -major vertex with $N(v) \cap V(K) \subseteq V(P_1)$. Since v is major, $j_v + 2 < k_v$. The subgraph induced by G on $V(K) \cup \{v\} \setminus \{f_{j_v+1}, \dots, f_{k_v-1}\}$ is a prism that is smaller than K . Therefore, $N(v) \cap V(K) \not\subseteq V(P_1)$. By symmetry between P_1 , P_2 , and P_3 , it follows that $N(v) \cap V(K) \not\subseteq V(P_2)$ and $N(v) \cap V(K) \not\subseteq V(P_3)$. This proves Theorem 4.4. $\square$

THEOREM 4.5. *Let x and y be K -major vertices that are not adjacent to each other. Then for some i, j with $1 \leq i < j \leq 3$, $N(x) \cap V(K) \subseteq V(P_i) \cup V(P_j)$ and $N(y) \cap V(K) \subseteq V(P_i) \cup V(P_j)$.*

Proof. We first prove that each of x and y has neighbors in exactly two of P_1 , P_2 , and P_3 .

CLAIM 4.1. *None of x and y has neighbors in all of P_1 , P_2 , P_3 .*

Assume that x and y both have neighbors in all three of P_1 , P_2 , and P_3 . Then there are three paths between x and y , each with interior in P_i^* for $i \in \{1, 2, 3\}$. These paths form a theta whose vertex set is contained in $(V(K) \setminus \{a_1, a_2, a_3, b_1, b_2, b_3\}) \cup \{x, y\}$

and is thus smaller than K , a contradiction. Therefore, at least one of x and y does not have neighbors in all three paths P_1 , P_2 , and P_3 .

We may assume by symmetry that $N(y) \cap V(P_1) = \emptyset$. Also, we may assume that x has neighbors in all of P_1 , P_2 , and P_3 , because otherwise Claim 4.1 holds. By Theorem 4.4, y has neighbors in both P_2 and P_3 . By Theorem 4.1, it has at least two neighbors in one of these paths, say P_2 . Then by Theorem 4.2, it has two nonadjacent neighbors in P_2 .

Since x and y both have neighbors in P_3 , there is a path R between them with interior in P_3^* . Let P be the path $h_{s_y}-P_2-a_2-a_1-P_1-b_1-b_2-P_2-h_{t_y}$. Assume that x has two nonadjacent neighbors in P . Then by Theorem 4.3, there are two paths R' and R'' between x and y whose interiors are anticomplete to each other and contained in $V(P)$. Since $V(P)$ is anticomplete to P_3^* , the paths $x-R-y$, $x-R'-y$, and $x-R''-y$ form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{a_3, b_3, h_{s_y+1}\}$, K' is smaller than K , a contradiction. Therefore, x does not have two nonadjacent neighbors in P . Since P_1 is a subpath of P , this implies that x does not have two nonadjacent neighbors in P_1 , so $j_x \geq k_x - 1$. Also, since $N(x) \cap P_1^* \neq \emptyset$, it follows that $N(x) \cap (V(P) \setminus V(P_1)) = \emptyset$. Thus $s_x > s_y$ and $t_x < t_y$.

If $s_x < t_y - 1$, then the paths $x-R-y$, $x-f_{k_x}-P_1-b_1-b_2-P_2-h_{t_y}-y$ and a path between x and y with interior in $\{h_{s_y}, \dots, h_{s_x}\}$ form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{h_{t_y-1}, a_3, b_3\}$, K' is smaller than K , a contradiction. Therefore, $s_x \geq t_y - 1$. Since $s_x \leq t_x$ and $t_x < t_y$, this proves that $s_x = t_y - 1 = t_x$. If $t_x > s_y + 1$, then the paths $x-R-y$, $x-f_{j_x}-P_1-a_1-a_2-P_2-h_{s_y}-y$ and a path between x and y with interior in $\{h_{t_x}, h_{t_y}\}$ form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{h_{s_y+1}, a_3, b_3\}$, K' is smaller than K , a contradiction. Therefore, $t_x \leq s_y + 1$. Since $t_x = s_x > s_y$, this proves that $s_x = t_x = s_y + 1$. Therefore, $s_y + 2 = t_y$ and $N(x) \cap V(P_2) = \{h_{s_y+1}\}$.

If y is not adjacent to h_{s_y+1} , then $|N(y) \cap V(P_2)| = 2$ and $N(y) \cap V(P_1) = \emptyset$, contradicting Theorem 4.2. Therefore, y is adjacent to h_{s_y+1} . If $j_x = k_x$, then the subgraph induced by G on $V(P_1) \cup V(P_2) \cup \{x\}$ is a theta smaller than K , a contradiction. Therefore, $j_x < k_x$, and since $j_x \geq k_x - 1$, it follows that $j_x = k_x - 1$. If $c_y < c_x$, then the triangles $\{h_{s_y}, h_{s_y+1}, y\}$ and $\{a_1, a_2, a_3\}$ and the paths $h_{s_y}-P_2-a_2$, $h_{s_y+1}-x-f_{j_x}-P_1-a_1$, and $y-g_{c_y}-P_3-a_3$ form a prism, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{b_1, b_2, b_3\}$, K' is smaller than K , a contradiction. Therefore $c_y \geq c_x$. If $c_x < c_y$, then the triangles $\{x, f_{j_x}, f_{k_x}\}$ and $\{a_1, a_2, a_3\}$ and the paths $x-g_{c_x}-P_3-a_3$, $f_{j_x}-P_1-a_1$, and $f_{k_x}-P_1-b_1-b_2-P_2-h_{t_y}-y-h_{s_y}-P_2-a_2$ form a prism, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{h_{s_y+1}, b_3, g_{c_y}\}$, K' is smaller than K , a contradiction. Therefore $c_x = c_y$. By the symmetry of the prism K , it follows by the same argument that $d_x = d_y$.

If $c_x = d_x$, then the induced subgraph with vertex set $V(P_2) \cup V(P_3) \cup \{x\}$ is a theta smaller than K , a contradiction. So $c_x < d_x$, which means that $c_y < d_y$. If $c_y < d_y - 1$, then the paths $x-g_{c_y}-y$, $x-g_{d_y}-y$, and $x-h_{s_y+1}-y$ form a theta smaller than K , a contradiction. Therefore, $c_y = d_y - 1$. But then $|N(y) \cap V(P_3)| = 2$ and $N(y) \cap V(P_1) = \emptyset$, which contradicts Theorem 4.2. This proves Claim 4.1.

By Theorem 4.4 and Claim 4.1, each of x and y has neighbors in precisely two of the paths P_1 , P_2 , and P_3 . Assume that $N(x) \cap V(K)$ and $N(y) \cap V(K)$ are not both contained in the union of the same two of $V(P_1)$, $V(P_2)$, and $V(P_3)$. By symmetry, we may assume that $N(x) \cap V(P_3) = \emptyset$ and that $N(y) \cap V(P_1) = \emptyset$.

Assume that $j_x < k_x - 1$ and $c_y < d_y - 1$; this means that x has two nonadjacent neighbors in P_1 and y has two nonadjacent neighbors in P_3 . Since both x and y have neighbors in P_2 , there exists a path R between x and y with interior in P_2^* . Then the three paths $x-f_{j_x}-P_1-a_1-a_3-P_3-g_{c_y}-y$, $x-f_{k_x}-P_1-b_1-b_3-P_3-g_{d_y}-y$, and $x-R-y$ form a

theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{a_2, b_2, f_{j_x+1}\}$, K' is smaller than K , a contradiction. This implies that either x does not have two nonadjacent neighbors in P_1 or y does not have two nonadjacent neighbors in P_3 . By symmetry we may assume that x does not have two nonadjacent neighbors in P_1 , so $j_x \geq k_x - 1$. By Theorem 4.2, $N(x) \cap V(P_1)$ does not consist of two adjacent vertices. Therefore, $|N(x) \cap V(P_1)| = 1$, and $j_x = k_x$.

Since x is K -major, by Theorem 4.1, $|N(x) \cap V(P_2)| \geq 2$, and by Theorem 4.2, $N(x) \cap V(P_2)$ does not consist of two adjacent vertices. Therefore, x has two nonadjacent neighbors in P_2 , and $s_x < t_x - 1$.

First assume that $|N(y) \cap V(P_2)| = 1$. Then $s_y = t_y$, and by Theorems 4.2 and 4.1, y has two nonadjacent neighbors in P_3 , so $c_y < d_y - 1$. If $s_y < t_x - 1$, then the paths $y-g_{d_y}-P_3-b_3-b_2-P_2-h_{t_x}-x$, $y-g_{c_y}-P_3-a_3-a_1-P_1-f_{j_x}-x$ and a path between x and y with interior in $V(h_{s_y}-P_2-h_{s_x})$ form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{h_{t_x-1}, a_2, b_1\}$, K' is smaller than K , a contradiction. Therefore, $s_y \geq t_x - 1$. From the symmetry, $s_y \leq s_x + 1$. So we have that $t_x - 1 \leq s_y \leq s_x + 1$ and $s_x < t_x - 1$, from which it follows that $s_x + 1 = s_y = t_x - 1$. If x is not adjacent to h_{s_x+1} , then $|N(x) \cap V(P_2)| = 2$ and $N(x) \cap V(P_3) = \emptyset$, contradicting Theorem 4.2. Therefore, x is adjacent to h_{s_x+1} . The triangles $\{h_{s_x}, h_{s_x+1}, x\}$ and $\{a_1, a_2, a_3\}$ and the paths $h_{s_x}-P_2-a_2$, $h_{s_x+1}-y-g_{c_y}-P_3-a_3$, and $x-f_{j_x}-P_1-a_1$ form a prism, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{b_1, b_2, b_3\}$, K' is smaller than K , a contradiction. This proves that $|N(y) \cap V(P_2)| > 1$.

By Theorem 4.2, y has more than two neighbors in P_2 , and $s_y < t_y - 1$. Let Q be the path $x-f_{j_x}-P_1-a_1-a_3-P_3-g_{c_y}-y$. If there are two paths A and B between x and y whose interiors are anticomplete to each other and contained in P_2^* , then the paths Q , A , and B form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{b_1, b_2, b_3\}$, K' is smaller than K , a contradiction. Therefore, there are no two paths between x and y whose interiors are anticomplete to each other and contained in P_2^* . Then by Theorem 4.3, either $N(y) \cap V(P_2) \subseteq \{h_1, \dots, h_{s_x+1}\}$ or $N(y) \cap V(P_2) \subseteq \{h_{t_x-1}, \dots, h_m\}$. By symmetry, we may assume that $N(y) \cap V(P_2) \subseteq \{h_1, \dots, h_{s_x+1}\}$, which means that $t_y \leq s_x + 1$.

If $t_y \leq s_x$, then the paths $y-h_{s_y}-P_2-a_2-a_1-P_1-f_{j_x}-x$, $y-g_{d_y}-P_3-b_3-b_2-P_2-h_{t_x}-x$ and a path between x and y with interior in $V(h_{t_y}-P_2-h_{s_x})$ form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{a_3, b_1, h_{s_x+1}\}$, K' is smaller than K , a contradiction. Therefore, $t_y > s_x$, and since $t_y \leq s_x + 1$, it follows that $t_y = s_x + 1$ and y is adjacent to h_{s_x+1} .

Assume that $s_y < s_x - 1$ and $t_x > t_y + 1$. Then the paths $x-f_{j_x}-P_1-a_1-a_2-P_2-h_{s_y}-y$, $x-h_{t_x}-P_2-b_2-b_3-P_3-g_{d_y}-y$ and a path between x and y with interior in $\{h_{s_x}, h_{s_x+1}\}$ form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{a_3, h_{s_x-1}, h_{t_x-1}\}$, K' is smaller than K , a contradiction. Therefore, either $s_y \geq s_x - 1$ or $t_x \leq t_y + 1$.

First assume that $s_y \geq s_x - 1$. Then since $s_x = t_y - 1$ and $s_y < t_y - 1$, this implies that $s_y = s_x - 1$. If y is adjacent to h_{s_x} , then the triangles $\{h_{s_x-1}, h_{s_x}, y\}$ and $\{a_1, a_2, a_3\}$ and the paths $h_{s_x-1}-P_2-a_2$, $h_{s_x}-x-f_{j_x}-P_1-a_1$, and $y-g_{c_y}-P_3-a_3$ form a prism, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{b_1, b_2, b_3\}$, K' is smaller than K , a contradiction. Therefore, y is not adjacent to h_{s_x} . Then since $t_y = s_y + 2$, this means that $|N(y) \cap V(P_2)| = 2$, which contradicts Theorem 4.2. This proves that $s_y < s_x - 1$, so we may now assume that $t_x \leq t_y + 1$. Since $t_y = s_x + 1$ and $t_x > s_x + 1$, this implies that $t_x = s_x + 2$. If x is adjacent to h_{s_x+1} , then the triangles $\{h_{s_x+1}, h_{t_x}, x\}$ and $\{b_1, b_2, b_3\}$ and the paths $h_{s_x+1}-y-g_{d_y}-P_3-b_3$, $h_{t_x}-P_2-b_2$, and $x-f_{k_x}-P_1-b_1$ form a prism, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{a_1, a_2, a_3\}$, K' is smaller than K , a contradiction. Therefore, x is not adjacent to h_{s_x+1} . Then since $t_x = s_x + 2$, this means that $|N(x) \cap V(P_2)| = 2$,

which contradicts Theorem 4.2. This proves Theorem 4.5. $\square$

THEOREM 4.6. *Two K -major vertices that are not adjacent have a common neighbor in K .*

Proof. Let x and y be K -major vertices that are not adjacent to each other. Assume that $N(x) \cap V(K) \cap N(y) = \emptyset$. From Theorem 4.5 we may assume that $(N(x) \cup N(y)) \cap V(K) \subset P_1^* \cup P_2^*$. By Theorem 4.4, each of P_1^* and P_2^* contains a neighbor of x and a neighbor of y . Let C be the cycle induced by G on $V(P_1) \cup V(P_2)$. A *gap* is a minimal path in C containing a neighbor of x and a neighbor of y . Every edge of C is in at most one gap, and any vertex common to two distinct gaps is an end of both. Since x and y have no common neighbor in C , there is an even number of gaps. For a gap P , let P^+ be the path between x and y whose interior is $V(P)$. Let $R_1, \dots, R_k$ be the gaps of C numbered in their order on C , such that the vertex set of the path of C sharing one end with R_i and the other end with R_{i+1} , and with interior disjoint from all the gaps, is anticomplete to y if i is odd and anticomplete to x if i is even. Call a gap a *small gap* if it consists of one edge and shares both of its ends with other gaps.

CLAIM 4.2. *There are at most six gaps.*

Suppose there are at least eight gaps. Assume that for some i , R_i is not a small gap. Then the paths R_{i+1}^+, R_{i+4}^+ , and R_{i+7}^+ form a theta, K' . It is smaller than K because each of R_i , $R_{i+2} \cup R_{i+3}$, and $R_{i+5} \cup R_{i+6}$ contains at least one vertex not in $V(K')$. Therefore, every gap is small. Then every edge of C is in a gap and every gap consists of one edge, so $V(C) \subseteq N(x) \cup N(y)$. But this contradicts the fact that x and y have no neighbors in $\{a_1, a_2, b_1, b_2\} \subseteq C$. This proves Claim 4.2.

CLAIM 4.3. *There are at most four gaps.*

Suppose there are six gaps. Assume that no gap is small. The paths R_1^+, R_3^+ , and R_5^+ form a theta. It is smaller than K because R_2, R_4 , and R_6 each contains at least one vertex not in the theta since each is not a small gap. So there is a small gap; by symmetry let it be R_1 . Assume next that none of R_2, R_4 , or R_6 is a small gap. Then the paths R_1^+, R_3^+ , and R_5^+ form a theta smaller than K . Therefore one of R_2, R_4 , or R_6 is small.

First assume that R_2 is a small gap. Then $R_1 \cup R_2$ is a 3-vertex path whose vertices are contained in $N(x) \cup N(y)$, and so it is a subpath of either P_1 or P_2 in K , because x and y have no neighbors in $\{a_1, a_2, b_1, b_2\}$. By symmetry assume it is a subpath of P_1 . Both x and y have a neighbor in P_2 , so there is a path Q between them with interior in P_2 . Let u be the vertex that is an end of both R_1 and R_2 . The three paths $u-R_1^+-y$, $u-R_2^+-y$, and $u-x-Q-y$ form a theta, K' . Since $V(K')$ is disjoint from $\{a_1, a_2, a_3, b_1, b_2, b_3\}$, it follows that the theta K' is smaller than K , a contradiction. This proves that R_2 is not a small gap. From the symmetry, we deduce that R_6 is not a small gap either.

Therefore R_4 is a small gap. From the argument in the previous paragraph with R_4 playing the role of R_1 , we deduce that neither R_3 nor R_5 is a small gap. Since R_1 is a small gap and R_5 is not, we deduce that $V(R_1) \cap V(R_6) \neq \emptyset$, and that $V(R_6)$ is anticomplete to $V(R_4)$. Let $V(R_1) \cap V(R_6) = \{v\}$. The paths $x-R_4^+-y-v$, $R_6^+ \setminus y$, and $R_1^+ \setminus y$ form a theta smaller than K , a contradiction.

This proves Claim 4.3.

CLAIM 4.4. *There are at most two gaps.*

Assume there are four gaps. Let $Q_1, \dots, Q_4$ be subpaths of C such that none of their interiors meet any gap, and such that for $1 \leq i \leq 3$, one end of Q_i is an end of R_i and the other is an end of R_{i+1} , while the ends of Q_4 are ends of R_1 and R_4 . The

neighbors of x in C are contained in $V(Q_1) \cup V(Q_3)$ and the neighbors of y in C are contained in $V(Q_2) \cup V(Q_4)$.

Assume that Q_1 and Q_3 both have length at most one. Then since x has neighbors in both P_1 and P_2 but no neighbors in $\{a_1, a_2, b_1, b_2\}$, one of Q_1 and Q_3 is a subpath of P_1 and the other is a subpath of P_2 . But then x has at most two neighbors in each of P_1 and P_2 , contradicting Theorems 4.1 or 4.2. This shows that at least one of Q_1 and Q_3 has length at least two. By symmetry, we may assume that Q_1 does. Therefore $V(R_1)$ is anticomplete to $V(R_2)$.

By the same argument applied to Q_2 and Q_4 , we deduce that at least one of Q_2 and Q_4 has length at least two. From the symmetry, we may assume that Q_2 has length at least two, and therefore $V(R_2)$ is anticomplete to $V(R_3)$. Since the paths R_1^+ , R_2^+ , and R_3^+ do not form a theta smaller than K , it follows that $|V(Q_3)| = |V(Q_4)| = 1$ and R_4 is a small gap. Let $V(Q_3) = \{q_3\}$ and $V(Q_4) = \{q_4\}$. Since $V(R_2)$ is anticomplete to both $V(R_1)$ and $V(R_3)$, the paths $x-q_3-q_4$, $R_1^+ \setminus y$, and $x-R_2^+-y-q_4$ form a theta smaller than K , a contradiction. This proves Claim 4.4.

Since x and y each have a neighbor in C , there are at least two gaps; so from Claim 4.4 there are two gaps. Since x and y each have neighbors in both P_1 and P_2 , there is a gap contained in P_1 and a gap contained in P_2 . The subgraph $C \setminus (R_1 \cup R_2)$ is the disjoint union of two paths, Q_1 and Q_2 , such that Q_1 contains no neighbors of x and Q_2 contains no neighbors of y . By Theorems 4.1 and 4.2, each of x and y has two nonadjacent neighbors in one of P_1 and P_2 , so each of x and y has a neighbor in C that is anticomplete to $V(R_1) \cup V(R_2)$. Therefore, there is a path S between x and y whose interior is anticomplete to $V(R_1) \cup V(R_2)$ and is contained in $V(K) \setminus (V(R_1) \cup V(R_2))$. The paths S , R_1^+ , and R_2^+ form a theta smaller than K , a contradiction. This proves Theorem 4.6. $\square$

A *broom* is a vertex $v \in V(G)$ such that, for some $\{i_1, i_2, i_3\} = \{1, 2, 3\}$,

- $N(v) \cap V(P_{i_1}) = \emptyset$,
- $|N(v) \cap V(P_{i_2})| = 1$, and
- $N(v) \cap V(P_{i_3})$ is the vertex set of a 3-vertex path.

THEOREM 4.7. *If x and y are two K -major vertices that are not adjacent to each other, then one of them is a broom, and for some i, j such that $1 \leq i < j \leq 3$, $N(x) \cap V(P_i) \cap N(y) \neq \emptyset$ and $N(x) \cap V(P_j) \cap N(y) \neq \emptyset$.*

Proof. By Theorem 4.5, we may assume that $N(x) \cap V(K) \subseteq P_1^* \cup P_2^*$ and $N(y) \cap V(K) \subseteq P_1^* \cup P_2^*$. By Theorem 4.4, x and y each have a neighbor both in P_1^* and in P_2^* . Also, by Theorem 4.6, x and y have a common neighbor in K , and by the symmetry between P_1 and P_2 we may assume that $f_r \in P_1^*$ is a neighbor of both x and y .

Let A be the path induced by G on $V(P_1) \cup V(P_2) \setminus \{f_{r-1}, f_r, f_{r+1}\}$.

CLAIM 4.5. *x and y do not both have a neighbor in $V(A) \cap V(P_1)$.*

Suppose both x and y have a neighbor in $V(A) \cap V(P_1)$. Then there is a path, S , between x and y with interior in $(V(A) \cap V(P_1)) \cup V(P_3)$. Since x and y each have a neighbor in P_2^* , there is a path R between them with interior in P_2^* . The paths S , $x-f_r-y$, and R form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{f_{r-1}, f_{r+1}, a_2, b_2\}$, K' is smaller than K , a contradiction. This proves Claim 4.5.

CLAIM 4.6. *If exactly one of x and y has a neighbor in $V(A) \cap V(P_1)$, then the theorem holds.*

Assume that exactly one of x and y has neighbors in $V(A) \cap V(P_1)$. By symmetry between x and y we may assume that $N(x) \cap V(A) \cap V(P_1) \neq \emptyset$ and $N(y) \cap V(A) \cap V(P_1) = \emptyset$. By Theorems 4.1 and 4.2, y has two nonadjacent neighbors in at least one of P_1 and P_2 . Since y has no neighbors in $V(A) \cap V(P_1)$, this means that it either has

two nonadjacent neighbors in P_2 , so $t_y > s_y + 1$, or it is adjacent to both f_{r-1} and f_{r+1} .

First assume that $t_y > s_y + 1$. Since $N(x) \cap V(A) \cap V(P_1) \neq \emptyset$ and $N(x) \cap V(P_2) \neq \emptyset$, each of x and y has two nonadjacent neighbors in A . If there are two paths S and R between x and y whose interiors are anticomplete to each other and are contained in $V(A)$, then the paths $x-f_{r-1}-y$, S , and R form a theta smaller than K , a contradiction. Therefore, by Theorem 4.3 we may assume that $N(x) \cap V(A) \subseteq V(a_1-P_1-f_{r-2}) \cup V(a_2-P_2-h_{s_y+1})$, which means that $t_x \leq s_y + 1$. Since $N(x) \cap V(A) \cap V(P_1) \neq \emptyset$, $j_x < r - 1$. If $s_x < t_y - 1$, then the paths $x-f_{r-1}-y$, $x-f_{j_x}-P_1-a_1-a_3-P_3-b_3-b_2-P_2-h_{t_y}-y$ and a path between x and y with interior in $V(h_{s_x}-P_2-h_{s_y})$ form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{a_2, b_1, f_{r-1}\}$, K' is smaller than K , a contradiction. Therefore, $s_x \geq t_y - 1$, and since $t_x \leq s_y + 1$ and $t_y > s_y + 1$, it follows that $s_y + 1 = t_y - 1$ and $N(x) \cap V(P_2) = \{h_{s_y+1}\}$. By Theorem 4.2, $|N(y) \cap V(P_2)| \neq 2$, so y is adjacent to h_{s_y+1} . Also by Theorem 4.2, $|N(y) \cap V(P_1)| \neq 2$, so y is either complete or anticomplete to $\{f_{r-1}, f_{r+1}\}$.

Suppose that y is $\{f_{r-1}, f_{r+1}\}$ -complete. If x is not adjacent to f_{r+1} , then the triangles $\{f_r, f_{r+1}, y\}$ and $\{b_1, b_2, b_3\}$ and the paths $f_r-x-f_{j_x}-P_1-a_1-a_3-P_3-b_3$, $f_{r+1}-P_1-b_1$, and $y-h_{t_y}-P_2-b_2$ form a prism, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{h_{s_y}, h_{s_y+1}, f_{r-1}\}$, K' is smaller than K , a contradiction. So x is adjacent to f_{r+1} , and the paths $x-f_{r+1}-y$, $x-h_{s_y+1}-y$ and a path between x and y with interior in $V(f_{j_x}-P_1-f_{r-1})$ form a theta smaller than K , a contradiction. Therefore, y is anticomplete to $\{f_{r-1}, f_{r+1}\}$, and so y is a broom and the theorem holds.

Since we have shown that the theorem holds when $t_y > s_y + 1$, we may now assume that $t_y \leq s_y + 1$ and that y is $\{f_{r-1}, f_{r+1}\}$ -complete. By Theorem 4.2, $|N(y) \cap V(P_2)| \neq 2$, so $t_y = s_y$. Since $N(x) \cap V(A) \cap V(P_1) \neq \emptyset$, either $j_x < r - 1$ or $k_x > r + 1$. By symmetry we may assume that $j_x < r - 1$.

Suppose that $|N(x) \cap V(P_2)| = 1$. If x is adjacent to f_{r+1} , then the path $x-f_{r+1}-y$, a path between x and y with interior in $V(f_{j_x}-P_1-f_{r-1})$, and a path between x and y with interior in P_2^* form a theta smaller than K , a contradiction. Therefore, x is not adjacent to f_{r+1} . If $s_x < s_y$, then the triangles $\{f_r, f_{r+1}, y\}$ and $\{b_1, b_2, b_3\}$ and the paths $f_r-x-f_{j_x}-P_1-a_1-a_3-P_3-b_3$, $f_{r+1}-P_1-b_1$, and $y-h_{s_y}-P_2-b_2$ form a prism, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{f_{r-1}, a_2, h_{s_x}\}$, K' is smaller than K , a contradiction. If $s_x > s_y$, then the triangles $\{f_r, f_{r+1}, y\}$ and $\{a_1, a_2, a_3\}$ and the paths $f_r-x-f_{j_x}-P_1-a_1$, $f_{r+1}-P_1-b_1-b_3-P_3-a_3$, and $y-h_{s_y}-P_2-a_2$ form a prism, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{f_{r-1}, b_2, h_{s_x}\}$, K' is smaller than K , a contradiction. This proves that $s_x = s_y$, and y is a broom, so the theorem holds. So we may assume that $|N(x) \cap V(P_2)| > 1$, and it follows from Theorem 4.2 that $s_x < t_x - 1$.

Let Q be the path between x and y with interior in $V(f_{j_x}-P_1-f_{r-1})$. If $t_x \leq s_y$, then the paths Q , $x-h_{t_x}-P_2-h_{s_y}-y$, and $x-h_{s_x}-P_2-a_2-a_3-P_3-b_3-b_1-P_1-f_{r+1}-y$ form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{f_r, a_1, b_2\}$, K' is smaller than K , a contradiction. Therefore, $t_x > s_y$. If $s_x \geq s_y$, then the paths Q , $x-h_{s_x}-P_1-h_{s_y}-y$, and $x-h_{t_x}-P_2-b_2-b_1-P_1-f_{r+1}-y$ form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{a_1, a_2, a_3\}$, K' is smaller than K , a contradiction. Therefore, $s_x < s_y$. If x is not adjacent to h_{s_y} , then the path $h_{s_y}-y-f_r-x$, a path between h_{s_y} and x with interior in $V(h_{s_y}-P_2-h_{s_x})$, and a path between h_{s_y} and x with interior in $V(h_{s_y}-P_2-h_{t_x})$ form a theta smaller than K , a contradiction. Thus x is adjacent to h_{s_y} , and y is a broom, so the theorem holds. This proves Claim 4.6.

CLAIM 4.7. *If neither x nor y has a neighbor in $V(A) \cap V(P_1)$, then the theorem holds.*

Suppose that neither x nor y has a neighbor in $V(A) \cap V(P_1)$. By Theorem 4.2, each of x and y is either complete or anticomplete to $\{f_{r-1}, f_{r+1}\}$. If both x and y are $\{f_{r-1}, f_{r+1}\}$ -complete, then the paths $x-f_{r-1}-y$, $x-f_{r+1}-y$ and a path between x and y with interior in P_2^* form a theta smaller than K , a contradiction. This means that at most one of x and y is $\{f_{r-1}, f_{r+1}\}$ -complete.

Assume that exactly one of x and y is $\{f_{r-1}, f_{r+1}\}$ -complete. By symmetry, we may assume that x is, and that y is anticomplete to $\{f_{r-1}, f_{r+1}\}$. By Theorems 4.1 and 4.2, $|N(y) \cap V(P_2)| \geq 3$, so $s_y + 1 < t_y$.

Suppose that $|N(x) \cap V(P_2)| > 1$; then by Theorem 4.2, $|N(x) \cap V(P_2)| \geq 3$ so $s_x + 1 < t_x$. If there are two paths S and R between x and y whose interiors are anticomplete to each other and contained in P_2^* , then S , R , and $x-f_r-y$ form a theta smaller than K , a contradiction. So from Theorem 4.3, we may assume that $t_x \leq s_y + 1$. From this it follows that $s_x < s_y$. Then the triangles $\{f_r, f_{r+1}, x\}$ and $\{b_1, b_2, b_3\}$ and the paths $f_r-y-h_{t_y}-P_2-b_2$, $f_{r+1}-P_1-b_1$, and $x-h_{s_x}-P_2-a_2-a_3-P_3-b_3$ form a prism, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{a_1, h_{s_y}, f_{r-1}\}$, K' is smaller than K , a contradiction. This proves that $|N(x) \cap V(P_2)| = 1$.

If $s_x < s_y$, then the triangles $\{f_{r-1}, f_r, x\}$ and $\{a_1, a_2, a_3\}$ and the paths $f_{r-1}-P_1-a_1$, $f_r-y-h_{t_y}-P_2-b_2-b_3-P_3-a_3$, and $x-h_{s_x}-P_2-a_2$ form a prism, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{h_{s_y}, b_1, f_{r+1}\}$, K' is smaller than K , a contradiction. Therefore $s_x \geq s_y$, and from the symmetry $s_x \leq t_y$.

If y is not adjacent to h_{s_x} , then $s_y < s_x$ and $t_y > s_x$. The path $y-f_r-x-h_{s_x}$ and two paths between y and h_{s_x} with interiors in $V(h_{s_y}-P_2-h_{s_x})$ and $V(h_{t_y}-P_2-h_{s_x})$ form a theta smaller than K , a contradiction. So y is adjacent to h_{s_x} . Since x is a broom, the theorem holds.

Since we have shown that the theorem holds when one of x and y is $\{f_{r-1}, f_{r+1}\}$ -complete, we may now assume that $\{x, y\}$ is anticomplete to $\{f_{r-1}, f_{r+1}\}$. Now $N(x) \cap V(P_1) = N(y) \cap V(P_1) = \{f_r\}$, so by Theorems 4.1 and 4.2, x and y each have two nonadjacent neighbors in P_2 . If there are two paths S and R between x and y whose interiors are anticomplete to each other and contained in P_2^* , then the paths S , R , and $x-f_r-y$ form a theta smaller than K , a contradiction. Therefore, by Theorem 4.3, we may assume that $t_x \leq s_y + 1$.

If $t_x \leq s_y$, then the paths $x-f_r-y$, $x-h_{t_x}-P_2-h_{s_y}-y$, and $x-h_{s_x}-P_2-a_2-a_3-P_3-b_3-b_2-P_2-h_{t_y}-y$ form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{a_1, b_1, h_{s_y+1}\}$, K' is smaller than K , a contradiction. Therefore $t_x > s_y$ so $t_x = s_y + 1$.

If $s_x < s_y - 1$ and $t_y > t_x + 1$, then the paths $x-f_r-y$, $x-h_{s_x}-P_2-a_2-a_3-P_3-b_3-b_2-P_2-h_{t_y}-y$ and a path between x and y with interior in $\{h_{s_y}, h_{t_x}\}$ form a theta, K' . Since $V(K') \subseteq V(K) \cup \{x, y\} \setminus \{a_1, b_1, h_{s_y-1}\}$, K' is smaller than K , a contradiction. Therefore, either $s_x \geq s_y - 1$ or $t_y \leq t_x + 1$. From the symmetry we may assume that $s_x \geq s_y - 1$. Then since $s_x < t_x - 1$ and $t_x = s_y + 1$, it follows that $s_x = s_y - 1$. If x is not adjacent to h_{s_y} , then $|N(x) \cap V(P_2)| = 2$, contradicting Theorem 4.2. So x is adjacent to h_{s_y} . Then x is a broom and the theorem holds. This proves Claim 4.7.

Now Theorem 4.7 follows from Claims 4.5, 4.6, and 4.7. $\square$

Let M be the subgraph of G induced on the set of K -major vertices. The following algorithm constructs a superset of $V(M)$ that is disjoint from $V(K)$ and is used for cleaning.

THEOREM 4.8. *There is an algorithm with the following specifications:*

- **Input:** A graph G .
- **Output:** A sequence of subsets $X_1, \dots, X_r$ of $V(G)$, with $r \leq 2|V(G)|^6$, such that for every smallest prism K in G , if K has a tidy frame, then one of $X_1, \dots, X_r$ is disjoint from $V(K)$ and contains all K -major vertices.
- **Running Time:** $O(|V(G)|^7)$.

Proof. The algorithm is as follows. For each triple of vertices (a, b, c) compute the set U of vertices complete to $N(a, b) \cup \{c\}$. List all subsets $W \subseteq V(G)$ of at most three vertices. For each pair of (a, b, c) and W , compute the subset

$$N(a, b) \cup \{c\} \cup U \setminus W.$$

Label the subsets generated $Y_1, \dots, Y_s$, and label the list of subsets $N(a, b)$ by $N_1, \dots, N_p$. Enumerate all quintuples of vertices (u, w, z_1, z_2, z_3) . For each, compute the subset

$$Z(u, w, z_1, z_2, z_3) = \{u\} \cup (N(u) \setminus \{w, z_1, z_2, z_3\}) \cup N(w, z_1) \cup N(w, z_2) \cup N(w, z_3).$$

Label the subsets generated $Z_1, \dots, Z_t$. Output the subsets

$$\emptyset, N_1, \dots, N_p, Y_1, \dots, Y_s, Z_1, \dots, Z_t.$$

This completes the description of the algorithm.

Since $p \leq |V(G)|^2$, $s \leq |V(G)|^6$, and $t \leq |V(G)|^5$, the number of subsets in the output sequence is $r \leq 2|V(G)|^6$. For each triple (a, b, c) , the subset U can be computed in quadratic time and the subsets W can be enumerated in cubic time. Computing $N(a, b) \cup \{c\} \cup U \setminus W$ takes linear time, so the time taken to generate the sequence $N_1, \dots, N_p, Y_1, \dots, Y_s$ is $O(|V(G)|^7)$. Since each of the $|V(G)|^5$ subsets Z can be computed in linear time, their computation does not affect the total running time.

Let K be a smallest prism in G , and assume K has a tidy frame. Let M be the subgraph induced by G on the set of K -major vertices. Suppose M is the complete graph. We may assume that there exists a set $\{a, b\} \subseteq V(K)$ that is not a subset of the vertex set of a 3-vertex subpath of K , because if no such choice is possible, then $V(M) = \emptyset$, which is in the output. We may also assume that, subject to these conditions, $\{a, b\}$ is chosen with $|N(a, b)|$ maximum. Then $N(a, b) \subseteq V(M)$. If $N(a, b) = V(M)$, then $V(M)$ is in the output sequence. Otherwise there is a vertex $c \in V(M) \setminus N(a, b)$. Then since $N(a, b) \cup \{c\} \subseteq V(M)$ and M is a complete graph, the set U of vertices complete to $N(a, b) \cup \{c\}$ contains $V(M) \setminus (N(a, b) \cup \{c\})$; so

$$V(M) \subseteq N(a, b) \cup \{c\} \cup U.$$

If $U \cap V(K)$ is not a subset of the vertex set of a 3-vertex path of K , then there are two vertices $a', b' \in U \cap V(K)$ such that $\{a', b'\}$ is not a subset of the vertex set of a 3-vertex path of K . Since $N(a, b) \cup \{c\} \subseteq N(a', b')$, this contradicts the choice of a and b . So $|U \cap V(K)| \leq 3$ and we can choose $W = U \cap V(K)$. Then the subset $N(a, b) \cup \{c\} \cup U \setminus W$ contains $V(M)$, and since $(N(a, b) \cup \{c\}) \cap V(K) = \emptyset$, $N(a, b) \cup \{c\} \cup U \setminus W$ is disjoint from $V(K)$. The subset in the output corresponding to this choice of (a, b, c) and W contains $V(M)$ and is disjoint from $V(K)$.

Therefore, we may assume that M is not a complete graph. Then there exist two vertices in $V(M)$ that are not adjacent to each other. By Theorem 4.7, one of these is a broom. So we may choose u to be a broom such that z_1, z_2, z_3 , and w are its neighbors in K and z_1 - z_2 - z_3 is a path of K . By Theorem 4.7, any nonneighbor of u in $V(M)$ is contained in the subset $N(w, z_1) \cup N(w, z_2) \cup N(w, z_3)$. Therefore the subset $Z(u, w, z_1, z_2, z_3)$ contains $V(M)$. Since w is in the interior of a different path of the prism K from that containing $\{z_1, z_2, z_3\}$, the subset $N(w, z_1) \cup N(w, z_2) \cup N(w, z_3)$ is disjoint from $V(K)$. Since $N(u) \cap V(K) = \{w, z_1, z_2, z_3\}$, $Z(u, w, z_1, z_2, z_3)$ is disjoint from $V(K)$. This proves Theorem 4.8. $\square$

5. The complete algorithm.

THEOREM 5.1. *There is an algorithm with the following specifications:*

- **Input:** A graph G .
- **Output:** Either
 - a theta in G , or
 - a determination that there is no smallest theta in G .
- **Running Time:** $O(|V(G)|^{22})$.

Proof. Here is the algorithm. Enumerate all 11-tuples $(a, b, s_1, s_2, s_3, m_1, m_2, m_3, t_1, t_2, t_3)$ of vertices of G such that

- a, b, s_1, s_2 , and s_3 are all distinct;
- a, b, t_1, t_2 , and t_3 are all distinct;
- a is not adjacent to b ; and
- a is complete to $\{s_1, s_2, s_3\}$ and b is complete to $\{t_1, t_2, t_3\}$.

We can find all such 11-tuples in time $O(|V(G)|^{11})$. For each 11-tuple do the following.

Remove $N(a) \setminus \{s_1, s_2, s_3\}$ and $N(b) \setminus \{t_1, t_2, t_3\}$ from $V(G)$, and for $i = 1, 2, 3$, if m_i is adjacent to a and b , then remove $N(m_i) \setminus \{a, b\}$ from $V(G)$. This can be done in linear time, and after this step every smallest theta in G with the 11-tuple as its frame has a tidy frame. Run the algorithm of Theorem 2.2, which takes $O(|V(G)|^4)$ time. Let $X_1, \dots, X_t$ be the subsets in the output; $t \leq |V(G)|^3$. For every triple of subsets $(X_{i_1}, X_{i_2}, X_{i_3})$ from this list, do the following.

Let G_1 be the graph induced by G on $V(G) \setminus X_{i_1}$. Find shortest paths S'_1 between s_1 and m_1 and T'_1 between m_1 and t_1 in G_1 . Next, let G_2 be the graph induced by G_1 on $V(G_1) \setminus X_{i_2}$. Find shortest paths S'_2 between s_2 and m_2 and T'_2 between m_2 and t_2 in G_2 . Finally, let G_3 be the graph induced by G_2 on $V(G_2) \setminus X_{i_3}$. Find shortest paths S'_3 between s_3 and m_3 and T'_3 between m_3 and t_3 in G_3 . Finding these paths takes quadratic time. Finally, test whether the following are true in the graph G :

- For $1 \leq i \leq 3$, $V(S'_i) \setminus \{m_i\}$ and $V(T'_i) \setminus \{m_i\}$ are disjoint and anticomplete to each other, and
- for $1 \leq i, j \leq 3$ with $i \neq j$, $V(S'_i) \cup V(T'_i)$ and $V(S'_j) \cup V(T'_j)$ are disjoint and anticomplete to each other.

If these conditions are satisfied, output that $\{a, b\}$ and the three paths $a-s_1-S'_1-m_1-T'_1-t_1-b$, $a-s_2-S'_2-m_2-T'_2-t_2-b$, and $a-s_3-S'_3-m_3-T'_3-t_3-b$ form a theta and stop. Testing these conditions takes quadratic time. So the time it takes to examine one 11-tuple is $O(|V(G)|^{11})$.

After examining all 11-tuples, output that G contains no smallest theta. The total running time is $O(|V(G)|^{22})$. Now we need to prove that this algorithm is correct. Suppose there is a smallest theta K in G ; let its half-paths be S_1, T_1, S_2, T_2, S_3 , and T_3 . Some 11-tuple chosen is the frame of K , and after the first step of the algorithm, K and every smallest theta in G with the same frame as K have a tidy frame.

By Theorem 2.2, we may choose X_{i_1} such that the graph G_1 contains K and there are no bad shortcuts across S_1 or T_1 in G_1 . Therefore, the subgraph induced by G_1 on $V(K) \setminus (V(S_1) \cup V(T_1)) \cup V(S'_1) \cup V(T'_1)$ is a theta, K_1 , in G_1 . Since S'_1 and T'_1 are shortest paths, $|V(K_1)| \leq |V(K)|$, so K_1 is also a smallest theta in G_1 . Since K_1 is not smaller than K in G it follows that $V(S'_1) \setminus \{m_1\}$ and $V(T'_1) \setminus \{m_1\}$ are anticomplete to each other. We observe that K_1 has the same frame as K , so it has a tidy frame.

Therefore, by Theorem 2.2, for some subset X_{i_2} the graph G_2 contains K_1 , and all shortcuts across S_2 or T_2 in G_2 are good shortcuts across K_1 . Then the subgraph induced by G_2 on $V(K_1) \setminus (V(S_2) \cup V(T_2)) \cup V(S'_2) \cup V(T'_2)$ is a theta, K_2 , in G_2 . Since

S'_2 and T'_2 are shortest paths, $|V(K_2)| \leq |V(K_1)|$, so K_2 is also a smallest theta in G_2 . Since K_2 is not smaller than K_1 in G it follows that $V(S'_2) \setminus \{m_2\}$ and $V(T'_2) \setminus \{m_2\}$ are anticomplete to each other. We observe that K_2 has the same frame as K , so it has a tidy frame.

Finally, by Theorem 2.2 again, for some subset X_{i_3} the graph G_3 contains K_2 , and all shortcuts across S_3 or T_3 in G_3 are good shortcuts across K_2 . Then the subgraph induced by G_3 on $V(K_2) \setminus (V(S_3) \cup V(T_3)) \cup V(S'_3) \cup V(T'_3)$ is a theta, K_3 , in G_3 . Since K_3 is not smaller than K_2 in G it follows that $V(S'_3) \setminus \{m_3\}$ and $V(T'_3) \setminus \{m_3\}$ are anticomplete to each other. We observe that K_3 is the theta that is output by the algorithm when it considers $X_{i_1}, X_{i_2}, X_{i_3}$. This proves that the output of the algorithm is a theta.

Conversely, if the algorithm outputs that there is a theta, then the properties of the 11-tuple and the conditions on the paths that the algorithm checks ensure that this output is actually a theta. $\square$

THEOREM 5.2. *There is an algorithm with the following specifications:*

- **Input:** A graph G .
- **Output:** Either
 - a prism in G , or
 - a determination that there is no smallest prism in G .
- **Running Time:** $O(|V(G)|^{35})$.

Proof. Here is the algorithm. Enumerate all 15-tuples of vertices of G ,

$$(a_1, a_2, a_3, m_1, m_2, m_3, b_1, b_2, b_3, a'_1, a'_2, a'_3, b'_1, b'_2, b'_3),$$

such that

- $a_1, a_2, a_3, b_1, b_2, b_3$ are distinct,
- $a_1, a_2, a_3, a'_1, a'_2, a'_3$ are distinct,
- $b_1, b_2, b_3, b'_1, b'_2, b'_3$ are distinct,
- G induces a triangle on each of $\{a_1, a_2, a_3\}$ and $\{b_1, b_2, b_3\}$, and
- for $i = 1, 2, 3$, a'_i is adjacent to a_i and b'_i is adjacent to b_i .

For each 15-tuple, do the following.

For $i = 1, 2, 3$ remove the subsets $N(a_i) \setminus \{a'_i, a_1, a_2, a_3\}$ and $N(b_i) \setminus \{b'_i, b_1, b_2, b_3\}$ from $V(G)$, and if $m_i = a'_i = b'_i$, then remove $N(m_i) \setminus \{a_i, b_i\}$ from $V(G)$. Now any smallest prism K that has $\{a_1, a_2, a_3, m_1, m_2, m_3, b_1, b_2, b_3\}$ as its frame and contains $\{a'_1, a'_2, a'_3, b'_1, b'_2, b'_3\}$ has a tidy frame.

Run the algorithm of Theorem 4.8 on G . This outputs $O(|V(G)|^6)$ subsets. For each subset Y output, do the following. Let G' be the graph induced by G on $V(G) \setminus Y$. Run the algorithm of Theorem 3.5, which takes $O(|V(G)|^5)$ time and outputs $O(|V(G)|^4)$ subsets. Let $X_1, \dots, X_t$ be the subsets in the output. For every triple of subsets $(X_{i_1}, X_{i_2}, X_{i_3})$ from this list, do the following.

Let G_1 be the graph induced by G' on $V(G') \setminus X_{i_1}$. Find shortest paths S'_1 between a_1 and m_1 and T'_1 between m_1 and b_1 in G_1 . Remove from $V(G_1)$ the set U_1 consisting of all vertices in $V(G_1) \setminus (S'^*_1 \cup T'^*_1 \cup \{a_1, b_1, m_1\})$ that have a neighbor in $S'^*_1 \cup T'^*_1$. Next, let G_2 be the graph induced by G_1 on $V(G_1) \setminus X_{i_2}$. Find shortest paths S'_2 between a_2 and m_2 and T'_2 between m_2 and b_2 in G_2 . Remove from G_2 the set U_2 consisting of all vertices in $V(G_2) \setminus (S'^*_2 \cup T'^*_2 \cup \{a_2, b_2, m_2\})$ that have a neighbor in $S'^*_2 \cup T'^*_2$. Finally, let G_3 be the graph induced by G_2 on $V(G_2) \setminus X_{i_3}$. Find shortest paths S'_3 between a_3 and m_3 and between m_3 and b_3 in G_3 . Finding these paths and removing these subsets take quadratic time.

Finally, for $1 \leq i < j \leq 3$ test whether $V(S'_i) \cup V(T'_i) \cup V(S'_j) \cup V(T'_j)$ is the vertex set of an induced cycle in the graph G . If so, then output that the triangles $\{a_1, a_2, a_3\}$

and $\{b_1, b_2, b_3\}$ and the paths $a_1-S'_1-m_1-T'_1-b_1$, $a_2-S'_2-m_2-T'_2-b_2$, and $a_3-S'_3-m_3-T'_3-b_3$ form a prism and stop. Testing this takes quadratic time. So the time it takes to examine one 15-tuple is $O(|V(G)|^{20})$.

After examining all 15-tuples, output that G contains no smallest prism. The total running time is $O(|V(G)|^{35})$.

Now we need to prove that this algorithm is correct. Suppose there is a smallest prism K in G ; let its half-paths be S_1, T_1, S_2, T_2, S_3 , and T_3 . For some 15-tuple chosen, $\{a_1, a_2, a_3, m_1, m_2, m_3, b_1, b_2, b_3\}$ is the frame of K , and $\{a'_1, a'_2, a'_3, b'_1, b'_2, b'_3\} \subseteq V(K)$. Therefore, after the first step of the algorithm runs for this 15-tuple, K has a tidy frame.

CLAIM 5.1. *Let J be an almost clean smallest prism in a graph H and let S' be a good shortcut across a half-path S of J . Let $U \subset V(H)$ be the set of vertices not in $V(S')$ that have a neighbor in S'^* . Then the subgraph J' induced by H on $(V(J) \setminus V(S)) \cup V(S')$ is an almost clean smallest prism in the graph induced by H on $V(H) \setminus U$.*

Since S' is a good shortcut, the subgraph J' induced by H on $(V(J) \setminus V(S)) \cup V(S')$ is a prism. Since a good shortcut is a shortest path, $|V(J')| \leq |V(J)|$, so J' is a smallest prism in H . We observe that since J' has the same frame as J , it has a tidy frame. Assume that there is a J' -major vertex $v \in V(H) \setminus U$. Then since $v \notin U$, $N(v) \cap S'^* = \emptyset$. Since v is J' -major, this means that $N(v) \cap (V(J') \setminus S'^*)$ is not a subset of the vertex set of a 3-vertex path in J' . Since $N(v) \cap (V(J') \setminus S'^*) = N(v) \cap (V(J) \setminus S^*)$, v is J -major, contradicting the fact that J is almost clean in H . Therefore, all J' -major vertices are contained in U . We observe also that U is disjoint from $V(J')$ because J' is a prism, so J' is an almost clean prism in the graph induced by H on $V(H) \setminus U$. This proves Claim 5.1.

By Theorem 4.8, we can choose Y such that G' contains no K -major vertices. Then K is almost clean in G' , so by Theorem 3.5, we can choose X_{i_1} such that the graph G_1 contains K and there are no bad shortcuts across S_1 or T_1 in G_1 . Let K_1 be the subgraph induced by G_1 on $V(K) \setminus (V(S_1) \cup V(T_1)) \cup V(S'_1) \cup V(T'_1)$. Since K_1 is not smaller than K in G , S'_1 is anticomplete to T'_1 . It now follows from Claim 5.1 applied twice (once for each half-path S_1 and T_1) that, after removing the set U_1 from $V(G_1)$, K_1 is an almost clean smallest prism in G_1 .

Therefore, by Theorem 3.5, we can choose X_{i_2} such that the graph G_2 contains K_1 and there are no bad shortcuts across S_2 or T_2 in G_2 . Let K_2 be the subgraph induced by G_2 on $V(K_1) \setminus (V(S_2) \cup V(T_2)) \cup V(S'_2) \cup V(T'_2)$. Since K_2 is not smaller than K_1 in G , S'_2 is anticomplete to T'_2 . Again it follows from Claim 5.1 applied twice (once for each half-path S_2 and T_2) that, after removing the set U_2 from $V(G_2)$, K_2 is an almost clean smallest prism in G_2 .

Finally, by Theorem 3.5, we can choose X_{i_3} such that the graph G_3 contains K_2 and there are no bad shortcuts across S_3 or T_3 in G_3 . Then it follows that the subgraph K_3 induced by G_3 on $V(K_2) \setminus (V(S_3) \cup V(T_3)) \cup V(S'_3) \cup V(T'_3)$ is a prism. Since K_3 is not smaller than K_2 in G , S'_3 is anticomplete to T'_3 . We observe that since G_3 is an induced subgraph of G , a prism in G_3 is a prism in G , and that K_3 is the prism output by the algorithm when it considers $X_{i_1}, X_{i_2}, X_{i_3}$. This proves that the algorithm outputs a prism.

Conversely, if the algorithm outputs that there is a prism, then the properties of the 15-tuple and the conditions on the paths that the algorithm checks ensure that this output is actually a prism. $\square$

Finally, we can put together the algorithms of Theorems 5.1 and 5.2 to construct

the complete algorithm.

THEOREM 5.3. *There is an algorithm with the following specifications:*

- **Input:** *A graph G .*
- **Output:** *Either*
 - *a theta or a prism in G , or*
 - *a determination that there is no theta or prism in G .*
- **Running Time:** $O(|V(G)|^{35})$.

Proof. First run the algorithm of Theorem 5.1. If it outputs a theta, then output this theta and stop. Otherwise run the algorithm of Theorem 5.2. If it outputs a prism, then output this prism and stop. Otherwise, output that there is no theta or prism in G .

The running time of this algorithm is the maximum of those of Theorems 5.1 and 5.2, which is $O(|V(G)|^{35})$. If this algorithm outputs either a theta or a prism, then it follows from Theorems 5.1 and 5.2 that it is correct. Conversely, if G contains a theta or a prism, then it contains a smallest theta or a smallest prism. Therefore, one of the algorithms of Theorem 5.1 or of Theorem 5.2 will output a theta or a prism, which will then be output by this algorithm. This proves Theorem 5.3. $\square$

REFERENCES

- [1] M. CHUDNOVSKY, G. CORNUÉJOLS, X. LIU, P. SEYMOUR, AND K. VUŠKOVIĆ, *Recognizing Berge graphs*, *Combinatorica*, 25 (2005), pp. 143–186.
- [2] M. CHUDNOVSKY AND P. SEYMOUR, *The three-in-a-tree problem*, submitted.
- [3] M. CONFORTI AND M. R. RAO, *Testing balancedness and perfection of linear matrices*, *Math. Programming*, 61 (1993), pp. 1–18.
- [4] F. MAFFRAY AND N. TROTIGNON, *Algorithms for perfectly contractile graphs*, *SIAM J. Discrete Math.*, 19 (2005), pp. 553–574.

AXIS-ALIGNED SUBDIVISIONS WITH LOW STABBING NUMBERS*

CSABA D. TÓTH†

Abstract. It is shown that for every subdivision of the d -dimensional Euclidean space, $d \geq 2$, into n axis-aligned boxes, there is an axis-parallel line that stabs at least $\Omega(\log^{1/(d-1)} n)$ boxes, and this bound is best possible. In general, it is also shown that for every integer k , $0 < k < d$, there is an axis-aligned k -flat that stabs at least $\Omega(\log^{1/\lfloor (d-1)/k \rfloor} n)$ boxes of the subdivision.

Key words. stabbing number, transversal, subdivision, Euclidean space

AMS subject classifications. 05B45, 05D15, 51M15, 52C22

DOI. 10.1137/06066031X

1. Introduction. An affine subspace in the d -dimensional Euclidean space $\mathbb{R}^d$ *stabs* a set $b \subset \mathbb{R}^d$ if it intersects the relative interior of b but does not contain b . The *stabbing number* (resp., *k-stabbing number*) of a family of objects in $\mathbb{R}^d$ is the maximum number of objects stabbed by a line (resp., k -dimensional affine subspace or *k-flat*). Geometric data structures of low stabbing numbers (e.g., spanning trees, triangulations, and R -trees) have versatile applications in computational geometry. They were used in efficient algorithms for simplex range searching [12], ray shooting [1], collision detection [1, 18], and dynamic point location [7]. They are also related to geometric tomography [10] and the shadow problem [4]. Data structures are often based on subdivisions of the Euclidean space (e.g., triangulations, Kd-trees, binary space partitions).

Following earlier results of Chazelle, Edelsbrunner, and Guibas [11], we address combinatorial problems about the stabbing number of subdivisions of the space. It was shown in [11] that for every subdivision of the plane into n convex cells, there is a line that stabs $\Omega(\log n / \log \log n)$ cells, and this bound is best possible, which means that the plane can be subdivided into n convex cells such that every line stabs $O(\log n / \log \log n)$ cells. This result was recently generalized to every dimension [24]: for every subdivision of $\mathbb{R}^d$ into n convex cells, with $d \geq 2$ fixed, there is a line that stabs $\Omega((\log n / \log \log n)^{1/(d-1)})$ cells, and this bound is best possible.

In this paper, we prove tight bounds on analogous problems in every dimension for the case that both the cells and the stabbing lines are axis aligned. For a subdivision B of $\mathbb{R}^d$ into axis-aligned boxes, we denote by $s_{d,k}(B)$ the *axis-aligned k-stabbing number* of B , which is the maximum number of boxes stabbed by an *axis-aligned k-flat*. In particular, the axis-aligned 1-stabbing number is also called the *axis-parallel stabbing number* of B . Denoting by $|B|$ the size (i.e., number of cells) of a subdivision B , we are interested in the minimum stabbing number over all axis-aligned subdivisions of size n . For an integer $n \in \mathbb{N}$, let

$$s_{d,k}(n) = \min\{s_{d,k}(B) : |B| = n \text{ and } B \text{ is an axis-aligned subdivision of } \mathbb{R}^d\}.$$

*Received by the editors May 19, 2006; accepted for publication (in revised form) February 29, 2008; published electronically July 16, 2008. A preliminary version of this work has been published in *Proceedings of the 9th Workshop on Algorithms and Data Structures (Waterloo, ON, 2005)*, Lecture Notes in Comput. Sci. 3608, Springer, New York, pp. 256–268.
<http://www.siam.org/journals/sidma/22-3/66031.html>

†Department of Mathematics and Statistics, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada (cdtoth@ucalgary.ca).

Contribution. A tight bound $s_{2,1}(n) = \Theta(\log n)$ easily follows from the Erdős–Szekeres theorem [15] and a result of de Berg and van Kreveld [8]. In section 2, we further extend this proof technique to arbitrary dimensions.

THEOREM 1. *For every $d, k \in \mathbb{N}$, $d/2 \leq k < d$, we have $s_{d,k}(n) = \Theta(\log n)$.*

In sections 3 and 4 we prove the main result of this paper, which is an asymptotically tight bound on the axis-parallel stabbing number of axis-aligned subdivisions in higher dimensions.

THEOREM 2. *For every $d \in \mathbb{N}$, we have*

$$s_{d,1}(n) = \Theta\left(\log^{1/(d-1)} n\right).$$

In other words, if d is fixed and every axis-parallel line stabs at most i boxes of a d -dimensional axis-aligned subdivision B , then the size of the subdivision is bounded by $|B| = \exp(O(i^{d-1}))$, and this bound is best possible. We prove the lower bound $s_{d,1}(n) = \Omega(\log^{1/(d-1)} n)$ in section 3 by generalizing the methods presented in section 2. The upper bound $s_{d,1}(n) = O(\log^{1/(d-1)} n)$ follows from an intricate construction that we describe in section 4. The lower bound also generalizes to k -stabbing numbers for every $k \in \mathbb{N}$, $0 < k < d$, and gives

$$s_{d,k}(n) = \Omega\left(\log^{1/\lfloor (d-1)/k \rfloor} n\right).$$

We conjecture this bound to be tight apart from a constant factor for every $d, k \in \mathbb{N}$, $0 < k < d$. The conjecture is confirmed in the cases that $k = 1$ (Theorem 2) and $k \geq d/2$ (Theorem 1), but it remains open for $1 < k < d/2$.

We develop nontrivial techniques to prove our upper and lower bounds. Theorem 1 indicates that Theorem 2 does not follow from a straightforward induction on the dimension. By Theorem 1, there is an axis-aligned subdivision B_n of size n in $\mathbb{R}^3$ such that any axis-aligned plane intersects only $O(\log n)$ boxes, and there exists an axis-aligned subdivision of size $O(\log n)$ in the plane such that every axis-parallel line stabs only $O(\log \log n)$ boxes. By Theorem 2, however, there is an axis-parallel line that stabs $\Omega(\sqrt{\log n})$ boxes of B_n . Our proof of Theorem 2 does use induction on the dimension and projections along the coordinate axes, but in a nontrivial way.

1.1. Related work. In a pivotal result, Chazelle and Welzl [12, 25] proved that for every fixed $d \in \mathbb{N}$, any n points in $\mathbb{R}^d$ can be connected by a straight line spanning tree such that every hyperplane stabs $O(n^{1-1/d})$ edges of the tree, and this bound is best possible. Matoušek [21] showed that for any r , $1 \leq r^d \leq n$, a set of n points in $\mathbb{R}^d$ can be partitioned into r^d subsets, each of size $\Theta(n/r^d)$, such that every hyperplane stabs the convex hull of $O(r^{d-1})$ subsets. If the points are uniformly distributed in a cube in $\mathbb{R}^d$, then a subdivision into r^d congruent cubes gives such a partition of the point set; Matoušek's partition, however, does not typically correspond to a subdivision of the space.

Researchers have also considered *subdivisions* with low stabbing numbers. Agarwal, Aronov, and Suri [2] proved that n points in $\mathbb{R}^2$ and in $\mathbb{R}^3$ can be triangulated (with Steiner points) such that every line stabs $O(\sqrt{n} \cdot \log n)$ simplices. They also presented point sets in general position where the stabbing number of any triangulation is $\Omega(\sqrt{n})$. The stabbing number of a *Delaunay* triangulation of n points in $\mathbb{R}^d$, however, may be $\Theta(n^{\lceil d/2 \rceil})$ in the worst case [23]. Bose and Devroye [9] showed recently that the expected stabbing number of the Delaunay triangulation of n points chosen uniformly at random from a compact convex set in the plane is $\Theta(\sqrt{n})$ with high probability.

Hershberger and Suri [19] studied stabbing numbers restricted to the interior of a simple polygon. They showed that every simple polygon with n vertices has a triangulation (with Steiner points) of size $O(n)$ such that any segment lying entirely in the polygon stabs $O(\log n)$ triangles. De Berg and van Kreveld [8] proved that every simple rectilinear polygon with n vertices can be subdivided into $O(n)$ rectangles such that any axis-parallel line segment lying entirely in the polygon stabs $O(\log n)$ rectangles. Both bounds are best possible. These results no longer hold if we drop the condition that the stabbing segment lies entirely in the polygon: a line may cross $\Theta(n)$ triangles in any triangulation of a comb polygon with n vertices (Figure 1(a)), and a horizontal line may cross $\Theta(n)$ rectangles in any rectangle decomposition of a rectilinear comb polygon with n vertices (Figure 1(b)).

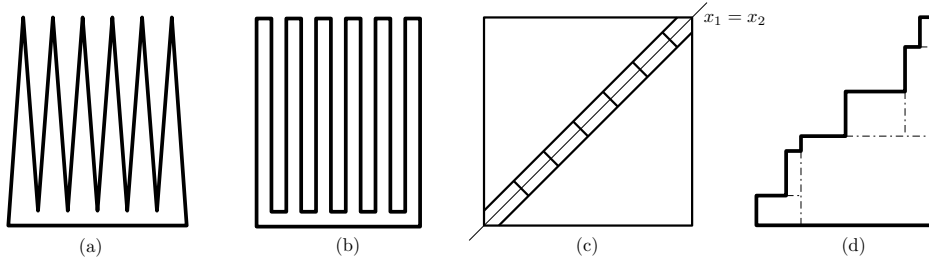


FIG. 1. (a) A comb polygon; (b) a rectilinear comb polygon; (c) a convex subdivision of a box where the axis-parallel stabbing number is 4; and (d) a staircase polygon with 5 reflex vertices.

Instead of the stabbing number of triangulations in $\mathbb{R}^d$, one can consider the stabbing number of lower-dimensional but noncrossing simplices on a given vertex set. Pach [22] asked: What is the minimum k -stabbing number of e distinct r -simplices with disjoint relative interiors and having n distinct vertices in $\mathbb{R}^d$, for $1 \leq r < k \leq d$? Some initial bounds were obtained by Dey and Pach [13].

Axis-aligned box configurations with low axis-aligned stabbing numbers were studied in various contexts. Agarwal et al. [3] constructed an axis-aligned rectangle hierarchy for an input of n disjoint boxes in the plane such that the 0-stabbing number of the rectangles is $O(\log^2 n)$; that is, any point lies in $O(\log^2 n)$ rectangles. De Berg et al. [7] studied similar questions for an input of *overlapping* objects in the plane. Gaur, Ibaraki, and Krishnamurti [17] gave a d -approximation algorithm for computing the minimum number of axis-aligned hyperplanes that jointly stab all input rectangles in $\mathbb{R}^d$.

In this paper, we give worst-case optimal combinatorial bounds for the axis-parallel stabbing numbers in every dimension. Note that it is easy to compute the axis-aligned k -stabbing number of any given d -dimensional axis-aligned subdivision of size n in $O(n^{d-k})$ time if $d, k \in \mathbb{N}$, $0 < k < d$, are constants. It is enough to construct the orthogonal projection of the subdivision into $(n-k)$ -flats in all $\binom{d}{k}$ directions and find the maximal multiplicity of a cell in $\mathbb{R}^{d-k}$. Each projection consists of up to $O(n^{d-k})$ cells. However, the computational complexity of this problem is not known if d and k are part of the input.

Fekete, Lübbecke, and Meijer [16] have recently proved that it is NP-complete to compute the minimum (axis-parallel) stabbing number of perfect matchings or triangulations for a given point set in the plane. Aronov and Fortune [6] and Aronov et al. [5] considered the average stabbing number of triangulations of polygonal scenes.

1.2. Further directions. In this paper, we prove lower and upper bounds on $s_{d,k}(n)$, the minimum axis-aligned k -stabbing number of axis-aligned subdivisions of

size n . One could consider, instead, the axis-aligned k -stabbing number of subdivisions of $\mathbb{R}^d$ into *convex* cells. For every $d \geq 2$ and $n \in \mathbb{N}$, it is easy to construct subdivisions of $\mathbb{R}^d$ into n full-dimensional convex cells with an axis-aligned k -stabbing number at most $d+2$. For example, consider a thin cylinder whose axis is the line $x_1 = x_2 = \dots = x_d$ and whose base is a regular d -gon; partition the cylinder into $n-d$ long cells and divide the complement of the cylinder into d unbounded convex cells (Figure 1(c)).

All our bounds on $s_{d,k}(n)$ also hold for the k -stabbing number (with k -flats of arbitrary orientations) of axis-aligned subdivisions, but we do not know whether they are tight or not for $d \geq 3$. For $d = 2$, Chazelle, Edelsbrunner, and Guibas [11] gave an $\Omega(\log n)$ bound on the stabbing number of axis-aligned subdivisions of size n in the plane and showed that this bound is best possible. A tight bound of $\Omega((\log n / \log \log n)^{1/(d-1)})$ was derived [11, 24] for the stabbing number of convex subdivisions in $\mathbb{R}^d$ for every $d \geq 2$. No nontrivial bound is known, however, for the k -stabbing number of convex subdivisions for $1 < k < d$.

1.3. Definitions. Every logarithm in this paper is of base 2, and we define the exponential function as $\exp(x) = 2^x$. For two real numbers $a, c \in \mathbb{R}$, let (a, c) and $[a, c]$ denote the open and closed intervals, respectively, between a and c . An *axis-aligned box* (for short, *box*) in $\mathbb{R}^d$ is the cross product of d (possibly infinite) closed intervals: $b = \prod_{i=1}^d [a_i, c_i]$, where $a_i \in \{-\infty\} \cup \mathbb{R}$ and $c_i \in \mathbb{R} \cup \{\infty\}$ for $i = 1, 2, \dots, d$. The interval $[a_i, c_i]$ is the i th *extent* of b .

An axis-aligned subdivision B in $\mathbb{R}^d$ is a set of interior disjoint boxes whose union is $\mathbb{R}^d$. We denote by $|B|$ the size of B . For a polyhedron P , let $B|P := \{b \cap P : b \in B, \text{int}(b \cap P) \neq \emptyset\}$ denote the restriction of B to P . For every $k \in \mathbb{N}$, $0 \leq k < d$, we say that a k -flat f *stabs* a d -dimensional box b if f intersects the interior of b . In our upper bound constructions, we define subdivisions restricted to a (bounding) box. Any axis-aligned subdivision of a box can, of course, be extended to fill the entire space with the same axis-aligned stabbing numbers.

2. k -stabbing numbers for $d/2 \leq k < d$. In this section, we study the axis-aligned k -stabbing numbers of axis-aligned subdivisions in the case that k is at least half of the dimension d . The tight bound $s_{2,1}(n) = \Theta(\log n)$ follows from an easy combination of previous results, which we present in subsection 2.1. Then in subsection 2.2, we generalize these arguments and establish Theorem 1.

2.1. Axis-aligned subdivisions in the plane. De Berg and van Kreveld [8] proved that the axis-parallel stabbing number of every axis-aligned subdivision of a staircase polygon with $n-1$ reflex vertices (see Figure 1(d)) is $\Omega(\log n)$; vice versa, every staircase polygon with $n-1$ reflex vertices can be subdivided into n rectangles so that the axis-parallel stabbing number of the subdivision is $O(\log n)$. (Note, however, that there are staircase polygons with $n-1$ reflex vertices, where a line of slope 1 stabs at least n cells of any convex subdivision.) We also evoke a result of Erdős and Szekeres [15].

THEOREM 3 (Erdős and Szekeres). *Every sequence of n reals, $n \in \mathbb{N}$, has a monotone subsequence of size $\lceil \sqrt{n} \rceil$.*

To show $s_{2,1}(n) = \Omega(\log n)$, consider an axis-aligned subdivision B of the plane into n boxes. Choose a *reference point* in the interior of every box such that all coordinates of all reference points are pairwise distinct. Let $\hat{B}$ be a subdivision of size $4n$ obtained from B by splitting every box of B into four boxes along the two axis-parallel lines passing through the reference point in the interior of that box. Order the reference points according to their x -coordinates. By Theorem 4, there

is a subsequence R of size $\lceil \sqrt{n} \rceil$ with strictly monotonically increasing or decreasing y -coordinates. We may assume without loss of generality that both coordinates are strictly monotonically increasing in R . Let π be an infinite monotonically increasing axis-aligned polyline that visits every point in R such that its lower-right vertices are the reference points in R . Denote by F the area below π , which is a staircase polygon whose reflex vertices are the reference points in R . The subdivision $\hat{B}|F$ is a subdivision of F into less than $4n$ boxes. By [8], an axis-parallel line stabs $\Omega(\log n)$ rectangles of $\hat{B}|F$. This line must stab at least $\Omega(\log n)$ boxes of B , as well.

To prove $s_{2,1}(n) = O(\log n)$, we subdivide the plane into n rectangles. For every $n \geq 3$, subdivide the plane into two polygonal regions by a monotonically increasing axis-aligned polyline with $n-2$ vertices. The two polygonal regions on the two sides of the polyline have $\lfloor (n-2)/2 \rfloor$ and $\lceil (n-2)/2 \rceil$ reflex vertices, respectively. By [8], they can be subdivided into $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$ boxes, respectively, with each subdivision having a stabbing number of at most $O(\log n)$. Let B be the union of the two subdivisions: it is a subdivision of the plane into n boxes with stabbing number $O(\log n)$, as required.

2.2. Proof of Theorem 1. De Bruijn (see Kruskal [20]) generalized the Erdős–Szekeres theorem (his result follows from Dilworth’s theorem [14], as well).

THEOREM 4 (De Bruijn). *Every n -element sequence of d -dimensional vectors has a subsequence of size $\lceil n^{2^{1-d}} \rceil$ which is monotone in each coordinate. This bound is tight in the worst case.*

The argument of de Berg and van Kreveld also generalizes to k -flats in $\mathbb{R}^d$ if $k \geq d/2$.

LEMMA 5. *For every fixed $d, k \in \mathbb{N}$, $d/2 \leq k < d$, we have $s_{d,k} = \Omega(\log n)$.*

Proof. Consider a subdivision of $\mathbb{R}^d$ into n axis-aligned boxes. Choose a sufficiently small axis-aligned *reference cube* in the interior of each box of B such that all extents of all reference cubes are pairwise nonoverlapping and disjoint from any coordinates of the vertices of the boxes in B . Pick a *reference point* in the interior of each reference cube. By Theorem 4, there is a coordinatewise strictly monotone sequence $R = (r_i : i = 1, 2, \dots, \lceil n^{2^{1-d}} \rceil)$ of reference points. For $1 \leq i \leq j \leq m$, let $u(i, j)$ denote the box spanned by the reference cubes containing r_i and r_j .

Let m be the power of 2 such that $m \leq \lceil n^{2^{1-d}} \rceil < 2m$. Build a binary hierarchy of height $\log m$ on the reference cubes $u(i, i)$, $i = 1, 2, \dots, m$. For $j = 0, 1, \dots, \log m$, let level j of the hierarchy consist of the set $U_j = \{u(i2^j + 1, i2^j + 2^j) : i = 0, 1, \dots, m2^{-j} - 1\}$ of pairwise disjoint boxes. That is, the reference cubes containing a single reference point $r_1, \dots, r_m$ form the 0th level $U_0 = \{u(i, i) : i = 1, 2, \dots, m\}$, and every box $u \in U_{j+1}$ at level $j+1$ is spanned by exactly two boxes of U_j . The root level, level $\log m$, consists of a single box $u(1, m)$, which we denote by $u_{\max}$. See Figure 2(a). Choose two axis-aligned k -flats X and Y such that they span the entire space $\mathbb{R}^d$. This is possible, since $k \geq d/2$. For any box u , let $x(u)$ (resp., $y(u)$) denote the maximum number of boxes of $B|u$ stabbed by a k -flat parallel to X (resp., Y) and incident to a reference point in $R \cap u$. We prove by induction on j that

$$(1) \quad x(u) + y(u) \geq j + 1 \quad \text{for every } u \in U_j, \quad j = 0, 1, \dots, \log m.$$

This immediately implies that $s_{d,k}(B) = \Omega(\log n)$, since

$$s_{d,k}(B) \geq \max(x(u_{\max}), y(u_{\max})) \geq \frac{x(u_{\max}) + y(u_{\max})}{2} \geq \frac{1 + \log m}{2} = \Omega(\log n).$$

We proceed by induction on j . Since $u(i, i)$ lies in the interior of a box of B , for every $i = 1, 2, \dots, m$, we have $x(i, i) = y(i, i) = 1$. This proves inequality (1) in the

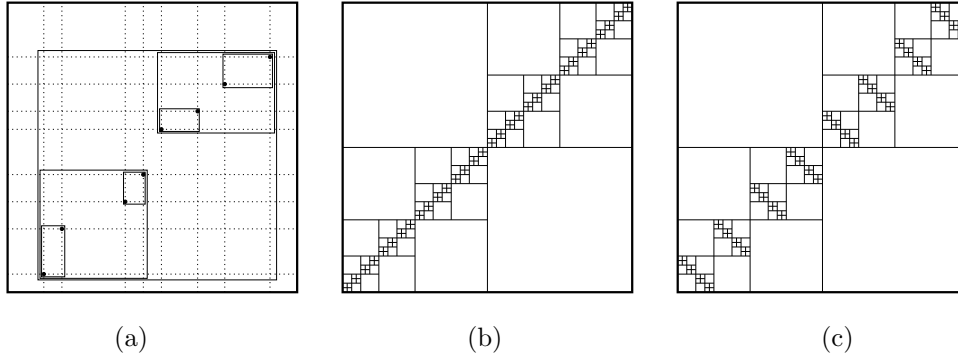


FIG. 2. (a) A sequence of eight reference points with strictly increasing coordinates (the corresponding subdivision is not shown in the figure); boxes of U_1 , U_2 , and U_3 are also shown. (b) Our upper bound construction. (c) An axis-aligned subdivision where the size of every monotone sequence of reference points is at most $\sqrt{n}$.

base case $j = 0$. For the induction step, suppose that $x(u') + y(u') \geq j' + 1$ holds for every $u' \in U_{j'}$, $0 \leq j' \leq j$. We prove for every box $u = u(i2^{j+1} + 1, i2^{j+1} + 2^{j+1}) \in U_{j+1}$ that $x(u) + y(u) \geq j + 2$. Box u contains two disjoint boxes, $u_1 = u(2i \cdot 2^j + 1, 2i \cdot 2^j + 2^j) \in U_j$ and $u_2 = u((2i + 1) \cdot 2^j + 1, (2i + 2) \cdot 2^j) \in U_j$. Let h_1 and v_1 be k -flats parallel to X and Y which attain $x(u_1)$ and $y(u_1)$, respectively. Similarly, let h_2 and v_2 be k -flats parallel to X and Y attaining $x(u_2)$ and $y(u_2)$. By induction, we have $x(u_1) + y(u_1) \geq j + 1$ and $x(u_2) + y(u_2) \geq j + 1$. If the total number of boxes of $B|u$ stabbed by h_1 and v_1 is at least $j + 2$, then $x(u) + y(u) \geq j + 2$ and the induction step is complete. Otherwise h_1 and v_1 each stab exactly $x(u_1)$ and $y(u_1)$ boxes of the subdivision $B|u$. Therefore, h_1 stabs the same boxes of the subdivisions $B|u_1$ and $B|u$. In other words, every box of $B|u$ stabbed by h_1 intersects $u_1 \subset u$. This implies that every box of $B|u$ stabbed by h_1 is disjoint from u_2 . Indeed, every box that intersects both u_1 and u_2 must contain both reference points $r_{2i \cdot 2^j + 2^j} \in u_1$ and $r_{(2i+1) \cdot 2^j + 1} \in u_2$. Since the reference points lie in distinct boxes of B , no box of B can intersect both u_1 and u_2 .

Note that h_1 and v_2 intersect in the interior of u (that is, there is a point $p \in h_1 \cap v_2 \cap \text{int}(u)$), and so the k -flat v_2 stabs at least one more box in $B|u$ than in $B|u_2$ (namely, the box containing p). Hence, v_2 stabs $y(u_2) + 1$ boxes of $B|u$, and the total number of boxes of $B|u$ stabbed by h_2 and v_2 is at least $x(u_2) + y(u_2) + 1 \geq j + 2$. We have $x(u) + y(u) \geq j + 2$, as required. $\square$

The upper bound $s_{d,k}(n) = O(\log n)$, $0 < k < d$, follows from a simple construction obtained by successively subdividing every box lying along a skew line (e.g., a diagonal).

LEMMA 6. For every $k, d \in \mathbb{N}$, $1 \leq k \leq d$, we have $s_{d,k}(n) = O(\log n)$.

Proof. Consider an axis-aligned cube C in $\mathbb{R}^d$, and let e be one of its diagonals. We construct inductively a sequence of axis-aligned subdivisions of C . We start with the trivial subdivision $B_0 = \{C\}$. If B_i is defined, we construct B_{i+1} as follows. For every box $b \in B_i$ whose interior intersects the diagonal e , choose a point $p_b \in \text{int}(b) \cap e$ and split b into 2^d boxes along the d axis-aligned hyperplanes incident to p_b (Figure 2(b)).

It is easy to see by induction that the diagonal e intersects the interior of 2^i boxes of B_i . In one induction step, each of these boxes is split into 2^d pieces, so $|B_{i+1}| - |B_i| = (2^d - 1)2^i$. This recursion gives $|B_i| = (2^d - 1)(2^i - 1) + 1$. On the other hand, every axis-aligned k -flat, $0 \leq k \leq d - 1$, stabs at most $i(2^k - 1) + 1$ boxes of B_i . For every $k, d \in \mathbb{N}$, $1 \leq k \leq d$, this gives $s_{d,k}(B_i) = O(i) = O(\log |B_i|)$. $\square$

3. Lower bound on the axis-parallel stabbing number. In this section, we prove that $s_{d,k}(n) = \Omega(\log^{1/\lfloor (d-1)/k \rfloor} n)$ for every $k, d \in \mathbb{N}$, $0 < k < d$. In particular, for $k = 1$ and every $d \in \mathbb{N}$, we have $s_{d,1}(n) = \Omega(\log^{1/(d-1)} n)$. For $k < d \leq 2k$, this is equivalent to Lemma 5. We proceed by induction on d using an argument analogous to the proof of Lemma 5.

Consider a subdivision B of $\mathbb{R}^d$ into n axis-aligned boxes. Similar to the proof of Lemma 5, choose sufficiently small axis-aligned *reference cubes* in the interior of each box of B such that all extents of all reference cubes are pairwise nonoverlapping and disjoint from any coordinates of the vertices of the boxes in B . Pick a *reference point* in the interior of the reference cube. Let $R = (r_i : i = 1, 2, \dots, \lceil n^{2^{1-d}} \rceil)$ be a coordinate-wise strictly monotone sequence of reference points, which exists by Theorem 4. Let m be the power of 2 such that $m \leq \lceil n^{2^{1-d}} \rceil < 2m$. For $1 \leq i \leq j \leq m$, let $u(i, j)$ denote the box spanned by the reference cubes containing r_i and r_j . Build the same binary hierarchy $U_0, U_1, \dots, U_{\log m}$ spanned by reference cubes as in the previous section.

For a box b and $0 < k < d$, let $x_k(b)$ denote the maximum number of boxes of $B|b$ stabbed by a k -flat parallel to the first k coordinate axes and incident to a reference point in $R \cap b$. In particular, let $x_k(i, j) = x_k(u(i, j))$ for $1 \leq i \leq j \leq m$. Let $V_k(b)$ denote the $(d - k)$ -flat orthogonal to the first k coordinate axes and incident to the first reference point in $R \cap b$ (for example, $V_{d-1}(b(i, j))$ is the hyperplane orthogonal to the x_1 -axis and incident to r_i for all j). We prove an inequality, which is analogous in some sense to inequality (1).

LEMMA 7. *If every $(n - k)$ -flat orthogonal to the first k coordinate axes stabs fewer than 2^α boxes of B for some integer α , then*

$$(2) \quad x_k(u) \geq \lfloor \ell / \alpha \rfloor + 1 \quad \text{for every } u \in U_\ell, \ell = 0, 1, \dots, \log m.$$

Before we prove Lemma 7, we show that it immediately implies the lower bound of Theorem 2.

LEMMA 8. *For every fixed $d, k \in \mathbb{N}$, $0 < k < d$, we have*

$$s_{d,k}(n) = \Omega\left(\log^{1/\lfloor (d-1)/k \rfloor} n\right).$$

Proof. For every $k \in \mathbb{N}$, we proceed by induction on d . The cases that $k < d \leq 2k$ are established in Lemma 5. Assume that $d > 2k$ and $s_{d',k} = \Omega(\log^{1/\lfloor (d'-1)/k \rfloor} n)$ for every $d' \in \mathbb{N}$, $k < d' < d$. We show that $s_{d,k}(n) = \Omega(\log^{1/\lfloor (d-1)/k \rfloor} n)$.

Consider an axis-aligned subdivision B of size n in $\mathbb{R}^d$. If an axis-aligned $(d - k)$ -flat W stabs $w \geq \exp(\log^{1-1/\lfloor (d-1)/k \rfloor} n) = \exp(\log^{\lfloor (d-k-1)/k \rfloor / \lfloor (d-1)/k \rfloor} n)$ boxes of B , then we apply the induction hypothesis for the subdivision induced by B in W : there is an axis-aligned k -flat in W that stabs at least $s_{d-k,k}(w) = \Omega(\log^{1/\lfloor (d-k-1)/k \rfloor} w) \geq \Omega(\log^{1/\lfloor (d-1)/k \rfloor} n)$ boxes of B .

If every axis-aligned $(d - k)$ -flat stabs fewer than $\exp(\log^{1-1/\lfloor (d-1)/k \rfloor} n)$ boxes of B , then apply Lemma 7 with $\alpha = \lfloor \log^{1-1/\lfloor (d-1)/k \rfloor} n \rfloor$ for the box $u(1, m)$. Hence,

$$x_k(u(1, m)) > \left\lceil \frac{\log m}{\lfloor \log^{1-1/\lfloor (d-1)/k \rfloor} n \rfloor} \right\rceil = \Omega\left(\frac{\log n}{\log^{1-1/\lfloor (d-1)/k \rfloor} n}\right) = \Omega\left(\log^{1/\lfloor (d-1)/k \rfloor} n\right).$$

Thus a k -flat parallel to the first k coordinate axes stabs at least $\Omega(\log^{1/\lfloor (d-1)/k \rfloor} n)$ boxes of B . $\square$

Proof of Lemma 7. We proceed by induction on ℓ . Since $u(i, i)$ lies in the interior of a box of B , we have $x_k(i, i) = 1$ for every $i = 1, 2, \dots, m$. This establishes inequality (2) in the base case $\ell = 0$. Note that $x_k(u)$ is an integer for every box u , so it is enough to prove (2) for $\ell = \alpha j$, $j \in \mathbb{N}$; otherwise $\lfloor (\ell - 1)/\alpha \rfloor = \lfloor \ell/\alpha \rfloor$. For the induction step, suppose that $x_k(u') \geq j' + 1$ holds for every $u' \in U_{\alpha j'}$, $0 \leq j' \leq j$. For every box $u = u(i2^{(j+1)\alpha} + 1, i2^{(j+1)\alpha} + 2^{(j+1)\alpha}) \in U_{(j+1)\alpha}$, we prove that $x_k(u) \geq j + 2$. Box u contains 2^α disjoint boxes of level $j\alpha$, namely, it contains the boxes $u_s = u(i2^{(j+1)\alpha} + s2^{\alpha j} + 1, i2^{(j+1)\alpha} + (s+1)2^{\alpha j}) \in U_{j\alpha}$ for $s = 0, 1, \dots, 2^\alpha - 1$.

For $s = 0, 1, \dots, 2^\alpha - 1$, let h_s denote a k -flat that is parallel to the first k coordinate axes and attains $x_k(u_s)$. By induction, we have $x_k(u_s) \geq j + 1$ for every s . If the total number of boxes of $B|u$ stabbed by h_s is at least $j + 2$ for any s , then we have $x_k(u) \geq j + 2$ and the induction step is complete. Otherwise h_s stabs exactly $j + 1$ boxes of the subdivision $B|u$ for every s . This means that every h_s stabs the same boxes of the subdivisions $B|u_s$ and $B|u$. In other words, every box of $B|u$ stabbed by h_s must intersect u_s . Let $b_s \in B$ be the box of B containing the intersection point $h_s \cap V(u_s)$ for $s = 0, 1, \dots, 2^\alpha - 1$. We have shown that box b_s intersects u_s .

However, every box that intersects both u_s and u_t , $0 \leq s < t \leq 2^\alpha - 1$, must contain both reference points $r_{i2^{(j+1)\alpha} + (s+1)2^{\alpha j}} \in u_s$ and $r_{i2^{(j+1)\alpha} + t2^{\alpha j} + 1} \in u_t$. Since distinct reference points lie in the interiors of distinct boxes of B , no box of B can intersect both u_s and u_t for $0 \leq s < t \leq 2^\alpha - 1$. Hence, the boxes b_s , $s = 0, 1, \dots, 2^\alpha - 1$, are distinct. We have shown that $V_k(u)$ stabs at least 2^α boxes of the subdivision $B|u$. This contradicts our assumption that every $(d - k)$ -flat orthogonal to the first k coordinate axes stabs fewer than 2^α boxes of B . We conclude that $x_k(u) \geq j + 2$. $\square$

4. Upper bound constructions. Every box in our constructions has integral vertices (all coordinates are integers). The bounding box of our constructions in $\mathbb{R}^d$ will be $[0, n]^d$, that is, the d -dimensional cube of edge length $n \in \mathbb{N}$. In this section, we construct an axis-aligned subdivision $D_d(n)$ of $[0, n]^d$ that contains at least n unit boxes and has axis-parallel stabbing number $s_{d,1}(D_d(n)) = O(\log^{1/(d-1)} n)$.

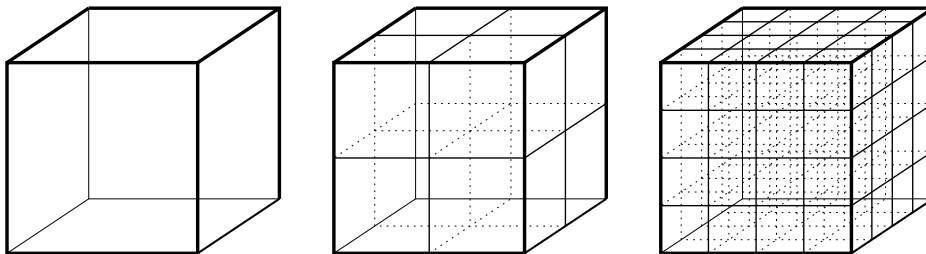


FIG. 3. The subdivision $\mathcal{C}_i$ of a cube for $\ell = 0, 1, 2$.

4.1. Preliminaries. It is convenient to describe our constructions in terms of a hierarchy of cube subdivisions. For a given $n \in \mathbb{N}$, let $\ell = \lceil \log n \rceil$ and consider the cube $E_{d,\ell} = [0, 2^\ell]^d$, which contains $[0, n]^d$. We define a hierarchy of subdivisions of $E_{d,\ell}$. Each level of the hierarchy is a subdivision $\mathcal{C}_i$, $i = 0, 1, \dots, \ell$, defined as follows. Let $\mathcal{C}_0 = \{E_{d,\ell}\}$, and for every $i = 1, 2, \dots, \ell$, we obtain $\mathcal{C}_i$ by partitioning each cube of $\mathcal{C}_{i-1}$ into 2^d congruent cubes (Figure 3). In particular, $\mathcal{C}_\ell$ is the subdivision of $E_{d,\ell}$ into unit cubes. The cubes in $\mathcal{C} = \bigcup_{i=0}^\ell \mathcal{C}_i$ are organized in a 2^d -ary tree: the parent

of a cube $u \in \mathcal{C}_i$, $1 \leq i \leq \ell$, is the unique cube $\hat{u} \in \mathcal{C}_{i-1}$ that contains u . Let e be the diagonal of $E_{d,\ell}$ incident to the origin. The cubes of $\mathcal{C}$ stabbed by e are called *diagonal cubes* (note that if e stabs a cube $u \in \mathcal{C}$, then $e \cup u$ is a diagonal of u). The 2^ℓ unit cubes stabbed by e are called *diagonal unit cubes*.

We define another subdivision of $[0, n]^d$, which we call $Q_d(n)$. If $n = 2^\ell$, $\ell \in \mathbb{N}$, then let $Q_d(2^\ell)$ be a subdivision of $E_{d,\ell}$ that contains all diagonal unit cubes of $\mathcal{C}$ and all maximal nondiagonal cubes of $\mathcal{C}$ (that is, every cube $u \in \mathcal{C}_i$, $1 \leq i \leq \ell$, such that its interior is disjoint from e , but its parent cube $\hat{u} \in \mathcal{C}_{i-1}$ is a diagonal cube). Figure 2(b) depicts $Q_2(64)$ in the plane. If n is not a power of 2, then let $Q_d(n) = \{[0, n]^d \cap u : u \in Q_d(\exp(\lceil \log n \rceil))\}$, which consist of the boxes of $Q_d(\exp(\lceil \log n \rceil))$ clipped in $[0, n]^d$.

Shadows. Our constructions of axis-aligned subdivisions rely heavily on the concept of *shadows*.

DEFINITION 9. Consider two axis-aligned boxes b and E in $\mathbb{R}^d$ such that $b \subset E$ and $b = \prod_{i=1}^d [a_i, c_i]$. For every sign vector $\mathbf{v} \in \{-, 0, +\}^d$, we define the $\mathbf{v}$ -shadow of b with respect to E as the set of points $p = (p_1, p_2, \dots, p_d) \in E$ such that (1) $a_i \leq p_i \leq c_i$ if $v_i = 0$; (2) $p_i \leq a_i$ if $v_i = -$; and (3) $c_i \leq p_i$ if $v_i = +$. (See Figure 4.)

We also introduce some terminology for certain frequently used shadows. For a vector $\mathbf{v}$, we denote by $|\mathbf{v}|$ the number of nonzero coordinates (the support of $\mathbf{v}$). For an integer s , $0 \leq s \leq d$, we say that a box is an s -shadow of b if it is a $\mathbf{v}$ -shadow of b for some $\mathbf{v}$ with $|\mathbf{v}| = s$. If $|\mathbf{v}| = 1$ and the j th coordinate of $\mathbf{v}$ is $v_j = \sigma \in \{-, +\}$, then we may also say that the $\mathbf{v}$ -shadow is the x_j^σ -shadow.

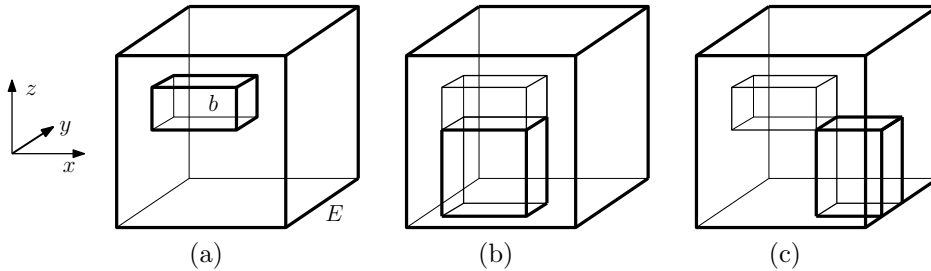


FIG. 4. (a) A box b in a cube E in $\mathbb{R}^3$. (b) The $(0, 0, -)$ -shadow of b , which can also be called the x_3^- -shadow. (c) The $(+, 0, -)$ -shadow of b .

It is clear that the $\mathbf{0}$ -shadow of the box b is itself. Since there are $2^s \binom{d}{s}$ distinct sign vectors $\mathbf{v} \in \{-, 0, +\}^d$ with $|\mathbf{v}| = s$, every box $b \subset E$ has $2^s \binom{d}{s}$ distinct s -shadows. Note, however, that some of the shadows might have empty interiors when the boundaries of E and b intersect. A key property of shadows is that the k -shadows, $k < d/2$, of any two boxes whose extents are disjoint in all d dimensions are also disjoint. The next proposition formulates this property.

PROPOSITION 10. Assume that we are given two axis-aligned boxes $b_1 \subset E$ and $b_2 \subset E$ whose extents have disjoint interiors in every dimension. We are also given two vectors $\mathbf{v}_1, \mathbf{v}_2 \in \{-, 0, +\}^d$ with $|\mathbf{v}_1| + |\mathbf{v}_2| < d$. Then the $\mathbf{v}_1$ -shadow of b_1 and the $\mathbf{v}_2$ -shadow of b_2 have disjoint interiors.

Proof. Assume, to the contrary, that a point $p \in E$ lies in the interior of both the $\mathbf{v}_1$ -shadow of b_1 and the $\mathbf{v}_2$ -shadow of b_2 . Then $d - |\mathbf{v}_1|$ coordinates of p lie in the interiors of the corresponding extents of b_1 , and $d - |\mathbf{v}_2|$ coordinates of p lie in the

interiors of the corresponding extents of b_2 . Since $|\mathbf{v}_1| + |\mathbf{v}_1| < d$, there is an index $1 \leq i \leq d$ such that the i th coordinate of p lies in the interior of the i th extent of both u_1 and u_2 . This contradicts our assumption, completing the proof. $\square$

Bricks. A $\mathbf{v}$ -shadow of a diagonal cube $u \subset E_{d,\ell}$ for a vector $\mathbf{v} \in \{-, 0, +\}^d$, $0 < |\mathbf{v}| < d$, is called a *brick* of $E_{d,\ell}$. In particular, if the j th coordinate of $\mathbf{v}$ is $v_j = \sigma \in \{-, +\}$, then the $\mathbf{v}$ -shadow of u is called an x_j^σ -*brick*. Every diagonal cube has $3^d - 2$ distinct bricks (some of which may have empty interiors). Bricks are crucial for our argument: it turns out that every box in our construction is either a diagonal unit cube or lies in some brick of $E_{d,\ell}$.

Side buffers. We define $2d$ subsets of the cube $E_{d,\ell}$, $d \geq 3$, called *side buffers*, one for each $((d-1)$ -dimensional) side of $E_{d,\ell}$. For every $j \in \{1, 2, \dots, d\}$ and $\sigma \in \{-, +\}$, let the side buffer F_j^σ be the union of the x_j^σ -shadows of all diagonal cubes of $\mathcal{C}$ of all possible sizes. (See Figure 5.) By Proposition 10, the $2d$ side buffers of $E_{d,\ell}$ have pairwise disjoint interiors if $d \geq 3$.

We introduce a shorthand notation to refer to each of the $2d$ sides of a d -dimensional box. We say that the x_j^σ -*side* of a box $\prod_{i=1}^d [a_i, c_i]$ is the side contained in the hyperplane $x_j = a_j$ if $\sigma = -$, and in the hyperplane $x_j = c_j$ if $\sigma = +$. In particular, let f_j^σ denote the x_j^σ -side of $E_{d,\ell}$. We say that the x_j^σ -*neighbor* of a cube $u \in \mathcal{C}_i$ is the cube of $\mathcal{C}_i$ adjacent to the x_j^σ -side of u .

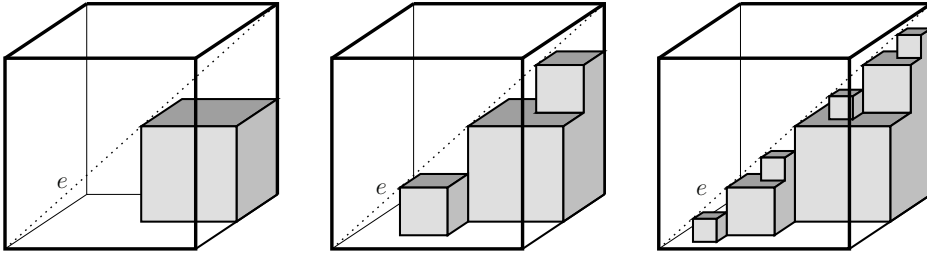


FIG. 5. The side buffer F_3^- of the cube $E_{3,\ell}$ for $\ell = 1, 2$, and 3 .

In $\mathbb{R}^3$, the diagonal unit cubes and the side buffers tile the entire cube $E_{3,\ell}$ for every $\ell \in \mathbb{N}$. For $d \geq 4$, however, they do not completely fill $E_{d,\ell}$. We show next that the remaining space in $E_{d,\ell}$, $d \geq 4$, can be subdivided into axis-aligned boxes, at most two of which are stabbed by any axis-parallel line.

LEMMA 11. Denote by $\hat{E}_{d,\ell}$ the complement of all diagonal unit cubes and the $2d$ side buffers of $E_{d,\ell}$. There is an axis-aligned subdivision $S_{d,\ell} \subset Q_d(2^\ell)$ of $\hat{E}_{d,\ell}$ whose axis-parallel stabbing number is at most 2.

Proof. A cube $u \in Q_d(2^\ell)$ can be of three possible types: it can be (i) a diagonal unit cube, (ii) contained in a side buffer, or (iii) interior disjoint from all diagonal unit cubes and side buffers. Cube u is a diagonal unit cube if and only if the diagonal e intersects the interior of u . Cube u is contained in a side buffer if and only if e intersects the interior of any 1-shadow of u (which must contain a diagonal cube in this case). Let $S_{d,\ell}$ be the set of cubes in $Q_d(2^\ell)$ of type (iii). Clearly, it is a subdivision of $\hat{E}_{d,\ell}$. It remains to show that its axis-parallel stabbing number is at most 2.

Suppose that an axis-parallel line L stabs a cube $u \in S_{d,\ell}$, where $u \in \mathcal{C}_i$ for some $i \in \{1, 2, \dots, \ell\}$. Consider its parent box $\hat{u} \in \mathcal{C}_{i-1}$, $u \subset \hat{u}$. The cube $\hat{u}$ must be a diagonal cube (by the definition of $Q_d(2^\ell)$). If we subdivide $\hat{u}$ into 2^d congruent subcubes, then one of the subcubes is u , but the diagonal e does not stab u (since u is not a diagonal box) or any of its d neighbors within $\hat{u}$ (since u is not contained

in any of the side buffers). So each of the d neighbors of u within $\hat{u}$ is contained in either a side buffer or it is another element of $S_{d,\ell}$. Hence, L stabs at most two boxes of $S_{d,\ell}$ within $\hat{u}$. Outside of $\hat{u}$, however, L cannot stab any boxes of $S_{d,\ell}$ because e stabs $\hat{u}$, and so each 1-shadow of $\hat{u}$ is contained in one of the side buffers. $\square$

4.2. The big picture. Using the terminology of the previous subsection, we can now give an overview of our upper bound construction. (A formal description follows in subsection 4.4.) In $\mathbb{R}^d$, we build a subdivision $D_d(n)$ of the cube $[0, n]^d$. We use an auxiliary subdivision $B_d(m)$ of the cube $[0, m]^d$, for $m = \exp(\log^{(d-2)/(d-1)} n)$, such that all diagonal unit cubes of $[0, m]^d$ and all their 1-shadows are elements of $B_d(m)$. The remaining space in $[0, m]^d$ is filled with some subdivision of low stabbing number (which we discuss in more detail in the next paragraph). The construction of $D_d(n)$ is done recursively in $\log^{1/(d-1)} n$ steps. In each step, we scale up a copy of the subdivision $B_d(m)$ by a factor of m and subdivide its m diagonal cubes with copies of $B_d(m)$. After $\log^{1/(d-1)} n$ recursive steps, we obtain the subdivision $D_d(n)$ of $[0, n]^d$. Note that each diagonal unit cube of $[0, n]^d$ is an element of $D_d(n)$. Every axis-parallel line L that stabs a diagonal unit cube of $[0, n]^d$ can stab at most $1 + 2 \log^{1/(d-1)} n$ boxes of $D_d(n)$, since each recursive step adds at most two new boxes along L (namely, the 1-shadows of the diagonal unit cube of a copy of $B_d(m)$). We will show in subsection 4.4 that this argument carries over to any axis-parallel line, as well, if the axis-parallel stabbing number of $B_d(m)$ is $O(\log^{1/(d-2)} n)$.

For the construction of $B_d(m)$, we use the subdivision $D_{d-1}(m)$ (which exists by induction on the dimension). Recall that the 1-shadows of the diagonal unit cubes lie in the side buffers. By Lemma 11, it is enough to subdivide each side buffer F_j^σ such that the x_j^σ -shadows of all diagonal unit cubes are elements of this subdivision for every j and σ . Note, also, that the orthogonal projections of the diagonal unit cubes of $[0, m]^d$ to the face $f_j^\sigma = [0, m]^{d-1}$ are the diagonal unit cubes of f_j^σ . We subdivide f_j^σ with the $(d-1)$ -dimensional subdivision $D_{d-1}(m)$ (which contains all diagonal unit cubes) and “lift” each $(d-1)$ -dimensional box in f_j^σ to a full-dimensional box to fill the side buffer F_j^σ . Unfortunately, the side buffers are typically nonconvex, so the lift-up of a box in f_j^σ to F_j^σ may be a nonconvex axis-aligned polyhedron. We construct a subdivision of each such polyhedron (with very small stabbing number) to obtain an axis-aligned box subdivision of each side buffer F_j^σ .

4.3. Extending a subdivision from a side to the adjacent side buffer.

Lifting a $(d-1)$ -dimensional axis-aligned subdivision to fill a side buffer of $E_{d,\ell}$ raises a technical problem, which we resolve in this subsection (Lemma 18). Imagine that D' is a $(d-1)$ -dimensional axis-aligned subdivision of side f_d^- of $E_{d,\ell}$ with $|D'| \geq n$ boxes and $s_{d-1,1}(D') = O(\log^{1/(d-2)} n)$. The extension of a $(d-1)$ -dimensional box $b' \in D'$ to the adjacent side buffer F_d^- of $E_{d,\ell}$ is $b = (b' \times [0, 2^\ell]) \cap F_d^-$. Unfortunately, b is not necessarily an axis-aligned box, since F_d^- is nonconvex whenever $\ell \geq 2$. In this subsection, we subdivide each polyhedron $b = (b' \times [0, 2^\ell]) \cap F_d^-$ into axis-aligned boxes such that the axis-parallel stabbing number of the resulting subdivision of F_d^- is only $O(\log^{1/(d-2)} n)$.

We formulate all our results in this subsection for the side $f = f_d^-$ and side buffer $F = F_d^-$ of $E_{d,\ell}$. This allows us to define the above-below relations with respect to the x_d -axis. In particular, every box lies above its x_d -shadow. By symmetry, all our results hold for any side f_j^σ and corresponding side buffer F_j^σ for $j \in \{1, 2, \dots, d\}$ and $\sigma \in \{-, +\}$. Let us fix some notation for the $(d-1)$ -dimensional boxes and subdivisions of the side f . For a d -dimensional box $b \subseteq E_{d,\ell}$, let b' denote its orthogonal

projection to f . For example, we have $f = E'_{d,\ell}$. Similarly, let e' denote the orthogonal projection of the diagonal e . Clearly, e' is the diagonal of f incident to the origin. The side f has a hierarchical subdivision $\mathcal{C}'_i$, $i = 0, 1, 2, \dots, \ell$, with $\mathcal{C}' = \bigcup_{i=0}^{\ell} \mathcal{C}'_i$. Note that the orthogonal projection of every diagonal cube of $\mathcal{C}$ is a $(d-1)$ -dimensional diagonal cube of $\mathcal{C}'$.

Properties of side buffers and bricks. We first study the intersection of a side buffer with an axis-parallel line.

PROPOSITION 12. *The intersection of F with every axis-parallel line is connected (i.e., it is either empty or a line segment).*

Proof. Consider an axis-parallel line L . We may assume that $L \cap F \neq \emptyset$; otherwise there is nothing left to prove. Note that L stabs at most one diagonal cube in $\mathcal{C}_i$ for each $i = 0, 1, \dots, \ell$. The diagonal cubes stabbed by L form a chain $u_0 \supset u_2 \supset \dots \supset u_{\kappa}$, for some $\kappa \leq \ell$.

Assume first that L is parallel to the x_d -axis. L stabs a diagonal cube if and only if L intersects its x_d^- -shadow. The intersection $L \cap F$ is, therefore, the intersection of L with the x_d^- -shadow of the smallest cube u_{κ} . The intersection of L with the x_d^- -shadow of u_{κ} is a line segment.

Assume next that L is orthogonal to the x_d -axis. If L intersects the x_d^- -shadow of a diagonal cube $u \in \mathcal{C}$, then u lies above L . The diagonal cubes lying above L form a chain $u_{\kappa} \supset u_{\kappa+1} \supset \dots \supset u_{\ell}$ for some $\kappa > 0$. This chain is nonempty, since $L \cap F \neq \emptyset$. Any diagonal cube strictly containing u_{κ} already intersects L (so its x_d^- -shadow is disjoint from L). The intersection $L \cap F$ is, therefore, the intersection of L with the x_d^- -shadow of the largest cube u_{κ} . The intersection of L with the x_d^- -shadow of u_{κ} is a line segment. $\square$

Next we study the intersection of an x_d^- -brick of $E_{d,\ell}$ and the side buffer F .

PROPOSITION 13. *Let $u_1, u_2 \in \mathcal{C}$ be two diagonal cubes of $E_{d,\ell}$. If an x_d^- -brick of u_1 and the x_d^- -shadow of u_2 are not interior disjoint, then $u_1 \subseteq u_2$ or $u_2 \subseteq u_1$. Furthermore if $u_2 \subseteq u_1$, then this x_d^- -brick of u_1 must be the x_d^- -shadow of u_1 .*

Proof. Consider two diagonal cubes, u_1 and u_2 . Assume first that they have disjoint interiors. Since e is the diagonal of both u_1 and u_2 , the cube u_2 is contained in either the $(-, -, \dots, -)$ - or the $(+, +, \dots, +)$ -shadow of u_1 . If u_2 lies in the $(-, -, \dots, -)$ -shadow of u_1 , then the x_d^- -shadow of u_2 also lies in the $(-, -, \dots, -)$ -shadow of u_1 , whose interior is disjoint from any x_d^- -brick of u_1 (by the definition of bricks). If u_2 lies in the $(+, +, \dots, +)$ -shadow of u_1 , then the x_d^- -shadow of u_2 lies in the union of the $(+, +, \dots, +, +)$ -, $(+, +, \dots, +, 0)$ -, and $(+, +, \dots, +, -)$ -shadows of u_1 , whose interiors are all disjoint from any x_d^- -brick of u_1 . This confirms that if an x_d^- -brick of u_1 and the x_d^- -shadow of u_2 overlap, then u_1 and u_2 also overlap. If two diagonal cubes of $\mathcal{C}$ overlap, then one must contain the other.

Assume now that $u_2 \subseteq u_1$. Then the x_d^- -shadow of u_2 is contained in the union of u_1 and its x_d^- -shadow. Hence, the x_d^- -shadow of u_2 may intersect the $\mathbf{0}$ - and x_d^- -shadows of u_1 , but it has disjoint interior from any other shadow of u_1 . $\square$

PROPOSITION 14. *If a $(d-1)$ -dimensional box b' lies in a brick of $f = E_{d-1,\ell}$, then $(b' \times [0, 2^{\ell}]) \cap F$ lies in a brick of $E_{d,\ell}$.*

Proof. We may assume without loss of generality that b' is an entire brick of f . Assume that b' is a $\mathbf{v}'$ -shadow of a diagonal cube $u' \in \mathcal{C}'$ of the side f with $0 < |\mathbf{v}'| < d-1$. Let $u = \prod_{i=1}^d [a_i, c_i]$ denote the diagonal cube of $E_{d,\ell}$ whose orthogonal projection is u' . Let $\mathbf{v} = (\mathbf{v}', -)$ be the d -dimensional vector, whose first $(d-1)$ -coordinates are the coordinates of $\mathbf{v}'$ and the d th coordinate is negative. We show that the $\mathbf{v}$ -shadow of u contains $(b' \times [0, 2^{\ell}]) \cap F$. Note that $b' \times [0, 2^{\ell}]$ is the union

of the $(\mathbf{v}', -)$ -, $(\mathbf{v}', 0)$ -, and $(\mathbf{v}', +)$ -shadows of u . By Proposition 13, the $(\mathbf{v}', +)$ - or $(\mathbf{v}', 0)$ -shadow of u may overlap with the x_d^- -shadow of another diagonal cube u_1 only if $u \subset u_1$. The x_d^- -shadow of every diagonal cube u_1 , with $u \subset u_1$, lies below the hyperplane $x_d = a_d$, but the $(\mathbf{v}', +)$ - and $(\mathbf{v}', 0)$ -shadows of u lie above $x_d = a_d$. We have shown that the $(\mathbf{v}', +)$ - and $(\mathbf{v}', 0)$ -shadows of u are interior disjoint from F . It follows that $(b' \times [0, 2^\ell]) \cap F$ is contained in the $\mathbf{v}$ -shadow of u . $\square$

Staircase polyhedra. A d -dimensional *staircase polyhedron* of size m is a polyhedron in $\mathbb{R}^d$, which is the union of m disjoint boxes $b_1, b_2, \dots, b_m$ such that their d th extents are consecutive intervals along the x_d -axis, their first $d-1$ extents form $(d-1)$ -dimensional boxes $b'_1, b'_2, \dots, b'_m$ such that $b'_1 \supset b'_2 \supset \dots \supset b'_m$, and they all share at least one common vertex. For example, a staircase polyhedron of size m in the plane is a staircase polygon with $m-1$ reflex vertices (Figure 1(d)). See Figure 6(a) for an example in 3-space. If the $(d-1)$ -dimensional boxes $b'_1, b'_2, \dots, b'_m$ share a κ -dimensional face g , $0 < \kappa < d-1$, then we can “factor” the staircase polyhedra P with g , that is, P is the cross-product of the κ -dimensional box g and a $(d-\kappa)$ -dimensional staircase polyhedron of size m .

PROPOSITION 15. *If b is an x_d^- -brick of $E_{d,\ell}$, then $b \cap F$ is a staircase polyhedron of size at most ℓ .*

Proof. Let b be the $\mathbf{v}$ -shadow of a diagonal cube u of $E_{d,\ell}$ for some $\mathbf{v} \in \{-, 0, +\}^d$ such that $0 < |\mathbf{v}| < d$ and $v_d = -$. If $|\mathbf{v}| = 1$, then b is the x_d^- -shadow of u , and hence $b \subseteq F$.

Assume now that $|\mathbf{v}| > 1$. By Proposition 13, b can overlap with the x_d^- -shadow of a diagonal cube $\hat{u}$ only if $u \subset \hat{u}$. The diagonal cubes containing u form a chain $u_0 \supset u_1 \supset \dots \supset u_\kappa = u$ of some length $\kappa \leq \ell$. The intersection $b \cap F$ is the union of the intersections of the x_d^- -shadows of the cubes $u_0, u_1, \dots, u_\kappa$ with b . So $b \cap F$ can be represented as a union of at most κ interior-disjoint boxes $b_1, b_2, \dots, b_\kappa$, where b_i is the intersection of b with the difference set of the x_d^- -shadows of u_i and u_{i-1} for $i = 1, 2, \dots, \kappa$. Here, the d th extents of the boxes $b_1, b_2, \dots, b_\kappa$ form consecutive intervals along the x_d -axis, and their projections $b'_1, b'_2, \dots, b'_\kappa$ form a chain $b'_1 \supset b'_2 \supset \dots \supset b'_\kappa$. Since b has at least one common edge with the x_d^- -shadow of u , the $(d-1)$ -dimensional boxes $b' \cap b'_i$, $i = 1, 2, \dots, \kappa$, have at least one common vertex. Therefore, $b \cap F$ is a staircase polyhedron of size κ , $\kappa \leq \ell$. $\square$

PROPOSITION 16. *A staircase polyhedron P of size m in $\mathbb{R}^d$ can be subdivided into axis-aligned boxes such that the axis-parallel stabbing number of the subdivision is at most $1 + \log m$.*

Proof. We proceed by induction on d and m . Our claim is obvious in the boundary cases that $(d = 1 \text{ and } m \in \mathbb{N})$ or $(m = 1 \text{ and } d \in \mathbb{N})$, since in these cases P is already a d -dimensional box, and the trivial subdivision $\{P\}$ has stabbing number 1. We prove the claim for $d \in \mathbb{N}$ and $m \in \mathbb{N}$, assuming that it holds for every $d' < d$ and $m' \leq m$ and that it also holds for d and every $m' < m$. Let P be a d -dimensional staircase polyhedron defined as the union of interior-disjoint boxes $b_1, b_2, \dots, b_m$.

If the boxes $b'_1, b'_2, \dots, b'_m$ share a κ -dimensional face g for $0 < \kappa < d-1$, then (as we have mentioned earlier) P is the cross-product of g and a $(d-\kappa)$ -dimensional staircase polyhedron P'' of size m . By induction, P'' has a subdivision S'' of axis-parallel stabbing number $1 + \log m$. We can construct the subdivision of P as the cross-product $S = S'' \times \{g\}$, which has the same axis-aligned stabbing number as S'' . Moreover, any line orthogonal to g stabs at most one box of S .

Suppose now that $b'_1, b'_2, \dots, b'_m$ share exactly one vertex, which we may assume to be the origin. We may assume further that P lies in the positive “quadrant” of

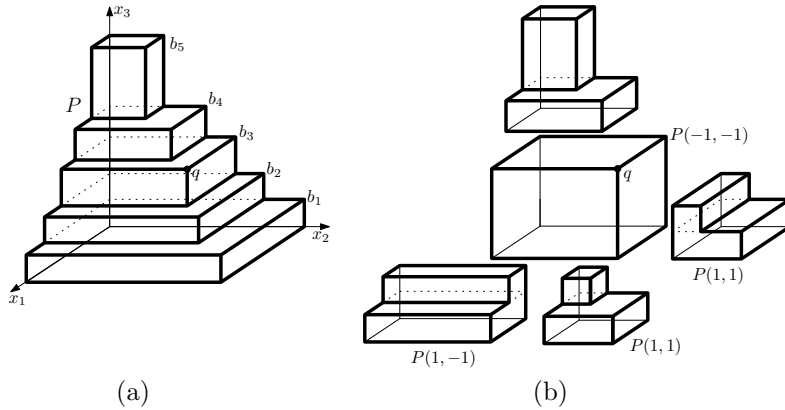


FIG. 6. (a) A staircase polyhedron P of size 5 in $\mathbb{R}^3$. (b) Its dissection along d axis-aligned hyperplanes incident to point q into $2^{d-1} + 1 = 5$ pieces.

$\mathbb{R}^d$. Refer to Figure 6. Let q be the vertex antipodal to the closest vertex to the origin in the box $b_{\lceil m/2 \rceil}$. Partition P along the d axis-aligned hyperplanes incident to q and subdivide the resulting $2^{d-1} + 1$ polyhedra by induction. Let us bound the axis-parallel stabbing number of the resulting subdivision S . The d hyperplanes incident to q partition P into *pieces*, and every axis-parallel line stabs at most two pieces. The only piece lying above the hyperplane $x_d = q_d$ is a staircase polyhedra of size $\lfloor m/2 \rfloor$. There are 2^{d-1} pieces below $x_d = q_d$, each lying in a $(\mathbf{v}', -)$ -shadow of the point q for some $\mathbf{v}' \in \{-, +\}^{d-1}$. Denote by $P(\mathbf{v}')$ the piece lying in the $(\mathbf{v}', -)$ -shadow of q . If κ entries of the vector $\mathbf{v}'$ are negative (and the remaining $d - 1 - \kappa$ entries are positive), then $P(\mathbf{v}')$ is a cross-product of a κ -dimensional box (corresponding to the coordinates with negative entries) and a $(d - \kappa)$ -dimensional staircase polyhedra of size $\lfloor m/2 \rfloor$ (corresponding to the positive entries and x_d). By induction, each piece has an axis-aligned subdivision with axis-parallel stabbing number at most $\log m$. Furthermore, any line parallel to a coordinate axis that corresponds to a negative entry stabs at most one box. Altogether, any axis-parallel line stabs at most $1 + \log m$ boxes of the subdivision S . $\square$

LEMMA 17. Let b be a box contained in a brick of $E_{d,\ell}$. Then the polyhedron $b \cap F$ has a subdivision S into axis-aligned boxes such that

- every axis-parallel line stabs at most $O(\log \ell)$ boxes of S ;
- if an axis-parallel line L is orthogonal to the x_d -axis and the segment $L \cap b$ strictly contains segment $(L \cap b) \cap F$, then L stabs at most one box of S .

Proof. By Proposition 15, we may assume that $P = b \cap F$ is a d -dimensional staircase polyhedron of size at most ℓ . If we subdivide $b \cap F$ as described in Proposition 16, then the first assertion of Lemma 17 would immediately hold. In order to guarantee the second assertion, we first partition $b \cap F$ into several pieces along at most $d - 1$ hyperplanes orthogonal to the x_d -axis, and then we apply Proposition 16 for each piece independently.

Let $b = \prod_{i=1}^d [a_i, c_i]$. Let $b_1, b_2, \dots, b_m$, $m \leq \ell$, denote the boxes that define the staircase polyhedron $P = b \cap F$. We may assume $m \geq 2$; otherwise S consists of the single box b_1 . Recall that $b'_1 \supset b'_2 \supset \dots \supset b'_m$, where b'_j denotes the $(d-1)$ -dimensional box spanned by the first $d-1$ extents of b_j . Hence, the i th extents of $b'_1, b'_2, \dots, b'_m$ also form a nested sequence for $i = 1, 2, \dots, d-1$. Let H_i be the hyperplane orthogonal to the x_d -axis above all boxes b_j whose i th extents are $[a_i, c_i]$ and below all boxes $b_{j'}$ whose i th extents are strictly contained in $[a_i, c_i]$. Refer to Figure 7.

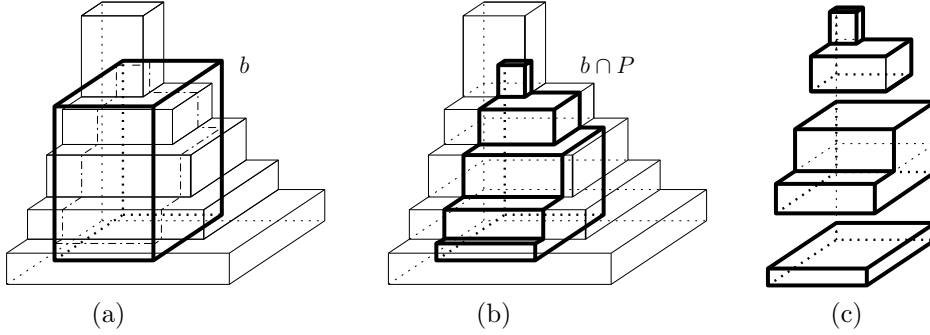


FIG. 7. (a) A box b and a staircase polyhedron P in $\mathbb{R}^3$. (b) The intersection $b \cap P$ and P . (c) The subdivision of $b \cap P$ into $d = 3$ pieces.

Partition P along all the hyperplanes H_i , $i = 1, 2, \dots, d-1$, into d pieces (some of which may be empty). Order the resulting pieces of P along the x_d -axis. The i th piece of P lies above $i-1$ hyperplanes. It is the cross-product of an i -dimensional staircase polyhedron P_i and a $(d-i)$ -dimensional box g_{d-i} , which is the cross-product of $d-i$ extents of b . Construct the subdivision of each piece, independently, as the cross-product of $\{g_{d-i}\}$ and the subdivision of P_i with stabbing number at most $1 + \log m \leq 1 + \log \ell$, as described in Proposition 16. Let S be the resulting subdivision of P .

Every line parallel to the x_d -axis stabs at most d pieces of P , and it stabs $O(\log \ell)$ boxes in the subdivision of each piece. If a line L is parallel to the x_i -axis, $1 \leq i \leq d-1$, and the segment $L \cap b$ strictly contains the segment $(L \cap b) \cap F$, then L must lie below the hyperplane H_i , and so L stabs at most one box of the subdivision S . $\square$

LEMMA 18. Let D' be an axis-aligned subdivision of side f of $E_{d,\ell}$ such that

- (a) D' contains every diagonal unit cube of C' , and
- (b) every box in D' either is a diagonal unit cube or lies in a brick of f .

Then there is an axis-aligned subdivision D of the side buffer F such that

- (a) D contains the x_d^- -shadow of every diagonal unit cube of $E_{d,\ell}$,
- (b) every box in D lies in a brick of $E_{d,\ell}$, and
- (c) the axis-aligned stabbing number of D is $s_{d,1}(D) = s_{d-1,1}(D') + O(\log \ell)$.

Proof. For every $(d-1)$ -dimensional diagonal unit cube $u' \in D'$, the lifting $(u' \times [0, 2^\ell]) \cap F$ is the x_d^- -shadow of a diagonal unit cube of C . Let the box $(u' \times [0, 2^\ell]) \cap F$ be an element of D for every diagonal unit cube $u' \in D'$. Next consider a box $b' \in D'$ which is a brick of $E_{d-1,\ell}$. By Proposition 14, the polyhedron $b = (b' \times [0, 2^\ell]) \cap F$ is contained in a brick of $E_{d,\ell}$. Subdivide b according to Lemma 17, and let all resulting boxes be elements of D . Then D is an axis-aligned subdivision of the side buffer F .

Every line parallel to the x_d -axis stabs at most one polyhedron $(b' \times [0, 2^\ell]) \cap F$, $b' \in D'$. By Lemma 17, it stabs at most $O(\log \ell)$ boxes of the subdivision D . Every line L orthogonal to the x_d -axis stabs at most $s_{d-1,1}(D')$ polyhedra $(b' \times [0, 2^\ell]) \cap F$. By Proposition 12, the intersection $L \cap F$ is a line segment, and so L intersects the boundary of F in at most two points, say p_1 and p_2 . Let b_1 and b_2 denote the polyhedra adjacent to p_1 and p_2 , respectively. By Lemma 17, b_1 and b_2 are subdivided so that $L \cap b_1$ and $L \cap b_2$ each stab $O(\log \ell)$ boxes. Any other polyhedron $(b' \times [0, 2^\ell]) \cap F$, $b' \in D'$, is subdivided so that L stabs only one resulting box. Altogether, L stabs at most $s_{d-1,1}(D') + 2O(\log \ell) = s_{d-1,1}(D') + O(\log \ell)$ boxes of D . $\square$

4.4. Upper bound construction. For every $d, n \in \mathbb{N}$, $d \geq 2$, we present an axis-aligned subdivision $D_d(n)$ that contains at least n boxes and has axis-parallel stabbing number $s_{d,1}(D_d(n)) = O(\log^{1/(d-1)} n)$. In the proof of Lemma 6 we have already described such a subdivision for $d = 2$. Here, we extend this construction for every $d \geq 3$. Along the way, we also define a family of auxiliary subdivisions $B_d(n)$ for every $d, n \in \mathbb{N}$, $d \geq 2$. We summarize the properties of the two families of subdivisions as follows.

LEMMA 19.

• For every $d \geq 3$ there is a constant $c_d > 0$ such that for every $n \in \mathbb{N}$ there is an axis-aligned subdivision $B_d(n)$ of the cube $[0, n]^d$ with the following properties:

- (i) $B_d(n)$ contains all n diagonal unit cube of $[0, n]^d$;
- (ii) every 1-shadow of every diagonal unit cube is an element of $B_d(n)$;
- (iii) every box of $B_d(n)$ either is a diagonal unit cube or lies in a brick of $[0, n]^d$;
- (iv) $s_{d,1}(B_d(n)) \leq c_d \log^{1/(d-2)} n$.

• For every $d \geq 2$ there is a constant $C_d > 0$ such that for every $n \in \mathbb{N}$ there is an axis-aligned subdivision $D_d(n)$ of the cube $[0, n]^d$ with the following properties:

- (a) $D_d(n)$ contains all n diagonal cubes of $[0, n]^d$;
- (b) every box of $D_d(n)$ is either a diagonal unit cube or lies in a brick of $[0, n]^d$;
- (c) $s_{d,1}(D_d(n)) \leq C_d \log^{1/(d-1)} n$.

Proof. We proceed by induction on d and n .

Base cases. For every $n \in \mathbb{N}$ and $d = 2$, Lemma 6 provides a construction $D_2(n)$. For every fixed $d \geq 3$, we proceed by induction on n . For an integer $d \geq 3$, let both $D_d(1)$ and $B_d(1)$ be the trivial subdivision $\{[0, 1]^d\}$. When we construct $B_d(n)$, we assume that we have already constructed the subdivision $D_{d-1}(n')$ for every $n' \in \mathbb{N}$. When we construct $D_d(n)$, we assume that we have already constructed the subdivision $B_d(n')$ for every $n' \in \mathbb{N}$.

Construction of $B_d(n)$. Let us assume first that $n = 2^\ell$ for some $\ell \in \mathbb{N}$. We construct the subdivision $B_d(n)$ of the cube $E_{d,\ell} = [0, 2^\ell]^d$. Let $B_d(n)$ contain all 2^ℓ diagonal unit cubes of $E_{d,\ell}$. Let $B_d(n)$ contain all cubes of the subdivision $S_{d,\ell}$ (which tile the complement of the diagonal unit cubes and all $2d$ side buffers, as described in Lemma 11). It remains to describe a subdivision of the $2d$ side buffers of $E_{d,\ell}$.

By symmetry, it suffices to describe the subdivision of a single side buffer $F = F_d^-$. Consider the $(d-1)$ -dimensional axis-aligned subdivision $D_{d-1}(2^\ell)$ of side $f = f_d^-$. By Lemma 18, it can be extended into an axis-aligned subdivision of F such that it contains all x_d^- -shadows of diagonal unit cubes of $E_{d,\ell}$, every box lies in a brick of $E_{d,\ell}$, and its axis-parallel stabbing number is $s_{d-1,1}(D_{d-1}(2^\ell)) + O(\log \ell) = O(\ell^{1/(d-2)} + \log \ell) = O(\log^{1/(d-2)} n + \log \log n) = O(\log^{1/(d-2)} n)$.

$B_d(2^\ell)$ clearly satisfies conditions (i), (ii), and (iii). For (iv), consider an axis-parallel line L . Line L stabs at most one diagonal cube, at most two boxes of $S_{d,\ell}$ by Lemma 11, and at most two side buffers. By the above argument, L stabs at most $O(\log^{1/(d-2)} n)$ boxes in the subdivision of each side buffer. This completes the construction of $B_d(2^\ell)$ for every $\ell \in \mathbb{N}$.

Finally, if n is not a power of 2, then take a subdivision $B_d(2^\ell)$ with $2^{\ell-1} < n < 2^\ell$, and clip it in the box $[0, n]^d \subset E_{d,\ell}$. Let $B_d(n)$ be the resulting subdivision. It contains all n diagonal unit cubes of $[0, n]^d$ and all their 1-shadows, and the stabbing number of $B_d(n)$ is at most that of $B_d(2^\ell)$, which is $O(\log^{1/(d-2)} n)$.

Construction of $D_d(n)$. First, let us assume that $n = \exp(\ell^{d-1})$ for some $\ell \in \mathbb{N}$ (and so $\log^{1/(d-1)} n$ is an integer). We obtain $D_d(n)$ in $\log^{1/(d-1)} n$ steps. Initially, let $\Delta(0) = \{[0, n]^d\}$ be the trivial subdivision. In step $i = 1, 2, \dots, \log^{1/(d-1)} n$, construct

$\Delta(i)$ by replacing each diagonal cube of $\Delta(i-1)$ with a scaled copy of the subdivision $B_d(\exp(\log^{(d-2)/(d-1)} n))$. Each step increases the number of diagonal cubes by a factor of $\exp(\log^{(d-2)/(d-1)} n)$. Let $D_d(n) = \Delta(\log^{1/(d-1)} n)$. It has

$$\begin{aligned} \left(\exp \left(\log^{(d-2)/(d-1)} n \right) \right)^{\log^{1/(d-1)} n} &= \exp \left(\log^{(d-2)/(d-1)} n \cdot \log^{1/(d-1)} n \right) \\ &= \exp(\log n) = n \end{aligned}$$

diagonal (unit) cubes.

Consider an axis-parallel line L . Assume that i , $1 \leq i \leq \log^{1/(d-1)} n$, is the largest integer such that L stabs some diagonal cube u of the subdivision $\Delta(i)$. Outside of u , the line L can stab at most $2(\log^{1/(d-1)} n - i)$ boxes, since each 1-shadow of every diagonal cube is an element of $B_d(\exp(\log^{(d-2)/(d-1)} n))$. Inside the cube u , the line L intersects $O(\log^{1/(d-2)} \exp(\log^{(d-2)/(d-1)} n)) = O(\log^{1/(d-1)} n)$ boxes by property (iv) of the subdivision $B_d(\log^{(d-2)/(d-1)} n)$. By construction and by property (iii), each box of $D_d(n)$ either is a cube of $Q_d(n)$ or lies in a brick of $[0, n]^d$.

Finally, if n is not of the form $\exp(\ell^{d-1})$, $\ell \in \mathbb{N}$, then take the subdivision $D_d(\exp(\ell^{d-1}))$ for $\exp((\ell-1)^{d-1}) < n < \exp(\ell^{d-1})$, and clip it in the box $[0, n]^d \subset E_{d, \ell^{d-1}}$. The resulting subdivision clearly contains n diagonal cubes, and its stabbing number is bounded by that of $D_d(\exp(\ell^{d-1}))$, which is $O(\log^{1/(d-1)}(\exp(\ell^{d-1}))) = O(\ell) = O(\log^{1/(d-1)} n)$. $\square$

REFERENCES

- [1] P. K. AGARWAL, *Ray shooting and other applications of spanning trees with low stabbing number*, SIAM J. Comput., 21 (1992), pp. 540–570.
- [2] P. K. AGARWAL, B. ARONOV, AND S. SURI, *Stabbing triangulations by lines in 3D*, in Proceedings of the 11th Symposium on Computational Geometry (Vancouver, BC), ACM Press, New York, 1995, pp. 267–276.
- [3] P. K. AGARWAL, M. DE BERG, J. GUDMUNDSSON, M. HAMMAR, AND H. J. HAVERKORT, *Box-trees and R-trees with near-optimal query time*, Discrete Comput. Geom., 28 (2002), pp. 291–312.
- [4] N. AMENTA AND G. M. ZIEGLER, *Shadows and slices of polytopes*, in Advances in Discrete and Computational Geometry, Contemp. Math. 223, AMS, Providence, RI, 1999, pp. 57–90.
- [5] B. ARONOV, H. BRÖNNIMANN, A. Y. CHANG, AND Y.-J. CHIANG, *Cost-driven octree construction schemes: An experimental study*, Comput. Geom. Theory Appl., 31 (2005), pp. 127–148.
- [6] B. ARONOV AND S. FORTUNE, *Approximating minimum-weight triangulations in three dimensions*, Discrete Comput. Geom., 21 (1999), pp. 527–549.
- [7] M. DE BERG, J. GUDMUNDSSON, M. HAMMAR, AND M. H. OVERMARS, *On R-trees with low query complexity*, Comput. Geom. Theory Appl., 24 (2003), pp. 179–195.
- [8] M. DE BERG AND M. VAN KREVELD, *Rectilinear decompositions with low stabbing number*, Inform. Process. Lett., 52 (1994), pp. 215–221.
- [9] P. BOSE AND L. DEVROYE, *On the stabbing number of a random Delaunay triangulation*, Comput. Geom. Theory. Appl., 36 (2007), pp. 89–105.
- [10] P. BOSE, F. HURTADO, H. MEIJER, S. RAMASWAMI, D. RAPPAPORT, V. SACRISTAN, T. C. SHERMER, AND G. T. TOUSSAINT, *Finding specified sections of arrangements: 2D results*, J. Math. Model. Algorithms, 1 (2002), pp. 3–16.
- [11] B. CHAZELLE, H. EDELSBRUNNER, AND L. J. GUIBAS, *The complexity of cutting complexes*, Discrete Comput. Geom., 4 (1989), pp. 139–181.
- [12] B. CHAZELLE AND E. WELZL, *Quasi-optimal range searching in space of finite VC-dimension*, Discrete Comput. Geom., 4 (1989), pp. 467–489.
- [13] T. K. DEY AND J. PACH, *Extremal problems for geometric hypergraphs*, Discrete Comput. Geom., 19 (1998), pp. 473–484.
- [14] R. DILWORTH, *A decomposition theorem for partially ordered sets*, Ann. of Math. (2), 51 (1950), pp. 161–166.

- [15] P. ERDŐS AND G. SZEKERES, *A combinatorial theorem in geometry*, Compos. Math., 2 (1935), pp. 463–470.
- [16] S. P. FEKETE, M. E. LÜBBECKE, AND H. MEIJER, *Minimizing the stabbing number of matchings, trees, and triangulations*, in Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms (New Orleans, LA), ACM Press, New York, SIAM, Philadelphia, 2004, pp. 437–446.
- [17] D. GAUR, T. IBARAKI, AND R. KRISHNAMURTI, *Constant ratio approximation algorithms for the rectangle stabbing problem and the rectilinear partitioning problem*, J. Algorithms, 43 (2002), pp. 138–152.
- [18] M. HELD, J. T. KLOSOWSKI, AND J. S. B. MITCHELL, *Evaluation of collision detection methods for virtual reality fly-throughs*, in Proceedings of the 7th Canadian Conference on Computational Geometry, Québec City, QC, 1995, pp. 205–210.
- [19] J. HERSHBERGER AND S. SURI, *A pedestrian approach to ray shooting: Shoot a ray, take a walk*, J. Algorithms, 18 (1995), pp. 403–431.
- [20] J. B. KRUSKAL, *Monotonic subsequences*, Proc. Amer. Math. Soc., 4 (1953), pp. 264–274.
- [21] J. MATOUŠEK, *Efficient partition trees*, Discrete Comput. Geom., 8 (1992), pp. 315–334.
- [22] J. PACH, *Notes on geometric graph theory*, in Discrete and Computational Geometry, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. 6, AMS, Providence, RI, 1991, pp. 273–285.
- [23] J. R. SHEWCHUK, *Stabbing Delaunay tetrahedralizations*, Discrete Comput. Geom., 32 (2004), pp. 339–343.
- [24] CS. D. TÓTH, *Convex subdivisions with low stabbing numbers*, submitted.
- [25] E. WELZL, *On spanning trees with low crossing numbers*, in Data Structures and Efficient Algorithms, Lecture Notes in Comput. Sci. 594, Springer-Verlag, Berlin, 1992, pp. 233–249.

THE VECTOR SPACE OVER $GF(2)$ GENERATED BY THE COMMON BASES OF TWO BINARY MATROIDS*

K. P. WEBB[†]

Abstract. A binary matroid pair consists of two binary matroids of the same rank defined on a common ground set. By excluding a particular binary matroid pair, we obtain a characterization of the vector space over $GF(2)$ generated by the incidence vectors for the common bases of a binary matroid pair.

Key words. binary matroid, matroid intersection

AMS subject classification. 05B35

DOI. 10.1137/050641417

1. Introduction. Let M and N be matroids with the same rank, defined on a common ground set S . We call the pair $P = (M, N)$ a *matroid pair*, and the common bases of M and N are called the *bases* of P . When M and N are both binary, P is said to be a *binary matroid pair*. The sets of bases of M, N , and P are denoted $\mathcal{B}(M)$, $\mathcal{B}(N)$, and $\mathcal{B}(P)$, respectively.

Given $S' \subseteq S$, the incidence vector $x \in \{0, 1\}^S$ for S' is defined by $x_e = 1$ if and only if $e \in S'$. We identify subsets of the ground set with their incidence vectors; for example, the convex hull generated by the incidence vectors for the elements of $\mathcal{B}(P)$ is referred to as the convex hull generated by $\mathcal{B}(P)$. Edmonds [3] proved that the convex hull generated by $\mathcal{B}(P)$ is the intersection of the convex hulls generated by $\mathcal{B}(M)$ and $\mathcal{B}(N)$. In this paper, we consider the incidence vectors to be defined over $GF(2) = \mathbb{Z}/2\mathbb{Z}$. We are interested in the vector space over $GF(2)$ generated by $\mathcal{B}(P)$, i.e., by the incidence vectors for the elements of $\mathcal{B}(P)$. This vector space is denoted $\text{Bin}(P)$. Similarly, $\text{Bin}(M)$ and $\text{Bin}(N)$ denote the vector spaces over $GF(2)$ generated by $\mathcal{B}(M)$ and $\mathcal{B}(N)$, respectively.

Edmonds' surprising result for the convex hull does not hold for the vector space over $GF(2)$: $\text{Bin}(P) \neq \text{Bin}(M) \cap \text{Bin}(N)$ in general. However, when M and N are binary matroids, by excluding a particular matroid pair we obtain a characterization of $\text{Bin}(P)$ in terms of $\text{Bin}(M)$ and $\text{Bin}(N)$. Some terminology is needed before stating our main result.

The rank of $S' \subseteq S$ in M is denoted $\text{rank}_M(S')$, and the matroid pair $P = (M, N)$ is *connected* if $\text{rank}_M(S') + \text{rank}_N(S \setminus S') > \text{rank}_M(S)$ whenever $\emptyset \neq S' \subsetneq S$. Recall that Edmonds' matroid intersection theorem [3] implies that if P is connected, then $\mathcal{B}(P) \neq \emptyset$. When $X, Y \subseteq S$ are such that there exists $B \in \mathcal{B}(P)$ with $X \subseteq B$ and $Y \subseteq S \setminus B$, then the matroid pair $(M/X \setminus Y, N/X \setminus Y)$ is called a *minor* of P and is denoted $P/X \setminus Y$.

Let $\mathcal{B}_M^0 \subseteq \text{Bin}(M)$ be the set of $x \in \text{Bin}(M)$ such that x is the sum of an *even* number of bases of M . As usual, we mean the sum of an even number of incidence vectors of bases of M . Similarly, let $\mathcal{B}_M^1$ contain all $x \in \text{Bin}(M)$ such that x is the

*Received by the editors September 28, 2005; accepted for publication (in revised form) March 11, 2008; published electronically July 16, 2008.

<http://www.siam.org/journals/sidma/22-3/64141.html>

[†]Department of Mathematics and Computer Science, University of Lethbridge, Lethbridge, Alberta T1K 3M4, Canada (kerri.webb@uleth.ca).

sum of an *odd* number of bases. For $x \in GF(2)^S$ and $S' \subseteq S$, let $x(S') = \sum_{s \in S'} x_s$. If $S_1, S_2, \dots, S_k$ are the connected components of M , then by a result of Rieder [6]

$$\mathcal{B}_M^0 = \{x \in GF(2)^S : x(S_i) = 0 \text{ for all } i \in \{1, \dots, k\}\},$$

$$\mathcal{B}_M^1 = \{x \in GF(2)^S : x(S_i) = \text{rank}_M(S_i) \pmod{2} \text{ for all } i \in \{1, \dots, k\}\}.$$

Since every element of $\text{Bin}(P)$ is the sum of either an even number or an odd number of bases of P , we have

$$(1.1) \quad \text{Bin}(P) \subseteq (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1).$$

Expanding the right-hand side of (1.1) gives

$$(1.2) \quad (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1) \subseteq \text{Bin}(M) \cap \text{Bin}(N).$$

We show in section 2 that $(\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$ can be properly contained in $\text{Bin}(M) \cap \text{Bin}(N)$ for simple reasons regarding the parity of the rank of components of M and N . Furthermore, if P is a *twined K_4 pair*, as defined in section 4, then $\text{Bin}(P)$ is properly contained in $(\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$. Our main theorem is the following.

THEOREM 1.1. *If M and N are binary matroids and $P = (M, N)$ is connected and does not have a twined K_4 pair as a minor, then*

$$\text{Bin}(P) = (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1).$$

The converse of Theorem 1.1 does not hold; a counterexample is given at the end of section 6. We use the matroid terminology and notation of Oxley [5] and assume that the reader is familiar with matroids and matroid representation.

2. Even or odd rank. By considering $(\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$, we show in this section that $\text{Bin}(P)$ can differ from $\text{Bin}(M) \cap \text{Bin}(N)$ due to rank parity. We first show that parity of the rank of the components of M and N affects the dimension of $(\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$.

PROPOSITION 2.1. *Let $P = (M, N)$ be a connected matroid pair on the ground set S . If c_M and c_N are the number of components of M and N , respectively, then $\mathcal{B}_M^0 \cap \mathcal{B}_N^0$ has dimension $|S| - c_M - c_N + 1$. Furthermore, if every component of M and N has even rank, then $(\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$ has dimension $|S| - c_M - c_N + 1$; otherwise, the dimension is $|S| - c_M - c_N + 2$.*

Proof. Let A_M and A_N be matrices over $GF(2)$ with columns indexed by S and rows given by the incidence vectors for the components of M and N , respectively. Let b_M and b_N be the column matrices such that $\mathcal{B}_M^1 = \{x \in GF(2)^S : A_M x = b_M\}$ and $\mathcal{B}_N^1 = \{x \in GF(2)^S : A_N x = b_N\}$. The existence of b_M and b_N follows from Rieder [6].

Note that $\mathcal{B}_M^0 \cap \mathcal{B}_N^0 = \{x \in GF(2)^S : A_M x = 0 \text{ and } A_N x = 0\}$. The rows of A_M and A_N together sum to zero, and the connectivity of P implies that no proper subset of these rows sums to zero. Since A_M and A_N together have $c_M + c_N$ rows, it follows that $\mathcal{B}_M^0 \cap \mathcal{B}_N^0$ has dimension $|S| - c_M - c_N + 1$. If each component of M and N has even rank, then $\mathcal{B}_M^1 = \mathcal{B}_M^0$ and $\mathcal{B}_N^1 = \mathcal{B}_N^0$ and thus $(\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$ has the same dimension as $\mathcal{B}_M^0 \cap \mathcal{B}_N^0$. If a component of M or N has odd rank, then for all $x \in \mathcal{B}(P)$ we have $x \in \mathcal{B}_M^1 \cap \mathcal{B}_N^1$ but $x \notin \mathcal{B}_M^0 \cap \mathcal{B}_N^0$. However, the sum of any two elements in $\mathcal{B}_M^1 \cap \mathcal{B}_N^1$ belongs to $\mathcal{B}_M^0 \cap \mathcal{B}_N^0$, and thus the dimension of $(\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$ in this case is one greater than the dimension of $\mathcal{B}_M^0 \cap \mathcal{B}_N^0$. $\square$

Next we show that the containment given in (1.2) can be strict.

PROPOSITION 2.2. *Let $P = (M, N)$ be a connected matroid pair on the ground set S , where M and N have even rank. If M and N both have a component of odd rank, then $(\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$ is properly contained in $\text{Bin}(M) \cap \text{Bin}(N)$.*

Proof. The definition of $\mathcal{B}_M^0, \mathcal{B}_M^1, \mathcal{B}_N^0$, and $\mathcal{B}_N^1$ implies that

$$(2.1) \quad \text{Bin}(M) \cap \text{Bin}(N) = (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^0 \cap \mathcal{B}_N^1).$$

We claim that there exists $x \in \mathcal{B}_M^1 \cap \mathcal{B}_N^0$ with $x \notin (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$. By (1.2), this implies the proposition.

As in the proof of Proposition 2.1, let $\mathcal{B}_M^1 = \{x \in GF(2)^S : A_M x = b_M\}$ and let $\mathcal{B}_N^1 = \{x \in GF(2)^S : A_N x = b_N\}$. Suppose y_M and y_N are row vectors over $GF(2)$ with $y_M A_M + y_N A_N = \mathbf{0}$. By the connectivity of P , the only nonzero solution to $y_M A_M + y_N A_N = \mathbf{0}$ is $y_M = y_N = \mathbf{1}$. Since M has even rank, $y_M = \mathbf{1}$ implies $y_M b_M = 0$. Therefore, there are no y_M, y_N satisfying both $y_M A_M + y_N A_N = \mathbf{0}$ and $y_M b_M = 1$. By Farkas' lemma, there exists $x \in GF(2)^S$ satisfying both $A_M x = b_M$ and $A_N x = 0$, and thus $\mathcal{B}_M^1 \cap \mathcal{B}_N^0 \neq \emptyset$, as claimed.

Since M and N both have a component of odd rank, $x \in \mathcal{B}_M^1 \cap \mathcal{B}_N^0$ implies $x \notin \mathcal{B}_M^0 \cup \mathcal{B}_N^1$, and thus $x \notin (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$. The proposition follows. $\square$

Proposition 2.2 together with (1.1) and (1.2) implies that $\text{Bin}(P) \neq \mathcal{B}(M) \cap \mathcal{B}(N)$ whenever M and N have even rank but each has a component of odd rank. However, by assuming Theorem 1.1, we can show that $\text{Bin}(P) = \mathcal{B}(M) \cap \mathcal{B}(N)$ in some cases.

COROLLARY 2.3. *If M and N are binary matroids of odd rank and the matroid pair $P = (M, N)$ is connected and does not have a twined K_4 pair as a minor, then*

$$\text{Bin}(P) = \text{Bin}(M) \cap \text{Bin}(N).$$

Proof. Since M and N have odd rank, both $\mathcal{B}_M^1 \cap \mathcal{B}_N^0$ and $\mathcal{B}_M^0 \cap \mathcal{B}_N^1$ are empty. Thus (2.1) implies that $\text{Bin}(M) \cap \text{Bin}(N) = (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$. The corollary now follows from Theorem 1.1. $\square$

3. Fundamental graphs. Let M be a matroid on the ground set S and let $B \in \mathcal{B}(M)$. For $X, Y \subseteq S$, the symmetric difference $(X \setminus Y) \cup (Y \setminus X)$ is denoted $X \Delta Y$. The *partial representation of M with respect to B* is the $\{0, 1\}$ matrix A with rows indexed by B and columns indexed by $S \setminus B$ such that $[A]_{u,v} = 1$ if and only if $B \Delta \{u, v\} \in \mathcal{B}(M)$. Note that when M is binary, the matrix $[I]A$ is a representation of M over $GF(2)$. The *fundamental graph for M with respect to B* is the labeled bipartite graph with vertex set S and edge set given by the adjacency matrix A . This graph is denoted $G(M, B)$, and abbreviated to $G(M)$ when the particular basis B is not significant.

The following proposition relates properties between a matroid and its fundamental graph. These properties can be proven by considering partial representations.

PROPOSITION 3.1. *Let M be a matroid on the ground set S with basis B .*

- (i) *The components of $G(M, B)$ correspond to the components of M .*
- (ii) *If $X \subseteq B$, $Y \subseteq S \setminus B$, then $G(M/X \setminus Y, B \setminus X)$ is the subgraph of $G(M, B)$ induced by $S \setminus (X \cup Y)$.*

Let $B, B' \in \mathcal{B}(M)$ have partial representations A and A' , respectively. If $B' = B \Delta \{u, v\}$, then A' is said to be obtained from A by *pivoting* on edge uv . If, in addition, M is binary, then $G(M, B')$ is said to be obtained from $G(M, B)$ by pivoting on edge uv .

For $u \in S$, let $N(u)$ be $\{v \in S : uv \in E(G(M, B))\}$. We recall that when M is binary, pivoting on uv in $G(M, B)$ results in swapping the labels on vertices u and v and

complementing the edges between $N(u) \setminus \{v\}$ and $N(v) \setminus \{u\}$. For $B, B' \in \mathcal{B}(M)$ with $|B \Delta B'| = 2k$, the basis exchange property of matroids guarantees a sequence of vertices $u_1, v_1, u_2, v_2, \dots, u_k, v_k$ such that B' is obtained from B by successively pivoting on edges $u_1v_1, u_2v_2, \dots, u_kv_k$. The fundamental graph $G(M, B')$ is similarly obtained from $G(M, B)$ by successively pivoting on each edge; this operation is referred to as pivoting on $B' \Delta B$ in $G(M, B)$.

PROPOSITION 3.2. *Let M be a binary matroid on the ground set S with basis B . If $B' \subseteq S$ with $B' \neq B$, then $B' \in \mathcal{B}(M)$ if and only if the subgraph of $G(M, B)$ induced by $B \Delta B'$ has an odd number of perfect matchings.*

Proof. Let A be the partial representation of M with respect to B , and let G' be the subgraph of $G(M, B)$ induced by $B \Delta B'$. Consider the matrix $X = [I|A]$, where the rows and columns of X are indexed by B and S , respectively, such that the indices correspond with the definition of A . Since M is binary, X is a representation over $GF(2)$ of M , and therefore B' is a basis of M if and only if the columns of X indexed by B' are a basis. By possibly permuting rows or columns, we may assume this column submatrix, X' , has the form

$$X' = \begin{matrix} & \begin{matrix} B' \cap B & B' \setminus B \end{matrix} \\ \begin{matrix} B' \cap B \\ B \setminus B' \end{matrix} & \begin{pmatrix} I & A'' \\ 0 & A' \end{pmatrix} \end{matrix}.$$

Note that A' is the incidence matrix for G' . If $|B'| \neq |B|$, then X' is not square and is therefore not a column basis of X . In this case, G' has a different number of vertices on each side of its vertex bipartition, and therefore G' has no perfect matchings. If $|B'| = |B|$, then X' is a column basis of X if and only if A' is nonsingular. Each nonzero term in the permutation expansion of the determinant of A' corresponds to a perfect matching in G' . Since we are working over $GF(2)$, a nonzero determinant occurs only in the case of an odd number of terms in this expansion. $\square$

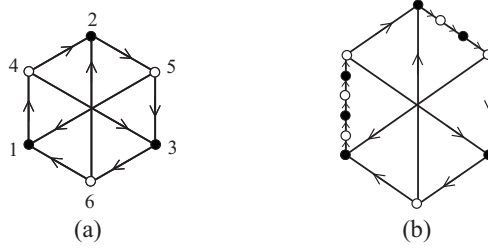
If $P = (M, N)$ is a binary matroid pair on the ground set S with basis B , then the *fundamental graph for P with respect to B* is the labeled directed graph, denoted $G(P, B)$, with vertex set S and edge set $E(G(M, B)) \cup E(G(N, B))$, where the edges of $G(M, B)$ are directed into B and the edges of $G(N, B)$ are directed out of B . When the particular basis B is not significant, $G(P, B)$ is abbreviated to $G(P)$. By Proposition 3.1(ii), the induced subgraphs of $G(P)$ correspond to minors of P .

Let $P = (M, N)$ be a binary matroid pair with $B \in \mathcal{B}(P)$. If $B' \in \mathcal{B}(P)$, then $G(P, B')$ is obtained by pivoting on $B \Delta B'$ in $G(M, B)$ to obtain $G(M, B')$, pivoting on $B \Delta B'$ in $G(N, B)$ to obtain $G(N, B')$, and then identifying the vertices of $G(M, B')$ and $G(N, B')$. This operation is referred to as pivoting on $B \Delta B'$ in $G(P, B)$.

If G is a directed graph and $W \subseteq V(G)$, then $\delta^-(W)$ denotes the edges of G directed from $V(G) \setminus W$ to W . Similarly, $\delta^+(W)$ denotes the edges directed from W to $V(G) \setminus W$. When $\delta^+(W) \neq \emptyset$ for all $\emptyset \neq W \subsetneq V(G)$, then G is *strongly connected*. Stating that a graph is strongly connected will imply that it is directed. We close this section with an observation of Fonlupt and Zemirline [4] relating the connectivity of a matroid pair to the connectivity of the fundamental graph for the pair.

PROPOSITION 3.3. *Let P be a matroid pair. If $G(P, B')$ is strongly connected for some $B' \in \mathcal{B}(P)$, then $G(P, B)$ is strongly connected for all $B \in \mathcal{B}(P)$. Furthermore, P is connected if and only if $G(P)$ is strongly connected.*

4. Twined K_4 pairs. Let M and N be binary matroids on the ground set $S = \{1, 2, 3, 4, 5, 6\}$ and suppose that M and N have the following partial representations

FIG. 4.1. (a) $G(P, B)$, (b) a graph in $\mathcal{G}$.

A_M and A_N with respect to their common basis $B = \{1, 2, 3\}$:

$$(4.1) \quad A_M = \begin{matrix} & 4 & 5 & 6 \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \end{matrix}, \quad A_N = \begin{matrix} & 4 & 5 & 6 \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix}.$$

If $P = (M, N)$, then $G(P, B)$ is the graph in Figure 4.1(a). Note that

$$\mathcal{B}(P) = \{\{1, 2, 3\}, \{1, 5, 6\}, \{2, 4, 6\}, \{3, 4, 5\}\},$$

and thus $|B' \cap \{4, 5, 6\}|$ is even for all $B' \in \mathcal{B}(P)$. Since $|S \cap \{4, 5, 6\}|$ is odd, we have $S \notin \text{Bin}(P)$. However, it can be verified that $S \in \mathcal{B}_M^0 \cap \mathcal{B}_N^0$, and thus $\text{Bin}(P) \neq (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$ for this matroid pair.

Let D be the arc set $\{(1, 4), (2, 5), (3, 6)\}$ of $G(P, B)$ in Figure 4.1. Note that $D = E(G(N, B))$. Let $\mathcal{G}$ be the set of graphs obtained from $G(P, B)$ by possibly subdividing arcs of D into directed paths of odd length, such that the orientation of the path is the same as the orientation of the original arc (see Figure 4.1(b)).

Let $G \in \mathcal{G}$ and let X, Y be the bipartition of $V(G)$. Note that the only induced subgraphs G' of G for which both $G'[\delta^+(X)]$ and $G'[\delta^+(Y)]$ have an odd number of perfect matchings are the three non-Hamiltonian directed cycles of G . Let C be such a cycle of G . It can be verified that pivoting on $V(C)$ in G , as described in section 3, results in a fundamental graph isomorphic to G . Thus, if P is a binary matroid pair with basis B such that $G(P, B) \in \mathcal{G}$, then by Proposition 3.2, $G(P, B') \in \mathcal{G}$ for all $B' \in \mathcal{B}(P)$.

DEFINITION 4.1. A binary matroid pair P with $G(P) \in \mathcal{G}$ is called a *twined K_4 pair*.

5. Vertex cycle space. Our proof of Theorem 1.1 uses the dimension of a vector space defined on the vertices of the fundamental graph for a matroid pair. In this section we define the required vector space and determine its dimension.

Let G be a directed graph, and for each circuit C in G , let $x_C \in GF(2)^{E(G)}$ and $y_C \in GF(2)^{V(G)}$ be the incidence vectors for $E(C)$ and $V(C)$, respectively. The (directed) *cycle space* of G is the vector space over $GF(2)$ generated by $\{x_C : C \text{ is a (directed) circuit in } G\}$. When G has c components, it is well known that the cycle space of G , denoted $\mathcal{C}(G)$, has dimension $|E(G)| - |V(G)| + c$. The cycle space and the directed cycle space are equal for strongly connected graphs (see Bondy and Murty [1]). We define the *vertex cycle space* of G , denoted $\mathcal{C}_V(G)$, as the vector space over $GF(2)$ generated by $\{y_C : C \text{ is a directed circuit in } G\}$.

THEOREM 5.1. Let G be a strongly connected bipartite graph with vertex bipartition $\{X, Y\}$. If c_X and c_Y are the number of connected components in $G[\delta^+(X)]$ and

$G[\delta^+(Y)]$, respectively, then the dimension of $\mathcal{C}_V(G)$ is

$$(5.1) \quad |X| + |Y| - c_X - c_Y + 1.$$

Proof. Let $f : \mathcal{C}(G) \rightarrow GF(2)^{X \cup Y}$ be the linear mapping with $f(x)_u = x(\delta^+(u))$ if $u \in X$, and $f(x)_u = x(\delta^-(u))$ if $u \in Y$. Then $f(x_C) = y_C$ for all directed circuits C in G , and therefore $\mathcal{C}_V(G)$ is the image of f . It follows that

$$(5.2) \quad \dim \mathcal{C}_V(G) = \dim \mathcal{C}(G) - \text{null} f = |E(G)| - |X| - |Y| + 1 - \text{null} f.$$

It remains to find $\text{null} f$.

For $Z \in \{X, Y\}$, let $G_Z = G[\delta^+(Z)]$. Note that

$$(5.3) \quad \begin{aligned} \mathcal{C}(G_Z) &= \{x \in GF(2)^{E(G_Z)} : x(\delta^+(u)) = 0 \text{ if } u \in Z, \\ &\quad x(\delta^-(u)) = 0 \text{ if } u \in V(G) \setminus Z\}. \end{aligned}$$

For $x \in GF(2)^{E(G)}$, let x^Z be x restricted to $E(G_Z)$. Conversely, given $x^X \in GF(2)^{E(G_X)}$ and $x^Y \in GF(2)^{E(G_Y)}$, let x be the associated vector in $GF(2)^{E(G)}$. We claim that the nullspace of f is

$$(5.4) \quad \{x \in GF(2)^{E(G)} : x^X \in \mathcal{C}(G_X) \text{ and } x^Y \in \mathcal{C}(G_Y)\}.$$

Certainly, if $x^X \in \mathcal{C}(G_X)$ and $x^Y \in \mathcal{C}(G_Y)$, then $x \in \mathcal{C}(G)$. By (5.3), $f(x) = 0$ in this case. Conversely, suppose $f(x) = 0$. Since $x \in \mathcal{C}(G)$, $x(\delta^+(u)) = x(\delta^-(u))$ for all $u \in X \cup Y$, and thus $x^X \in \mathcal{C}(G_X)$ and $x^Y \in \mathcal{C}(G_Y)$ by (5.3). Hence the nullspace of f is (5.4), as claimed. Thus

$$\text{null} f = \dim \mathcal{C}(G_X) + \dim \mathcal{C}(G_Y) = |E(G)| - 2(|X| + |Y|) + c_X + c_Y,$$

and the theorem now follows from (5.2). $\square$

6. Characterization by fundamental graphs. For a binary matroid pair P , we use Theorem 5.1 to characterize $\text{Bin}(P)$ in terms of $\mathcal{C}_V(G(P))$. This characterization will be the main tool in our proof of Theorem 1.1, and will also be used to show that the converse of Theorem 1.1 is false.

First note that $\mathcal{C}_V(G(P, B))$ does not depend on the particular basis B of P .

LEMMA 6.1. *Let $P = (M, N)$ be a connected binary matroid pair on the ground set S . If $B \in \mathcal{B}(P)$, then $\mathcal{C}_V(G(P, B)) = \mathcal{B}_M^0 \cap \mathcal{B}_N^0$.*

Proof. If C is a directed circuit in $G = G(P, B)$, then $E(C) \cap E(G(M, B))$ is a matching in $G(M, B)$, and thus $V(C)$ intersects each component in $G(M, B)$ with even parity. Rieder's characterization of $\mathcal{B}_M^0$ then implies that $V(C) \in \mathcal{B}_M^0$. Similarly, $V(C) \in \mathcal{B}_N^0$. Since C was arbitrary, $\mathcal{C}_V(G) \subseteq \mathcal{B}_M^0 \cap \mathcal{B}_N^0$.

By Proposition 3.1(i), $G[\delta^-(B)]$ and $G[\delta^+(B)]$ have c_M and c_N components, respectively, and by Proposition 3.3, G is strongly connected. Theorem 5.1 therefore implies that $\mathcal{C}_V(G)$ has dimension $|S| - c_M - c_N + 1$. Hence, by Proposition 2.1, $\mathcal{C}_V(G)$ and $\mathcal{B}_M^0 \cap \mathcal{B}_N^0$ have the same dimension, and the lemma follows. $\square$

THEOREM 6.2. *If $P = (M, N)$ is a connected binary matroid pair, then*

$$(6.1) \quad \text{Bin}(P) = (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$$

if and only if $\mathcal{C}_V(G(P)) \subseteq \text{Bin}(P)$.

Proof. Recall that $\text{Bin}(P) \subseteq (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$, and let $G = G(P, B)$ for some $B \in \mathcal{B}(P)$. If (6.1) holds, then $\mathcal{C}_V(G) \subseteq \text{Bin}(P)$ by Lemma 6.1. Conversely, suppose

that $\mathcal{C}_V(G) \subseteq \text{Bin}(P)$. If every component of M and every component of N has even rank, then $\mathcal{B}_M^1 \cap \mathcal{B}_N^1 = \mathcal{B}_M^0 \cap \mathcal{B}_N^0$ and (6.1) holds by Lemma 6.1. If a component of M or N has odd rank, then $B \notin \mathcal{B}_M^0 \cap \mathcal{B}_N^0$, and thus by Lemma 6.1, the dimension of $\text{Bin}(P)$ is at least one greater than the dimension of $\mathcal{B}_M^0 \cap \mathcal{B}_N^0$. By Proposition 2.1, the dimension of $\text{Bin}(P)$ is then at least the dimension of $(\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$. This yields (6.1). $\square$

When $P = (M, N)$ is a connected matroid pair on the ground set S , and M and N are simple matroids with c_M and c_N components, respectively, Fonlupt and Zemirline [4] show that the convex hull generated by $\mathcal{B}(P)$, over $\mathbb{R}$, has dimension $|S| - c_M - c_N + 1$. Theorem 6.2 together with Proposition 2.1 implies the following corollary about the dimension of $\text{Bin}(P)$.

COROLLARY 6.3. *Let $P = (M, N)$ be a connected binary matroid pair on the ground set S with $\mathcal{C}_V(G(P)) \subseteq \text{Bin}(P)$. If M and N have c_M and c_N components, respectively, then $\text{Bin}(P)$ has dimension $|S| - c_M - c_N + 1$ when every component of M and N has even rank, and dimension $|S| - c_M - c_N + 2$ otherwise.*

While Theorem 1.1 provides sufficient conditions for the equality of $\text{Bin}(P)$ and $(\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$, the converse of the theorem is false. For example, let $P = (M, N)$ be the binary matroid pair on the ground set $S = \{1, 2, \dots, 8\}$ with basis $B = \{1, 2, 3, 4\}$ and fundamental graph $G(P, B)$ as shown in Figure 6.1.

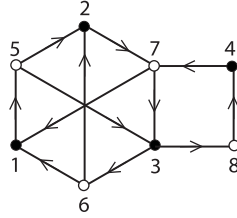


FIG. 6.1. $G(P, B)$ has a twined K_4 pair as a minor.

There are exactly seven directed circuits in $G(P, B)$, and it can be verified that the vertex set for each circuit is contained in $\text{Bin}(P)$. Thus, by Theorem 6.2, $\text{Bin}(P) = (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$. However, $G(P, B) \setminus \{4, 8\} \in \mathcal{G}$, and thus $P/\{4\} \setminus \{8\}$ is a twined K_4 pair.

7. Proof of the main theorem. In this section, Theorem 6.2 is used to prove the main theorem, Theorem 1.1. We first require some terminology. Given subgraphs G_1 and G_2 of a graph G , let $G_1 \cup G_2$ denote the subgraph with vertices $V(G_1) \cup V(G_2)$ and edges $E(G_1) \cup E(G_2)$. If C and C' are directed circuits in the directed graph G , then the components of the subgraph with vertices $V(C) \cap V(C')$ and edges $E(C) \cap E(C')$ are called the *intersection paths* between C and C' . Note that an intersection path may be an isolated vertex. If there are at least two intersection paths between C and C' , and the order of the paths traveling on C is the reverse of the order of the paths traveling on C' , then we say that C and C' are *twined* circuits. For example, the two Hamilton circuits in the graph of Figure 4.1(b) are twined and have three intersection paths. We require one lemma about twined circuits.

LEMMA 7.1. *Let C and C' be directed circuits in the directed graph G . If there are at least two intersection paths between C and C' , then there exists a directed circuit in $C \cup C'$ that is twined with C .*

Proof. Let C and C' have k intersection paths $P_1, P_2, \dots, P_k$. Assume that the order of the paths traveling on C is $P_1, P_2, \dots, P_k$. Let α be the permutation of

$\{1, 2, \dots, k\}$ such that $\alpha(1) = 1$ and $P_{\alpha(k)}, P_{\alpha(k-1)}, \dots, P_{\alpha(1)}$ is the order of the paths on C' . If $\alpha(i) = i$ for all i , then C and C' are twined. Otherwise, let n be minimum such that $\alpha(n) > \alpha(n+1)$ and let $m < n$ be maximum such that $\alpha(m) < \alpha(n+1)$. The directed path in C from the first vertex of $P_{\alpha(m)}$ to the first vertex of $P_{\alpha(n+1)}$, followed by the directed path in C' from the first vertex of $P_{\alpha(n+1)}$ to the first vertex of $P_{\alpha(m)}$, forms a directed circuit C'' in $C \cup C'$ that is twined with C . The number of intersection paths between C'' and C is $n+1-m \geq 2$. $\square$

We now prove our main theorem, which we restate here.

THEOREM 1.1. *If M and N are binary matroids and the matroid pair $P = (M, N)$ is connected and does not have a twined K_4 pair as a minor, then*

$$\text{Bin}(P) = (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1).$$

Proof. Suppose that $\text{Bin}(P) \neq (\mathcal{B}_M^0 \cap \mathcal{B}_N^0) \cup (\mathcal{B}_M^1 \cap \mathcal{B}_N^1)$. By Theorem 6.2, there is a directed circuit C in $G(P)$ with $V(C) \notin \text{Bin}(P)$. Over all $B \in \mathcal{B}(P)$ and all directed circuits C in $G(P, B)$ with $V(C) \notin \text{Bin}(P)$, choose B and C to minimize the length of C , denoted $|C|$. Let G be the subgraph of $G(P, B)$ induced by $V(C)$, and let G_M and G_N be the subgraphs of G induced by the edges of $G(M, B)$ and $G(N, B)$, respectively. Given a directed circuit C' in G with chord e , let C'_e denote the unique directed circuit in $C' \cup \{e\}$ with $e \in E(C'_e)$.

CLAIM 7.1.1. *If C' is a directed circuit in G and C' is twined with C , then C' is a Hamilton circuit of G and $|E(C') \setminus E(C)|$ is odd.*

Subproof. Suppose that the directed circuit C' in G is twined with C and let $E(C') \setminus E(C) = \{e_1, \dots, e_k\}$, where the edges are labeled to respect their cyclic order on C' . Since C and C' are twined,

$$(7.1) \quad V(C') = V(C) \Delta V(C_{e_1}) \Delta V(C_{e_2}) \Delta \dots \Delta V(C_{e_k}).$$

Since G does not have parallel edges, $|C_e| < |C|$ for all $e \in E(G) \setminus E(C)$. In particular, $|C_{e_i}| < |C|$ for all $i \in \{1, \dots, k\}$, and thus the minimality of $|C|$ implies

$$(7.2) \quad V(C_{e_i}) \in \text{Bin}(P) \quad \text{for all } i \in \{1, \dots, k\}.$$

Since $V(C) \notin \text{Bin}(P)$, (7.1) and (7.2) imply that $V(C') \notin \text{Bin}(P)$. The minimality of $|C|$ implies that $|C'| = |C|$ and therefore C' is a Hamilton circuit of G . If k is even, say $k = 2l$, then since C' is a Hamilton circuit,

$$(7.3) \quad V(C) = \bigtriangleup_{i=1}^l V(C_{e_{2i}}).$$

Together with (7.2), equation (7.3) contradicts the assumption that $V(C) \notin \text{Bin}(P)$. Thus k is odd and the claim follows. $\square$

CLAIM 7.1.2. *There exists a directed circuit in G that is twined with C .*

Subproof. Consider the perfect matching $T_M = E(G_M) \cap E(C)$ of G_M and the perfect matching $T_N = E(G_N) \cap E(C)$ of G_N . Since $V(C) \notin \text{Bin}(P)$ and $B \in \mathcal{B}(P)$, $V(C) \Delta B$ is not a basis of P . Without loss of generality, we may assume that $V(C) \Delta B \notin \mathcal{B}(N)$. By Proposition 3.2, G_N has a perfect matching T'_N with $T'_N \neq T_N$. Since T_M and T'_N are both perfect matchings of G , the union of T_M and T'_N is a set of disjoint directed circuits which spans $V(G)$. Assume that $T_M \cup T'_N$ is the disjoint union of m directed circuits $C^1, \dots, C^m$. Since $V(C) = V(C^1) \Delta \dots \Delta V(C^m)$ and $V(C) \notin \text{Bin}(N)$, there exists $i \in \{1, \dots, m\}$ with $V(C^i) \notin \text{Bin}(N)$. If $m > 1$, then $|C^i| < |C|$, and thus C^i contradicts the choice of C . It follows that $m = 1$, and

$T_M \cup T'_N$ is therefore a directed circuit C' with $V(C') = V(C)$. Since $T_N \neq T'_N$, the circuits C' and C are distinct and therefore have at least two intersection paths. By Lemma 7.1, G has a directed circuit that is twined with C . $\square$

Over all directed circuits in G that are twined with C , choose the circuit C' to minimize $|E(C') \setminus E(C)|$. Let $k = |E(C') \setminus E(C)|$ and let $P_1, P_2, \dots, P_k$ be the intersection paths between C and C' , where the indices respect the cyclic order of the paths on C . For $i \in \{1, \dots, k\}$, let $x_i, y_i \in V(G)$ be the first and last vertices, respectively, of P_i . Then

$$E(C) \setminus E(C') = \{y_1 x_2, y_2 x_3, \dots, y_k x_1\},$$

$$E(C') \setminus E(C) = \{y_1 x_k, y_2 x_1, \dots, y_k x_{k-1}\}.$$

Without loss of generality we may assume that $x_1 \in B$. Since k is odd (Claim 7.1.1), it follows that $x_i \in B$ for all i . The intersection paths between C and C' therefore have odd length, and to prove that the matroid pair minor corresponding to G is a twined K_4 pair, it remains to show that $G = C \cup C'$ and $k = 3$.

Given a directed path Q in G and $x, y \in V(Q)$ with $xy \in E(G) \setminus E(Q)$, call xy a *forward arc* of Q if x precedes y along Q , and call xy a *backward arc* of Q otherwise.

CLAIM 7.1.3. *If $e \in E(G) \setminus E(C \cup C')$, then e is a backward arc of an intersection path between C and C' .*

Subproof. Let $x, y \in V(G)$ be such that $xy \in E(G) \setminus E(C \cup C')$. Since C and C' are Hamiltonian in G , there exist $i, j \in \{1, \dots, k\}$ such that $x \in V(P_i)$ and $y \in V(P_j)$. Without loss of generality we may assume that $i = 1$.

If $j \neq 1$, let $C'' = C'_{xy}$. Then C'' is not a Hamilton circuit of G , but C and C'' are twined with $j \geq 2$ intersection paths. This contradicts Claim 7.1.1; hence we may assume that $x, y \in V(P_1)$.

Suppose that xy is a forward arc of P_1 . Let C'' be the directed circuit formed by the union of xy , the path from y to y_2 in C , the arc $y_2 x_1$ in C' , and the path from x_1 to x in P_1 . Then C'' is not a Hamilton circuit of G , but C and C'' are twined with two intersection paths. This again contradicts Claim 7.1.1, and thus xy is a backward arc of P_1 . $\square$

CLAIM 7.1.4. *Let $xy \in E(G) \setminus E(C)$ be such that C_{xy} is induced, and let $x', y' \in V(G)$ be such that $x'y, xy' \in E(C)$. Furthermore, assume that $k > 3$ if $xy \in E(C')$. Then $B' = B \Delta V(C_{xy})$ is a basis of P , and $C_{x'y'}$ is a directed circuit in $G(P, B')$.*

Subproof. Since C_{xy} is induced, $B' \in \mathcal{B}(P)$ by Proposition 3.2. We consider how successively pivoting on each edge of C_{xy} affects G . Recall that pivoting in G implies separate pivoting in $G(M, B)$ and in $G(N, B)$. Recall also that if $uv \in E(G(M, B))$, then pivoting on uv deletes the edge $u'v'$ if and only if $u'v' \in E(G(M, B))$, $u' \neq u, v' \neq v$, and both $u'v \in E(G(M, B))$ and $uv' \in E(G(M, B))$. Similarly, pivoting on uv adds the edge $u'v'$ if and only if $u'v' \notin E(G(M, B))$, $u' \neq u, v' \neq v$, and both $u'v \in E(G(M, B))$ and $uv' \in E(G(M, B))$.

Let $uv \in E(C_{xy})$, and let Q be the path in C from y' to x' . Let $u'v' \in E(Q)$ be such that the two edges $u'v'$ and uv are either both in $G(M, B)$ or both in $G(N, B)$. By Claim 7.1.3, either $u'v \notin E(G)$ or $uv' \notin E(G)$. Thus, by the properties of pivoting in G , the edge $u'v'$ remains after pivoting on uv . Since C_{xy} is induced, if $z \in V(C_{xy}) \setminus \{u, v\}$, then $zv \notin E(G)$ and $uz \notin E(G)$. Thus pivoting on uv does not change $N(z)$. Hence, after successively pivoting on each edge of C_{xy} , all edges of Q remain.

Note that by Claim 7.1.3, $x'y' \notin E(G)$. Let $uv \in E(C_{xy})$ with $uv \neq xy$. If $xy \notin E(C')$, then neither $x'v$ nor uy' is an edge of G by Claim 7.1.3, so pivoting on

uv does not add edge $x'y'$. If $xy \in E(C')$, then since $k > 3$, either $x'v \notin E(G)$ or $uy' \notin E(G)$ by Claim 7.1.3, and again pivoting on uv does not add edge $x'y'$. However, since $x'y, xy', xy \in E(G)$, pivoting on xy adds edge $x'y'$. It follows that after pivoting on each edge of C_{xy} , the edge $x'y'$ has been added. Since $C_{x'y'} = Q \cup x'y'$, we conclude that the directed circuit $C_{x'y'}$ is a subgraph of $G(P, B')$. $\square$

CLAIM 7.1.5. $G = C \cup C'$.

Subproof. Suppose instead that $xy \in E(G) \setminus E(C \cup C')$. By Claim 7.1.3, xy is a backward arc of an intersection path P_j between C and C' . Choose xy such that $|C_{xy}|$ is minimum and let $x', y' \in V(G)$ be such that $x'y, xy' \in E(C)$.

The minimality of $|C_{xy}|$ together with Claim 7.1.3 implies that C_{xy} is an induced subgraph of G . Let $B' = B \Delta V(C_{xy})$. By Claim 7.1.4, $B' \in \mathcal{B}(P)$ and $C_{x'y'}$ is a directed circuit in $G(P, B')$.

By the minimality of $|C|$, both $V(C_{x'y'})$ and $V(C_{xy})$ are contained in $\text{Bin}(P)$. Since $V(C) = V(C_{x'y'}) \Delta V(C_{xy})$, this contradicts the assumption that $V(C) \notin \text{Bin}(P)$. Thus $E(G) \setminus E(C \cup C') = \emptyset$ and $G = C \cup C'$. $\square$

CLAIM 7.1.6. *There are exactly three intersection paths between C and C' .*

Subproof. Suppose instead that $k > 3$ and consider $C_{y_2x_1}$. By Claim 7.1.5, $C_{y_2x_1}$ is induced in G , and therefore, by Proposition 3.2, $B' = B \Delta V(C_{y_2x_1})$ is a basis of P . Since $y_2x_3, y_kx_1 \in E(C)$, Claim 7.1.4 implies that $C_{y_kx_3}$ is a directed circuit in $G(P, B')$. By the minimality of $|C|$, both $V(C_{y_kx_3})$ and $V(C_{y_2x_1})$ are contained in $\text{Bin}(P)$. Since $V(C) = V(C_{y_kx_3}) \Delta V(C_{y_2x_1})$, this contradicts the assumption that $V(C) \notin \text{Bin}(P)$. Thus $k = 3$. $\square$

Claims 7.1.5 and 7.1.6 imply that the minor of P corresponding to G is a twined K_4 pair, thus completing the proof of Theorem 1.1. $\square$

We close by noting that Theorem 1.1 can be applied to binary matroid pairs that are not connected. For example, let $P = (M, N)$ be a binary matroid pair on the ground set S where M and N have rank r , and suppose that P is not connected. If there exists a partition $\{S_1, S_2\}$ of S such that $\text{rank}_M(S_1) + \text{rank}_N(S_2) < r$, then $\mathcal{B}(P)$ is empty and thus $\text{Bin}(P)$ is trivial. If there exists a partition $\{S_1, S_2\}$ of S such that $\text{rank}_M(S_1) + \text{rank}_N(S_2) = r$, then $B \in \mathcal{B}(P)$ if and only if $B \cap S_1$ is a basis of $(M \setminus S_2, N \setminus S_2)$ and $B \cap S_2$ is a basis of $(M/S_1, N/S_1)$ (see Fonlupt and Zemirline [4]). Determining $\text{Bin}(P)$ in this case thus reduces to determining the vector space over $GF(2)$ generated by the common bases of two smaller pairs of matroids.

Acknowledgments. I thank Jim Geelen for many helpful discussions, and thank the referees for many helpful suggestions.

REFERENCES

- [1] J. A. BONDY AND U. S. R. MURTY, *Graph Theory with Applications*, American Elsevier, New York, 1976.
- [2] G. CORNUÉJOLS, *Combinatorial Optimization: Packing and Covering*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 74, SIAM, Philadelphia, PA, 2001.
- [3] J. EDMONDS, *Submodular functions, matroids, and certain polyhedra*, in Combinatorial Structures and Their Applications, Proc. Calgary Internat. Conf., Calgary, Alta., Gordon and Breach, New York, 1970, pp. 69–87.
- [4] J. FONLUPT AND A. ZEMIRLINE *On the number of common bases of two matroids*, Discrete Math., 45 (1983), pp. 217–228.
- [5] J. G. OXLEY, *Matroid Theory*, Oxford University Press, New York, 1992.
- [6] J. RIEDER, *The lattices of matroid bases and exact matroid bases*, Arch. Math., 56 (1991), pp. 616–623.
- [7] K. P. WEBB, *Counting Bases*, Ph.D. thesis, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON, 2004.

DETERMINANTS AND LONGEST CYCLES OF GRAPHS*

VLADIMIR EJOV[†], JERZY A. FILAR[†], WALTER MURRAY[‡], AND GIANG T. NGUYEN[†]

Abstract. We consider the Hamiltonian cycle problem on a given graph G . With such a graph we can associate a family $\mathcal{F}$ of probability transition matrices of Markov chains whose entries represent the probabilities of traversing corresponding arcs of the graph. When the underlying graph is Hamiltonian, we show the transition probability matrix induced by a Hamiltonian cycle maximizes—over $\mathcal{F}$ —the determinant of a matrix that is a rank-one correction of the generator matrix of a Markov chain. In the case when the graph does not possess a Hamiltonian cycle, the above maximization yields a transition matrix of a chain with a longest simple cycle (in G) comprising that chain’s unique ergodic class. These problems also have analogous eigenvalue interpretations.

Key words. Markov chains, Hamiltonian cycle

AMS subject classifications. 60J10, 05C45, 11C20

DOI. 10.1137/070693898

1. Introduction. Briefly stated, the Hamiltonian cycle problem (HCP) is as follows: *Given a graph, find a simple cycle that contains all vertices of the graph (Hamiltonian cycle) or prove that a Hamiltonian cycle does not exist.* With respect to this property—*Hamiltonicity*—graphs possessing one or more Hamiltonian cycles are called Hamiltonian.¹ By a *graph* of order N we will mean a simple N -vertex graph (without self-loops and multiple edges) that can be both symmetric (every edge admits the two-way traffic) or directed (*digraph*, with one-way traffic along each edge).

This paper evolved from a line of research [6], [4], [1], [2] (surveyed in [5]) that exploits Markov chains to study properties of the HCP. More specifically, we consider the HCP as an optimization problem over the space $\mathcal{F}$ of probability transition matrices of Markov chains whose entries represent the probabilities of traversing corresponding arcs of a graph. We prove that when the underlying graph is Hamiltonian the transition probability matrix induced by a Hamiltonian cycle maximizes, over $\mathcal{F}$, the determinant of a matrix that is a rank-one correction of the (negative) generator matrix of the corresponding Markov chain. In the case when the graph does not possess a Hamiltonian cycle, the above maximization yields a transition matrix of a chain with a longest simple cycle (in G) comprising that chain’s single, nonexhaustive, ergodic class. These problems also have analogous eigenvalue interpretations.

In notation made precise later, the main result states that for all possible transition probability matrices $\mathbf{P} \in \mathcal{F}$ generated by the given graph G

$$\det[\mathbf{I} - \mathbf{P} + \alpha\mathbf{J}] \leq \alpha kN,$$

*Received by the editors June 7, 2007; accepted for publication (in revised form) March 31, 2008; published electronically July 16, 2008.

<http://www.siam.org/journals/sidma/22-3/69389.html>

[†]Centre for Industrial and Applied Mathematics, University of South Australia, Mawson Lakes, SA 5095, Australia (Vladimir.Ejov@unisa.edu.au, Jerzy.Filar@unisa.edu.au, Giang.Nguyen@unisa.edu.au). The research of these authors was supported by Australian Research Council Discovery grant DP0666632 and the Australian Research Council Linkage International grant LX0881972.

[‡]Department of Management Science, Stanford University, Stanford, CA 94305-4023 (walter@SOL-walter.stanford.edu). The research of this author was supported by Department of the Army grant W911NF-07-2-0027 and Office of Naval Research grant N000140810191.

¹The name of the problem is due to Sir William Hamilton’s investigations of the existence of such cycles on the dodecahedron graph [7].

where $\alpha > 0$ is a fixed parameter, k is the length of the longest simple cycle in G , N is the number of nodes, and $\mathbf{J}$ is an $N \times N$ matrix of 1's. If $\mathbf{e}$ denotes the column vector of 1's, then $\mathbf{J}$ is a rank-one matrix $\mathbf{e}\mathbf{e}^T$. Furthermore, the above upper bound is attained at some $\mathbf{P} \in \mathcal{F}$.

The paper is organized as follows. In section 2 we introduce some basic concepts and the necessary notation. In section 3 we discuss a numerical example that motivates the theoretical results derived in section 4. Conclusions are supplied in section 5.

2. Preliminaries and notation. Consider a connected graph G of size N . Let $V(G)$ and $E(G)$ be the set of nodes and the set of arcs in G , respectively, where $|V(G)| = N$. Let $\mathcal{F}$ be the set of all probability transition matrices $\mathbf{P}$ of Markov chains whose entries represent the probabilities of traversing corresponding arcs of the graphs. Naturally, every element in $\mathbf{P}$ is nonnegative and $\mathbf{P}$ is stochastic, which means every row sums to 1. That is,

$$\mathcal{F} := \left\{ \mathbf{P} \mid p_{ij} \geq 0; \sum_j p_{ij} = 1 \forall i; p_{ij} = 0 \forall (i, j) \notin E(G) \right\}.$$

Let $\mathcal{D} \subset \mathcal{F}$ be the set of all probability transition matrices $\mathbf{P}$ of deterministic Markov chains,

$$\mathcal{D} := \{ \mathbf{P} \in \mathcal{F} \mid p_{ij} \in \{0, 1\} \},$$

and let $\mathcal{DS} \subset \mathcal{F}$ be the set of all probability transition matrices $\mathbf{P}$ of doubly stochastic Markov chains,

$$\mathcal{DS} := \left\{ \mathbf{P} \in \mathcal{F} \mid \sum_i p_{ij} = 1 \forall j \right\}.$$

Remark 2.1. Note that all Hamiltonian cycles (if any) of a given graph G correspond to probability transition matrices of Markov chains that are both irreducible² and doubly stochastic. These probability transition matrices are also called Hamiltonian cycles.

3. Determinant of rank-one corrected generator matrix. Define

$$(3.1) \quad \mathbf{A}(\mathbf{P}) := \mathbf{I} - \mathbf{P} + \frac{1}{N}\mathbf{J},$$

where $\mathbf{J} = \mathbf{e}\mathbf{e}^T$. Recall that in the literature on Markov chains, the matrix $\mathbf{P} - \mathbf{I}$ is often called the *generator matrix* of the chain whose probability transition matrix is $\mathbf{P}$.

Example 3.1. Consider a complete graph G_4 on 4 nodes (without self-loops). Take three deterministic Markov chains with probability transition matrices $\mathbf{P}_1$, $\mathbf{P}_2$, and $\mathbf{P}_3$:

$$\mathbf{P}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{P}_3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

²A Markov chain is irreducible if all states belong to a single ergodic class.

It is easy to verify that

$$\mathbf{A}(\mathbf{P}_1) = \mathbf{I} - \mathbf{P}_1 + \frac{1}{4}\mathbf{J} = \frac{1}{4} \begin{bmatrix} 5 & -3 & 1 & 1 \\ 1 & 5 & -3 & 1 \\ 1 & 1 & 5 & -3 \\ -3 & 1 & 1 & 5 \end{bmatrix},$$

$$\mathbf{A}(\mathbf{P}_2) = \mathbf{I} - \mathbf{P}_2 + \frac{1}{4}\mathbf{J} = \frac{1}{4} \begin{bmatrix} 5 & -3 & 1 & 1 \\ -3 & 5 & 1 & 1 \\ 1 & 1 & 5 & -3 \\ 1 & 1 & -3 & 5 \end{bmatrix},$$

and

$$\mathbf{A}(\mathbf{P}_3) = \mathbf{I} - \mathbf{P}_3 + \frac{1}{4}\mathbf{J} = \frac{1}{4} \begin{bmatrix} 5 & -3 & 1 & 1 \\ 1 & 5 & -3 & 1 \\ -3 & 1 & 5 & 1 \\ 1 & -3 & 1 & 5 \end{bmatrix}.$$

Calculating the corresponding determinants yields

$$\det \mathbf{A}(\mathbf{P}_1) = 4, \quad \det \mathbf{A}(\mathbf{P}_2) = 0, \quad \text{and} \quad \det \mathbf{A}(\mathbf{P}_3) = 3.$$

Incidentally, the easiest way to verify that $\det \mathbf{A}(\mathbf{P}_2) = 0$ is to multiply $\mathbf{A}(\mathbf{P}_2)$ by the column vector $(1, 1, -1, -1)^T$.

It will be shown later that it is not a coincidence that $\mathbf{P}_1$ corresponds to a Hamiltonian cycle and the determinant of $\mathbf{A}(\mathbf{P}_1)$ is exactly equal to the number of nodes in that longest of all possible simple cycles. Analogously, $\mathbf{P}_3$ corresponds to a Markov chain that traces out a 3-cycle and a transient state and $\det \mathbf{A}(\mathbf{P}_2)$ is again equal to the number of nodes in the cycle containing the home node. While $\mathbf{P}_2$ may appear to be a counterexample to this pattern, it can be excluded because it contains more than one ergodic class.

4. Results. When a Markov chain traces out a Hamiltonian cycle, its probability transition matrix $\mathbf{P}$ will be denoted by $\mathbf{P}_H$. Our results will be proved under a little more generality by replacing the coefficient $\frac{1}{N}$ of the matrix $\mathbf{J}$ by a parameter $\alpha > 0$. Thus, for $\mathbf{P} \in \mathcal{F}$, define

$$\mathbf{A}_\alpha(\mathbf{P}) = \mathbf{I} - \mathbf{P} + \alpha\mathbf{J}, \quad \alpha > 0.$$

For the special value $\alpha = \frac{1}{N}$, we write

$$\mathbf{A}(\mathbf{P}) := \mathbf{A}_{\frac{1}{N}}(\mathbf{P}) = \mathbf{I} - \mathbf{P} + \frac{1}{N}\mathbf{J},$$

as in (3.1).

THEOREM 4.1. *For any graph G*

$$\det_{\mathbf{P} \in \mathcal{F}} \mathbf{A}_\alpha(\mathbf{P}) \leq \alpha kN,$$

where k is the size of the longest cycle in the graph.

In order to prove Theorem 4.1, we will require several results. Lemma 4.2 states the relationship between $\det \mathbf{A}_\alpha(\mathbf{P})$ and the eigenvalues of $\mathbf{P}$. Propositions 4.3, 4.4, and 4.5 prove Theorem 4.1 for three disjoint exhaustive classes of deterministic Markov chains: Hamiltonian cycles, Markov chains with two or more independent cycles (corresponding to two or more ergodic classes), and Markov chains with only a single cycle of length $k < N$, respectively. Proposition 4.6 shows that for any $\mathbf{P} \in \mathcal{F}$, the determinant of $\mathbf{A}_\alpha(\mathbf{P})$ is bounded above by the corresponding determinant of some deterministic Markov chain.

LEMMA 4.2. *Consider any graph G and $\mathbf{P} \in \mathcal{F}$ that has eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_{N-1}$ and $\lambda_N = 1$. Then*

$$\det_{\mathbf{P} \in \mathcal{F}} \mathbf{A}_\alpha(\mathbf{P}) = \alpha N \prod_{i=1}^{N-1} (1 - \lambda_i).$$

Proof. Let μ_i 's be the eigenvalues of $\mathbf{A}_\alpha(\mathbf{P})$. Then

$$\det_{\mathbf{P} \in \mathcal{F}} \mathbf{A}_\alpha(\mathbf{P}) = \prod_{i=1}^N \mu_i.$$

We shall show that

$$\prod_{i=1}^N \mu_i = \alpha N \prod_{i=1}^{N-1} (1 - \lambda_i).$$

We start with the case $\mathbf{P} \in \mathcal{DS}$. For any $\mathbf{P} \in \mathcal{DS}$, $\ker(\mathbf{J})$ is $\mathbf{P}$ -invariant. This follows from the fact that $\mathbf{JP} = \mathbf{PJ} = \mathbf{J}$. If $\mathbf{v} \in \ker(\mathbf{J})$, then $\mathbf{A}_\alpha(\mathbf{P})\mathbf{v} = \mathbf{v} - \mathbf{Pv} + \mathbf{0} \in \ker(\mathbf{J})$. Therefore, the eigenvectors $\mathbf{u}_1, \dots, \mathbf{u}_N$ for $\mathbf{A}_\alpha(\mathbf{P})$ can be chosen so that $\mathbf{u}_N = \mathbf{e} = [1, 1, \dots, 1]$ and the remaining eigenvectors $\mathbf{u}_1, \dots, \mathbf{u}_{N-1}$ span $\ker(\mathbf{J})$. We observe that $\mathbf{e}$ is an eigenvector of $\mathbf{P}$ with eigenvalue 1 and $\mu_N = \alpha N$ since $\mathbf{A}_\alpha(\mathbf{P})\mathbf{e} = \alpha N\mathbf{e}$. As $\mathbf{u}_1, \dots, \mathbf{u}_{N-1}$ span the $\ker(\mathbf{J})$, then $\mu_i = 1 - \lambda_i$, $i = 1, \dots, N-1$, where $\lambda_1, \dots, \lambda_{N-1}$ are the remaining eigenvalues of $\mathbf{P}$.

Hence,

$$\det \mathbf{A}_\alpha(\mathbf{P}) = \alpha N \prod_{i=1}^{N-1} (1 - \lambda_i).$$

Consider now the general case $\mathbf{P} \in \mathcal{F}$. We recall that $\mathbf{e}$ is an eigenvector of $-\mathbf{P}$ with eigenvalue -1 . Let $\mathbf{u}$ be an eigenvector of $-\mathbf{P} + \alpha\mathbf{J}$ with eigenvalue $-\lambda$, $\lambda \neq 1$. Then $\mathbf{u} = \mathbf{v} + \mu\mathbf{e}$ for some $\mathbf{v} \in \ker(\mathbf{J})$. We can then show that there exists vector $\mathbf{w} = \mathbf{v} + \beta\mathbf{e}$ such that $\mathbf{Pw} = \lambda\mathbf{w}$. Indeed, take $\beta = \frac{\mu(1-\alpha N-\lambda)}{1-\lambda}$. Then

$$\begin{aligned} \mathbf{Pw} &= \mathbf{P}(\mathbf{v} + \mu\mathbf{e} - \mu\mathbf{e} + \beta\mathbf{e}) = (\mathbf{P} - \alpha\mathbf{J} + \alpha\mathbf{J})(\mathbf{v} + \mu\mathbf{e} - \mu\mathbf{e}) + \mathbf{P}\beta\mathbf{e} \\ &= \lambda\mathbf{v} + \mu(\lambda + \alpha N - 1)\mathbf{e} + \beta\mathbf{e} = \lambda\mathbf{w}. \end{aligned}$$

Thus, every eigenvalue $\lambda \neq 1$ of $\mathbf{P}$ is also an eigenvalue of $\mathbf{P} - \alpha\mathbf{J}$, and vice versa, which completes the proof. $\square$

PROPOSITION 4.3. *For any Hamiltonian cycle $\mathbf{P}_H \in \mathcal{D}$*

$$\det \mathbf{A}_\alpha(\mathbf{P}_H) = \alpha N^2, \quad \alpha > 0.$$

Proof. Because the arguments used are not the same, we will present the proof for Proposition 4.3 for two separate cases: (i) $\alpha = \frac{1}{N}$ and (ii) $\alpha \neq \frac{1}{N}$.

(i) Consider $\alpha = \frac{1}{N}$. From Lemma 4.2,

$$\begin{aligned} \det \mathbf{A}(\mathbf{P}_H) &= \prod_{i=1}^N \mu_i = \prod_{i=1}^{N-1} (1 - \lambda_i) \\ &= 1 - \sum_{i=1}^{N-1} \lambda_i + \sum_{\substack{i > j \\ i, j=1}}^{N-1} \lambda_i \lambda_j + \cdots + (-1)^{N-1} \prod_{i=1}^{N-1} \lambda_i \\ &= 1 - q_1(\lambda) + q_2(\lambda) + \cdots + (-1)^{N-1} q_{N-1}(\lambda), \end{aligned}$$

where q_i is the i th elementary symmetric polynomial in $\lambda_1, \dots, \lambda_{N-1}$, $i = 1, \dots, N-1$.

Let $Q_1(\lambda), \dots, Q_{N-1}(\lambda), Q_N(\lambda)$ be elementary symmetric polynomials in all of $\mathbf{P}_H$'s eigenvalues $\lambda_1, \dots, \lambda_N$. Define $q_0 = Q_0 = 1$. Hence, $q_1 = Q_1 - q_0$, $q_2 = Q_2 - q_1, \dots, q_{N-1} = Q_{N-1} - q_{N-2}$. Since the characteristic polynomial $c_x(\mathbf{P}_H) = 1 - x^N$ (e.g., see [3]), it implies that $Q_1 = Q_2 = \cdots = Q_{N-1} = 0$. Therefore, $q_1 = Q_1 - q_0 = -1$; $q_2 = Q_2 - q_1 = 1, \dots, q_{N-1} = Q_{N-1} - q_{N-2} = (-1)^{N-1}$. Hence,

$$\det \mathbf{A}(\mathbf{P}_H) = 1 + 1 + \cdots + 1 = N.$$

In particular,

$$\prod_{i=1}^{N-1} (1 - \lambda_i) = N.$$

(ii) Consider $\alpha \neq \frac{1}{N}$. As in the proof of Lemma 4.2, we can choose first $N-1$ eigenvectors $\mathbf{u}_1, \dots, \mathbf{u}_{N-1}$ of $\mathbf{A}_\alpha(\mathbf{P}_H)$ spanning $\ker(\mathbf{J})$ with eigenvalues $1 - \lambda_i$, $i = 1, \dots, N-1$, and the last eigenvector $\mathbf{u}_N = \mathbf{e}$ with corresponding eigenvalue αN . Therefore,

$$\det \mathbf{A}_\alpha(\mathbf{P}_H) = \alpha N \prod_{i=1}^{N-1} (1 - \lambda_i) = \alpha N^2. \quad \square$$

Remark 4.1. Since $\mathbf{P}\mathbf{e} = \mathbf{e}$, then $\mathbf{P}$ has an eigenvalue with value 1. Let $\lambda_1, \dots, \lambda_{N-1}$ denote the remaining eigenvalues of $\mathbf{P}$ and note that the corresponding eigenvectors may be chosen to be orthogonal to $\mathbf{e}$. Without loss of generality, we can assume that a Hamiltonian cycle is the cycle $1 \rightarrow 2 \rightarrow 3 \rightarrow \cdots \rightarrow N \rightarrow 1$. It follows that $\mathbf{I} - \mathbf{P}_H$ is a nonsingular upper-Hessenberg matrix plus a rank-one modification. Consequently, it can have at most one zero eigenvalue, and it follows that $\mathbf{P}_H$ has only one unit eigenvalue.³

PROPOSITION 4.4. *Let $\mathbf{P} \in \mathcal{D}$ be a Markov chain that contains two or more independent cycles. Then*

$$\det \mathbf{A}_\alpha(\mathbf{P}) = 0, \quad \alpha > 0.$$

Proof. Without loss of generality, assume that the first cycle is $1 \rightarrow 2 \rightarrow \cdots \rightarrow k_1 \rightarrow 1$ and the second is $k_1 + 1 \rightarrow k_1 + 2 \rightarrow \cdots \rightarrow k_1 + k_2 \rightarrow k_1 + 1$. Note that the size of the first cycle is k_1 and that of the second is k_2 .

³This argument avoids the appeal to the celebrated Perron–Frobenius theorem.

Consider the column vector

$$\mathbf{v} = [1, \dots, 1, -\frac{k_1}{k_2}, \dots, -\frac{k_1}{k_2}, 0, \dots, 0],$$

where the first k_1 elements are unity, the next k_2 elements are $-k_1/k_2$, and the last $N - k_1 - k_2$ elements are zeros. Then $\mathbf{I}\mathbf{v} = \mathbf{v}$, $\mathbf{P}\mathbf{v} = \mathbf{v}$, $\mathbf{J}\mathbf{v} = \mathbf{0}$, so $(\mathbf{I} - \mathbf{P} + \alpha\mathbf{J})\mathbf{v} = \mathbf{0}$, and hence

$$\det \mathbf{A}_\alpha(\mathbf{P}) = \det [\mathbf{I} - \mathbf{P} + \alpha\mathbf{J}] = 0. \quad \square$$

PROPOSITION 4.5. *Let $\mathbf{P} \in \mathcal{D}$ be a Markov chain that contains just a single cycle of length $k < N$. Then*

$$\det \mathbf{A}_\alpha(\mathbf{P}) = \alpha k N, \quad \alpha > 0.$$

Proof. Without loss of generality, assume that $\mathbf{P}$ contains the cycle $1 \rightarrow 2 \rightarrow \dots \rightarrow k \rightarrow 1$, and for $i \geq k+1$, $\mathbf{P}_{i,j} = 1$ for one value of j from the set $\{i+1\} \cup \{1, \dots, k\}$, and $\mathbf{P}_{i,l} = 0$ for all $l \neq j$.

Hence $\mathbf{P}$ is a matrix of the generic form

$$(4.1) \quad \left[\begin{array}{ccc|ccc} 0 & 1 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & 1 & \\ 1 & & & & & 0 \\ \hline & & & 1 & & \\ & & & & 0 & 0 \\ & & & & 0 & 1 \\ & & & & & \ddots \\ 1 & & & & & & 0 \end{array} \right].$$

We will present the proof for Proposition 4.5 for two separate cases: $\alpha = \frac{1}{N}$ and $\alpha \neq \frac{1}{N}$.

(i) Consider $\alpha = \frac{1}{N}$. Let $\mathbf{E} = \{\mathbf{e}_i\}$ be the standard basis of $\mathbb{R}^N$ where for each $\mathbf{e}_i$, the i th entry is unity and all other elements are zeros. Consider a new basis $\mathbf{V}$ of $\mathbb{R}^N$ that consists of $\mathbf{v}_1 = \mathbf{e}_2 - \mathbf{e}_1, \dots, \mathbf{v}_{k-1} = \mathbf{e}_k - \mathbf{e}_1$; $\mathbf{v}_k = \mathbf{e}$; $\mathbf{v}_{k+1} = \mathbf{e}_{k+1}, \dots, \mathbf{v}_N = \mathbf{e}_N$.

Let $\mathbf{B} = -\mathbf{P} + \frac{1}{N}\mathbf{J}$. First, we show that vectors $\mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_N$ form a basis of $\ker(\mathbf{B}^{N-k+1})$. It is clear that $\mathbf{v}_k = \mathbf{e} \in \ker \mathbf{B}$, and $\mathbf{J}\mathbf{u} = \mu\mathbf{e}$ for any $\mathbf{u} \in \mathbb{R}^N$ and some $\mu \in \mathbb{R}$, since $\text{rank } \mathbf{J} = 1$ and $\mathbf{J}\mathbf{e} = \mathbf{e}$. Therefore, by (4.1)

$$\mathbf{B}\mathbf{v}_{k+1} = -\mathbf{P}\mathbf{e}_{k+1} + \frac{1}{N}\mathbf{J}\mathbf{e}_{k+1} = \frac{1}{N}\mathbf{e},$$

and similarly for $k+2 \leq i \leq N$

$$\mathbf{B}\mathbf{v}_i = \gamma_i \mathbf{v}_{i-1} + \frac{1}{N}\mathbf{e} \quad \text{for } \gamma_i \in \{-1, 0\}.$$

If now we consider $\mathbf{B}^m \mathbf{v}_i$, then (modulo multiple of $\mathbf{e}$) the longest possible chain could be $\mathbf{v}_N \xrightarrow{\mathbf{B}} \mathbf{v}_{N-1} \xrightarrow{\mathbf{B}} \mathbf{v}_{N-2} \xrightarrow{\mathbf{B}} \dots \xrightarrow{\mathbf{B}} \mathbf{v}_{k+1} \xrightarrow{\mathbf{B}} \mathbf{0}$. At the last step any multiple of $\mathbf{e}$ also vanishes. Thus $\mathbf{B}$, restricted to the span of basic vectors $\{\mathbf{v}_k, \dots, \mathbf{v}_N\}$, is a nilpotent

linear operator $\tilde{\mathbf{B}}$, with $\tilde{\mathbf{B}}^{N-k+1} = 0$. Hence, 0 is an eigenvalue of $\mathbf{B}$ of multiplicity at least $N - k + 1$.

Since every $\mathbf{v}_i$, $i = k, \dots, N$, is obviously an eigenvector of $\mathbf{I}$ with eigenvalue 1,

$$\det \mathbf{A}(\mathbf{P}) = \prod_{i=1}^{k-1} (1 - \lambda_i),$$

where λ_i 's are the eigenvalues of the $(k-1) \times (k-1)$ leading principal minor of $\mathbf{B}$ in the $\mathbf{V}$ basis.

It remains to show that $\lambda_1, \dots, \lambda_{k-1}$ coincide with eigenvalues of a $\mathbf{P}_{H_k}$ of the simple cycle of size k with eigenvalue 1 excluded.

Indeed,

$$\mathbf{J}\mathbf{v}_i = 0 \quad \text{for } i = 1, \dots, k-1.$$

Hence,

$$\mathbf{B}\mathbf{v}_1 = -\mathbf{P}(\mathbf{e}_2 - \mathbf{e}_1) = -\mathbf{e}_1 + \mathbf{e}_k + \mathbf{v}_1^*,$$

where $\mathbf{v}_1^* \in \text{span}\{\mathbf{v}_k, \dots, \mathbf{v}_N\}$.

Analogously,

$$\mathbf{B}\mathbf{v}_i = -\mathbf{P}(\mathbf{e}_{i+1} - \mathbf{e}_1) = -\mathbf{e}_i + \mathbf{e}_k + \mathbf{v}_i^*,$$

where $\mathbf{v}_i^* \in \text{span}\{\mathbf{v}_k, \dots, \mathbf{v}_N\}$, $i = 2, \dots, k-1$.

Therefore, we observe that the $(k-1) \times (k-1)$ leading principal minor for $\mathbf{B}$, namely, $[\mathbf{B}]_{k-1}$, is the same as the principal leading minor $[\mathbf{M}]_{k-1}$ of $-\mathbf{P}_{H_k} + \frac{1}{k}\mathbf{J}_k$ if we choose bases $\mathbf{V}$ and $\{\mathbf{e}_1^{(k)} - \mathbf{e}_1, \dots, \mathbf{e}_{k-1}^{(k)} - \mathbf{e}_1, \mathbf{e}_k\}$, respectively, where $\mathbf{J}_k$ is the $k \times k$ matrix of units; $\mathbf{e}_j^{(k)}$ is the j th vector of the standard unit basis of $\mathbb{R}^k$ and $\mathbf{e}^{(k)} = (1, 1, \dots, 1)^T \in \mathbb{R}^k$.

Hence, for $r > 0$, principal minors of powers of $\mathbf{B}^r$ and $\mathbf{M}^r$ coincide and so the eigenvalues $\lambda_1, \dots, \lambda_{k-1}$ of $[\mathbf{B}]_{k-1}$ and of $\mathbf{P}_{H_k}$ (excluding eigenvalue 1) also coincide. For $\mathbf{P}_{H_k}$, according to Proposition 4.3, with $N = k$,

$$\prod_{i=1}^{k-1} (1 - \lambda_i) = k.$$

Therefore,

$$\det \mathbf{A}(\mathbf{P}) = \prod_{i=1}^{k-1} (1 - \lambda_i) = k.$$

(ii) Consider $\alpha \neq \frac{1}{N}$. Let $\mathbf{V}$ be a new basis in $\mathbb{R}^N$ that consists of $\mathbf{v}_1 = \mathbf{e}_2 - \mathbf{e}_1, \dots, \mathbf{v}_{k-1} = \mathbf{e}_k - \mathbf{e}_1$; $\mathbf{v}_k = \mathbf{e}$, $\mathbf{v}_{k+1} = \mathbf{e}_{k+1} + \mu_{k+1}\mathbf{e}, \dots, \mathbf{v}_N = \mathbf{e}_N + \mu_N\mathbf{e}$, where for $1 \leq j \leq N - k$,

$$(4.2) \quad \mu_{k+j} = \begin{cases} \frac{\alpha}{1 - \alpha N} & \text{if } \mathbf{P}_{k+j-1, k+j} = 0, \\ \frac{\alpha + \mu_{k+j-1}}{1 - \alpha N} & \text{if } \mathbf{P}_{k+j-1, k+j} = 1. \end{cases}$$

It follows that $\mathbf{B}_\alpha \mathbf{v}_k = (-1 + \alpha N) \mathbf{v}_k$ and that vectors $\mathbf{v}_{k+1}, \dots, \mathbf{v}_N$ form a basis of $\ker(\mathbf{B}_\alpha^{N-k})$, where $\mathbf{B}_\alpha = -\mathbf{P} + \alpha \mathbf{J}$.

Indeed,

$$\mathbf{B}_\alpha \mathbf{v}_k = (-\mathbf{P} + \alpha \mathbf{J}) \mathbf{e} = -\mathbf{e} + \alpha N \mathbf{e} = (-1 + \alpha N) \mathbf{v}_k.$$

By (4.1) and (4.2)

$$\mathbf{B}_\alpha \mathbf{v}_{k+1} = (-\mathbf{P} + \alpha \mathbf{J}) \left(\mathbf{e}_{k+1} + \frac{\alpha}{1 - \alpha N} \mathbf{e} \right) = \alpha \mathbf{e} - \frac{\alpha}{1 - \alpha N} \mathbf{e} + \frac{\alpha^2 N}{1 - \alpha N} \mathbf{e} = 0,$$

and for $k + 2 \leq i < N$

$$\mathbf{B}_\alpha \mathbf{v}_i = \gamma_i \mathbf{v}_{i-1} \quad \text{for } \gamma_i \in \{-1, 0\}.$$

Thus $\mathbf{B}_\alpha$, restricted to the span of basic vectors $\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_N\}$, is a nilpotent linear operator $\tilde{\mathbf{B}}_\alpha$, with $\tilde{\mathbf{B}}_\alpha^{N-k} = 0$. Hence, 0 is an eigenvalue of $\mathbf{B}_\alpha$ of multiplicity $N - k$. We note that $\mathbf{e}$ is an eigenvector of $\mathbf{A}_\alpha(\mathbf{P})$ with eigenvalue αN and $\mathbf{v}_1, \dots, \mathbf{v}_{k-1} \in \ker(\mathbf{J})$. Hence,

$$\det \mathbf{A}_\alpha(\mathbf{P}) = \alpha N \prod_{i=1}^{k-1} (1 - \lambda_i),$$

where λ_i 's are the eigenvalues of the $(k-1) \times (k-1)$ leading principal minor of $\mathbf{B}_\alpha$.

By an argument completely analogous to that used to establish part (i),

$$\prod_{i=1}^{k-1} (1 - \lambda_i) = k. \quad \square$$

Given a probabilistic solution we show how to obtain a deterministic solution that is at least as good.

PROPOSITION 4.6. *For any $\mathbf{P} \in \mathcal{F}$, there exists $\mathbf{D} \in \mathcal{D}$ such that $\det \mathbf{A}_\alpha(\mathbf{D}) \geq \det \mathbf{A}_\alpha(\mathbf{P})$.*

Proof. Let $\hat{\mathbf{P}} \in \{\arg \max_{\mathbf{P} \in \mathcal{F}} \det \mathbf{A}_\alpha(\mathbf{P})\}$ with minimal total number of randomizations in its rows compared to the other Markov chains in $\{\arg \max_{\mathbf{P} \in \mathcal{F}} \det \mathbf{A}_\alpha(\mathbf{P})\}$. If $\hat{\mathbf{P}} \in \mathcal{D}$, there is nothing to prove.

Suppose $\hat{\mathbf{P}} \notin \mathcal{D}$. Then $\hat{\mathbf{P}}$ contains a row, say row 1, where a randomization occurs. Let this row be $[\dots \ a \ \dots \ b \ \dots \ c \ \dots]$, $0 < a, b, \dots < 1$. Consider a Markov chain $\hat{\mathbf{D}}_\nu \in \mathcal{F}$ that coincides with $\hat{\mathbf{P}}$ in rows $2, \dots, N$ and whose first row is

$$[\dots \ \nu \ \dots \ \frac{(1-\nu)b}{1-a} \ \dots \ \frac{(1-\nu)c}{1-a} \ \dots], \quad 0 \leq \nu \leq 1.$$

Note that $\hat{\mathbf{D}}_{\nu=a} = \hat{\mathbf{P}}$. As $\det \mathbf{A}(\hat{\mathbf{D}}_\nu)$ is a linear function in ν , $\max_{\nu \in [0,1]} \det \mathbf{A}(\hat{\mathbf{D}}_\nu)$ occurs either at $\nu = 0$ or $\nu = 1$, or else $\det \mathbf{A}(\hat{\mathbf{D}}_\nu) = \det \mathbf{A}(\hat{\mathbf{P}})$ for all $\nu \in [0,1]$. In each of these situations, the Markov chain $\hat{\mathbf{D}} = \hat{\mathbf{D}}_{\nu=0}$ or $\hat{\mathbf{D}}_{\nu=1}$ has at least one more zero than $\hat{\mathbf{P}}$, and $\det \mathbf{A}(\hat{\mathbf{D}}) \geq \det \mathbf{A}(\hat{\mathbf{P}})$ so $\hat{\mathbf{D}} \in \{\arg \max_{\mathbf{P} \in \mathcal{F}} \det \mathbf{A}(\mathbf{P})\}$. Thus, by induction on the number of randomizations, the problem is reduced to $\mathbf{P} \in \mathcal{D}$. $\square$

Proof of Theorem 4.1. Propositions 4.3, 4.4, 4.5, and 4.6 complete the proof for Theorem 4.1, which gives an upper bound for $\det \mathbf{A}_\alpha(\mathbf{P})$, $\alpha > 0$. $\square$

Corollary 4.7 states the results specifically for $\alpha = \frac{1}{N}$, showing that for Hamiltonian graphs, this upper bound can only be achieved at Hamiltonian cycles and is larger than the upper bound for non-Hamiltonian graphs.

COROLLARY 4.7.

(i) For any graph G

$$\det_{\mathbf{P} \in \mathcal{F}} \mathbf{A}(\mathbf{P}) \leq N.$$

(ii) If G is a Hamiltonian graph, then

$$(a) \max_{\mathbf{P} \in \mathcal{F}} \det \mathbf{A}(\mathbf{P}) = N,$$

$$(b) \det \mathbf{A}(\mathbf{P}) = N \Leftrightarrow \mathbf{P} = \mathbf{P}_H.$$

(iii) If G is a non-Hamiltonian graph, then $\det_{\mathbf{P} \in \mathcal{F}} \mathbf{A}(\mathbf{P}) \leq N - 1$.

Proof. (i) and (ii.a) These statements follow directly from Theorem 4.1 for $\alpha = \frac{1}{N}$.

(ii.b) Suppose there exists $\mathbf{P} \in \mathcal{F}$ such that $\det \mathbf{A}(\mathbf{P}) = N$ and $\mathbf{P}$ is not a $\mathbf{P}_H$. Then there exists a reduction $\mathbf{P} \rightarrow \mathbf{P}_1 \rightarrow \cdots \rightarrow \mathbf{P}_r \rightarrow \mathbf{D}$ as in Proposition 4.6. Now, because $\det \mathbf{A}(\mathbf{D}) = N$, it follows from Propositions 4.3–4.5 that $\mathbf{D} = \mathbf{P}_H$. In the preceding sequence, at every step, $\det \mathbf{A}(\mathbf{P}_i) = N$, $i = 1, \dots, r$. Consider $\mathbf{P}_r$. Since it is the last step of the reduction, there is exactly one row where a randomization occurs. Let this row be $[\dots \ a \ \dots \ b \ \dots \ c \ \dots]$, $0 < a, b, \dots < 1$, and $\mathbf{P}_H$ has a unity in the place of a .

If we replace the above row with

$$[\dots \ \nu \ \dots \ \frac{(1-\nu)b}{1-a} \ \dots \ \frac{(1-\nu)c}{1-a} \ \dots], \quad 0 \leq \nu \leq 1,$$

$\det \mathbf{A}(\mathbf{P}_r) = N$ for all ν . Set $\nu = 0$. Then the above row becomes

$$[\dots \ 0 \ \dots \ \frac{b}{1-a} \ \dots \ \frac{c}{1-a} \ \dots],$$

and we can further apply the reduction procedure as in Proposition 4.6. This procedure will terminate with a deterministic Markov chain with the same value of $\det \mathbf{A}$, but it is not of the above Hamiltonian form $\mathbf{P}_H$. Since it is a deterministic Markov chain whose $\det \mathbf{A} = N$, it is also a Hamiltonian cycle. So we now have two Hamiltonian cycles sharing exactly $N - 1$ arcs, and this is a contradiction.

(iii) In a non-Hamiltonian graph, the length of the longest possible cycle is $N - 1$. It follows directly from Theorem 4.1 that for any $\mathbf{P} \in \mathcal{F}$, $\det \mathbf{A}(\mathbf{P}) \leq N - 1$. $\square$

To emphasize that the preceding results provide an interesting upper bound for a function of eigenvalues λ_i 's of a stochastic matrix $\mathbf{P} \in \mathcal{F}$, we restate Corollary 4.7 above. Of course, if the graph G is complete, it follows that this upper bound applies to any stochastic matrix.

COROLLARY 4.8. Consider any graph G , and let $\lambda_1, \dots, \lambda_{N-1}, \lambda_N = 1$ be the eigenvalues of $\mathbf{P}$. Then the following hold:

(i) For $\mathbf{P} \in \mathcal{F}$,

$$\prod_{i=1}^{N-1} (1 - \lambda_i) \leq N.$$

(ii) For a Hamiltonian graph G ,

$$\max_{\mathbf{P} \in \mathcal{F}} \prod_{i=1}^{N-1} (1 - \lambda_i) = \max_{\mathbf{P} \in \mathcal{DS}} \prod_{i=1}^{N-1} (1 - \lambda_i) = N.$$

(iii) For a non-Hamiltonian graph G ,

$$\max_{\mathbf{P} \in \mathcal{F}} \prod_{i=1}^{N-1} (1 - \lambda_i) \leq N - 1.$$

Remark 4.2. In [1] and [2], Remark 2.1 was exploited by perturbing probability transition matrices in $\mathcal{F}$ by a small parameter $\varepsilon > 0$ to ensure that all Markov chains are irreducible. Then it was shown that for ε sufficiently small, any Hamiltonian Markov chain $\mathbf{P}_H$ is a global minimizer of the following optimization problem, namely,

$$g_{11}^\varepsilon(\mathbf{P}_H) = \min_{\mathbf{P} \in \mathcal{DS}} [g_{11}^\varepsilon(\mathbf{P})],$$

where $g_{11}^\varepsilon(\mathbf{P})$ is the $(1, 1)$ -element of the fundamental matrix $\mathbf{G}(\mathbf{P})$.

The observation that led to the present paper is that for $\mathbf{P} \in \mathcal{DS}$ the objective function $g_{11}^\varepsilon(\mathbf{P})$ of the above minimization problem is, of course, a ratio of two determinants,

$$(4.3) \quad g_{11}^\varepsilon(\mathbf{P}) = \frac{\text{cof}([\mathbf{A}_\varepsilon(\mathbf{P})]_{11})}{\det \mathbf{A}_\varepsilon(\mathbf{P})},$$

where $\mathbf{A}_\varepsilon(\mathbf{P}) = (\mathbf{I} - \mathbf{P} + \frac{1}{N}\mathbf{J}) = [\mathbf{G}_\varepsilon(\mathbf{P})]^{-1}$, and where $\text{cof}([\mathbf{A}_\varepsilon(\mathbf{P})]_{11})$ is the cofactor of the $(1, 1)$ -element of $\mathbf{A}_\varepsilon(\mathbf{P})$ and $\det \mathbf{A}_\varepsilon(\mathbf{P})$ is its determinant.

Apart from no longer having the need for the ε perturbation, the global minimum of the new problem has a better separation from the next best minimum.

5. Conclusions. We note that the matrix

$$\mathbf{A}_\alpha(\mathbf{P}) = \mathbf{I} - \mathbf{P} + \alpha\mathbf{J}, \quad \alpha > 0,$$

is merely a “rank-one modification” of the generator $\mathbf{P} - \mathbf{I}$ of the Markov chain $\mathbf{P}$. Of course, the case $\alpha = 0$ was considered before but did not yield useful results as $\det[\mathbf{I} - \mathbf{P}] \equiv 0$ for all $\mathbf{P} \in \mathcal{F}$.

Clearly, the preceding results suggest an interesting optimization problem,

$$\max_{\mathbf{P} \in \mathcal{DS}} \det \mathbf{A}_\alpha(\mathbf{P}),$$

the solution of which solves the original Hamiltonian cycle problem. Of course, in the above, the smaller feasible region of doubly stochastic Markov chains $\mathcal{DS}$ was taken instead of $\mathcal{F}$ as all Hamiltonian cycles (if any) induce probability transition matrices in Γ that are doubly stochastic. The numerical tractability of the above nonlinear programming problem is currently being investigated.

Acknowledgment. The authors would like to thank N. Litvak, M. Haythorpe, and A. Eshragh for numerous useful discussions.

REFERENCES

- [1] V. S. BORKAR, V. EJOV, AND J. A. FILAR, *Directed graphs, Hamiltonicity and doubly stochastic matrices*, Random Structures Algorithms, 25 (2004), pp. 376–395.
- [2] V. S. BORKAR, V. EJOV, AND J. A. FILAR, *On the Hamiltonicity gap and doubly stochastic matrices*, Random Structures Algorithms, to appear.
- [3] V. EJOV, J. A. FILAR, S. K. LUCAS, AND J. L. NELSON, *Solving the Hamiltonian cycle problem using symbolic determinants*, Taiwanese J. Math., 10 (2006), pp. 327–338.

- [4] E. FEINBERG, *Constrained discounted Markov decision process with Hamiltonian cycles*, Math. Oper. Res., 25 (2000), pp. 130–140.
- [5] J. FILAR, *Controlled Markov chains, graphs and Hamiltonicity*, Foundations and Trends in Stochastic Systems, 1 (2007), pp. 1–85.
- [6] J. A. FILAR AND D. KRASS, *Hamiltonian cycles and Markov chains*, Math. Oper. Res., 19 (1994), pp. 223–237.
- [7] R. J. WILSON, *Introduction to Graph Theory*, Longman, New York, 1996.

ISOMETRIC EMBEDDINGS OF SUBDIVIDED COMPLETE GRAPHS IN THE HYPERCUBE*

LAURENT BEAUDOU[†], SYLVAIN GRAVIER[†], AND KAHINA MESLEM[‡]

Abstract. Isometric subgraphs of hypercubes are known as partial cubes. These graphs have first been investigated by Graham and Pollak [*Bell System Tech. J.*, 50 (1971), pp. 2495–2519] and Djoković [*J. Combinatorial Theory Ser. B*, 14 (1973), pp. 263–267]. Several papers followed with various characterizations of partial cubes. In this paper, it is proven that a subdivision of a complete graph of order n ($n \geq 4$) is a partial cube if and only if it is isomorphic to $S(K_n)$ or there exist $n - 1$ nonsubdivided edges of K_n adjacent to a common vertex in the subdivision and the other edges of K_n are subdivided an odd number of times. As a corollary, we build partial cubes with arbitrary graph as a minor.

Key words. isometric embedding, partial cube, subdivision of a graph

AMS subject classifications. 05C75, 05C12

DOI. 10.1137/070681909

Introduction. Isometric subgraphs of hypercubes are called *partial cubes*. They were first investigated by Graham and Pollak [11] and Djoković [7] in the early 1970s. Later, several algorithmic characterizations were shown, most of them using a relation defined on the set of edges (the Θ relation introduced by Djoković [7] and Winkler [17]). This relation easily led to a recognition algorithm for partial cubes running in $O(mn)$ (m is the number of edges and n the number of vertices). This complexity was then improved for specific classes of graphs by Brešar, Imrich, and Klavžar [2], and recently the general complexity was pulled down to $O(n^2)$ by Eppstein [8].

The search for structural characterizations was motivated by the work of Brešar and Klavžar [3] using subdivided wheels to answer in negative a question of Chepoi and Tardif [6] whether partial cubes are precisely bipartite graphs with convex intervals. Then Gravier, Klavžar, and Mollard [12] characterized the partial cubes that are subdivided wheels (see Theorem 1.1). Until now, only low-density graphs have been considered, so that we decided to study the other side of the problem, namely, high-density graphs, and their best known representatives: complete graphs. This is the starting point of this paper. For more information on partial cubes, see the book [13]. We also mention recent relevant works in [9, 10, 15].

In this paper we prove a conjecture due to Aïder, Gravier, and Meslem [1] which characterizes all the subdivisions of a complete graph K_n that are a partial cube. Either all its edges are subdivided exactly once, or one of the vertices has no incident subdivided edge and all other edges are subdivided an odd number of times. First, using the characterization of partial cubes through expansions over proper covers, it is shown that these graphs are partial cubes. Then, through a case study, it is shown

*Received by the editors February 5, 2007; accepted for publication (in revised form) April 2, 2008; published electronically July 16, 2008.

<http://www.siam.org/journals/sidma/22-3/68190.html>

[†]Institut Fourier, ERTé Maths à Modeler, UMR 5582 CNRS/Université Joseph Fourier, 100 rue des maths, BP 74, 38402 St. Martin d'Hères, France (laurent.beaudou@ujf-grenoble.fr, sylvain.gravier@ujf-grenoble.fr).

[‡]Laboratoire LAID3, Faculté des Mathématiques, U.S.T.H.B. El Alia Bab-Ezzouar 16 111, Algiers, Algeria (kahina_m2000@yahoo.fr).

that there is always a vertex of the complete graph that can be isometrically removed. Thus the result is proven by induction.

1. Preliminary definitions and main result. We consider only finite, simple, loopless, connected, and undirected graphs $G = (V, E)$ where V is the set of vertices and E is the set of edges.

For a graph G , the *distance* $d_G(u, v)$ between vertices u and v is defined as the number of edges on a shortest uv -path (or uv -geodesic). A subgraph H of G is called *isometric* in G if $d_G(u, v) = d_H(u, v)$ for all distinct vertices u and v in $V(H)$.

The vertex set of the n -cube (or *hypercube*) Q_n consists of n -tuples $(b_1, b_2, \dots, b_n)$ with $b_i \in \{0, 1\}$. Two vertices are adjacent in Q_n if the corresponding tuples differ precisely in one place. Q_n is a bipartite graph. A *partial cube* is an isometric subgraph of Q_n . A graph G admits an *isometric embedding* in the hypercube if it is isomorphic to a partial cube and by extension is called partial cube.

By replacing edges of a graph G by paths, we obtain a *subdivided graph* of G . The *subdivision* of a graph G , denoted $S(G)$, is the subdivided graph obtained from G by adding exactly one vertex to each edge of G (see Figure 1.1(b)). A *principal vertex* is a vertex of a subdivided graph of G that was not added during the subdivision (it was already in G). Given S a set of principal vertices of G , we denote by $\langle S \rangle$ the subgraph induced in G by the vertices of S and all the subdivision vertices added between each pair of vertices in S . A vertex v in a subdivided graph of G which is adjacent to all the principal vertices that are its neighbors in G is called *universal* in the subdivided graph of G . It means that all the edges of G incident to v have not been subdivided.

W_k denotes the k -wheel, that is, the graph obtained as a combination of the one vertex graph K_1 and the cycle C_k . We denote the central vertex of W_k by u and the remaining vertices by $w_1, \dots, w_k$. $W_k(m_1, \dots, m_k; n_1, \dots, n_k)$ is the graph obtained by subdividing edges of W_k , where m_i is the number of vertices added on the edge $w_i w_{i+1}$, and n_i the number of vertices added on the inner edge uw_i (see Figure 1.1(a)). Gravier, Klavžar, and Mollard proved the following theorem.

THEOREM 1.1 (see [12]). *Let $k \geq 3$. Then a subdivided k -wheel W is a partial cube if and only if W is isomorphic to $W_k(m_1, \dots, m_k; 0, \dots, 0)$, where m_i is odd for $i = 1, \dots, k$, or $W = W_3(1, 1, 1; 1, 1, 1)$.*

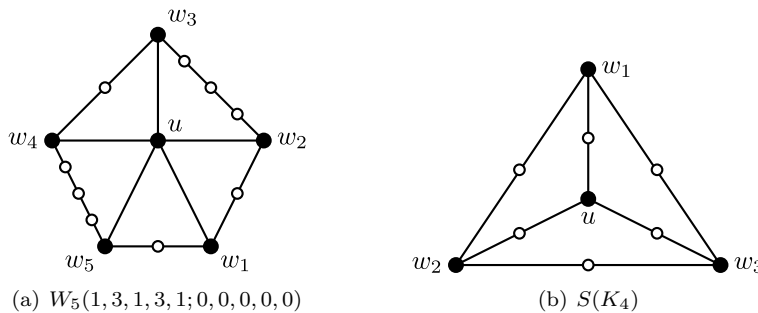


FIG. 1.1.

Let G' be a connected graph. A *proper cover* G'_1, G'_2 consists of two isometric subgraphs G'_1, G'_2 of G' such that $G' = G'_1 \cup G'_2$ and $G'_0 = G'_1 \cap G'_2$ is a nonempty subgraph, called the *intersection* of the cover. Additionally, there are no edges between $G'_1 \setminus G'_2$ and $G'_2 \setminus G'_1$. The *expansion* of G' with respect to the proper cover G'_1, G'_2 is the graph G constructed as follows. Let G_i be an isomorphic copy of G'_i for $i = 1, 2$,

and, for any vertex u' in G'_0 , let u_i be the corresponding vertex in G_i for $i = 1, 2$. Then G is obtained from the disjoint union $G_1 \cup G_2$, where for each u' in G'_0 the vertices u_1 and u_2 are joined by an edge. See Figure 1.2 for a visual explanation.

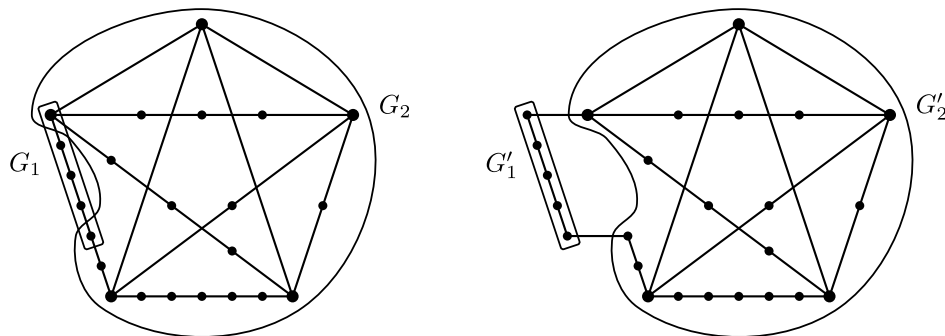


FIG. 1.2. Expansion through a proper cover.

Chepoi [4] proved that a graph is a partial cube if and only if it can be obtained from K_1 by a sequence of such expansions.

Our aim is to demonstrate the following theorem conjectured in [1].

THEOREM 1.2. *Let G be a subdivided graph of a complete graph K_n ($n \geq 4$). G is a partial cube if and only if G is isomorphic to $S(K_n)$ or G contains a universal vertex u and the number of added vertices to each edge not incident to u in K_n is odd.*

2. Proof of the main result. In this section, we provide the validity of Theorem 1.2. We will use some specific terminology. Let G be a subdivided K_n . We have to note that in our proof, we use only principal vertices. The proof is mostly based on paths joining principal vertices in G . The subdivided edge between x and y , principal vertices of G , is denoted $P(x, y)$. An edge joining two principal vertices in G , x and y , is said to be *plain* (it has not been subdivided). For u a principal vertex of G , we denote by $G \setminus u$ the subdivided graph of K_{n-1} induced by $V(G) \setminus U$, where U consists of vertex u and all nonprincipal vertices lying on subdivided edges between u and other principal vertices. For each x and y principal vertices in G , we say that x sees y , and we denote $x \rightarrow y$ if the subdivided edge joining these vertices in G is geodesic.

Notice that in our figures, a line (resp., a dotted line) represents a geodesic (resp., nonisometric) path between two principal vertices in G . A bold line represents a plain edge. A dashed line represents a subdivided edge with undetermined status.

We first recall several former results.

PROPOSITION 2.1 (see [14]). *For any $n \geq 1$, $S(K_n)$ is a partial cube.*

LEMMA 2.2 (see [1]). *Let G be a subdivided graph of K_n ($n \geq 4$), where each edge in K_n is an isometric path in G . G is a partial cube if and only if G contains a universal vertex and the other edges of K_n have exactly one added vertex or G is isomorphic to $S(K_n)$.*

2.1. Proof of the sufficient condition. With these two results, we are now able to complete the proof of the sufficient condition.

Let $n \geq 4$ and let G be a subdivided graph of K_n isomorphic to $S(K_n)$, or with a universal vertex and all its other edges subdivided an odd number of times. Let us show it is a partial cube.

According to Proposition 2.1, if G is isomorphic to $S(K_n)$, then it is a partial cube. Thus, we just have to prove that if G has a universal vertex u and an odd number of added vertices to each edge of K_n not incident to u , then it is a partial cube.

Let us suppose that G is such a graph. Considering only principal vertices, we can enumerate them from w_1 to w_n and assume that w_n is the universal vertex. For $1 \leq i < j \leq n-1$, let us denote by m_{ij} the number of subdivision vertices added between vertices w_i and w_j . The m_{ij} 's are odd and define all of the graph G . We show that G is a partial cube by induction on $\sum_{1 \leq i < j \leq n-1} m_{ij}$. If $m_{ij} = 1$ for all $1 \leq i < j \leq n-1$, then by Lemma 2.2, G is a partial cube. We can thus assume that $m_{12} \geq 3$. By the induction hypothesis the graph G' which has the same values for the m_{ij} 's but $m'_{12} = m_{12} - 2$ is a partial cube. Let us denote by $w'_1, w'_{1,1}, w'_{1,2}, \dots, w'_{1,m'_{12}}, w'_2$ the vertices (not only principal ones) along the subdivided edge $w'_1 w'_2$ (if $m_{12} = 3$, set $w'_{1,2} = w'_2$). Then G can be obtained from G' by an expansion with respect to the proper cover $G'_1 = \{w'_1, w'_{1,1}, \dots, w'_{1,(m'_{12}+1)/2+1}\}$ and $G'_2 = G' \setminus G'_1 \cup \{w'_1, w'_{1,(m'_{12}+1)/2+1}\}$. Both subgraphs are isometric in G and there is no edge from $G'_1 \setminus G'_2$ to $G'_2 \setminus G'_1$. Hence the sufficient condition is proven.

See Figure 1.2 for an example when $n = 5$ and $m_{12} = 7$.

Remark 2.3. Each expansion adds one dimension to the hypercube of the embedding. As the graph with m_{ij} 's equal to 1 is embedded in Q_{n-1} (see [1]), the graph G can be embedded in the hypercube of dimension $n-1 + (\sum_{1 \leq i < j \leq n-1} (m_{ij} - 1))/2$.

2.2. Proof of the necessary condition. We will proceed by induction. Subdivided graphs of K_4 and K_5 have already been studied in other articles: the theorem holds for K_4 since it is contained in Theorem 1.1 (a 3-wheel is isomorphic to K_4); and for K_5 , we have the following proposition.

PROPOSITION 2.4 (see [1]). *Let G be a subdivided graph of K_5 . G is a partial cube if and only if G is isomorphic to $S(K_5)$ or G contains a universal vertex u and the number of the added vertices to each edge not incident to u in K_5 is odd.*

2.2.1. Useful results. We now demonstrate results to be used for the later proof.

PROPOSITION 2.5. *A geodesic between two principal vertices in G goes through at most 3 principal vertices.*

Proof. For a contradiction, assume that a geodesic path between x and y principal vertices of G goes through at least four principal vertices (x and y included). We now consider the first four principal vertices of this xy -geodesic: x, x_1, x_2 , and x_3 . They induce in G an isometric subgraph (for each pair of vertices in this subgraph, a geodesic is contained in the subgraph) $\langle x, x_1, x_2, x_3 \rangle$ which is a subdivided graph of K_4 (see Figure 2.1(a)). Thus, this subgraph $\langle x, x_1, x_2, x_3 \rangle$ is a partial cube as an isometric subgraph of a partial cube. Then, by Theorem 1.1, either $\langle x, x_1, x_2, x_3 \rangle$ is isomorphic to $S(K_4)$ (impossible since the edge between x and x_3 is subdivided strictly more than once, or the considered path would not be a geodesic) or there is a universal vertex in it. But $P(x, x_2)$ cannot be a plain edge because $d_G(x, x_2) = d_G(x, x_1) + d_G(x_1, x_2) \geq 2$. In addition, $P(x_1, x_3)$ cannot be a plain edge for the same reason, so that there is no universal vertex in this subgraph which is in contradiction with Theorem 1.1. $\square$

From now on, for any x, y , and v principal vertices of G , we will denote $x \xrightarrow{v} y$ if the path along x, v , and y is an xy -geodesic. Therefore, given x and y principal vertices of G , either $x \rightarrow y$ or there exists v principal vertex of G such that $x \xrightarrow{v} y$.

LEMMA 2.6. *If $P(x, y)$ is plain and $y \rightarrow v$, then an xv -geodesic is contained in $\langle x, y, v \rangle$.*

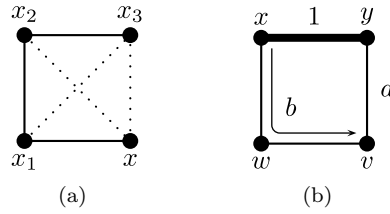


FIG. 2.1.

Proof. For a contradiction let us suppose that $x \xrightarrow{w} v$ for some w principal vertex distinct from x , y , and v , with $x \not\xrightarrow{y} v$ and $x \not\xrightarrow{y} v$. We denote by a the distance between y and v , and by b the distance between x and v (see Figure 2.1(b)). Then we must have $b < a + 1$ (else, $x \xrightarrow{y} v$). Moreover, if $b \leq a - 1$, the sequence y, x, w, v would be a yv -geodesic going through 4 principal vertices, which is impossible by Proposition 2.5. Finally, we have $b = a$ and thus an odd cycle of length $2a + 1$, which is also impossible (partial cubes are bipartite). $\square$

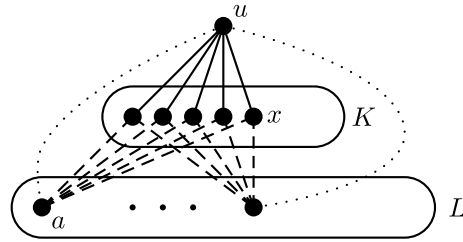


FIG. 2.2.

For any principal vertex u of G , we define two sets K_u and L_u partitioning the remaining principal vertices of G : K_u contains principal vertices x such that $u \rightarrow x$ and L_u contains other principal vertices a (there exists $x \in K_u$ such that $u \xrightarrow{x} a$). We will denote them by K and L when there is no doubt (see Figure 2.2).

DEFINITION 2.7. We define types related to u for the vertices of K_u as follows (see also Figure 2.3):

- x has type $\mathcal{L}$ if there exists y in K_u such that $u \xrightarrow{x} y$.
- x has type $\mathcal{I}$ if there exists y in L_u such that $u \xrightarrow{x} y$.
- x has type $\mathcal{C}$ if there exist y, z in K_u such that $y \xrightarrow{x} z$.
- x has type Λ if there exist y, z in L_u such that $y \xrightarrow{x} z$.
- x has type $\mathcal{R}$ if there exist y in K_u and z in L_u such that $y \xrightarrow{x} z$.

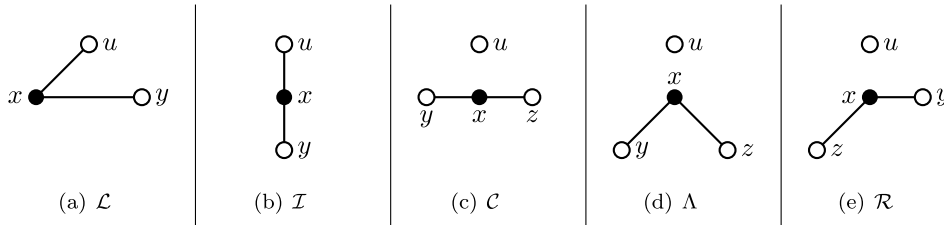


FIG. 2.3.

Remark 2.8. If there is no principal vertex v in G such that $G \setminus v$ is isometric, then for any principal vertex u in G , every vertex of K_u has at least one of these types. Moreover, there is no principal vertex x with only type $\mathcal{L}$ ($u \xrightarrow{x} y$) because $u \rightarrow y$, and $G \setminus x$ would be isometric.

LEMMA 2.9. *If there is no vertex v in G such that $G \setminus v$ is isometric, then for any u in G , no vertex of K_u has type $\mathcal{C}$.*

Proof. Let u be a principal vertex of G . Let us suppose that there exists a vertex x in K with type $\mathcal{C}$: $y \xrightarrow{x} z$, where y and z are in K . The subdivided $K_4 - \langle u, x, y, z \rangle$ —is isometric in G , and thus a partial cube. It is not isomorphic to $S(K_4)$ since $y \xrightarrow{x} z$, so there must be a universal vertex. Neither z nor y is a universal vertex in this subdivided K_4 since $d_G(y, z) \geq 2$. The vertex u is not universal either, otherwise $y \xrightarrow{u} z$. Consequently x is the universal vertex in $\langle u, x, y, z \rangle$, and $P(u, z)$ and $P(u, y)$ are subdivided exactly once.

CLAIM 2.10. *For each vertex t in K distinct from x , y , and z , if y sees t , then $P(x, t)$ is plain.*

Let $t \in K$ be a vertex seen by y , and let us denote $a = d_G(y, t)$. On the one hand, $a - 2 < d_G(u, t) < a + 2$, otherwise a ut -geodesic or a yt -geodesic would go through four principal vertices (forbidden by Proposition 2.5). On the other hand, $d_G(u, t) \notin \{a - 1, a + 1\}$, otherwise $\langle u, y, t, x \rangle$ would contain an odd cycle, and G would not be a partial cube. Then $d_G(u, t) = a$ (see Figure 2.4).

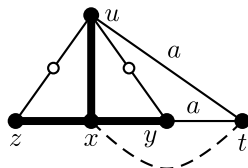


FIG. 2.4.

Consider now an xt -geodesic. It does not go through y , otherwise y would have type $\mathcal{C}$ and thus $P(u, y)$ would be plain, and $\langle y, u, x \rangle$ would be an odd cycle (more precisely a triangle). Consequently, $d_G(x, t) \leq a$. Furthermore, an xt -geodesic has a length at least $a - 1$, otherwise $y \not\rightarrow t$ (a shorter path would exist through x). $d_G(x, t) \neq a$, otherwise $\langle x, t, y \rangle$ would be an odd cycle in G , which is a contradiction. Consequently, $d_G(x, t) = a - 1$. Moreover, we can assume that $x \rightarrow t$, otherwise there would be a yt -geodesic going through four principal vertices of G , which is forbidden by Proposition 2.5. $\langle x, y, t, u \rangle$ is an isometric subdivision of K_4 , which is not isomorphic to $S(K_4)$ ($P(x, y)$ is plain). There must be a universal vertex which can only be x (else, it would lead to triangles). Thus $P(x, t)$ is plain and $a = 2$. Claim 2.10 is proven.

Now, let us consider a type of vertex y in K .

- The vertex y is not of type $\mathcal{C}$, otherwise we would have $P(u, y)$ plain that would lead to a triangle which is forbidden.
- If the vertex y has type $\mathcal{I}$, then there exists a vertex v in L such that $u \xrightarrow{y} v$. An xv -geodesic is contained in $\langle x, y, v \rangle$ by Lemma 2.6. Then the subdivided K_4 , $\langle u, y, x, v \rangle$, is isometric in G and different from $S(K_4)$. The vertices u and v are not universal since $P(u, v)$ is not isometric. The same holds for x , otherwise we would not have $u \xrightarrow{y} v$. y is not universal since $P(u, y)$ is subdivided once; this is a contradiction. Consequently, y is not a vertex of type $\mathcal{I}$.

• Let us show that the vertex y is not of type Λ . If there exist two vertices v and w in L , such that $v \xrightarrow{y} w$, then an xv -geodesic (resp., xw -geodesic) is contained in $\langle x, v, y \rangle$ (resp., $\langle x, w, y \rangle$) thanks to Lemma 2.6. Consequently, the subdivided K_4 , $\langle x, y, v, w \rangle$, is isometric in G and is not isomorphic to $S(K_4)$ (because $v \xrightarrow{y} w$). The vertex v (resp., w) is not universal, otherwise (x, y, v) (resp., (x, y, w)) is a triangle in G . The vertex x is not universal, otherwise $P(v, y)$ and $P(w, y)$ must be subdivided and $v \xrightarrow{y} w$. If y is a universal vertex in $\langle x, y, v, w \rangle$, then the distance between u and v is less than or equal to 3. If it is 3, we would have a uv -geodesic going through u, x, y, v principal vertices, which is forbidden. If it is 2, an odd cycle will arise, which is also forbidden. If it is 1, v would not be in L , which contradicts our hypothesis. Finally, we conclude that y has not the type Λ .

• The vertex y is not of type $\mathcal{R}$. If y has type $\mathcal{R}$, then either $t \xrightarrow{y} v$ or $x \xrightarrow{y} v$, where t is a vertex in K different from x, y , and z (it would give rise to a zv -geodesic going through four principal vertices z, x, y , and v), and v is a vertex in L . In the first case, using Claim 2.10 a tv -geodesic would go through t, x, y, v principal vertices (see Figure 2.5(a)). This is a contradiction. In the second case, we cannot have $u \xrightarrow{y} v$, otherwise y would have type $\mathcal{I}$ treated above. Neither can we have $u \xrightarrow{x} v$ (else, the sequence u, x, y, v would be a geodesic going through four principal vertices). Then there exists t in K such that $u \xrightarrow{t} v$ (see Figure 2.5(b)). Let a be the distance between v and y . We denote by b the distance between u and v . Then $b \leq a + 1$, else we would have $u \xrightarrow{y} v$, which contradicts precedent conclusions (y would have type $\mathcal{I}$). Moreover, $a - 1 \leq b$, else we would have a yv -geodesic going through y, u, t, v principal vertices, which is forbidden. Finally, to avoid an odd cycle, we must have $b = a$. But $d(x, v) = a + 1$, and thus an xv -geodesic goes through x, u, t, v principal vertices. This is forbidden by Proposition 2.5. Therefore, y cannot have type $\mathcal{R}$.

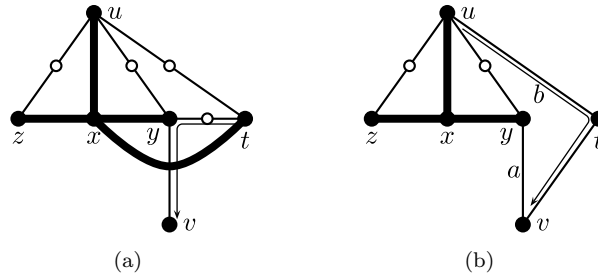


FIG. 2.5.

We conclude that the vertex y has none of the types $\mathcal{C}, \mathcal{I}, \Lambda, \mathcal{R}$. This contradicts Remark 2.8. Consequently, there is no vertex with type $\mathcal{C}$. $\square$

2.2.2. Proof of the induction step. We can now suppose that the partial cube G is a subdivision of K_n with $n \geq 6$. Moreover, by Lemma 2.2 we can assume that G is not isomorphic to $S(K_n)$ and that the theorem holds for any $m < n$.

Case 1. There exists u a principal vertex in G such that $G \setminus u$ is isometric. As $G \setminus u$ is isometric in G which is a partial cube, $G \setminus u$ is also a partial cube. With the induction hypothesis, there exists $x \in G \setminus u$ a universal vertex in $G \setminus u$, or $G \setminus u$ is isomorphic to $S(K_{n-1})$.

We first consider the case when $G \setminus u$ is isomorphic to $S(K_{n-1})$.

• If $u \rightarrow v$ for all $v \in G \setminus u$: As G is not isomorphic to $S(K_n)$ there exists $v \in G \setminus u$ such that $P(u, v)$ is not subdivided exactly once. Let y, z be vertices in $G \setminus \{u, v\}$. Then $\langle u, v, y, z \rangle$ is clearly isometric in G (there are no shorter paths outside from $\langle u, v, y, z \rangle$). Therefore, it is a partial cube, and a subdivided K_4 not isomorphic to $S(K_4)$ since $P(u, v)$ is not subdivided exactly once. Then, by Theorem 1.1, u is the only possible universal vertex in $\langle u, v, y, z \rangle$, and $P(u, v)$, $P(u, y)$, and $P(u, z)$ are plain edges. This is true for any pair of vertices y, z . Therefore, u is a universal vertex in G (see Figure 2.6(a)). Hence, the theorem is proven in this case.

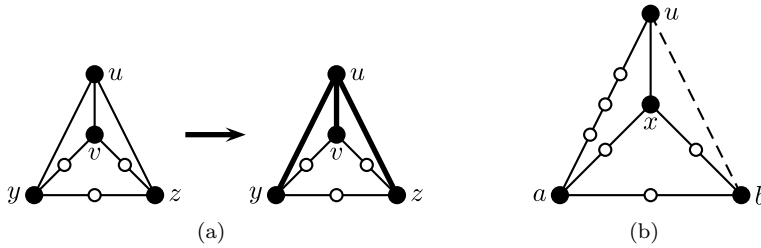


FIG. 2.6.

• If there exists a principal vertex $a \in G$ such that $u \not\rightarrow a$: Let x be a principal vertex of $G \setminus u$ that minimizes $d_G(u, x)$. There exists a principal vertex b such that $b \in G \setminus \{u, a, x\}$ ($n \geq 6$). We can assume that $u \xrightarrow{x} a$, and either $u \rightarrow b$ or $u \xrightarrow{x} b$. These four vertices induce an isometric subgraph of G which is a subdivided graph of K_4 . But $P(u, a)$ is subdivided at least three times (or it would be isometric). Therefore, it neither is isomorphic to $S(K_4)$ nor has a universal vertex (see Figure 2.6(b)). Hence, there is no such case.

We then consider the case when $G \setminus u$ has a universal vertex x .

CLAIM 2.11. x is universal in G .

For this, we consider y and z two principal vertices of G closest to u (distinct from x). Therefore, a uy -geodesic and a uz -geodesic are contained in $\langle x, u, y, z \rangle$. Moreover, as $P(x, y)$ and $P(x, z)$ are plain, a yz -geodesic is contained in $\langle x, y, z \rangle$. In addition, we have either $u \rightarrow x$, or $u \xrightarrow{y} x$. Finally $\langle u, x, y, z \rangle$ is an isometric subdivision of K_4 which is not $S(K_4)$ (there is a plain edge). The only possible universal vertex is x , otherwise there would be odd cycles (more precisely triangles), so that $P(u, x)$ is plain and x is universal in G .

Case 2. For every principal vertex u in G , $G \setminus u$ is not isometric. In this case, we know that there exist three principal vertices u, v_K, v_L such that $u \xrightarrow{v_K} v_L$. We will now denote K for K_u and L for L_u . These are clearly two nonempty sets since $v_K \in K$ and $v_L \in L$.

By Lemma 2.9 there is no vertex with type $\mathcal{C}$ in K .

LEMMA 2.12. There are no vertices x and y in K such that $P(x, u)$ and $P(x, y)$ are both plain.

Proof. Suppose on the contrary that it is possible. Thus, we consider a type of y , which cannot be $\mathcal{C}$ by Lemma 2.9.

y has type $\mathcal{I}$: Then there exists $b \in L$ such that $u \xrightarrow{y} b$. The distance between u and b is then $d(y, b) + 2$. Then the sequence u, x, y, b would also be a ub -geodesic going through four principal vertices. This is forbidden by Proposition 2.5. Thus y cannot be of type $\mathcal{I}$.

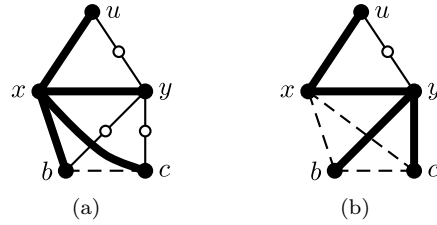


FIG. 2.7.

y has type Λ : Then there exist b and c in L such that $b \xrightarrow{y} c$. By Lemma 2.6, we know that an xb -geodesic and an xc -geodesic are contained in $\langle x, y, b, c \rangle$. Then the subdivision of K_4 induced by $\langle x, y, b, c \rangle$ is isometric and distinct from $S(K_4)$ so it has a universal vertex which can either be x or y (else it would lead to a triangle). If x is universal, then $P(b, y)$ and $P(c, y)$ are both subdivided once because they are geodesics (they cannot be plain to avoid triangles; see Figure 2.7(a)). This contradicts $b \xrightarrow{y} c$; there is a shorter path through x . If y is universal (see Figure 2.7(b)), let us suppose $d_G(u, c) = 3$; then u, x, y, c is a uc -geodesic going through four principal vertices, which is impossible by Proposition 2.5; therefore, $d_G(u, c) = 2$ (it cannot be equal to 1 because $c \in L$). The geodesic of length 2 and the walk of length 3 between u and c induce an odd cycle of length 5. We conclude that there is no y with type Λ .

y has type $\mathcal{R}$: Then there exists $b \in L$ and $t \in K$ such that $t \xrightarrow{y} b$.

- If $t = x$, we denote by l the distance between y and b . Then $d_G(x, b) = l + 1$. A ub -geodesic cannot go through x (it would go through four principal vertices, which is forbidden by Proposition 2.5) or y (it would have type $\mathcal{I}$). Therefore, there exists a vertex z in K such that $u \xrightarrow{z} b$; we denote by p the distance between u and b . Then $p > l$, else the sequence x, u, z, b would be a geodesic going through four principal vertices. In addition, $p \leq l + 1$, else we would have $u \xrightarrow{y} b$ and y would have type $\mathcal{I}$. Therefore $p = l + 1$, which is also impossible because it induces an odd cycle going through b, y, u, z of length $2l + 3$ (see Figure 2.8).

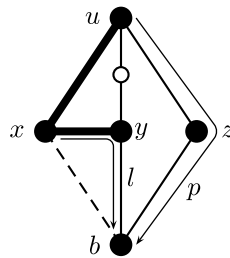


FIG. 2.8.

- If $t \neq x$, let l be the distance between y and t . We then consider $d_G(x, t)$, denoted by p (see Figure 2.9). On the one hand, $p \leq l$, else we would have a geodesic path of length $l + 1$ between x and t , going through y , and thus it would have type $\mathcal{C}$, which is forbidden by Lemma 2.9. Furthermore, p is not equal to l since there would be an odd cycle (x, t, y) of length $2l + 1$. On the other hand we know that $p > l - 2$ since $t \rightarrow y$. Hence, $p = l - 1$, which implies that $l \geq 2$ (and hence $P(y, t)$ is not plain). From this we can also conclude that $x \rightarrow t$, because if we had $x \xrightarrow{y} t$,

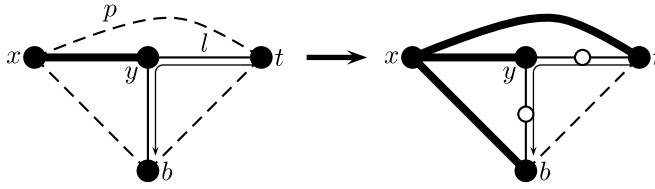


FIG. 2.9.

the path (y, x, v, t) of length l would then be a geodesic going through four principal vertices, which is impossible (by Proposition 2.5). Finally, $\langle y, x, b, t \rangle$ is isometric in G . It cannot be $S(K_4)$ because $P(x, y)$ is plain, and, as $P(y, t)$ is not plain, x must be the universal vertex in this subgraph. Then x would have type $\mathcal{C}$, which is impossible (by Lemma 2.9).

This means y cannot not have any of the required types. This ends the proof of Lemma 2.12. $\square$

PROPOSITION 2.13. *Each vertex v of L sees exactly one vertex of K .*

Proof. Existence: We just have to consider a uv -geodesic: it goes through K by visiting a vertex x which is seen by v .

Uniqueness: Suppose on the contrary that v sees two distinct vertices x and y of K . We can assume that $u \xrightarrow{x} v$. Furthermore, as it exists we can consider a 4-tuple (a, u, x, y) that realizes the following: $u \xrightarrow{x} a$, $a \rightarrow y$, and $d_G(u, x) + d_G(x, a) + d_G(a, y)$ is minimum for G and u .

CLAIM 2.14. $\langle u, a, x, y \rangle$ is an isometric subdivision of K_4 in G .

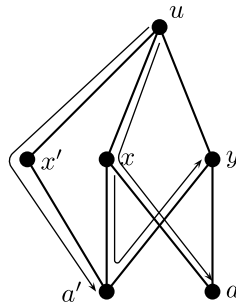


FIG. 2.10.

Suppose on the contrary that it is not isometric. There exists a' principal vertex of G distinct from u and a such that $x \xrightarrow{a'} y$. By Lemma 2.9, we can assume that $a' \in L$ (otherwise it would have type $\mathcal{C}$). Moreover, $d_G(u, x) + d_G(x, a') + d_G(a', y) < d_G(u, x) + d_G(x, a) + d_G(a, y)$. As (a, u, x, y) is a minimum for this quantity, it means that ua' -geodesics do not go through x or y . There exists a vertex x' in K distinct from x and y , such that $u \xrightarrow{x'} a'$ (see Figure 2.10). We then have the following: $u \xrightarrow{x'} a'$, $a' \rightarrow y$, and $d_G(u, x') + d_G(x', a') + d_G(a', y) < d_G(u, x) + d_G(x, a) + d_G(a, y)$. This is a contradiction with the assumption that proves our Claim 2.14.

Theorem 1.1 implies that x or y is universal in $\langle x, y, a, u \rangle$ (it cannot be isomorphic to $S(K_4)$ since $u \neq a$). Let us assume x is this universal vertex. Then $P(u, x)$ and $P(x, y)$ are both plain, which is impossible by Lemma 2.12.

We then have proven Proposition 2.13: each vertex v of L sees exactly one vertex x of K . In addition, we have $u \xrightarrow{x} v$, which is the only ua -geodesic. $\square$

We will now prove that $|K| = 1$.

Suppose on the contrary that there exist x and y distinct vertices of K . Both must have a type $\mathcal{R}$, $\mathcal{I}$, or Λ (by Remark 2.8 and Lemma 2.9). Each one of these types implies that both x and y see at least one vertex in L . Let a be a vertex of L such that x sees a . Then we consider a shortest path from a to y . It is not the subdivided edge $P(y, a)$ because of Proposition 2.13. It does not go through x , else $\langle a, y, u, x \rangle$ would induce an isometric subdivision of K_4 and x would be universal: $P(x, u)$ and $P(x, y)$ would then be plain, which is forbidden by Lemma 2.12. So there exists b in L such that $a \xrightarrow{b} y$. As b sees y we can assume that $u \xrightarrow{y} b$ (by Proposition 2.13). We then consider that (x, a, y, b) are the vertices such that $x, y \in K$, $a, b \in L$, $u \xrightarrow{x} a$, $u \xrightarrow{y} b$, and $a \xrightarrow{b} y$, which minimize the quantity $d_G(x, a) + d_G(a, b) + d_G(b, y)$.

CLAIM 2.15. $\langle u, x, y, a, b \rangle$ is isometric.

The only pairs of vertices that could have their geodesics out of the subdivided K_5 , $\langle u, x, y, a, b \rangle$, are (x, y) and (x, b) . Let us consider (x, y) and suppose that $x \xrightarrow{t} y$. Then $t \notin K$, else it would have type $\mathcal{C}$, and $t \notin L$, else it would see two vertices of K (forbidden by Proposition 2.13). Therefore, $t = u$, which is in the subdivided K_5 considered. Consequently, if $\langle u, x, y, a, b \rangle$ is not isometric, then any shortest xb -path does not belong to it. According to Proposition 2.13, the vertex b does not see any other vertex in K than y . Then there exists a vertex a' in L such that $b \xrightarrow{a'} x$. Since x is the unique vertex of K that a' sees (Proposition 2.13), $u \xrightarrow{x} a'$ (see Figure 2.11). Therefore, $u \xrightarrow{x} a'$, $u \xrightarrow{y} b$, and $b \xrightarrow{a'} x$. Furthermore, $d_G(x, a') + d_G(a', b) + d_G(b, y) < d_G(x, a) + d_G(a, b) + d_G(b, y)$, which contradicts the assumption $(d_G(x, a) + d_G(a, b) + d_G(b, y))$ is minimum). Hence, our claim is proven.

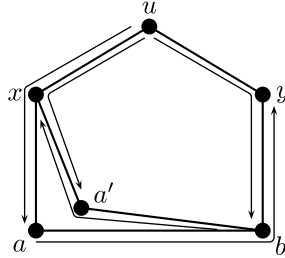


FIG. 2.11.

Hence, this subdivision of K_5 is a partial cube, and by Proposition 2.4 it either is isomorphic to $S(K_5)$ or has a universal vertex. It cannot be isomorphic to $S(K_5)$ because $P(u, a)$ would then be a ua -geodesic, and it is not since $a \in L_u$. Hence, it must have a universal vertex which can neither be u (it does not see a), nor a or b (they do not see u), nor x or y (then we would have $P(x, u)$ and $P(x, y)$ plain, which is forbidden by Lemma 2.12). Thus, this subgraph is not a partial cube, which contradicts our hypotheses.

We conclude $|K| = 1$, and then u sees only one vertex in G , which means that no geodesic goes through it. It implies that $G \setminus u$ is isometric. This is in contradiction with the assumption that for every principal vertex u in G , $G \setminus u$ is not isometric.

Hence Theorem 1.2 is proven.

2.3. A corollary. With this theorem we can construct many partial cubes starting from an arbitrary graph.

COROLLARY 2.16. *Let G be a subdivision of a graph of order n such that each edge contains an odd number of added vertices. K is a graph obtained from G by joining a vertex u adjacent to each principal vertex of G . Then K is a partial cube.*

Proof. Let D be the maximum number of subdivision of an edge in G . Then we add edges to obtain a subdivision of K_n . Each added edge is subdivided $2Dn + 1$ times so that the distances between vertices are unchanged. Then, by Theorem 1.2, the graph K' obtained by adding a vertex u adjacent to each principal vertex of G is a partial cube. Consequently we can remove the added edges without changing distances between principal vertices. We obtain that K is a partial cube. $\square$

Remark 2.17. This corollary shows that any minor can be found in a partial cube. Thus there cannot be a characterization of partial cube based upon forbidden minors.

Conclusion. A brief summary of the proof could be the following. We first prove that if a vertex can be removed isometrically, we then have a universal vertex thanks to the induction hypothesis. Then we still have to prove that we can always remove a vertex. We consider that if every vertex is needed, they can be split into two sets L and K such that all the vertices of K have a type among $\mathcal{C}, \Lambda, \mathcal{R}, \mathcal{I}$. We prove that none can have type $\mathcal{C}$. After that, we exhibit an isometric subdivision of K_4 to show that any vertex of L cannot see two vertices of K . We conclude by finding an isometric subdivision of K_5 that cannot have any universal vertex and is distinct from $S(K_5)$.

In the end, we have a structural characterization of every subdivision of complete graphs that are partial cubes.

At this point, one could ask questions with respect to two distinct directions. On the one hand, we could wonder what kind of graphs are isometrically embeddable in powers of K_3 and more generally in Hamming graphs. Our thought is that most of these graphs are also embeddable in hypercubes.

On the other hand, the study of subdivided graphs is a first step toward studying ℓ_1 -graphs (graphs which admit a λ -scale embedding in hypercubes; see [16, 5] for more information about ℓ_1 -graphs).

Acknowledgment. The authors wish to thank the anonymous referee for his precious advise and comments.

REFERENCES

- [1] M. AİDER, S. GRAVIER, AND K. MESLEM, *Isometric embeddings of subdivided connected graphs in the hypercube*, in Proceedings of the 5th Cracow Conference on Graph Theory (USTRON '06), Electron. Notes Discrete Math. 24, Elsevier, Amsterdam, 2006, pp. 145–151.
- [2] B. BREŠAR, W. IMRICH, AND S. KLAVŽAR, *Fast recognition algorithms for classes of partial cubes*, Discrete Math., 131 (2003), pp. 51–61.
- [3] B. BREŠAR AND S. KLAVŽAR, *On partial cubes and graphs with convex intervals*, Comment. Math. Univ. Carolin., 43 (2002), pp. 537–545.
- [4] V. CHEPOI, *Isometric subgraphs of Hamming graphs and d -convexity*, Cybernetics, 24 (1988), pp. 6–11.
- [5] V. CHEPOI, M. DEZA, AND V. GRISHUKHIN, *Clin d'oeil on ℓ_1 -embeddable planar graphs*, Discrete Appl. Math., 80 (1997), pp. 3–19.
- [6] V. CHEPOI AND C. TARDIF, *personal communication*, 1994.
- [7] D. DJOKOVIĆ, *Distance preserving subgraphs of hypercubes*, J. Combinatorial Theory Ser. B, 14 (1973), pp. 263–267.
- [8] D. EPPSTEIN, *Recognizing partial cubes in quadratic time*, in Proceedings of the 19th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2008), SIAM, Philadelphia, PA, 2008, pp. 1258–1266.

- [9] D. EPPSTEIN, *The lattice dimension of a graph*, European J. Combin., 26 (2005), pp. 585–592.
- [10] D. EPPSTEIN, *Cubic partial cubes from simplicial arrangements*, Electron. J. Combin., 13 (2006), article 79.
- [11] R. GRAHAM AND H. POLLAK, *On the addressing problem for loop switching*, Bell System Tech. J., 50 (1971), pp. 2495–2519.
- [12] S. GRAVIER, S. KLAVŽAR, AND M. MOLLARD, *Isometric embeddings of subdivided wheels in hypercubes*, Discrete Math., 269 (2003), pp. 287–293.
- [13] W. IMRICH AND S. KLAVŽAR, *Product Graphs: Structure and Recognition*, Wiley, New York, 2000.
- [14] S. KLAVŽAR AND A. LIPOVEC, *Partial cubes as subdivision graphs and as generalized Petersen graphs*, Discrete Math., 263 (2003), pp. 157–165.
- [15] S. KLAVŽAR AND A. LIPOVEC, *Edge-critical isometric subgraphs of hypercubes*, Ars Combin., 70 (2004), pp. 139–147.
- [16] S. SHECTOROV, *On scale embeddings of graphs into hypercubes*, European J. Combin., 14 (1993), pp. 117–130.
- [17] P. WINKLER, *Isometric embeddings in products of complete graphs*, Discrete Appl. Math., 7 (1984), pp. 221–225.

OPTIMAL DISTANCE LABELING FOR INTERVAL GRAPHS AND RELATED GRAPH FAMILIES*

CYRIL GAVOILLE[†] AND CHRISTOPHE PAUL[‡]

Abstract. A distance labeling scheme is a distributed graph representation that assigns labels to the vertices and enables answering distance queries between any pair (x, y) of vertices by using only the labels of x and y . This paper presents an optimal distance labeling scheme with labels of $\mathcal{O}(\log n)$ bits for the n -vertex interval graphs family. It improves by $\log n$ factor the best known upper bound of [M. Katz, N. A. Katz, and D. Peleg, *Distance labeling schemes for well-separated graph classes*, in Proceedings of the 17th Annual Symposium on Theoretical Aspects of Computer Science, Lecture Notes in Comput. Sci. 1770, Springer-Verlag, Berlin, 2000, pp. 516–528]. Moreover, the scheme supports constant time distance queries, and if the interval representation of the input graph is given and the intervals are sorted, then the set of labels can be computed in $\mathcal{O}(n)$ time. Our result is tight as we show that the length of any label is at least $3 \log n - \mathcal{O}(\log \log n)$ bits. This lower bound derives from a new estimator of the number of unlabeled n -vertex interval graphs, that is, $2^{\Omega(n \log n)}$. To our knowledge, interval graphs are thereby the first known nontrivial hereditary family with $2^{\Omega(n f(n))}$ unlabeled elements and with a distance labeling scheme with $f(n)$ bit labels.

Key words. graph representation, distance in graphs, distributed data-structure

AMS subject classifications. 05C62, 05C12, 68P30

DOI. 10.1137/050635006

1. Introduction. From an efficient graph representation, we could expect a small amount of memory space and fast elementary routines to answer queries like adjacency (see, e.g., [30] for an overview). Among the classical type of queries, let us mention those arising from the field of communication networks and distributed computing, like routing, connectivity, and distance queries. Peleg [26] introduced the concept of *informative labeling schemes* which generalizes the *implicit graph representation* proposed in [19]. Like classical graph representations, informative labeling schemes have the property of being a distributed data-structure: labels are assigned to the vertices and queries can be answered using only the labels of the involved vertices. Schemes providing compact labels play an important role for localized distributed data-structures (see [14] for a survey).

This paper focuses on the *distance labeling scheme* [27], which, for a graph family $\mathcal{F}$, is a pair $\langle L, f \rangle$ of functions such that for any graph $G = (V, E) \in \mathcal{F}$:

- $L(v, G)$ is a binary label associated with the vertex v in the graph G , and
- for all $x, y \in V$, $f(L(x, G), L(y, G)) = \text{dist}_G(x, y)$, the distance in G between x and y .

A scheme is an $\ell(n)$ -distance labeling if, for every n -vertex graph $G \in \mathcal{F}$, the length of the labels is at most $\ell(n)$ bits. The first results on the distance labeling

*Received by the editors July 2, 2005; accepted for publication (in revised form) April 14, 2008; published electronically July 25, 2008. This work has been supported by the French research grant ANR-06-BLAN-0148 “Graph Decomposition and Algorithm - GRAAL.” An extended abstract of this work appears in *Proceedings of 11th Annual European Symposium on Algorithms (ESA '03)*, Lecture Notes in Comput. Sci. 2832.

<http://www.siam.org/journals/sidma/22-3/63500.html>

[†]Laboratoire Bordelais de Recherche en Informatique, Université Bordeaux I, France (gavoille@labri.fr).

[‡]CNRS - Laboratoire d’Informatique, de Robotique et de Micro-électronique de Montpellier, Université Montpellier II, France (paul@lirmm.fr).

scheme [15] show that the family of n -vertex graphs requires $\Theta(n)$ bit labels while n -vertex trees need only $\Theta(\log^2 n)$ bit labels.¹ The case of dynamic tree networks has been studied in [23, 22, 24, 21]. The variant of approximated distance labeling schemes has been considered in [12, 32, 33, 34, 35, 29, 16, 25, 31]. Efficient distance labeling schemes are known for a number of restricted graph families. $\mathcal{O}(\sqrt{n} \log n)$ bit labels are enough for planar graphs, which is not tight with respect to the $\Omega(n^{1/3})$ lower bound [15]. For bounded degree graphs, $\Omega(\sqrt{n})$ is a lower bound on the label length. Finally, $\mathcal{O}(\log^2 n)$ -distance labeling schemes are known for interval and permutation graphs [20], for distance hereditary graphs [13], for bounded tree-width graphs, and, more generally, for bounded clique-width graphs [9].

A trivial observation is that a distance labeling scheme for a family $\mathcal{F}$ requires labels at least as large as labels of an adjacency labeling scheme for $\mathcal{F}$. A first natural question is: (1) *Could it be the same size for a large enough graph family?* As shown in [19], an information-theoretic lower bound coming from the number of n -vertex graphs in the family plays an important role for the label length. The unsolved *implicit graph representation conjecture* of [19] asks whether any hereditary family with $2^{nf(n)}$ graphs of n vertices enjoys an $\mathcal{O}(f(n))$ -adjacency labeling; $\Omega(f(n))$ is a lower bound. Moreover, up to now, none of the hereditary graph families is known to support an $o(\log^2 n)$ -distance labeling scheme.² It should be noticed that among the graph families listed above, the $\Omega(\log^2 n)$ lower bound on label length for n -vertex trees applies to neither interval nor permutation graphs. So the second question is: (2) *Can we improve the $\mathcal{O}(\log^2 n)$ upper bound [20] for either interval or permutation graphs?* This paper positively answers the above questions by considering the family of interval graphs. Interestingly, using some ideas from this paper, [2] positively answers the case of permutation graphs.

A graph $G = (V, E)$ is an *interval graph* (see, e.g., [4]) iff it is the intersection graph of a family $\mathcal{I}$ of intervals on the real line: vertices of V are in bijection with the intervals of $\mathcal{I}$ and two vertices x and y are adjacent iff their intervals $\mathcal{I}(x)$ and $\mathcal{I}(y)$ intersect. $\mathcal{I}$ is a *realizer* of the interval graph G . Obviously storing the bounds of each interval defines an $\mathcal{O}(\log n)$ -adjacency labeling scheme, as it can be shown that all the interval boundaries of $\mathcal{I}$ can be chosen in $\{1, \dots, 2n\}$. A graph is a *proper interval graph* iff it is the intersection graph of an inclusion-free realizer; i.e., none of the intervals of the realizer is contained in another interval.

We show that any $\ell(n)$ -distance labeling on the n -vertex interval graph family requires labels of length $\ell(n) \geq 3 \log n - \mathcal{O}(\log \log n)$ bits. This lower bound derives from a new estimator on the number of interval graphs. We also obtain an asymptotically tight upper bound of $5 \log n + 3$ bits for the label length. To that aim, we reduce any interval graph G to a proper interval graph G' such that the distance in G between any two vertices x, y can be retrieved from the distance in G' and the adjacency in G between x and y . It surprisingly follows that the core of the distance labeling problem of interval graphs consists in the design of a $2 \lceil \log n \rceil$ -distance labeling scheme for proper interval graphs. We also prove that the bound for proper interval graphs is asymptotically tight. Moreover, once the labels have been assigned, the distance computation from the labels takes a constant number of additions and comparisons on $\mathcal{O}(\log n)$ -bit integers. Even more interestingly, if the list of intervals of the graph is

¹In this paper all the logarithms are in base two.

²It is not difficult to construct a family of diameter two graphs whose adjacency can be decided with $\mathcal{O}(\log n)$ bit labels (some bipartite graphs, for instance), so supporting an $\mathcal{O}(\log n)$ distance labeling scheme. However, “diameter two” is not a hereditary property.

given sorted, then the preprocessing step to set all the labels runs optimally in $\mathcal{O}(n)$ time. Our scheme extends to circular-arc graphs.

At this step, it is worth remarking that any $\ell(n)$ -distance labeling scheme on a family $\mathcal{F}$ converts trivially to a nondistributed data-structure for $\mathcal{F}$ of $\mathcal{O}(\ell(n)n/\log n)$ space. The time complexity of distance queries remains unchanged, assuming that a cell of space can store $\Omega(\log n)$ bits of data. Therefore, as a byproduct we can compute in $\mathcal{O}(n)$ time an $\mathcal{O}(n)$ space data-structure for interval graphs supporting constant time distance queries. This latter formulation implies the result of [5]. However, both approaches differ in essence. The technique of [5] consists in building a one-to-one mapping from the vertices of the input graph to the nodes of a rooted tree, say, T . Then, distances are computed as follows. Let $l(v)$ be the level of v in T (i.e., the distance from the root), and let $A(i, v)$ be the i th ancestor of v (i.e., the i th node on the path from v to the root). If $l(x) > l(y) + 1$, then $\text{dist}(x, y) = l(x) - l(y) - 1 + d_1(z, x)$, where $z = A(l(x) - l(y) - 1, x)$, and where $d_1(z, x)$ is the distance between two nodes whose levels differ by at most 1. The distance d_1 is 1, 2, or 3 depending on the interval representation of the involved vertices. Answering a query is mainly based on the efficient parallel implementation of *level ancestor queries* on trees (to compute z) of [3]. However, this clever scheme cannot be converted into a distributed data-structure as ours can be for the following reason. As the tree has to support level ancestor queries, it implies that any node, if represented with a local label, can extract any of its ancestors with its label. In particular, x and y can extract from their label their nearest common ancestor and its level, so x and y can compute their distance in T . By the lower bound of [15], this cannot be done in less than $\Omega(\log^2 n)$ bit labels. So, access to a global data-structure is inherent in the approach of [5].

Section 2 deals with lower bounds. We first show that the number $I(n)$ of n -vertex labeled interval graphs is at least $2^{2n \log n - o(n \log n)}$, which is asymptotically tight. This first bound is then used to prove the lower bound on the length label for any distance labeling scheme of n -vertex interval graphs. A lower bound for proper interval graphs is also given. Section 3 describes a $(5 \log n + 3)$ -distance labeling scheme for interval graphs and its extension to circular-arc graphs. Section 4 deals with the relationship between distance labeling schemes and the notion of universal distance matrix.

2. Lower bounds. This section establishes lower bounds on the label length required to get an exact distance labeling scheme for interval graphs and proper interval graphs. We first prove a lower bound on the number of interval graphs from which, using the technical composition tool of an α -linkable subfamily (introduced in subsection 2.2), we derive the announced lower bounds on distance labeling schemes.

2.1. On the number of interval graphs. Let $I(n)$ be the number of labeled interval graphs with n vertices. Computing $I(n)$ is an unsolved graph-enumeration problem. Cohen, Komlós, and Muller [7] gave the probability $p(n, m)$ that a labeled n -vertex m -edge random graph is an interval graph under conditions on m . They computed $p(n, m)$ for $m < 4$ and showed, more generally, that

$$p(n, m) = \exp\left(-\frac{32c^6}{3}\right), \quad \text{where } \lim_{n \rightarrow +\infty} \frac{m}{n^{5/6}} = c.$$

As the total number of labeled n -vertex m -edge graphs is $\binom{\binom{n}{2}}{m}$, a formula of $p(n, m) \cdot \binom{\binom{n}{2}}{m}$ follows for the number of labeled interval graphs with $m = \Theta(n^{5/6})$ edges. Unfortunately, using this formula results in $I(n) \geq 2^{\Omega(n^{5/6} \log n)} = 2^{o(n)}$, which is a

lower bound too weak for our needs. The exact number of interval graphs is given up to 30 vertices in [17]. Actually, the generating functions for interval and proper interval graphs (labeled and unlabeled) are known [17], but only an asymptotic of $2^{2n+o(n)}$ for unlabeled n -vertex proper interval graphs can be estimated from these equations. In conclusion, Hanlon [17] left open whether the asymptotic on the number of unlabeled interval graphs is $2^{\mathcal{O}(n)}$ or $2^{\Omega(n \log n)}$.

As the number of labeled interval graphs is clearly at least $n! = 2^{(1-o(1))n \log n}$ (just consider a labeled path), the open question of Hanlon is whether $I(n) = 2^{(c-o(1))n \log n}$ for some constant $c > 1$. Hereafter we show that $c = 2$, which is optimal.

THEOREM 1. *The number $I(n)$ of labeled n -vertex interval graphs satisfies $\frac{1}{n} \log I(n) \geq 2 \log n - \log \log n - \mathcal{O}(1)$. It follows that there are $2^{\Omega(n \log n)}$ unlabeled n -vertex interval graphs.*

Proof. There are at least $I(n)/n!$ unlabeled interval graphs. First let us show that the lower bound on $\frac{1}{n} \log I(n)$ stated in Theorem 1 implies that there are $2^{\Omega(n \log n)}$ unlabeled n -vertex interval graphs. Indeed, using the fact that $\log(n!) = n \log n - \Theta(n)$, we have

$$\begin{aligned} \frac{1}{n} \log I(n) &\geq 2 \log n - \log \log n - \mathcal{O}(1) \\ \Rightarrow \log I(n) &\geq 2n \log n - n \log \log n - \mathcal{O}(n) \\ \Rightarrow \log I(n) - \log(n!) &\geq 2n \log n - \log(n!) - n \log \log n - \mathcal{O}(n) \\ \Rightarrow \log(I(n)/n!) &\geq n \log n - n \log \log n - \mathcal{O}(n) \\ \Rightarrow I(n)/n! &\geq 2^{n \log n - o(n \log n)}. \end{aligned}$$

So there are $2^{\Omega(n \log n)}$ unlabeled interval graphs. Let us now prove the lower bound on $I(n)$.

Let $\mathcal{S}_{p,k}$ denote the set of all the sequences $(S_1, \dots, S_p)$ of integer sets such that, for every $i \in \{1, \dots, p\}$, it holds that $|S_i| = k$ and $S_i \subset \{0, \dots, (p-i)(k+1) + k\}$. As there are $\binom{(p-i)(k+1)+k+1}{k}$ ways to choose S_i , it follows that $|\mathcal{S}_{p,k}| = \prod_{i=1}^p \binom{(p-i)(k+1)+k+1}{k} = \prod_{i=1}^p \binom{(p-i+1)(k+1)}{k} = \prod_{i=1}^p \binom{i(k+1)}{k}$.

From every $S = (S_1, \dots, S_p) \in \mathcal{S}_{p,k}$ we construct an interval graph G_S of vertex set $V(G_S) = A \cup B \cup C$ with $A = \{a_{i,j} \mid 1 \leq i \leq p, 1 \leq j \leq k\}$, $B = \{b_0, \dots, b_p\}$, and $C = \{c_1, \dots, c_k\}$. The edge set is defined by the intersection of the intervals associated to each vertex of G_S . For every $v \in V(G_S)$, the interval of v is $\mathcal{I}(v)$ and is defined as follows: for $i \in \{1, \dots, p\}$ and $j \in \{1, \dots, k\}$ (in the following, we denote by $s_{i,j}$ the j th smallest integer of S_i)

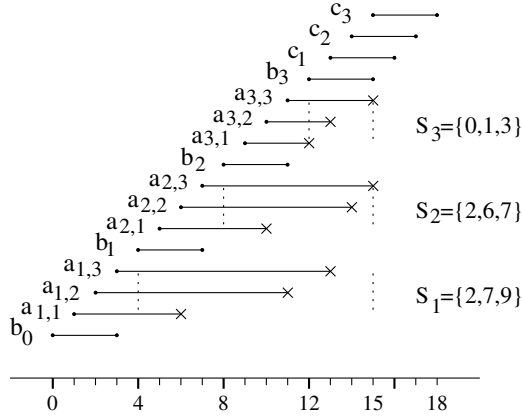
- $\mathcal{I}(a_{i,j}) = [(i-1)(k+1) + j, i(k+1) + s_{i,j}]$;
- $\mathcal{I}(b_i) = [i(k+1), i(k+1) + k]$ and $\mathcal{I}(b_0) = [0, k]$; and
- $\mathcal{I}(c_j) = [p(k+1) + j, p(k+1) + j + k]$.

As depicted in Figure 1 (where $p = k = 3$), for each fixed i , the right boundaries of the $a_{i,j}$'s must differ and must fall into the area delimited by the two vertical dotted lines. The intervals associated with the b_i 's and c_j 's are all of length k .

We consider the set $\mathcal{F}_{p,k}$ composed of all labeled graphs G_S obtained for $S \in \mathcal{S}_{p,k}$ and such that b_0 is labeled 1 and c_j is labeled $j+1$ for every $j \in \{1, \dots, k\}$. Clearly, $\mathcal{F}_{p,k}$ is composed of labeled connected interval graphs on $n = (p+1)(k+1)$ vertices. Thus, $I((p+1)(k+1)) \geq |\mathcal{F}_{p,k}|$. We will first show that

$$(1) \quad \forall p, k \geq 1, \quad |\mathcal{F}_{p,k}| = (p(k+1))! \cdot |\mathcal{S}_{p,k}|$$

and then show how to fix $p = p(n)$ and $k = k(n)$ to get the desired lower bound on $I(n)$.

FIG. 1. The interval representation of G_S for $S = (S_1, S_2, S_3)$.

By construction $\mathcal{F}_{p,k}$ contains at most $|\mathcal{S}_{p,k}|$ unlabeled graphs, and each one has at most $(|A| + |B \setminus \{b_0\}|)! = (p(k+1))!$ ways of labeling, since $b_0, c_1, \dots, c_k$ are labeled $1, \dots, k+1$. Hence $|\mathcal{F}_{p,k}| \leq (p(k+1))! |\mathcal{S}_{p,k}|$. To prove equality, we describe hereafter a two-step procedure that from any $G \in \mathcal{F}_{p,k}$ (1) retrieves in a unique way a sequence S such that G is isomorphic to G_S and (2) gives the label sequence of the vertices of G in a given order (for instance, in the increasing order of their left boundary in the interval representation of G_S). These two steps ensure that $|\mathcal{F}_{p,k}| \geq (|A| + |B \setminus \{b_0\}|)! |\mathcal{S}_{p,k}|$.

For that, let us consider any labeled graph $G \in \mathcal{F}_{p,k}$. Let S be a sequence such that G is isomorphic to G_S . (Obviously, at this step, S and the interval representation of G_S are unknown.) We say that a vertex v of G is *identified* if one can decide whether $v = a_{i,j}$ or $v = b_i$ or $v = c_j$ (and then give for each case the corresponding indices i and j). So identifying all the vertices of G completes step (2). As we will see later, identifying a vertex v also allows the recovery of its interval representation and eventually its associated integer $s_{i,j}$, completing step (1) as well.

For every $i \in \{0, \dots, p\}$, let L_i denote the set of vertices of G_S whose left boundary is $\leq i(k+1)$, i.e., at most equal to the left boundary of b_i . To show that all the vertices of G can be identified, it suffices to show that all the vertices of L_p can be identified in G . Indeed, $G_S \setminus L_p$ consists of the vertices of C , and each $c_j \in C$ can be identified in G as its label is $j+1$. So, let us show by induction that all the vertices of L_p can be identified in G .

$L_0 = \{b_0\}$ can be identified in G as its label is 1. Assume that L_{i-1} can be identified, $i > 0$, and let us show how to identify the vertices of L_i . As $L_{i-1} \subset L_i$, we just have to identify $L_i \setminus L_{i-1} = \{a_{i,1}, \dots, a_{i,k}\} \cup \{b_i\}$.

Let $G_i = (G_S \setminus L_{i-1}) \cup \{b_{i-1}\}$. In the graph G_i , we have the following properties that come trivially from the interval definition of G_S :

1. the neighbors of b_{i-1} are $a_{i,1}, \dots, a_{i,k}$;
2. every vertex of $V(G_i) \setminus \{b_{i-1}\} \setminus C$ is of degree $> k$; and
3. the degree of $a_{i,j}$ is $k + j + s_{i,j}$.

Note that by assumption b_{i-1} and L_{i-1} are identified, so the above properties also hold for the graph $(G \setminus L_{i-1}) \cup \{b_{i-1}\}$. Combining properties 1 and 3, the $a_{i,j}$'s can be identified as the degree of $a_{i,j}$'s are distinct and increasing with j . Moreover, as the graph G_{i+1} can be obtained from G_i by removing the identified vertices $b_{i-1}, a_{i,1}, \dots, a_{i,k}$,

it follows (applying properties 1 and 2 to G_{i+1}) that b_i can be identified: it is the only degree k vertex of $G_{i+1} \setminus C$. Thus L_i 's vertices can be identified in G . Therefore, all the vertices of L_p , and thus all the vertices of G , can be identified. This completes step (2).

Recovering the sequence S is easy once $a_{i,j}$'s have been identified in G : it suffices to compute G_i and, by Property 3, to recover $s_{i,j}$ from the degree of $a_{i,j}$. This completes step (1) of the procedure and thus proves equation (1).

Using the fact that $\binom{a}{b} \geq (a/b)^b$, we have

$$|\mathcal{S}_{p,k}| \geq \prod_{i=1}^p \left(\frac{i(k+1)}{k} \right)^k = \left(1 + \frac{1}{k} \right)^{pk} (p!)^k \geq (p!)^k.$$

Thanks to (1) and using the fact that $x! \geq (x/e)^x$, we have

$$|\mathcal{F}_{p,k}| \geq \left(\frac{p(k+1)}{e} \right)^{p(k+1)} \left(\frac{p}{e} \right)^{pk} \geq \left(\frac{p^2 k}{e^2} \right)^{pk}.$$

As $I(n)$ is increasing, for every $n \geq (p+1)(k+1)$ we therefore have

$$\log I(n) \geq \log |\mathcal{F}_{p,k}| \geq 2pk \log p + pk \log k - \mathcal{O}(pk).$$

Let us choose $k = \lfloor \log n \rfloor - 1$ and $p = \lfloor n/\log n \rfloor - 1$ so that $n \geq (p+1)(k+1)$. Observe that $pk \geq n(1 - 4/\log n)$. We have

$$\begin{aligned} \log I(n) &\geq 2pk \log(n/\log n) + pk \log \log n - \mathcal{O}(n) \\ &\geq 2pk \log n - pk \log \log n - \mathcal{O}(n) \\ &\geq 2n(1 - 4/\log n) \log n - n(1 - 4/\log n) \log \log n - \mathcal{O}(n) \\ &\geq 2n \log n - n \log \log n - \mathcal{O}(n). \end{aligned}$$

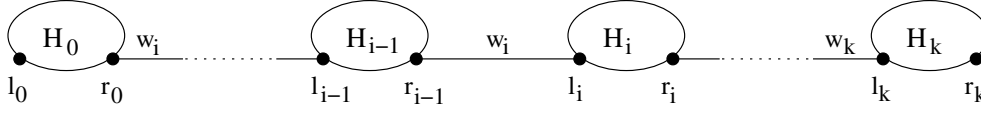
This ends the proof of Theorem 1. $\square$

Each vertex can be described by its interval representation, and since there are at most $2n$ interval boundaries ranging in $[1, 2n]$, adjacency labels of length $\lceil 2 \log(2n) \rceil$ suffice. Concatenating such labels according to the vertex ordering of the graph yields an upper bound on the number of labeled n -vertex graphs, i.e., $I(n) \leq \lceil 2 \log(2n) \rceil^n \leq 2n \log n + \mathcal{O}(n)$. It follows that $\frac{1}{n} \log I(n) \leq 2 \log n + \mathcal{O}(1)$, and so the lower bound of Theorem 1 on $\frac{1}{n} \log I(n)$ is tight up to an additive factor of $\log \log n$.

2.2. Lower bounds on distance labeling schemes. The *eccentricity* of a vertex x in a graph G is the maximum of the distances between x and any vertex of G . An α -graph, for integer $\alpha \geq 1$, is a graph G having a pair of vertices (l, r) , possibly with $l = r$, each one of eccentricity at most α . Let $\mathcal{S} = (H_0, H_1, \dots, H_k)$ be a sequence of α -graphs, and let (l_i, r_i) denote the pair of vertices that defines the α -graph H_i for $i \in \{0, \dots, k\}$. For each nonnull integer sequence $W = (w_1, \dots, w_k)$, we denote by $\mathcal{S}^W$ the graph obtained by attaching a path of length w_i between the vertices r_{i-1} and l_i for every $i \in \{1, \dots, k\}$ (see Figure 2).

A subfamily $\mathcal{H} \subset \mathcal{F}$ of graphs is α -linkable if $\mathcal{H}$ consists only of α -graphs and if $\mathcal{S}^W \in \mathcal{F}$ for every sequence $\mathcal{S}$ of graphs of $\mathcal{H}$ and for every nonnull integer sequence W .

The following lemma shows a lower bound on the length of the labels required by a distance labeling scheme on any graph family having an α -linkable subfamily. The bound is related to the number of labeled graphs contained in the subfamily. As we

FIG. 2. Linking a sequence of α -graphs.

shall see, interval graphs support a large 1-linkable subfamily (we mean large with respect to n), whereas proper interval graphs support a large 2-linkable subfamily.

LEMMA 1. Let $\mathcal{H}$ be an α -linkable subfamily of a graph family $\mathcal{F}$. Let $H(\nu)$ denote the number of labeled ν -vertex graphs of $\mathcal{H}$. Then, every distance labeling scheme on the family of n -vertex graphs of $\mathcal{F}$ requires a label of length at least $\frac{1}{\nu} \log H(\nu) + \log \nu - 9$, where $\nu = \lfloor n/(\alpha \log n) \rfloor$.

Let us first sketch the proof. It uses a sequence $\mathcal{S}$ of $k = \Theta(\log n)$ graphs H_i taken from an arbitrary α -linkable subfamily $\mathcal{H}$. Each graph H_i has $\nu = \Theta(n/\log n)$ vertices, and consecutive H_i 's from $\mathcal{S}$ are spaced with paths of length $\Theta(n/\log n)$. Intuitively, the term $\frac{1}{\nu} \log H(\nu)$ measures the minimum label length required to compute the distance between any two vertices within the same H_i 's (i.e., whether they are adjacent or at distance at least two). The $\log \nu$ term is required to compute the distance between vertices of distinct H_i 's. The difficulty is to show that some vertices require both pieces of information, observing that one can distribute information on the vertices in a nontrivial way. For instance, the two extreme vertices of a path of length w_i do not require $\log w_i$ bit labels, but only $\frac{1}{2} \log w_i$ bits: each extremity can store one half of the binary word representing w_i and merge their labels for a distance query.

Proof of Lemma 1. Let p, q, k be nonnull integers. Let $\mathcal{W}$ be the set of all the integer sequences $(w_1, \dots, w_k)$ such that $w_i \in \{1 + (j-1)(2\alpha+1) \mid 1 \leq j \leq q\}$. Since a sequence may contain repeated elements, $|\mathcal{W}| = q^k$. Let $\mathcal{H}$ be an arbitrary α -linkable subfamily of $\mathcal{F}$. Let $H(p)$ be the number of labeled p -vertex graphs of $\mathcal{H}$. For every $W \in \mathcal{W}$, we denote by $\mathcal{H}_p(W)$ the set of all graphs $\mathcal{S}^W$ where $\mathcal{S}$ is a sequence of $k+1$ graphs of $\mathcal{H}$, each with exactly p vertices (see Figure 2).

The number of vertices of any graph of $\mathcal{H}_p(W)$ is $(k+1)p + (\sum_{i=1}^k (w_i - 1)) \leq (k+1)p + k(q-1)(2\alpha+1) + 1 \leq (k+1)(p+3\alpha q)$ as $\alpha \geq 1$. Since $\mathcal{H}$ is an α -linkable subfamily of $\mathcal{F}$, any graph $\mathcal{S}^W$ belongs to $\mathcal{F}$, and thus, for every $n \geq (k+1)(p+3\alpha q)$, we get $\bigcup_{W \in \mathcal{W}} \mathcal{H}_p(W) \subseteq \mathcal{F}_n$, where $\mathcal{F}_n$ denotes the family of all n -vertex graphs of $\mathcal{F}$.

Let $\langle L, f \rangle$ be any distance labeling scheme on $\mathcal{F}_n$. We consider the set $\mathcal{L}_W$ of all the $(k+1)$ -tuples of labels for all the graphs $\mathcal{S}^W \in \mathcal{H}_p(W)$, each tuple being formed by taking a vertex label successively in $H_0, H_1, \dots, H_k$. More formally,

$$\begin{aligned} \mathcal{L}_W = & \bigcup_{\substack{\mathcal{S}^W \in \mathcal{H}_p(W) \\ \text{with } H=(H_0, \dots, H_k)}} \{ (L(u_0, H_0), \dots, L(u_k, H_k)) \mid (u_0, \dots, u_k) \\ & \in V(H_0) \times \dots \times V(H_k) \}. \end{aligned}$$

Finally, let $\mathcal{L} = \bigcup_{W \in \mathcal{W}} \mathcal{L}_W$.

CLAIM 1. L assigns on a graph of $\mathcal{F}_n$ a label of length at least $\frac{1}{k+1} \log |\mathcal{L}| - 1$.

Indeed, assume that the length of any label assigned by L on some graph of $\mathcal{F}_n$ is at most t bits. Then, as there are $2^{t+1} - 1$ binary strings of length at most t , there are $(2^{t+1} - 1)^{k+1}$ distinct $(k+1)$ -tuples of such labels. Since $\mathcal{L}$ is a set of $(k+1)$ -tuples, thus $|\mathcal{L}| \leq (2^{t+1} - 1)^{k+1}$, implying that $t \geq \log(|\mathcal{L}|^{1/(k+1)} + 1) - 1 \geq \frac{1}{k+1} \log |\mathcal{L}| - 1$.

CLAIM 2. $|\mathcal{L}_W|^p \geq H(p)^{k+1}$. We can assume each graph $H_j \in \mathcal{S}$ of any $\mathcal{S}^W \in \mathcal{H}_p(W)$ is defined on a copy of the same sorted set of vertices. Therefore, a set L_S of $(k+1)$ -tuples can be associated with each $\mathcal{S}^W$,

$$L_S = \{(L(v_j, H_0), \dots, L(v_j, H_k)) \mid j \in [1, p]\},$$

each $(k+1)$ -tuple being composed by the labels of an arbitrary vertex in each H_i . We have $\bigcup_{\mathcal{S}^W \in \mathcal{H}_p(W)} L_S \subseteq (\mathcal{L}_W)^p$. Let us show that $L_S \neq L_{S'}$ for all $\mathcal{S} \neq \mathcal{S}'$. Indeed, if $\mathcal{S} \neq \mathcal{S}'$, there exist two vertices, say, v_{j_1} and v_{j_2} , adjacent in the i th graph H_i of $\mathcal{S}$ but not adjacent in $H'_i \in \mathcal{S}'$ ($\text{dist}_{\mathcal{S}^W}(v_{j_1}, v_{j_2}) \geq 2$). The distance decoder f must therefore return different values for the corresponding pair of labels, i.e., $f(L(v_{j_1}, H_i), L(v_{j_2}, H_i)) \neq f(L(v_{j_1}, H'_i), L(v_{j_2}, H'_i))$. As both labels, $L(v_{j_1}, H_i)$ and $L(v_{j_1}, H'_i)$ ($L(v_{j_2}, H_i)$ and $L(v_{j_2}, H'_i)$, respectively), are located in the same respective $(k+1)$ -tuple of L_S and $L_{S'}$, $L_S \neq L_{S'}$ follows. It implies that

$$\left| \bigcup_{\mathcal{S}^W \in \mathcal{H}_p(W)} L_S \right| \geq |\mathcal{H}_p(W)|.$$

Since $\mathcal{H}_p(W)$ contains any sequence of $(k+1)$ p -vertex graphs of $\mathcal{H}$, $|\mathcal{H}_p(W)| = H(p)^{k+1}$ ($H(p)$ counting the p -vertex graphs of $\mathcal{H}$). Combining the previous inequalities and inclusions, Claim 2 follows: we have $|\mathcal{L}_W|^p \geq H(p)^{k+1}$.

CLAIM 3. For all distinct $W, W' \in \mathcal{W}$, $\mathcal{L}_W \cap \mathcal{L}_{W'} = \emptyset$.

First observe that if W and W' differ at the i th entry, with $w_i \neq w'_i$, then $[w_i, w_i + 2\alpha] \cap [w'_i, w'_i + 2\alpha] = \emptyset$. Indeed, by construction of the integers of $\mathcal{W}$, $|w_i - w'_i| \geq 2\alpha + 1$. We observe that $f(L(u, H_{i-1}), L(v, H_i)) \in [w_i, w_i + 2\alpha]$ for every $i \in \{1, \dots, k\}$ and for all $(u, v) \in V(H_{i-1}) \times V(H_i)$. Note also that every tuple of labels $T \in \mathcal{L}_W$ contains a label taken from H_{i-1} and a label from H_i . So, as the ranges of distances are disjoint, if W and W' differ by the i th entry, then every element $T \in \mathcal{L}_W$ and every element $T' \in \mathcal{L}_{W'}$ must differ by the $(i-1)$ th or the i th entry, proving the last claim.

From the last two claims and since $|\mathcal{W}| = q^k$, $|\mathcal{L}| \geq \sum_{W \in \mathcal{W}} |\mathcal{L}_W| \geq H(p)^{(k+1)/p} q^k$ holds. From Claim 1, the label length is at least (note that $k/(k+1) \geq 1 - 2/(k+2)$ for $k \geq 1$)

$$t \geq \frac{1}{p} \log H(p) + \frac{k}{k+1} (\log q) - 1 > \frac{1}{p} \log H(p) + \left(1 - \frac{2}{k+2}\right) (\log q) - 1.$$

Let $\nu = \lfloor n/(\alpha \log n) \rfloor$, $p = q = \nu$, and $k = \lfloor \alpha \log n/(1+3\alpha) \rfloor - 1$. We have $n \geq (k+1)(p+3\alpha q)$, as required. So,

$$\begin{aligned} t+1 &\geq \frac{1}{\nu} \log H(\nu) + \left(1 - \frac{2}{\alpha \log n/(1+3\alpha)}\right) \log \nu \\ &\geq \frac{1}{\nu} \log H(\nu) + \log \nu - \frac{2(1+3\alpha)}{\alpha} \frac{\log \nu}{\log n} \\ &\geq \frac{1}{\nu} \log H(\nu) + \log \nu - 8, \quad \text{as } n \geq \nu \text{ and } \alpha \geq 1, \text{ completing the proof.} \quad \square \end{aligned}$$

Let us now show that interval graphs and proper interval graphs both enjoy a large enough α -linkable subfamily of labeled graphs. The lower bounds on distance labeling schemes for these two graph families therefore follow from Lemma 1 and Theorem 1.

THEOREM 2. *Any distance labeling scheme on the family of n -vertex interval graphs requires a label of length at least $3 \log n - 4 \log \log n$.*

Proof. Let $I(\nu)$ be the number of labeled interval graphs with ν vertices. Let $\mathcal{H}$ be that family of labeled ν -vertex interval graphs enjoying a universal vertex (adjacent to any other vertex) where the universal vertex is labeled ν . Any graph of $\mathcal{H}$ is a 1-graph (set the universal vertex as $l = r$). And note that $\mathcal{H}$ is a 1-linkable subfamily of the family of interval graphs (see Figure 3). Moreover, it is straightforward to see that any $(\nu - 1)$ -vertex labeled interval graph is an induced subgraph of some $H \in \mathcal{H}$. It follows that the number $H(\nu)$ of ν -vertex labeled graphs of $\mathcal{H}$ is precisely $I(\nu - 1)$.

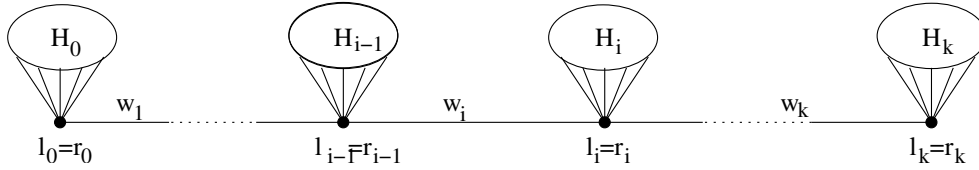


FIG. 3. Each H_i is a 1-linkable ν -interval graph with a universal vertex $l_i = r_i$.

By Lemma 1 applied to $\mathcal{H}$ and the interval graph family, the maximum length t of a label of length is at least

$$t \geq \frac{1}{\nu} \log H(\nu) + \log \nu - \mathcal{O}(1) \geq \frac{1}{\nu} \log I(\nu - 1) + \log \nu - \mathcal{O}(1).$$

From Theorem 1, we have

$$\begin{aligned} \frac{1}{\nu} \log I(\nu - 1) &= \frac{\nu - 1}{\nu} (2 \log(\nu - 1) - \log \log(\nu - 1) - \mathcal{O}(1)) \\ &\geq 2 \log(\nu - 1) - \log \log \nu - \mathcal{O}(1). \end{aligned}$$

Since $\mathcal{H}$ is a 1-graph family, we have $\alpha = 1$. Therefore, Lemma 1 sets $\nu = \lfloor n / \log n \rfloor$. This implies that $\nu - 1 \geq \frac{n}{2 \log n}$.

$$\begin{aligned} \frac{1}{\nu} \log I(\nu - 1) &\geq 2 \log \left(\frac{n}{2 \log n} \right) - \log \log \lfloor n / \log n \rfloor - \mathcal{O}(1) \\ &\geq 2 \log n - 3 \log \log n + \log \log \log n - \mathcal{O}(1). \end{aligned}$$

Gathering everything together, we obtain

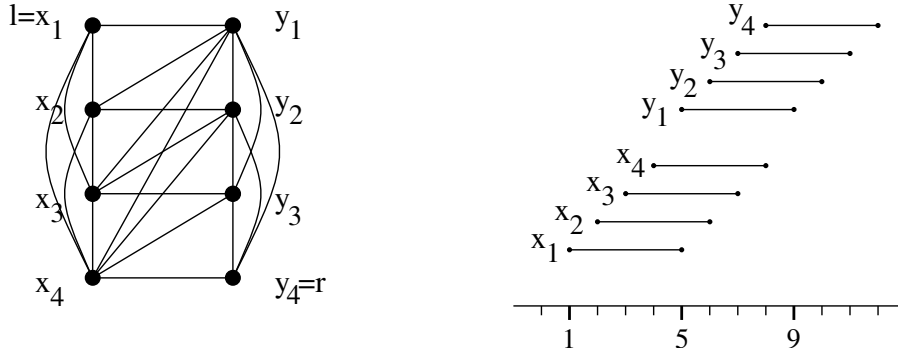
$$\begin{aligned} t &\geq 2 \log n - 3 \log \log n + \log \log \log n + \log \lfloor n / \log n \rfloor - \mathcal{O}(1) \\ &\geq 3 \log n - 4 \log \log n, \end{aligned}$$

completing the proof. $\square$

THEOREM 3. *Any distance labeling scheme on the family of n -vertex proper interval graphs requires a label of length at least $2 \log n - 2 \log \log n - \mathcal{O}(1)$.*

Proof. Let Q_p be the intersection graph of the interval family $\{\mathcal{I}(x_i) = [i, i + p], \mathcal{I}(y_i) = [i + p, i + 2p - 1], 1 \leq i \leq p\}$. (See Figure 4.) Since any interval is of length p , Q_p is a proper interval graph with $2p$ vertices.

Let $\mathcal{H}$ be the family of labeled graphs isomorphic to Q_p such that the vertices x_1 and y_1 are labeled 1 and 2, respectively. Let $Q \in \mathcal{H}$. It is not difficult to see by

FIG. 4. The graph Q_4 and its interval representation.

induction that once x_i and y_i have been identified in Q , then x_{i+1} and y_{i+1} can be identified, too. It suffices to remove all the vertices $x_1, \dots, x_{i-1}$ and $y_1, \dots, y_{i-1}$ from H (and so to form a graph isomorphic to $Q_{p-(i-1)}$), and observe that in this latter graph (1) x_{i+1} is the neighbor of x_i , distinct from y_i , with minimum degree; and (2) y_{i+1} is the only nonneighbor of x_i that is adjacent to x_{i+1} . Therefore, the cardinality of $\mathcal{H}$ is $(2p-2)!$.

Now, setting $l = x_1$ and $r = y_p$, l and r are both of eccentricity 2, as the vertex x_p is a universal vertex. So $\mathcal{H}$ is a family of 2-graphs. Moreover $\mathcal{H}$ is a 2-linkable subfamily of the proper interval graphs.

Let $\nu = 2p$. The cardinality $H(\nu)$ of $\mathcal{H}$ is $(\nu-2)!$. We note that $\frac{1}{\nu} \log(\nu-2)! \geq \log \nu - \mathcal{O}(1)$. By Lemma 1, for $\alpha = 2$ and $\nu = \lfloor n/(2 \log n) \rfloor$, the length of the maximum label for n -vertex proper interval graphs is at least

$$\frac{1}{\nu} \log H(\nu) + \log \nu - \mathcal{O}(1) \geq \log \nu + \log \nu - \mathcal{O}(1) \geq 2 \log n - 2 \log \log n - \mathcal{O}(1),$$

ending the proof. $\square$

3. Upper bounds on distance labeling schemes. This section deals with optimal distance labeling schemes for proper interval graphs, interval graphs, and circular-arc graphs. We first show how to complete any interval graph G into a proper interval graph G' so that the distance between two vertices x and y in G can be retrieved from the distance in G' and the adjacency in G between x and y . Then, we present a distance labeling scheme for proper interval graphs. This scheme is based on an adjacency labeling scheme of an auxiliary graph. It turns out that the auxiliary graph of any proper interval graph enjoys an optimal adjacency labeling scheme. The $\mathcal{O}(\log n)$ -distance labeling scheme for interval graphs follows. Finally, an extension of this scheme is proposed for the family of circular-arc graphs.

From now on, the input interval graph G is supposed to be an n -vertex connected graph, and $\mathcal{I}$ is given as a realizer. We then denote by $\mathcal{I}(x)$ the interval associated with vertex x and by $L(x)$ and $R(x)$, respectively, the left and the right boundaries of $\mathcal{I}(x)$. As in [5], the intervals of $\mathcal{I}$ are assumed to be sorted according to the left boundaries. We will also assume w.l.o.g. that the left boundaries are pairwise distinct and range in $\{1, \dots, 2n\}$. These hypotheses are not restrictive since in $\mathcal{O}(n)$ time one can scan all the intervals and their boundaries from $\min_x L(x)$ to $\max_x L(x)$ and compute another realizer for G with sorted and distinct boundaries in $\{1, \dots, 2n\}$.

3.1. Reduction to proper interval graphs. Proper interval graphs are interval graphs having an inclusion-free realizer. There are several recognition algorithms that run in linear time, that is, $\mathcal{O}(n + m)$ time, where n is the number of vertices and m the number of edges; cf. [8, 10, 18]. Moreover, these algorithms output an inclusion-free realizer.

It is well known that the family of proper interval graphs can be characterized as the family of $K_{1,3}$ -free³ interval graphs [37] (see Figure 5) or as the family of intersection graphs of a family of unit length intervals [28].



FIG. 5. The $K_{1,3}$ is an interval graph but not a proper interval graph.

Reducing the interval graph to a proper interval graph, the interval realizer we obtain yields a slightly different characterization of proper interval graphs. Such a realizer is called a *proper realizer*.

DEFINITION 1. A realizer $\mathcal{I}$ of an interval graph is proper if for every pair of intervals $\mathcal{I}(x)$ and $\mathcal{I}(y)$ such that $\mathcal{I}(x) \subseteq \mathcal{I}(y)$, $R(x) = R(y)$.

Let us notice that an inclusion-free realizer of intervals is a proper realizer; indeed, only special inclusions are allowed in a proper realizer. The reverse is also true: any interval graph having a proper realizer is $K_{1,3}$, i.e., is a proper interval graph. Let $\mathcal{I}$ be a proper realizer of an interval graph G . Assume by contradiction that G contains a set of four vertices that induces a $K_{1,3}$; say, a, b, c are three independent vertices adjacent to a common neighbor x . W.l.o.g. assume that $L(a) < R(a) < L(b) < R(b) < L(c) < R(c)$. It follows that $L(x) \leq R(a)$ and $L(c) \leq R(x)$. Thereby we have $\mathcal{I}(b) \subset \mathcal{I}(x)$ and $R(b) \neq R(x)$, which is a contradiction; the realizer $\mathcal{I}$ is supposed to be proper.

Our reduction is based on the notion of an *enclosing neighbor*.

DEFINITION 2. Let $\mathcal{I}$ be a nonproper realizer of an interval graph $G = (V, E)$. Let $x \in V$ such that $\mathcal{I}(x) \subset \mathcal{I}(y)$ and $R(x) < R(y)$ for some vertex y . The enclosing neighbor of x , denoted by $N_e(x)$, is a neighbor of x for which $\mathcal{I}(x) \subset \mathcal{I}(N_e(x))$ such that the following conditions are satisfied:

- $R(N_e(x))$ is maximal, and
- $\mathcal{I}(N_e(x))$ is maximal for the inclusion.

As the left boundaries of $\mathcal{I}$ are pairwise disjoint, the enclosing neighbor, if it exists, is unique for every vertex. From a realizer $\mathcal{I}$, we can construct a proper realizer $\mathcal{I}'$ as follows: if x has an enclosing neighbor, then we set $\mathcal{I}'(x) = [L'(x) = L(x), R'(x) = R(N_e(x))]$, and we set $\mathcal{I}'(x) = \mathcal{I}(x)$ otherwise.

Let us remark that

1. $R'(x) \geq R(x)$;
2. by maximality of its right boundary $\mathcal{I}(N_e(x)) = \mathcal{I}'(N_e(x))$; and
3. G is a subgraph of $G' = (V, E' \subseteq E)$, the graph realized by $\mathcal{I}'$.

We say that a vertex x has been *extended* in G' if $R'(x) > R(x)$.

LEMMA 2. The realizer $\mathcal{I}'$ is proper. It can be computed from $\mathcal{I}$ in $\mathcal{O}(n)$ time.

³That is, the graph does not contain a $K_{1,3}$ graph as an induced subgraph.

Proof. Assume there is a pair of vertices u and v such that $\mathcal{I}'(u) \subset \mathcal{I}'(v)$ and $R'(u) < R'(v)$. W.l.o.g. assume that $\mathcal{I}'(v)$ is maximal for inclusion. Notice that $R(v) = R'(v)$; otherwise v would have an enclosing neighbor w that would imply $\mathcal{I}'(v) \subset \mathcal{I}'(w)$, contradicting the maximality of v . As $R(u) \leq R'(u)$, we have $L(v) < L(u) < R(u) < R'(v)$. Now, $R(v) = R'(v)$ implies $\mathcal{I}(u) \subset \mathcal{I}(v)$, meaning that u has an enclosing neighbor. It follows that $R'(u) \geq R(v) = R'(v)$, contradicting $R'(u) < R'(v)$. So we proved that for any pair of vertices u and v such that $\mathcal{I}'(u) \subset \mathcal{I}'(v)$, we have $R'(u) = R'(v)$; thereby $\mathcal{I}'$ is proper.

The realizer $\mathcal{I}'$ can be computed as follows: scan the intervals of $\mathcal{I}$ according to their left boundary ordering and output the ordered set of vertices $S = \{x_1, \dots, x_i, \dots\}$ so that x_{i+1} has minimum left boundary among $\{x \mid L(x_i) < L(x) \text{ and } R(x_i) < R(x)\}$. Now for any vertex $v \notin S$, its enclosed neighbor is the maximum $x_i \in S$ such that $L(x_i) < L(v)$. Two scans of the vertices suffice to compute those enclosed neighbors. As by construction the vertices of S are maximal for the inclusion, they do not have enclosed neighbors. It follows that computing $\mathcal{I}'$ requires only $\mathcal{O}(n)$ time. $\square$

The main result of this subsection shows that distances in G can be retrieved from the distances in G' whose realizer $\mathcal{I}'$ is proper. To prove Theorem 4, the lemma below is required.

LEMMA 3. *If x and y are two vertices such that $R(x) = R'(x) \leq R(y)$, then $\text{dist}_G(x, y) = \text{dist}_{G'}(x, y)$.*

Proof. Let us first remark that since G is a subgraph of G' , then $\text{dist}_G(x, y) \geq \text{dist}_{G'}(x, y)$. So assume that $\text{dist}_G(x, y) = k > 1$. Let us first remark that any path of G' that does not use any extended vertex is also a path of G . Let $P = [x = u_0, u_1, \dots, u_j, \dots, u_k = y]$ be a shortest x, y -path in G' using a minimum number of extended vertices. Assume there exists some $j > 0$ such that u_j has been extended (i.e., $\mathcal{I}(u_j) \neq \mathcal{I}'(u_j)$) and choose j minimum. By assumption $\mathcal{I}(u_{j-1}) = \mathcal{I}'(u_{j-1})$ holds and implies that u_{j-1} and u_j are adjacent in G . Since $\mathcal{I}(u_j) \subset \mathcal{I}(N_e(u_j))$, substituting in P vertex u_j by vertex $N_e(u_j)$ yields a shortest x, y -path P' in G' using fewer extended vertices than P , which is a contradiction; P does not contain any extended vertex. $\square$

DEFINITION 3. *For any vertex x , the maximum neighbor of x , denoted by $N_m(x)$, is a neighbor of x such that $R(N_m(x))$ is maximum.*

Like enclosing neighbors, maximum neighbors are not extended in G' ; i.e., for any x , $R'(N_m(x)) = R(N_m(x))$. Otherwise we would have $\mathcal{I}(N_m(x)) \subset \mathcal{I}(v)$ for some vertex v and $R(N_m(x)) < R(v)$. Since $\mathcal{I}(x) \cap \mathcal{I}(N_m(x)) \neq \emptyset$, we also have $\mathcal{I}(x) \cap \mathcal{I}(v) \neq \emptyset$. Since $R(N_m(x)) < R(v)$, by definition, we should have $v = N_m(x)$, which is a contradiction.

Let us define $\text{adj}_G(x, y)$ as 1 if x and y are adjacent in G and as 0 otherwise.

THEOREM 4. *Let x and y be two vertices such that $R(x) \leq R(y)$. Then,*

$$(2) \quad \text{dist}_G(x, y) = \text{dist}_{G'}(N_m(x), y) + 1 - \text{adj}_G(x, y).$$

Proof. If x and y are adjacent, then $N_m(x)$ and y are also adjacent in both G and G' ; therefore, the result holds. So assume that x and y are not adjacent. Since $N_m(x)$ has not been extended in G' , by Lemma 3, $\text{dist}_{G'}(N_m(x), y) = \text{dist}_G(N_m(x), y)$. So we have to prove that $\text{dist}_G(x, y) = 1 + \text{dist}_G(N_m(x), y)$.

Let us first remark that $\text{dist}_G(N_m(x), y) \leq \text{dist}_G(x, y)$ since $R(x) < L(y)$ and $R(x) < R(N_m(x))$. The only case to be considered is $\text{dist}_G(N_m(x), y) = k > 1$. Let $P = [N_m(x), v_1, \dots, v_{k-1}, y]$ be a shortest $N_m(x), y$ -path. Let us assume that

$\text{dist}_G(x, y) = k$, then there exists a path $Q = [x, u_1, u_2, \dots, u_{k-1}, y]$ of length k . By definition of $N_m(x)$, we have $R(u_1) < R(N_m(x))$. It implies that $N_m(x)$ is adjacent to u_2 ; therefore, $[N_m(x), u_2, \dots, u_{k-1}, y]$ is a path of length $k - 1$, which is a contradiction. $\square$

Given a distance labeling scheme of a proper interval graph, (2) directly describes the decoder function of a distance labeling scheme for the family of interval graphs. The label $L(x, G)$ of each vertex should contain

- $R(x)$ and $L(x)$, the boundaries of $\mathcal{I}(x)$ that enable us to test $\text{adj}_G(x, y)$; and
- $L'(x, G')$ and $L'(N_m(x), G')$, the labels of x and $N_m(x)$ in the distance labeling scheme of the proper interval graph G' that enable us to compute $\text{dist}_{G'}(x, y)$.

COROLLARY 1. *If there exists a distance labeling scheme for n -vertex proper interval graphs using labels of length $s(n)$, then the n -vertex interval graph family enjoys a distance labeling scheme using labels of length $2 \lceil \log(2n) \rceil + 2s(n)$.*

3.2. Distance labeling scheme of proper interval graphs. Let us now assume that a proper realizer $\mathcal{I}$ for a graph $G = (V, E)$ is given (so G is a proper interval graph). Let x_0 be the vertex with the minimum left boundary. Notice that x_0 also has a minimum right boundary. We define the *layer partition* $V_0, V_1, \dots, V_k$ (see Figure 6) as the partition of vertices into distance layers from x_0 :

$$\forall i \geq 0, \quad V_i = \{v \mid \text{dist}_G(x_0, v) = i\}.$$

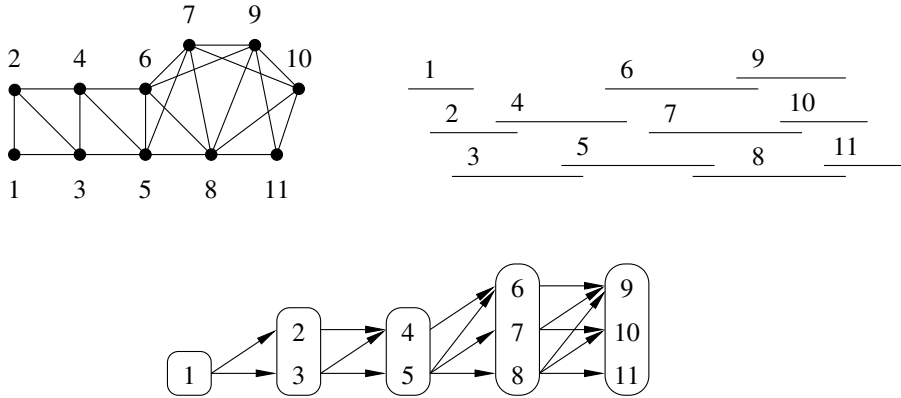


FIG. 6. A proper interval graph with an interval representation and the associated layer partition.

Given the sorted list of intervals, the layer partition can be computed in $\mathcal{O}(n)$ time. Let $\lambda(x)$ denote the unique index i such that $x \in V_i$. Let H be the digraph on the vertex set V composed of all the arcs xy such that $\lambda(x) < \lambda(y)$ and $(x, y) \in E$. Note that H is acyclic.

CLAIM 4. *The transitive closure H^t of H is a poset (partially ordered set).*

Let $\text{adj}_{H^t}(x, y)$ be the boolean function such that

$$\text{adj}_{H^t}(x, y) = 1 \text{ iff } xy \text{ is an arc of } H^t.$$

Notice that since H^t is a digraph, $\text{adj}_{H^t}(x, y) = 1$ implies that $\text{adj}_{H^t}(y, x) = 0$. Also by definition of H^t , if $\text{adj}_{H^t}(x, y) = 1$, then $\lambda(x) < \lambda(y)$ holds.

THEOREM 5. *For all distinct vertices x and y such that $\lambda(x) \leq \lambda(y)$,*

$$(3) \quad \text{dist}_G(x, y) = \lambda(y) - \lambda(x) + 1 - \text{adj}_{H^t}(x, y).$$

To prove Theorem 5, we need the following preliminary result.

LEMMA 4. *For all $i \geq 1$, let x_{i-1} be the vertex of V_{i-1} having maximum right boundary.⁴ Then $\{x_{i-1}\} \cup V_i$ is a clique of G .*

Proof. Let us first show that x_{i-1} dominates V_i . It holds for $i = 1$ by definition of the layer partition. Let us consider the case for larger i . First any vertex $v \in V_i$ has a neighbor in V_{i-1} and x_{i-1} has a maximum right boundary among V_{i-1} 's vertices. This implies that $L(x_{i-1}) < L(v) < R(x_{i-1})$; in other words, v neighbors x_{i-1} .

Let us show now that V_i induces a clique. Since $V_0 = \{x_0\}$, for any vertex $v \in V_1$, $\mathcal{I}(v)$ contains $R(x_0)$; both V_0 and V_1 are cliques. For larger i , we already proved that any $v \in V_i$ satisfies $L(x_{i-1}) < L(v) < R(x_{i-1})$. Since the realizer of G is assumed to be proper, we also have $L(v) < R(x_{i-1}) \leq R(v)$. Therefore, for any $v \in V_i$, $\mathcal{I}(v)$ contains $R(x_{i-1})$. This implies that V_i is a clique. $\square$

Proof of Theorem 5. Let x and y be two vertices such that $\lambda(x) = i$ and $\lambda(y) = i + h$ with $h \geq 0$. It is obvious that $h \leq \text{dist}_G(x, y)$. Note that by definition of the layer partition, the statement holds for $i = 0$.

Let us consider $i > 0$. Lemma 4 implies the result in cases $h = 0$ and $h = 1$. So assume the statement holds for $h - 1 \geq 0$, and let us prove it for h . By Lemma 4, x is adjacent to x_i and x_{i+h-1} is adjacent to y . Therefore, $[x, x_i, \dots, x_{i+h-1}, y]$ is a path of length $h + 1$ and $\text{dist}_G(x, y) \leq \lambda(y) - \lambda(x) + 1$. Inequality holds iff x is adjacent to some vertex $z \in V_{i+1}$ such that $\text{dist}_G(z, y) = h - 1$. In that case, we have $\text{adj}_{H^t}(x, z) = 1$, and, by induction, hypothesis $\text{adj}_{H^t}(z, y) = 1$ holds. Since H^t is transitive, $\text{dist}_G(x, y) = h$ iff $\text{adj}_{H^t}(x, y) = 1$. This ends the proof. $\square$

Therefore, to compute the distance between any pair of vertices, we have to test whether $\text{adj}_{H^t}(x, y) = 1$. For this reason H^t can be considered as the *graph of errors* associated with the layer partition. Let us now prove that in H^t the adjacency can be encoded using short labels (namely, $2 \log n$ bits labels).

THEOREM 6. *There exists a total ordering π of the vertices such that*

$$(4) \quad \text{adj}_{H^t}(x, y) = 1 \text{ iff } \lambda(x) < \lambda(y) \text{ and } \pi(y) < \pi(x).$$

Moreover, given H^t and $\mathcal{I}$, the ordering π can be computed in $\mathcal{O}(n)$ time.

Proof. We first claim that such an ordering π is obtained by an elimination ordering of the vertices such that at each step, the sink of the subgraph of H induced by the noneliminated vertices that has a minimum left bound is removed. This can be implemented by using a depth-first search of H : the ordering π is the postorder traversal of the search (i.e., the order in which the vertices have been removed from the LIFO-stack).

In the example in Figure 6, we obtain the following ordering:

$$\pi = 9, 6, 4, 2, 10, 7, 11, 8, 5, 3, 1.$$

It is easy to check that if $\text{adj}_{H^t}(x, y) = 1$, then $\lambda(x) < \lambda(y)$ and $\pi(y) < \pi(x)$. So assume that $\text{adj}_{H^t}(x, y) = 0$ and $\lambda(x) < \lambda(y)$ (which implies $L(x) < L(y)$).

⁴For $i > 0$, x_i is the maximum neighbor of x_{i-1} and $x_0 \in V_0$.

CLAIM 5. *Let z be a vertex such that $\lambda(y) < \lambda(z)$. If $\text{adj}_{H^t}(x, z) = 1$, then $\text{adj}_{H^t}(y, z) = 1$.*

Since $\text{adj}_{H^t}(x, z) = 1$, the graph G contains a x, z -shortest path of length $\lambda(z) - \lambda(x)$. Let $P = [x, \dots, v_{\lambda(y)-1}, v_{\lambda(y)}, v_{\lambda(y)+1}, \dots, z]$ be such a path with $v_i \in V_i$ for $i \in \{\lambda(y) - 1, \lambda(y), \lambda(y) + 1\}$ and possibly $x = v_{\lambda(y)-1}$ and $z = v_{\lambda(y)+1}$. Note that $\text{adj}_{H^t}(x, y) = 0$ implies that $\text{adj}_{H^t}(v_{\lambda(y)-1}, y) = 0$. It follows that $v_{\lambda(y)-1}$ and y are nonadjacent in G . Now if $\text{adj}_{H^t}(y, z) = 0$, it similarly implies that $\text{adj}_{H^t}(y, v_{\lambda(y)+1}) = 0$, and thus y and $v_{\lambda(y)+1}$ are also nonadjacent vertices of G . As P is a shortest path, $v_{\lambda(y)-1}$ and $v_{\lambda(y)+1}$ are nonadjacent in G . As by construction $\lambda(v_{\lambda(y)}) = \lambda(y)$, y is adjacent in G to $\lambda(v_{\lambda(y)})$. It follows that in G the set of vertices $\{y, v_{\lambda(y)-1}, v_{\lambda(y)}, v_{\lambda(y)+1}\}$ induce a $K_{1,3}$, which is a contradiction; G is a proper interval graph. Claim 5 follows: if $\text{adj}_{H^t}(x, z) = 1$, then $\text{adj}_{H^t}(y, z) = 1$.

Note also that at the step when y becomes a sink, there remains no vertex z such that $\lambda(y) < \lambda(z)$ and $\text{adj}_{H^t}(x, z) = 1$. If x is also a sink, since $L(x) < L(y)$, the priority rule implies $\pi(x) < \pi(y)$, and we are done. So assume x is not a sink. Therefore, there exists z such that $\text{adj}_{H^t}(x, z) = 1$. By Claim 5, z satisfies $\lambda(z) \leq \lambda(y)$. Moreover, we show that $L(z) < L(y)$. Otherwise z and y would both belong to $V_{\lambda(y)}$, and a shortest path of length $\lambda(y) - \lambda(x)$ between x and z would imply a shortest path of the same length between x and y , contradicting $\text{adj}_{H^t}(x, y) = 0$. Finally, both z and y are sinks, but $L(z) < L(y)$ implies that z is eliminated before y . Recursively applying this argument enables us to prove that x will be eliminated before y by the algorithm, and so $\pi(x) < \pi(y)$.

Let us now analyze the time complexity of the computation of the π ordering. Notice that the digraph H can be stored in $\mathcal{O}(n)$ space. The set of vertices is sorted in an array with respect to the left boundary of their interval. Since the layers and also the neighborhood in H of any vertex appear consecutively in this array, each vertex can be associated with its first and last neighbor in the array. The priority rule implies that when a vertex v is popped by the algorithm, any vertex $u \in V_{\lambda(v)}$ such that $L(u) < L(v)$ has already been popped. Assuming that the layer of any vertex is stored and that for each layer, the index of the last popped vertex is maintained, the next vertex to be pushed can be found in $\mathcal{O}(1)$. Therefore, π can be computed in $\mathcal{O}(n)$ time. $\square$

In passing, let us point out an interesting combinatorial property of poset H^t . The *dimension* of a poset P on a set V is the minimum number d of total orders $\rho_1, \dots, \rho_d$ on V such that $x <_P y$ iff $x <_{\rho_i} y$ for every i [11, 36]. By choosing $\rho_1 = \pi$ and ρ_2 as any linear extension of λ , Theorem 6 immediately implies the following corollary.

COROLLARY 2. *H^t is a poset of dimension two.*

We now have the material to define the distance labeling scheme for proper interval graphs. To each vertex x of G , we assign the label $L(x, G) = \langle \lambda(x), \pi(x) \rangle$. The distance decoder is given by (3) and (4). Clearly, $\lambda(x), \pi(x) \in \{1, \dots, n\}$, so the main result follows.

THEOREM 7. *The family of n -vertex proper interval graphs enjoys a distance labeling scheme using labels of length $2 \lceil \log n \rceil$, and the distance decoder has constant time complexity. Moreover, given the sorted list of intervals, all the labels can be computed in $\mathcal{O}(n)$ time.*

3.3. Optimized distance labeling scheme for interval graphs. According to Corollary 1 and Theorem 7, the family of interval graphs enjoys a $(6 \log n + \mathcal{O}(1))$ -distance labeling scheme with constant time decoder. We shall see that these labels can be compacted to $5 \log n + \mathcal{O}(1)$ bits. Let us recall that for any vertex x of an

interval graph G , we have to know $R(x)$ and $L(x)$, the boundaries of the interval $\mathcal{I}(x)$, plus the labels of x and $N_m(x)$ in the extended proper interval graph G' .

The improvement is based on the observation that $\lambda(x)$ and $\lambda(N_m(x))$ differ at most by one ($\lambda(N_m(x)) - \lambda(x) \in \{0, 1\}$). Indeed, there is no edge in G between a pair of vertices that does not belong to the same distance layer or consecutive distance layers. Therefore, a single bit, say, $b(x)$, suffices to recover $\lambda(N_m(x))$ from $\lambda(x)$. Therefore, the label of a vertex x can be defined as follows:

$$L(x, G) = \langle L(x), R(x), \lambda(x), \pi(x), b(x), \pi(N_m(x)) \rangle.$$

As $1 \leq L(x) < R(x) \leq 2n$, so $2 \lceil \log n \rceil + 2$ bits suffice to store the first two integers. The next four components can be stored with $3 \lceil \log n \rceil + 1$ bits, summing up to $5 \lceil \log n \rceil + 3$. Thus, we have the following theorem.

THEOREM 8. *The family of n -vertex interval graphs enjoys a distance labeling scheme using labels of length $5 \lceil \log n \rceil + 3$, and the distance decoder has constant time complexity. Moreover, given the sorted list of intervals, all the labels can be computed in $\mathcal{O}(n)$ time.*

Proof. We have already discussed the length of the labels. The distance decoder takes a constant time. Concerning the time complexity of constructing all the labels, let us first observe that in $\mathcal{O}(n)$ time one can compute $N_m(x)$ for each x as follows: Since $\mathcal{I}'$ is proper, a maximum neighbor of a given vertex x is precisely the vertex y with maximum left boundary satisfying $L(y) \leq R(x)$. Since we made the assumption that the intervals have been given sorted with respect to their left boundaries, a single scan enables us to compute the maximum neighbor of each vertex.

Since the distance labels in proper interval graphs can be computed in $\mathcal{O}(n)$ time, and since G' can be obtained within the same complexity, all the labels of G can be computed in $\mathcal{O}(n)$ time. $\square$

3.4. A scheme for circular-arc graphs. Circular-arc graphs are a natural generalization of interval graphs in which vertices are mapped to arcs of a circle rather than intervals on the real line. This section shows that the distance labeling scheme problem for circular-arc graphs reduces to the interval graphs case. Consider a circular-arc graph G . We can associate with every vertex x of G a range $I(x) = [\theta_l(x), \theta_r(x)]$ of angles where $\theta_l(x)$ and $\theta_r(x)$ are taken clockwise in $[0, 2\pi)$. The range $I(x)$ can be seen as a “cyclic” interval in which $[\theta_l(x), \theta_r(x)]$ for $\theta_l(x) > \theta_r(x)$ stands for $[\theta_l(x), 2\pi) \cup [0, \theta_r(x)]$. The vertices x and y are adjacent if $I(x) \cap I(y) \neq \emptyset$. See an example in Figure 7.

A neighbor y of x is a *right neighbor* (resp., *left neighbor*) iff $\theta_l(y) \in I(x)$ and $\theta_r(y) \notin I(x)$ (resp., $\theta_l(y) \notin I(x)$ and $\theta_r(y) \in I(x)$). Notice that some neighbor of x may be neither a right nor a left neighbor. A path $x_0, x_1, \dots, x_k$ is a *right-path* from x_0 to x_k if, for every $i \in \{0, \dots, k-1\}$, x_{i+1} is a right neighbor of x_i . Intuitively, a right-path is obtained starting from x_0 and successively clockwise encountering the vertices $x_1, \dots, x_k$.

LEMMA 5. *Every shortest path between nonadjacent vertices x and y is either a right-path from x to y or a right-path from y to x .*

Proof. Assume that $P = [x = x_0, x_1, \dots, x_k = y]$ is a shortest path from x to y but is neither a right-path from x_0 to x_k nor a right-path from x_k to x_0 .

W.l.o.g. assume that x_1 is a right neighbor of x_0 . The case that x_1 is not a right neighbor of x_0 can be treated similarly by exchanging the roles of θ_r and θ_l boundaries. Since the path from x_0 to x_k is not a right-path, there is an integer i , $1 \leq i < k$, such

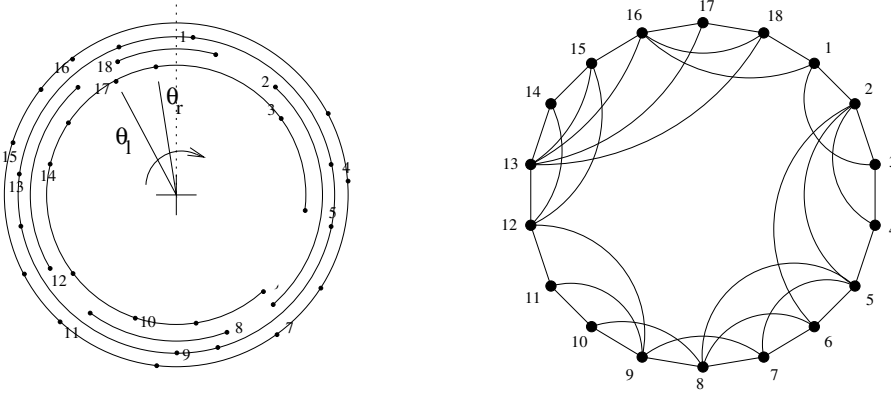


FIG. 7. A circular-arc graph with its arc representation. Vertices are labeled from 1 to 18 according to their value θ_l .

that x_{i+1} is not a right neighbor of x_i and such that the subpath from x_0 to x_i is a right-path.

- If x_{i+1} is not a left neighbor of x_i , then either $I(x_i)$ is included in $I(x_{i+1})$ or vice-versa. In both cases, the vertex with the smallest range can be removed from the path, henceforth proving the existence of a shorter path, which is a contradiction.
- If x_{i+1} is a left neighbor of x_i , then $\theta_l(x_i) \in I(x_{i-1}) \cap I(x_{i+1})$, proving that x_{i-1} and x_{i+1} are adjacent. By the way, the path P can be shortened to $[x = x_0, \dots, x_{i-1}, x_{i+1}, \dots, x_k = y]$, which is a contradiction. $\square$

Let $x_1, \dots, x_n$ be the vertices of G ordered such that $\theta_l(x_i) \leq \theta_l(x_{i+1})$ for every $i \in \{1, \dots, n-1\}$. We associate with G a new intersection graph $\tilde{G}$ with vertex set $V(\tilde{G}) = \bigcup_{1 \leq i \leq n} \{I(x_i^1), I(x_i^2)\}$, where, for every $1 \leq i \leq n$,

$$I(x_i^1) = \begin{cases} [\theta_l(x_i), \theta_r(x_i)] & \text{if } \theta_l(x_i) \leq \theta_r(x_i), \\ [\theta_l(x_i), \theta_r(x_i) + 2\pi] & \text{if } \theta_l(x_i) > \theta_r(x_i), \end{cases}$$

$$I(x_i^2) = \begin{cases} [\theta_l(x_i) + 2\pi, \theta_r(x_i) + 2\pi] & \text{if } \theta_l(x_i) \leq \theta_r(x_i), \\ [\theta_l(x_i) + 2\pi, \theta_r(x_i) + 4\pi] & \text{if } \theta_l(x_i) > \theta_r(x_i). \end{cases}$$

Intuitively, $\tilde{G}$ is obtained from G by listing all the intervals according to increasing angle θ_l , starting from the arc of x_1 , and turning clockwise around twice. So, each vertex x_i in G appears twice in $\tilde{G}$ as x_i^1 and x_i^2 . It is straightforward to check that $\tilde{G}$ is an interval graph.

LEMMA 6. For every $i < j$, $\text{dist}_G(x_i, x_j) = \min \{ \text{dist}_{\tilde{G}}(x_i^1, x_j^1), \text{dist}_{\tilde{G}}(x_j^1, x_i^2) \}$.

Proof. Let $P = [x_i, x_{t_1}, \dots, x_{t_k}, x_j]$ be a shortest path between x_i and x_j in G . From Lemma 5 P is a right-path from x_i to x_j or a right-path from x_j to x_i .

In the former case, we have $i \leq t_1 \leq \dots \leq t_k \leq j$. Consider the path $P' = [x_i^1, x_{t_1}^1, \dots, x_{t_k}^1, x_j^1]$ of $\tilde{G}$. Because P' can be constructed from P by turning clockwise from x_i to x_j in G , P' is isomorphic to P , showing that $\text{dist}_G(x_i, x_j) \leq \text{dist}_{\tilde{G}}(x_i^1, x_j^1)$. We observe that any shortest path from x_i^1 to x_j^1 in $\tilde{G}$ must be a right-path. By construction, every right-path of length l between a given pair of vertices in $\tilde{G}$ is isomorphic to some right-path of same length between the corresponding vertices

in G . So a shorter path between x_i^1 and x_j^1 would provide a shorter path between x_i and x_j in G . Therefore, P' is a shortest path in $\tilde{G}$ and $\text{dist}_G(x_i, x_j) = \text{dist}_{\tilde{G}}(x_i^1, x_j^1)$.

Assume now that P is a right-path from x_j to x_i . P is obtained by turning clockwise from x_j to x_i . P is isomorphic to a right-path P'' in $\tilde{G}$ between x_j^1 and x_i^2 . Indeed, since $i < j$, $\theta_l(x_i) \leq \theta_l(x_j) < \theta_l(x_i) + 2\pi$. Similarly to the previous case, P'' is a shortest path, proving that, in this case, $\text{dist}_G(x_i, x_j) = \text{dist}_{\tilde{G}}(x_j^1, x_i^2)$ and completing the proof of the lemma. $\square$

Therefore, a distance labeling scheme for interval graphs can be transformed into a scheme for a circular-arc graph family by doubling the number of vertices and the label length.

THEOREM 9. *The family of n -vertex circular-arc graphs enjoys a distance labeling scheme using labels of length $\mathcal{O}(\log n)$, and the distance decoder has constant time complexity. Moreover, given the sorted list of intervals, all the labels can be computed in $\mathcal{O}(n)$ time.*

4. Universal distance matrix. Given a family $\mathcal{F}$ of graphs, [6, 1] considered the size of the smallest *induced universal graph* to give a measure of the density of $\mathcal{F}$. A graph $\mathcal{G}$ is an induced universal graph with respect to $\mathcal{F}$ if any graph of $\mathcal{F}$ is an induced subgraph of $\mathcal{G}$. In this section, we similarly define the *universal distance matrix* of a family $\mathcal{F}$, which is a square matrix $\mathcal{U}_n$ containing the distance matrix of every (unweighted) n -vertex graph of $\mathcal{F}$ as an *induced submatrix*. A matrix $B = (B_{i,j})_{1 \leq i,j \leq p}$ is an induced submatrix of a matrix $A = (A_{i,j})_{1 \leq i,j \leq q}$ if there exists a sequence $(s_1, \dots, s_p)$ of indices such that $B_{i,j} = A_{s_i, s_j}$ for all $i, j \in \{1, \dots, p\}$. Intuitively, a small universal distance matrix for $\mathcal{F}$ indicates that many distance patterns repeat in many different graphs of $\mathcal{F}$.

Proposition 1 will be a tool for proving lower and upper bounds on the dimension of universal distance matrix. The proof technique is similar to the proof given by [19] for the relationship between adjacency schemes and induced universal graphs [1, 6].

PROPOSITION 1. *If $\mathcal{F}$ enjoys an ℓ -distance labeling scheme, then $\mathcal{F}$ has a universal distance matrix of dimension $2^{\ell+1} - 1$. If $\mathcal{F}$ has a universal distance matrix of dimension d , then $\mathcal{F}$ enjoys a $\lfloor \log d \rfloor$ -distance labeling scheme.*

Proof. For every integer $i \geq 1$, let $\phi(i)$ denote the i th binary string in the lexicographic order $\phi(1) = \epsilon$, $\phi(2) = 0$, $\phi(3) = 1$, $\phi(4) = 00$, $\phi(5) = 01$, $\phi(6) = 10$, $\phi(7) = 11$, $\phi(8) = 000, \dots$. We note that the length of $\phi(i)$ is $\lfloor \log i \rfloor$. We denote by ϕ^{-1} the inverse function of ϕ .

A universal distance matrix M for $\mathcal{F}$ can be defined from any distance labeling scheme $\langle L, f \rangle$ for $\mathcal{F}$ as follows: $M_{i,j} = f(\phi(i), \phi(j))$ if there exists a graph $G \in \mathcal{F}$ with two vertices u and v such that $L(u, G) = \phi(i)$ and $L(v, G) = \phi(j)$, and $M_{i,j} = \infty$ otherwise.

With every $G \in \mathcal{F}$, one can associate the subset $I_G = \{\phi^{-1}(L(u, G)) \mid u \in V(G)\}$ of indices. By construction, for all $i, j \in I_G$, $M_{i,j} = f(\phi(i), \phi(j))$, which is the distance in G between the vertices labeled $\phi(i)$ and $\phi(j)$. So M contains as an induced submatrix the distance matrix of every $G \in \mathcal{F}$: M is universal. Moreover, its dimension is smaller than the largest integer contained in a subset I_G , which cannot exceed the number of binary strings of length at most ℓ , i.e., $2^{\ell+1} - 1$.

Now, assume M is a universal distance matrix of dimension d for $\mathcal{F}$. Then a distance labeling scheme $\langle L, f \rangle$ can be constructed as follows: if the distance matrix of G is the induced submatrix of M with a sequence of indices I_G , then we label the vertex corresponding to the index $i \in I_G$ by the binary string $\phi(i)$. The length of

this label is the length of $\phi(i)$. The length is therefore bounded by the length of $\phi(d)$, that is, $\lfloor \log d \rfloor$ bits. Now, given two binary labels λ and λ' , extracted from the same graph, one can compute the distance by searching in M (which depends only on $\mathcal{F}$) the entry $M_{\phi^{-1}(\lambda), \phi^{-1}(\lambda')}$. As the two labels are extracted from the same graph, this value is the distance between the corresponding vertices. $\square$

From Proposition 1 and previously known bounds on the distance labeling scheme [15], trees have a universal distance matrix of dimension $2^{\Theta(\log^2 n)}$, whereas any universal distance matrix for cubic graphs must be of dimension at least $2^{\Omega(\sqrt{n})}$. By the way, the dimension of the smallest universal distance matrix can be considered as a complexity measure of a graph family. From Proposition 1 and Theorems 3 and 2 we obtain the following corollaries.

COROLLARY 3. *The dimension of the smallest universal distance matrix of n -vertex proper interval graphs is at most $\mathcal{O}(n^2)$ and at least $\Omega(n^2/\log^2 n)$.*

COROLLARY 4. *The dimension of the smallest universal distance matrix of n -vertex interval graphs is at most $\mathcal{O}(n^5)$ and at least $\Omega(n^3/\log^4 n)$.*

REFERENCES

- [1] S. ALSTRUP AND T. RAUHE, *Improved labeling scheme for ancestor queries*, in Proceedings of the 13th ACM-SIAM Symposium on Discrete Algorithms (SODA), ACM, New York, SIAM, Philadelphia, 2002, pp. 947–953.
- [2] F. BAZZARO AND C. GAVOILLE, *Localized and compact data-structure for comparability graphs*, Discrete Math., (2008), to appear.
- [3] O. BERKMAN AND U. VISHKIN, *Finding level-ancestors in trees*, J. Comput. System Sci., 48 (1994), pp. 214–230.
- [4] A. BRANDSTÄDT, V.B. LE, AND J. SPINRAD, *Graph Classes, A Survey*, SIAM Monogr. Discrete Math. Appl. 3, SIAM, Philadelphia, 1999.
- [5] D.Z. CHEN, D.T. LEE, R. SRIDHAR, AND C.N. SEKHARAN, *Solving the all-pair shortest path query problem on interval and circular-arc graphs*, Networks, 31 (1998), pp. 249–257.
- [6] F.R.K. CHUNG, *Universal graphs and induced-universal graphs*, J. Graph Theory, 14 (1990), pp. 443–454.
- [7] J.E. COHEN, J. KOMLÓS, AND T. MUELLER, *The probability of an interval graph, and why it matters*, in Relations Between Combinatorics and Other Parts of Mathematics, Sympos. Pure Math. 34, AMS, Providence, RI, 1979, pp. 97–115.
- [8] D.G. CORNEIL, H. KIM, S. NATARAJAN, S. OLARIU, AND A.P. SPRAGUE, *Simple linear time algorithm of unit interval graphs*, Inform. Process. Lett., 55 (1995), pp. 99–104.
- [9] B. COURCELLE AND R. VANICAT, *Query efficient implementation of graphs of bounded clique width*, Discrete Appl. Math., 131 (2003), pp. 129–150.
- [10] C.M. HERRERA DE FIGUEIREDO, J. MEIDANIS, AND C. PICININ DE MELLO, *A linear-time algorithm for proper interval recognition*, Inform. Process. Lett., 56 (1995), pp. 179–184.
- [11] B. DUSHNIK AND E.W. MILLER, *Partially ordered sets*, Amer. J. Math., 63 (1941), pp. 600–610.
- [12] C. GAVOILLE, M. KATZ, N.A. KATZ, C. PAUL, AND D. PELEG, *Approximate distance labeling schemes*, in Proceedings of the 9th Annual European Symposium on Algorithms (ESA), Lecture Notes in Comput. Sci. 2161, Springer-Verlag, New York, 2001, pp. 476–488.
- [13] C. GAVOILLE AND C. PAUL, *Distance labeling scheme and split decomposition*, Discrete Math., 273 (2003), pp. 115–130.
- [14] C. GAVOILLE AND D. PELEG, *Compact and localized distributed data structures*, Distributed Computing, 16 (2003), pp. 111–120.
- [15] C. GAVOILLE, D. PELEG, S. PÉRENNES, AND R. RAZ, *Distance labeling in graphs*, in Proceedings of the 12th ACM-SIAM Symposium on Discrete Algorithms (SODA), ACM, New York, SIAM, Philadelphia, 2001, pp. 210–219.
- [16] A. GUPTA, A. KUMAR, AND R. RASTOGI, *Traveling with a pez dispenser (or, routing issues in MPLS)*, SIAM J. Comput., 34 (2005), pp. 453–474.
- [17] P. HANLON, *Counting interval graphs*, Trans. Amer. Math. Soc., 272 (1982), pp. 383–426.
- [18] P. HELL, J. BANG-JENSEN, AND J. HUANG, *Local tournaments and proper circular arc graphs*, in Algorithms, Lecture Notes in Comput. Sci. 450, Springer-Verlag, Berlin, 1990, pp. 101–108.
- [19] S. KANNAN, M. NAOR, AND S. RUDICH, *Implicit representation of graphs*, SIAM J. Discrete Math., 5 (1992), pp. 596–603.

- [20] M. KATZ, N.A. KATZ, AND D. PELEG, *Distance labeling schemes for well-separated graph classes*, in Proceedings of the 17th Annual Symposium on Theoretical Aspects of Computer Science (STACS), Lecture Notes in Comput. Sci. 1770, Springer-Verlag, Berlin, 2000, pp. 516–528.
- [21] A. KORMAN, *General compact labeling schemes for dynamic trees*, in Proceedings of the 19th International Symposium on Distributed Computing (DISC), Lecture Notes in Comput. Sci. 3724, Springer-Verlag, Berlin, 2005, pp. 457–471.
- [22] A. KORMAN AND D. PELEG, *Labeling schemes for weighted dynamic tree*, in Proceedings of the 30th International Colloquium on Automata, Languages and Programming (ICALP), Lecture Notes in Comput. Sci. 2719, Springer-Verlag, Berlin, 2003, pp. 369–383.
- [23] A. KORMAN, D. PELEG, AND Y. RODEH, *Labeling schemes for dynamic tree networks*, in Proceedings of the 19th Annual Symposium on Theoretical Aspects of Computer Science (STACS), Lecture Notes in Comput. Sci. 2285, Springer-Verlag, Berlin, 2002, pp. 76–87.
- [24] A. KORMAN, D. PELEG, AND Y. RODEH, *Labeling schemes for dynamic tree networks*, Theory Comput. Syst., 37 (2004), pp. 49–75.
- [25] M. MENDEL AND S. HAR-PELED, *Fast construction of nets in low-dimensional metrics and their applications*, SIAM J. Comput., 35 (2006), pp. 1148–1184.
- [26] D. PELEG, *Informative labeling schemes for graphs*, in Proceedings of the 25th International Symposium on Mathematical Foundations of Computer Science (MFCS), Lecture Notes in Comput. Sci. 1893, Springer-Verlag, Berlin, 2000, pp. 579–588.
- [27] D. PELEG, *Proximity-preserving labeling schemes*, J. Graph Theory, 33 (2000), pp. 167–176.
- [28] F.S. ROBERTS, *Indifference graphs*, in Proof Techniques in Graph Theory, Academic Press, New York, 1969, pp. 139–146.
- [29] A. SLIVKINS, *Distance estimation and object location via rings of neighbors*, in Proceedings of the 24th Annual ACM Symposium on Principles of Distributed Computing (PODC), ACM, New York, 2005, pp. 41–50.
- [30] J. P. SPINRAD, *Efficient Graph Representations*, Fields Inst. Monogr. 19, AMS, Providence, RI, 2003.
- [31] K. TALWAR, *Bypassing the embedding: Algorithms for low dimensional metrics*, in Proceedings of the 36th Annual ACM Symposium on Theory of Computing (STOC), ACM, New York, 2004, pp. 281–290.
- [32] M. THORUP, *Compact oracles for reachability and approximate distances in planar digraphs*, in Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science (FOCS), IEEE Computer Society, Los Alamitos, CA, 2001, pp. 242–251.
- [33] M. THORUP, *Compact oracles for reachability and approximate distances in planar digraphs*, J. ACM, 51 (2004), pp. 993–1024.
- [34] M. THORUP AND U. ZWICK, *Approximate distance oracles*, in Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC), ACM, New York, 2001, pp. 183–192.
- [35] M. THORUP AND U. ZWICK, *Approximate distance oracles*, J. ACM, 52 (2005), pp. 1–24.
- [36] W. T. TROTTER, *Combinatorics and Partially Ordered Sets: Dimension Theory*, The Johns Hopkins University Press, Baltimore, MD, 1992.
- [37] G. WEGNER, *Eigenschaften der Neven homologisch-einfacher Familien im R^n* , Ph.D. thesis, University of Göttingen, Göttingen, Germany, 1967.

A UNIFIED VIEW OF GRAPH SEARCHING*

DEREK G. CORNEIL[†] AND RICHARD M. KRUEGER[†]

Abstract. Graph searching is perhaps one of the simplest and most widely used tools in graph algorithms. Despite this, few theoretical results are known about the vertex orderings that can be produced by a specific search algorithm. A simple characterizing property, such as is known for LexBFS, can aid greatly in devising algorithms, writing proofs of correctness, and showing impossibility results. This paper unifies our view of graph search algorithms by showing simple, closely related characterizations of various well-known search paradigms, including BFS and DFS. Furthermore, these characterizations naturally lead to other search paradigms, namely, maximal neighborhood search and LexDFS.

Key words. graph searching, graph traversal, depth-first search, breadth-first search, LexBFS

AMS subject classifications. 05C85, 68R10, 68W05

DOI. 10.1137/050623498

1. Introduction. Graph searching is fundamental. Most graph algorithms employ some mechanism for systematically visiting all vertices and edges in the given graph. After choosing an initial vertex, a search of a connected graph visits each of the vertices and edges of the graph such that a new vertex is visited only if it is adjacent to some previously visited vertex. At any point there may be several vertices that may possibly be visited next; this generic graph search does not specify which vertex to choose. The *breadth-first search* (BFS) and *depth-first search* (DFS) algorithms are the traditional strategies for determining the next vertex to visit (see [7] for a recent overview of these searches). BFS traverses an untraversed edge incident with the least recently visited vertex, whereas DFS traverses an untraversed edge incident with the most recently visited vertex. Both searches were used in the 19th century to solve maze traversals. More recently, DFS, as popularized by Tarjan [14], has been used for such problems as graph connectivity, planarity, topological ordering, and strong connectivity of digraphs. Applications of BFS include shortest path problems, network flows, and recognition of various graph classes.

In the mid 1970s, Rose, Tarjan, and Lueker [11] introduced a variant of BFS for determining perfect elimination orderings in chordal graphs. Their *lexicographic breadth-first search* (LexBFS) has garnered much attention recently and has become an important search algorithm used for recognition of various graph families, diameter approximation, and finding key structures in certain graph classes [3]. One of the key points regarding LexBFS, and one of the reasons for its usefulness, is that there is a very simple characterization of LexBFS orderings. This characterization is critical to the proofs of correctness of various algorithms. In light of the power of such a tool, it is somewhat surprising that the traditional searches of BFS and DFS are not known to have similar characterizations. Other than LexBFS, only the *maximum cardinality search* (MCS) of Tarjan and Yannakakis [15] has a known characterization [1].

*Received by the editors January 28, 2005; accepted for publication (in revised form) March 26, 2008; published electronically July 25, 2008. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Ontario Graduate Scholarship Program.
<http://www.siam.org/journals/sidma/22-4/62349.html>

[†]Department of Computer Science, University of Toronto, Toronto, Ontario M5S 3G4, Canada (dgc@cs.toronto.edu, krueger@cs.toronto.edu).

To provide a common underpinning for viewing graph searches, we examine what is meant by a “generic graph search.” We describe a graph search similarly to Tarjan [14]. Given a possibly disconnected graph G whose vertices are initially unexplored, we start from some vertex s of G and choose one of its edges to follow. Traversing this edge leads to a new vertex. We continue, at each step selecting an untraversed edge leading from a vertex already reached. Traversing this edge leads either to a new vertex or to one previously reached. Whenever we run out of edges leading from old vertices, we choose some new unreached vertex and begin again from there. Each edge is traversed exactly once, and we eventually reach all vertices in the graph. How we choose amongst all the edges we may next traverse determines the type of search we perform (i.e., whether it is a BFS or a DFS). The order in which we reach the vertices exhibits very different properties depending on this selection rule.

From a characterization point of view, we concentrate on the following situation. Suppose in the ordering of vertices we have three vertices, a , b , and c , where a was reached before b , which was reached before c , and we have an edge ac present in our graph, but the edge ab is absent (see Figure 1.1). The question is, “how could vertex b have been chosen before vertex c by our search?” With what we know so far about this situation, it seems that c should have been chosen before b . What can we infer about the graph such that the search could have chosen b before c ?

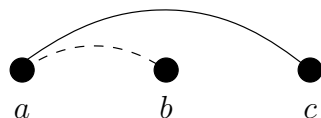


FIG. 1.1. A curious triple of vertices in a search ordering.

In the case of LexBFS, the answer is known. There must exist another vertex d chosen earlier than a such that db is an edge, but dc is not an edge of G [11, 8]. In fact, Brandstädt, Dragan, and Nicolai [1] showed that this condition is a characterization of a vertex ordering being generated by LexBFS.

Experience with LexBFS has shown that such a characterizing property is instrumental in the applicability of the search to various problems. For example, the LexBFS characterization is fundamental in proving that the reverse of a LexBFS ordering of a chordal graph is a perfect elimination ordering [11]. Furthermore, many of the recent LexBFS-based algorithms use multiple sweeps of LexBFS to reveal structure of the given graph (see [3] for several examples). Each subsequent search uses information from the previous sweep(s) to break ties of which vertex to visit next. The related proofs of correctness rely heavily on the LexBFS characterization. The characterization can also be used to show that there is some LexBFS ordering that certifies a graph as being cocomparability, even though we don’t know how to find it efficiently [4].

In light of this simple characterization result for LexBFS (and similar results for MCS) and recognizing its usefulness, it is somewhat surprising that it appears that this question has never been asked regarding the more traditional search paradigms of BFS and DFS. In section 2 we further explore the question “how could b have been chosen before c by a graph search?” We present answers to this question for each of the known search paradigms and show that those answers in fact characterize the vertex orderings realizable by the graph search. We will also introduce two new search paradigms, a depth-first version of LexBFS we call LexDFS, and a generalization we

call maximal neighborhood search (MNS). To illustrate the usefulness of these new characterizations, we examine in section 3 which searches of a given graph G are valid searches of powers of G . In particular, we would like to obtain a search ordering of G^k without the necessity of computing G^k (i.e., in time linear in the size of G). We conclude with some remarks about the impact of these new characterizations and their extension to other areas.

1.1. Notation and definitions. We now define some notation to facilitate our discussion. We let $G = (V, E)$ be a graph on vertex set V and edge set E . We will assume, for simplicity, that all graphs in this paper are connected and undirected; extension of our results to disconnected graphs is straightforward. Let $n = |V|$. We define $\sigma = (v_1, \dots, v_n)$ to be a linear ordering of V , which will typically be the left-to-right order in which we number the vertices in a search. We let $\sigma_i = (v_1, \dots, v_i)$ be a prefix of the ordering σ for $1 \leq i \leq n$. We use the notation $\sigma(i) = v_i$ and $\sigma^{-1}(v_i) = i$ to refer to a specific vertex in σ . The ordering σ implicitly imposes a total order on V ; we say $v_1 <_\sigma v_2 <_\sigma \dots <_\sigma v_n$ and omit the subscript σ when the ordering is clear. We let $G_i = G[v_1, \dots, v_i]$ be the subgraph of G induced on a prefix of σ , and let $N_H(v)$ be the neighborhood of v in a subgraph H of G . When referring to searches, $s = v_1$ is the start vertex.

A graph is said to be *chordal* if every cycle of length at least 4 has at least one chord. A vertex is *simplicial* in a graph G if its neighborhood is a clique in G . An ordering $(v_1, \dots, v_n)$ of V is a *perfect elimination ordering* of G if v_i is simplicial in G_i . A graph is chordal if and only if it has a perfect elimination ordering [6]. A graph is *bipartite* if its vertex set can be partitioned into two disjoint sets X, Y such that all edges are of the form xy for some $x \in X$ and $y \in Y$. A graph is *chordal bipartite* if it is bipartite and every cycle of length at least 6 has a chord. The *complement* of a graph $G = (V, E)$, written $\overline{G} = (V, \overline{E})$, is defined by $\overline{E} = \{xy : x, y \in V \text{ with } x \neq y \text{ and } xy \notin E\}$. The k th *power* of a graph G , written $G^k = (V, E^k)$, for some positive integer k , is obtained from G by including an edge between every pair of vertices that are of distance at most k in G (hence G^1 is precisely G itself).

We draw particular attention to the fact that the search algorithms in this paper generate vertex orderings from left to right. In many applications it may be useful to reverse the ordering generated by search, and some of the LexBFS literature considers this reversal as part of the search algorithm.

2. Characterizing graph search.

2.1. Generic search. We begin by formalizing the generic search idea, similar to Tarjan's description in [14], and present it as Algorithm 1. We use a generic "set" data structure S to store the candidate vertices we may next visit during our search. We note that we impose no selection criteria on which member of S we choose in line 3 beyond that it must be in S .

We call any ordering of V that can be generated by Algorithm 1 a *search ordering* of the graph G . As an example, we notice that $(1, 6, 2, 3, 4, 5)$ is not a valid search ordering of the 3-sun in Figure 2.1, whereas $(1, 2, 4, 3, 5, 6)$ is a valid search ordering of this graph. The reader will notice that this search ordering is neither a BFS nor a DFS, which is permissible by the generic selection criterion in line 3. We may opt to use a more restrictive data structure to obtain other properties for our search, such as replacing set S with a queue to obtain BFS. Note also that, as stated, a given vertex may be added to S multiple times as each of its neighbors is numbered. How S deals with this can change the nature of the search. For example, using a variant of a stack

ALGORITHM 1: Generic search.

input : a graph $G = (V, E)$ and start vertex $s \in V$
output: an ordering σ of V

```

1  $S \leftarrow \{s\}$ 
2 for  $i \leftarrow 1$  to  $n$  do
3   pick and remove an unnumbered vertex  $v$  from  $S$ 
4    $\sigma(i) \leftarrow v$  // This assigns to  $v$  the number  $i$ 
5   foreach unnumbered vertex  $w$  adjacent to  $v$  do
6      $\mid$  add  $w$  to  $S$ 

```

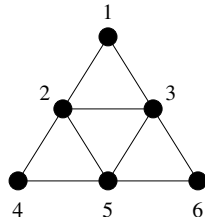
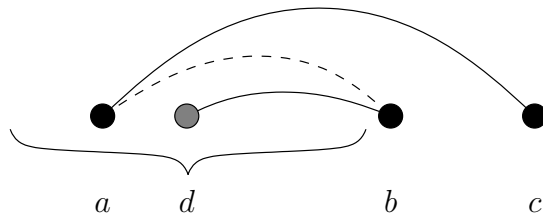


FIG. 2.1. The 3-sun, used as an example for our graph searches.

where the add operation moves a vertex to the top of the stack if that vertex was added previously will yield DFS.

We now proceed to answer the fundamental question raised above about our search orderings: how could b have been chosen before c ? We give a simple answer, which turns out to be the characterizing property of search orderings.



(S): Given an ordering σ of V , if $a < b < c$ and $ac \in E$ and $ab \notin E$, then there exists a vertex $d < b$ such that $db \in E$.

We note that (S) makes no statement to the relative order of a and d in σ . It simply says that there must be a vertex adjacent to b visited sometime earlier than b in σ . We will now show that (S) characterizes search orderings. That is to say that any ordering that the generic search algorithm can produce has this property, and that any vertex ordering with this property can be produced by the generic search algorithm when we break ties suitably.

THEOREM 2.1. *For an arbitrary graph G , an ordering σ of V is a search ordering of G if and only if σ has property (S).*

Proof. For the forward direction, suppose σ is an ordering generated by Algorithm 1. Let $a < b < c$ be a triple of vertices in σ with $ac \in E$ and $ab \notin E$. At the point when b is chosen, both b and c must be in S . Since b is in S , some neighbor of b must have already been chosen; call this neighbor d . Thus $d < b$ and $db \in E$. Since the triple $a < b < c$ was arbitrary, property (S) holds on σ .

For the converse, suppose $\sigma = (v_1, \dots, v_n)$ is an ordering respecting property (S). Suppose for contradiction that $(v_1, \dots, v_{i-1})$ can be obtained by Algorithm 1, but $(v_1, \dots, v_i)$ cannot, for some i , $1 < i \leq n$. That is, v_i is the first vertex in σ that cannot be chosen next by the algorithm. Consider the next iteration of the algorithm after choosing v_{i-1} . Since v_i cannot be chosen next, $v_i \notin S$. Let u be a vertex that the algorithm may choose next. Then u must be in S . Let w be the neighbor of u which caused u to be added to S . Since $v_i \notin S$, we know that $wv_i \notin E$. By applying property (S) to the vertex triple $w < v_i < u$, we know there exists a vertex $d < v_i$ with $dv_i \in E$. Hence $v_i \in S$, contradicting our assumption that v_i could not be chosen by Algorithm 1. $\square$

In fact, the perceptive reader will note that vertex c is irrelevant when restricting our attention to connected graphs. However, the presence of c is required when extending to more general settings.

2.2. BFS. We now proceed to the familiar BFS algorithm. BFS is a restriction of generic search in that we use a queue to determine which vertex we may choose next at our selection step. BFS does not, however, determine which order to push the neighbors of a numbered vertex onto the queue (see LexBFS in section 2.3 for a search that does). Algorithm 2 should be familiar to most readers.

ALGORITHM 2: Breadth-first search (BFS).

input : a graph $G = (V, E)$ and start vertex $s \in V$
output: an ordering σ of V

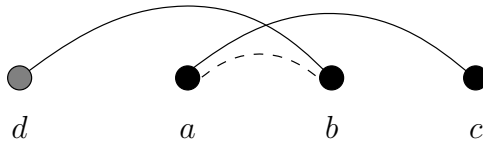
```

1 initialize queue to  $\{s\}$ 
2 for  $i \leftarrow 1$  to  $n$  do
3   dequeue  $v$  from beginning of queue
4    $\sigma(i) \leftarrow v$  // This assigns to  $v$  the number  $i$ 
5   foreach unnumbered vertex  $w$  adjacent to  $v$  do
6     if  $w$  is not already in queue then
7       enqueue  $w$  to end of queue

```

We mention that $(1, 2, 3, 4, 5, 6)$ is a valid BFS ordering of the 3-sun in Figure 2.1. Note that $(1, 2, 3, 6, 5, 4)$ is not a BFS ordering of this graph, though it does give a distance layering of the graph with respect to vertex 1.

We key in on our central question: How could b have been chosen before c in a BFS? The answer is a simple restriction to the answer we found for generic search.



(B): Given an ordering σ of V , if $a < b < c$ and $ac \in E$ and $ab \notin E$, then there exists a vertex $d < a$ such that $db \in E$.

This property is merely property (S) with the added restriction that $d < a$. Interestingly, we find that this is sufficient to characterize all BFS orderings.

THEOREM 2.2. *For an arbitrary graph G , an ordering σ of V is a BFS ordering of G if and only if σ has property (B).*

Proof. For the forward direction, suppose σ is a BFS ordering. Let $a < b < c$ be a triple of vertices in σ with $ac \in E$ and $ab \notin E$. If b was removed from (and thus placed on) the queue before c , it must have a neighbor d that was numbered at least as early as a . Since a does not have b as a neighbor, d must be earlier than a .

For the reverse direction, suppose for contradiction that $\sigma = (v_1, \dots, v_n)$ is an ordering satisfying property (B) such that $(v_1, \dots, v_{i-1})$ can be produced by BFS, but $(v_1, \dots, v_i)$ cannot, for some i , $1 < i \leq n$. That is, v_i cannot be chosen next by any valid execution of BFS on G . Let u be a vertex that may be chosen next by some execution of BFS. Since u is dequeued in this extension, it must have been enqueued when some neighbor v_j with $j < i$ was previously numbered. Then $v_j v_i \notin E$; otherwise some valid execution of BFS can enqueue v_i before u . Applying (B) to the triple $v_j < v_i < u$, there must exist a vertex $d < v_j$ with $dv_i \in E$. Thus v_i is enqueued before, and dequeued before, u , contradicting our assumption above. $\square$

To illustrate an application of this property, we now prove the correctness of a generalization of a well-known technique for extracting a shortest s, w -path (for arbitrary w) given an ordering with property (B) starting with s . For notation, when given a vertex ordering σ of V and a subset $T \subseteq V$, we say $\sigma|_T$ is the restriction of σ to T .

COROLLARY 2.3. *Let σ be an arbitrary ordering of the vertices of a graph G , and let $H = G[T]$ be a connected induced subgraph of G such that $\sigma|_T$ satisfies (B). Suppose s is the first vertex in $\sigma|_T$, and let $w = w_0$ be an arbitrary vertex in T . For $i > 0$, define w_i to be the leftmost (in $\sigma|_T$) neighbor of w_{i-1} , until s is reached. Then $(s = w_k, \dots, w_1, w_0 = w)$ is a shortest s, w -path in H .*

Proof. Suppose, for contradiction, that $(s = w_k, \dots, w_1, w_0 = w)$ is not a shortest s, w -path in H . Let $(s = x_{k'}, \dots, x_1, x_0 = w)$ be a shorter s, w -path, with $k' < k$. Note that $x_{k'} < w_{k'}$. Choose the smallest $i > 0$ such that $x_i < w_i$. Then $w_{i-1} \leq x_{i-1}$. Since $x_i x_{i-1} \in E$ and w_i is the leftmost neighbor of w_{i-1} , we know that $x_i w_{i-1} \notin E$ (and thus $w_{i-1} \neq x_{i-1}$). Consider the triple $x_i < w_{i-1} < x_{i-1}$. Since $x_i x_{i-1} \in E$ but $x_i w_{i-1} \notin E$, property (B) applied to $\sigma|_T$ says there is a vertex d in T such that $d < x_i$ with $dw_{i-1} \in E$. This contradicts w_i being the leftmost neighbor of w_{i-1} , and hence our original path was in fact shortest. $\square$

2.3. LexBFS. In the studies of LexBFS, the answer to our key question, property (LB) below, has long been exploited as a useful characterization of LexBFS orderings. For completeness, we present the traditional LexBFS labeling algorithm as Algorithm 3.

As an example, consider Algorithm 3 on the graph in Figure 2.1 from the start vertex 1. The ordering $(1, 2, 3, 5, 4, 6)$ is a valid LexBFS ordering of the 3-sun. Note that once 1, 2, 3 are picked, 5 must be chosen over 4 since $label(5) = \{4, 3\}$ while

ALGORITHM 3: Lexicographic breadth-first search (LexBFS).

input : a graph $G = (V, E)$ and start vertex $s \in V$

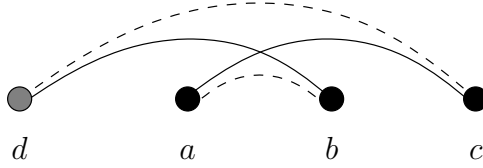
output: an ordering σ of V

```

1 assign the label  $\epsilon$  to all vertices
2  $label(s) \leftarrow \{n+1\}$ 
3 for  $i \leftarrow 1$  to  $n$  do
4   pick an unnumbered vertex  $v$  with lexicographically largest label
5    $\sigma(i) \leftarrow v$  // This assigns to  $v$  the number  $i$ 
6   foreach unnumbered vertex  $w$  adjacent to  $v$  do
7     append  $(n-i)$  to  $label(w)$ 
```

$label(4) = \{4\}$. This ordering is also a BFS ordering of the 3-sun, but the BFS ordering $(1, 2, 3, 4, 5, 6)$ is not a LexBFS ordering.

We now restate the answer to our key question for LexBFS.



(LB): Given an ordering σ of V , if $a < b < c$ and $ac \in E$ and $ab \notin E$, then there exists a vertex $d < a$ such that $db \in E$ and $dc \notin E$.

This property says that if we choose b before c , but there is an earlier vertex a that would normally pull c before b , then there must be an even earlier vertex d that caused b to be pulled ahead of c . Though this has long been known to be a property of LexBFS orderings [11, 8], the reverse direction showing (LB) characterizes LexBFS orderings has only recently been shown [1].

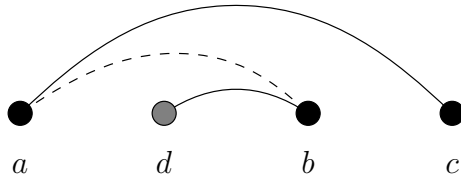
THEOREM 2.4 (see [11, 8, 1]). *For an arbitrary graph G , a vertex ordering σ is a LexBFS ordering of G if and only if σ has property (LB).*

Remarkably, the requirement of dc not being an edge is the only difference between lexicographic and traditional BFS. We obtain the immediate corollary, without any mention of the respective algorithms, that a LexBFS ordering is a BFS ordering, since (B) subsumes (LB).

2.4. DFS. We now turn our attention to DFS. DFS is very different from BFS in that it progresses forward through the graph as much as possible, backtracking only when necessary, whereas BFS first systematically explores close vertices before venturing out to the far reaches of the graph. One also notes that DFS produces two vertex orderings: a discovery ordering, which is the order vertices are first visited, and a finishing ordering, indicating the order in which each vertex's neighborhood is fully explored with all incident edges traversed and causing a backtracking step. In this paper we concentrate on the discovery orderings of DFSs, though we note an interesting relationship between the two orderings in section 4.

We start by presenting a DFS algorithm exploiting a stack data structure to store candidate vertices to visit next. We present DFS as Algorithm 4 in a style similar to our generic search algorithm in section 2.1. We note that line 6 would typically be included in the push operation of a suitable stack-type data structure, but we include the step here in our algorithm for clarity.

We again concentrate on our key question: If given a DFS ordering with $a < b < c$ with ac an edge but ab not an edge, how could DFS have chosen b over c ? The answer is property (D), which we will show characterizes DFS orderings.



(D): Given an ordering σ of V , if $a < b < c$ and $ac \in E$ and $ab \notin E$, then there exists a vertex d with $a < d < b$ such that $db \in E$.

ALGORITHM 4: Depth-first search (DFS).

input : a graph $G = (V, E)$ and start vertex $s \in V$
output: an ordering σ of V

```

1 initialize stack to  $\{s\}$ 
2 for  $i \leftarrow 1$  to  $n$  do
3   pop  $v$  from top of stack
4    $\sigma(i) \leftarrow v$  // This assigns to  $v$  the number  $i$ 
5   foreach unnumbered vertex  $w$  adjacent to  $v$  do
6     if  $w$  already in stack then remove  $w$  from stack
7     push  $w$  on top of stack

```

Notice that this property restricts (S) only in that d must also appear to the right of a . We contrast this with BFS property (B), where d was to the left of a . In this sense, DFS is the opposite search paradigm to BFS.

THEOREM 2.5. *For an arbitrary graph G , an ordering σ of V is a DFS ordering of G if and only if σ has property (D).*

Proof. For the forward direction, suppose σ is an ordering produced by DFS. Suppose we have a triple of vertices $a < b < c$ in σ with $ac \in E$ and $ab \notin E$. Since b is numbered (and $b \neq s$, the initial vertex), b must have been pushed on the *stack* when one of its neighbors was previously numbered. Let d be the rightmost neighbor of b appearing before b in σ . Then $d < b$ and $d \neq a$. Suppose $d < a$. Since a was numbered after d , c is pushed on the stack above b . As there are no other neighbors of b before b is numbered, c is higher on the stack than b and must be numbered before b , a contradiction. Thus we must have $a < d$, and property (D) holds.

Now for the reverse direction, suppose $\sigma = (v_1, \dots, v_n)$ is an ordering respecting property (D). Suppose $(v_1, \dots, v_{i-1})$ can be obtained by DFS, but $(v_1, \dots, v_i)$ cannot. That is, v_i is the first vertex in σ that cannot be chosen next by DFS. Let u be a vertex that DFS may choose next. Then $v_i < u$. Let v_j be the rightmost vertex in σ_{i-1} such that $v_j u \in E$ (that is, v_j is the parent vertex of u in the DFS-tree). Then $v_j v_i \notin E$ (otherwise v_i could be chosen next) and, more generally, $v_k v_i \notin E$ for all k , $j \leq k < i$. By applying property (D) to the triple $v_j < v_i < u$, we know there exists a vertex d between v_j and v_i such that $dv_i \in E$, a contradiction to v_i being ineligible to be chosen next. $\square$

We can apply (D) repeatedly to obtain the following simple corollary.

COROLLARY 2.6. *Given a graph G and an arbitrary DFS ordering σ of V , if $a < b < c$ and $ac \in E$, then there exists a path P from a to b whose vertices increase in σ . That is, there exists an a, b -path $P = (a = d_k, \dots, d_1, d_0 = b)$ in G with $\sigma^{-1}(d_i) < \sigma^{-1}(d_{i-1})$ for each $d_i \in P \setminus \{b\}$.*

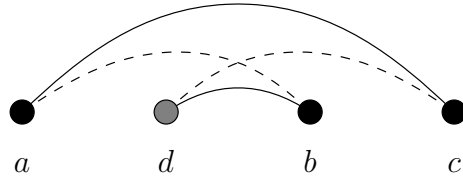
Most previous results regarding DFS rely on properties of the DFS-tree. The *DFS-tree* T is defined on the vertices of G and includes exactly those edges traversed to number a new vertex. Formally, $v_i v_j \in E(T)$ with $i < j$ if and only if $i = \max\{k \mid v_k v_j \in E(G) \text{ and } k < j\}$. We say u is the *parent* of v in T if $uv \in E(T)$ and $\sigma^{-1}(u) < \sigma^{-1}(v)$. The *nontree edges* (or *back edges*) are the remaining edges in G . In T , u is an *ancestor* of v if there is a path from u to v all of whose edges go from parent to child. We say two vertices are *related* if one is an ancestor of the other.

Many data structure textbooks (for example, [2]) mention the following fundamental property of DFS (which follows from Corollary 2.6).

(D2): In a DFS-tree of an undirected graph, the two endpoints of each back (non-tree) edge are related.

Consider a back edge uv in a DFS-tree. The four-vertex condition (D) is stronger than (D2) in that it says all vertices between u and v (inclusive) are related to u .

2.5. LexDFS. Looking at the relationship between BFS and LexBFS, one naturally asks whether there is a “lexicographic analogue” for DFS. We have noted that LexBFS enforces the requirement of “ $dc \notin E$ ” on top of BFS. It is natural to question whether a similar requirement makes sense for DFS. We state this restriction of property (D) below.



(LD): Given an ordering σ of V , if $a < b < c$ and $ac \in E$ and $ab \notin E$, then there exists a vertex d with $a < d < b$ such that $db \in E$ and $dc \notin E$.

We turn to the task of devising an algorithm that realizes this property. Considering our 3-sun example in Figure 2.1, once vertices 1 and 2 are chosen, the only way to avoid violating (LD), and hence the only valid ordering, is $(1, 2, 3, 5, 6, 4)$. Clearly this is a DFS ordering, and we observe that the DFS ordering $(1, 2, 4, 5, 6, 3)$ violates (LD). The LexBFS ordering $(1, 2, 3, 5, 4, 6)$ is also disallowed.

These examples suggest that our algorithm should choose vertices that are adjacent to as many vertices we have more recently numbered as possible. As it is to be related to LexBFS, our algorithm should also be in some sense similar to our LexBFS algorithm. (We discuss the property that actually relates LexBFS and LexDFS in section 2.6.) We present Algorithm 5 as our LexDFS algorithm.

ALGORITHM 5: Lexicographic depth-first search (LexDFS).

input : a graph $G = (V, E)$ and start vertex $s \in V$

output: an ordering σ of V

```

1 assign the label  $\epsilon$  to all vertices
2  $label(s) \leftarrow \{0\}$ 
3 for  $i \leftarrow 1$  to  $n$  do
4   pick an unnumbered vertex  $v$  with lexicographically largest label
5    $\sigma(i) \leftarrow v$  // This assigns to  $v$  the number  $i$ 
6   foreach unnumbered vertex  $w$  adjacent to  $v$  do
7     prepend  $i$  to  $label(w)$ 
```

Examining Algorithm 5, our answer to the “what is LexDFS?” question becomes quite convincing. There is just one difference between this algorithm and the LexBFS algorithm. For LexDFS, we *prepend* the new digit i to $label(w)$ at line 7, whereas, for LexBFS, we *append* the new label digit $n - i$. This elementary change converts the algorithm from a BFS into a DFS. We now show that our algorithm correctly defines LexDFS.

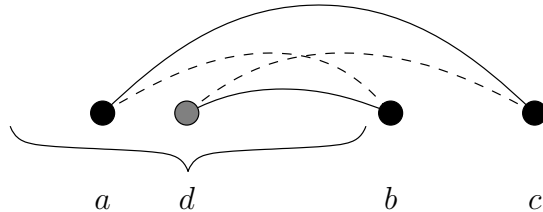
THEOREM 2.7. *For an arbitrary graph G , an ordering σ of V is a LexDFS ordering of G if and only if σ has property (LD).*

Proof. For the forward direction, suppose σ is an ordering computed by some execution of LexDFS. We define $label_u(v)$ to represent the label assigned to vertex v in this execution immediately before vertex u was numbered. Let $a < b < c$ be a triple of vertices in σ with $ac \in E$ and $ab \notin E$. Since b was numbered before c , $label_b(b) \geq label_b(c)$. We note that the digits in each label appear in decreasing order. Since $\sigma^{-1}(a)$ appears in $label_b(c)$ and does not appear in $label_b(b)$, $label_b(b)$ must include some digit $j > \sigma^{-1}(a)$ which does not appear in $label_b(c)$. Let $d = \sigma(j)$. Then $d < b$, $db \in E$, and $dc \notin E$. Since $j > \sigma^{-1}(a)$, $a < d$, and thus property (LD) is satisfied.

Now for the reverse direction, suppose $\sigma = (v_1, \dots, v_n)$ is an ordering respecting property (LD). Suppose $(v_1, \dots, v_{i-1})$ can be obtained by LexDFS, but $(v_1, \dots, v_i)$ cannot. That is, v_i is the first vertex in σ that cannot be chosen next by LexDFS. Let u be a vertex that LexDFS may choose next. Since u must be chosen over v_i , at this point of execution $label(u)$ must be lexicographically larger than $label(v_i)$: let j be the largest digit in $label(u)$ that is not in $label(v_i)$ (i.e., v_j is the rightmost vertex in σ_{i-1} with $v_j u \in E$ and $v_j v_i \notin E$). By applying property (LD) to the vertex triple $v_j < v_i < u$, we know there exists a vertex d between v_j and v_i such that $dv_i \in E$ and $du \notin E$. Then $\sigma^{-1}(d)$ is in $label(v_i)$ but not in $label(u)$. Since $v_j < d$, $label(v_i)$ is in fact larger than $label(u)$, a contradiction to v_i being ineligible to be chosen next. $\square$

When examining a new search algorithm, we should concern ourselves with its implementation. Algorithm 5 does not immediately have a linear-time implementation using those ideas used for LexBFS. The partitioning idea exploited for LexBFS cannot efficiently be applied to LexDFS (leading to an $O(m \log n)$ time bound), as its depth-first properties can pull in an arbitrary vertex to be the next to be numbered. We refer to [10] for an $O(\min\{n^2, m \log n\})$ implementation; recently an $O(m \log \log n)$ implementation has been discovered [13] employing van Emde Boas trees to keep track of the relative lexicographic order of unnumbered vertices. It is an open question whether there is a linear-time implementation of LexDFS.

2.6. Maximal neighborhood search. We now consider the question of “what fact relates LexBFS and LexDFS?” What useful property does this nonedge dc give us? Equivalently, we ask what type of search we get when adding the restriction $dc \notin E$ to the generic search property (S). We formalize our discussion around this property.



(M): Given an ordering σ of V , if $a < b < c$ and $ac \in E$ and $ab \notin E$, then there exists a vertex d with $d < b$ and $db \in E$ and $dc \notin E$.

Surprisingly, this property has already been used in the literature. Tarjan and Yannakakis [15] mentioned that both MCS (where at each step we choose a vertex that is adjacent to the maximum number of vertices previously numbered) and LexBFS respect our property (M) (which [15] calls “P”), and they showed that any ordering

with property (M) is the reverse of a perfect elimination ordering of a chordal graph. But the paper does not consider whether there is a natural search algorithm that is characterized by this property. We answer that question here by introducing *maximal neighborhood search* (MNS) as Algorithm 6.

The idea behind MNS is simple: at each step, pick a vertex whose neighborhood in the part of the graph already explored is maximal (with respect to set inclusion). This is a straightforward generalization of the LexBFS, LexDFS, and MCS algorithms. We obtain each of these searches from Algorithm 6 by further specifying how to order incomparable neighborhoods. For LexBFS, we choose a vertex adjacent to as many *earliest* chosen vertices as possible. For LexDFS, we choose a vertex adjacent to as many *most recently* chosen vertices as possible. For MCS, we choose a vertex whose neighborhood has *maximum* cardinality.

ALGORITHM 6: Maximal neighborhood search (MNS).

input : a graph $G = (V, E)$ and start vertex $s \in V$
output: an ordering σ of V

- 1 assign the label $\emptyset$ to all vertices
- 2 $label(s) \leftarrow \{n + 1\}$
- 3 **for** $i \leftarrow 1$ **to** n **do**
- 4 pick an unnumbered vertex v with maximal label (under set inclusion)
- 5 $\sigma(i) \leftarrow v$ // This assigns to v the number i
- 6 **foreach** *unnumbered vertex* w *adjacent to* v **do**
- 7 add i to $label(w)$

We now show that the vertex orderings produced by Algorithm 6 are characterized by property (M).

THEOREM 2.8. *For an arbitrary graph G , an ordering σ of V is an MNS ordering if and only if σ has property (M).*

Proof. For the forward direction, suppose σ is an MNS ordering. Suppose $a < b < c$ is a triple in σ that violates (M). That is, $ac \in E$ and $ab \notin E$, yet there is no vertex $d < b$ with $db \in E$ and $dc \notin E$. Consider the point in the execution when $b = v_i$ is chosen. Then $N_{G_{i-1}}(b) \subsetneq N_{G_{i-1}}(c)$, so b could not have been chosen next by MNS, a contradiction. Hence (M) must hold for σ .

For the converse, suppose $\sigma = (v_1, \dots, v_n)$ is an ordering respecting property (M). Suppose $(v_1, \dots, v_{i-1})$ can be obtained by MNS, but $(v_1, \dots, v_i)$ cannot; that is, v_i is the first vertex in σ that cannot be chosen next by MNS. Then there must be another vertex u with $N_{G_{i-1}}(u) \supsetneq N_{G_{i-1}}(v_i)$. Let $x \in N_{G_{i-1}}(u) - N_{G_{i-1}}(v_i)$ (such an x must exist). We apply (M) to the triple $x < v_i < u$, giving a vertex $d < v_i$ with $dv_i \in E$ and $du \notin E$. This contradicts $N_{G_{i-1}}(u) \supsetneq N_{G_{i-1}}(v_i)$, and hence our assumption that v_i could not be chosen by MNS was incorrect. $\square$

Thus, for chordal graphs, every MNS ordering is the reverse of a perfect elimination ordering. A natural question is whether every perfect elimination ordering of a chordal graph is the reverse of an MNS ordering. This question is studied by Shier [12] and answered in the negative: the reverse of many perfect elimination orderings are not search orderings, and thus cannot be found by MNS. However, [12] shows how to generalize search orderings by considering connected components in the unvisited portion of the graph to capture the entire family of perfect elimination orderings.

3. Application to powers of graphs. All searches give a certain amount of insight into a graph's structure, including the insight gained from our four-vertex characterizations. At times we wish to use this insight to gain information about some graph arising from a given graph via some function. Examples of such functions include the graph's complement, some power of the graph, the line graph, or the product or join of the graph. Often, we want to obtain insight into the structure of a derived graph without incurring the possibly large cost of computing the function.

Search is often a convenient and efficient way to do this. For example, it has been previously noted that a LexBFS ordering (and thus a BFS ordering) of the complement of a graph can be found without computing the complement, given a slight modification of the standard LexBFS algorithm [9]. Dahlhaus, Gustedt, and McConnell [5] have shown how to find a DFS ordering of the complement of a graph in time linear in the original graph. We can use such results to obtain information about the derived graph (such as whether the graph's complement is chordal) more efficiently than actually computing the derived graph.

In this section, we consider the problem of computing some type of search ordering of some power of the graph, but without resorting to computing that power. One possibility is to compute a search ordering of the original graph, and ask whether it is a search ordering of the power. We will show that the aforementioned search characterizations are invaluable in obtaining straightforward proofs of both positive and negative results for this problem.

We start with the simple question of whether a BFS ordering of a graph remains a BFS ordering for all powers of a graph.

THEOREM 3.1. *For any graph G , every BFS ordering of G is a BFS ordering of G^k for all $k \geq 1$.*

Proof. Let σ be a BFS ordering of G . The theorem is trivially satisfied when $k = 1$. For $k > 1$, we assume for induction that σ is a BFS ordering for all powers of G strictly less than k and consider σ on G^k . Suppose $a < b < c$ is a triple in σ on G^k with $ac \in E^k$ and $ab \notin E^k$.

If $a < b < c$ is such a triple in σ on G^{k-1} , then by (B) and our induction hypothesis, there exists a vertex $d < a$ with $db \in E^{k-1} \subseteq E^k$, hence satisfying (B) on G^k .

Suppose instead that $a < b < c$ is not such a triple in σ on G^{k-1} . Then $ac \notin E^{k-1}$. But then there must be a vertex e with $ae \in E^{k-1}$ and $ec \in E$. Since $ab \notin E^k$ but $ae \in E^{k-1}$, we must have $eb \notin E$. We have 3 cases for the position of e in σ relative to a and b .

Case 1. $e < a$. In σ on G , since $ec \in E$ and $eb \notin E$, we apply (B) to the triple $e < b < c$, giving us a vertex $d < e < a$ such that $db \in E \subseteq E^k$, satisfying (B) on G^k .

Case 2. $a < e < b$. We first consider σ on G . We know that $ec \in E$ and $eb \notin E$, so by applying (B) to the triple $e < b < c$, we get a vertex $d < e$ such that $db \in E$. If $d < a$, we're done; otherwise, we have $a < d$. Since $db \in E$ but $ab \notin E^k$, we must have $ad \notin E^{k-1}$. We apply (B) and our inductive hypothesis to the triple $a < d < e$ in σ on G^{k-1} , giving us a vertex $d_2 < a$ such that $d_2d \in E^{k-1}$. Since $d_2d \in E^{k-1}$ and $db \in E$, d_2 and b are of distance at most k in G . Thus $d_2b \in E^k$, satisfying (B) on G^k .

Case 3. $b < e$. In σ on G^{k-1} , since $ae \in E^{k-1}$ and $ab \notin E^{k-1}$, we apply (B) and our inductive hypothesis to the triple $a < b < e$, giving us a vertex $d < a$ with $db \in E^{k-1} \subseteq E^k$, satisfying (B) on G^k .

In each case, we show the existence of a vertex $d < a$ with $db \in E^k$, and hence (B) is satisfied by σ on G^k . Thus σ is a BFS ordering of G^k . $\square$

While a BFS ordering survives through the powers of a graph, and clearly generic search orderings also survive, one naturally wonders about DFS. We now show that a DFS ordering of G is not necessarily a DFS ordering in a given power of G . Again, the characterizing property of DFS will play an crucial role in our proof.

THEOREM 3.2. *For every $k \geq 2$, there exists a graph G such that no DFS ordering of G is a DFS ordering of G^k .*

Proof. Fix $k \geq 2$. We first show the existence of a graph G such that no DFS ordering starting from a distinguished start vertex s is a DFS ordering of G^k .

Let G be the graph consisting of two vertices s and r connected by three paths of length $k + 1$. This graph is diagrammed in Figure 3.1.

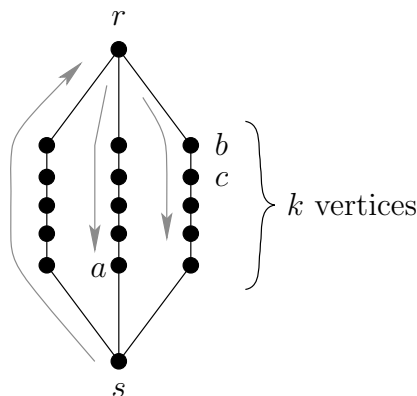


FIG. 3.1. Given k , a graph G where no DFS from s is a DFS of G^k . The possible DFS path through G is shown in gray.

Let σ be an arbitrary DFS ordering of G starting from s . The DFS will proceed from s along one of the three paths to reach r . From r , the DFS will follow a second path towards s to reach a adjacent to s . As s is already discovered, the DFS will recurse back to r and travel the third path towards s . Let b be the first vertex visited on this third path, and let c be the next vertex on the path.

Thus, in σ , we have $s < r < a < b < c$. Moreover, b appears in σ immediately following a (when a is reached, all neighbors of a are already numbered, so DFS recurses to r , where b must be the next vertex to be numbered).

Now consider σ in G^k . One easily verifies that a and c are of distance k in G , and a and b are of distance $k + 1$ in G . Hence $ac \in E^k$ and $ab \notin E^k$. But since a and b are next to each other in σ , there can be no vertex between a and b . Therefore, the triple $a < b < c$ violates the condition (D). Hence σ cannot be a DFS ordering of G^k .

To prove the theorem, we consider two copies of G joined at s by an edge: call this graph G' . Consider any DFS ordering σ' of G' : it may start in either copy of G , possibly at s itself. We observe that if σ' starts in one copy of G , σ' restricted to the other copy of G appears as a DFS ordering of G starting at s . Following the argument above, we can identify three vertices $a < b < c$ in the second copy of G .

We note that no vertex from the other copy of G is within distance k of b , and thus there can be no edge ab in $(G')^k$. The triple $a < b < c$ violates (D) as above, so σ' is not a DFS of $(G')^k$. Since σ' was chosen arbitrarily, no DFS ordering of G' can be a DFS ordering of $(G')^k$. $\square$

Failing for DFS, one may wonder about LexBFS. Brandstädt, Dragan, and Nicolai [1] have shown that the reverse of a LexBFS ordering of a chordal graph G is a common

perfect elimination ordering of all odd powers of G . One may be tempted to think that a LexBFS ordering of a chordal graph is a LexBFS ordering of the odd powers, but we show this is not the case.

THEOREM 3.3. *For each $k \geq 2$, there exists a chordal graph G such that no LexBFS ordering of G is a LexBFS ordering of G^k .*

Proof. Let $k \geq 2$ be fixed. If k is even, we consider the graph G_k in Figure 3.2(a); otherwise, if k is odd, we consider the graph H_k in Figure 3.2(b). We note that each of these graphs is chordal. We will show that no LexBFS ordering of the graph starting from vertex 1 will be a LexBFS ordering of the k th power, then join two copies of the graph at vertex 1 to prove the theorem for a LexBFS ordering starting from an arbitrary vertex.

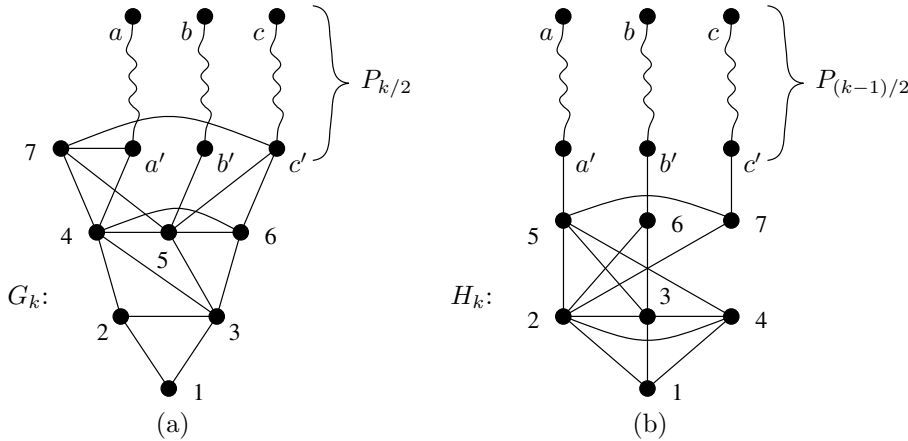


FIG. 3.2. Constructions for (a) k even and (b) k odd, where no LexBFS ordering starting from 1 is a LexBFS ordering of G^k .

Suppose k is even. We note that any LexBFS ordering σ of the graph G_k which starts at 1 must have $7 < a' < b' < c'$. Then $a < b < c$ in σ . We examine the triple $a < b < c$ in G_k^k . We note that $ac \in E(G_k^k)$ since $d_{G_k}(a, c) = 2(k/2 - 1) + 2 = k$, and $ab \notin E(G_k^k)$ since $d_{G_k}(a, b) = 2(k/2 - 1) + 3 = k + 1$. It is easy to check that there is no vertex $d < a$ with $db \in E(G_k^k)$ and $dc \notin E(G_k^k)$: any vertex left of a within distance k of b in G_k must be within distance k of c . Thus property (LB) is violated in G_k^k , showing σ is not a LexBFS ordering of G_k^k .

Now suppose k is odd. Considering the graph in Figure 3.2(b), we note that any LexBFS ordering σ of H_k that starts at 1 must have $a' < b' < c'$. Then $a < b < c$ in σ . We note that $d_{H_k}(b', 3) = 2$ and $d_{H_k}(c', 3) = 3$, and $d_{H_k}(b', 6) = 1$ and $d_{H_k}(c', 6) = 3$, but for all other vertices $x < a'$ in σ , $d_{H_k}(b', x) \geq d_{H_k}(c', x)$. We examine the triple $a < b < c$ in H_k^k . We have $ac \in E(H_k^k)$, since $d_{H_k}(a, c) = 2(\frac{k-1}{2} - 1) + 3 = k$, and $ab \notin E(H_k^k)$, since $d_{H_k}(a, b) = 2(\frac{k-1}{2} - 1) + 4 = k + 1$. But since all vertices $d < a$ in σ within distance k of b are also within distance k of c , there is no candidate vertex d to satisfy property (LB) for a, b, c . Thus σ is not a LexBFS ordering of H_k^k .

To complete the proof, we construct a graph G consisting of two copies of G_k if k is even (respectively, two copies of H_k if k is odd), with the vertices labeled 1 in each copy joined by an edge. Any LexBFS ordering of G when restricted to at least one of the copies of G_k (or H_k) will appear as a LexBFS ordering with start

vertex 1. Applying the argument above will identify vertices $a < b < c$ that violate property (LB) in G^k , completing our proof. $\square$

We note that the graph in Figure 3.2(b) is in fact strongly chordal. A graph is *strongly chordal* if it is chordal and every cycle $C = (x_1, \dots, x_{2k})$ of even length has a chord $x_i x_j$ where the distance between the endpoints x_i and x_j in C is odd. An immediate corollary of Theorem 3.4 is that for each $k \geq 1$, there exists a strongly chordal graph G such that no LexBFS ordering of G is a LexBFS ordering of G^{2k+1} .

LexBFS does, however, work for a particular class of graphs, as shown by the next theorem.

THEOREM 3.4. *Let G be a chordal bipartite graph. If σ is a LexBFS ordering of G , then σ is a LexBFS ordering of the graph G^2 .*

Proof. Let σ be a LexBFS ordering of a chordal bipartite graph $G = B(X, Y, E)$. Let $l(x)$ denote the BFS layer of x in σ ; $l(x) = 0$ if x is the first vertex of σ . We first state a few simple facts.

FACT 1. *If $x \in X$ and $xy \in E$, then $y \in Y$.*

FACT 2. *If $x \in X$ and $xy \in E^2 - E$, then $y \in X$.*

FACT 3. *If $x \in X$, $y \in Y$, and $z \in X$, then $l(x) = l(y) + 2i - 1 = l(z) + 2j$ for some (possibly negative) integers i, j .*

FACT 4. *If there is some induced P_5 a, b, c, d, e with $l(a) = l(e) = l(b) - 1 = l(d) - 1 = l(c) - 2$, then there exists an induced C_{2k} , $k \geq 3$, in G .*

To prove our theorem, suppose we have $a < b < c$ with $ac \in E^2$ and $ab \notin E^2$. We will show the existence of a vertex $z < a$ with $zb \in E^2$ and $zc \notin E^2$.

Case 1. $ac \in E$. By (LB) on $a < b < c$, there is a $d < a$ with $db \in E$ and $dc \notin E$. If $dc \notin E^2$, d satisfies the theorem. So assume $dc \in E^2 - E$. Since $ab \notin E^2$, we must have $da \notin E$. By Facts 2 and 3, and since $l(d) \leq l(a) = l(c) - 1$, we have $l(d) = l(c) - 2$. By (LB) on $d < a < b$, there is an $e < d$ with $ea \in E$ and $eb \notin E$. Since $l(e) = l(d)$, $ed \notin E$. By (LB) on $e < d < a$, there must be some $f < e$ with $fd \in E$, $fa \notin E$ and $l(f) = l(d) - 1 = l(c) - 3$. Thus $fb \in E^2$ and $fc \notin E^2$, satisfying our requirement.

Case 2. $ac \notin E$. Then there is some d with $da, dc \in E$; choose the leftmost (in σ) such d . Since $ab \notin E^2$, $db \notin E$. We have 4 cases for the position of d relative to a, b , and c .

Case 2.1. $c < d$. By (LB) on $a < c < d$, there is an $e < a$ with $ec \in E$ and $ed \notin E$. Since d is leftmost, $ea \notin E$. By (LB) on $e < a < c$, there is an $f < e$ with $fa \in E$ and $fc \notin E$. Then f, a, d, c, e is an induced P_5 such that Fact 4 gives us an induced cycle of length at least 6, contradicting G being chordal bipartite.

Case 2.2. $b < d < c$. By (LB) on $a < b < d$, there is an $e < a$ with $eb \in E$ and $ed \notin E$. Since $ab \notin E^2$, $ea \notin E$. For generality, we discard our assumptions that $b < d$ and $ed \notin E$. Now, by (LB) on $e < a < b$, there is some $f < e$ with $fa \in E$ and $fb \notin E$, which means $l(f) = l(a) - 1 = l(c) - 3$. If $fe \in E$, then $fb \in E^2$ and since $fc \notin E^2$, f satisfies the theorem. Otherwise $fe \notin E$, so applying (LB) to $f < e < a$ gives us a $g < f$ with $ge \in E$, $ga \notin E$. Since $l(g) \leq l(f) = l(c) - 3$, we know $gc \notin E^2$, so g fulfills the theorem since $gb \in E^2$.

Case 2.3. $a < d < b$. By (LB) on $d < b < c$, there is some $e < d$ with $eb \in E$ and $ec \notin E$. Since $ab \notin E^2$, $ea \notin E$. If $e < a$, we are in the generalized Case 2.2 above, and can find a vertex g to satisfy the theorem. When $a < e$, we can apply (LB) to $a < e < d$ to give an $f < a$ with $fe \in E$, $fd \notin E$. Thus $fb \in E^2$ and f satisfies the theorem unless there exists some g with $fg, gc \in E$. Suppose such a g exists. Then $l(f) = l(g) - 1 = l(c) - 2$ and $dg \notin E$. If $ag \notin E$, then a, d, c, g, f forms a P_5 and

Fact 4 gives us an induced cycle of length at least 6, contradicting G being chordal bipartite. Thus $ag \in E$ and $d < g$ (since d is leftmost). By (LB) on $f < d < g$, there is an $h < f$ with $hd \in E$ and $hg \notin E$. Then $l(h) = l(f)$ and $hf \notin E$. But then f, g, c, d, h is a P_5 such that Fact 4 gives us an induced cycle of length at least 6, contradicting G being chordal bipartite. Thus no such g can exist, and f satisfies the theorem.

Case 2.4. $d < a$. By (LB) on $d < b < c$, there is an $e < d$ with $eb \in E$, $ec \notin E$. Then $l(e) = l(b) - 1 = l(c) - 1$, so $ec \notin E^2$, thus completing the proof. $\square$

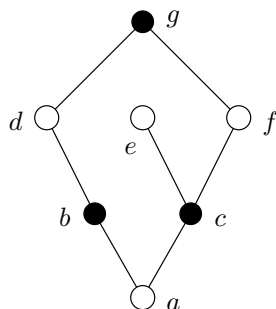


FIG. 3.3. A bipartite graph with a LexBFS ordering that is not a LexBFS ordering of the square.

Note that this theorem cannot be strengthened to bipartite graphs, as shown by the graph in Figure 3.3. Call this graph B . The ordering $\sigma = (a, b, c, d, e, f, g)$ is clearly a LexBFS ordering of B ; however, the triple of vertices $d < e < f$ violates property (LB) in B^2 . We have $df \in E^2$ and $de \notin E^2$, but every vertex in σ left of d that is adjacent to e in B^2 is also adjacent to f in B^2 . Attaching a second copy of B via an edge at a ensures no LexBFS ordering of the graph is a LexBFS ordering of its square.

4. Concluding remarks. We have examined six different search algorithms from the perspective of our key question: If a search visits a vertex a before b before c , and there is an edge ac but no edge ab , how could the search have visited b before c ? In this examination, we have discovered characterizing properties for the classic breadth-first and depth-first searches, and introduced a new class of maximal neighborhood searches. We summarize the characterizations of the six search paradigms and the hierarchy they form in Figure 4.1.

From this picture, one might be tempted to conjecture that the set of LexBFS orderings of a graph might be exactly the intersection of the set of MNS orderings and the set of BFS orderings. Clearly any LexBFS ordering must be both a BFS ordering and an MNS ordering, but the converse does not hold. For example, the ordering $(3, 1, 2, 5, 4)$ is both a BFS ordering and an MNS ordering of the gem in Figure 4.2, but is not a LexBFS ordering of this graph. Similarly for DFS, $(1, 2, 3, 4, 5)$ is both a DFS ordering and an MNS ordering of this graph, but not a LexDFS ordering.

We mentioned that the DFS algorithm produces two different orderings of the vertices of a graph: the discovery ordering and the finishing ordering. The following proposition relates these two orderings in the case of undirected graphs.

PROPOSITION 4.1. *For an arbitrary undirected graph G , an ordering σ of V is the reverse of a DFS finishing ordering if and only if σ has property (D).*

In other words, for undirected graphs, every ordering is the reverse of a finishing

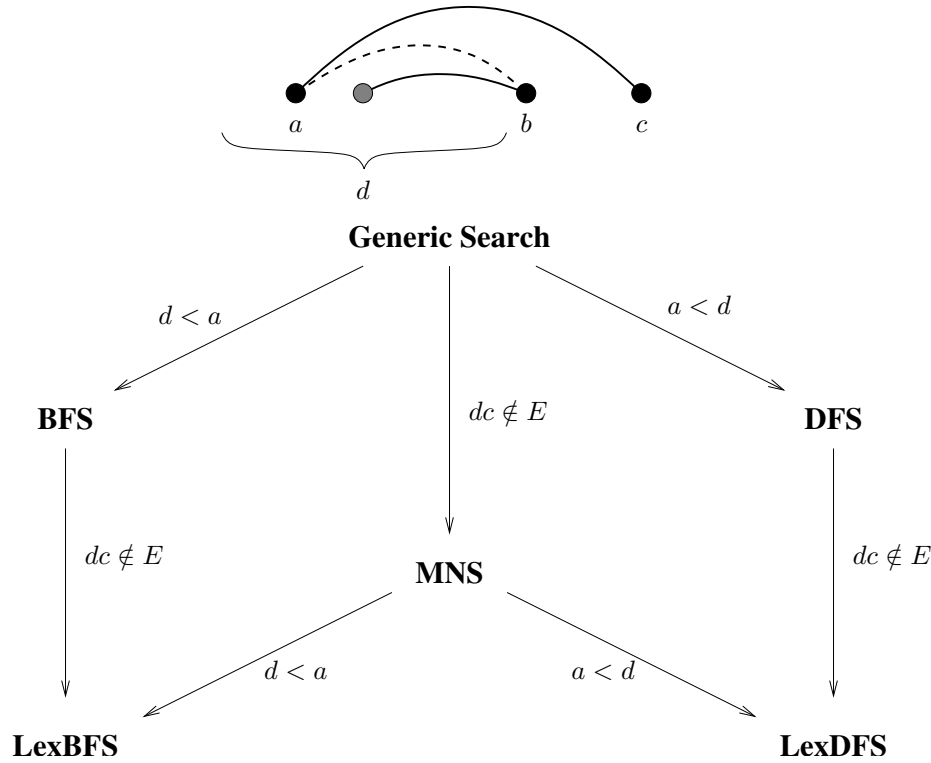


FIG. 4.1. Summary of search characterizations. For example, vertex ordering σ is a LexDFS ordering if and only if σ satisfies the condition for generic search together with $dc \notin E$ and $a < d$.

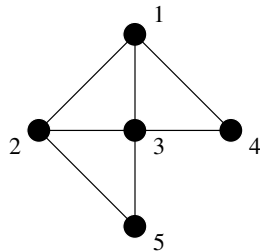


FIG. 4.2. The gem, used as an example to distinguish sets of orderings.

ordering if and only if it is a discovery ordering. Note that a similar proposition does not hold for LexDFS.

Though we have concentrated on simple graphs in this paper, these ideas can easily be extended to more general environments. Our characterizations make no mention of multiple edges, but are equally applicable to searches on multigraphs. A search ordering implicitly imposes an orientation on the edges of a simple graph: it is straightforward to apply the same ideas to directed graphs to obtain similar characterizations. We can also extend such characterizations beyond graphs: these ideas also hold for searches on hypergraphs. We refer to [10] for details.

Such characterizing properties can be a very powerful and useful tool for studying graphs. Beyond giving us a better understanding of how a search reveals the structure

of a graph, it allows us to employ (and analyze) multiple sweeps of a search to gather additional information about a graph. Multisweep LexBFS has recently been applied to a variety of problems [3], resulting in very simple, and in many cases intuitive, algorithms. Will the results presented in this paper lead to multisweep algorithms involving other searches or combinations of different searches?

Acknowledgments. We wish to thank the anonymous referees for their valuable comments on the paper; in particular, one referee noticed Proposition 4.1 and kindly allowed us to include it in the paper. We also thank Marty Golumbic for his careful reading of an earlier draft of the paper.

REFERENCES

- [1] A. BRANDSTÄDT, F. F. DRAGAN, AND F. NICOLAI, *LexBFS-orderings and powers of chordal graphs*, Discrete Math., 171 (1997), pp. 27–42.
- [2] T. H. CORMEN, C. STEIN, R. L. RIVEST, AND C. E. LEISERSON, *Introduction to Algorithms*, 2nd ed., McGraw-Hill, New York, 2001.
- [3] D. G. CORNEIL, *Lexicographic breadth first search—a survey*, in Graph-Theoretic Concepts in Computer Science, Proceedings of WG2004, Lecture Notes in Comput. Sci. 3353, Springer, Berlin, 2004, pp. 1–19.
- [4] D. G. CORNEIL, S. OLARIU, AND L. STEWART, *LBFS orderings and cocomparability graphs*, in Proceedings of the Tenth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, 1999, pp. S883–S884.
- [5] E. DAHLHAUS, J. GUSTEDT, AND R. M. MCCONNELL, *Partially complemented representations of digraphs*, Discrete Math. Theor. Comput. Sci., 5 (2002), pp. 147–168.
- [6] D. R. FULKERSON AND O. A. GROSS, *Incidence matrices and interval graphs*, Pacific J. Math., 15 (1965), pp. 835–855.
- [7] H. GABOW, *Searching*, in Handbook of Graph Theory, J. Gross and J. Yellen, eds., CRC Press, Boca Raton, FL, 2004, pp. 953–984.
- [8] M. C. GOLUMBIC, *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, New York, 1980; second edition, Ann. Discrete Math. 57, Elsevier, Amsterdam, 2004.
- [9] M. HABIB, R. MCCONNELL, C. PAUL, AND L. VIENNOT, *Lex-BFS and partition refinement, with applications to transitive orientation, interval graph recognition and consecutive ones testing*, Theoret. Comput. Sci., 234 (2000), pp. 59–84.
- [10] R. KRUEGER, *Graph Searching*, Ph.D. thesis, University of Toronto, 2005.
- [11] D. J. ROSE, R. E. TARJAN, AND G. S. LUEKER, *Algorithmic aspects of vertex elimination on graphs*, SIAM J. Comput., 5 (1976), pp. 266–283.
- [12] D. R. SHIER, *Some aspects of perfect elimination orderings in chordal graphs*, Discrete Appl. Math., 7 (1984), pp. 325–331.
- [13] J. SPINRAD, *Efficient implementation of lexicographic depth first search*, in preparation.
- [14] R. TARJAN, *Depth-first search and linear graph algorithms*, SIAM J. Comput., 1 (1972), pp. 146–160.
- [15] R. E. TARJAN AND M. YANNAKAKIS, *Simple linear-time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs*, SIAM J. Comput., 13 (1984), pp. 566–579.

A SIMPLE LINEAR TIME LexBFS COGRAPH RECOGNITION ALGORITHM*

ANNA BRETSCHER[†], DEREK CORNEIL[†], MICHEL HABIB[‡], AND CHRISTOPHE PAUL[§]

Abstract. Recently lexicographic breadth first search (LexBFS) has been shown to be a very powerful tool for the development of linear time, easily implementable recognition algorithms for various families of graphs. In this paper, we add to this work by producing a simple two LexBFS sweep algorithm to recognize the family of cographs. This algorithm extends to other related graph families such as P_4 -reducible, P_4 -sparse, and distance hereditary. It is an open question whether our cograph recognition algorithm can be extended to a similarly easy algorithm for modular decomposition.

Key words. graph algorithms, cograph recognition, linear time algorithms, LexBFS, P_4 -free graphs, modular decomposition

AMS subject classification. 05C85

DOI. 10.1137/060664690

1. Introduction. We introduce a new¹ linear time algorithm to recognize *cographs*, namely, graphs with no induced path on four vertices (P_4). Although several optimal cograph recognition algorithms already exist, our algorithm is of interest for several reasons: it is conceptually and practically simpler than previous algorithms; it returns both positive and negative certificates; it is *lexicographic breadth first search* (LexBFS) based, introducing a new variant LexBFS[−]; it is the foundation of recognition algorithms for other related families of graphs; and finally, it could potentially be generalized to a simple linear time graph modular decomposition algorithm.

Cographs or *complement reducible* graphs have the property of being defined both recursively and by a simple forbidden induced subgraph characterization, namely, no induced P_4 [9]. Cographs are the family of graphs constructed from a single vertex under the closure of the operations of union and complementation or, equivalently, union and join. As shown in [9] these operations uniquely define a tree representation referred to as a *cotree*. The certificate returned by our algorithm is either a cotree representation if the graph is a cograph or an induced P_4 .

Cographs arise in applications such as examination scheduling problems [26], automatic clustering of index terms [17], and recently in the recognition of read-once functions [16]. Surprisingly, despite the structural simplicity of cographs, constructing linear time recognition algorithms has been challenging. The first linear time algorithm to recognize cographs and construct the cotree was achieved by Corneil, Perl, and Stewart [10] in 1985. The algorithm examines each vertex in turn, inserts it into the cotree if the graph is a cograph, and returns an induced P_4 otherwise. A

*Received by the editors July 10, 2006; accepted for publication (in revised form) March 26, 2008; published electronically July 25, 2008. The authors wish to thank the Natural Sciences and Engineering Council of Canada and the French ANR project *Graph Decompositions and Algorithms* for financial assistance in support of this research.

<http://www.siam.org/journals/sidma/22-4/66469.html>

[†]Department of Computer Science, University of Toronto, Toronto, ONT M5S 3G4, Canada (bretscher@utsc.utoronto.ca, dgc@cs.toronto.edu).

[‡]LIAFA, CNRS-Université Paris 7-Denis Diderot, Paris Cedex 05, F-75251 France (michel.habib@liafa.jussieu.fr).

[§]LIRMM, CNRS-Université Montpellier II, Montpellier Cedex 5, 34 392, France (paul@lirmm.fr).

¹A variation of our algorithm, using three LexBFS sweeps instead of two, was outlined in [4].

parallel cograph recognition algorithm appeared in [13]. Recently, Habib and Paul [19] have used vertex splitting to produce a new linear time recognition algorithm for cographs. Despite being optimal, each of these algorithms requires more complex data structures or algorithmic techniques than our new algorithm.

The motivation for studying cographs lies not only in their simplicity but also in their extension to other families and to *modular decomposition*. In fact, our cograph algorithm has already been extended to recognize *distance hereditary*, *P_4 -reducible*, and *P_4 -sparse* graphs [3]. *Distance hereditary* graphs are characterized by the property that the distance between any two connected vertices of an induced subgraph equals their distance in the original graph. Their relationship to cographs stems from the fact that for any vertex v and integer k , the set of vertices at distance k from v induces a cograph. The family of *P_4 -reducible* graphs requires that each vertex belongs to at most one induced P_4 , and the *P_4 -sparse* graphs are defined by the restriction that any set of five vertices induces at most one P_4 . These two graph families belong to the *P_4 -heirarchy*, which is a set of graph families defined by various restrictions on the number or types of allowed induced P_4 s. Additional families in the *P_4 -heirarchy*, parity graphs and house, hole, domino (HHD)-free graphs, are related to cographs and may also be recognizable by similar algorithms.

There are several methods for *decomposing* or revealing the structure of a graph. A graph is decomposed by splitting the graph into smaller subgraphs in a well-defined manner. The ultimate goal is to decompose a graph into its most basic building blocks or *prime subgraphs*. An important consequence of the decomposition process is the construction of a tree representation of the graph (see, for example, [22]). Such a tree representation of a graph can lead to the simplification of difficult graph related problems, such as transitive orientation [24] and NP-hard problems such as maximum size clique, maximum size stable set, minimum coloring, and minimum clique covering, for restricted families of graphs.

Modular decomposition is one such decomposition scheme. Unfortunately, algorithms for modular decomposition are notorious for being either linear time and impractical or less efficient. The optimal *theoretical* algorithms (see [23, 14, 24, 12]) are linear in time; however, the best *practical* algorithms are only $\mathcal{O}(|V| + |E| \log |V|)$ [14, 20] and $\mathcal{O}(|V| + |E| \cdot \alpha(|V|, |E|))$ [14], where $\alpha(|V|, |E|)$ is the inverse Ackermann function. Recent results are proving hopeful that a simple linear time algorithm may exist for finding the modular decomposition of a graph. The motivation for this discussion lies in the relationship between cographs and modular decomposition.

Cographs are exactly those graphs that are completely decomposable with respect to modular decomposition; i.e., they are the graphs whose prime subgraphs are single vertices. Additionally, the cotree is exactly the modular decomposition tree for cographs. The cotree is algorithmically significant as a tool for dynamic programming algorithms, thereby reducing NP-hard problems such as coloring, clique detection, hamiltonicity, etc., on arbitrary graphs to fast polynomial time problems on cographs [9]. As we will show, our results provide a very natural cotree construction algorithm from two LexBFS orderings of the vertices. Since the algorithm was easily extended to construct tree representations of distance hereditary, *P_4 -reducible*, and *P_4 -sparse* graphs, a natural question is whether there exists a generalized algorithm to construct the modular decomposition tree for any graph. Note that any linear time modular decomposition algorithm, such as [23, 14, 24, 12], immediately yields a somewhat complicated, linear time cograph recognition algorithm.

As previously noted, our algorithm is based on two LexBFS orderings of the ver-

tex set. The real beauty of LexBFS lies in its straightforward implementation and linear running time. Not only can it be implemented in linear time for an input graph G , but using only the edges of G one can also do a LexBFS of the complement of G in time linear in the size of G . Furthermore, more sophisticated vertex selection methods can be easily implemented with minor modifications to the original LexBFS implementation. We will define one of these variations (LexBFS⁻) in section 3. LexBFS based recognition algorithms fall into a simple algorithmic paradigm of “*order*” the vertex set and “*check*” for a graph characterizing property. The following algorithm overview for our cograph recognition algorithm illustrates the paradigm.

Algorithm sketch RECOGNIZE_COGRAPH(G)

Input: Graph G .

Output: Cotree T if G is a cograph, an induced P_4 otherwise.

 Compute σ a LexBFS of G

 Given σ , use LexBFS⁻ to compute $\bar{\sigma}$, a LexBFS of $\bar{G}$

if $\sigma, \bar{\sigma}$ satisfy the *neighborhood subset property* (NSP) on $G, \bar{G}$, respectively,

return (COTREE)

else

return (REPORT_ P_4)

end RECOGNIZE_COGRAPH

This technique of using LexBFS to construct a vertex ordering that has a graph characterizing property was first introduced by Rose, Tarjan, and Lueker [25] in 1976 for chordal graph recognition. Their algorithm produces a vertex ordering that is then tested for a property specific to chordal graphs, namely, a perfect elimination ordering [25]. More recently, this paradigm has been used to recognize interval graphs [11, 18], unit interval graphs [7, 21], and bipartite permutation graphs [5, 21]. See [6] for a survey of these algorithms as well as other applications of LexBFS.

The remainder of this paper is organized as follows. We begin by defining cograph and cotree specific notation and theory, followed by a discussion of LexBFS and the newly defined variant LexBFS⁻. The remaining sections discuss the algorithm and prove its correctness. The first of these sections proves the *neighborhood subset property* (NSP), a relationship between the neighborhoods of two sets drawn from the neighborhood and nonneighborhood with respect to a given vertex. This property leads naturally to the recognition algorithm, which is followed by an overview of the implementation details. We conclude by summarizing our results and indicating possible directions for future work.

2. Background. We will assume standard graph theoretic notation and definitions as in [28]; however, here we will highlight those of particular importance to this paper.

All graphs in this paper are undirected. Let U be a subset of V . Then $V - U$ or $V \setminus U$ is the set of vertices belonging to V and not U . The *cardinality* of U is denoted by $|U|$ and $G[U]$ is the subgraph of G , induced by U . We will denote the complement graph of $G = (V, E)$ by $\bar{G} = (V, \bar{E})$, where $\bar{E} = \{uv | u, v \in V, u \neq v, \text{ and } (u, v) \notin E\}$. The neighborhood $N(v)$, of vertex v , is the set of vertices adjacent to v , and the complement neighborhood $\bar{N}(v)$ is the set of vertices not adjacent to v in G .

We refer to a chordless path on k vertices by P_k . Suppose that $x_1x_2 \dots x_{k-1}x_k$ is a P_k . We say that x_1 and x_k are the endpoints of P_k and that $x_2, \dots, x_{k-1}$ are

the midpoints. Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, the *union* of G_1 and G_2 is the graph $G = (V_1 \cup V_2, E_1 \cup E_2)$. The *join* of G_1 and G_2 is the graph $G = (V_1 \cup V_2, E_1 \cup E_2 \cup E)$, where $E = \{u_1 u_2 | u_1 \in V_1 \text{ and } u_2 \in V_2\}$.

Given a rooted tree T and two vertices x and y in T , we say that x is an *ancestor* of y , and y is a *descendant* of x , if x lies on the path from y to the root of T . For a set of leaves S of T , we say that the *lowest common ancestor* (lca) of S is the internal node v of T such that v is the root of the smallest rooted subtree of T containing S . We say that the children of a node are *siblings*. A subset $M \subseteq V$ is a *module* of $G = (V, E)$ if for all vertices $x, y \in M$ and $v \in V - M$, $xv \in E$ if and only if $yv \in E$.

2.1. Cographs and cotrees. Figure 1 illustrates a cograph G and an embedding of the corresponding cotree T_G . Leaves of T_G represent the vertex set V , and each internal node signifies the union (0) or join (1) operations on the children. The significance of the 0(1) nodes is captured by the fact that $xy \in E$ if and only if the $\text{lca}(x, y)$ is a 1 node, as depicted in Figure 1. For the remainder of this section, assume that G refers to a cograph and T_G to its cotree.

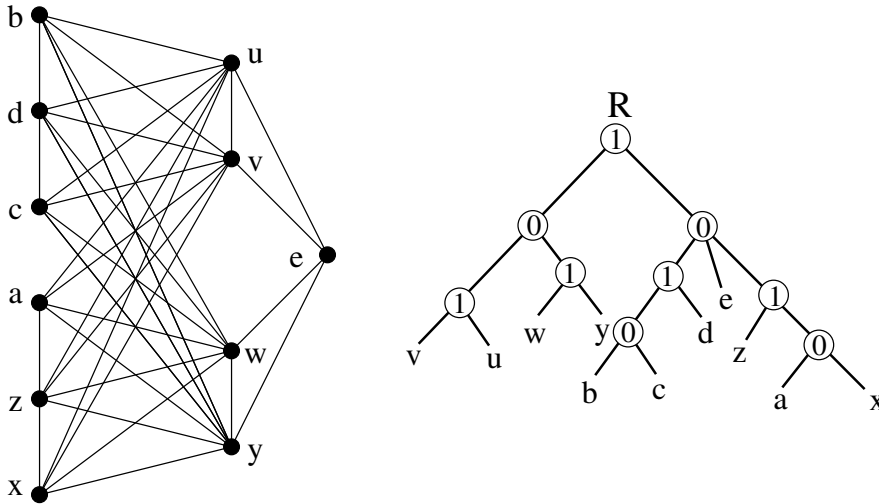


FIG. 1. (Left) A cograph G . (Right) An embedding of the cotree T_G of G .

The algorithm relies on the following properties of cographs:

1. *Complement property.* $\overline{G}$ is a cograph; i.e., cographs are closed under complementation.
2. *Complement cotree property.* The cotree $\overline{T}$ of $\overline{G}$ is exactly T with 0 and 1 nodes interchanged.
3. *Hereditary property.* Cographs are hereditary in the sense that any induced subgraph is also a cograph.
4. *Induced subtree property.* An induced rooted subtree t of T is the cotree of an induced subgraph of G .

Generally, given an internal 0 or 1 node u , we let T^u refer to the induced subtree rooted at u . However, we will need to be able to refer to more specific subtrees of a cotree. Consider the root R of T . Let x be a leaf of T . Then P_{xR} will denote the directed path of T from x to R . Note that with the exception of x , the path alternates between 0 and 1 nodes. Let $(0_1^x, \dots, 0_k^x)$ (resp., $(1_1^x, \dots, 1_{k'}^x)$) refer to the 0 nodes (resp.,

the 1 nodes) from x to R on P_{xR} . Notice that each node 0_i^x or 1_i^x has one child on P_{xR} and at least one child not on P_{xR} .

We now define the subtrees rooted on a path P_{yR} for each y . Consider 0_i^y on P_{yR} and only the children of 0_i^y that do *not* lie on P_{yR} . We can define the set of subtrees T_{0i}^y to be the induced subtrees rooted at these children of 0_i^y . The analogous definition holds for T_{1j}^y and 1_j^y . Figure 2 highlights four such sets, namely, $T_{01}^x, T_{02}^x, T_{11}^x$, and T_{12}^x , for the cotree presented in Figure 1.

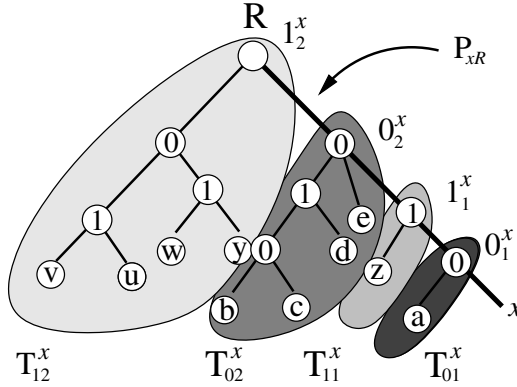


FIG. 2. Each shaded subtree indicates a different subtree T_{0i}^x or T_{1j}^x rooted on the path P_{xR} .

We conclude this section with an observation about the induced subtrees T_{0i}^y and T_{1i}^y of a cotree.

OBSERVATION 1. *For any vertex y of a cograph, the leaves of T_{0i}^y (or T_{1i}^y) define a module. More generally, a set M of vertices is a module if and only if there exists a set U of sibling nodes of the cotree such that $M = \cup_{u \in U} \{x \mid x \text{ is a leaf of } T^u\}$.*

3. LexBFS and LexBFS⁻. We begin by describing a lexicographic breadth first search (LexBFS) and then develop the new variant, LexBFS⁻.

3.1. LexBFS. For a graph $G = (V, E)$, a *LexBFS* of G , denoted $\text{LexBFS}(G)$, is a breadth first search in which preference is given to vertices whose neighbors have been visited *earliest* in the search [25]. We denote this ordering by the function $\sigma : V \rightarrow \mathbb{N}$, where $\sigma(y) = i$ indicates that y is the i th vertex of σ and inversely, $\sigma^{-1}(i) = y$. We use $u <_\sigma v$ to indicate that $\sigma(u) < \sigma(v)$.

There are several different paradigms for implementing LexBFS. The term *lexicographic* stems from the original *labeling* paradigm in which vertices pass a label to each unnumbered neighbor. The next vertex is chosen such that it has the lexicographically largest label. Another conceptualization of LexBFS involves pivots and partitioning [18]. The partitioning paradigm will prove more useful for illustrating the variant LexBFS⁻. Algorithm LEXBFS below is a variation of the version appearing in [18].

The partitioning paradigm considers a pivot (the current vertex) and all unnumbered vertices. A vertex is considered to be visited if it has been a pivot, i.e., numbered. The algorithm begins with a list L of cells initialized to a single cell containing the vertex set. Each step consists of choosing a pivot α in the leftmost (first) cell of L . The pivot is removed from its cell and the unnumbered neighborhood of α is defined to be $N'(\alpha) = \{v \mid v \in N(\alpha) \text{ and } v \text{ unnumbered}\}$. For each cell $P \in L$, all

Algorithm LEXBFS(G, τ)

Input: A graph $G = (V, E)$ and an initial ordering τ of the vertices.

Output: An ordering σ of the vertices of G .

```

1.  $L \leftarrow (V)$ 
2.  $i \leftarrow 1$ 
3. while  $\exists P_i \neq \emptyset$  in  $L = (P_1, \dots, P_k)$  do
4.   Let  $P_l$  be the leftmost nonempty cell
5.   Remove the first vertex  $x$  (smallest with respect to  $\tau$ ) from  $P_l$ 
6.    $\sigma(x) \leftarrow i$ 
7.    $i \leftarrow i + 1$ 
8.   for each cell  $P_j, j \geq l$  do
9.     Let  $P' = \{v \mid v \in N(x) \cap P_j\}$ ;
10.    if  $P'$  is nonempty and  $P' \neq P_j$ , then
11.      Remove  $P'$  from  $P_j$ 
12.      Insert  $P'$  to the left of  $P_j$  in  $L$ 
13.    end for
14. end while
15. return  $(\sigma)$ 
end LEXBFS

```

members of $N'(\alpha)$ belonging to P are removed from P and inserted into a new cell P' positioned in L immediately to the left of P .

For example, if x is the first vertex selected by a LexBFS on G (see Figure 1 and Table 3.1), then the vertex set is split into two cells $\boxed{y \ u \ v \ w \ z}$ and $\boxed{d \ e \ c \ a \ b}$, where $\boxed{y \ u \ v \ w \ z}$ is moved ahead of $\boxed{d \ e \ c \ a \ b}$ since each vertex in $\{y, u, v, w, z\}$ is adjacent to x , whereas each vertex in $\{d, e, c, a, b\}$ is not adjacent to x .

The LexBFS terminates when all vertices have been visited. It outputs an ordering σ of the vertices. The numbering convention in this paper follows that of [11], where the ordering σ is the order in which vertices are visited during the LexBFS. Throughout the paper, when we refer to the “leftmost vertex” or “first” vertex in a set, the ordering is with respect to σ . Table 3.1 walks through a LexBFS of the cograph of Figure 1.

The following lemma is a very useful characterization of LexBFS orderings.

LEMMA 3.1 (see [15, 1]). Four point condition. *An ordering σ is a LexBFS of G if and only if for any $x <_\sigma y <_\sigma z$ such that $xz \in E$ but $xy \notin E$, there exists v such that $v <_\sigma x$, $vy \in E$, and $vz \notin E$.*

In addition, when restricted to cographs, LexBFS orderings can be shown to be *umbrella-free*. An *umbrella* of an ordering σ is a triple of vertices x, y, z such that $x <_\sigma y <_\sigma z$, where $xz \in E$, $xy \notin E$, and $yz \notin E$.

LEMMA 3.2. *Any LexBFS σ of a cograph G is umbrella-free (there does not exist $x <_\sigma y <_\sigma z$ such that $xz \in E$, $xy \notin E$, and $yz \notin E$).*

Proof. Let x, y, z be an umbrella such that x is leftmost in σ . By the four-point condition, there exist $u <_\sigma x$ such that $uy \in E$ but $uz \notin E$. Note that $ux \notin E$; otherwise, $\{u, x, y, z\}$ would induce a P_4 . But then u, x, y is also an umbrella, contradicting the choice of x . $\square$

We now introduce the notion of *slices*. Notice that when a pivot is selected from the leftmost cell, *any* vertex in the cell could be the pivot as each vertex in the cell

TABLE 3.1

Step by step LexBFS of G from Figure 1. The resulting ordering is $\sigma : x y w z u v a d c b e$.

$\sigma(\alpha)$	α	$N'(\alpha)$	Cells
			x d y u e v w c a z b
1	x	{y u v w z}	y u v w z d e c a b
2	y	{w z d e c a b}	w z u v d e c a b
3	w	{z d e c a b}	z u v d e c a b
4	z	{u v a}	u v a d e c b
5	u	{v a d e c b}	v a d e c b
6	v	{a d e c b}	a d e c b
7	a	{ }	d e c b
8	d	{c b}	c b e
9	c	{ }	b e
10	b	{ }	e
11	e	{ }	

has the same set of numbered neighbors. We say that these vertices form a *slice*. The following definitions formalize this notion.

Given an ordering σ of the vertices, for each vertex x , we can define subsets of the neighborhood of x ,

$$N_i(x) = \{u \mid u \in N(x) \text{ and } \sigma(u) < i\},$$

which contain the neighbors of x appearing before the i th vertex in the ordering σ . When the i th pivot of a LexBFS is selected, $N_i(x)$ is the same for each vertex x in the leftmost cell P_i (the following definition establishes that these vertices form a slice).

DEFINITION 3.3. A slice (or set of tied vertices) of a LexBFS ordering σ is a set of consecutive vertices $S = \{u \mid i \leq \sigma(u) \leq j\}$ such that for any $u \in S$, $N_i(u) = N_i(\sigma^{-1}(i))$, and for any v , $j < \sigma(v)$, or $\sigma(u) < i$, $N_i(v) \neq N_i(\sigma^{-1}(i))$.

Examples of slices of a LexBFS are the whole set V and the neighborhood of the first vertex visited by the search. Notice that there are exactly $|V|$ slices, one for each vertex of the search, or one for each occurrence of a “leftmost” cell in the LexBFS procedure. Consider the LexBFS obtained in Table 3.1. The different bracket styles indicate the slices constructed during the sweep and highlight those at a common nesting depth:

$$(3.1) \quad \sigma : [x(y\{w\langle z \rangle\}\{u\langle v \rangle\})(a)(d\{c\langle b \rangle\}\{e\})].$$

As shown by the following lemma, slices of a LexBFS ordering exhibit nice structural properties.

LEMMA 3.4. Let S be a slice of a LexBFS σ of a graph G . Then σ restricted to the vertices of S is a LexBFS of the subgraph induced by those vertices.

Since slices form the foundation of the recognition algorithm, we will need notation to refer to any slice of a LexBFS. The slice starting at vertex x will be denoted by $S(x)$. A *subslice* is a slice nested within a slice. In the following, some of the identified subslices may be empty. Notice that $S(x) \cap N(x)$ is a subslice of $S(x)$ which we denote by $S^A(x)$. Once the final vertex of $S^A(x)$ has been used as a pivot, the remaining vertices of $S(x)$ are grouped into cells $S(x) \cap \overline{N}(x) = S_1(x), S_2(x), \dots$, where all vertices in $S_i(x)$ have the same neighborhood in $S^A(x)$. The $S_i(x)$ are called x -cells. We let x_i be the first vertex, as chosen by σ , in $S_i(x)$. Note that $S_1(x)$ is a slice in

σ , whereas $S_j(x), j \geq 2$, could be the union of slices since there may be some edges from vertices in some $S_i(x), i < j$, to vertices in S_j . Since x_i is identified by σ , the four-point condition immediately establishes the following observation.

OBSERVATION 2. *Consider v -cell $S_j(v)$ as identified by σ . If some vertex in $S_j(v)$ other than v_j has a neighbor in $S_i(v)$ for $i < j$, then v_j also has a neighbor in $S_h(v), h \leq i$. Thus to find the leftmost vertex in $\overline{N}(v)$ that is adjacent to some vertex in $S_j(v)$, we have only to look at v_j 's neighbors.*

For cographs, each $S_i(x)$ is a maximal slice of $S(x)$, since, as shown in Lemma 3.6 below, any edge between two x -cells in $S(x) \cap \overline{N}(x)$ implies the existence of a P_4 . Equation (3.2) illustrates the subslices of $S(x)$ as defined by the LexBFS ordering of Table 3.1. (The $S_i(x)$ are slices since G is a cograph.) In particular, $S^A(x) = \{y, w, z, u, v\}$ and $S^N(x) = \langle S_1(x), S_2(x) \rangle$, where $S_1(x) = \{a\}$ and $S_2(x) = \{d, c, b, e\}$:

$$(3.2) \quad \mathbf{x} \ [\mathbf{y} \ \mathbf{w} \ \mathbf{z} \ \mathbf{u} \ \mathbf{v}] \ [\mathbf{a}] \ [\mathbf{d} \ \mathbf{c} \ \mathbf{b} \ \mathbf{e}].$$

Again it is noted that in a LexBFS ordering, the slice $S(x)$ is defined for each vertex x and that the sets $S^A(x)$ and $S^N(x)$ may be empty sets. Consider a, u , and c in (3.1); $S(a) = \{a\}$, so $S^A(a) = \emptyset$ and $S^N(a) = \emptyset$; $S^N(u)$ is empty in $S(u)$, while for vertex c the set $S^A(c)$ is empty.

DEFINITION 3.5. *If S is a consecutive set of vertices in a LexBFS ordering σ and x is the first vertex of S , then the left neighborhood of S is*

$$N_{<}(S) = \{y | y <_{\sigma} x \text{ and } y \in N(z) \ \forall z \in S\}.$$

LEMMA 3.6. *Let $G = (V, E)$ be a cograph and σ a LexBFS sweep of G . Then for any vertex v and $i < j$*

$$(3.3) \quad \forall x \in S_i(v), \forall y \in S_j(v), xy \notin E \text{ and}$$

$$(3.4) \quad N_{<}(S_i(v)) \supset N_{<}(S_j(v)).$$

Proof. We begin by showing that for any $x \in S_i(v)$ and $y \in S_j(v)$, $xy \notin E$. Assume that $xy \in E$ and i is the smallest index such that there exist $j > i$, $y \in S_j(v)$, and $xy \in E$. As $y \notin S_i(v)$ there exists $u \in S^A(v)$ such that $ux \in E$ and $uy \notin E$. As both x and y are nonadjacent to v , $\{v, u, x, y\}$ induces a P_4 , contradicting that G is a cograph, and thus (3.3) holds. Finally, Lemma 3.2 and (3.3) show that there does not exist $u <_{\sigma} x$ such that $uy \in E$ but $ux \notin E$, thereby proving (3.4). $\square$

A property similar to Lemma 3.6, stated differently, has been observed in [13].

3.2. LexBFS⁻. In the above implementation of LexBFS, Algorithm LEXBFS selects the pivot x from the leftmost cell, and for each cell P_j , the vertices in $N(x) \cap P_j$ are removed from P_j and inserted into a new cell P' that is inserted to the *left* of P_j . Notice that inserting P' to the *right* of P_j is equivalent to inserting the *nonneighbors* in P_j , $\overline{N}(x) \cap P_j$, to the left of the neighbors of x in P_j , i.e., equivalent to performing a LexBFS of the complement of G . Thus, a LexBFS of $\overline{G}$ can be accomplished in linear time *without* computing $\overline{G}$, i.e., using only the edge set of G and Algorithm LEXBFS with a small modification to line 12, namely,

$$12^- \text{Insert } P' \text{ to be the right of } P_j \text{ in } L.$$

A *LexBFS minus*, or LexBFS⁻, of a graph G is a modified version of a LexBFS of $\overline{G}$. The modification arises in how the pivots are selected. The idea is to have an initial LexBFS ordering τ of G . LexBFS⁻(G, τ) does a LexBFS of $\overline{G}$ and selects the

next pivot to be the vertex in the leftmost cell that was numbered earliest in τ . To implement this “*tie-breaking mechanism*,” simply pass τ as the initial ordering of the vertex set. Notice that the leftmost vertex in the leftmost cell *is* the vertex that was numbered earliest in τ . Therefore, a $\text{LexBFS}^-(G, \tau)$ can easily be implemented in linear time with minor modifications to Algorithm LEXBFS. It follows that a minus sweep on $\overline{G}$ can also be accomplished in linear time. Such an ordering resulting from $\text{LexBFS}^-(G, \sigma)$ will be denoted $\overline{\sigma}^-$.

The slice-cell notation for a LexBFS ordering $\overline{\sigma}$ of $\overline{G}$ follows that of a LexBFS of G . We will denote the slice starting at vertex x and its subslice of adjacent vertices by $\overline{S}(x)$, $\overline{S}^A(x)$, respectively. The remaining x -cells are $\overline{S}^N(x) = \langle \overline{S}_1(x), \dots, \overline{S}_k(x) \rangle$. For example, given the graph G in Figure 1 and $\sigma : xyzwzuvadcbe$, LexBFS^- yields

$$\overline{\sigma}^- : \mathbf{x} [\mathbf{a} \mathbf{d} \mathbf{e} \mathbf{c} \mathbf{b}] [\mathbf{z}] [\mathbf{y} \mathbf{u} \mathbf{v} \mathbf{w}],$$

where $\overline{S}^A(x) = \{a, d, e, c, b\}$ and $\overline{S}_1(x) = \{z\}$, $\overline{S}_2(x) = \{y, u, v, w\}$.

The next lemma highlights an important observation about modules and LexBFS^- sweeps.

LEMMA 3.7. *Let M be a module of a graph G , let σ be a LexBFS of G , and let $\overline{\sigma}^-$ be the ordering resulting from $\text{LexBFS}^-(G, \sigma)$. The first vertex of M in σ is v if and only if v is the first vertex of M in $\overline{\sigma}^-$. Furthermore, $S(v)$ and $\overline{S}(v)$ are the inclusion minimal slices of σ , respectively, $\overline{\sigma}^-$, containing M .*

Proof. Let v be the first vertex of M visited by $\overline{\sigma}^-$. At the stage v is selected in $\overline{\sigma}^-$, M is contained in the slice $\overline{S}(v)$. Since v breaks the tie, v must be the first vertex of $\overline{S}(v)$ to appear in σ and therefore the first vertex of M to appear in σ .

Assume v is the first vertex of M in σ . Let $\overline{S}$ be the inclusion minimal slice of $\overline{\sigma}^-$ containing M , and let u be the first vertex of $\overline{S}$. If $u \neq v$, then $u \in \overline{S} \setminus M$. But this implies that there exists a slice $\overline{S}' \subseteq \overline{S} \setminus \{u\}$ containing M , contradicting the minimality of $\overline{S}$. $\square$

Motivated by Lemma 3.6, we wish to have a property appearing in σ or $\overline{\sigma}^-$ that indicates the presence of a P_4 in G . To set the stage for this, we are interested in whether the first vertex of a specific P_4 in a given LexBFS is an endpoint or a midpoint of the P_4 . In particular, note that a LexBFS of a P_4 starting at a midpoint of the P_4 satisfies the conditions of Lemma 3.6, whereas a LexBFS starting at an endpoint does not. First, we prove a result about the presence of P_4 s in an arbitrary LexBFS.

LEMMA 3.8. *Let $S(v)$ be the minimal slice of LexBFS σ that contains $P_4 p$. If v does not belong to any P_4 s in G , then $p \cap S^A(v)$ consists of the midpoints of p .*

Proof. By the minimality of $S(v)$, we know that v is not universal to p , since $S^A(v)$ is a slice of $S(v)$. If v is not adjacent to any vertices in p , then, again by the minimality of $S(v)$, there is some vertex $x \in S^A(v)$ that is adjacent to some, but not all, vertices in p . Thus there is an edge $yz \in p$ such that $xy \in E$, $xz \notin E$; but now we have a P_4 on v, x, y, z , contradicting v not belonging to a P_4 . Thus v is adjacent to some, but not all, vertices of p . Now it is easy to check that for v to avoid being in a P_4 , v must be adjacent to the midpoints of p , as required. $\square$

We now use this lemma to study the relative starting points of a $P_4 p$ in σ and of $\overline{p}$ in $\overline{\sigma}^-$.

COROLLARY 3.9. *Let σ be an arbitrary LexBFS of graph G , and let vertex w be the first vertex of σ that is in a P_4 . Let p be an arbitrary P_4 that contains w . If w is an endpoint of p , then in $\overline{\sigma}^-$, the first vertex of $\overline{p}$ is a midpoint of $\overline{p}$.*

Proof. Let $\overline{S}(v)$ be the minimal slice of $\overline{\sigma}^-$ containing $\overline{p}$. If $v \in \overline{p}$, then $v = w$ and the corollary clearly holds. Otherwise, since w is the leftmost vertex in σ that

belongs to a P_4 , and $w \in \overline{S}(v)$, we know that $v <_\sigma w$ and thus v does not belong to a P_4 . The result now follows by applying Lemma 3.8 to $\overline{\sigma}^-$. In particular, all such P_4 s $\overline{p}$ in $\overline{S}(v)$, regardless of whether w is an endpoint of p , have a midpoint as their first vertex is $\overline{\sigma}^-$. $\square$

The previous results deal only with the “first” P_4 s. We now show that these results cannot be generalized to all P_4 s. For example, not all P_4 s with an endpoint as the first vertex in σ have a midpoint as the first vertex of the complement P_4 in $\overline{\sigma}^-$, as demonstrated by C_5 123451, $\sigma : 1\ 2\ 5\ 3\ 4$, and $\overline{\sigma}^- : 1\ 3\ 4\ 5\ 2$. The $P_4 : 2345$ in G and the $P_4 : 3524$ in $\overline{G}$ have endpoints 2 and 3, respectively, as the first vertex in σ and $\overline{\sigma}^-$, respectively. Furthermore, it is possible for a P_4 to have both endpoints as midpoints of p and $\overline{p}$, respectively. To see this, consider a “bull” (a vertex v adjacent to the midpoints of a P_4) and a LexBFS starting at v . Note that the “bull” is self-complimentary.

4. Correctness of the algorithm. Recall that in the cograph recognition algorithm overview given previously, the algorithm consists of a LexBFS sweep σ , followed by a LexBFS $^-$ of G using σ as the input vertex ordering. A certificate in the form of a cotree or an induced P_4 is returned. This section first presents a characterization of cographs in terms of a property of σ and $\overline{\sigma}^-$. We then consider the case when the input graph G does not satisfy this property and a P_4 is returned, followed by the case where G does satisfy the property and a cotree is constructed. We show how LexBFS orderings can be used to characterize the family of cographs.

4.1. The neighborhood subset property. Motivated by Lemma 3.6, this section introduces the neighborhood subset property (NSP) and a related theorem characterizing cographs.

DEFINITION 4.1. *The local neighborhood of a v -cell $S_i(v)$, denoted by $N^l(S_i(v))$, of a slice $S(v)$ for a LexBFS σ is the set of vertices in $S(v)$ before v_i adjacent to at least one vertex in $S_i(v)$.*

Notice that in the case of cographs, it directly follows from Lemma 3.6, (3.3), that the above definition can be restricted.

OBSERVATION 3. *If σ is a LexBFS of a cograph, then $N^l(S_i(v)) = N_{<}(S_i(v)) \cap S^A(v)$.*

DEFINITION 4.2. *A LexBFS sweep satisfies the NSP if and only if*

$$\forall v \in V, \forall i < j, \text{ such that } S_j(v) \neq \emptyset, N^l(S_i(v)) \supset N^l(S_j(v)).$$

As pointed out in the previous section, a LexBFS of a P_4 starting at a midpoint of the P_4 does not violate the NSP. However, as shown in the following theorem, by looking at both σ and $\overline{\sigma}^-$ we are guaranteed to find a P_4 that does violate the property.

THEOREM 4.3 (neighborhood subset theorem). *Let σ be a LexBFS ordering of a graph $G = (V, E)$, and let $\overline{\sigma}^-$ be the sweep LexBFS $^-(G, \sigma)$. Then G is a cograph if and only if σ and $\overline{\sigma}^-$ satisfy the NSP.*

Proof. ($\Rightarrow$) Assume that G is a cograph. Lemma 3.6 and Observation 3 prove that σ satisfies the NSP. As $\overline{G}$ is a cograph, applying Lemma 3.6 and Observation 3 on $\overline{G}$ and $\overline{\sigma}^-$ proves that the NSP holds for $\overline{\sigma}^-$.

($\Leftarrow$) Assume that G is not a cograph; i.e., there exists an induced P_4 . First notice that the complement $\overline{p}$ of a P_4 p is a P_4 and that the midpoints of p are precisely the endpoints of $\overline{p}$. The following properties are the core arguments of the proof.

CLAIM 1. Let $p = abcd$ be an induced P_4 of G . Notice that the endpoints are a and d . If $S(a)$ contains p , then there exists a pair of integers i, j such that $S_i(a)$ and $S_j(a)$ fail the NSP.

Proof. Since a is adjacent to b but not adjacent to c or d , $b \in S^A(a)$ and there exists a pair of integers i, j such that $c \in S_i(a)$ and $d \in S_j(a)$. Since b is adjacent to c but not to d , i and j are distinct integers. If $c <_\sigma d$ (i.e., $i < j$) as $cd \in E$, $bc \in E$, and $bd \notin E$, $N^l(S_i(a)) \not\supseteq N^l(S_j(a))$. Otherwise, $d <_\sigma c$ (i.e., $j < i$), and the four-point condition applied on $b <_\sigma d <_\sigma c$ implies the existence of a vertex $x \in S^A(a)$, $x <_\sigma b$, such that $xd \in E$ and $xc \notin E$. Thereby $N^l(S_j(a)) \not\supseteq N^l(S_i(a))$, failing the NSP. $\square$

CLAIM 2. Let $p = abcd$ be an induced P_4 of G . If $S(v)$ with v not in any P_4 is the inclusion-minimal slice of LexBFS σ containing p , then there exists a pair of integers i, j such that $S_i(v)$ and $S_j(v)$ fail the NSP.

Proof. By Lemma 3.8, $p \cap S^A(v) = \{b, c\}$; thus $a \in S_i(v)$ and $d \in S_j(v)$. Finally, since $b \in N(a) \setminus N(d)$ and $c \in N(d) \setminus N(a)$, the local neighborhoods $N^l(S_i(v))$ and $N^l(S_j(v))$ fail the NSP. $\square$

We now return to the proof that if G is not a cograph, then the NSP fails. Since G is not a cograph, there exists a P_4 p in G or, equivalently, $\bar{p}$ in $\bar{G}$. We let w be the leftmost vertex in σ that is in some P_4 and let $p = abcd$ be any P_4 containing w . Examine $\bar{p}$ in $\bar{G}$ and let $\bar{S}(v)$ be the inclusion-minimal slice containing $\bar{p}$. Three distinct cases have to be considered:

1. Vertex v is an endpoint of $\bar{p}$. Then Claim 1 applies.
2. Vertex v is a midpoint, say, a , of $\bar{p}$. Thus a is the leftmost vertex of p in σ and is an endpoint of p . Let $S(u)$ be the inclusion minimal slice of σ containing p . If $u = w = a$, Claim 1 applies; otherwise, by the choice of w , u is not in any P_4 , and Claim 2 applies.
3. Vertex v does not belong to $\bar{p}$. By the LexBFS⁻ rule, v is to the left of $\{a, b, c, d\}$ in σ and thus is not in any P_4 , by our choice of p . Thus Claim 2 applies. $\square$

We now show that the NSP for σ can be tested in linear time.

LEMMA 4.4. Given the list of local neighborhoods, $N^l(S_i(v))$, $i > 0$, for each vertex v , sorted with respect to the same ordering (say, σ), the NSP of σ can be tested in $\mathcal{O}(n + m)$.

Proof. For each vertex v and for each i , first check that $|N^l(S_{i+1}(v))| < |N^l(S_i(v))|$; if not, report failure. Otherwise, simultaneously walk the lists $N^l(S_i(v))$ and $N^l(S_{i+1}(v))$. If the current vertex x of $N^l(S_{i+1}(v))$ is not the current vertex of $N^l(S_i(v))$, then walk $N^l(S_i(v))$ until x is reached. If x is not reached, then the inclusion $N^l(S_{i+1}(v)) \subset N^l(S_i(v))$ is not satisfied and failure is reported. If $N^l(S_{i+1}(v))$ is walked without reporting failure, then report success.

Since each local neighborhood is searched at most twice, the above algorithm runs in time linear in the total size of the local neighborhoods. Since there are at most n slices, there are at most n local neighborhoods. The size of each local neighborhood of a v -cell $S_i(v)$ is bounded by the number of “left” edges from vertices in $S_i(v)$. Therefore, summing the sizes of all the neighborhood sets results in $\mathcal{O}(n + m)$ complexity. $\square$

4.2. Reporting a P_4 from σ or $\bar{\sigma}^-$ (when G fails the NSP). We now consider the certificate returned by the algorithm, an induced P_4 , when the input graph is not a cograph.

LEMMA 4.5. Let σ be a LexBFS ordering failing the NSP for v -cells of $S^N(v)$. If j is the smallest integer such that $S_j(v)$ and $S_{j+1}(v)$ fail the NSP, then one of the

following induced P_4 s is in G :

$$vw y v_{j+1} \quad \text{or} \quad v_j w y v_{j+1} \quad \text{or} \quad y v w v_j,$$

where v_i is the first vertex chosen by σ in $S_i(v)$, $w <_\sigma v_j$, $w \in (N(v_j) \setminus N(v_{j+1})) \cap S^A(v)$, and $y <_\sigma v_{j+1}$, $y \in (N^l(v_{j+1}) \setminus N^l(v_j))$.

Proof. By the definition of v -cells, there exists $w \in S^A(v)$ such that $w \in N(v_j) \setminus N(v_{j+1})$. We now show that there exists $y <_\sigma v_{j+1}$ such that $y \in N^l(v_{j+1}) \setminus N^l(v_j)$. Since $S_j(v), S_{j+1}(v)$ fail the NSP, there exists $x \in S_{j+1}(v)$ adjacent to $z \in N^l(v_{j+1}) \setminus N^l(v_j)$. If $v_{j+1}z \in E$, then set $y = z$. Otherwise, by the four-point condition on z, v_{j+1}, x , there exists $z' <_\sigma z$ such that $z'v_{j+1} \in E, z'x \notin E$. Thus by the definition of $S_{j+1}(v)$, $z' \in \bar{N}(v)$. If $z' \in S_j(v)$, then set $y = z'$. If $z' \in N^l(v_j)$ and thus $z' <_\sigma v_j$, we have contradicted the choice of $j, j+1$. Without loss of generality, we assume y is the *rightmost* such vertex. There are two cases; y is either a neighbor of v or a nonneighbor.

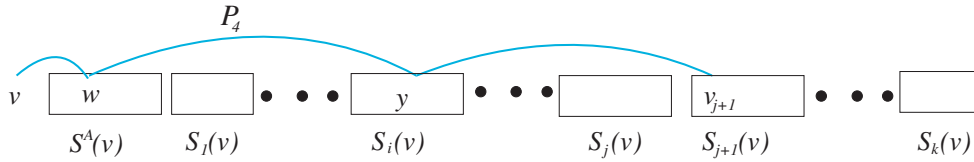


FIG. 3.

1. y is a nonneighbor of v . There exists $i < j+1$ such that $y \in S_i(v)$ (see Figure 3). Since j is minimum, w adjacent to v_j implies that w is also adjacent to y . Note that v is a neighbor of w ; v is not a neighbor of y or v_{j+1} ; w is a neighbor of y but not v_{j+1} ; and y is a neighbor of v_{j+1} . Hence, $\{v, w, y, v_{j+1}\}$ is an induced P_4 .
2. y is a neighbor of v . Then $y \in S^A(v)$, and since y is rightmost, v_j and v_{j+1} are not adjacent (see Figure 4). It follows that the subgraph induced by the set of vertices $\{v, w, y, v_j, v_{j+1}\}$ contains an induced P_4 . More precisely, if w and y are adjacent, $v_j w y v_{j+1}$ is an induced P_4 ; otherwise $v_j w v y v_{j+1}$ is an induced P_5 , containing the induced P_4 $y v w v_j$, thereby completing the proof. $\square$

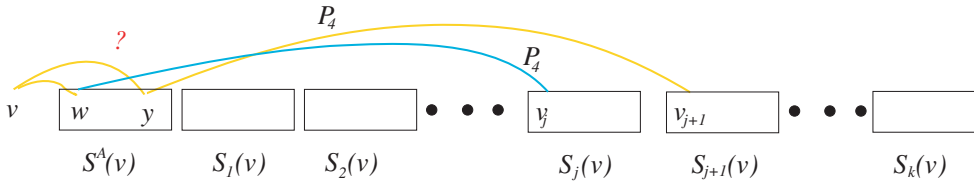


FIG. 4.

We now have the following simple algorithm $\text{Report\_}P_4$ to output an induced P_4 once the neighborhood subset test returns a vertex v and an index $j \geq 1$ such that the first pair of v -cells failing the NSP are $S_j(v)$ and $S_{j+1}(v)$. Note that our description is for σ only; see the comments at the end of section 5 for the minor modifications needed to produce a P_4 , if the NSP failure is in $\bar{\sigma}^-$.

Algorithm REPORT- $P_4(\sigma, \bar{\sigma}^-)$

Input: Vertex v and smallest index j such that $S_j(v), S_{j+1}(v)$ fail the NSP in σ

Output: An induced P_4 .

Choose $w \in (N^l(v_j) \setminus N^l(v_{j+1})) \cap S^A(v)$

Choose y to be the rightmost vertex such that $y \in N^l(v_{j+1}) \setminus N^l(v_j)$

if $yv \notin E$

return ($vw y v_{j+1}$)

else if $wy \in E$

return ($v_j w y v_{j+1}$)

else

return ($y v w v_j$)

end REPORT- P_4

LEMMA 4.6. *Reporting a P_4 , given a pair of v -cells $S_j(v)$ and $S_{j+1}(v)$ of σ failing the NSP, takes at most $\mathcal{O}(\Delta)$ time, where Δ is the maximum degree of the graph.*

Proof. Assume that the local neighborhoods are sorted according to the LexBFS ordering σ . To locate vertices $w \in (N^l(v_j) \setminus N^l(v_{j+1})) \cap S^A(v)$ and $y \in N^l(v_{j+1}) \setminus N^l(v_j)$, simply walk the local neighborhoods $N^l(v_j)$ and $N^l(v_{j+1})$, each of which are no larger than Δ . The adjacency test between y and w takes Δ time simply by walking the adjacency list of w or of y . $\square$

4.3. Constructing the cotree from $\sigma, \bar{\sigma}^-$ (when G satisfies the NSP).

We now assume that the graphs we are operating on are cographs. We will refer to an arbitrary cograph as G and its cotree as T with root R . We begin by relating slices $S_i(x)$ to modules of G and subtrees T_{0i}^x of G . This relationship results in a simple cotree construction algorithm.

Let x be an arbitrary vertex of G and consider the LexBFS(G) starting at x , resulting in σ , and the ordering $\bar{\sigma}^-$ produced by LexBFS $^-(G, \sigma)$. The following observation follows directly from the neighborhood subset theorem.

OBSERVATION 4. *Each $S_i(x)$ and $\bar{S}_j(x)$ induces a module in G .*

Furthermore, there is a well-defined relationship between the leaves of an induced subtree in T and the vertices in each $S_i(x)$ or $\bar{S}_j(x)$.

LEMMA 4.7. *Let x be the first vertex of a LexBFS σ of G . For any i , the slice $S_i(x)$ contains exactly the leaves of T_{0i}^x .*

Proof. Consider σ and the stage of the sweep where x and each vertex of $S^A(x)$ have been numbered. Since x is the first vertex of σ , only nodes adjacent to x have been numbered. If a vertex in $S^A(x)$ is adjacent to the leaves of T_{0i}^x , then by the neighborhood subset theorem, the vertex is adjacent to all leaves in T_{0j}^x , where $j < i$; see Figure 2. Therefore, the leaves of T_{01}^x must belong to the slice $S_1(x)$. By the neighborhood subset theorem, no leaf in T_{0i}^x is adjacent to any leaf in $T_{0j}^x, j \neq i$. Consequently, we can use the same argument to show that the next slice $S_2(x)$ contains exactly the leaves of T_{02}^x and, inductively, that the i th slice $S_i(x)$ contains the leaves of the subtree T_{0i}^x . $\square$

COROLLARY 4.8. *Let x be the first vertex of a LexBFS σ of a cograph G , and let τ be any LexBFS of $\bar{G}$ that starts at x (for example, $\bar{\sigma}^-$). Then, in τ , for any j , $\bar{S}_j(x)$ contains exactly the leaves of T_{1j}^x .*

Proof. Since the cotree of $\bar{G}$ is exactly the cotree of G with 1 and 0 nodes interchanged, Lemma 4.7 applied to τ implies that $v \in \bar{S}_i(x)$ if and only if $v \in$

T_{1i}^x . $\square$

Lemma 4.7 and Figure 2 motivate the following corollary.

COROLLARY 4.9. *Given σ and $\bar{\sigma}^-$ of a cograph G , if x is the first vertex of σ , then the path P_{xR} can be reconstructed from the sequences $S^N(x)$ and $\bar{S}^N(x)$.*

Proof. Lemma 4.7 applied to σ and to $\bar{\sigma}^-$ shows that each $S_i(x)$ and $\bar{S}_i(x)$ slice defines the set of leaves of each T_{0i}^x and T_{1i}^x subtree. To embed P_{xR} , it remains only to determine which of 0_1^x and 1_1^x is the parent of the leaf x in the cotree. Let u (resp., v) be the first vertex of $S_1(x)$ (resp., $\bar{S}_1(x)$). It is straightforward to see that 0_1^x is the parent of leaf x if and only if u is adjacent to v . $\square$

These results, along with the hereditary property of cographs, motivate the following preliminary recursive cotree construction algorithm.

Algorithm TREE (module M , G)

Input: A module M of cograph G .

Output: A Cotree T of M .

```

if  $M = \{v\}$ 
    return ( $\{v\}$ )
for any  $v \in M$ , compute:
    [ $v, S^A(v), S_1(v), \dots, S_k(v)$ ] with respect to  $\sigma$ , an arbitrary LexBFS of  $G[M]$ 
    [ $v, \bar{S}^A(v), \bar{S}_1(v), \dots, \bar{S}_h(v)$ ] with respect to  $\bar{\sigma}^-$ , the LexBFS $^-$  of  $\bar{G}[M]$ 
    with respect to  $\sigma$ 
    if  $S^A(v) = \emptyset$  ( $\bar{S}^A(v)$ , resp.),
        return ( $v \textcircled{0} \text{ TREE}(S_1(v), G)$ ) (resp., ( $v \textcircled{1} \text{ TREE}(\bar{S}_1(v), G)$ ))
    else for any  $a \in S_1(v)$  and  $b \in \bar{S}_1(v)$ 
        if  $ab \in E$ , then
            ( $\star$ ) return ( $v \textcircled{0} \text{ TREE}(S_1(v), G) \textcircled{1} \text{ TREE}(\bar{S}_1(v), G) \textcircled{0} \text{ TREE}(S_2(v), G) \dots$ )
        else
            ( $\star\star$ ) return ( $v \textcircled{1} \text{ TREE}(\bar{S}_1(v), G) \textcircled{0} \text{ TREE}(S_1(v), G) \textcircled{1} \text{ TREE}(\bar{S}_2(v), G) \dots$ )

```

Note that the notation “ $x \textcircled{1} T_1 \textcircled{0} T_2 \dots$ ” is parsed as “ $((x \textcircled{1} T_1) \textcircled{0} T_2) \dots$ ” and means that x and T_1 are children of a 1 node, this 1 node and T_2 are the children of a 0 node, etc. In addition, we point out that if T_1 is rooted at a 1 node, the $\textcircled{1}$ operation will *merge* the two 1 nodes together and likewise for the 0 node case. The correctness of the algorithm follows directly from Lemma 4.7, Observation 4, Corollary 4.9, and the properties of cographs.

LEMMA 4.10. *If G is a cograph, Algorithm TREE($V(G)$, G) returns T , the cotree of G .*

Proof. The proof is a simple induction argument. Certainly, a single leaf is the cotree of the graph on one vertex. So now consider when M is the entire vertex set $V(G)$. Corollary 4.9 proves that v and the $\textcircled{1}$ and $\textcircled{0}$ nodes in line ($\star$) or ($\star\star$) construct the path P_{vR} for the cotree of the cograph G . The remainder of the cotree consists of the subtrees whose parent nodes lie on P_{vR} . Since each slice $S_i(v)$ or $\bar{S}_i(v)$ is a module and furthermore, by the hereditary property of cographs, induces a cograph, by the inductive assumption, TREE($S_i(v)$, G) will return the cotree representing $G[S_i(v)]$. In addition, by Lemma 4.7, the cotree returned by TREE($S_i(v)$, G) is exactly the subtree T_{0i}^v of T . An analogous argument using Corollary 4.8 holds for $\bar{S}_j(v)$, completing the proof. $\square$

For complexity reasons, it is impractical to compute LexBFS and LexBFS⁻ orderings of each module in G . To meet the linear time bound, TREE(V(G), G) needs to be altered to use only the two orderings σ and $\bar{\sigma}^-$, a LexBFS of G and a LexBFS⁻ of G using σ . Lemma 4.11 and Corollary 4.12 show that in certain situations, the two sweeps are sufficient.

LEMMA 4.11. *Given σ and $\bar{\sigma}^-$ of G and $S_i(v)$, for arbitrary v and i , if $S_i(v) = S(y)$ contains exactly the leaves of T_{0i}^v , then $S_j(y)$ contains exactly the leaves of T_{0j}^y for all j .*

Proof. Since $S(y) = S_i(v) = T_{0i}^v$, $S_i(v)$ induces a module and a cograph. Therefore, applying Lemma 4.7 to $S(y)$ implies that $S_j(y)$ contains exactly the leaves of T_{0j}^y for all j . $\square$

Complementing Lemma 4.11 results in Corollary 4.12.

COROLLARY 4.12. *Given σ and $\bar{\sigma}^-$ of G and $\bar{S}_i(v)$ for arbitrary v and i , if $\bar{S}_i(v) = \bar{S}(y)$ contains exactly the leaves of T_{1i}^v , then $\bar{S}_j(y)$ contains exactly the leaves of T_{1j}^y for all j .*

This raises the following questions: If $S(y) = S_i(v) = T_{0i}^v$, does $\bar{S}_j(y)$ contain exactly the leaves of T_{1j}^y for each j ? Or, if $\bar{S}(y) = \bar{S}_i(v) = T_{1i}^v$, does $S_j(y)$ contain exactly the leaves of T_{0j}^y for each j ? We now answer these questions with the following observation and lemmas.

Since the leaves of an induced subtree of a cotree induce a module, the next observation follows immediately from Lemma 3.7.

OBSERVATION 5. *Given σ and $\bar{\sigma}^-$ of G , if z is the leftmost vertex with respect to σ in T_{0i}^y then $T_{0i}^y \subseteq S(z)$ and $T_{0i}^y \subseteq \bar{S}(z)$. The same is true for z , the leftmost vertex with respect to σ in T_{1j}^y .*

Lemma 4.13 claims that if $S_i(v)$ contains the leaves of a subtree T_{0i}^v , then in $\bar{\sigma}^-$, the only vertices in the inclusion minimal slice containing $S_i(v)$ either belong to $S_i(v)$ or are independent of $S_i(v)$.

LEMMA 4.13. *Consider σ and $\bar{\sigma}^-$ of G and any vertex v such that $S(v)$ is a module and any index i such that $S_i(v)$ contains exactly the leaves of T_{0i}^v . Let $\bar{S}(u)$ be the inclusion-minimal slice containing $S_i(v)$; then $\bar{S}(u) - S_i(v)$ is a subset of $\bar{S}^A(u)$ and for all $z \in \bar{S}(u) - S_i(v)$, z is independent of the vertices in $S_i(v)$.*

Proof. By Lemma 3.7 and the fact that $S(v)$ induces a module, $S(v) \subseteq \bar{S}(v)$. Since each vertex in $S_i(v)$ is not adjacent to v and $S_i(v) \subseteq \bar{S}(u) \subseteq \bar{S}(v)$, every vertex in $\bar{S}(u)$ must be uniform with respect to v . Therefore, for all $y \in \bar{S}(u)$, y and v cannot be adjacent, implying that $\text{lca}(y, v) = 0$. Now consider any vertex $z \in \bar{S}(u) - S_i(v)$. Since $\text{lca}(z, v) = 0$, then $z \in T$, where T is rooted at the child of a 0 node (not 0_i^v) on P_{vR} . Hence, the lca of z with any leaf in T_{0i}^v is a 0 node. Since $u \in S_i(v)$ and every vertex in $\bar{S}(u) - S_i(v)$ is not adjacent to u , it must be the case that $\bar{S}(u) - S_i(v) \in \bar{S}^A(u)$ and for all $z \in \bar{S}(u) - S_i(v)$, z is independent of the vertices in $S_i(v) = T_{0i}^v$. $\square$

Lemma 4.13 indicates that we can recurse on $\bar{S}(u)$ to construct the 1-rooted subtrees of T_{0i}^v . Similarly, Corollary 4.14 shows that the complement case holds as well, i.e., that if $\bar{S}_j(v)$ contains exactly the leaves of T_{1j}^v , then we can recurse on the inclusion-minimum slice containing $\bar{S}_j(v)$ in σ to construct the 0 rooted subtrees of T_{1j}^v . The proof of the corollary is very similar to that of the lemma.

COROLLARY 4.14. *Consider σ and $\bar{\sigma}^-$ of G and any vertex v such that $\bar{S}(v)$ is a module and any index j such that $\bar{S}_j(v)$ contains exactly the leaves of T_{1j}^v . Let $S(u)$ be the inclusion-minimal slice containing $\bar{S}_j(v)$; then $S(u) - \bar{S}_j(v)$ is a subset of $S^A(u)$ and for all $y \in S(u) - \bar{S}_j(v)$, y is universal to the vertices in $\bar{S}_j(v)$.*

LEMMA 4.15. *Consider σ and $\bar{\sigma}^-$ of G . For any vertex v in V , $S_i(v)$ contains exactly the leaves of T_{0i}^v and $\bar{S}_j(v)$ contains exactly the leaves of T_{1j}^v .*

Proof. Throughout this proof we will apply Observation 4 to conclude that the various $S_j(v), \bar{S}_j(v)$ are modules. By Lemma 4.7 and Corollary 4.8, if v is the first vertex of σ , the claim holds. The proof will be by induction on the nested depth of the subslices. Assume that $S(y)$ is a module and that $S_i(y)$ contains exactly the leaves of T_{0i}^y and $\bar{S}_j(y)$ contains exactly the leaves of T_{1j}^y . We will consider each case separately.

Let v be the first vertex of $S_i(y)$, so $S(v) = S_i(y)$. By Lemma 4.11, each $S_k(v)$ contains exactly the vertices of the subtrees T_{0k}^v of T_{0i}^y . We now consider the 1-rooted subtrees of T_{0i}^y . Since v is the first vertex in $S_i(y)$, Observation 5 implies that $S_i(y) \subseteq \bar{S}(v)$. Let the set of vertices in $\bar{S}(v)$ but *not* in $S_i(y)$ be denoted by X . Lemma 4.13 applied to $S(y)$ and $S_i(y)$ implies that $X \in \bar{S}^A(v)$ and, furthermore, that every vertex in X is independent of the vertices in $\bar{S}(v) \cap S_i(y)$. Therefore, we can use an argument similar to that in the proof of Corollary 4.8 to show that each $\bar{S}_j(v)$ contains exactly the leaves of the 1-rooted subtrees T_{1j}^v of T_{1j}^y .

Similarly, if v is the first vertex of $\bar{S}_j(y)$, then by Corollary 4.12, each $\bar{S}_k(v)$ contains exactly the leaves of the subtrees T_{1k}^v of T_{1j}^y . Observation 5 implies that given $\bar{S}_j(y) = T_{1j}^y$ and that v is the first vertex of $\bar{S}_j(y)$, $\bar{S}_j(y) \subseteq S(v)$. Corollary 4.14 requires that every vertex in $S(v) - \bar{S}_j(y)$ belongs to $S^A(v)$ and is universal to the vertices in $\bar{S}_j(y)$. Therefore, we can use an argument analogous to that in the proof of Lemma 4.7 to show that each $S_m(v)$ contains exactly the leaves of the 0-rooted subtrees T_{0m}^v of T_{1j}^y . $\square$

We now have the tools to prove that a recursive cotree construction algorithm can be achieved using only σ and $\bar{\sigma}^-$.

LEMMA 4.16. *Given σ and $\bar{\sigma}^-$ of a cograph G , the cotree T of G can be constructed.*

Proof. Certainly the claim holds for the trivial case of a single vertex. We will argue inductively by considering the largest slice possible and showing that the recursion holds.

Let x be the first vertex of σ and therefore of $\bar{\sigma}^-$ as well, and consider each $S_i(x)$. By Lemma 4.7, $S_i(x)$ contains exactly the leaves of T_{0i}^x . In addition, since $S(x) = V$, $S(x)$ induces a module. Let u be the first vertex of $S_i(x)$. By Lemmas 4.11 and 4.15, each $S_j(u)$ contains the leaves of the subtree T_{0j}^u , and each $\bar{S}_h(u)$ contains exactly the leaves of T_{1h}^u . Therefore, we can build the path P_u , from u to the root of T_{0i}^x as in the proof of Corollary 4.9 and by induction, build each subtree T_{0j}^u and T_{1h}^u rooted on P_u .

The argument for the complement is nearly identical. By Corollary 4.8, $\bar{S}_m(x)$ contains exactly the leaves of T_{1m}^x . If u is the first vertex of $\bar{S}_m(x)$, then by Corollary 4.12 and Lemma 4.15, each $\bar{S}_j(u)$ contains the leaves of the subtree T_{1j}^u and each $S_h(u)$ contains exactly the leaves of T_{0h}^u . Therefore, again, we can build the path from u to the root of T_{1m}^x as in the proof of Corollary 4.9 and by induction, build each subtree T_{1j}^u and T_{0h}^u . $\square$

Notice that at each stage of building a path in the cotree and recursing on the subtrees, for each vertex v , we need only to know the first vertex of each of the $S_i(v)$ and $\bar{S}_j(v)$ subslices. We will denote the first vertex of the i th subslice by $v[i]$ and the first vertex of the j th subslice in the complement by $\bar{v}[j]$. Each list of first vertices is generated during the appropriate LexBFS sweep at no extra cost to LexBFS complexity. COTREE illustrates the algorithm.

Algorithm COTREE (vertex v)

Input: Vertex v of a slice $S(v)$ of cograph G with respect to σ .

Output: A Cotree T of $v \bigcup_{\forall i} S_i(v) \bigcup_{\forall j} \bar{S}_j(v)$.

```

  if ( $v[1] = \text{nil}$  and  $\bar{v}[1] = \text{nil}$ ),
    return (  $\{v\}$  )
  if ( $v[1] = \text{nil}$ ),
    return (  $v \textcircled{1}$  COTREE( $\bar{v}[1]$ ) )
  if ( $\bar{v}[1] = \text{nil}$ ),
    return (  $v \textcircled{0}$  COTREE( $v[1]$ ) )
  if  $v[1]\bar{v}[1] \in E$ , then
    (★) return (  $v \textcircled{0}$  COTREE( $v[1]$ )  $\textcircled{1}$  COTREE( $\bar{v}[1]$ )  $\textcircled{0}$  COTREE( $v[2]$ )  $\dots$  )
  else
    (★★) return (  $v \textcircled{1}$  COTREE( $\bar{v}[1]$ )  $\textcircled{0}$  COTREE( $v[1]$ )  $\textcircled{1}$  COTREE( $\bar{v}[2]$ )  $\dots$  )

```

LEMMA 4.17. *Given $v[i]$ and $\bar{v}[j]$ for each vertex v and for all i, j , Algorithm COTREE(v) belongs to $\mathcal{O}(n + m)$.*

Proof. Notice that each vertex v is passed to COTREE at most once and that for each v , the size of the lists $v[i]$ and $\bar{v}[i]$ is bounded by the degree of v . Further, one iteration of the algorithm requires $\mathcal{O}(\deg(v))$ time to execute line (★) or (★★). The test that $v[1]\bar{v}[1] \in E$ can be completed in constant time. To see this, note that $v[1]$ must belong to $\bar{S}^A(v)$ and, further, by the LexBFS⁻ tie breaking mechanism, must be the first vertex in $\bar{S}^A(v)$. This implies that $v[1]$ will be the first vertex of $N^l(\bar{S}_1(v))$ (with respect to the adjacencies of G) if and only if $v[1]\bar{v}[1] \in E$. The local neighborhood for each $\bar{S}_i(v)$ with respect to the adjacencies of G can be generated during the LexBFS⁻ at no extra cost.

Therefore, summing for all v , the complexity is $\mathcal{O}(n + m)$. $\square$

COROLLARY 4.18. *Let x be the first vertex of a pair of LexBFS $\sigma, \bar{\sigma}^-$ of a cograph G . Then Algorithm COTREE(x) produces a cotree of G in $\mathcal{O}(n + m)$ time.*

5. Putting things together. Theorem 4.3 ensures the correctness of Algorithm RECOGNIZE_COGRAPH. The subroutines Algorithm COTREE and Algorithm REPORT_P₄ have already been described.

Algorithm RECOGNIZE_COGRAPH(G)

Input: Graph G .

Output: Cotree T if G is a cograph, an induced P_4 otherwise.

Let τ be an arbitrary ordering of V

Compute LexBFS(G, τ) resulting in σ of G , with first vertex x

Given σ , compute LexBFS⁻(G, σ), a LexBFS of $\bar{G}$, resulting in $\bar{\sigma}^-$

if $\sigma, \bar{\sigma}^-$ satisfy the NSP on $G, \bar{G}$, respectively,

 return (COTREE(x))

else

 return (REPORT_P₄($\sigma, \bar{\sigma}^-$))

end RECOGNIZE_COGRAPH

To show the linearity of the above algorithm, it remains to explain how the NSP can be checked in linear time. Although it can be done during the LexBFS search itself as explained in [3], it is simpler to do the test as a postprocessing. This test is

divided into two steps.

Step 1

Sort the adjacency list of every vertex in G with respect to σ

Step 2

for each $v \in V$ (following σ increasing) where $v_2 \neq NIL$ **do**

$A \leftarrow N_{<}(v_1) \cap S^A(v)$

$i \leftarrow 1$

while $v_{i+1} \neq NIL$ **do**

$B \leftarrow N_{<}(v_{i+1}) \cap S(v)$

if B not included in A , **then**

return (NSP fails on σ, v, i)

else $A \leftarrow B; i \leftarrow i + 1$

end while

end for

return (NSP succeeds on σ)

5.1. Complexity analysis. First we note, by Observation 2, that it suffices to consider $N_{<}(v_{i+1})$ instead of $N^l(S_{i+1}(v))$. Sorting the adjacency lists according to σ as specified in step 1 takes $\mathcal{O}(|E|)$ time; simply create a new set of adjacency lists, walk through the ordering, and for each vertex v , append v to the end of the adjacency list of each vertex $y \in N(v)$, i.e., for each y in the original adjacency list of v .

Each time the neighborhood of a vertex is used in the previous procedure, the edges to visit are at the beginning of its adjacency list and can be deleted from this list after having been processed. The computations of the sets A and B can be done easily using standard partition refinement techniques as developed in [20, 18]. Hence the overall complexity of this procedure is $\mathcal{O}(m)$.

Therefore, testing the NSP on σ can be done in $\mathcal{O}(n + m)$ time complexity. Observing that the NSP of a LexBFS ordering $\bar{\sigma}^-$ on $\bar{G}$ can be checked on G by reversing the order inclusion to be tested, it is straightforward to check the NSP of $\bar{\sigma}^-$ on $\bar{G}$ in linear time $\mathcal{O}(n + m)$. Similar arguments hold for finding a P_4 in G when $\bar{\sigma}^-$ on $\bar{G}$ does not fulfill the NSP. It can be done in $\mathcal{O}(\Delta)$ time, where Δ is the maximum degree of the graph G .

6. Concluding remarks. We conclude by first stating the major result of this paper followed by a discussion of future research directions directly related to these results.

MAIN RESULT. *There exists a simple, linear time, LexBFS based recognition algorithm for the family of cographs that returns a certificate in the form of a cotree or an induced P_4 .*

While optimal cograph recognition algorithms already exist, we feel our new LexBFS based recognition algorithm improves upon previous algorithms for several reasons. It follows a simple paradigm: *order* the vertices, *check* for a characterizing property, and *return* an appropriate certificate. In addition, the algorithm is easily implemented and returns a certificate in both the positive and negative cases. Finally we feel that the algorithm has potential to be generalized and extended to other families and, possibly, modular decomposition. We realize that our new algorithm does not have one very nice feature of the original linear time cograph recognition algorithm [10], in that it is not incremental. Unfortunately, LexBFS does not lend itself

to incremental algorithms, and, as such, our algorithm cannot be used in dynamic situations.

Concerning the algorithm itself, it is natural to wonder if LexBFS is the only sweep that can be used in this way to recognize cographs. First we note that BFS is too weak since it does not guarantee the nesting of the local neighborhoods of the $\{S_i(v)\}$, and thus with such an algorithm, a cograph might not satisfy the NSP. If we try to generalize LexBFS to other maximal neighborhood searches (MNS) such as LexDFS (see [8]) and maximum cardinality search (MCS), we risk having some vertex of $S^N(v)$ before a vertex of $S^A(v)$. Note that using the 4-vertex condition of MNS [8] (namely, an ordering σ is an MNS of G if and only if, for any $x <_\sigma y <_\sigma z$ such that $xz \in E$ but $xy \notin E$, there exists v such that $v <_\sigma y$, $vy \in E$, and $vz \notin E$), it is easy to show that Lemma 3.1 holds, namely, every MNS of a cograph is umbrella-free.

Similar to P_4 -reducible and P_4 -sparse graphs, other families in the P_4 -hierarchy such as P_4 -lite, P_4 -extendible, and P_4 -tidy (see [2] for definitions) may be well suited to a recognition algorithm that is an extension of our cograph recognition algorithm. In addition, distance hereditary graphs have a simple recognition algorithm that is again based on this cograph algorithm, motivating the question of whether families of graphs related to distance hereditary graphs, such as parity graphs, $(k, +)$ -distance hereditary graphs, or HHD-free graphs, have similar recognition algorithms.

Finally, the most interesting and most challenging question is whether there exists a modular decomposition algorithm that generalizes our cograph recognition algorithm. This is a natural question to ask since cographs are the family of completely decomposable graphs with respect to modular decomposition.

Note added in proof. The cograph recognition algorithm presented in this paper does extend to a simple, linear time algorithm for modular decomposition; however, an even easier linear time modular decomposition algorithm has recently appeared (see [27]).

Acknowledgments. We are grateful to the anonymous referees and Marty Golumbic for their comments and suggestions.

REFERENCES

- [1] A. BRANDSTÄDT, F.F. DRAGAN, AND F. NICOLAI, *LexBFS orderings and powers of chordal graphs*, Discrete Math., 171 (1997), pp. 27–42.
- [2] A. BRANDSTÄDT, V.B. LE, AND J.P. SPINRAD, *Graph Classes: A Survey*, SIAM Monogr. Discrete Math. Appl. 3, SIAM, Philadelphia, 1999.
- [3] A. BRETSCHER, *LexBFS Based Recognition Algorithms for Cographs and Related Graphs Families*, Ph.D. thesis, Department of Computer Science, University of Toronto, Toronto, Canada, 2004.
- [4] A. BRETSCHER, D.G. CORNEIL, M. HABIB, AND C. PAUL, *A simple linear time LexBFS cograph recognition algorithm*, in Graph Theoretical Concepts in Computer Science (WG), Lecture Notes in Computer Science 2880, 2003, pp. 119–130.
- [5] J.-M. CHANG, C.-W. HO, AND M.-T. KO, *LexBFS ordering in asteroidal triple-free graphs*, in Proceedings of the 10th International Symposium on Algorithms and Computation (ISAAC), Lecture Notes in Comput. Sci. 1741, Springer-Verlag, Berlin, 1999, pp. 163–172.
- [6] D.G. CORNEIL, *Lexicographic breadth first search—a survey*, in Proceedings of the 30th International Workshop on Graph Theory (WG2004), Lecture Notes in Comput. Sci. 3353, Springer-Verlag, Berlin, 2004, pp. 1–19.
- [7] D.G. CORNEIL, *A simple 3-sweep LBFS algorithm for the recognition of unit interval graphs*, Discrete Appl. Math., 138 (2004), pp. 171–179.
- [8] D.G. CORNEIL AND R. KRUEGER, *A unified view of graph searching*, SIAM J. Discrete Math., 22 (2008), pp. 1259–1276.

- [9] D.G. CORNEIL, H. LERCHS, AND L.K. STEWART BURLINGHAM, *Complement reducible graphs*, Discrete Appl. Math., 3 (1981), pp. 163–174.
- [10] D.G. CORNEIL, Y. PERL, AND L.K. STEWART, *A linear time recognition algorithm for cographs*, SIAM J. Comput., 14 (1985), pp. 926–934.
- [11] D. G. CORNEIL, S. OLARIU, AND L. STEWART, *The ultimate interval graph recognition algorithm?*, in Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), ACM, New York, SIAM, Philadelphia, 1998, pp. 175–180.
- [12] A. COURNIER AND M. HABIB, *A new linear time algorithm for modular decomposition*, in Proceedings of the 19th International Colloquium on Trees in Algebra and Programming (CAAP), Lecture Notes in Comput. Sci. 787, Springer-Verlag, Berlin, 1994, pp. 68–84.
- [13] E. DAHLHAUS, *Efficient parallel recognition algorithms for cographs and distance hereditary graphs*, Discrete Appl. Math., 57 (1995), pp. 29–45.
- [14] E. DAHLHAUS, J. GUSTEDT, AND R.M. MCCONNELL, *Efficient and practical modular decomposition algorithm*, in Proceedings of the 8th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), ACM, New York, SIAM, Philadelphia, 1997, pp. 26–35.
- [15] M.C. GOLUMBIC, *Algorithmic Graph Theory and Perfect Graphs*, 2nd ed., Ann. Discrete Math. 57, Elsevier Science B.V., Amsterdam, 2004.
- [16] M.C. GOLUMBIC, A. MINTZ, AND U. ROTICS, *Factoring and recognition of read-once functions using cographs and normality, and the readability of functions associated with partial k-trees*, Discrete Appl. Math., 154 (2006), pp. 1465–1677.
- [17] C.C. GOTLIEB AND S. KUMAR, *Semantic clustering of index terms*, J. ACM, 15 (1968), pp. 493–513.
- [18] M. HABIB, R.M. MCCONNELL, C. PAUL, AND L. VIENNOT, *LexBFS and partition refinement, with applications to transitive orientation, interval graph recognition and consecutive ones testing*, Theoret. Comput. Sci., 234 (2000), pp. 59–84.
- [19] M. HABIB AND C. PAUL, *A simple linear time algorithm for cograph recognition*, Discrete Appl. Math., 145 (2005), pp. 183–197.
- [20] M. HABIB, C. PAUL, AND L. VIENNOT, *Partition refinement: An interested algorithmic tool-kit*, Internat. J. Found. Comput. Sci., 10 (1999), pp. 147–170.
- [21] P. HELL AND J. HUANG, *Certifying LexBFS recognition algorithms for proper interval graphs and proper interval bigraphs*, SIAM J. Discrete Math., 18 (2005), pp. 554–570.
- [22] B. JAMISON AND S. OLARIU, *P-components and the homogeneous decomposition of graphs*, SIAM J. Discrete Math., 8 (1995), pp. 448–463.
- [23] R.M. MCCONNELL AND J.P. SPINRAD, *Linear-time modular decomposition and efficient transitive orientation of comparability graphs*, in Proceedings of the 6th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), ACM, New York, SIAM, Philadelphia, 1994, pp. 536–545.
- [24] R.M. MCCONNELL AND J.P. SPINRAD, *Modular decomposition and transitive orientation*, Discrete Math., 201 (1999), pp. 536–545.
- [25] D.J. ROSE, R.E. TARJAN, AND G.S. LUEKER, *Algorithmic aspects on vertex elimination on graphs*, SIAM J. Comput., 5 (1976), pp. 266–283.
- [26] S. SHEW, *A Cograph Approach to Examination Scheduling*, Master’s thesis, Department of Computer Science, University of Toronto, Toronto, Canada, 1986.
- [27] M. TEDDER, D. CORNEIL, M. HABIB, AND C. PAUL, *Simpler linear-time modular decomposition via recursive factorizing permutations*, in Proceedings of the 35th International Colloquium on Automata, Languages and Programming (ICALP), Lecture Notes in Comput. Sci. 5125, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 634–645.
- [28] D.B. WEST, *Introduction to Graph Theory*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2001.

ON COMPUTING THE DISTINGUISHING NUMBERS OF PLANAR GRAPHS AND BEYOND: A COUNTING APPROACH*

V. ARVIND[†], CHRISTINE T. CHENG[‡], AND NIKHIL R. DEVANUR[§]

Abstract. A vertex k -labeling of graph G is *distinguishing* if the only automorphism that preserves the labels of G is the identity map. The *distinguishing number* of G , $D(G)$, is the smallest integer k for which G has a distinguishing k -labeling. In this paper, we apply the principle of inclusion-exclusion and develop recursive formulas to count the number of inequivalent distinguishing k -labelings of a graph. Along the way, we prove that the distinguishing number of a planar graph can be computed in time polynomial in the size of the graph.

Key words. distinguishing numbers, distinguishing labelings, planar graphs, efficient counting algorithms

AMS subject classifications. 05C78, 05C85, 05C25

DOI. 10.1137/07068686X

1. Introduction. A vertex k -labeling of graph G is a mapping $\phi : V(G) \rightarrow \{1, 2, \dots, k\}$. It is said to be *distinguishing* if the only automorphism that preserves the labels of G is the identity map. The *distinguishing number* of G , $D(G)$, is the minimum number of labels needed so that G has a distinguishing labeling. The notion of distinguishing numbers for graphs was first introduced and developed by Albertson and Collins [2]. Their focus was on determining the relationships between a graph's automorphism group and its distinguishing number. Their work has since been extended in many directions by researchers for graphs and groups (e.g., [3, 5, 8, 6, 7, 10, 12, 16, 17, 20, 21]).

Let (G, ϕ) denote the labeled version of G under the labeling ϕ . Given two distinguishing k -labelings ϕ and ϕ' of G , we say that ϕ and ϕ' are *equivalent* if there is some automorphism of G that maps (G, ϕ) to (G, ϕ') . We are interested in computing $D(G, k)$ —the number of inequivalent k -distinguishing labelings of G —which was first considered by Arvind and Devanur [4] and Cheng [9] to determine the distinguishing numbers of trees. Our motivations for studying this parameter are as follows. First, $D(G) = \min\{k : D(G, k) > 0\}$, so if we can compute $D(G, k)$ efficiently, then we can also determine $D(G)$ efficiently. The usual way of proving that $D(G) = k^*$ is to present a distinguishing k^* -labeling of G and then argue that G has no distinguishing labelings that use $k < k^*$ labels. Counting the number of inequivalent distinguishing k -labelings of G provides us with an altogether different method for solving $D(G)$. Second, when G is connected, solving for $D(G, k)$ allows us to determine the distinguishing number of H , where $H = \alpha G$ (i.e., H consists of α copies of G). To distinguish H , each copy of G must be assigned a distinguishing labeling. Additionally, no two copies of G can be assigned equivalent distinguishing labelings. Hence, $D(H) =$

*Received by the editors March 30, 2007; accepted for publication (in revised form) February 26, 2008; published electronically September 4, 2008. A preliminary version of this paper appeared in *Proceedings of the 8th Nordic Combinatorial Conference*, Aalborg University, Aalborg, Denmark, 2004.

<http://www.siam.org/journals/sidma/22-4/68686.html>

[†]The Institute of Mathematical Sciences, Chennai 600 113, India (arvind@imsc.res.in).

[‡]Department of Electrical Engineering and Computer Science, University of Wisconsin–Milwaukee, Milwaukee, WI 53211 (ccheng@uwm.edu).

[§]College of Computing, Georgia Institute of Technology, Atlanta, GA 30332 (nikhil@cc.gatech.edu).

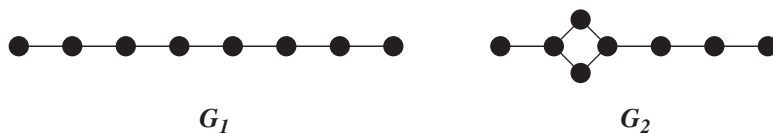


FIG. 1. An example of two graphs with the same number of vertices, isomorphic automorphism groups, and identical distinguishing numbers but a different number of inequivalent k -labelings: $D(G_1, k) = k^4(k^4 - 1)/2$, but $D(G_2, k) = k^7(k - 1)/2$.

$\min\{k : D(G, k) \geq \alpha\}$. Finally, researchers have noted that two labels are sufficient for distinguishing many nonrigid graphs (e.g., [3, 12, 16]). The number of inequivalent distinguishing k -labelings of graphs provides one more level of granularity that enables us to differentiate between graphs with the same distinguishing numbers. For example, consider the two graphs shown in Figure 1. They have the same number of vertices, their automorphism groups are isomorphic, and they can be distinguished with two labels. Yet, $D(G_1, k) = k^4(k^4 - 1)/2$, but $D(G_2, k) = k^7(k - 1)/2$, so with two labels at most 120 copies of G_1 can be distinguished compared to 64 copies for G_2 . In this sense, G_1 is less symmetric than G_2 because k labels can distinguish more copies of G_1 than G_2 for any $k \geq 2$.

To solve for $D(G, k)$, we apply two of the most common techniques for counting—the principle of inclusion-exclusion (PIE) and recursion. We show that when G 's automorphisms are known and the size of its automorphism group $\text{Aut}(G)$ is $O(\log n)$, where n is the number of vertices in G , then a straightforward application of PIE can determine $D(G, k)$ efficiently. We then modify the technique so that when $\text{Aut}(G)$ is isomorphic to Z_t (the cyclic group of order t), D_t (the dihedral group of order $2t$), $Z_t \times Z_2$, or $D_t \times Z_2$, the value of $D(G, k)$ can be computed in time polynomial in n, t , and $\log k$. Consequently, we are able to prove that if G is a triconnected planar graph, then $D(G, k)$ and $D(G)$ can be determined efficiently. Next, by viewing G via a tree decomposition T_G that is made up of G 's cut vertices, separating pairs, and triconnected components, we show that $D(G, k)$ can be determined recursively. To implement this technique efficiently for a family of graphs, several ingredients are necessary including efficient algorithms for testing graph isomorphism and finding the automorphisms of a graph's triconnected components. Since these algorithms exist for planar graphs, we arrive at the main result of the paper—that when G is a planar graph, then $D(G, k)$ and $D(G)$ can be computed efficiently.

In their introductory paper, Alberston and Collins [2] raised the issue of determining the computational complexity of $\text{DIST} = \{(G, k) | G \text{ has a distinguishing } k\text{-labeling}\}$. Currently, the best known result about DIST , which is due to Russell and Sundaram [19], is that it belongs to AM, the set of languages for which there are Arthur and Merlin games. This result essentially follows from the fact that testing graph rigidity is in AM. When G is restricted to certain graph families, however, DIST can belong to P . For example, distinguishing numbers of cycles, hypercubes [5, 8], and acyclic graphs [4, 9] can be computed efficiently. Our main result extends this further— DIST belongs to P when G is a planar graph. Our work complements that of Fukuda, Negami, and Tucker [12] on triconnected planar graphs where they show that, except for seven graphs, all graphs in this family have a distinguishing number of at most 2.

In the next section, we give basic results that will be used throughout the paper. In section 3, we show how the PIE can be used to determine $D(G, k)$. In section 4, we develop recursive formulas for a tree decomposition of G that eventually lead to the

computation of $D(G, k)$. We conclude in section 5. We note that our algorithms for computing $D(G, k)$ have G and k as input; hence, when we say that they are efficient, we mean that they run in time polynomial in the size of G and $\log k$. Additionally, these algorithms involve addition and multiplication. In cases where the numbers used are functions of k , their values never exceed k^n , where n is the number of nodes in graph G ; i.e., the numbers have at most $n \log k$ bits. Thus, in our analysis, we assume that each addition takes $O(n \log k)$ time and each multiplication takes $O(n^2 \log^2 k)$ time in the worst case.

2. Basic notions. Suppose that ϕ and ϕ' are two distinguishing labelings of G . Since (labeled) graph isomorphism is an equivalence relation, we shall say that ϕ and ϕ' are *equivalent* if $(G, \phi) \cong (G, \phi')$; that is, there is an automorphism of G that maps (G, ϕ) to (G, ϕ') . Let $\mathcal{L}(G, k)$ denote the set of all distinguishing k -labelings of G , $L(G, k)$ the size of $\mathcal{L}(G, k)$, and $D(G, k)$ the number of equivalence classes of $\mathcal{L}(G, k)$. Below, we establish the relationships between $D(G)$, $D(G, k)$, and $L(G, k)$.

LEMMA 2.1. *Let G be a graph and $\text{Aut}(G)$ its automorphism group.*

- i. $D(G) = \min\{k : L(G, k) > 0\} = \min\{k : D(G, k) > 0\}$.
- ii. $D(G, k) = L(G, k)/|\text{Aut}(G)|$.

Proof. If there is a distinguishing k -labeling of G , then the set $\mathcal{L}(G, k)$ must have at least one labeling and one equivalence class. It follows that the smallest k for which this is true must be the distinguishing number of G , proving the first part of the lemma.

To prove the second part, note that $\text{Aut}(G)$ is a group that acts on $\mathcal{L}(G, k)$. By definition, each $\phi \in \mathcal{L}(G, k)$ is preserved by only one automorphism in $\text{Aut}(G)$ —the identity automorphism. Hence, according to the orbit-stabilizer lemma, the size of the equivalence class of $\mathcal{L}(G, k)$ that contains ϕ (i.e., the orbit of ϕ) is $|\text{Aut}(G)|$. Consequently, the number of equivalence classes of $\mathcal{L}(G, k)$ is $L(G, k)/|\text{Aut}(G)|$. $\square$

Throughout this paper, we shall make use of Lemma 2.1 by viewing the problem of finding a graph's distinguishing number as a counting problem. While it may seem that computing $D(G, k)$ to find $D(G)$ requires more work than needed, the lemma below (first proved in [9]) shows that it does not if we need to distinguish multiple copies of G .

LEMMA 2.2. *Let G be a graph whose g connected components are $G_1, G_2, \dots, G_g$. Let ϕ be a labeling of G . Then ϕ is distinguishing if and only if the following two conditions hold:*

- i. ϕ when restricted to G_i is distinguishing for $i = 1, \dots, g$.
- ii. If $G_i \cong G_j$, $i \neq j$, then $(G_i, \phi|_{G_i}) \not\cong (G_j, \phi|_{G_j})$ for every pair of $i, j \in \{1, \dots, g\}$.

The following is immediate.

LEMMA 2.3. *Let G be a connected graph. If H consists of α copies of G (i.e., $H = \alpha G$), then $D(H) = \min\{k : D(G, k) \geq \alpha\}$.*

2.1. Blocks, cut vertices, separating pairs, and triconnected components. Let $G = (V, E)$ be a connected graph. Recall that G is r -connected if $|V| > r$ and, for any $X \subseteq V$ such that $|X| < r$, removing the vertices in X from G does not disconnect G ; i.e., $G - X$ remains connected. Suppose that we are interested in determining if G has some property (e.g., if it is planar). A common technique is to first decompose G into its blocks—which are either edges or 2-connected (or biconnected) subgraphs of G —and then decompose the blocks into its “triconnected components” [14]—which are parallel edges (or bonds), cycles, or 3-connected graphs.¹ It is then the

¹Unlike blocks, however, the triconnected components of a graph need not be one of its subgraphs.

triconnected components which are initially studied; the results are then assembled to infer the properties of the blocks, which in turn infer the property of G . We shall apply this technique in section 4 to determine $D(G, k)$. In particular, we shall make use of a tree T_G that captures the relationships among the cut vertices, separating pairs, and triconnected components of G to assemble the information for computing $D(G, k)$.

A *block* of G is a maximally connected subgraph of G that does not contain a cut vertex. Thus, a block of G is either an edge or a maximal biconnected subgraph of G . Furthermore, any two blocks of G have at most one vertex in common, and this vertex is a cut vertex of G . The *block-cut vertex graph* of G is a bipartite graph where one partite set consists of *b-vertices*, which correspond to the blocks of G , and the other partite set consists of *c-vertices*, which correspond to the cut vertices of G . A *b-vertex* is adjacent to a *c-vertex* if and only if the block associated with the *b-vertex* contains the cut vertex associated with the *c-vertex*. It is well known that the block-cut vertex graph of G is a tree whose leaves are *b-vertices*, and so it has a unique center. Moreover, it can be constructed in time linear in the size of G [1].

Every block of G that is biconnected can similarly be represented by a tree via its triconnected components and separating pairs. To do so, the definitions of 3-connectedness and separating pairs have to be extended to multigraphs. Our discussion closely follows the paper of Hopcroft and Tarjan [14]. Let B be a biconnected multigraph and $\{x, y\}$ be a pair of vertices in B . The set $\{x, y\}$ partitions the edge set of B in the following way: Two edges belong to the same class if and only if they lie in a path that contains neither x nor y except possibly as endpoints. The classes are called the *separation classes of B with respect to $\{x, y\}$* . If there are at least two separation classes, then the pair $\{x, y\}$ is a *separating pair of B* except when (i) there are exactly two separation classes, and one class consists of a single edge, or (ii) there are exactly three classes, each consisting of a single edge. If B is a biconnected multigraph and has no separating pairs, then B is said to be *triconnected*.

Let $\{x, y\}$ be a separating pair of B , and let the separation classes of B with respect to $\{x, y\}$ be $E_1, \dots, E_m$. An immediate consequence of the definition of separating pairs is that the classes can be divided into two groups $E' = \cup_{i=1}^k E_i$ and $E'' = \cup_{i=k+1}^m E_i$ so that both E' and E'' have at least two edges. Let $B' = (V(E'), E' \cup \{(x, y)\})$ and $B'' = (V(E''), E'' \cup \{(x, y)\})$. The graphs B' and B'' are called *split graphs of B with respect to (x, y)* , and the edges (x, y) added to both graphs are called *virtual edges*. To *split B* is to replace B by two of its split graphs. Hopcroft and Tarjan suggested denoting the i th splitting operation via the pair $\{x, y\}$ by $s(x, y, i)$ and labeling the (x, y) edges added to B' and B'' by i to differentiate this split from other splits.

Suppose that B is split, its split graphs are split, and so on until there are no more splits possible. The remaining graphs are called the *split components of B* . Clearly, they all must be triconnected; they can be grouped together as follows: the triple bonds $\mathcal{B}_{b3}$, the (simple) triangles $\mathcal{B}_t$, and the rest of the triconnected (simple) graphs $\mathcal{B}_{tg}$. Since there are many ways of splitting B , the split components of B are not necessarily unique (e.g., consider a 4-cycle). Nonetheless, this lack of uniqueness can be fixed by an operation called *merge*, which is the reverse of split. Let $B_1 = (V_1, E_1)$ and $B_2 = (V_2, E_2)$ be two split components of B that contain virtual edge $e = (x, y)$ labeled i . The graph $(V_1 \cup V_2, E_1 - \{e\} \cup E_2 - \{e\})$ is called the *merge graph of B_1 and B_2* . To *merge B_1 and B_2* is to create their merge graph. As before, the operation is denoted by $m(x, y, i)$ to differentiate it from other merge operations. So suppose that the split components of B are contained in $\mathcal{B}_{b3} \cup \mathcal{B}_t \cup \mathcal{B}_{tg}$. Merge the triple bonds in $\mathcal{B}_{b3}$ as much as possible to obtain a set of bonds $\mathcal{B}_b$. Merge the triangles in $\mathcal{B}_t$ as much as possible to obtain a set of cycles $\mathcal{B}_p$. The set of graphs in

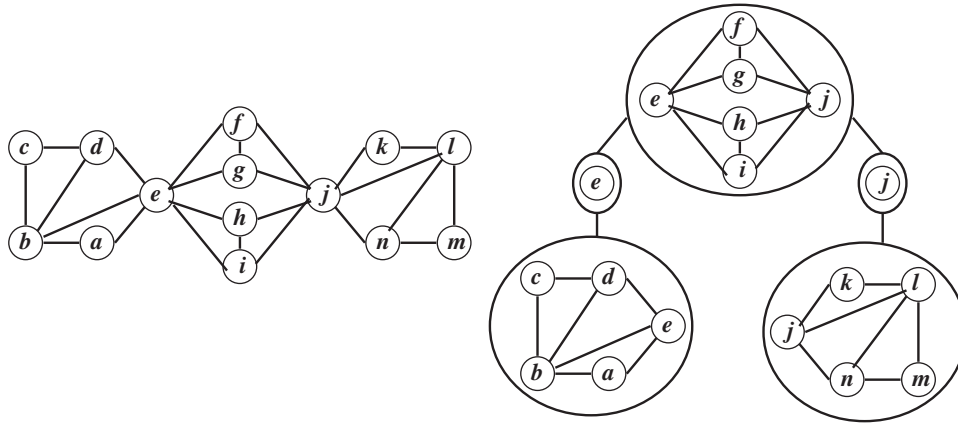


FIG. 2. A graph and its block-cut vertex graph.

$\mathcal{B}_b \cup \mathcal{B}_p \cup \mathcal{B}_{tg}$ is called the *triconnected components* of B . For example, a cycle has only one triconnected component—itsself—because the triangles obtained by splitting the cycle can be merged. The following has been proven in [14].

LEMMA 2.4. *Let B be a biconnected multigraph with $m_B \geq 3$ edges. The total number of edges in the split components of B is at most $3m_B - 6$. Additionally, the triconnected components of B are unique and can be found in time linear in the size of B .*

Lemma 2.4 implies that the order in which the split and merge operations are applied to decompose B to its triconnected components is not important—the same components are obtained.

The biconnected multigraph B can now be represented by its *triconnected component-separating pair graph*, which is a bipartite graph where one partite set consists of *t*-vertices that correspond to the triconnected components of B and the other partite set consists of *s*-vertices that correspond to B 's separating pairs which exist as virtual edges in B 's triconnected components. A *t*-vertex is adjacent to an *s*-vertex if and only if the triconnected component associated with the *t*-vertex contains the separating pair associated with the *s*-vertex. It is easy to verify that this bipartite graph must again be a tree, all of its leaves are *t*-vertices, and consequently it has a unique center. Moreover, because the triconnected components of B can be found in linear time, the tree can also be constructed in linear time.

Building a tree decomposition of G . Let G be a (simple) connected graph. Let us now build a tree decomposition of G , T_G , that incorporates the triconnected component-separating pair graph of each block of G into the block-cut vertex graph of G . Initially set T_G to be the block-cut vertex graph of G . Then, for each *b*-vertex z whose associated block is B , replace z with B 's triconnected component-separating pair graph T_B . Attach T_B to each neighbor y of z in the following manner. Let a be the cut vertex associated with y . Node a is part of one or more triconnected components and separating pairs of B . It is straightforward to check that the vertices in T_B associated with these components and pairs form a subtree which has a unique center because all of the leaves of the subtree are *t*-vertices. Connect the center of this subtree in T_B to y . For example, in Figure 2, the center of the block-cut vertex graph is a *b*-vertex which has two *c*-vertices as neighbors: One is associated with e and the other with j . When this *b*-vertex is replaced with its associated triconnected component-

From here onwards, we shall treat T_G as a rooted tree. For each node v in T_G , let T_v denote the subtree of T_G rooted at v . We define $G(T_v)$ recursively as follows: When v is a leaf node of T_G and, hence, a t -vertex, let $G(T_v)$ be the triconnected component of G associated with v . Otherwise, v is a c -, s -, or a t -vertex and has at least one child. When v is a c -vertex, construct $G(T_v)$ by taking the union of the graphs in $\{G(T_w) : w \text{ a child of } v\}$. When v is an s -vertex whose associated separating pair is $\{x, y\}$, let w_x and w_y denote the children of v whose associated cut vertices are x and y , respectively. (Note that they may not exist.) Construct $G(T_v)$ by first merging the graphs in $\{G(T_w) : w \text{ is a child of } v \text{ and is a } t\text{-vertex}\}$ via $\{x, y\}$ and then appending $G(T_{w_x})$ and $G(T_{w_y})$ (if they exist) to x and y , respectively. Finally, when v is a t -vertex whose associated triconnected component is H , construct $G(T_v)$ by first merging the graphs in $\{G(T_w) : w \text{ is a child of } v \text{ and is an } s\text{-vertex}\}$ to H and then taking the union of the resulting graph with the graphs in $\{G(T_w) : w \text{ is a child of } v \text{ and is a } c\text{-vertex}\}$.

We make a few observations about $G(T_v)$. First, if w_1 and w_2 are children of v in T_G , the vertices and edges that $G(T_{w_1})$ and $G(T_{w_2})$ have in common belong to $G(T_v)$. Thus, $G(T_v)$ consists of the structure associated with v together with connected graphs “hanging” from it; these connected graphs are exactly from the set $\{G(T_w) : w \text{ is a child of } v\}$. Second, when $v = r(T_G)$, then $G(T_v) = G$. Finally, some of the $G(T_v)$ ’s may not be subgraphs of G since some of them may contain one or more copies of the edge (x, y) , where $\{x, y\}$ is a separating pair of G but (x, y) is not an edge of G .³

In our later discussions, we will mostly be interested in the automorphisms of $G(T_v)$ that fix a cut vertex, a separating pair, or a triconnected component, and so we use $\text{Aut}(G(T_v); *)$ to denote the set of automorphisms of $G(T_v)$ that fix the structures in $*$. For example, let $\{x, y\}$ be a separating pair in $G(T_v)$. The automorphisms of $G(T_v)$ in $\text{Aut}(G(T_v); x, y)$ fix the vertices x and y , while those in $\text{Aut}(G(T_v); xy)$ fix the edge (x, y) . When H is a triconnected component in $G(T_v)$, the automorphisms of $G(T_v)$ in $\text{Aut}(G(T_v); H)$ map H to itself (i.e., the set $V(H)$ to itself), and the automorphisms in $\text{Aut}(G(T_v); H, x, y)$ map H to itself and, additionally, vertices x and y to themselves. From Lemma 2.5, we have the next lemma.

LEMMA 2.6. *Let G be a connected graph and T_G its tree decomposition. Then $\text{Aut}(G) = \text{Aut}(G; A)$, where A is the structure associated with $r(T_G)$.*

From the construction of T_G , we also have the next two lemmas.

LEMMA 2.7. *Let v be a c -vertex in T_G and a be its associated cut vertex. Let w be a child of v in T_G . The following are true:*

- (i) *If w is an s -vertex, then it is associated with some separating pair $\{a, b\}$ and $\text{Aut}(G(T_w); a) = \text{Aut}(G(T_w); a, b)$;*
- (ii) *if w is a t -vertex and its associated triconnected component is H , then H contains a and $\text{Aut}(G(T_w); a) = \text{Aut}(G(T_w); H, a)$.*

Proof. Since v and w are adjacent in T_G and v is a c -vertex while w is an s - or a t -vertex, there is a block B that contains cut vertex a and the structure associated with w . As we noted in the construction of T_B , a must be part of one or more separating pairs and triconnected components in B , and the vertices associated with these pairs and components form a subtree in T_B . Let us call this subtree $T_{B,a}$. Since w was chosen so that it is the center of $T_{B,a}$, the structure associated with w contains a .

By the way the block-cut vertex graph of G is constructed, B must be the only block in $G(T_w)$ that contains a . Hence, every automorphism of $G(T_w)$ that fixes a

³We note, though, that, in computing for $D(G, k)$, we can ignore the “other” copies of (x, y) as their multiplicity does not affect any of our computations.

must map the separating pairs and triconnected components of B that contain a to similar separating pairs and triconnected components. That is, the actions of every automorphism in $\text{Aut}(G(T_w); a)$ induces an automorphism on $T_{B,a}$. But $T_{B,a}$ has a unique center— w —which means that every automorphism in $\text{Aut}(G(T_w); a)$ must map the structure associated with w to itself. The lemma follows. $\square$

LEMMA 2.8. *Let v be an s -vertex in T_G and $\{x, y\}$ be its associated separating pair. Let w be a child of v in T_G . If w is a t -vertex whose associated triconnected component is H , then $\text{Aut}(G(T_w); x, y) = \text{Aut}(G(T_w); H, x, y)$ and $\text{Aut}(G(T_w); xy) = \text{Aut}(G(T_w); H, xy)$.*

Proof. Since v and w are adjacent in T_G and v is an s -vertex while w is a t -vertex, there is again a block B that contains the structures associated with both vertices. By the way the triconnected component-separating pair graph of B is constructed, it must be the case that H is the only triconnected component in $G(T_w)$ that contains $\{x, y\}$. Hence, every automorphism in $\text{Aut}(G(T_w); x, y)$ must map H to itself, and so $\text{Aut}(G(T_w); x, y) = \text{Aut}(G(T_w); H, x, y)$. By the same reasoning, $\text{Aut}(G(T_w); xy) = \text{Aut}(G(T_w); H, xy)$. $\square$

The following lemmas will also be useful later.

LEMMA 2.9. *The tree T_G can be constructed in $O(n^2 + nm)$ time, where n is the number of vertices and m the number of edges in G .*

Proof. Constructing G 's block-cut vertex graph and rooting it at its center takes $O(n + m)$ time. Creating the separating pairs-triconnected components graph T_B of block B takes $O(n_B + m_B)$ time, where n_B and m_B are the number of nodes and edges, respectively, in block B . Connecting T_B to T_G takes $O(c_B(n_B + m_B))$, where c_B is the number of cut vertices in block B . Thus, doing this for all blocks B takes $O(n^2 + nm)$ time since $\sum_B n_B \leq n + m$ and $\sum_B m_B = m$. $\square$

LEMMA 2.10. *Let B be a block of G with n_B vertices and $m_B \geq 3$ edges. Let $\mathcal{H}$ be the set that contains all of the triconnected components of B . For each $H \in \mathcal{H}$, let $\mathcal{S}_H$ denote the set containing the separating pairs of G in H used in the construction of T_G . Then $\sum_{H \in \mathcal{H}} |\mathcal{S}_H| = O(m_B)$ and $\sum_{H \in \mathcal{H}} |V(H)| = O(m_B)$.*

Proof. Suppose that the split operation was applied g times to B until no more splits are possible. Let $\mathcal{H}'$ contain the resulting split components. For each $H' \in \mathcal{H}'$, define $\mathcal{S}_{H'}$ as in the lemma. We note that, when B is split into two components, the separating pair used to create the split becomes part of both components. That is, each split operation contributes a value of 2 to $\sum_{H' \in \mathcal{H}'} |\mathcal{S}_{H'}|$. Hence, $\sum_{H' \in \mathcal{H}'} |\mathcal{S}_{H'}| = 2g$. Now, according to Lemma 2.4, the total number of edges in the split components in $\mathcal{H}'$ is at most $3m_B - 6$. Since a split component in $\mathcal{H}'$ has at least three edges, $g \leq m_B - 2$ and so $\sum_{H' \in \mathcal{H}'} |\mathcal{S}_{H'}| = O(m_B)$. Next, notice that $|V(H')| \leq |E(H')|$ for each $H' \in \mathcal{H}'$ so $\sum_{H' \in \mathcal{H}'} |V(H')| \leq 3m_B - 6$. Finally, because $\sum_{H \in \mathcal{H}} |\mathcal{S}_H| \leq \sum_{H' \in \mathcal{H}'} |\mathcal{S}_{H'}|$ and $\sum_{H \in \mathcal{H}} |V(H)| \leq \sum_{H' \in \mathcal{H}'} |V(H')|$, the lemma follows. $\square$

Finally, we note that we call T_G a tree decomposition of G because it really is a tree decomposition as defined by Robertson and Seymour (see Chapter 12 in [11] for an introduction). That is, if v is a node of T_G and V_v contains the vertices of the structure in G associated with v , it should be clear from our construction that the following are true: (i) $V_v \subseteq V(G)$ for each v , (ii) $\cup_v V_v = V(G)$, (iii) every edge of G has two of its endpoints in some V_v , and (iv) whenever y and z are neighbors of v , then $V_y \cap V_z \subseteq V_v$. In our discussion, however, it is important that we keep track of the actual structure associated with v and not just the vertices in V_v .

3. Counting the distinguishing k -labelings of graphs via PIE. Given a graph G and its automorphisms, we begin by applying the PIE to count its distinguish-

ing k -labelings. Unfortunately, the technique requires the computation of $\Omega(2^{|Aut(G)|})$ terms and so becomes impractical when G has many automorphisms. We show how the method can be modified when $Aut(G)$ is isomorphic to certain groups. In particular, we prove that, when G is a triconnected planar graph, $L(G, k)$, $D(G, k)$, and $D(G)$ can be computed in time polynomial in $\log k$ and the size of G .

Suppose that $Aut(G) = \{\pi_0, \pi_1, \dots, \pi_{g-1}\}$, where π_0 is the identity automorphism. Let ϕ be some k -labeling of G . We say that an automorphism π_i of G *preserves* ϕ if $\phi(v) = \phi(\pi_i(v))$ for each v of G . Clearly, π_0 preserves ϕ , and if no other automorphism of G preserves ϕ , then ϕ is a distinguishing k -labeling of G . Let $P \subseteq Aut(G)$ and $N_{\geq}(P)$ denote the number of k -labelings of G that are preserved by all of the automorphisms in P . Let $N_{=}(P)$ equal the number of k -labelings of G that are preserved by all of the automorphisms in P but no others. Thus, $L(G, k) = N_{=}(\{\pi_0\})$. According to the PIE,

$$(1) \quad N_{=}(\{\pi_0\}) = \sum_{\{\pi_0\} \subseteq P \subseteq Aut(G)} (-1)^{|P|-1} N_{\geq}(P).$$

Next, we describe a method for computing $N_{\geq}(P)$ for each $P \subseteq Aut(G)$. Suppose that $\pi_i \in P$. A k -labeling ϕ is preserved by π_i if and only if ϕ assigns the same label to v and to $\pi_i(v)$ for each vertex v in G . In fact, if there is a sequence of vertices $v_1, v_2, \dots, v_r$ such that $v_j = \pi_i(v_{j-1})$ for $j = 2, \dots, r$, then ϕ must assign all of these r vertices the same label. By extending this idea further, we arrive at the following lemma.

LEMMA 3.1. *Let $\pi_i \in Aut(G)$ and ϕ be a k -labeling of G . Let G_{π_i} be the graph whose node set is $V(G)$ and whose edge set consists of the pairs $(v, \pi_i(v)), v \in V(G)$. The automorphism π_i preserves ϕ if and only if, for each connected component in G_{π_i} , ϕ assigns the same label to all of the vertices in that component. Consequently, when $P \subseteq Aut(G)$, the automorphisms in P preserve ϕ if and only if, for each connected component in $\cup_{\pi_i \in P} G_{\pi_i}$, ϕ assigns the same label to all of the vertices in that component.*

An immediate implication of the lemma is if $\cup_{\pi_i \in P} G_{\pi_i}$ has t connected components and there are k labels available, then $N_{\geq}(P) = k^t$. We are now ready to prove the next result.

THEOREM 3.2. *Let G be a graph on n vertices and k be a positive integer. Suppose that all of the automorphisms of G are given. Then $L(G, k)$ can be computed in $O(n^3 \log^2 k + 2^{|Aut(G)|}(n \times |Aut(G)| + n \log k))$ time.*

Proof. Begin by computing and storing the values $k, k^2, k^3, \dots, k^n$. Set $L(G, k)$ to 0. For each subset P such that $\{\pi_0\} \subseteq P \subseteq Aut(G)$, (i) construct $\cup_{\pi_i \in P} G_{\pi_i}$ and find the number of its connected components t by using a breadth-first search and (ii) add $(-1)^{|P|-1} k^t$ to $L(G, k)$. According to (1), at the end of this algorithm the value of $L(G, k)$ is the number of distinguishing k -labelings of G . Computing the powers of k can be done in $O(n^3 \log^2 k)$ steps. Each iteration of the for loop takes at most $O(n \times |Aut(G)| + n \log k)$ time, where the first term in the sum accounts for the time it takes to construct $\cup_{\pi_i \in P} G_{\pi_i}$ and find its connected components and the latter term accounts for adding k^t to $L(G, k)$. Since there are $2^{|Aut(G)|-1}$ subsets P to consider, computing $L(G, k)$ takes $O(n^3 \log^2 k + 2^{|Aut(G)|}(n \times |Aut(G)| + n \log k))$ time. $\square$

COROLLARY 3.3. *Let G be a graph with n vertices and k be a positive integer. Suppose that all of the automorphisms of G are given. If $|Aut(G)| = O(\log n)$, then $L(G, k)$ can be computed in time polynomial in n and $\log k$.*

The reason why implementing the PIE formula for $L(G, k)$ can take exponential time is because there are $\Omega(2^{|Aut(G)|}) N_{\geq}(P)$ terms in the formula. Below, we demonstrate that the technique can be modified when $Aut(G)$ is isomorphic to certain groups. We consider the case when $Aut(G) \cong \Gamma$, where (i) $\Gamma = Z_t$, the cyclic group of order t , (ii) $\Gamma = D_t$, the dihedral group of order $2t$, and (iii) $\Gamma = Z_t \times Z_2$ or $D_t \times Z_2$. All will be useful when we discuss triconnected planar graphs in the next subsection. A key feature of these results is that $|Aut(G)| = O(t)$, and yet the number of $N_{\geq}(P)$ terms that must be computed to derive $L(G, k)$ is polynomial in t and not exponential in t . Before we proceed, we first prove the following lemma.

LEMMA 3.4. *Let $P \subseteq Aut(G)$ and $\langle P \rangle$ be the subgroup generated by P . Every k -labeling of G preserved by all of the automorphisms in P is also preserved by all of the automorphisms in $\langle P \rangle$.*

Proof. Let ϕ be a k -labeling of G preserved by all of the automorphisms in P . Let $\pi \in \langle P \rangle$. Since $Aut(G)$ is finite, we can write π as $\sigma_r * \sigma_{r-1} * \cdots * \sigma_1$, where $r \in \mathbf{Z}^+$ and each $\sigma_i \in P$. Since each σ_i preserves ϕ , for each vertex u of G ,

$$\phi(u) = \phi(\sigma_1(u)) = \phi(\sigma_2(\sigma_1(u))) = \cdots = \phi(\sigma_r(\cdots(\sigma_2(\sigma_1(u))))).$$

That is, $\pi = \sigma_r * \sigma_{r-1} * \cdots * \sigma_1$ preserves ϕ as well. $\square$

In the subsequent discussion, when $Aut(G) \cong \Gamma$, we shall denote the automorphisms of G as π_σ , where $\sigma \in \Gamma$, and let $\pi_\sigma * \pi_{\sigma'} = \pi_{\sigma * \sigma'}$.

When $Aut(G) \cong Z_t$. Let Z_t be the cyclic group of order t and ρ be one of its generators. Its elements are ρ^0 (the identity), $\rho, \rho^2, \dots, \rho^{t-1}$, where $\rho^i * \rho^j = \rho^{i+j \bmod t}$.

THEOREM 3.5. *Let $Aut(G) \cong Z_t$, where the prime factorization of t is $\prod_{i=1}^s p_i^{r_i}$. Suppose that a generator π_ρ of $Aut(G)$ is given. Let $P^* = \{\pi_{\rho^i} : i \in \{t/p_1, t/p_2, \dots, t/p_s\}\}$. Then*

$$L(G, k) = \sum_{P \subseteq P^*} (-1)^{|P|} N_{\geq}(P).$$

Proof. To prove the theorem, we will show that a k -labeling ϕ of G is distinguishing if and only if no automorphism in P^* preserves ϕ . One direction is obvious: If ϕ is distinguishing, all nontrivial automorphisms of G do not preserve ϕ . Since P^* contains only nontrivial automorphisms of G , the result follows. So suppose that ϕ is not distinguishing. It must be preserved by some π_{ρ^j} , $j \neq 0$. Let $g = \gcd(t, j) = \prod_{i=1}^s p_i^{t_i}$, where $0 \leq t_i \leq r_i$. We know that $\rho^g \in \langle \rho^j \rangle$. Since $j < t$, we also know that g must divide one of the numbers in $\{t/p_1, t/p_2, \dots, t/p_s\}$, say, t/p_1 ; i.e., $\rho^{t/p_1} \in \langle \rho^g \rangle$. By Lemma 3.4, it follows that if π_{ρ^j} preserves ϕ , then π_{ρ^g} also preserves ϕ , which implies that $\pi_{\rho^{t/p_1}}$ does so as well. That is, some automorphism in P^* preserves ϕ . By applying the PIE, $L(G, k) = \sum_{P \subseteq P^*} (-1)^{|P|} N_{\geq}(P)$. $\square$

When $Aut(G) \cong D_t$. Let D_t be the dihedral group of order $2t$. If we let the generators of D_t be the rotation ρ and reflection τ , then the elements of D_t are ρ^0 (the identity), $\rho^1, \dots, \rho^{t-1}$, $\tau\rho^0, \tau\rho^1, \dots, \tau\rho^{t-1}$, where $\tau^2 = \rho^0$, $\tau\rho^i = \rho^{-i}\tau$, and $\rho^i * \rho^j = \rho^{i+j \bmod t}$.

THEOREM 3.6. *Let $Aut(G) \cong D_t$, where the prime factorization of t is $\prod_{i=1}^s p_i^{r_i}$. Suppose that generators π_ρ and π_τ of $Aut(G)$ are given. Let $P^* = \{\pi_{\rho^i} : i \in \{t/p_1, t/p_2, \dots, t/p_s\}\}$. Then*

$$N_{=}(\{\pi_{\rho^0}, \pi_{\tau\rho^i}\}) = \sum_{\{\pi_{\rho^i}\} \subseteq P \subseteq \{\pi_{\tau\rho^i}\} \cup P^*} (-1)^{|P|-1} N_{\geq}(P)$$

and

$$(2) \quad L(G, k) = \sum_{P \subseteq P^*} (-1)^{|P|} N_{\geq}(P) - \sum_{i=0}^{t-1} N_{=}(\{ \pi_{\rho^0}, \pi_{\tau \rho^i} \}).$$

Proof. We shall first prove that a k -labeling ϕ of G that is preserved by at least two nontrivial automorphisms of G is also preserved by some automorphism in the set $P^* = \{ \pi_{\rho^i} : i \in \{t/p_1, t/p_2, \dots, t/p_s\} \}$. If one of the automorphisms that preserves ϕ is preserved by π_{ρ^j} , $j \neq 0$, then by the proof of Theorem 3.5 it must also be preserved by some automorphism in P^* . If the two automorphisms that preserve ϕ are $\pi_{\tau \rho^i}$ and $\pi_{\tau \rho^j}$, where $i < j$, then $\pi_{\tau \rho^i} * \pi_{\tau \rho^j} = \pi_{\rho^{j-i}}$ also preserves ϕ . Once again, some automorphism in P^* must preserve ϕ .

To prove (2), we now consider the set of all k -labelings of G . Let sets A , B , and C consist of all k -labelings of G preserved by π_{ρ^0} only, by π_{ρ^0} and $\pi_{\tau \rho^i}$ for some $i \in \{0, 1, \dots, t-1\}$ only, and by some automorphism in P^* , respectively. Any k -labeling of G must belong to exactly one of the three sets because: (i) if it is distinguishing, it belongs to set A and if not to $B \cup C$; (ii) if it is preserved by exactly one nontrivial automorphism of G , and it is of the form $\pi_{\tau \rho^i}$, it belongs to set B ; otherwise, it belongs to set C ; (iii) finally, if it is preserved by at least two nontrivial automorphisms of G , then it belongs to set C . That is, $A \cup B \cup C$ contains all of the k -labelings of G , and no two of them have a k -labeling of G in common. Thus, $L(G, k) = |A| = k^n - |B| - |C|$.

By the way we defined set B , $|B| = \sum_{i=0}^{t-1} N_{=}(\{ \pi_{\rho^0}, \pi_{\tau \rho^i} \})$. Consider a k -labeling of G that is preserved by $\pi_{\tau \rho^i}$. From our earlier argument, we can assume that such a k -labeling is preserved by $\pi_{\tau \rho^i}$ only or by $\pi_{\tau \rho^i}$ and some other automorphism in P^* , in addition to being preserved by π_{ρ^0} . According to the PIE, this means that $N_{=}(\{ \pi_{\rho^0}, \pi_{\tau \rho^i} \}) = \sum_{\{ \pi_{\tau \rho^i} \} \subseteq P \subseteq \{ \pi_{\tau \rho^i} \} \cup P^*} (-1)^{|P|-1} N_{\geq}(P)$. Finally, C consists of all of the k -labelings of G preserved by at least one of the automorphisms in P^* . So, according to the PIE, $k^n - |C| = \sum_{P \subseteq P^*} (-1)^{|P|} N_{\geq}(P)$. Hence, $L(G, k) = \sum_{P \subseteq P^*} (-1)^{|P|} N_{\geq}(P) - \sum_{i=0}^{t-1} N_{=}(\{ \pi_{\rho^0}, \pi_{\tau \rho^i} \})$, which proves (2). $\square$

Example. Consider the cycle on n vertices C_n , where n is a prime number. Then $\text{Aut}(C_n) = D_n$ and $P^* = \{ \pi_{\rho} \}$. To solve for $L(C_n, k)$, we need the following values: $N_{\geq}(\emptyset)$, $N_{\geq}(\pi_{\rho})$, $N_{\geq}(\{ \pi_{\tau \rho^i}, \pi_{\rho} \})$, and $N_{\geq}(\{ \pi_{\tau \rho^i} \})$. Every k -labeling of C_n should be counted in $N_{\geq}(\emptyset)$ so $N_{\geq}(\emptyset) = k^n$. To solve for $N_{\geq}(\pi_{\rho})$, recall that we considered $G_{\pi_{\rho}}$, which is a graph that has only one component. Hence, $N_{\geq}(\pi_{\rho}) = k$. Similarly, $N_{\geq}(\{ \pi_{\tau \rho^i}, \pi_{\rho} \}) = k$. Finally, $G_{\pi_{\tau \rho^i}}$ consists of $(n+1)/2$ components, since any reflection of C_n fixes one vertex v and maps the equidistant vertices from v to each other. Thus, $N_{\geq}(\{ \pi_{\tau \rho^i} \}) = k^{(n+1)/2}$. From (2),

$$\begin{aligned} L(C_n, k) &= N_{\geq}(\emptyset) - N_{\geq}(\{ \pi_{\rho} \}) - \sum_{i=0}^{n-1} (N_{\geq}(\{ \pi_{\tau \rho^i} \}) - N_{\geq}(\{ \pi_{\tau \rho^i}, \pi_{\rho} \})) \\ &= k^n - k - nk^{(n+1)/2} + nk \\ &= k^n - nk^{(n+1)/2} + (n-1)k \\ &= k(k^{(n-1)/2} - 1)(k^{(n-1)/2} - (n-1)). \end{aligned}$$

Consequently, $D(C_n, k) = k(k^{(n-1)/2} - 1)(k^{(n-1)/2} - (n-1))/2n$. When $n = 5$, for example, $D(C_5, 1) = D(C_5, 2) = 0$, but $D(C_5, 3) = 12$, so $D(C_5) = 3$.

When $\text{Aut}(G) \cong Z_t \times Z_2$ or $D_t \times Z_2$. We state the following theorem without proof because the arguments are just extensions of those in Theorems 3.5 and 3.6.

THEOREM 3.7. *Suppose that the prime factorization of t is $\prod_{i=1}^s p_i^{r_i}$, the group $Z_t \times Z_2 = \{(\rho^i, \sigma^j), i \in \{0, 1, \dots, t-1\}, j \in \{0, 1\}\}$, and the group $D_t \times Z_2 = \{(\rho^i, \sigma^j), (\tau \rho^i, \sigma^j), i \in \{0, 1, \dots, m-1\}, j \in \{0, 1\}\}$. When t is odd, set $P_0^* = \{\pi_{(\rho^0, \sigma)}\}$; otherwise, set $P_0^* = \{\pi_{(\rho^0, \sigma)}, \pi_{(\rho^{t/2}, \sigma)}\}$. Let $P^* = P_0^* \cup \{\pi_{(\rho^i, \sigma^0)} : i \in \{t/p_1, t/p_2, \dots, t/p_s\}\}$.*

(i) *When $\text{Aut}(G) \cong Z_t \times Z_2$,*

$$L(G, k) = \sum_{P \subseteq P^*} (-1)^{|P|} N_{\geq}(P).$$

(ii) *When $\text{Aut}(G) \cong D_t \times Z_2$, and for $b = 0$ or 1 ,*

$$N_{=}(\{\pi_{(\rho^0, \sigma^0)}, \pi_{(\tau \rho^i, \sigma^b)}\}) = \sum_{\{\pi_{(\tau \rho^i, \sigma^b)}\} \subseteq P \subseteq \{\pi_{(\tau \rho^i, \sigma^b)}\} \cup P^*} (-1)^{|P|-1} N_{\geq}(P)$$

and

$$L(G, k) = \sum_{P \subseteq P^*} (-1)^{|P|} N_{\geq}(P) - \sum_{b=0}^1 \sum_{i=0}^{t-1} N_{=}(\{\pi_{(\rho^0, \sigma^0)}, \pi_{(\tau \rho^i, \sigma^b)}\}).$$

Remark. Since the number of prime factors of t is $O(\log t)$, the number of $N_{\geq}(P)$ terms in the formula for computing $L(G, k)$ is $O(t)$ when $\text{Aut}(G) \cong Z_t$ or $Z_t \times Z_2$ and $O(t^2)$ when $\text{Aut}(G) \cong D_t$ or $D_t \times Z_2$.

3.1. When G is a triconnected planar graph. What is interesting about the family of triconnected planar graphs is that the automorphism groups of the graphs are only of limited kinds.

LEMMA 3.8 (see [18]). *Let G be a triconnected planar graph. The automorphism group of G is isomorphic to a subgroup of one of the following groups: A_4 , A_5 , S_4 , $A_4 \times Z_2$, $A_5 \times Z_2$, $S_4 \times Z_2$, Z_t , D_t , $Z_t \times Z_2$, or $D_t \times Z_2$, for some integer t .*

Since a subgroup of a dihedral group is a cyclic group or a dihedral group, clearly the subgroups of $D_t \times Z_2$ are cyclic, dihedral, or isomorphic to $Z_{t'} \times Z_2$ or $D_{t'} \times Z_2$, where $t' \leq t$. In other words, the automorphism group of a triconnected planar graph is either bounded by a constant or is isomorphic to one of four groups only.

Additionally, because triconnected planar graphs have only unique embeddings on the plane up to equivalence,⁴ finding all of their automorphisms can also be done efficiently. We sketch one such method next. Let G be a triconnected planar graph with n vertices and m edges. Let $e = (u, v)$ be an edge of G . Let us designate its direction as being from u to v and one of the faces F that it borders as its right face. Create a copy of G , $G_{e,F}$, which specially marks e and its direction—and face F . For any edge $e' = (u', v')$ whose direction and right face F' are fixed, create an analogous graph $G_{e',F'}$, and by using a planar graph isomorphism testing algorithm determine if $G_{e,F}$ and $G_{e',F'}$ are isomorphic (where the marked edge and face of $G_{e,F}$ are mapped to the marked edge and face of $G_{e',F'}$). If so, then there is an automorphism of G that maps e to e' and F to F' ; moreover, by visiting the faces of $G_{e,F}$ and $G_{e',F'}$ in the same order, the rest of π can be determined in time linear in the size of G . Since there is a linear time isomorphism testing algorithm for planar graphs [15], each iteration of the for loop takes $O(n)$ time. Since there are $O(m)$ iterations, then in

⁴A triconnected planar graph can have two planar embeddings, one of which is a mirror image of the other.

$O(nm) = O(n^2)$ time all of the automorphisms of G can be determined. Furthermore, because each edge has two directions and two faces bordering it, the algorithm above also shows that $|Aut(G)| \leq 4m = O(n)$ when G is a triconnected planar graph.

To solve for $L(G, k)$ for triconnected planar graphs, we do the following: If $|Aut(G)| \leq 5!$, use Theorem 3.2. Otherwise, determine if $Aut(G)$ is cyclic, dihedral, or isomorphic to a direct product of a cyclic group and Z_2 or to a direct product of a dihedral group and Z_2 . If $Aut(G)$ is cyclic or dihedral, apply Theorem 3.5 or 3.6, respectively; otherwise, apply Theorem 3.7.

THEOREM 3.9. *Let G be an n -vertex triconnected planar graph. Computing $L(G, k)$ and $D(G, k)$ can be done in $O(n^3 \log^2 k + n^3 \log n + n^3 \log k)$ time. Therefore, computing $D(G)$ takes $O(n^3 \log^3 n)$ time.*

Proof. As we stated earlier, if G has at most $5!$ automorphisms, we use Theorem 3.2 to solve for $L(G, k)$ and $D(G, k)$. Otherwise, we need to determine to which of the four groups $Aut(G)$ is isomorphic. In particular, $Aut(G)$ falls into case i , where $i = 1$ if the group is cyclic, $i = 2$ if the group is isomorphic to $Z_t \times Z_2$ for some t , $i = 3$ if the group is dihedral, and $i = 4$ if the group is isomorphic to $D_t \times Z_2$ for some t . We note that there is some overlap in the four cases because, if t is odd, $Z_t \times Z_2 \cong Z_{2t}$ and $D_{2t} \times Z_2 \cong D_{4t}$. Thus, when we say that $Aut(G)$ belongs to case 2 or 4, we shall assume that t is even. We describe our algorithm $TriconnectCount(G, k)$ in Figure 4.

In the first part of our algorithm, we determine the case to which $Aut(G)$ belongs by considering the order of each element in $Aut(G)$. It is easy to verify the following facts: (i) If $Aut(G)$ has an element with order $|Aut(G)|$, it must be cyclic; (ii) if $Aut(G)$ has only three elements with order 2 (and $3 < |Aut(G)|/2$), then it belongs to case 2; (iii) if $Aut(G)$ has between $|Aut(G)|/2$ and $|Aut(G)|/2 + 1$ of its elements with order 2, it belongs to case 3. Once the appropriate case for $Aut(G)$ is determined, we set the value of t .

The second part of the algorithm begins by computing the prime factors of t , finding an element $\pi \in Aut(G)$ such that the order of π is t , and then computing $P^* = \{\pi^i : i \in \{t/p_1, t/p_2, \dots, t/p_s\}\}$. If $Aut(G)$ is cyclic or dihedral, P^* is indeed the one needed in Theorems 3.5 and 3.6, respectively, to compute $L(G, k)$. In cases 2 and 4, two more elements are missing in P^* . To understand what they are, we note that since t is even π would be of the form $\pi_{(\rho, \sigma^b)}$, where $b = 0$ or 1 and ρ and σ are generators of Z_t and Z_2 , respectively. If we set $p_1 = 2$, then $(\pi_{(\rho, \sigma^b)})^{t/2} = \pi_{(\rho^{t/2}, \sigma^0)}$, or $\pi_{(\rho^{t/2}, \sigma^1)}$, and $(\pi_{(\rho, \sigma^b)})^{t/p_i} = \pi_{(\rho^{t/p_i}, \sigma^0)}$ for $i = 2, \dots, s$. At this point, the two missing elements in P^* have order 2; they can be distinguished from the other elements of $Aut(G)$ with order 2 because they commute with every other element of $Aut(G)$ (i.e., they belong to the center of $Aut(G)$), whereas the others do not. By updating P^* , we now obtain the appropriate P^* in Theorem 3.7. Finally, for cases 3 and 4, we place all elements of $Aut(G)$ with order 2 not in P^* into set T . It is easy to check that the rest of the algorithm computes $L(G, k)$ correctly since they follow directly from the theorems we have established.

Computing and storing the powers of k take $O(n^3 \log^2 k)$ time. Finding all of the automorphisms of G takes $O(n^2)$ time. It is easy to verify that, in the rest of the algorithm, the bottleneck is in computing the value of $L(G, k)$ when $|Aut(G)| > 5!$. By applying the same analysis we used in Theorem 3.2 and noting that $|P^*| = O(\log t)$ and $|T| = O(t)$, computing $L(G, k)$ takes $O(t^2(n \log t + n \log k))$ time. Finally, because G is a triconnected graph, $|Aut(G)| = O(n)$, and so $t = O(n)$. Hence, the total run time of $TriconnectCount(G, k)$ is $O(n^3 \log^2 k + n^3 \log n + n^3 \log k)$. Once we have the value for $L(G, k)$, we simply divide it by $|Aut(G)|$ to determine $D(G, k)$. To find $D(G)$, do a binary search over the range $[1, n]$ to determine the smallest k for which $D(G, k) >$

TriconnectCount(G, k)

Input: A triconnected planar graph G with n vertices, a positive integer k .

Output: The value of $L(G, k)$.

Compute and store the values $k, k^2, k^3, \dots, k^n$.

Find all the automorphisms of G .

If $|Aut(G)| \leq 5!$

$L(G, k) \leftarrow \sum_{\{\pi_0\} \subseteq P \subseteq Aut(G)} (-1)^{|P|-1} N_{\geq}(P)$
return($L(G, k)$)

else

compute the order of each automorphism $\pi \in Aut(G)$

if there is an automorphism whose order is $|Aut(G)|$

CASE $\leftarrow 1$, $t \leftarrow |Aut(G)|$,

else

if there are only 3 automorphisms with order 2

CASE $\leftarrow 2$, $t \leftarrow |Aut(G)|/2$,

else

if there are between $|Aut(G)|/2$ and $|Aut(G)|/2 + 1$ elements with order 2

CASE $\leftarrow 3$, $t \leftarrow |Aut(G)|/2$,

else

CASE $\leftarrow 4$, $t \leftarrow |Aut(G)|/4$.

Compute the prime factors of t : $p_1, p_2, \dots, p_s$.

Find an automorphism $\pi \in Aut(G)$ whose order is t .

Compute $P^* = \{\pi^i : i \in \{t/p_1, t/p_2, \dots, t/p_s\}\}$.

If CASE = 2 or 4

add to P^* the two other automorphisms of G which belong to the center of $Aut(G)$.

If CASE = 3 or 4

let T consist of all automorphisms in $Aut(G)$ that is not in P^* whose order is 2.

$L(G, k) \leftarrow \sum_{P \subseteq P^*} (-1)^{|P|} N_{\geq}(P)$.

If CASE = 1 or 2

return($L(G, k)$)

else

while $T \neq \emptyset$

pick $\pi' \in T$ and delete π' from T

$L(G, k) \leftarrow L(G, k) - \sum_{\{\pi'\} \subseteq P \subseteq \{\pi'\} \cup P^*} (-1)^{|P|-1} N_{\geq}(P)$

return ($L(G, k)$).

FIG. 4. The algorithm for computing the number of distinguishing k -labelings of a triconnected planar graph.

0 to find $D(G)$. This adds an extra $\log n$ factor to the run time of *TriconnectCount*(G, k). $\square$

4. Computing $D(G, k)$ via recursion. In this section, we shall generalize the recursive technique (discovered independently by Arvind and Devanur [4] and by Cheng [9]) that was used to compute the distinguishing numbers of trees. The main idea behind the technique is quite simple. Let T be a tree rooted at r . Let T_v denote the subtree of T rooted at vertex v . Start by setting $D(T_v, k) = k$ for each leaf v since a single node has k distinguishing k -labelings. Then, for $i = \text{height}(T) - 1$ to 0, do the following: For all nodes v at depth i , compute $D(T_v, k)$ based on the values computed for $D(T_w, k)$, w a child of v in T . Thus, at the end of the algorithm

$D(T_v, k)$, which equals $D(T, k)$, is determined. To apply the above technique to a connected graph G , we will view G as a rooted tree using the tree decomposition T_G described in section 2.1. Additionally, we will also consider a generalized version of the distinguishing k -labelings of a graph which we shall define shortly. Finally, we will develop recursive formulas that relate the number of (generalized) inequivalent distinguishing k -labelings of $G(T_v)$ with those of $G(T_w)$, w a child of v in T_G .

Let Γ be a subgroup of $\text{Aut}(G)$. We say that a labeling ϕ of G is Γ -distinguishing if no nontrivial automorphism in Γ preserves ϕ and that two labelings ϕ and ϕ' of G are *equivalent with respect to Γ* if some automorphism in Γ maps (G, ϕ) to (G, ϕ') . Let $\mathcal{L}(G, k; \Gamma)$ be the set consisting of the Γ -distinguishing k -labelings of G , $L(G, k; \Gamma)$ be the size of $\mathcal{L}(G, k; \Gamma)$, and $D(G, k; \Gamma)$ be the number of equivalence classes of $\mathcal{L}(G, k; \Gamma)$ with respect to Γ . When $\Gamma = \text{Aut}(G; *)$ as defined in section 2.1, we shall refer to $\mathcal{L}(G, k; \Gamma)$, $L(G, k; \Gamma)$, and $D(G, k; \Gamma)$ as $\mathcal{L}(G, k; *)$, $L(G, k; *)$, and $D(G, k; *)$, respectively. Finally, when (x, y) is an edge of G , we will at times differentiate between the cases when a k -labeling of G assigns x and y the same or different colors. When we do so, we will place a subscript next to $\mathcal{L}$, L , and D ; the subscript is 1 if x and y are assigned the same color and is 2 otherwise. Thus, $\mathcal{L}_1(G, k; xy)$ consists of all k -labelings of G in $\mathcal{L}(G, k; xy)$ that assigned x and y the same color, etc. It is easy to verify that the following version of Lemma 2.1 remains true.

LEMMA 4.1. *Let G be a graph and Γ be a subgroup of $\text{Aut}(G)$. Then $D(G, k; \Gamma) = L(G, k; \Gamma)/|\Gamma|$.*

Given a connected graph G , we showed in section 2.1 how to construct a tree decomposition of G , T_G . The construction started with G 's block-cut vertex graph. Each b -vertex whose associated block is B is then replaced with B 's triconnected component-separating pair graph T_B and then connected to the rest of the block-cut vertex graph. Thus, T_G is made up of c -, s -, and t -vertices which represent the cut vertices, separating pairs, and triconnected components, respectively, of G . We shall now describe recursive formulas for $D(G(T_v), k; *)$ based on the type of vertex v in T_G .

THEOREM 4.2. *Let v be a c -vertex in T_G and a be the cut vertex in G associated with v . Suppose that, when all of the graphs in $\mathcal{G} = \{G(T_w) : w \text{ is a child of } v \text{ in } T_G\}$ are fixed at a , there are g isomorphic classes and the i th isomorphic class contains m_i copies of the connected graph G_i ; i.e., $\mathcal{G} = m_1 G_1 \cup m_2 G_2 \cup \dots \cup m_g G_g$. Then*

$$D(G(T_v), k; a) = k \prod_{i=1}^g \binom{D(G_i, k; a)/k}{m_i}.$$

Proof. By the way T_G was constructed, if $\mathcal{G} = m_1 G_1 \cup m_2 G_2 \cup \dots \cup m_g G_g$, then $G(T_v)$ is made up of $\sum_{i=1}^g m_i$ connected components all hanging from vertex a . It is easy to verify that ϕ is a labeling in $\mathcal{L}(G(T_v), k; a)$ if and only if ϕ assigns inequivalent labelings from $\mathcal{L}(G_i, k; a)$ to the m_i copies of G_i , $i = 1, \dots, g$, and the labels assigned to vertex a by all of the labelings are the same. This means that an equivalence class of $\mathcal{L}(G(T_v), k; a)$ is defined by (i) the label assigned to a and (ii) the set of m_i equivalence classes from $\mathcal{L}(G_i, k; a)$ that contain the labelings of the m_i copies of G_i , $i = 1, \dots, g$. There are k possible labels for a . Once the label for a is chosen, say, l , there are $D(G_i, k; a)/k$ different equivalence classes of $\mathcal{L}(G_i, k; a)$ which assign vertex a the same label. This is so because the number of equivalence classes of $\mathcal{L}(G_i, k; a)$ where a is assigned the label l must be the same for every possible value of l . It follows that there are $\binom{D(G_i, k; a)/k}{m_i}$ different sets of m_i equivalence classes that can contain the labelings assigned to the m_i copies of G_i for $i = 1, \dots, g$. By the product rule of counting, the theorem is established. $\square$

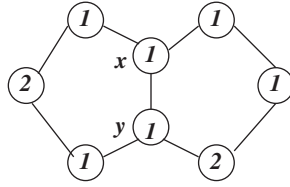


FIG. 5. In this graph, (x, y) is a separating pair. Notice that the labeling of the 5-cycle on the left does not destroy the automorphism of the 5-cycle that flips the graph along the edge (x, y) , but the labeling for the entire graph is distinguishing.

The next two theorems deal with the case when v is an s -vertex.

THEOREM 4.3. *Let v be an s -vertex in T_G and $\{x, y\}$ be the separating pair associated with v . If it exists, let w_x denote the child of v that is a c -vertex associated with x . Similarly, if it exists, let w_y denote the child of v that is a c -vertex associated with y . Suppose that, when all of the graphs in $\mathcal{G} = \{G(T_w) : w \text{ is a } t\text{-vertex and a child of } v \text{ in } T_G\}$ are fixed at x and y , there are g isomorphic classes and the i th isomorphic class has m_i copies of the connected graph G_i ; i.e., $\mathcal{G} = m_1 G_1 \cup m_2 G_2 \cup \dots \cup m_g G_g$. Then $D(G(T_v), k; x, y)$ equals*

$$k^2 \max\{D(G(T_{w_x}), k; x)/k, 1\} \max\{D(G(T_{w_y}), k; y)/k, 1\} \prod_{i=1}^g \binom{D(G_i, k; x, y)/k^2}{m_i}.$$

Proof. Once again, it is easy to verify that $\phi \in \mathcal{L}(G(T_v), k; x, y)$ if and only if ϕ assigns inequivalent labelings from $\mathcal{L}(G_i, k; x, y)$ to the m_i copies of G_i , $i = 1, \dots, g$, and the labels assigned to x and to y by all of the labelings are the same. Thus, an equivalence class of $\mathcal{L}(G(T_v), k; x, y)$ is defined by (i) the labels assigned to x and y , (ii) the equivalence classes of $\mathcal{L}(G(T_{w_x}), k; x)$ and $\mathcal{L}(G(T_{w_y}), k; y)$ that contain the labelings of $G(T_{w_x})$ and $G(T_{w_y})$, respectively, and (iii) the set of m_i equivalence classes of $\mathcal{L}(G_i, k; x, y)$ that contain the labelings of the m_i copies of G_i , $i = 1, \dots, g$. There are k^2 labels available for x and y . Once the labels are chosen, say, l_x and l_y , the number of equivalence classes of $\mathcal{L}(G_i, k; x, y)$ where x and y are assigned the said labels is $D(G_i, k; x, y)/k^2$. This is so because, in any labeling in $\mathcal{L}(G_i, k; x, y)$, the labels of the vertices in G_i other than x and y are the ones that actually destroy the nontrivial automorphisms of $\text{Aut}(G_i; x, y)$. Consequently, the number of labelings and the number of equivalence classes of $\mathcal{L}(G_i, k; x, y)$ with x and y assigned l_x and l_y must be the same regardless of the pair (l_x, l_y) . Of the $D(G_i, k; x, y)/k^2$ equivalence classes of $\mathcal{L}(G_i, k; x, y)$ that are being considered, m_i must be chosen to contain the labelings of the m_i copies of G_i . Similarly, the number of equivalence classes of $\mathcal{L}(G(T_{w_x}), k; x)$ where x is assigned the label l_x and the number of equivalence classes of $\mathcal{L}(G(T_{w_y}), k; y)$ where y is assigned the label l_y are $D(G(T_{w_x}), k; x)/k$ and $D(G(T_{w_y}), k; y)/k$, respectively. By the product rule of counting, the theorem is established. $\square$

We will also need to compute $D(G(T_v), k; xy)$. Unlike our previous characterizations, however, it is not necessarily the case that when $\phi \in \mathcal{L}(G(T_v), k; xy)$, then ϕ assigns inequivalent labelings from $\mathcal{L}(G_i, k; xy)$ to each copy of G_i . Figure 5 shows one such exception. Our approach this time is to consider the equivalence classes of $\mathcal{L}(G(T_v), k; x, y)$ and count those that *do not* belong to $\mathcal{L}(G(T_v), k; xy)$.

Consider an arbitrary graph H with an edge (x, y) , and suppose that $\text{Aut}(H; xy) \neq \text{Aut}(H; x, y)$. Let $\text{Aut}(H; x \rightarrow y, y \rightarrow x)$ denote the set of automorphisms of H that

map x to y and y to x . Notice that $\text{Aut}(H; xy)$ is the disjoint union of $\text{Aut}(H; x, y)$ and $\text{Aut}(H; x \rightarrow y, y \rightarrow x)$. Moreover, $\text{Aut}(H; x, y)$ is a subgroup of $\text{Aut}(H; xy)$, and $\text{Aut}(H; x \rightarrow y, y \rightarrow x)$ is a coset of $\text{Aut}(H; x, y)$.

Suppose that ϕ_1 and ϕ_2 belong to the same equivalence class in $\mathcal{L}(H, k; x, y)$. If some automorphism in $\text{Aut}(H; x \rightarrow y, y \rightarrow x)$ preserves ϕ_1 , then it is easy to verify that another automorphism in $\text{Aut}(H; x \rightarrow y, y \rightarrow x)$ also preserves ϕ_2 . That is, for each equivalence class in $\mathcal{L}(H, k; x, y)$, either all of the labelings do not destroy some automorphism of $\text{Aut}(H; x \rightarrow y, y \rightarrow x)$ (and so do not belong to $\mathcal{L}(H, k; xy)$) or all do (and so belong to $\mathcal{L}(H, k; xy)$). Let $\mathcal{B}(H, k; x, y)$ be the set that contains all equivalence classes of $\mathcal{L}(H, k; x, y)$ whose labelings do not belong to $\mathcal{L}(H, k; xy)$, and denote its size as $B(H, k; x, y)$. Our discussion will focus on computing $B(H, k; x, y)$ because this is the number of “bad” equivalence classes of $\mathcal{L}(H, k; x, y)$ in that they do not carry over as equivalence classes of $\mathcal{L}(H, k; xy)$.

Suppose that $\pi \in \text{Aut}(H; x \rightarrow y, y \rightarrow x)$. It is easy to verify that (i) when all of the labelings in an equivalence class of $\mathcal{L}(H, k; x, y)$ destroy all of the automorphisms in $\text{Aut}(H; xy)$, then π maps these labelings to the labelings of another equivalence class of $\mathcal{L}(H, k; x, y)$; (ii) on the other hand, when all of the labelings in an equivalence class of $\mathcal{L}(H, k; x, y)$ do not destroy the automorphisms in $\text{Aut}(H; xy)$, then π maps these labelings to themselves. In other words, under the action of $\text{Aut}(H; xy)$ the equivalence classes of $\mathcal{L}(H, k; x, y)$ either get paired up or stay singleton. The ones that get paired up are precisely the equivalence classes of $\mathcal{L}(H, k; xy)$; i.e., each equivalence class of $\mathcal{L}(H, k; xy)$ is made up of two equivalence classes of $\mathcal{L}(H, k; x, y)$. We shall say that such a pair of equivalence classes are *partners* in $\mathcal{L}(H, k; x, y)$. The ones that stay single are the equivalence classes in $\mathcal{B}(H, k; x, y)$. We have proved the following lemma.

LEMMA 4.4. *Let H be a graph with edge (x, y) and $\text{Aut}(H; xy) \neq \text{Aut}(H; x, y)$. Then $D(H, k; x, y) = 2D(H, k; xy) + B(H, k; x, y)$.*

We also need the following lemma.

LEMMA 4.5. *Let H be a graph with edge (x, y) and $\text{Aut}(H; xy) \neq \text{Aut}(H; x, y)$. Then $D_2(H, k; xy) = \binom{k}{2} D(H, k; x, y) / k^2$, so $D_1(H, k; xy) = D(H, k; xy) - \binom{k}{2} D(H, k; x, y) / k^2$.*

Proof. Consider the labelings in $\mathcal{L}_2(H, k; xy)$. Since x and y are assigned different labels, they immediately destroy all of the automorphisms in $\text{Aut}(H; x \rightarrow y, y \rightarrow x)$. Thus, we can construct a labeling in $\mathcal{L}_2(H, k; xy)$ by first choosing distinct labels for x and y , then choosing the equivalence classes of $\mathcal{L}(H, k; x, y)$ that will contain the labeling of H , and finally picking the labeling of H from the equivalence class. Hence, $L_2(H, k; xy) = k(k-1) \times D(H, k; x, y) / k^2 \times |\text{Aut}(H, k; x, y)|$, so

$$\begin{aligned} D_2(H, k; xy) &= L_2(H, k; xy) / |\text{Aut}(H, k; xy)| \\ &= k(k-1) D(H, k; x, y) / k^2 \frac{|\text{Aut}(H, k; x, y)|}{|\text{Aut}(H, k; xy)|} \\ &= k(k-1) D(H, k; x, y) / 2k^2, \end{aligned}$$

where the last equation follows from the fact that $|\text{Aut}(H; x, y)| = |\text{Aut}(H; x \rightarrow y, y \rightarrow x)|$ because $\text{Aut}(H; x \rightarrow y, y \rightarrow x)$ is a coset of $\text{Aut}(H; x, y)$ and so $|\text{Aut}(H; xy)| = 2|\text{Aut}(H; x, y)|$. Finally, since $D(H, k; xy) = D_1(H, k; xy) + D_2(H, k; xy)$, the formula for $D_1(H, k; xy)$ in the lemma follows. $\square$

THEOREM 4.6. *Let v be an s -vertex in T_G and $\{x, y\}$ be the separating pair associated with v . If it exists, let w_x denote the child of v that is a c -vertex associated with x . Similarly, if it exists, let w_y denote the child of v that is a c -vertex associated with y . Suppose that, when all of the graphs in $\mathcal{G} = \{G(T_w) :$*

w is a t -vertex and a child of v in T_G are fixed at x and y , there are g isomorphic classes and the i th isomorphic class has m_i copies of the connected graph G_i ; i.e., $\mathcal{G} = m_1 G_1 \cup m_2 G_2 \cup \dots \cup m_g G_g$. If $\text{Aut}(G(T_v); xy) = \text{Aut}(G(T_v); x, y)$, then $D(G(T_v), k; xy) = D(G(T_v), k; x, y)$. Otherwise, $\text{Aut}(G(T_v); x \rightarrow y, y \rightarrow x) \neq \emptyset$, so $G(T_{w_x}) \cong G(T_{w_y})$, and

$$D(G(T_v), k; xy) = [D(G(T_v), k; x, y) - B(G(T_v), k; x, y)]/2,$$

where $B(G(T_v), k; x, y)$ equals

$$k \max\{D(G(T_{w_x}), k; x)/k, 1\} \prod_{i=1}^g \sum_{l=0}^{\lfloor m_i/2 \rfloor} \binom{D_1(G_i, k; xy)/k}{l} \binom{[D(G_i, k; x, y) - 2D(G_i, k; xy)]/k}{m_i - 2l}.$$

Proof. When $\text{Aut}(G(T_v); xy) = \text{Aut}(G(T_v); x, y)$, $\mathcal{L}(G(T_v), k; xy) = \mathcal{L}(G(T_v), k; x, y)$, and so it follows that $D(G(T_v), k; xy) = D(G(T_v), k; x, y)$. Otherwise, $\text{Aut}(G(T_v); x \rightarrow y, y \rightarrow x) \neq \emptyset$. Hence, if they exist, $G(T_{w_x}) \cong G(T_{w_y})$ and $\text{Aut}(G_i; xy) \neq \text{Aut}(G_i; x, y)$ for $i = 1, \dots, g$. In computing $D(G(T_v), k; x, y)$, we noted that there are three sets of parameters that describe the equivalence classes of $\mathcal{L}(G(T_v), k; x, y)$: (i) the labels assigned to x and y , (ii) the equivalence classes of $\mathcal{L}(G(T_{w_x}), k; x)$ that contain the labelings of $G(T_{w_x})$ and $G(T_{w_y})$, and (iii) the set of m_i equivalence classes of $\mathcal{L}(G_i, k; x, y)$ that contain the labelings of the m_i copies of G_i for $i = 1, \dots, g$. We shall extend them to characterize the equivalence classes in $\mathcal{B}(G(T_v), k; x, y)$ —i.e., the equivalence classes of $\mathcal{L}(G(T_v), k; x, y)$ whose labelings are preserved by some automorphism in $\text{Aut}(G(T_v); x \rightarrow y, y \rightarrow x)$.

CLAIM. An equivalence class of $\mathcal{L}(G(T_v), k; x, y)$ belongs to $\mathcal{B}(G(T_v), k; x, y)$ if and only if

- (i) the labels assigned to x and y are the same for every labeling in the class,
- (ii) the equivalence classes that contain the labelings of $G(T_{w_x})$ and $G(T_{w_y})$ are the same, and
- (iii) for $i = 1, \dots, g$, the set of m_i equivalence classes of $\mathcal{L}(G_i, k; x, y)$ that contain the labelings of the m_i copies of G_i can be partitioned into l_i groups of size 2 and $m_i - 2l_i$ groups of size 1 for some $0 \leq l_i \leq \lfloor m_i/2 \rfloor$ so that the pairs of equivalence classes that belong to a group of size 2 are partners in $\mathcal{L}(G_i, k; x, y)$ and the equivalence classes that belong to a group of size 1 are in $\mathcal{B}(G_i, k; x, y)$.

Proof of claim. Let $\phi \in \mathcal{L}(G(T_v), k; x, y)$ belong to an equivalence class that satisfies conditions (i), (ii), and (iii) above. For $i = 1, \dots, g$, denote the m_i copies of G_i as $G_{i,1}, \dots, G_{i,m_i}$. Without loss of generality, assume that the equivalence classes that contain the labelings of $G_{i,2j-1}$ and $G_{i,2j}$ are partners in $\mathcal{L}(G_i, k; x, y)$ for $j = 1, \dots, l_i$ and the equivalence classes that contain the labelings of each of the remaining copies of G_i belong to $\mathcal{B}(G_i, k; x, y)$. Thus, condition (iii) implies that there are automorphisms⁵ in $\text{Aut}(G_i; x \rightarrow y, y \rightarrow x)$ that map $(G_{i,2j-1}, \phi)$ to $(G_{i,2j}, \phi)$ and vice versa for $j = 1, \dots, l_i$; similarly, there is also some automorphism in $\text{Aut}(G_i; x \rightarrow y, y \rightarrow x)$ that preserves $(G_{i,j}, \phi)$ for $j = 2l_i + 1, \dots, m_i$. Furthermore, condition (ii) implies that there are some automorphisms in $\text{Aut}(G(T_{w_x}); x)$ that map

⁵Technically, we are referring to the automorphisms implicitly defined by the isomorphisms that map $(G_{i,2j-1}, \phi)$ to $(G_{i,2j}, \phi)$ and vice versa since $G_{i,2j-1}$ and $G_{i,2j}$ are copies of G_i . We shall keep this usage throughout the proof for ease of discussion.

$(G(T_{w_x}), \phi)$ to $(G(T_{w_y}), \phi)$ and vice versa. By combining these automorphisms, we conclude that some automorphism in $\text{Aut}(G(T_v); x \rightarrow y, y \rightarrow x)$ preserves ϕ ; that is, the equivalence class that contains ϕ belongs to $\mathcal{B}(G(T_v), k; x, y)$.

On the other hand, suppose that $\phi \in \mathcal{L}(G(T_v), k; x, y)$ and some automorphism $\pi \in \text{Aut}(G(T_v); x \rightarrow y, y \rightarrow x)$ preserves ϕ . Clearly, π maps x to y and vice versa, and so condition (i) is true. It also maps $G(T_{w_x})$ to $G(T_{w_y})$ and vice versa, and so condition (ii) is true. For $i = 1, \dots, g$, if π maps $G_{i,j}$ to itself, then ϕ must have assigned $G_{i,j}$ a labeling that destroys all automorphisms in $\text{Aut}(G_i; x, y)$ since $\phi \in \mathcal{L}(G(T_v), k; x, y)$ but is still preserved by some automorphism in $\text{Aut}(G_i; x \rightarrow y, y \rightarrow x)$ since $\pi \in \text{Aut}(G(T_v); x \rightarrow y, y \rightarrow x)$. That is, the equivalence class that contains the labeling of $G_{i,j}$ belongs to $\mathcal{B}(G_i, k; x, y)$. If π maps $G_{i,j}$ to $G_{i,j'}$, $j \neq j'$, then there is some automorphism in $\text{Aut}(G_i; x \rightarrow y, y \rightarrow x)$ that maps $(G_{i,j}, \phi)$ to $(G_{i,j'}, \phi)$ so that the equivalence classes that contain the labelings assigned by ϕ to $G_{i,j}$ and $G_{i,j'}$ are partners in $\mathcal{L}(G_i, k; x, y)$. Finally, if π maps $G_{i,j}$ to $G_{i,j'}$ and $G_{i,j'}$ to $G_{i,j''}$, where j, j' and j'' are distinct, then the equivalence classes that contain the labelings assigned by ϕ to $G_{i,j}$ and $G_{i,j'}$ are partners and those of $G_{i,j'}$ and $G_{i,j''}$ are partners as well. But this implies that the equivalence classes that contain the labelings assigned by ϕ to $G_{i,j}$ and $G_{i,j''}$ are exactly the same, contradicting the assumption that ϕ assigned inequivalent labelings to the m_i copies of G_i since $\phi \in \mathcal{L}(G(T_v), k; x, y)$. Thus, condition (iii) must be true.

By combining conditions (i) and (iii), we note that each group of size 2 in condition (iii) corresponds to an equivalence class of $\mathcal{L}_1(G_i, k; xy)$. Moreover, because no two of the equivalence classes that contain the labelings of the m_i copies of G_i are identical, distinct groups of size 2 correspond to distinct equivalence classes of $\mathcal{L}_1(G_i, k; xy)$. By using the claim, we can now compute $B(G(T_v), k; x, y)$. An equivalence class in $\mathcal{B}(G(T_v), k; x, y)$ is defined by (i) the label assigned to x and y , (ii) the equivalence class of $\mathcal{L}(G(T_{w_x}), k; x)$ that contains the labelings of $G(T_{w_x})$ and $G(T_{w_y})$, and (iii) the set of l_i equivalence classes of $\mathcal{L}_1(G_i, k; xy)$ and the set of $m_i - 2l_i$ equivalence classes in $\mathcal{B}(G_i, k; x, y)$ which together contain the labelings of the m_i copies of G_i for $i = 1, \dots, g$. There are k ways of assigning the same labels to x and y , say, l . Once l is fixed, there are $D(G(T_{w_x}), k; x)/k$ choices for the equivalence class of (ii), $\binom{D_1(G_i, k; xy)/k}{l_i}$ choices for the set of l_i equivalence classes, and $\binom{B(G_i, k; x, y)/k}{m_i - 2l_i}$ choices for the set of $m_i - 2l_i$ equivalence classes of (iii). Thus, $B(G(T_v), k; x, y)$ equals

$$k \binom{D(G(T_{w_x}), k; x)/k}{1} \prod_{i=1}^g \sum_{l=0}^{\lfloor m_i/2 \rfloor} \binom{D_1(G_i, k; xy)/k}{l} \binom{[D(G_i, k; x, y) - 2D(G_i, k; xy)]/k}{m_i - 2l}.$$

Once $B(G(T_v), k; x, y)$ is computed, we can determine $D(G(T_v), k; xy)$ by using Lemma 4.4. $\square$

An important implication of Theorems 4.3 and 4.6 and Lemma 4.5 is that both $D(G(T_v), k; x, y)$ and $D(G(T_v), k; xy)$ can be computed given the values of $D(G(T_{w_x}), k; x)$ and $D(G(T_{w_y}), k; y)$ and those of $D(G_i, k; x, y)$ and $D(G_i, k; xy)$ for $i = 1, \dots, g$.

Let us now consider the case when v is a t -vertex. Let H be the triconnected component associated with v . We need to solve for $D(G(T_v), k; H, A)$, where A is the structure associated with the parent of v in T_G (if the parent exists). Our goal is to create a formula for $D(G(T_v), k; H, A)$ that is dependent on H and the values of $D(G(T_w), k; *)$ only, where w is a child of v in T_G and $*$ is the structure associated with

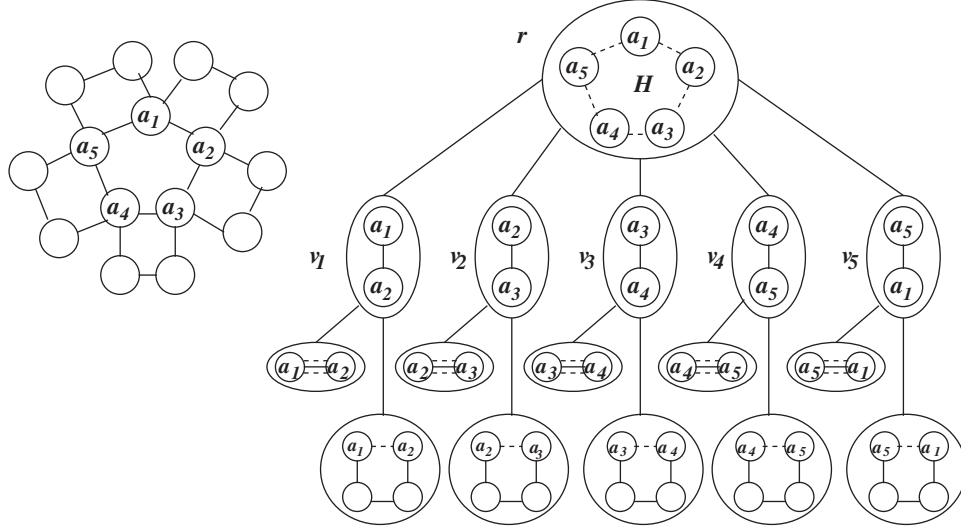


FIG. 6. Graph G and its tree decomposition. Again, $r = r(T_G)$. Note that $D(G(T_r), k; H) = D(G, k)$.

w , so that the formula can be computed efficiently. Our approach follows section 3 closely; the difference is that in our current setting $G(T_v)$ is made up of H together with components hanging off of the cut vertices and separating pairs of G in H , whereas in section 3 we dealt only with the graph H . We demonstrate our approach by solving for $D(G(T_v), k; H)$; others can be solved similarly. To aid us in our discussion, we shall use the graph in Figure 6 for illustration. In particular, we will compute for $D(G(T_r), k; H)$, which equals $D(G, k)$.

Let $\mathcal{C}_H$ contain the cut vertices of G in H whose corresponding vertices in T_G are children of v . Let $\mathcal{S}_H$ contain the separating pairs of G in H used in creating T_G whose corresponding vertices in T_G are children of v . When $a \in \mathcal{C}_H$ and w is the child of v that is associated with a , we shall refer to $G(T_w)$ as G_a for ease of notation. We do the same for each pair $\{x, y\} \in \mathcal{S}_H$. We begin by considering $\mathcal{L}^*(G(T_v), k)$, the set that contains all of the k -labelings ϕ of $G(T_v)$, so that ϕ when restricted to G_a belongs to $\mathcal{L}(G_a, k; a)$ for every cut vertex $a \in \mathcal{C}_H$ and ϕ when restricted to $G_{x,y}$ belongs to $\mathcal{L}(G_{x,y}, k; x, y)$ for every separating pair $\{x, y\} \in \mathcal{S}_H$. For instance, a labeling that assigns all of the nodes of G in Figure 6 the same label belongs to $\mathcal{L}^*(G(T_r), k)$ because every $G_{a_i, a_{i+1}}$ fixed at a_i and a_{i+1} has no nontrivial automorphisms. Clearly, every labeling in $\mathcal{L}(G(T_v), k; H)$ also belongs to $\mathcal{L}^*(G(T_v), k)$; otherwise, some nontrivial automorphism in $\text{Aut}(G(T_v); H)$ is not destroyed. We shall use PIE to weed out the labelings in $\mathcal{L}^*(G(T_v), k)$ that are not part of $\mathcal{L}(G(T_v), k; H)$.

Let $\text{Aut}_v(H)$ consist of all of the automorphisms in $\text{Aut}(G(T_v); H)$ when restricted to H . Suppose that $P \subseteq \text{Aut}_v(H)$. Let $N_{\geq}(P)$ denote the number of labelings ϕ in $\mathcal{L}^*(G(T_v), k)$ so that, for every $\sigma \in P$, some extension of σ in $\text{Aut}(G(T_v); H)$ preserves ϕ . Define $N_{=}(P)$ similarly except that aside from the automorphisms in P no other automorphism in $\text{Aut}_v(H)$ has extensions that preserve ϕ . We now state the formulas for $L(G(T_v), k; H)$ and $D(G(T_v), k; H)$.

THEOREM 4.7. *Let v be a t -vertex in T_G and H be the triconnected component associated with v . Let $\text{Aut}_v(H)$ consist of the automorphisms in $\text{Aut}(G(T_v); H)$ when*

restricted to H and σ_0 be the identity automorphism in $\text{Aut}_v(H)$. Then

$$L(G(T_v), k; H) = N_{=}(\{\sigma_0\}) = \sum_{\{\sigma_0\} \subseteq P \subseteq \text{Aut}_v(H)} (-1)^{|P|-1} N_{\geq}(P)$$

and

$$D(G(T_v), k; H) = \frac{1}{|\text{Aut}_v(H)|} \sum_{\{\sigma_0\} \subseteq P \subseteq \text{Aut}_v(H)} \frac{(-1)^{|P|-1} N_{\geq}(P)}{\prod_{a \in \mathcal{C}_H} |\text{Aut}(G_a; a)| \prod_{\{x, y\} \in \mathcal{S}_H} |\text{Aut}(G_{x, y}; x, y)|}.$$

Proof. It is straightforward to verify that $N_{=}(\{\sigma_0\}) = \sum_{\{\sigma_0\} \subseteq P \subseteq \text{Aut}_v(H)} (-1)^{|P|-1} N_{\geq}(P)$ is the number of labelings in $\mathcal{L}^*(G(T_v), k)$ that is preserved by some extension of σ_0 and by no other automorphism in $\text{Aut}_v(H)$. But by the way we defined $\mathcal{L}^*(G(T_v), k)$, if some extension of σ_0 preserves a labeling of $\mathcal{L}^*(G(T_v), k)$, that extension must be the identity automorphism of $G(T_v)$. The first equation of the theorem follows. Now $|\text{Aut}(G(T_v); H)| = |\text{Aut}_v(H)| \times \prod_{a \in \mathcal{C}_H} |\text{Aut}(G_a; a)| \times \prod_{\{x, y\} \in \mathcal{S}_H} |\text{Aut}(G_{x, y}; x, y)|$ since every automorphism in $\text{Aut}(G(T_v); H)$ can be decomposed into an automorphism in $\text{Aut}_v(H)$ and automorphisms of the connected components that hang off of H fixed at the cut vertices or separating pairs that connect them to H . Hence, dividing $L(G(T_v), k; H)$ by $|\text{Aut}(G(T_v); H)|$ gives us the second equation of the theorem. $\square$

COROLLARY 4.8. *When $\text{Aut}_v(H)$ is cyclic, dihedral, or isomorphic to $Z_t \times Z_2$ or $D_t \times Z_2$ for some integer t , the formulas in Theorems 3.5, 3.6, and 3.7 for $N_{=}(\{\sigma_0\})$, where σ_0 is the identity automorphism, still apply.*

Suppose that, instead of $D(G(T_v), k; H)$, we wish to compute $D(G(T_v), k; H, a)$ (or $D(G(T_v), k; H, x, y)$ or $D(G(T_v), k; H, xy)$). Let Γ denote the subgroup of $\text{Aut}_v(H)$ that fixes a (or x and y or xy). By replacing $\text{Aut}_v(H)$ with Γ , the formula in Theorem 4.7 still holds.

Next, we describe a method for computing $N_{\geq}(P)$, $P \subseteq \text{Aut}_v(H)$. For each $\sigma \in P$, let H_σ be the graph whose vertex set is $V(H)$ and edge set is $\{(v, \sigma(v)) : v \in V(H)\}$. Let SP_σ be the graph whose vertex set consists of xy and yx whenever $\{x, y\} \in \mathcal{S}_H$ and whose edge set is $\{(xy, \sigma(x)\sigma(y)), (yx, \sigma(y)\sigma(x)) : \{x, y\} \in \mathcal{S}_H\}$. In $\cup_{\sigma \in P} SP_\sigma$, let $SP_P(xy)$ denote the component that contains the vertex xy . Notice that if a and a' are part of the same component in $\cup_{\sigma \in P} H_\sigma$, then both a and a' are cut vertices or both are not; and when both of them are, then $G_a \cong G_{a'}$. Similarly, when xy and $x'y'$ are part of the same component in $\cup_{\sigma \in P} SP_\sigma$, $G_{x, y} \cong G_{x', y'}$. Finally, when $yx \in SP_P(xy)$, then, for every $x'y' \in SP_P(xy)$, $y'x' \in SP_P(xy)$ as well. The following can easily be verified: $\phi \in \mathcal{L}^*(G(T_v), k)$ is counted in $N_{\geq}(P)$ if and only if (i) whenever two vertices are part of the same component in $\cup_{\sigma \in P} H_\sigma$, ϕ assigns them the same label; (ii) whenever two cut vertices a and a' are part of the same component in $\cup_{\sigma \in P} H_\sigma$, ϕ when restricted to G_a and $G_{a'}$ belong to the same equivalence class of $\mathcal{L}(G_a, k; a)$; and (iii) when xy and $x'y'$ are part of the same component in $\cup_{\sigma \in P} SP_\sigma$, ϕ when restricted to $G_{x, y}$ and $G_{x', y'}$ belong to the same equivalence class of $\mathcal{L}(G_{x, y}, k; x, y)$.

In SP_σ , we have chosen to represent the separating pair $\{x, y\}$ as two vertices xy and yx to capture situations in which an automorphism in $\text{Aut}_v(H)$ maps x to y and y to x . However, such a representation can introduce redundancies in the sense that two different components in SP_σ may be capturing the same relationships between the same sets of separating pairs. As such, we shall say that a collection

of components $\{SP_P(x_i y_i), i = 1, \dots, g\}$ forms a *partition* of $\cup_{\sigma \in P} SP_\sigma$ if for every separating pair $\{x, y\} \in \mathcal{S}_H$ exactly one component in the collection contains xy or yx or both. In Figure 6, let $\pi_\rho \in \text{Aut}_r(H)$ be the rotation that maps a_i to a_{i+1} , $i = 1, \dots, 5$, and $\pi_\tau \in \text{Aut}_r(H)$ be the reflection that fixes a_1 . A partition for SP_{π_ρ} contains only one component, while a partition for SP_{π_τ} contains three components—e.g., $SP(a_1 a_5)$, $SP(a_5 a_4)$, and $SP(a_4 a_3)$, where $a_3 a_4 \in SP(a_4 a_3)$.

THEOREM 4.9. *Let $\{\sigma_0\} \subseteq P \subseteq \text{Aut}_v(H)$. Suppose that $\cup_{\sigma \in P} H_\sigma$ has t components, where the i th component contains the vertex a_i , and the collection $\{SP_P(x_i y_i), i = 1, \dots, g\}$ forms a partition of $\cup_{\sigma \in P} SP_\sigma$. Let $\kappa(a_i) = D(G_{a_i}, k; a_i)$ if a_i is a cut vertex and is equal to k otherwise. Let $\kappa(x_i, y_i) = [D(G_{x_i, y_i}, k; x_i, y_i) - 2D(G_{x_i, y_i}, k; x_i y_i)]/k$ if $y_i x_i \in SP_P(x_i y_i)$ and is equal to $D(G_{x_i, y_i}, k; x_i, y_i)/k^2$ otherwise. Then*

$$N_{\geq}(P) = \prod_{i=1}^t \kappa(a_i) \times \prod_{i=1}^g \kappa(x_i, y_i) \times \prod_{a \in \mathcal{C}_H} |\text{Aut}(G_a; a)| \times \prod_{\{x, y\} \in \mathcal{S}_H} |\text{Aut}(G_{x, y}; x, y)|.$$

Proof. To create a labeling ϕ that is counted in $N_{\geq}(P)$, we do the following. (1) If a_i is not a cut vertex, pick a label that will be assigned to it and all of the vertices in the same component as a_i in $\cup_{\sigma \in P} H_\sigma$. (2) If a_i is a cut vertex, pick an equivalence class of $\mathcal{L}(G_{a_i}, k; a_i)$ that will contain the labelings of G_{a_i} and all of the G_u 's, where u and a_i are in the same component of $\cup_{\sigma \in P} H_\sigma$. Then for a_i and each of the vertices u , pick a labeling from the equivalence class just chosen. (3) For each $\{x_i, y_i\}$, pick an equivalence class of $\mathcal{L}(G_{x_i, y_i}, k; x_i, y_i)$ that will contain the labelings of G_{x_i, y_i} and all $G_{u, w}$'s, where uw and $x_i y_i$ are in the same component of $\cup_{\sigma \in P} SP_\sigma$. Additionally, this equivalence class must respect the labels that have already been assigned to x_i and y_i in step 2. Then for $\{x_i, y_i\}$ and each $\{u, w\}$, pick a labeling from the equivalence class just chosen.

There are k ways of doing step 1 and $D(G_{a_i}, k; a_i) \times |\text{Aut}(G_{a_i}; a_i)|^{j_i}$ ways of doing step 2, where j_i is the number of vertices in the same component as a_i . Thus, there are $\prod_{i=1}^t \kappa(a_i) \times \prod_{a \in \mathcal{C}_H} |\text{Aut}(G_a; a)|$ ways of doing steps 1 and 2 since, whenever the a_i and u are in the same component of $\cup_{\sigma \in P} H_\sigma$, both vertices are cut vertices or both are not and $\text{Aut}(G_{a_i}; a_i) \cong \text{Aut}(G_u; u)$.

To do step 3, we need to differentiate between the case when $y_i x_i$ also belongs to $SP_P(x_i y_i)$ and when it doesn't. In the former case, ϕ when restricted to G_{x_i, y_i} must destroy all of the automorphisms of G_{x_i, y_i} when x_i and y_i are fixed but is preserved by some automorphism of the graph that maps x_i to y_i and vice versa. In other words, the equivalence class containing the labeling belongs to $\mathcal{B}(G_{x_i, y_i}, k; x_i, y_i)$. Since the labels of x and y have already been chosen in step 1 or 2 (note that they had to be the same), from Lemma 4.4 there are exactly $[D(G_{x_i, y_i}, k; x_i, y_i) - 2D(G_{x_i, y_i}, k; x_i y_i)]/k$ equivalence classes to choose from in step 3. On the other hand, when $y_i x_i$ does not belong to $SP_P(x_i y_i)$, ϕ when restricted to G_{x_i, y_i} must simply belong to $\mathcal{L}(G_{x_i, y_i}, k; x_i, y_i)$, and so, once the labels of x_i and y_i have been chosen in steps 1 and 2, there are exactly $D(G_{x_i, y_i}, k; x_i, y_i)/k^2$ equivalence classes to choose from in step 3. Finally, when the equivalence classes have been selected, then there are $|\text{Aut}(G_{x_i, y_i}; x_i, y_i)|^{j_i}$ labelings that can be assigned to the $G_{u, w}$'s, $uw \in SP_P(x_i y_i)$, where j_i is the number of distinct separating pairs in $SP(x_i y_i)$. There are $\prod_{i=1}^g \kappa(x_i, y_i) \prod_{\{x, y\} \in \mathcal{S}_H} |\text{Aut}(G_{x, y}; x, y)|$ ways of doing step 3 because, again, whenever $x_i y_i$ and uw are in the same component in $\cup_{\sigma \in P} SP_\sigma$, $\text{Aut}(G_{x_i, y_i}; x_i, y_i) \cong \text{Aut}(G_{u, w}; u, w)$. The theorem follows. $\square$

By using the formula for $N_{\geq}(P)$ above, we can now simplify the second formula in Theorem 4.7 as

$$D(G(T_v), k; H) = \frac{1}{|Aut_v(H)|} \sum_{\{\sigma_0\} \subseteq P \subseteq Aut_v(H)} (-1)^{|P|-1} \prod_{i=1}^t \kappa(a_i) \prod_{i=1}^g \kappa(x_i, y_i).$$

Hence, $D(G(T_v), k; H)$ (and $D(G(T_v), k; H, a)$, $D(G(T_v), k; H, x, y)$, and $D(G(T_v), k; H, xy)$) can be computed once the automorphisms in $Aut_v(H)$ and the values of $D(G(T_w), k; *)$ are known, where w is a child of v in T_G and $*$ is the structure associated with w .

FindDist(G, k). Let us now describe our algorithm *FindDist*(G, k) for computing $D(G, k)$. First, construct T_G and root it at $r(T_G)$. Then, for $i = \text{height}(T_G)$ to 0, do the following for each vertex v at depth i . When v is a c -vertex, compute $D(G(T_v), k; a)$, where a is the cut vertex associated with v using Theorem 4.2. When v is an s -vertex, compute $D(G(T_v), k; x, y)$ and $D(G(T_v), k; xy)$, where $\{x, y\}$ is the separating pair associated with v using Theorems 4.3 and 4.6. When v is a t -vertex and H is the triconnected component associated with v , compute $D(G(T_v), k; H, *)$, where $*$ is the structure associated with $p(v)$, the parent of v , if it exists. (That is, if $p(v)$ is a c -vertex, compute $D(G(T_v), k; H, a)$, where a is the cut vertex associated with $p(v)$; if $p(v)$ is an s -vertex, compute $D(G(T_v), k; H, x, y)$ and $D(G(T_v), k; H, xy)$, where $\{x, y\}$ is the separating pair associated with $p(v)$; if v has no parent, compute $D(G(T_v), k; H)$.) Do all computations according to Theorem 4.7, Corollary 4.8, and Theorem 4.9. Finally, if the root node $r(T_G)$ is a c -vertex or a t -vertex, return the value computed at depth 0; otherwise, if $r(T_G)$ is an s -vertex, return $D(G(T_v), k; xy)$.

It is not obvious that at every iteration the algorithm can compute $D(G(T_v), k; *)$ based on the values obtained for v 's children in the previous iterations. We shall now show that this, in fact, is the case.

- *When v is a c -vertex.* Let a be the cut vertex associated with v . From Theorem 4.2, $D(G(T_w), k; a)$ is needed. If w is an s -vertex and associated with some separating pair $\{a, b\}$, $D(G(T_w), k; a, b)$ was computed in the previous iterations. But from Lemma 2.7, $Aut(G(T_w); a, b) = Aut(G(T_w); a)$; i.e., a labeling of $G(T_w)$ destroys all automorphisms in $Aut(G(T_w); a, b)$ if and only if it destroys all automorphisms in $Aut(G(T_w); a)$. Hence, $D(G(T_w), k; a, b) = D(G(T_w), k; a)$. If w is a t -vertex, w is associated with some triconnected component H , and so $D(G(T_w), k; H, a)$ was computed in the previous iteration. From Lemma 2.7, $Aut(G(T_w); H, a) = Aut(G(T_w); a)$ and, consequently, $D(G(T_w), k; H, a) = D(G(T_w), k; a)$.
- *When v is an s -vertex.* Let $\{x, y\}$ be the separating pair associated with v . This time, according to Theorems 4.3 and 4.6, if w is a c -vertex associated with x (or y), $D(G(T_w), k; x)$ (or $D(G(T_w), k; y)$) is needed and was clearly computed in the previous iterations by the algorithm. On the other hand, if w is a t -vertex whose associated triconnected component is H , $D(G(T_w), k; x, y)$ and $D(G(T_w), k; xy)$ are needed. Now $D(G(T_w), k; H, x, y)$ and $D(G(T_w), k; H, xy)$ were computed in the previous iterations. But from Lemma 2.8, $Aut(G(T_w); H, x, y) = Aut(G(T_w); x, y)$ and $Aut(G(T_w); H, xy) = Aut(G(T_w); xy)$. It follows that $D(G(T_w), k; H, x, y) = D(G(T_w), k; x, y)$ and $D(G(T_w), k; H, xy) = D(G(T_w), k; xy)$.
- *When v is a t -vertex.* Let H be the triconnected component associated with v . According to Theorem 4.9, if w is a c -vertex associated with a , $D(G(T_w), k; a)$ is needed, and if w is an s -vertex associated with the separating pair $\{x, y\}$,

$D(G(T_w), k; x, y)$ and $D(G(T_w), k; xy)$ are needed. All of these values were computed in the previous iterations.

From Theorems 4.2, 4.3, 4.6, 4.7, and 4.9, we know that all of the $D(G(T_v), k; *)$ values computed by the algorithm are correct. Now the algorithm returned the value $D(G(T_v), k; A)$, where $v = r(T_G)$ and A is the structure associated with $r(T_G)$. But $G(T_{r(T_G)}) = G$, and, according to Lemma 2.6, $\text{Aut}(G; A) = \text{Aut}(G)$. It follows that the algorithm returned $D(G, k)$.

THEOREM 4.10. *When G is a connected graph, $\text{FindDist}(G, k)$ returns the value $D(G, k)$.*

Example 1. Consider the graph and its tree decomposition in Figures 2 and 3. By using the formulas given in this section, it is easy to verify the following: $D(G(T_{v_3}), k; e) = k^5$, $D(G(T_{v_1}), k; e, j) = k^3(k-1)/2$, and $D(G(T_{v_1}), k; ej) = k^2(k-1)^2/4$. According to Theorem 4.3,

$$\begin{aligned} D(G(T_r), k; e, j) &= k^2 [D(G(T_{v_3}), k; e)/k]^2 \binom{D(G(T_{v_1}), k; e, j)/k^2}{2} \\ &= k^2 (k^4)^2 \binom{k(k-1)/2}{2} \\ &= (k+1)k^{11}(k-1)(k-2)/8. \end{aligned}$$

Next, let us compute $B(G(T_r), k; e, j)$. We have

$$B(G(T_{v_1}), k; e, j) = D(G(T_{v_1}), k; e, j) - 2D(G(T_{v_1}), k; ej) = k^2(k-1)/2.$$

Notice that $D_1(G(T_{v_1}), k; ej) = 0$ because any labeling that assigns e and j the same label cannot destroy the automorphism that maps e to j , j to e , f to itself, and g to itself. According to Theorem 4.6,

$$\begin{aligned} B(G(T_r), k; e, j) &= k D(G(T_{v_3}), k; e)/k \sum_{l=0}^1 \binom{D_1(G(T_{v_1}), k; e, j)/k}{l} \binom{B(G(T_{v_1}), k; e, j)/k}{2-2l} \\ &= k k^4 \sum_{l=0}^1 \binom{0}{l} \binom{k(k-1)/2}{2-2l} \\ &= k^5 \left[\binom{0}{0} \binom{k(k-1)/2}{2} + \binom{0}{1} \binom{k(k-1)/2}{0} \right] \\ &= (k+1)k^6(k-1)(k-2)/8. \end{aligned}$$

Consequently, $D(G, k) = D(G(T_r), k; ej) = [D(G(T_r), k; e, j) - B(G(T_r), k; e, j)]/2 = (k+1)k^6(k-1)(k-2)(k^5-1)/16$. Since $D(G, 1) = D(G, 2) = 0$ and $D(G, 3) > 0$, $D(G) = 3$.

Example 2. This time, consider the graph in Figure 6, and let us determine $D(G(T_r), k; H)$. Since $\text{Aut}_r(H) \cong D_5$, we can make use of the computations we made in the example after Theorem 3.6. It is easy to verify that $D(G_{a_i, a_{i+1}}, k; a_i, a_{i+1}) = k^4$ and $D(G_{a_i, a_{i+1}}, k; a_i a_{i+1}) = (k^4 - k^2)/2$ for $i = 1, \dots, 5$. Let $\alpha = \prod_{i=1}^5 |\text{Aut}(G_{a_i, a_{i+1}}; a_i, a_{i+1})|$. From Theorem 4.9, we have $N_{\geq}(\emptyset)/\alpha = k^5 \times (k^2)^5 = k^{15}$, $N_{\geq}(\{\pi_\rho\})/\alpha = k \times k^2 = k^3$, $N_{\geq}(\{\pi_{\tau\rho^i}\})/\alpha = k^3 \times (k^2)^2 (k^4 - 2(k^4 - k^2)/2)/k = k^8$, and $N_{\geq}(\{\pi_{\tau\rho^i}, \pi_\rho\})/$

$\alpha = k \times (k^4 - 2(k^4 - k^2)/2)/k = k^2$. Thus,

$$\begin{aligned} D(G(T_r), k; H) &= \frac{1}{10} \left[N_{\geq}(\emptyset)/\alpha - N_{\geq}(\{\pi_\rho\})/\alpha - \sum_{i=0}^4 (N_{\geq}(\{\pi_{\tau\rho^i}\})/\alpha - N_{\geq}(\{\pi_{\tau\rho^i}, \pi_\rho\})/\alpha) \right] \\ &= \frac{1}{10} \left[k^{15} - k^3 - \sum_{i=0}^4 (k^8 - k^2) \right] \\ &= \frac{1}{10} [k^{15} - 5k^8 - k^3 + 5k^2]. \end{aligned}$$

Since $D(G, 1) = 0$ and $D(G, 2) > 0$, we conclude that $D(G) = 2$.

In order for $\text{FindDist}((G, k))$ to run efficiently for a family of graphs, we note that a few ingredients are necessary. It must have an efficient graph isomorphism testing algorithm to determine the isomorphism classes of $\mathcal{G}$ in Theorems 4.2, 4.3, and 4.6. There must also be an efficient algorithm that can determine the automorphisms of its triconnected components which are needed in Theorem 4.7. Finally, there must be a way to apply the PIE formula in Theorem 4.7 to its triconnected components in an efficient manner. Since the family of planar graphs satisfies these criteria, we can now proceed to prove the main result of the paper. We first show that $D(G(T_v), k; *)$, when v is a t -vertex, can be computed efficiently when the appropriate values are known.

LEMMA 4.11. *Let G be a connected planar graph on n vertices and k be a positive integer. Let v be a t -vertex in T_G , H be the triconnected component on n_H vertices associated with v , and $\mathcal{S}_H$ contain the separating pairs of G in H used in the construction of T_G . Suppose that all of the automorphisms in $\text{Aut}_v(H)$ and the values of $D(G(T_w), k; *)$, w a child of v in T_G , are known. Then $D(G(T_v), k; *)$ can be computed in $O(n_H^2 n^2 \log^2 k (n_H + |\mathcal{S}_H|))$ time.*

Proof. As in Theorems 4.7 and 4.9, we will prove the theorem for $D(G(T_v), k; H)$; others can be argued similarly. Let $\alpha = \prod_{a \in \mathcal{C}_H} |\text{Aut}(G_a; a)| \times \prod_{\{x, y\} \in \mathcal{S}_H} |\text{Aut}(G_{x, y}; x, y)|$. First, let us consider the time it takes to compute $N_{\geq}(P)/\alpha$. Constructing $\cup_{\sigma \in P} H_\sigma$ and finding its t connected components can be done in $O(|P| \times n_H)$. Similarly, constructing $\cup_{\sigma \in P} SP_\sigma$ and finding a collection $\{SP_P(x_i y_i), i = 1, \dots, g\}$ that forms a partition of $\cup_{\sigma \in P} SP_\sigma$ takes $O(|P| \times |\mathcal{S}_H|)$ time. Finally, since $N_{\geq}(P) \leq k^n$, every multiplication in $\prod_{i=1}^t \kappa(a_i) \times \prod_{i=1}^g \kappa(x_i y_i)$ takes at most $O(n^2 \log^2 k)$ time; that is, finding the said product takes $O((t + g)n^2 \log^2 k) = O((n_H + |\mathcal{S}_H|)n^2 \log^2 k)$ time. Therefore, computing $N_{\geq}(P)/\alpha$ takes $O((n_H + |\mathcal{S}_H|)(|P| + n^2 \log^2 k))$ time.

Since G is a connected planar graph, each of its blocks B are as well. It is easy to verify that, when the split operation is applied to B , the resulting split graphs remain planar. Hence, the triconnected component associated with v is either a bond, a cycle, or a triconnected planar graph, and $\text{Aut}_v(H)$ belongs to one of the groups mentioned in Lemma 3.8. From Corollary 4.8, we noted that we can apply the formulas in Theorems 3.5, 3.6, and 3.7 for $N_{\geq}(\{\sigma_0\})$. Consequently, we can use $\text{TriconnectCount}(H, k)$ to compute $D(G(T_v), k; H)$ by doing two modifications—we replace every occurrence of $N_{\geq}(P)$ with $N_{\geq}(P)/\alpha$ (keeping in mind that the k -labelings counted in $N_{\geq}(P)$ belong to $\mathcal{L}^*(G(T_v), k)$ and are computed by using Theorem 4.9) and return the value $L(H, k)/|\text{Aut}_v(H)|$ instead of $L(H, k)$. It is easy to check that the value returned is $D(G(T_v), k; H)$. By applying the same analysis in Theorem 3.9, the bottleneck of the algorithm is in computing the $O(n_H^2) N_{\geq}(P)/\alpha$ terms. By using the result from the previous paragraph and the fact that $|P| \leq \log n_H$, the run time of the algorithm is $O(n_H^2 (n_H + |\mathcal{S}_H|)(\log n_H + n^2 \log^2 k))$ time. $\square$

Finding $D(G(T_v), k; *)$ can take significantly more time than finding $D(H, k)$, because $N_{\geq}(P)$ in the former case has to be explicitly computed from the $D(G(T_w), k; *)$ values whereas, in the latter case, $N_{\geq}(P)$ is always some power of k which we precomputed ahead of time.

THEOREM 4.12. *Let G be an n -vertex connected planar graph and k be a positive integer. $\text{FindDist}(G, k)$ can be implemented in $O(n^5 \log^2 k)$ time. Consequently, computing $D(G)$ takes $O(n^5 \log^3 n)$ time.*

Proof. To implement $\text{FindDist}(G, k)$, we need to do some preprocessing steps.

1. *Isomorphism testing.* For each c - and s -vertex, determine the isomorphism classes of the graphs in $\mathcal{G}$ fixed at the appropriate vertices as described in Theorems 4.2, 4.3, and 4.6. Additionally, for each s -vertex whose associated separating pair is $\{x, y\}$, determine if $G(T_{w_x})$ fixed at x is isomorphic to $G(T_{w_y})$ fixed at y . For each t -vertex, group together its children based on their type—i.e., if they are c -vertices or s -vertices—and then for each group determine the isomorphism classes of $\{G(T_w) : w \text{ is a child of } v \text{ and a } c\text{-vertex (or an } s\text{-vertex)}\}$ fixed at the appropriate vertices. By using the linear-time planar graph isomorphism testing algorithm, the isomorphism testing tasks at each v can be done in $O(\deg(v)^2 n)$ time, where $\deg(v)$ is the degree of vertex v in T_G . The total time to do all of the isomorphism testing then takes $O(n^3)$ time since the size of T_G is $O(n)$.

2. *Finding automorphisms.* For each t -vertex v whose associated triconnected component is H , find all of the automorphisms in $\text{Aut}(H)$ by using the algorithm described in section 3.1. Apply the information obtained from the first preprocessing step to determine which of the automorphisms also belong to $\text{Aut}_v(H)$ (or $\text{Aut}_v(H; *)$, where $*$ depends on $p(v)$). That is, for each $\pi \in \text{Aut}(H)$, verify that π maps cut vertices and separating pairs to cut vertices and separating pairs, respectively; moreover, π also maps the subgraphs hanging from these vertices and pairs to isomorphic structures. Finding $\text{Aut}(H)$ takes $O(n_H^2)$ time, where n_H is the number of vertices in H . Determining if $\pi \in \text{Aut}(H)$ also belongs to $\text{Aut}_v(H)$ takes $O(n_H + |\mathcal{S}_H|)$ time, so finding $\text{Aut}_v(H)$ takes $O(n_H^2 + |\text{Aut}(H)|(n_H + |\mathcal{S}_H|)) = O(n_H^2 + n_H |\mathcal{S}_H|)$ time. Finding $\text{Aut}_v(H)$ for all t -vertices v then takes $O(n^2)$ time, since from Lemma 2.10 $\sum_H n_H = O(n)$ and $\sum_H |\mathcal{S}_H| = O(n)$.

Let us now consider $\text{FindDist}(G, k)$. Constructing T_G takes $O(n^2)$ time. The number of arithmetic operations for computing $D(G(T_v), k; *)$ when v is a c -vertex and an s -vertex is $O(\deg(v))$ and $O(\deg^2(v))$, respectively. Since the values involved in each operation is at most k^n , the total amount of work at v takes $O(\deg^2(v)n^2 \log^2 k)$ time. By summing this up for all c - and s -vertices in T_G , the total work takes $O(n^4 \log^2 k)$ time because the size of T_G is $O(n)$. When v is a t -vertex, according to Lemma 4.11, computing $D(G(T_v), k; *)$ takes $O(n_H^2 n^2 \log^2 k (n_H + |\mathcal{S}_H|))$ time. Thus, for all t -vertices, the total work is $O(n^5 \log^2 k)$ time. From our analysis, the bottleneck of $\text{FindDist}(G, k)$ is in processing the t -vertices. The run time in the theorem follows. $\square$

COROLLARY 4.13. *Let G be an n -vertex planar graph. Then $D(G)$ can be computed in $O(n^5 \log^3 n)$ time.*

Proof. By using the linear-time planar graph isomorphism testing algorithm, group together isomorphic connected components of G . So suppose that $G = m_1 G_1 \cup m_2 G_2 \cup \dots \cup m_g G_g$, where each G_i has n_i vertices. For each G_i , find k_i so that $k_i = \min\{k : D(G_i, k) \geq m_i\}$. This can be done by implementing $\text{FindDist}(G_i, k)$ $O(\log n)$ times. Finally, $D(G) = \max_i \{k_i\}$, so $D(G)$ can be found in $O(\sum_{i=1}^g n_i^5 \log^2 n_i \log n) = O(n^5 \log^3 n)$ time. $\square$

5. Final comments. In this paper, we considered $D(G, k)$, the number of inequivalent distinguishing k -labelings of graph G . We have applied the principle of inclusion-exclusion and developed recursive formulas to compute its value for a fixed k . When a graph G is planar, we showed that these techniques led to an algorithm for computing $D(G)$ that runs in time polynomial in the size of G . There are other interesting aspects about $D(G, k)$ as well, as noted in the next theorem.

THEOREM 5.1. *Let G be a graph on n vertices. Then $D(G, k)$ is a polynomial in k whose degree is n and whose constant term is 0. If G has no nontrivial automorphisms, then $D(G, k) = k^n$; otherwise, the sum of the coefficients of $D(G, k)$ is 0.*

Proof. In Theorem 3.2, $N_{\geq}(P)$ equals k^n when $P = \{\pi_0\}$ and k^t , $t < n$, otherwise, so $D(G, k)$ must be a polynomial in k whose degree is n . Furthermore, $D(G, 0) = 0$, so the constant term in $D(G, k)$ must be 0. When G has no nontrivial automorphism, every k -labeling of G is distinguishing and no two are equivalent; hence, $D(G, k) = k^n$. When G has some nontrivial automorphisms, $D(G, 1) = 0$, so it must be the case that the sum of the coefficients of $D(G, k)$ is 0. $\square$

We now call $D(G, k)$ the *distinguishing polynomial* of G . An interesting research direction would be to study this polynomial along the lines of the more famous chromatic polynomial of G .

Next, consider Lemma 3.4 and its implications. According to the lemma, if S is the largest subset of $\text{Aut}(G)$ that preserves a k -labeling ϕ of G , then S must be a subgroup of $\text{Aut}(G)$. This suggests that, instead of considering all of the subsets of $\text{Aut}(G)$ as we do in PIE, we should just consider the subgroup lattice $(\mathcal{S}, \leq)$ of $\text{Aut}(G)$, where $\mathcal{S}$ is the set that contains all of the subgroups of $\text{Aut}(G)$. For each $S \in \mathcal{S}$, define $N_{\geq}(S)$ and $N_{=}(S)$ as we did in the PIE formulation. Since $N_{\geq}(S) = \sum_{S' \geq S} N_{=}(S')$ for every $S \in \mathcal{S}$, according to the principle of Möbius inversion, the following must be true.

THEOREM 5.2. *Let $(\mathcal{S}, \leq)$ be the subgroup lattice of $\text{Aut}(G)$ and $\mu(*, *)$ be the Möbius function of $(\mathcal{S}, \leq)$. Let S_0 be the subgroup of $\text{Aut}(G)$ consisting of the identity automorphism of G . Then*

$$L(G, k) = N_{=}(S_0) = \sum_{S \in \mathcal{S}: S \geq S_0} \mu(S_0, S) N_{\geq}(S).$$

By using the formula above, the number of $N_{\geq}(S)$ terms we have to compute to determine $L(G, k)$ is $|\mathcal{S}|$ as opposed to $\Omega(2^{|\text{Aut}(G)|})$ in the PIE formulation. This, of course, comes at a price—we must now find all of the subgroups of $\text{Aut}(G)$, determine the structure of $(\mathcal{S}, \leq)$, and then compute the $\mu(S_0, S)$ values. We leave it up to the reader to apply Theorem 5.2 when $\text{Aut}(G) \cong Z_t, D_t, Z_t \times Z_2$, or $D_t \times Z_2$.

Another direction in which we use Lemma 3.4 is a generalization of Theorem 3.5.

THEOREM 5.3. *Let S^* be a subset of $\text{Aut}(G)$ that does not contain the identity automorphism. Suppose that every nontrivial subgroup of G contains at least one element of S^* . Then*

$$L(G, k) = \sum_{S \subseteq S^*} (-1)^{|S|} N_{\geq}(S).$$

In the above theorem, the number of $N_{\geq}(S)$ terms to compute is $2^{|S^*|}$ so the smaller $|S^*|$ is, the better. Finding the smallest S^* , however, is nontrivial; it is the *hitting set problem* and is known to be NP-hard in general [13]. Again, we leave it up

to the reader to formulate a generalization of Theorem 3.6 similar to Theorem 5.3. Finally, we pose the following open problem.

Open problem: Let $\mathcal{G}$ be a family of graphs for which there is an efficient algorithm for testing graph isomorphism. Can the distinguishing number of graphs in $\mathcal{G}$ be computed efficiently? More specifically, is there a polynomial-time algorithm for computing $D(G)$ if G is a bounded degree graph or a bounded genus graph?

Acknowledgments. The second author thanks Jeb Willenbring for discussions she had with him on section 3 of the paper and Tom Tucker for providing her with copies of his papers [12, 20]. The authors are also grateful to the anonymous referees for their comments, which improved the presentation of the paper.

REFERENCES

- [1] A. AHO, J. HOPCROFT, AND J. ULLMAN, *The Design and Analysis of Algorithms*, Addison-Wesley, Reading, MA, 1974.
- [2] M. ALBERTSON AND K. COLLINS, *Symmetry breaking in graphs*, Electron. J. Combin., 3 (1996), p. R18.
- [3] M. ALBERTSON, *Distinguishing Cartesian powers of graphs*, Electron. J. Combin., 12 (2005), p. N17.
- [4] V. ARVIND AND N. DEVANUR, *Symmetry breaking in trees and planar graphs by vertex coloring*, in Proceedings of the 8th Nordic Combinatorial Conference, Aalborg University, Aalborg, Denmark, 2004.
- [5] W. BOGSTAD AND L. COWEN, *The distinguishing number of the hypercube*, Discrete Math., 283 (2004), pp. 29–35.
- [6] M. CHAN, *The distinguishing number of the direct product and wreath product action*, J. Algebraic Combin., 24 (2006), pp. 331–345.
- [7] M. CHAN, *The maximum distinguishing number of a group*, Electron. J. Combin., 13 (2006), p. R70.
- [8] M. CHAN, *The distinguishing number of the augmented cube and hypercube powers*, Discrete Math., 308 (2007), pp. 2330–2336.
- [9] C. CHENG, *On computing the distinguishing numbers of trees and forests*, Electron. J. Combin., 13 (2006), p. R11.
- [10] K. COLLINS AND A. TRENK, *The distinguishing chromatic number*, Electron. J. Combin., 13 (2006), p. R16.
- [11] R. DIESTEL, *Graph Theory*, 3rd ed., Springer-Verlag, New York, 2006.
- [12] T. FUKUDA, S. NEGAMI, AND T. TUCKER, *3-Connected Planar Graphs Are 2-Distinguishable with Few Exceptions*, preprint.
- [13] M. GAREY AND D. JOHNSON, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, San Francisco, 1979.
- [14] J. HOPCROFT AND R. TARJAN, *Dividing a graph into triconnected components*, SIAM J. Comput., 2 (1973), pp. 135–158.
- [15] J. HOPCROFT AND J. WONG, *Linear time algorithm for isomorphism of planar graphs*, in Proceedings of 6th ACM Symposium on Theory of Computing, 1974, pp. 172–184.
- [16] W. IMRICH AND S. KLAVŽAR, *Distinguishing Cartesian powers of graphs*, J. Graph Theory, 53 (2006), pp. 250–260.
- [17] S. KLAVŽAR, T.-L. WONG, AND X. ZHU, *Distinguishing labelings of group action on vector spaces and graphs*, J. Algebra, 303 (2006), pp. 626–641.
- [18] P. MANI, *Automorphismen von polyedrischen graphen*, Math. Ann., 1971, pp. 279–303.
- [19] A. RUSSELL AND R. SUNDARAM, *A note on the asymptotics and computational complexity of graph distinguishability*, Electron. J. Combin., 5 (1998), p. R23.
- [20] T. TUCKER, *Distinguishability of Maps*, preprint.
- [21] J. TYMOCZKO, *Distinguishing numbers for graphs and groups*, Electron. J. Combin., 11 (2004), p. R63.

LARGE NEARLY REGULAR INDUCED SUBGRAPHS*

NOGA ALON[†], MICHAEL KRIVELEVICH[‡], AND BENNY SUDAKOV[§]

Abstract. For a real $c \geq 1$ and an integer n , let $f(n, c)$ denote the maximum integer f such that every graph on n vertices contains an induced subgraph on at least f vertices in which the maximum degree is at most c times the minimum degree. Thus, in particular, every graph on n vertices contains a regular induced subgraph on at least $f(n, 1)$ vertices. The problem of estimating $f(n, 1)$ was posed long ago by Erdős, Fajtlowicz, and Staton. In this paper we obtain the following upper and lower bounds for the asymptotic behavior of $f(n, c)$: (i) For fixed $c > 2.1$, $n^{1-O(1/c)} \leq f(n, c) \leq O(cn/\log n)$. (ii) For fixed $c = 1 + \varepsilon$ with $\varepsilon > 0$ sufficiently small, $f(n, c) \geq n^{\Omega(\varepsilon^2/\ln(1/\varepsilon))}$. (iii) $\Omega(\ln n) \leq f(n, 1) \leq O(n^{1/2} \ln^{3/4} n)$. An analogous problem for not necessarily induced subgraphs is briefly considered as well.

Key words. induced subgraphs, regular subgraphs, Ramsey-type problem

AMS subject classification. 05C35

DOI. 10.1137/070704927

1. Introduction. All graphs considered here are finite and simple. For a graph $G = (V, E)$, let $\Delta(G)$, $\delta(G)$, and $d(G) = \frac{2|E|}{|V|}$ denote its maximum degree, minimum degree, and average degree, respectively. The *density* of G is $p = |E|/\binom{|V|}{2}$; clearly, this is a number between 0 and 1. For $U \subseteq V$, let $G[U]$ denote the subgraph of G induced on U .

DEFINITION 1. A graph G is c -nearly regular if $\Delta(G) \leq c \cdot \delta(G)$.

For a graph $G = (V, E)$ and a constant $c \geq 1$, let

$$f(G, c) = \max\{|U| : G[U] \text{ is a } c\text{-nearly regular graph}\}.$$

Define

$$f(n, c) = \min\{f(G, c) : |V(G)| = n\}.$$

Thus, every graph G on n vertices contains a c -nearly regular induced subgraph on at least $f(n, c)$ vertices. In particular, for $c = 1$ every such G contains a strictly regular induced subgraph on at least $f(n, 1)$ vertices.

The problem of estimating $f(n, 1)$ was posed by Erdős, Fajtlowicz, and Staton (cf. [3] or [2, page 85]). By the known estimates for graph Ramsey numbers (cf.,

*Received by the editors October 9, 2007; accepted for publication (in revised form) April 28, 2008; published electronically September 4, 2008.

<http://www.siam.org/journals/sidma/22-4/70492.html>

[†]Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel, and Institute for Advanced Study, Princeton, NJ 08540 (nogaa@tau.ac.il). The research of this author was supported in part by the Israel Science Foundation, by a USA-Israeli BSF grant, by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University, and by the Von Neumann Fund.

[‡]Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel (krivelev@post.tau.ac.il). The research of this author was supported in part by USA-Israel BSF grant 2006322 and by grant 526/05 from the Israel Science Foundation.

[§]Department of Mathematics, UCLA, Los Angeles, CA 90095, and Institute for Advanced Study, Princeton, NJ 08540 (bsudakov@math.ucla.edu). The research of this author was supported in part by NSF CAREER award DMS-0546523, NSF grants DMS-0355497 and DMS-0635607, by a USA-Israeli BSF grant, and by the State of New Jersey.

e.g., [6]), every graph on n vertices contains either a clique or an independent set of order $\Omega(\ln n)$. This implies that $f(n, 1) \geq \Omega(\ln n)$. Erdős, Fajtlowicz, and Staton conjectured that the ratio $f(n, 1)/\ln n$ tends to infinity as n tends to infinity. We are unable to prove or disprove this conjecture and can only obtain several bounds, which are listed in the following results. The first deals with the case of large c .

PROPOSITION 1.1. *There exists an absolute constant b so that for all $K \geq 2.1$, $f(n, K) \geq n^{1-b/K}$.*

The problem of obtaining a nontrivial lower bound for values of c close to 1 is more interesting. Here we first deal with the case of graphs with positive density and show that any such graph must contain a nearly regular subgraph on a linear number of vertices.

THEOREM 1.2. *Let $\varepsilon > 0$ be a small real, and let p satisfy $0 < p < 1$. Then, for every sufficiently large n , any graph $G = (V, E)$ on n vertices with density at least p contains an induced $(1 + \varepsilon)$ -nearly regular subgraph on at least*

$$0.5 \left(\frac{\varepsilon}{6} \right)^{\frac{144}{\varepsilon^2} \ln(\frac{1}{p})} \cdot n$$

vertices.

For general (possibly sparse) graphs we have the following.

THEOREM 1.3. *Let $\varepsilon > 0$ be a sufficiently small constant. Then*

$$f(n, 1 + \varepsilon) \geq n^{\frac{\varepsilon^2}{250 \ln(1/\varepsilon)}}$$

for all sufficiently large n .

Our upper bounds for $f(n, c)$ are rather far from the lower bounds. For the strictly regular case we prove the following.

THEOREM 1.4. $f(n, 1) \leq O(n^{1/2} \log^{3/4} n)$.

This is a slight improvement of an earlier estimate of Bollobás (cf. [2]), who showed that for every $\epsilon > 0$, $f(n, 1) \leq c(\epsilon)n^{1/2+\epsilon}$. For the nearly regular case we have the following.

PROPOSITION 1.5. *For every constant $K \geq 2$, $f(n, K) \leq 7K \frac{n}{\log n}$.*

The lower bounds are proved in the next section, and the upper bounds are presented in section 3. We conclude in section 4 with a few open problems and a brief discussion of an analogous problem for not necessarily induced subgraphs. Throughout this paper we assume, whenever needed, that the number of vertices n of the graphs discussed is sufficiently large. To simplify the presentation, we make no attempt to optimize the absolute constants, and we omit all floor and ceiling signs whenever these are not crucial. We also use the following standard asymptotic notation: For two functions $f(n)$, $g(n)$ of a natural valued parameter n , we write $f(n) = o(g(n))$ whenever $\lim_{n \rightarrow \infty} f(n)/g(n) = 0$; $f(n) = O(g(n))$ if there exists a constant $C > 0$ such that $f(n) \leq Cg(n)$ for all n ; and $f(n) = \Omega(g(n))$ if $g(n) = O(f(n))$.

2. Lower bounds. For a graph $G = (V, E)$ and a subset $U \subseteq V$, the number of edges of G spanned by U in G is denoted by $e_G(U)$; the number of edges between disjoint subsets U, W of vertices of G is denoted by $e_G(U, W)$.

2.1. Large c . In this subsection we prove Proposition 1.1, which provides a lower bound for $f(n, c)$ when c is a relatively large constant.

We need the following rather standard argument, allowing one to pass from a graph with a large average degree to one with a large minimum degree.

PROPOSITION 2.1. *Let $K > 1, \alpha < 1/2$ be constants. Then every graph $G = (V, E)$ on $|V| = n$ vertices with $\Delta(G) \leq Kd(G)$ contains an induced (K/α) -nearly regular subgraph G^* with at least $\frac{1-2\alpha}{K-2\alpha}n$ vertices and at least $\frac{K-2K\alpha}{2K-4\alpha}nd$ edges.*

Proof. Denote the average degree $d(G)$ of G by d . We can obviously assume that $d > 0$. Start with G and delete repeatedly vertices of degree less than αd until there are none left. Denote the resulting graph by G^* . Then G^* is an induced subgraph of G , satisfying $\Delta(G^*) \leq \Delta(G)$, $\delta(G^*) \geq \alpha d$, implying

$$\Delta(G^*) \leq \Delta(G) \leq Kd \leq \frac{K}{\alpha} \delta(G^*),$$

and thus G^* is a (K/α) -nearly regular graph. We now estimate the number of vertices of G^* . Denote the latter by t . While creating G^* from G , we deleted less than $(n-t)\alpha d$ edges, and thus

$$\Delta(G^*) \geq d(G^*) > \frac{2(|E(G)| - (n-t)\alpha d)}{t} = \frac{nd - 2(n-t)\alpha d}{t}.$$

But $\Delta(G^*) \leq \Delta(G) \leq Kd$, implying

$$\frac{nd - 2(n-t)\alpha d}{t} \leq Kd.$$

Solving the above inequality for t , we get $t \geq \frac{1-2\alpha}{K-2\alpha}n$, supplying the required lower bound for the number of vertices of G^* . To bound the number of its edges, note that the number of vertices deleted is $n-t$, and hence the number of edges deleted is at most $(n-t)\alpha d$, leaving at least

$$\frac{1}{2}nd - (n-t)\alpha d = \left(\frac{1}{2} - \alpha\right)nd + t\alpha d \geq \left(\frac{1}{2} - \alpha\right)nd + \frac{1-2\alpha}{K-2\alpha}n\alpha d = \frac{K-2K\alpha}{2K-4\alpha}nd,$$

as needed. $\square$

Remark 1. It is instructive to observe that the above argument breaks down completely for $\alpha \geq 1/2$. Therefore, when estimating $f(n, c)$ from below for small c , in particular for $c < 2$, we will adapt a different strategy.

We proceed with the following result, whose proof resembles that of one of the results in [4].

PROPOSITION 2.2. *Let $K > 1$ be a constant. Every graph $G = (V, E)$ on $|V| = |V(G)| = n$ vertices contains an induced subgraph G^* on at least $n^{1+\log_2(1-\frac{1}{K})}$ vertices, for which $\Delta(G^*) \leq Kd(G^*)$.*

For large K the above estimate behaves like $n^{1-\Theta(\frac{1}{K})}$. Therefore, the assertions of Propositions 2.1 and 2.2 imply that of Proposition 1.1.

Proof of Proposition 2.2. Set $G_0 = G$, $k^* = \log_2 n$. For $i = 0, \dots, k^*$ repeat the following loop. Set

$$n_i = |V(G_i)|, \quad \Delta_i = \Delta(G_i), \quad d_i = d(G_i).$$

If $\Delta_i \leq Kd_i$, abort the loop. Otherwise, delete repeatedly vertices of degree at least $Kd_i/2$ from G_i until there are none left. Let G_{i+1} be the resulting graph and increment i .

Denote by G^* the resulting graph of the above described process. Observe that at iteration i we delete at most $|E(G_i)|/(Kd_i/2) = (n_i d_i/2)/(Kd_i/2) = n_i/K$ vertices,

and thus $n_{i+1} \geq (1 - 1/K)n_i$. It follows that

$$|V(G^*)| \geq \left(1 - \frac{1}{K}\right)^{k^*} n = n^{1 + \log_2(1 - \frac{1}{K})}.$$

If G^* was created when the above loop was aborted due to $\Delta(G_i) \leq Kd(G_i)$, then obviously the obtained graph satisfies the assertion of the theorem. Otherwise, G^* was obtained after k^* iterations. At each such iteration i , we have $\Delta_{i+1} \leq Kd_i/2$ and $d_i \leq \Delta_i/K$, implying $\Delta_{i+1} \leq \Delta_i/2$. Therefore, in this case

$$\Delta(G^*) \leq \Delta_0 \cdot \left(\frac{1}{2}\right)^{k^*} < n \cdot \left(\frac{1}{2}\right)^{\log_2 n} = 1,$$

implying that G^* has no edges, and thus $\Delta(G^*) = d(G^*) = 0$, and G^* can again serve as the required graph. $\square$

2.2. Small c . Next we treat the more challenging case where the constant c in $f(n, c)$ is very close to 1. Throughout this subsection ε denotes a small positive real.

We start with several lemmas.

LEMMA 2.3. *Let $G = (V, E)$ be a graph on n vertices with density p , and let $\varepsilon > 0$. Then G contains an induced subgraph G' of density $p' \geq p$ on a set of*

$$n' \geq \varepsilon^{\frac{2}{\varepsilon} \ln(\frac{1}{p})} n$$

vertices so that every set of $t \geq \varepsilon n'$ vertices of G' spans at most $\binom{t}{2} p'(1 + \varepsilon)$ edges.

Proof. Set $G_0 = G$. For $i = 0, 1, \dots$ repeat the following loop. Set

$$n_i = |V(G_i)|, \quad m_i = |E(G_i)|, \quad p_i = \frac{m_i}{\binom{n_i}{2}}.$$

If G_i contains a subset $U \subseteq V(G_i)$ of at least εn_i vertices such that $e_{G_i}(U) \geq \binom{|U|}{2} p_i(1 + \varepsilon)$, then set $G_{i+1} := G_i[U]$, $i := i + 1$.

Observe that after k iterations of the above loop, the density p_k of the current graph G_k satisfies $p_k \geq (1 + \varepsilon)^k p_0$. Thus, if the loop is repeated at least $\frac{2}{\varepsilon} \ln(\frac{1}{p})$ times, we have

$$p_k \geq (1 + \varepsilon)^{\frac{2}{\varepsilon} \ln(\frac{1}{p})} \cdot p > 1,$$

which is a contradiction. It follows that the above process concludes after less than $\frac{2}{\varepsilon} \ln(\frac{1}{p})$ iterations. The resulting graph G_k has n_k vertices and m_k edges. Observe that at each iteration the number of vertices of the new graph is at least an ε -proportion of the number of vertices of the previous graph. Therefore,

$$n_k \geq \varepsilon^k |V(G)| \geq \varepsilon^{\frac{2}{\varepsilon} \ln(\frac{1}{p})} n.$$

We can thus take $G' = G_k$, $n' = n_k$ to complete the proof. $\square$

LEMMA 2.4. *Let $G = (V, E)$ be a graph on $|V| = n$ vertices with $m = |E|$ edges and density $p = m/\binom{n}{2} \geq n^{-a}$ for some constant $0 < a < 1$. Suppose that*

$$(1) \quad \text{every } t \geq \varepsilon n \text{ vertex in } G \text{ spans at most } \binom{t}{2} p(1 + \varepsilon) \text{ edges.}$$

Then for every subset $U \subseteq V$ of cardinality $|U| = \varepsilon n$ in G , there are at most

$$\varepsilon n^2 p(1 + 2\sqrt{\varepsilon})$$

edges between U and its complement in G .

Proof. Assume that $U \subset V$ contradicts the above statement. Denote

$$e_1 = e_G(U), \quad e_2 = e_G(U, V - U), \quad e_3 = e_G(V - U).$$

Then $e_1 + e_2 + e_3 = m = \binom{n}{2}p$.

Choose uniformly at random a subset $X \subset V - U$ of cardinality $|X| = x = \sqrt{\varepsilon}|V - U| = \sqrt{\varepsilon}(1 - \varepsilon)n$. Then the expected number of edges of G spanned by $U \cup X$ is

$$\begin{aligned} E[e_G(U \cup X)] &= e_1 + \frac{x}{|V - U|}e_2 + \frac{x(x-1)}{|V - U|(|V - U| - 1)}e_3 \\ &> e_1 + \frac{x}{n - \varepsilon n}e_2 + \frac{x^2}{(n - \varepsilon n)^2} \left(1 - \frac{1}{x}\right) e_3 \\ &= e_1 + \sqrt{\varepsilon}e_2 + \varepsilon e_3 - O(n) = e_1 + \sqrt{\varepsilon}e_2 + \varepsilon(m - e_1 - e_2) - O(n) \\ &\geq (\sqrt{\varepsilon} - \varepsilon)e_2 + \varepsilon m - O(n) \geq (\sqrt{\varepsilon} - \varepsilon)\varepsilon n^2 p(1 + 2\sqrt{\varepsilon}) + \varepsilon m - O(n) \\ &= (\sqrt{\varepsilon} - \varepsilon)\varepsilon(1 + 2\sqrt{\varepsilon})n^2 p + \frac{\varepsilon n^2 p}{2} - O(n) =: A. \end{aligned}$$

On the other hand, by the assumption on G , every such set $U \cup X$ satisfies

$$\begin{aligned} e_G(U \cup X) &\leq \left(\varepsilon n + \sqrt{\varepsilon}(1 - \varepsilon)n\right)p(1 + \varepsilon) \\ &\leq \frac{\varepsilon n^2 p}{2}(1 + \sqrt{\varepsilon} - \varepsilon)^2(1 + \varepsilon) = \frac{\varepsilon n^2 p}{2}(1 + 2\sqrt{\varepsilon} - \varepsilon - 2\varepsilon^{3/2} + \varepsilon^2)(1 + \varepsilon) \\ &= \frac{\varepsilon n^2 p}{2}(1 + 2\sqrt{\varepsilon} + O(\varepsilon^{5/2})) =: B. \end{aligned}$$

Let us compare the asymptotic (in small ε) behavior of the two quantities A and B defined above. We have

$$\begin{aligned} A &= \frac{\varepsilon n^2 p}{2}(1 + (2\sqrt{\varepsilon} - 2\varepsilon)(1 + 2\sqrt{\varepsilon})) - O(n) \\ &= \frac{\varepsilon n^2 p}{2}(1 + 2\sqrt{\varepsilon} + 2\varepsilon - 4\varepsilon^{3/2}) - O(n). \end{aligned}$$

Since $n = o(n^2 p)$, we have that $A > B$ for ε small enough, which is a contradiction. $\square$

LEMMA 2.5. *Let $G = (V, E)$ be a graph on $|V| = n$ vertices with $m = |E|$ edges and density $p = m/\binom{n}{2} \geq n^{-a}$ for some constant $0 < a < 1$. Suppose that (1) holds. Then, for all sufficiently large n , G contains an induced subgraph G^* on at least $(1 - \varepsilon - 2\sqrt{\varepsilon})n > n/2$ vertices with maximum degree $\Delta(G^*) \leq (1 + 3\sqrt{\varepsilon})pn$ and minimum degree $\delta(G^*) \geq (1 - 2\sqrt{\varepsilon})np$. In particular, G^* is c -nearly regular for $c = (1 + 6\sqrt{\varepsilon})$.*

Proof. Let U be a set of εn vertices of highest degrees in G (ties are broken arbitrarily). Set $H = G[V - U]$. We claim that all vertex degrees in H are at most

$np(1 + 3\sqrt{\varepsilon})$. If this is not so, then the degrees of all vertices of U in G are at least $np(1 + 3\sqrt{\varepsilon})$, implying (through condition (1))

$$\begin{aligned} e_G(U, V - U) &\geq \varepsilon n^2 p(1 + 3\sqrt{\varepsilon}) - 2e_G(U) \geq \varepsilon n^2 p(1 + 3\sqrt{\varepsilon}) - \varepsilon^2 n^2 p(1 + \varepsilon) \\ &> \varepsilon n^2 p(1 + 2\sqrt{\varepsilon}) \end{aligned}$$

for small enough ε , thus contradicting Lemma 2.4. Therefore, H is an induced subgraph of G on $|V(H)| = (1 - \varepsilon)n$ vertices, of maximum degree $\Delta(H) \leq np(1 + 3\sqrt{\varepsilon})$, still satisfying condition (1), and having

$$\begin{aligned} |E(H)| &\geq \binom{n}{2} p - \frac{\varepsilon^2 n^2}{2} p(1 + \varepsilon) - \varepsilon n^2 p(1 + 2\sqrt{\varepsilon}) \\ (2) \quad &\geq \binom{n}{2} p - 2\varepsilon n^2 p \end{aligned}$$

edges.

We now repeatedly delete from H vertices of degree less than $np(1 - 2\sqrt{\varepsilon})$ until there are no such vertices, or until we have deleted $2\sqrt{\varepsilon}n$ of them. Assume the latter case and denote the set of $2\sqrt{\varepsilon}n$ deleted vertices by W . Then the set $V(H) - W$ has $|V(H) - W| = (1 - \varepsilon - 2\sqrt{\varepsilon})n$ vertices and by (2) spans at least

$$\begin{aligned} e_H(V(H) - W) &\geq |E(H)| - 2\sqrt{\varepsilon}n \cdot np(1 - 2\sqrt{\varepsilon}) \\ &\geq \binom{n}{2} p - 2\varepsilon n^2 p - 2\sqrt{\varepsilon}n^2 p(1 - 2\sqrt{\varepsilon}) \\ &= \binom{n}{2} p - 2\sqrt{\varepsilon}n^2 p + 2\varepsilon n^2 p = \frac{n^2 p}{2} (1 - 4\sqrt{\varepsilon} + 4\varepsilon) - O(np) \end{aligned}$$

edges. On the other hand, by condition (1), the set $V(H) - W$ satisfies

$$\begin{aligned} e_H(V(H) - W) &\leq \binom{(1 - 2\sqrt{\varepsilon} - \varepsilon)n}{2} p(1 + \varepsilon) \leq (1 - 2\sqrt{\varepsilon} - \varepsilon)^2 (1 + \varepsilon) \frac{n^2 p}{2} \\ &= \frac{n^2 p}{2} (1 - 4\sqrt{\varepsilon} + 3\varepsilon + O(\varepsilon^{3/2})). \end{aligned}$$

Comparing the above two estimates for $e_H(V(H) - W)$, we get a contradiction for small enough ε .

It follows that the above deletion process stops before $2\sqrt{\varepsilon}n$ vertices have been deleted. Denote the resulting graph by G^* . Then G^* has $|V(G^*)| \geq n - \varepsilon n - 2\sqrt{\varepsilon}n > \frac{n}{2}$ vertices and has maximum degree $\Delta(G^*) \leq np(1 + 3\sqrt{\varepsilon})$ and minimum degree $\delta(G^*) \geq np(1 - 2\sqrt{\varepsilon})$. Hence

$$\Delta(G^*) \leq \frac{1 + 3\sqrt{\varepsilon}}{1 - 2\sqrt{\varepsilon}} \delta(G^*) < (1 + 6\sqrt{\varepsilon}) \delta(G^*),$$

completing the proof of the lemma. $\square$

We are now ready to prove Theorems 1.2 and 1.3.

Proof of Theorem 1.2. By Lemma 2.3 (with $\frac{\varepsilon^2}{36}$ playing the role of ε), G contains an induced subgraph G' of density $p' \geq p$ on

$$n' \geq \left(\frac{\varepsilon^2}{36}\right)^{\frac{72}{\varepsilon^2} \ln(\frac{1}{p})} \cdot n = \left(\frac{\varepsilon}{6}\right)^{\frac{144}{\varepsilon^2} \ln(\frac{1}{p})} \cdot n$$

vertices, such that every set of $t \geq \frac{\varepsilon^2}{36} n'$ vertices of G' spans at most $\binom{t}{2} p' (1 + \frac{\varepsilon^2}{36})$ edges. Since $p' \geq p > n^{-1/2}$ (as $p > 0$ is a constant and n is large), Lemma 2.5 implies that G' contains an induced subgraph G^* on at least $n'/2$ vertices which is c -nearly-regular for $c = 1 + 6\sqrt{\frac{\varepsilon^2}{36}} = 1 + \varepsilon$, as needed. $\square$

Proof of Theorem 1.3. Set

$$\varepsilon_0 = \frac{\varepsilon^2}{36}, \quad a = \frac{\varepsilon_0}{3 \ln(1/\varepsilon_0)}.$$

Let $G = (V, E)$ be a graph on n vertices. Denote by $p = |E|/\binom{n}{2}$ the density of G . The average degree of G is at most np and by Turán's theorem G contains an independent set U of size $n/(np + 1)$. Therefore we can assume that $p \geq n^{-a}$, as otherwise,

$$\begin{aligned} \frac{n}{np + 1} &\geq \frac{n}{n^{1-a} + 1} \geq (1 - o(1))n^a = (1 - o(1))n^{\frac{\varepsilon_0}{3 \ln(1/\varepsilon_0)}} \\ &= (1 - o(1))n^{\frac{\varepsilon^2}{108 \ln(36/\varepsilon^2)}} > n^{\frac{\varepsilon^2}{250 \ln(1/\varepsilon)}}, \end{aligned}$$

where here we used the assumptions that ε is sufficiently small and n is sufficiently large. This gives an induced 0-regular subgraph of G , and we can thus indeed assume that $p \geq n^{-a}$.

By Lemma 2.3, G contains an induced subgraph G' of density $p' \geq p$ on

$$n' \geq \varepsilon_0^{\frac{2}{\varepsilon_0} \ln(n^a)} \cdot n = \varepsilon_0^{\frac{2}{3 \ln(1/\varepsilon_0)} \ln n} \cdot n = n^{\frac{1}{3}}$$

vertices, in which the density of the induced subgraph on any set of at least $\varepsilon_0 n'$ vertices does not exceed $p'(1 + \varepsilon_0)$. By Lemma 2.5, G' (and hence G) contains an induced subgraph on at least $0.5n' \geq 0.5n^{1/3}$ vertices, which is $(1 + 6\sqrt{\varepsilon_0}) = (1 + \varepsilon)$ -nearly regular, completing the proof. $\square$

Remark 2. Note that we have actually proved the following result, which is stronger than the assertion of Theorem 1.3: Every graph on n vertices contains either an independent set of size at least $n^{\frac{\varepsilon^2}{250 \ln(1/\varepsilon)}}$, or a $(1 + \varepsilon)$ -nearly regular induced subgraph on at least $0.5n^{1/3}$ vertices.

3. Upper bounds.

3.1. The strictly regular case.

Proof of Theorem 1.4. Fix an integer k satisfying

$$k \geq Cn^{1/2} \ln^{3/4} n,$$

where $C > 0$ is a sufficiently large constant to be set later.

We will work with the following model of random graphs on n vertices which we denote by $G(n, \bar{p})$. Let $\bar{p} = (p_1, \dots, p_n)$, where

$$p_i = \frac{1}{4} + \frac{i}{2n}, \quad i = 1, \dots, n.$$

Then $G(n, \bar{p})$ is the probability space of graphs with vertex set $[n] = \{1, \dots, n\}$, where for each pair $1 \leq i \neq j \leq n$, (i, j) is an edge of $G(n, \bar{p})$ with probability $p_i p_j$, independently of all other pairs. Notice that the probability of each individual pair (i, j) to be an edge of $G(n, \bar{p})$ is strictly between $1/16$ and $9/16$.

PROPOSITION 3.1. *Let $X_1, \dots, X_t$ be independent Bernoulli random variables, where $\Pr[X_i = 1] = \rho_i$, $i = 1, \dots, t$. Let $X = X_1 + \dots + X_t$. Assume that $1/16 \leq \rho_i \leq 9/16$ for $i = 1, \dots, t$. Then for every integer $0 \leq s \leq t$, $\Pr[X = s] \leq c_0/\sqrt{t}$ for some absolute constant $c_0 > 0$.*

Proof. For every $1 \leq i \leq t$, we represent X_i as a product $X_i = Y_i \cdot Z_i$, where $\{Y_i, Z_i\}$ is a collection of mutually independent Bernoulli random variables defined by $\Pr[Y_i = 1] = \frac{9}{16}$, $\Pr[Z_i = 1] = \frac{16}{9}\rho_i$, $i = 1, \dots, t$.

Set $I_0 = \{1 \leq i \leq t : Z_i = 1\}$. Since $\rho_i \geq 1/16$ for all $1 \leq i \leq t$, we have that $\Pr[Z_i = 1] \geq \frac{1}{9}$ and $E[|I_0|] \geq \frac{t}{9}$. By standard large deviation arguments, $|I_0| \geq t/10$ with probability $1 - o(\frac{1}{\sqrt{t}})$. Thus

$$\begin{aligned} \Pr[X = s] &= \sum_{I \subseteq [t]} \Pr[I_0 = I] \cdot \Pr\left[\sum_{i \in I} Y_i = s\right] = \sum_{I \subseteq [t]} \Pr[I_0 = I] \Pr[B(|I|, 9/16) = s] \\ &\leq \Pr[|I_0| < t/10] + \sum_{|I| \geq t/10} \Pr[I_0 = I] \Pr[B(|I|, 9/16) = s]. \end{aligned}$$

Here $B(n, p)$ denotes the binomial random variable with parameters n and p . From known estimates on binomial random variables, we obtain that $\Pr[B(r, \frac{9}{16}) = s] \leq \frac{c}{\sqrt{r}} \leq \frac{4c}{\sqrt{t}}$ for every $r \geq \frac{t}{10}$, where $c > 0$ is an absolute constant. Plugging this estimate into the inequality above, we get the claimed result. $\square$

LEMMA 3.2. *Let $U \subseteq [n]$ be a fixed set of $|U| = k$ vertices. Then the probability that in $G(n, \bar{p})$, with $\bar{p}$ as defined above, the graph $G[U]$ is a regular graph is at most $n \left(\frac{c_1}{k}\right)^{k/2}$ for some absolute constant $c_1 > 0$.*

Proof. Fix the degree of regularity d of the regular subgraph $G[U]$ (this can be done in n ways). Let $U = \{u_1, u_2, \dots, u_k\}$. We bound the probability that the induced subgraph $G[U]$ is d -regular as follows.

Expose the edges of $G[U]$ by first exposing the edges from u_1 to $U - \{u_1\}$, then from u_2 to $U - \{u_1, u_2\}$, etc. If vertex u_i gets t_i neighbors in $\{u_1, \dots, u_{i-1}\}$, then u_i should have exactly $d - t_i$ neighbors in $\{u_{i+1}, \dots, u_k\}$. Recall that all edge probabilities in $G(n, \bar{p})$ are between $1/16$ and $9/16$. Thus Proposition 3.1 applies, and the probability of the latter event is at most $c_0/\sqrt{k-i}$. Multiplying these probabilities for $i = 1, \dots, k$, we derive that the probability that U is a d -regular graph is at most

$$\prod_{i=1}^{k-1} \frac{c_0}{\sqrt{k-i}} = \frac{(c_0)^{k-1}}{\sqrt{(k-1)!}} \leq \left(\frac{c_1}{k}\right)^{k/2},$$

where $c_1 > 0$ is an absolute constant and the last inequality follows by applying the Stirling formula. $\square$

We can now complete the proof of Theorem 1.4. We bound the probability that in $G(n, \bar{p})$ there exists a set U of size $|U| = k$ such that $U = (u_1 < \dots < u_{k/\ln k} = a < \dots < u_{k-k/\ln k} = b < \dots < u_k)$ and $G[U]$ is a regular graph by considering two possible cases depending on the difference $t = b - a$ between a and b .

Case 1. $t \leq c_2 k^{3/2}$, where $c_2 > 0$ is a small positive constant to be determined later.

The probability that there exists such a U is at most

$$(3) \quad \binom{n}{2k/\ln k} n^2 \binom{t}{k-2k/\ln k} \cdot n \left(\frac{c_1}{k}\right)^{k/2}.$$

Indeed, there are less than $\binom{n}{2k/\ln k}$ ways to choose the vertices

$$u_1, \dots, u_{k/\ln k-1}, u_{k-k/\ln k+1}, \dots, u_k,$$

less than n^2 ways to choose the vertices a and b so that the difference between them, t , is at most $c_2 k^{3/2}$, and less than $\binom{t}{k-2k/\ln k}$ ways to choose the vertices $u_{k/\ln k+1}, \dots, u_{k-k/\ln k-1}$. For each such choice, the probability that the induced subgraph on $\{u_1, u_2, \dots, u_k\}$ is regular is at most $n \left(\frac{c_1}{k}\right)^{k/2}$, by Lemma 3.2.

A simple computation shows that the expression in (3) is (much) smaller than, say, $1/n^2$ for an appropriate choice of c_2 . Indeed, since $n \leq k^2$ and $\binom{t}{k-2k/\ln k} \leq \binom{2t}{k} \leq (2et/k)^k$, this expression is at most

$$n^3 k^{4k/\ln k} \left(\frac{2et}{k}\right)^k \left(\frac{c_1}{k}\right)^{k/2} \leq n^3 2^k e^{5k} c_2^k c_1^{k/2} = n^3 [4e^{10} c_2^2 c_1]^{k/2},$$

implying the required estimate by choosing, for example, $c_2 = \frac{1}{3e^5 \sqrt{c_1}}$.

Case 2. $t \geq c_2 k^{3/2}$.

Let $U = (u_1 < \dots < u_{k/\ln k} = a < \dots < u_{k-k/\ln k} = b < \dots < u_k)$. Denote the first (smallest) $k/2$ vertices of U by U_1 , and the last $k/2$ vertices by U_2 . Observe that if $G[U]$ is a regular graph, then the two induced subgraphs $G[U_1]$ and $G[U_2]$ have the same average degree. This is highly improbable. Indeed, by the definition of $G(n, \bar{p})$ the probability of each pair inside U_1 to become an edge is strictly less than the probability of each pair inside U_2 to become an edge. In addition, since $b - a = t$, the probability that each pair $i, j \in U_1$ with $i < j, i \leq a$ is an edge is less than the probability that each pair $i', j' \in U_2$ with $i' < j', j' \geq b$ is an edge by $\Omega(t/n)$ —this is because in this case,

$$\begin{aligned} \left(\frac{1}{4} + \frac{i'}{2n}\right) \left(\frac{1}{4} + \frac{j'}{2n}\right) - \left(\frac{1}{4} + \frac{i}{2n}\right) \left(\frac{1}{4} + \frac{j}{2n}\right) &\geq \Omega\left(\frac{i' + j' - i - j}{2n}\right) \\ &\geq \Omega\left(\frac{j' - i}{2n}\right) \geq \Omega\left(\frac{t}{n}\right). \end{aligned}$$

It follows that the expected number of edges inside U_2 exceeds that inside U_1 by $\Omega(\frac{k^2 t}{n \ln k})$. By Chernoff's inequality (cf., e.g., [1, Appendix A]), we obtain that the probability that $G[U_1]$ and $G[U_2]$ have the same average degree in $G(n, \bar{p})$ is at most $\exp\{-c_3 k^2 t^2 / (n^2 \ln^2 k)\}$ for some absolute constant $c_3 > 0$.

Thus, recalling our assumption on t , we derive that the probability that there is U as above for which $G[U]$ is a regular graph is at most

$$\binom{n}{k} e^{-\frac{c_3 k^2 t^2}{n^2 \ln^2 k}} \leq \left(\frac{en}{k}\right)^k e^{-\frac{c_3 c_2^2 k^5}{n^2 \ln^2 k}} = \left(\frac{en}{k} e^{-\frac{c_3 c_2^2 k^4}{n^2 \ln^2 k}}\right)^k.$$

Finally, as $k \geq Cn^{1/2} \ln^{3/4} n$ and $\ln k = \Theta(\ln n)$, we can choose $C > 0$ to be large enough so that the above expression is (much) smaller than $1/n^2$.

Combining the two cases, we conclude that the probability that for any fixed k which is at least $Cn^{1/2} \ln^{3/4} n$ our graph contains an induced regular subgraph on k vertices is smaller than $2/n^2 < 1/n$, and as there are less than n choices for k , this shows that with positive probability the graph contains no such subgraph. This completes the proof of Theorem 1.4. $\square$

3.2. The nearly regular case.

Proof of Proposition 1.5. We will prove that for every (large) n there exists a graph G on n vertices in which every K -nearly regular induced subgraph has at most $7Kn/\log n$ vertices.

Assume first that n is of the form $n = (s+1)2^s$ for a positive integer s . Notice that $s = (1 - o(1)) \log_2 n$. Take a set V of n vertices and partition it into $s+1$ disjoint equally sized subsets $V_0, \dots, V_s$, $|V_i| = 2^s$. Now we define G as follows. For $i = 0, \dots, s$ the set V_i spans 2^{s-i} disjoint cliques of size 2^i each. There are no edges between the cliques inside V_i and no edges between distinct subsets $V_i \neq V_j$ in G .

Assume now that a subset $U \subseteq V$ spans a graph $G[U]$ satisfying $\delta(G[U]) = d$, $\Delta(G[U]) \leq Kd$. Observe that the degrees of all vertices from V_i in G are $2^i - 1$, and thus if $2^i - 1 < d$, then $U \cap V_i = \emptyset$. Let now $2^i \geq d+1$. Since $\Delta(G[U]) \leq Kd$, U has at most $Kd+1$ vertices in each clique spanned by V_i , implying $|U \cap V_i| \leq 2^{s-i}(Kd+1)$. Therefore,

$$\begin{aligned} |U| &= \sum_{i=0}^s |U \cap V_i| = \sum_{i: 2^i \geq d+1} |U \cap V_i| \\ &\leq \sum_{i=\lceil \log_2(d+1) \rceil}^s 2^{s-i}(Kd+1) < (Kd+1)2^{s-\lceil \log_2(d+1) \rceil+1} \\ &\leq \frac{2(Kd+1)}{d+1} 2^s = \frac{2(Kd+1)}{d+1} \frac{n}{s+1} \leq (2+o(1))K \frac{n}{\log n}, \end{aligned}$$

implying the desired result.

For n not of the form $n = (s+1)2^s$, choose a minimal s satisfying $(s+1)2^s \geq n$. Let $n' = (s+1)2^s$. It is easy to verify that $n' \leq 3n$. Now we can apply the above construction to create a graph G' on n' vertices in which every K -nearly induced subgraph has at most $(2+o(1))Kn'/\log n' \leq 7Kn/\log n$ vertices, and then take G to be an arbitrary induced subgraph of G' on exactly n vertices. $\square$

4. Open problems. The most intriguing open problem is that of obtaining a better estimate for $f(n, 1)$. In particular, the conjecture of Erdős, Fajtlowicz, and Staton that $f(n, 1)/\ln n$ tends to infinity as n tends to infinity remains open. The values of $f(n, 1)$ for $n \leq 17$ have been determined by the authors of [5] and by McKay, and these are indeed larger than the bounds that follow from the corresponding Ramsey numbers.

Our upper and lower bounds for $f(n, c)$ for $c > 1$ are also rather far from each other, and it will be nice to better understand the behavior of this function.

One can also study a variant of the problems considered here that deals with not necessarily induced subgraphs. Of course, every graph contains a regular subgraph on all vertices (the subgraph with no edges), and hence in this case it is natural to look for regular or nearly regular subgraphs with a large number of *edges*. For every two positive integers n, m with $m \leq \binom{n}{2}$ and a real $c \geq 1$, let $g(n, m, c)$ denote the largest g so that every graph with n vertices and m edges contains a (not necessarily

induced) c -nearly regular subgraph with at least g edges. The problems of determining or estimating the behavior of this function seems interesting. Here we can establish tighter estimates than the ones obtained for the induced case.

Consider first the case $c = 1$. Since the complete graph on n vertices can be covered by n matchings (and by $(n-1)$ for even n), it follows that every graph with m edges contains a matching of size at least m/n . This implies that $g(n, m, 1) \geq m/n$. The star $K_{1, n-1}$ shows that for some values of m and n this is essentially tight, and that $g(n, m, 1) = 1$ for all $1 \leq m < n$. By a simple application of Szemerédi's regularity lemma it can be shown (see [8]) that for every fixed $p > 0$ there is a $\delta = \delta(p) > 0$ so that $g(n, pn^2, 1) \geq \delta n^2$. This bound was significantly improved by Rödl and Wysocka [9], who proved that every graph with n vertices and pn^2 edges contains an r -regular subgraph with $r \geq \alpha p^3 n$ for some positive constant α .

For a larger constant c observe, first, that complete bipartite graphs show that for $m \geq n$, $g(n, m, c) \leq O(c(m/n)^2) = O(cd^2)$, where $d = 2m/n$ is the average degree of a graph with n vertices and m edges. Indeed, a complete bipartite graph $K_{k, n-k}$ with $k \leq n/2$ has average degree $d = \Theta(k)$. Every c -nearly regular subgraph in it has minimum degree at most k , and hence maximum degree at most ck . Thus it cannot have more than $k \cdot ck = ck^2$ edges. Therefore, for every fixed $c > 1$ there exists some $C = C(c)$ so that $g(m, n, c) \leq C(m/n)^2$ for all $m > n$. We can show that for $c > 2$ this is tight, up to a constant factor; namely, for any $c > 2$ there is a $b = b(c) > 0$ so that $g(n, m, c) \geq b(m/n)^2$ for all $m > n$. For simplicity we present the proof only for $c = 5$; the proof for any other $c > 2$ is similar.

THEOREM 4.1. *Let $G = (V, E)$ be a graph with $|V| = n$ vertices, $|E| = m > n$ edges, and average degree $d = d(G) = 2m/n$. Then G contains a 5-nearly regular subgraph with at least $\frac{d^2}{2^{12}}$ edges.*

Proof. We apply the method of Pyber [7] together with a few extra twists. Clearly we may assume that $d \geq 2^6$. First, omit from G repeatedly vertices of degree smaller than $d/2$, as long as there are such vertices. As this process can only increase the average degree, it ends with a nonempty graph G' with minimum degree at least $d/2$. Now take a spanning bipartite subgraph of G' with the maximum number of edges. It is easy to see and well known that the degree of every vertex in this bipartite subgraph is at least half its degree in G' , giving a bipartite graph H with minimum degree at least $d/4$. Put $H_1 = H$. Let A and B denote the two vertex classes of H , where $|A| \geq |B|$. Let $A_1 \subseteq A$ be a nonempty subset of A which satisfies $|N_{H_1}(A_1)| \leq |A_1|$ and A_1 is minimal with respect to containment (subject to the condition above and to being nonempty). Clearly there is such an A_1 , as $|N_{H_1}(A)| \leq |A|$ and $|N_{H_1}(v)| \geq d/4 > 1$ for all $v \in A$. Put $N_{H_1}(A_1) = B_1$ and note that by the minimality of A_1 , $|A_1| = |B_1|$. By the minimality, again, and by Hall's theorem, there is a matching M_1 saturating A_1 and B_1 . Let H_2 be the graph obtained from H_1 by deleting all edges of M_1 . Now let $A_2 \subseteq A_1$ be a nonempty, minimal subset of A_1 satisfying $|N_{H_2}(A_2)| \leq |A_2|$. As before, there is such a set, as $|N_{H_2}(A_1)| \leq |N_{H_1}(A_1)| = |A_1|$. The minimality shows, again, that in fact $|N_{H_2}(A_2)| = |A_2|$, and that there is a matching M_2 saturating A_2 and $N_{H_2}(A_2) = B_2$. Proceeding in this manner we define a sequence of sets

$$\emptyset \neq A_{d/4} \subseteq A_{d/4-1} \subseteq \cdots \subseteq A_2 \subseteq A_1 \subseteq A$$

and

$$\emptyset \neq B_{d/4} \subseteq B_{d/4-1} \subseteq \cdots \subseteq B_2 \subseteq B_1 \subseteq B$$

where $|A_i| = |B_i|$ for all i , and a sequence of pairwise edge-disjoint matchings $M_{d/4}, \dots, M_2, M_1$, where M_i is a perfect matching between A_i and B_i . Note that, indeed,

this process does not terminate before these $d/4$ phases, as initially all degrees in H are at least $d/4$, and with the omission of each matching the degrees drop by 1.

For convenience we assume, from now on, that d is a power of 2 (otherwise, simply consider only the first $d'/4$ sets A_i, B_i and matchings M_i , where $d' > d/2$ is the largest power of 2 that does not exceed d). Note that $|A_{d/8}| > d/8$, since every vertex of $B_{d/4}$ is incident with an edge of each of the matchings M_i for $d/8 \leq i \leq d/4$, and all these edges are incident with vertices of $A_{d/8}$. We consider two possible cases.

Case 1. For every i , $0 \leq i \leq \log_2 d - 4$,

$$|A_{d/2^{i+4}}| > 2|A_{d/2^{i+3}}|.$$

In this case,

$$|A_1| > 2^{\log_2 d - 3} |A_{d/8}| \geq \frac{d^2}{64},$$

and the matching M_1 is a regular subgraph with more than $d^2/64$ edges, supplying the desired result (with room to spare).

Case 2. There is an i , $0 \leq i \leq \log_2 d - 4$, such that

$$|A_{d/2^{i+4}}| \leq 2|A_{d/2^{i+3}}|.$$

In this case, take the minimum i for which this holds. Then

$$|A_{d/2^{i+3}}| > 2^i |A_{d/8}| \geq 2^{i-3} d.$$

Let H' be the graph consisting of the $\frac{d}{2^{i+4}}$ matchings M_j for $\frac{d}{2^{i+4}} \leq j < \frac{d}{2^{i+3}}$. The vertices of H' are all those saturated by the largest matching among those, namely, $M_{\frac{d}{2^{i+4}}}$. Then the maximum degree in H' is exactly $\frac{d}{2^{i+4}}$ (as every vertex of $A_{d/2^{i+3}}$ has that degree), and the average degree is at least half of that, since each of the $\frac{d}{2^{i+4}}$ matchings M_j above is of size at least half that of the largest one, which is spanning. As in H' the degree of every vertex of $A_{d/2^{i+3}}$ is exactly $\frac{d}{2^{i+4}}$, and the total number of edges of H' is at least

$$|A_{d/2^{i+3}}| \cdot \frac{d}{2^{i+4}} \geq 2^{i-3} d \cdot \frac{d}{2^{i+4}} = 2^{-7} d^2.$$

Thus, H' is a graph with maximum degree that exceeds the average degree by a factor of at most $K = 2$. We can now apply Proposition 2.1 with $K = 2$ and $\alpha = 0.4$ to conclude that H' contains a $K/\alpha = 5$ -nearly regular subgraph with at least

$$\frac{K - 2K\alpha}{2K - 4\alpha} |E(H')| \geq \frac{1}{6} 2^{-7} d^2$$

edges, completing the proof. $\square$

Acknowledgments. We thank Domingos Dellamonica and Vojta Rödl for pointing an error in an early version of the paper.

REFERENCES

- [1] N. ALON AND J. H. SPENCER, *The Probabilistic Method*, 2nd ed., Wiley, New York, 2000.
- [2] F. R. K. CHUNG AND R. L. GRAHAM, *Erdős on Graphs: His Legacy of Unsolved Problems*, A. K. Peters, Ltd., Wellesley, MA, 1998.

- [3] P. ERDŐS, *On some of my favourite problems in various branches of combinatorics*, in Proceedings of the Fourth Czechoslovakian Symposium on Combinatorics, Graphs and Complexity (Prachatice, 1990), Ann. Discrete Math. 51, North-Holland, Amsterdam, 1992, pp. 69–79.
- [4] P. ERDŐS AND M. SIMONOVITS, *Some extremal problems in graph theory*, in Combinatorial Theory and Its Applications I (Proc. Colloq., Balatonfüred, 1969), North-Holland, Amsterdam, 1970, pp. 377–390.
- [5] S. FAJTLOWICZ, T. MCCOLGAN, T. REID, AND W. STATON, *Ramsey numbers for induced regular subgraphs*, Ars Combin., 39 (1995), pp. 149–154.
- [6] R. L. GRAHAM, B. L. ROTHSCILD, AND J. H. SPENCER, *Ramsey Theory*, 2nd ed., Wiley, New York, 1990.
- [7] L. PYBER, *Regular subgraphs of dense graphs*, Combinatorica, 5 (1985), pp. 347–349.
- [8] L. PYBER, V. RÖDL, AND E. SZEMERÉDI, *Dense graphs without 3-regular subgraphs*, J. Combin. Theory Ser. B, 63 (1995), pp. 41–54.
- [9] V. RÖDL AND B. WYSOCKA, *Note on regular subgraphs*, J. Graph Theory, 24 (1997), pp. 139–154.

PERFECT CODES OF LENGTH n WITH KERNELS OF DIMENSION $n - \log(n + 1) - 3^*$

OLOF HEDEN[†]

Abstract. Perfect 1-error correcting binary codes are considered. Those of length n and with a kernel of dimension $n - \log(n + 1) - 3$ are shown to be obtainable by the Phelps construction.

Key words. perfect codes, tilings, Phelps codes

AMS subject classification. 94B25

DOI. 10.1137/060658643

1. Introduction. A perfect e -error correcting q -ary code of length n is a subset C of a direct product of n alphabets F_q with q elements,

$$C \subseteq F_q^n = F_q \times F_q \times \dots \times F_q,$$

with the property that every word of F_q^n is at distance at most e from a unique word of the code C . The distance we will be concerned with here is the *Hamming distance*:

$$d(\bar{c}, \bar{d}) = |\{i \mid c_i \neq d_i\}|, \quad \bar{c} = (c_1, c_2, \dots, c_n), \quad \bar{d} = (d_1, d_2, \dots, d_n).$$

In the case where q is a prime power and $e \geq 2$, there are only two possibilities for perfect codes: the two Golay codes [7] that have the parameters $(n, q, e) = (23, 2, 3)$, respectively, $(n, q, e) = (11, 3, 2)$; see [22] or [16].

The situation in the case $e = 1$ and $q = 2$ is quite different. There are rather many different constructions of perfect codes in this case, more than 20, and the number of different perfect codes is large [21]. Krotov [13] proved that the number of different perfect binary 1-error correcting codes of length n is at least¹

$$2^{2^{\frac{n+1}{2} - \log_2(n+1)}} \cdot 3^{2^{\frac{n-3}{4}}} \cdot 2^{2^{\frac{n+5}{4} - \log_2(n+1)}}.$$

Two binary codes C and C' are *equivalent* if there is a word $\bar{d}$ and a permutation π of the set of coordinate positions such that

$$C' = \pi(\bar{d} + C).$$

It is remarkable that, even if we count the number of equivalence classes of perfect codes, the number of perfect codes will be extremely large. For example, in the case of perfect codes of length 31, it follows from the formula above, by using the trivial estimate $31! \leq (2^5)^{31}$ for the number of permutations of the set of 31 coordinate positions, that the number of equivalence classes of perfect codes will be at least

$$\frac{2^{2^{\frac{31+1}{2} - \log_2(31+1)}} \cdot 3^{2^{\frac{31-3}{4}}} \cdot 2^{2^{\frac{31+5}{4} - \log_2(31+1)}}}{2^{155} \cdot 2^{31}} = 2^{2040} \cdot 3^{128}.$$

*Received by the editors May 2, 2006; accepted for publication (in revised form) May 8, 2008; published electronically September 4, 2008.

<http://www.siam.org/journals/sidma/22-4/65864.html>

[†]Department of Mathematics, Royal Institute of Technology (KTH), S-100 44 Stockholm, Sweden (olohed@math.kth.se).

¹This lower bound has been improved; see [15]. Their new formula is, however, far too complicated to be described here.

The case where $e = 1$ and q is an odd prime power is less studied than the binary case, but it is known that, also in this case, the situation is similar to the binary case.

Here we are concerned only with perfect 1-error correcting binary codes. These will below be called *perfect codes*. We will just consider a special class of perfect codes and show that they all can be constructed by using the Phelps construction from 1984 [18]. The Phelps construction was one of the first constructions of nonlinear perfect codes. Phelps construction is very useful for producing interesting examples of perfect codes.

The special class of perfect codes that we will consider is defined by the dimension of its kernel. This concept will now be defined.

A word $\bar{p}$ is a *period* of a set of words C if

$$\bar{p} + \bar{c} \in C \quad \text{for every } \bar{c} \in C.$$

The *kernel* of a set of words C , $\ker(C)$, is the set of all periods of C . The kernel is always a subspace of $\mathbb{Z}_2^n$ and if the all zero word $\bar{0} = (0, 0, \dots, 0)$ belongs to C , then $\ker(C)$ is a subset of C .

Our main result is the following: *Every perfect 1-error correcting binary code of length n and with a kernel of dimension $n - \log(n + 1) - 3$ can be constructed by using the Phelps construction.*

Let us give a remark on the dimension k of the kernel of perfect codes in general. It is now well known [3], [19] which the possibilities are for the integer k . Let the *rank* r of a perfect code C be the dimension of the linear space $\langle C \rangle$, spanned by the words of C . Then, for any perfect code C of length n and containing the all zero word $\bar{0}$,

$$(1) \quad k \geq \begin{cases} (n-1)/2 & \text{if } r = n - \log(n+1) + 1, \\ 2^{n-r} & \text{if } n - \log(n+1) + 2 \leq r \leq n, \end{cases}$$

and

$$(2) \quad r \leq k - 1 + 2^{n - \log(n+1) - k}.$$

Further, for all triples (n, r, k) satisfying all of the inequalities (1) and (2), there are, with just a very few exceptions, perfect codes of length n , rank r , and with a kernel of dimension k ; see, e.g., [3] and [9]. So the remark must be that our theorem covers just a very particular case.

Let us also remark that it is not very difficult to prove; see, e.g., [1], that every perfect code of rank $n - \log(n + 1) + 1$ can be constructed by using the Vasil'ev construction [23]. If the rank of the perfect code equals $n - \log(n + 1) + 2$, then it is known [2] that the perfect code is a Phelps code. In [24], Zinoviev and Zinoviev enumerated all of the perfect codes of length 15 and of rank 13. There they raised the problem whether or not there are perfect codes of length 15 and of rank 14 that are not obtainable by the Solov'eva-Phelps construction; see [20] and [17]. In [10], it was shown that there actually exist non-Phelps codes of length 15 and rank 14. It must also be remarked that in [25], all of the extended perfect codes of length 16 and rank 14 were enumerated by Zinoviev and Zinoviev. Further, it was proved in [8] that every non-full rank perfect code is obtainable by the Krotov construction [14].

Finally, in the last section, section 5, we give a *description of all of the perfect codes of length $n = 2^m - 1$, rank $r = n - \log(n + 1) + 4$ and with a kernel of dimension $k = n - \log(n + 1) - 3$* . In fact, we show that every such code can be described by a family of $t + 1 = 2^{m-4}$ perfect codes $C_0, C_1, \dots, C_t$ of length 7 such that the words in

the union of these codes span $\mathbb{Z}_2^7$. Further, each such family describe a perfect code with the above given values of length, rank, and a dimension of kernel.

We must remark that these particular results of section 5 are almost immediate consequences of the results on the tilings of [5]. In our description of this class of perfect codes, equivalence is not considered. But we wish to give the following remark: Hessler [12] defined two perfect codes C and C' , both of length n , to be *linear equivalent* if there is a nonsingular linear map of the vector space $\mathbb{Z}_2^n$ that maps the words of C onto the set of the words of C' . Hessler also gave a simple characterization when two perfect codes are linear equivalent. From that result and our description in section 5, it follows that the set of the perfect codes of length n , rank $n - \log(n+1) + 4$ and with a kernel of dimension $n - \log(n+1) - 3$ just consist of one linear equivalence class.

2. Some preliminary results. The *dual code* $C^\perp$ of a linear code C of length n consists of those words $\bar{d} = (d_1, d_2, \dots, d_n) \in \mathbb{Z}_2^n$ such that

$$c_1 d_1 + c_2 d_2 + \dots + c_n d_n \equiv 0 \pmod{2}$$

for every word $\bar{c} = (c_1, c_2, \dots, c_n) \in C$.

A *simplex code* of length $n = 2^m - 1$ is a subspace S of $\mathbb{Z}_2^n$, with the property that every nonzero word of S has weight $(n+1)/2$. The following lemma is well known [11].

LEMMA 1. *The dual code of the linear span $\langle C \rangle$, of the words of any perfect code C of length n , constitutes a simplex code S of length n and dimension $n - r$, where $r = \text{rank}(C)$.*

Let S be any simplex code of length n and dimension $n - r$. If we define an equivalence relation on the set of coordinate positions $\{1, 2, \dots, n\}$ by

$$i \sim j \quad \text{if} \quad s_i = s_j \quad \text{for every} \quad \bar{s} = (s_1, s_2, \dots, s_n) \in S,$$

then this equivalence relation induces a partition $\mathcal{P}$ of the set $\{1, 2, \dots, n\}$ into mutually disjoint equivalence classes $I_0, I_1, \dots, I_t$. It can be proved; see, e.g., [4], that with a suitable enumeration of these sets,

$$(3) \quad |I_0| + 1 = |I_1| = |I_2| = \dots = |I_t| = s, \quad \text{where} \quad t = 2^{n-r} - 1 \quad \text{and} \quad s = \frac{n+1}{t+1}.$$

Any perfect code C of rank r is equivalent to a perfect code C' , of which the simplex code $\langle C' \rangle^\perp$ gives the following partition $\mathcal{P}$:

$$(4) \quad I_0 = \{1, 2, \dots, s-1\}, I_1 = \{s, s+1, \dots, 2s-1\}, \dots, I_t = \{ts, ts+1, \dots, n\}$$

of the set of coordinate positions $\{1, 2, \dots, n\}$. To simplify our notation we will always assume that the dual code of the linear span of the perfect code we consider will induce this partition of the set of coordinate positions.

If a perfect code C of length n and of rank r is under consideration, then we use the following notation of the words of $\mathbb{Z}_2^n$:

$$(x_1, x_2, \dots, x_n) = (\bar{x}_0 | \bar{x}_1 | \dots | \bar{x}_t),$$

where

$$\bar{x}_0 = (x_1, x_2, \dots, x_{s-1}), \quad \text{and} \quad \bar{x}_i = (x_{is}, x_{is+1}, \dots, x_{(i+1)s-1}), \quad i = 1, 2, \dots, t.$$

The *weight* $w(\bar{c})$ of a word $\bar{c} = (c_1, c_2, \dots, c_n)$ is the number of nonzero coordinates c_i , $i = 1, 2, \dots, n$, or equivalently, $w(\bar{c}) = d(\bar{c}, \bar{0})$.

For any word $\bar{c} \in C$ we define the *parity weight vector* of a word $\bar{c} = (\bar{c}_0 \mid \bar{c}_1 \mid \dots \mid \bar{c}_t)$ to be the binary word

$$w_{\mathcal{P}}(\bar{c}) = (h_1, h_2, \dots, h_t) \in \mathbb{Z}_2^t, \quad \text{where } h_i \equiv w(\bar{c}_i) \pmod{2}, \quad i = 1, 2, \dots, t,$$

where $\mathcal{P}$ is the partition of the coordinate set described above.

One of the ingredients of a Phelps code is a MDS-code. We now show that every perfect code of length n and rank $r < n$ defines a MDS-code.

Let F_s denote the finite field with s elements, where s is a power of the integer two. The set of the elements of F_s constitutes an alphabet that we denote by $\mathcal{A}$:

$$(5) \quad \mathcal{A} = \{\beta_0 = 0, \beta_1, \beta_2, \dots, \beta_{s-1}\}.$$

Let $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_{s-1}$ denote the following words of length $s-1$:

$$\bar{e}_1 = (1, 0, 0, \dots, 0, 0), \quad \bar{e}_2 = (0, 1, 0, \dots, 0, 0), \quad \dots, \quad \bar{e}_{s-1} = (0, 0, 0, \dots, 0, 1).$$

We define a map φ_0 from $\mathcal{A}$ to $\mathbb{Z}_2^{s-1}$ by

$$\varphi_0(\alpha) = \begin{cases} (0, 0, \dots, 0) & \text{if } \alpha = 0, \\ \bar{e}_i & \text{if } \alpha = \beta_i, \quad i = 1, 2, \dots, s-1. \end{cases}$$

We also define maps φ_j from $\mathcal{A}$ to $\mathbb{Z}_2^s$, for $j = 1, 2, \dots, t$, by just adding an extra parity check:

$$\varphi_j(\alpha) = \begin{cases} (0, 0, \dots, 0 \mid 0) & \text{if } \alpha = 0, \\ (\bar{e}_i \mid 1) & \text{if } \alpha = \beta_i, \quad i = 1, 2, \dots, s-1. \end{cases}$$

We extend these maps to a map φ from $\mathcal{A}^{t+1}$ to $\mathbb{Z}_2^{s(t+1)-1}$ by

$$\varphi((\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_t)) = (\varphi_0(\alpha_0) \mid \varphi_1(\alpha_1) \mid \varphi_2(\alpha_2) \mid \dots \mid \varphi_t(\alpha_t)).$$

Consider any perfect code C , with corresponding simplex code $S = \langle C \rangle^\perp$. Let $M(C)$ be defined to be the set of words

$$M(C) = \{\bar{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_t) \in \mathcal{A}^{t+1} \mid \varphi(\bar{\alpha}) \in C\}.$$

Just to avoid any possibility of misunderstanding, we give an example.

Example 1. Let $s = 4$, $t = 3$, and $\mathcal{A} = \{0, \beta_1, \beta_2, \beta_3\}$. Then, e.g.,

$$\varphi((\beta_1, 0, \beta_3, \beta_2)) = (1, 0, 0 \mid 0, 0, 0, 0 \mid 0, 0, 1, 1 \mid 0, 1, 0, 1).$$

Further there is no 4-tuple $\bar{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \alpha_3) \in \mathcal{A}^4$ such that

$$\varphi(\bar{\alpha}) = (1, 1, 1 \mid 1, 1, 1, 1 \mid 1, 1, 1, 1 \mid 1, 1, 1, 1).$$

LEMMA 2. *For any perfect code C of length n and rank r such that $n - \log(n+1) < r < n$ and containing the all zero word, the code $M(C)$ is a MDS-code in $\mathcal{A}^{t+1}$ of size $|\mathcal{A}|^t$ and with minimum distance 2, where $t = 2^{n-r}$.*

Proof. We prove that $|M(C)| = s^t$ and that the minimum distance of $M(C)$ will be equal to 2. This is equivalent to saying that $M(C)$ is a MDS-code of length $t+1$ over an alphabet, with s elements and with a minimum distance equal to 2.

Assume that there are words $\bar{\alpha} = (\alpha_0, \dots, \alpha_t)$ and $\bar{\alpha}' = (\alpha'_0, \dots, \alpha'_t)$ in $M(C)$ such that $d(\bar{\alpha}, \bar{\alpha}') = 1$. Let i be the coordinate position in which $\bar{\alpha}$ and $\bar{\alpha}'$ differ. Then

$$d(\varphi(\bar{\alpha}), \varphi(\bar{\alpha}')) = d(\varphi(\alpha_i), \varphi(\alpha'_i)) = w(\varphi_i(\alpha_i) - \varphi_i(\alpha'_i)) = 2,$$

which is impossible as the minimum distance of C is three.

Now consider any word

$$\bar{\beta} = (0, \alpha_1, \alpha_2, \dots, \alpha_t) \in \mathcal{A}^{t+1}.$$

There is a unique word $\bar{c} = (\bar{c}_0 | \bar{c}_1 | \dots | \bar{c}_t)$ of C such that

$$(6) \quad d(\varphi(\bar{\beta}), \bar{c}) \leq 1.$$

Let S be the dual space of $\langle C \rangle$. From Lemma 1 and the definition of the map φ , it follows that $\varphi(\bar{\beta})$ belongs to $S^\perp$. By the definition of the simplex code S , the word $\bar{c}$ will belong to $S^\perp$. Hence also $\bar{c} - \varphi(\bar{\beta}) \in S^\perp$. By the above equation (6), this word has weight at most one. Thus, we may conclude that the only possibility is that

$$\bar{c} - \varphi(\bar{\beta}) = (\varphi_0(\alpha_0), 0, 0, \dots, 0)$$

for some element $\alpha_0 \in \mathcal{A}$. This gives that

$$(\varphi_0(\alpha_0), \varphi_1(\alpha_1), \varphi_2(\alpha_2), \dots, \varphi_t(\alpha_t)) = \bar{c} \in C$$

or equivalently that

$$(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_t) \in M(C).$$

The lemma is proved. $\square$

3. Tilings and perfect codes. One of the most fruitful approaches to perfect codes goes via tilings; see [5], [6]. A *tiling* of $\mathbb{Z}_2^n$ is a couple (V, A) of two subsets V and A of $\mathbb{Z}_2^n$ such that every word $\bar{x}$ of $\mathbb{Z}_2^n$ uniquely can be written as a sum of a word $\bar{v} \in V$ and a word $\bar{a} \in A$.

We will consider the words in the sets V and A as column vectors. To each of these sets we will associate matrices $\mathbf{V}$ and $\mathbf{A}$, respectively. If the words of V and A are enumerated as follows:

$$V = \{\bar{v}_1, \bar{v}_2, \dots, \bar{v}_{|V|}\} \quad \text{and} \quad A = \{\bar{a}_1, \bar{a}_2, \dots, \bar{a}_{|A|}\},$$

then these columns are placed in that order in the matrices $\mathbf{V}$ and $\mathbf{A}$. Sometimes we will exclude a zero column in these matrices. This will be declared each time it will occur.

The next lemma says that, to every perfect code C , we may associate a tiling.

Let C be any perfect code of length n , and let G be any linear subspace of the kernel of C . Consider the matrix $\mathbf{V}$, in which the rows constitute a base for the dual space of G . Let A be the set of columns

$$(7) \quad A = \{\mathbf{V}\bar{c}^T \mid \bar{c} \in C\},$$

and let V denote the set of columns of the matrix $\mathbf{V}$ together with the all zero word.

LEMMA 3. *If C is a perfect code of length n with A and V defined as above, then the couple (V, A) is a tiling of $\mathbb{Z}_2^g$, where $g = \dim(G^\perp)$.*

This lemma is well known [5]; the proof is a trivial exercise of verifications.

We will say that the tiling (V, A) , described in the above Lemma 3, is the *tiling associated with the perfect code C* . Let us just give two examples, that actually will be used in the sections 4 and 5, respectively.

Example 2. The *Hamming code H* of length 7 is defined to be the set of words $\bar{c} \in \mathbb{Z}_2^7$ satisfying

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \bar{c}^T = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The rows of the matrix above span the dual space of the Hamming code. If V denotes the set of columns of the 3×7 matrix above, complemented with the zero column, and A the zero column of length three, then, perhaps trivially, the couple (V, A) constitutes a tiling of $\mathbb{Z}_2^3$.

Example 3. Let C be any perfect code of length n . Let A consist of all of the words of weight less than or equal to one, and let V consist of the words of the code C . Then $|V| = 2^{n-\log(n+1)}$ and $|A| = n + 1$. From the definition of a perfect code, it follows immediately that the couple (V, A) constitutes a tiling of $\mathbb{Z}_2^n$.

We now consider any tiling (V, A) of $\mathbb{Z}_2^n$ such that $\bar{0} \in V$. Let $\mathbf{V}$ denote the matrix which as columns has the nonzero columns of V , and let $n = |V| - 1$. We will denote by $C(V, A)$ the set of words

$$C(V, A) = \{ \bar{c} \in \mathbb{Z}_2^n \mid \mathbf{V} \bar{c}^T \in A \}.$$

The following lemma is well known; see, e.g., [5], and is also a converse of Lemma 3.

LEMMA 4. *For any tiling (V, A) of $\mathbb{Z}_2^g$ such that $\bar{0} \in V$, the code $C(V, A)$ is a perfect code of length $n = 2^m - 1$, where $m = \log(|V|)$.*

We will say that $C(V, A)$ is the *perfect code associated to the tiling (V, A)* .

For any tiling (V, A) of $\mathbb{Z}_2^g$, with $|V| = n$, we will let $C^0(V, A)$ denote the set

$$C^0(V, A) = \{ \bar{c} \in \mathbb{Z}_2^n \mid w(\bar{c}) \equiv 0 \pmod{2} \text{ and } \mathbf{V} \bar{c}^T \in A \},$$

where $\mathbf{V}$ denotes a matrix, which as columns has the columns of the set V .

LEMMA 5. *For any tiling (V, A) of $\mathbb{Z}_2^g$, the code $C^0(V, A)$ is an extended perfect code of length n .*

The proof of this lemma is an easy exercise in coding theory, and further, the result can also partly be seen as an extension of Lemma 4.

We shall say that $C^0(V, A)$ is the *extended perfect code of even weight words associated to the tiling (V, A)* . Note that different orderings of the columns in the matrix $\mathbf{V}$ will give equivalent perfect codes.

Consider any tiling (V, A) . Let $Im(C(V, A))$ denote the following set of columns:

$$Im(C(V, A)) = \{ \mathbf{V} \bar{c}^T \in A \mid \bar{c} \in C(V, A) \}.$$

Similarly the set $Im(C^0(V, A))$ is defined.

LEMMA 6. *For any tiling (V, A) of $\mathbb{Z}_2^g$ such that $\bar{0} \in V$,*

$$|Im(C(V, A))| = |Im(C^0(V, A))| = 2^{\text{rank}(V) - \log(|V|)}.$$

If $\text{rank}(V) = g$, then

$$\text{Im}(C(V, A)) = \text{Im}(C^0(V, A)) = A.$$

Proof. We prove the statements of the lemma just for the set $\text{Im}(C(V, A))$. The proof for the set $\text{Im}(C^0(V, A))$ is quite similar.

The code $C(V, A)$ is a perfect code of length $|V| - 1$, and thus

$$(8) \quad |C(V, A)| = 2^{|V|-1-\log|V|}.$$

Let D be the null space of $\mathbf{V}$. Then

$$C(V, A) = (\bar{p}_1 + D) \cup (\bar{p}_2 + D) \cup \cdots \cup (\bar{p}_z + D)$$

for some words $\bar{p}_1, \bar{p}_2, \dots, \bar{p}_z$. The integer z will be equal to

$$(9) \quad z = |C(V, A)|/|D|,$$

and the words $\bar{p}_1, \bar{p}_2, \dots, \bar{p}_z$ will have the properties: $\mathbf{V}\bar{p}_i^T \in A$ for $i = 1, 2, \dots, z$, and $\mathbf{V}\bar{p}_i^T \neq \mathbf{V}\bar{p}_j^T$ for $i \neq j$. Hence, from the definition of $\text{Im}(C(V, A))$, it thus follows that $|\text{Im}(C(V, A))| = z$. As the subspace D of $\mathbb{Z}_2^{|V|-1}$ is the null space of the matrix $\mathbf{V}$, we get that $\dim(D) = |V| - 1 - \text{rank}(\mathbf{V})$. By combining the equations (8) and (9), we now get the first statement of the lemma.

If V has rank g , then any vector of $\mathbb{Z}_2^g$ belongs to the column space of $\mathbf{V}$; see any elementary text book in linear algebra. From the definition of the code $C(V, A)$ and the set $\text{Im}(C(V, A))$, the second statement of the lemma now follows. $\square$

The next two lemmas will be crucial in our investigations. In both lemmas we will let $\bar{v}_1, \bar{v}_2, \bar{v}_3, \dots$ denote the columns number 1, 2, 3, etc., of the matrix $\mathbf{V}$.

LEMMA 7. Consider any tiling (V, A) of $\mathbb{Z}_2^g$ and assume that the all zero word belongs to both A and V . Let

$$W = \{i \mid \bar{0} \neq \bar{v}_i \in \text{Im}(C(A, V))\}.$$

If the word $\bar{e}_i$ of weight one belongs to the linear span $\langle C(V, A) \rangle$, then $i \in W$. If further V has maximal rank, then for any $i \in W$, the word $\bar{e}_i$ belongs to $\langle C(V, A) \rangle$.

Proof. As the all zero column belongs to A , we conclude that the dual space of the row space of $\mathbf{V}$ belongs to $C(V, A)$. Remind, that in this case, we assume that the matrix $\mathbf{V}$ does not contain the all zero column.

Let $\bar{g}_1, \bar{g}_2, \dots, \bar{g}_r$ be a set of base vectors for $\langle C(V, A) \rangle$ such that $\mathbf{V}\bar{g}_i^T = \bar{a}_i$, for $i = 1, 2, \dots, t$, and such that $\bar{g}_{t+1}, \bar{g}_{t+2}, \dots, \bar{g}_r$ constitute a base for the dual space of the row space of $\mathbf{V}$. We note that

$$\mathbf{V}\bar{g}_i^T \neq \mathbf{V}\bar{g}_j^T \quad \text{for any } 1 \leq i < j \leq t.$$

This implies by Lemma 6 that if $\mathbf{V}$ has maximal rank, then

$$(10) \quad |A \setminus \{\bar{0}\}| = t.$$

Assume that $\bar{e}_i \in \langle C(V, A) \rangle$, that is,

$$\bar{e}_i = \sum_{j=1}^r \lambda_j \bar{g}_j.$$

Then

$$\bar{v}_i = \mathbf{V} \left(\sum_{j=1}^r \lambda_j \bar{g}_j \right)^T = \sum_{j=1}^t \lambda_j \bar{a}_j, \quad \text{where } \bar{a}_j \neq \bar{0}, \quad j = 1, 2, \dots, t.$$

This implies that

$$(\lambda_1, \lambda_2, \dots, \lambda_t, 0, 0, \dots, 0) \in C(A, V),$$

and consequently that $i \in W$.

Next assume that $\mathbf{V}$ has maximal rank. Then, as $\bar{0} \in A$, we get from Lemma 6 that to any $\bar{a} \in A$ there is a word $\bar{c} \in C(V, A)$ such that $\bar{a} = \mathbf{V}\bar{c}^T$.

Assume that i is any element of W . Then, for some word $(\lambda_1, \dots, \lambda_t) \in C(A, V)$, compare equation (10),

$$\mathbf{V}\bar{e}_i^T = \bar{v}_i = \sum_{j=1}^t \lambda_j \bar{a}_j = \mathbf{V} \left(\sum_{j=1}^t \lambda_j \bar{g}_j \right)^T.$$

This implies that

$$\mathbf{V} \left(\sum_{j=1}^t \lambda_j \bar{g}_j - \bar{e}_i \right)^T = \bar{0},$$

and thus

$$\sum_{j=1}^t \lambda_j \bar{g}_j - \bar{e}_i = \sum_{j=t+1}^r \lambda_j \bar{g}_j.$$

The lemma is proved. $\square$

Similarly one may prove the following.

LEMMA 8. Consider any tiling (V, A) of $\mathbb{Z}_2^g$ such that the all zero word belongs to V . Let

$$W = \{i \mid \bar{0} \neq \bar{v}_i \in \text{Im}(C^0(A, V))\}.$$

If the word $\bar{e}_i$ of weight one belongs to $\langle C(V, A) \rangle$, then $i \in W$. If further V has maximal rank, then for any $i \in W$, the word $\bar{e}_i$ belongs to $\langle C(V, A) \rangle$.

Proof. In the case where $\bar{0} \in A$, then

$$\text{Im}(C^0(A, V)) = \text{Im}(C(A, V)),$$

and the lemma follows immediately from the previous lemma.

If $\bar{0} \notin A$, then consider any word $\bar{c}$ of the perfect code $C(V, A) = \{\bar{c} \mid \mathbf{V}\bar{c}^T \in A\}$. Let $\bar{a} = \mathbf{V}\bar{c}^T$.

As $\langle \bar{c} + C(V, A) \rangle = \langle C(V, A) \rangle$ we note that

$$(11) \quad \bar{e}_i \in \langle C(V, A) \rangle \Leftrightarrow \bar{e}_i \in \langle \bar{c} + C(V, A) \rangle.$$

However, the code $\bar{c} + C(V, A)$ is the perfect code associated to the tiling $(V, \bar{a} + A)$. Further, as all of the words of $C^0(A, V)$ have even weight, we note that

$$\text{Im}(C^0(A, V)) = \text{Im}(C^0(\bar{a} + A, V)).$$

Hence

$$(12) \quad \bar{v}_i \in \text{Im}(C^0(A, V)) \Leftrightarrow \bar{v}_i \in \text{Im}(C^0(\bar{a} + A, V)).$$

As $\bar{0} \in \bar{a} + A$, the lemma now follows from (11), (12), and the first case considered above. $\square$

4. Proof of the theorem. In the first part of this section we consider any perfect code C of length n , containing the all zero word, and with a kernel of dimension k .

Let $\mathbf{V}$ be a matrix in which the rows $\bar{r}_i$, for $i = 1, 2, \dots, n - k$ constitute a set of base vectors for the dual space of the kernel of C . It is well known that, for any perfect code C , the dual space of the linear span $\langle C \rangle$ belongs to the kernel of C [11], and by definition, that

$$\langle C \rangle^\perp \subseteq \ker(C)^\perp.$$

Let r denote the rank of C . We will always assume that the first $n - r$ rows $\bar{r}_1, \bar{r}_2, \dots, \bar{r}_{n-r}$ of $\mathbf{V}$ constitute a base of $\langle C \rangle^\perp$.

Let (V, A) be the tiling of $\mathbb{Z}_2^{n-k}$ associated to the perfect code C . The element a_{ij} in the matrix $\mathbf{A} = (a_{ij})$ will then, by Lemma 6 as V has maximal rank, be the dot product of the row $\bar{r}_i$, with at least one word of C . It now follows, from the fact that the first $n - r$ rows of the matrix $\mathbf{V}$ all belong to the dual space of C , that the first $n - r$ rows of the matrix $\mathbf{A}$ must be rows of just zeros:

$$\mathbf{A} = \begin{pmatrix} \bar{0} & \bar{0} & \cdots & \bar{0} \\ \bar{v}_1 & \bar{v}_2 & \cdots & \bar{v}_{|A|} \end{pmatrix}.$$

In the matrix above, the vector $\bar{0}$ is a column vector. Also the vectors $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_{|A|}$ are column vectors, and we will let V' denote the set of these vectors. (Similar notations are used below.) If we combine this fact with the fact that (V, A) is a tiling, we observe that *any possible column vector of length $n - r$ appears at least once in the submatrix of the matrix $\mathbf{V}$ consisting of the first $n - r$ rows of $\mathbf{V}$.*

As we are interested in the structure of a perfect code, we may apply any permutation π on the set of coordinate positions and thus, without any loss of generality, we may assume that the matrix $\mathbf{V}$ can be described as the matrix

$$\mathbf{V} = \left(\begin{array}{ccc|ccc|ccc|ccc|ccc|ccc} 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ 0 & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 & 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ \hline \bar{a}_1^{(0)} & \cdots & \bar{a}_{s_0}^{(0)} & \bar{a}_1^{(1)} & \cdots & \bar{a}_{s_1}^{(1)} & \bar{a}_1^{(2)} & \cdots & \bar{a}_{s_2}^{(2)} & \bar{a}_1^{(3)} & \cdots & \bar{a}_{s_3}^{(3)} & \cdots & \bar{a}_1^{(t)} & \cdots & \bar{a}_{s_t}^{(t)} \end{array} \right).$$

We must also remark that here the set V consists of the columns of the matrix $\mathbf{V}$ and the all zero column. Further, $C = C(V, A)$.

Let, for $i = 1, 2, \dots, t$, $\mathbf{A}^{(i)}$ denote the matrix which as columns has the columns in the set $A^{(i)} = \{\bar{a}_1^{(i)}, \bar{a}_2^{(i)}, \dots, \bar{a}_{s_i}^{(i)}\}$ and $\mathbf{A}^{(0)}$ a matrix with columns in the set $A^{(0)} = \{\bar{a}_1^{(0)}, \bar{a}_2^{(0)}, \dots, \bar{a}_{s_0}^{(0)}\}$. We will later use the abbreviation

$$\mathbf{V} = \left(\begin{array}{c|c|c|c|c|c} 0 & 1 & 2 & 3 & \cdots & t \\ \hline \mathbf{A}^{(0)} & \mathbf{A}^{(1)} & \mathbf{A}^{(2)} & \mathbf{A}^{(3)} & \cdots & \mathbf{A}^{(t)} \end{array} \right)$$

for the above matrix.

We will also consider the $(n - r) \times (2^{n-r} - 1)$ matrix, denoted by $\mathcal{H}$, which as columns has all possible nonzero words of length $n - r$, and in the same order as they appear as columns in the matrix $\mathbf{V}$.

As noted above, the first $n - r$ rows of the matrix $\mathbf{A}$ consist just of zeros. Thus, as a consequence of the fact that (V, A) is a tiling, we may conclude that also each couple $(V', A^{(i)})$, for $i = 1, 2, \dots, t$, as well as the couple $(V', A^{(0)} \cup \{\bar{0}\})$, are tilings of $\mathbb{Z}_2^{r-k}$.

As for any tiling (V, A) of a set $\mathbb{Z}_2^g$, it is true that $|A| = 2^g / |V|$ we may conclude that, compare equation (3),

$$s_0 + 1 = s_1 = s_2 = s_3 = \dots = s_t.$$

The all zero word belongs to the perfect code C , and hence one of the columns $\bar{v}_i$ is the all zero column, which we assume to be the column $\bar{v}_{|A|}$.

From now on we assume that the kernel of C has dimension $n - \log(n + 1) - 3$. The set A will then consist of eight columns. Further, the nonzero rows of the matrix $\mathbf{A}$ must be linearly independent, and thus the matrix

$$\mathbf{V}' = \begin{pmatrix} \bar{v}_1 & \bar{v}_2 & \dots & \bar{v}_7 \end{pmatrix}$$

has maximal rank. Let V' denote the set of columns of this matrix $\mathbf{V}'$ together with the all zero column.

LEMMA 9. For $i = 1, 2, \dots, t$ and for any word $\bar{d}_i \in C^0(A^{(i)}, V')$,

$$(13) \quad \mathbf{A}^{(i)} \bar{d}_i^T = \bar{0}.$$

Further, for any word $\bar{d} \in C(A^{(0)}, V')$,

$$(14) \quad \mathbf{A}^{(0)} \bar{d}^T = \bar{0}.$$

Proof. We will apply Lemma 7 and Lemma 8.

The length of the perfect code $C(V', A^{(i)})$ will, for each $i = 1, 2, \dots, t$, be equal to 7. For this length of perfect codes, there are only linear perfect codes, all equivalent to the perfect code given in Example 2. Hence, the linear span of the words of $C(V', A^{(i)})$ does not contain any word of weight one. As already observed above, the matrix $\mathbf{V}'$ has maximal rank. This implies, by Lemma 8 and with notation from that lemma, that $W = \emptyset$. This proves equation (13).

The proof of equation (14) is quite similar and will thus be omitted. $\square$

LEMMA 10. For the kernel of $C = C(V, A)$, the following is true:

$$C(A^{(0)}, V') \times C^0(A^{(1)}, V') \times C^0(A^{(2)}, V') \times \dots \times C^0(A^{(t)}, V') \subseteq \ker(C(V, A)).$$

Proof. It is sufficient to prove that, for any $i = 0, 1, 2, \dots, t$,

$$\{\bar{0}\} \times \{\bar{0}\} \times \dots \times \{\bar{0}\} \times C^0(A^{(i)}, V') \times \{\bar{0}\} \times \dots \times \{\bar{0}\} \subseteq \ker(C(V, A)).$$

Now consider any word $\bar{c} \in C(V, A)$ and any word $\bar{d}_i \in C^0(A^{(i)}, V')$. Let $\bar{d} = (\bar{0}|\bar{0}|\dots|\bar{0}|\bar{d}_i|\bar{0}|\dots|\bar{0})$, where $1 \leq i \leq t$. As $\mathbf{V}\bar{c}^T \in A$, we get from Lemma 9 that

$$\mathbf{V}(\bar{c} + \bar{d})^T = \mathbf{V}\bar{c}^T + \bar{0}^T \in A,$$

and hence $\bar{c} + \bar{d} \in C$. Thus $\bar{d}$ is a period of $C = C(V, A)$. The case $\bar{d} = (\bar{d}_0|\bar{0}|\dots|\bar{0})$ is treated in the same way. The lemma is proved. $\square$

We are now ready to prove our main theorem.

THEOREM 1. *Every perfect code C of length n , containing the all zero word, of rank $n - \log(n+1) + r$, where $r \geq 2$, and with a kernel of dimension $n - \log(n+1) - 3$ is equivalent to some Phelps code.*

Proof. Let H be the Hamming code that as parity check matrix has the matrix $\mathcal{H}$. We first find a partition of the perfect code C into mutually disjoint subsets $C_{\bar{h}}$, $\bar{h} \in H$.

Let $\mathcal{P}$ denote the partition (4) given by the dual code of $\langle C \rangle$, compare section 2.

The code H' will be the set of parity weight vectors that we get from C :

$$H' = \{w_{\mathcal{P}}(\bar{c}) \mid \bar{c} \in C\}.$$

Define the subset $C_{\bar{h}}$ of C , for $\bar{h} \in H'$, to be the set of those words of C that has the parity weight vector $\bar{h}$:

$$C_{\bar{h}} = \{\bar{c} \in C \mid w_{\mathcal{P}}(\bar{c}) = \bar{h}\}, \quad \bar{h} \in H'.$$

As the first $n - r$ rows of the matrix $\mathbf{A}$ just consists of zeros, we get that, for the matrix $\mathcal{H}$, defined above, and for any word $\bar{h}$ of H' :

$$\bar{h} \in H' \quad \Rightarrow \quad \mathcal{H}\bar{h}^T = \bar{0}^T.$$

Hence H' must be a subset of the Hamming code H .

We now show that indeed H' is that Hamming code H . So assume that $\bar{h}$ is such that $\mathcal{H}\bar{h}^T = \bar{0}^T$. Let $\bar{x}$ be any word of length n such that $w_{\mathcal{P}}(\bar{x}) = \bar{h}$. Then

$$\mathbf{V}\bar{x}^T = (y_1, y_2, \dots, y_{n-k})^T, \quad \text{where } y_i = 0 \quad \text{for } i = 1, 2, \dots, n - r.$$

As the couple $(A^{(0)}, V')$ is a tiling, we get that there are columns $\bar{a}_i^{(0)}$ of $A^{(0)}$ and $\bar{v}_j$ of V' such that

$$(y_{n-r+1}, y_{n-r+2}, \dots, y_{n-k})^T = \bar{a}_i^{(0)} + \bar{v}_j.$$

This implies that the word $\bar{x}$ differs in position i from a word $\bar{c} \in C$. As the coordinate position i belongs to the set of positions $\{1, 2, \dots, s_0 - 1\}$, we note that

$$w_{\mathcal{P}}(\bar{c}) = \bar{h}, \quad \text{and hence } \bar{c} \in C_{\bar{h}}.$$

We may conclude that

$$(15) \quad \mathcal{H}\bar{h}^T = \bar{0}^T \quad \Rightarrow \quad C_{\bar{h}} \neq \emptyset.$$

This proves that H' actually is a Hamming code.

By (15), for any $\bar{h} \in H$, we know that $C_{\bar{h}} \neq \emptyset$. Hence we may choose a fixed word $\bar{c}_{\bar{h}}$ in each of the subsets $C_{\bar{h}}$, $\bar{h} \in H$, of C . Let us, for each $\bar{h} \in H$, define a code $M_{\bar{h}}$ by (notation as in section 2)

$$(16) \quad M_{\bar{h}} = M(C + \bar{c}_{\bar{h}}).$$

We note that

$$\langle C \rangle = \langle C + \bar{c}_{\bar{h}} \rangle,$$

and hence, by Lemma 2, the code $M_{\bar{h}}$ is a MDS-code.

To prove our theorem, we also need a family of t partitions of the sets $\mathbb{Z}_2^{s_i}$, $i = 1, 2, \dots, t$, into extended perfect codes, and one partition of the set $\mathbb{Z}_2^{s_0}$ into perfect codes. The partition of $\mathbb{Z}_2^{s_0}$ consists of the translates of the perfect code $C(A^{(0)}, V')$. For each $i = 1, 2, 3, \dots, t$, we get a partition of $\mathbb{Z}_2^{s_i}$ by using the translates of the perfect codes $C^0(A^{(i)}, V')$.

Let φ be defined as in section 2. We conclude, from equation (16) and Lemma 10, that the set $C_{\bar{h}}$, for each $\bar{h} \in H$, equals the set

$$C_{\bar{h}} = \cup_{\bar{\alpha} \in M_{\bar{h}}} (\bar{c}_{\bar{h}} + \varphi(\bar{\alpha}) + C(A^{(0)}, V') \times C^0(A^{(1)}, V') \times \dots \times C^0(A^{(t)}, V')).$$

That C is a union of these sets $C_{\bar{h}}$, for $\bar{h} \in H$, is equivalent to the fact that C is a Phelps code; see [18] for a definition of that most well-known construction of perfect codes. The theorem is proved. $\square$

5. The perfect codes of rank $n - \log(n + 1) + 4$ and with kernels of dimension $n - \log(n + 1) - 3$. The results in this section are not directly connected to Phelps codes but are consequences of the general considerations in this paper and concern perfect codes of length n , with kernels of dimension $n - \log(n + 1) - 3$ and of rank $n - \log(n + 1) + 4$. Although they follow almost directly from the results of [5], we think the results are beautiful enough to be presented somewhere. We state the result as a theorem and use the same notation as in the previous section. Below, $\mathbf{I}_7$ denotes the 7×7 -identity matrix and $\mathbf{0}_{n \times m}$ the $n \times m$ zero matrix, and t will denote the integer $(n + 1)/16 - 1$.

THEOREM 2. *A perfect code C , with $\bar{0} \in C$, of length n and of rank $r = n - \log(n + 1) + 4$ has a kernel of dimension $k = n - \log(n + 1) - 3$ if and only if C is equivalent to a perfect code $C(V, A)$ satisfying the following three conditions:*

(i)

$$\mathbf{V} = \left(\begin{array}{c|c|c|c|c} 0 & 1 & 2 & \dots & t \\ \hline \mathbf{A}^{(0)} & \mathbf{A}^{(1)} & \mathbf{A}^{(2)} & \dots & \mathbf{A}^{(t)} \end{array} \right) \quad \text{and} \quad \mathbf{A} = \left(\begin{array}{c|c} \mathbf{0}_{s \times 7} & \mathbf{0}_{s \times 1} \\ \hline \mathbf{I}_7 & \mathbf{0}_{7 \times 1} \end{array} \right).$$

(ii) *The set of the columns in each of the matrices $\mathbf{A}^{(i)}$, $i = 1, 2, \dots, t$, constitutes the set of the words of a perfect code C_i of length 7. The set of the columns of the matrix $\mathbf{A}^{(0)}$ constitutes the set of nonzero words of a perfect code C_0 of length 7 such that $\bar{0} \in C_0$.*

(iii) *The rank of the union of the words in the perfect codes $C_0, C_1, C_2, \dots, C_t$ will be equal to seven.*

Note that, from the theorem above, it immediately follows that there are no perfect codes with the parameters $(n, r, k) = (16, 16, 8)$.

Proof. Assume that C satisfies the conditions of the theorem. Then each code $C(V', A^{(i)})$ will be a perfect code of length 7. As the matrix $\mathbf{V}'$ in this case just consists of the columns of weight 1, we get that the set of columns in each matrix $\mathbf{A}^{(i)}$ constitutes the set of the words of the code $C(V', A^{(i)})$; compare Example 3 of section 3. The theorem now follows from Lemma 6. $\square$

REFERENCES

- [1] S. V. AVGUSTINOVICH, O. HEDEN, AND F. I. SOLOV'EVA, *The classification of some perfect codes*, preprint TRITA-MAT-2001-09, KTH Stockholm, Sweden, 2001.
- [2] S. V. AVGUSTINOVICH, O. HEDEN, AND F. I. SOLOV'EVA, *The classification of some perfect codes*, Des. Codes Cryptogr., 31 (2004), pp. 313–318.

- [3] S. V. AVGUSTINOVICH, O. HEDEN, AND F. I. SOLOV'eva, *On the ranks and kernels problem for perfect codes*, Probl. Inf. Transm., 39 (2003), pp. 341–345.
- [4] A. BONISOLI, *Every equidistant linear code is a sequence of dual Hamming codes*, Ars Combin., 18 (1984), pp. 181–186.
- [5] G. COHEN, S. LITSYN, A. VARDY, AND G. ZÉMOR, *Tilings of binary spaces*, SIAM J. Discrete Math., 9 (1996), pp. 393–412.
- [6] T. ETZION AND A. VARDY, *On perfect codes and tilings: Problems and solutions*, SIAM J. Discrete Math., 11 (1998), pp. 205–223.
- [7] M. J. E. GOLAY, *Notes on digital coding*, Proc. IRE, 37(1949), p. 657.
- [8] O. HEDEN, *On the classification of perfect binary 1-error correcting codes*, preprint TRITA-MAT-2002-01, KTH, Stockholm, Sweden, 2002.
- [9] O. HEDEN, *A full rank perfect code of length 31*, Des. Codes Cryptogr., 38 (2006), pp. 125–129.
- [10] O. HEDEN AND T. WESTERBÄCK, *Non Phelps codes of length 15 and rank 14*, Australas. J. Combin., 38 (2007), pp. 141–153.
- [11] F. HERGERT, *Algebraische Methoden für Nichtlineare Codes*, Thesis, Darmstadt University of Technology, Darmstadt, Germany, 1985.
- [12] M. HESSLER, *Perfect codes as isomorphic spaces*, Discrete Math., 306 (2006), pp. 1981–1987.
- [13] D. S. KROTOV, *Lower bounds on the number of m -quasigroups of order 4 and the number of perfect binary codes*, Discrete Anal. Oper. Res., 1 (2000), pp. 47–53.
- [14] D. S. KROTOV, *Combining construction of perfect binary codes*, Prob. Inf. Transm., 36 (2000), pp. 74–79.
- [15] D. S. KROTOV AND S. V. AVGUSTINOVICH, *On the number of 1-perfect binary codes: A lower bound*, IEEE Trans. Inform. Theory, 54 (2008), pp. 1760–1765.
- [16] V. K. LEONTIEV AND V. A. ZINOVIEV, *Nonexistence of perfect codes over galois fields*, Probl. Control Inform. Theory, 2 (1973), pp. 123–132.
- [17] K. T. PHELPS, *A combinatorial construction of perfect codes*, SIAM J. Alg. Discrete Methods, 4 (1983), pp. 398–403.
- [18] K. T. PHELPS, *A general product construction for error correcting codes*, SIAM J. Alg. Discrete Methods, 5 (1984), pp. 224–228.
- [19] K. T. PHELPS AND M. VILLANUEVA, *On perfect codes: Rank and kernel*, Des. Codes Cryptogr., 27 (2002), pp. 183–194.
- [20] F. I. SOLOV'eva, *On binary nongroup code*, Methody Discretnogo Analiza, 37 (1981), pp. 65–76 (in Russian).
- [21] F. I. SOLOV'eva, *On Perfect Codes and Related Topics*, Com²Mac Lecture Note Ser. 13, Postech, Pohang, Korea, 2004.
- [22] A. TIETÄVÄINEN, *On the nonexistence of perfect codes over finite fields*, SIAM J. Appl. Math., 24 (1973), pp. 88–96.
- [23] Y. L. VASIL'EV, *On nongroup close-packed codes*, Prob. Cybernet., 8 (1962), pp. 375–378.
- [24] V. A. ZINOVIEV AND D. V. ZINOVIEV, *Binary perfect codes of length 15 by the generalized concatenated construction*, Prob. Inf. Transm., 40 (2004), pp. 27–39.
- [25] V. A. ZINOVIEV AND D. V. ZINOVIEV, *Binary extended perfect codes of length 16 and rank 14*, Prob. Inf. Transm., 42 (2006), pp. 123–138.

INCIDENCE MATRICES OF PROJECTIVE PLANES AND OF SOME REGULAR BIPARTITE GRAPHS OF GIRTH 6 WITH FEW VERTICES*

C. BALBUENA†

Abstract. Let q be a prime power and $r = 0, 1, \dots, q - 3$. Using the Latin squares obtained by multiplying each entry of the addition table of the Galois field of order q by an element distinct from zero, we obtain the incidence matrices of projective planes and the incidence matrices of $(q - r)$ -regular bipartite graphs of girth 6 and $q^2 - rq - 1$ vertices in each partite set. Moreover, in this work two Latin squares of order $q - 1$ with entries belonging to $\{0, 1, \dots, q\}$, not necessarily the same, are defined to be quasi row-disjoint if and only if the Cartesian product of any two rows contains at most one pair (x, x) with $x \neq 0$. Using these quasi row-disjoint Latin squares we find $(q - 1)$ -regular bipartite graphs of girth 6 with $q^2 - q - 2$ vertices in each partite set. Some of these graphs have the smallest number of vertices known so far among the regular graphs with girth 6.

Key words. incidence matrices, Latin squares, projective plane, girth, cages

AMS subject classifications. 05C35, 51A45

DOI. 10.1137/070688225

1. Introduction. Throughout this paper, only undirected simple graphs without loops or multiple edges are considered. Unless otherwise stated, we follow the book by Godsil and Royle [12] and the book by Lint and Wilson [14] for terminology and definitions.

Let $G = (V(G), E(G))$ be a graph with vertex set $V = V(G)$ and edge set $E = E(G)$. The *girth* of a graph G is the number $g = g(G)$ of edges in a smallest cycle. The *degree* of a vertex $v \in V$ is the number of vertices adjacent to v . A graph is called *regular* if all the vertices have the same degree. A *cage* is a k -regular graph with girth g having the smallest possible number of vertices. Simply counting the number of vertices in the distance partition with respect to a vertex yields a lower bound $n_0(k, g)$ on the number of vertices in a cage $n(k, g)$, with the precise form of the bound depending on whether g is even or odd.

$$(1.1) \quad n_0(k, g) = \begin{cases} 1 + k + k(k - 1) + \dots + k(k - 1)^{(g-3)/2} & \text{if } g \text{ is odd;} \\ 2(1 + (k - 1) + \dots + (k - 1)^{g/2-1}) & \text{if } g \text{ is even.} \end{cases}$$

Biggs [6] calls an *excess* of a k -regular graph G the difference $|V(G)| - n_0(k, g)$. Cages have been studied intensely since they were introduced by Tutte [20] in 1947. Erdős and Sachs [9] proved the existence of a graph for any value of the regularity k and the girth g ; thus most of the work carried out has been focused on constructing a smallest one [1, 3, 5, 8, 10, 11, 17, 15, 16, 22]. Biggs is the author of an impressive report on distinct methods for constructing cubic cages [7]. Royle [19] keeps a website in which all the cages known so far appear. More details about constructions on cages can be found in the survey by Wong [22] or in the survey by Holton and Sheehan [13].

*Received by the editors April 15, 2007; accepted for publication (in revised form) May 14, 2008; published electronically September 4, 2008. This research was supported by the Ministry of Education and Science, Spain, and the European Regional Development Fund (ERDF) under project MTM2005-08990-C02-02.

<http://www.siam.org/journals/sidma/22-4/68822.html>

†Departament de Matemàtica Aplicada III, Universitat Politècnica de Catalunya, Barcelona, Spain (m.camino.balbuena@upc.edu).

It is conjectured that cages with even girth are bipartite [18, 22]. If the vertices of a bipartite graph are ordered in such a way that the vertices of one partite set come first, then the adjacency matrix can be written in the form

$$(1.2) \quad A = \begin{pmatrix} 0 & N \\ N^\top & 0 \end{pmatrix}.$$

An *incidence graph* is a bipartite graph in which the elements of one partite set are declared as *lines* and the elements of the other partite set are declared as *points*. The terminology for incidence graphs is geometric. A point and a line are said to be *incident* if they are adjacent, thus the submatrix N of (1.2) is called an *incidence matrix* of the bipartite graph. If the number of points and the number of lines coincide, then N is clearly a square matrix. An incidence matrix N defines a *partial plane* when

- any line has at least two points, and
- two points are incident with at most one line.

Consequently, two lines of a partial plane have at most one point in common. The corresponding bipartite graph is called the *incidence graph of the partial plane*. Clearly the incidence graph of a partial plane has even girth $g \geq 6$.

In this work we are interested in the construction of k -regular bipartite graphs of girth 6 with small excess. When the degree $k = p^n + 1$, where $q = k - 1 = p^n$ is a prime power, $(k, 6)$ -cages are obtained as the incidence graphs of *projective planes of order* $q = k - 1$. A projective plane is a partial plane satisfying the following conditions:

- (i) Any two lines meet at a single point.
- (ii) Any two points lie on a single line.
- (iii) There exists a set of four points of which no three are collinear.

As a consequence of these conditions, it follows that a projective plane has $q^2 + q + 1$ lines and $q^2 + q + 1$ points, thus it is said to have *order* q . The incidence graph of a projective plane is a $(q + 1)$ -regular bipartite graph with diameter 3 and girth six. It has $q^2 + q + 1$ vertices in each partite set, and then this graph is a $(q + 1, 6)$ cage because its order achieves the lower bound given in (1.1).

One way for obtaining a projective plane when q is a prime power is to consider the 3-dimensional vector space $\mathbb{F}_q^3$ over the field $\mathbb{F}_q$ with q elements. The points are the 1-dimensional subspaces and the lines are the 2-dimensional subspaces. A point is incident with a line if the 1-dimensional is contained in the 2-dimensional subspace. It is not hard to see that there are $(q^3 - 1)/(q - 1) = q^2 + q + 1$ points and the same number of lines, so this structure is a projective plane denoted by $PG(2, q)$. For more details on this construction of projective planes, see, for example, [4, 12, 14].

In this work we present an alternative way to construct projective planes of order a prime power, and other partial planes whose incidence graphs are k -regular bipartite graphs of girth 6 with small excess. Let us recall that a *Latin square* of order q is a $q \times q$ matrix with entries from a set of q symbols such that each symbol occurs exactly once in each row and exactly once in each column. Using the Latin squares obtained by multiplying each entry of the addition table of the Galois field of order q by an element distinct from zero, we obtain the incidence matrices of projective planes and the incidence matrices of $(q - r)$ -regular bipartite graphs of girth 6 and $q^2 - rq - 1$ vertices in each partite set. q -regular bipartite graphs with q a prime power and $2(q^2 - 1)$ vertices of girth 6 have recently been obtained in [1]. We also improve this result for $r = 1$, finding $(q - 1)$ -regular bipartite graphs of girth 6 with $2(q^2 - q - 2)$ vertices. To do this we use a family of Latin squares called *quasi row-disjoint*, a property that will be introduced in the last section. It follows from (1.1)

that the $(q-1)$ -regular bipartite graphs constructed in this work have an excess $|V(G)| - n_0(q-1, 6) = 2(2q-5)$. The $(q-r)$ -regular bipartite graphs for $r \geq 2$ have an excess of $|V(G)| - n_0(q-r, 6) = 2(rq + q - r^2 - r - 2) = 2(q-r)(r+1) - 4$. In fact, some of these graphs have the smallest number of vertices known so far among the regular graphs with girth 6.

2. C_4 -free set of matrices. Throughout this work let $[[n]]$ denote the set of nonnegative integers $\{0, 1, \dots, n\}$ and $(n) = [[n]] \setminus \{0\} = \{1, \dots, n\}$. Let A be a matrix of order $\alpha \times \beta$ whose elements are subsets of $[[n]]$. Let $P_x(A)$ be a matrix of zeros and ones of the same order as A that satisfies

$$(P_x(A))_{ij} = 1 \text{ if and only if } x \in A_{ij}.$$

Thus, $P_x(A)$ is called the *position matrix of the symbol x in A* . The position matrices of all the symbols in A different from zero give rise to the following $(0, 1)$ -matrix $\mathcal{P}(A)$ of order $\alpha \times n\beta$ called *position matrix of A* :

$$[\mathcal{P}(A)] = [P_1(A) \cdots P_n(A)].$$

Let $\{A^1, A^2, \dots, A^s\}$ be a family of s matrices of the same number of columns whose elements are subsets of $[[n]]$. Then the $(0, 1)$ -matrix spanned by the position matrices of all of them

$$(2.1) \quad \begin{pmatrix} \mathcal{P}(A^1) \\ \mathcal{P}(A^2) \\ \vdots \\ \mathcal{P}(A^s) \end{pmatrix} = \begin{pmatrix} P_1(A^1) & \cdots & P_n(A^1) \\ P_1(A^2) & \cdots & P_n(A^2) \\ & \vdots & \\ P_1(A^s) & \cdots & P_n(A^s) \end{pmatrix}$$

is said to be the *position matrix of the family $\mathcal{F} = \{A^1, A^2, \dots, A^s\}$* . The next example shows two matrices of order 2×2 whose elements are subsets of $[[2]]$ and the position matrix of them. From now on, if there is no confusion the 1-sets will be indicated as integers.

$$(2.2) \quad A^1 = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}, \quad A^2 = \begin{pmatrix} \{1, 2\} & 0 \\ 0 & \{1, 2\} \end{pmatrix},$$

$$\begin{pmatrix} \mathcal{P}(A^1) \\ \mathcal{P}(A^2) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

As already mentioned in the introduction, our main aim is to obtain incidence matrices of bipartite k -regular graphs of girth 6 with small excess. Such incidence matrices shall be incidence matrices of partial planes. A basic known fact is that a $(0, 1)$ -matrix N is the incidence matrix of a partial plane if N is free of any submatrix J_2 of order 2×2 full of ones. Then, to achieve our purpose we propose the following definitions.

DEFINITION 2.1. A family of matrices $\mathcal{F} = \{A^1, A^2, \dots, A^s\}$ of the same number of columns whose elements are subsets of $[[n]]$ is said to be C_4 -free if the position matrix of $\mathcal{F}$ does not contain a submatrix J_2 of order 2×2 full of ones.

Notice that if the position matrix of $\{A^1, A^2\}$ in (2.2) is free of J_2 , then according to Definition 2.1, the set $\{A^1, A^2\}$ is C_4 -free.

Let I_n be the identity matrix and denote by SI_n the matrix obtained from I_n replacing each one with a subset S of $(n]$. It is also clear that the position matrix of SI_n is the incidence matrix of a partial plane consisting of n parallel lines. The above example (2.2) is a particular case of the following helpful remark.

Remark 1. Let A be any C_4 -free matrix of n columns whose entries are 1-sets of $[[n]]$ and S a subset of $(n]$. Then the set $\{A, SI_n\}$ is C_4 -free.

Proof. Note that $P_x(SI_n) = I_n$ for each symbol x in S . Therefore, the position matrix of the set $\{A, SI_n\}$ is J_2 -free because the entries of A are 1-sets. $\square$

Our immediate concern is to obtain some method for constructing families of C_4 -free matrices in such a way that the position matrix of all of them will be a J_2 -free $(0, 1)$ -matrix.

Let us recall that two Latin squares L^1 and L^2 of order n with n distinct symbols are *orthogonal* if all the n^2 pairs of $L^1 \times L^2$ are distinct. The $(1, 3)$ *conjugate* of a Latin square L of order n on n symbols denoted by $(1, 3)L$ is defined by $((1, 3)L)_{ij} = a$ if and only if $(L)_{aj} = i$; that is to say, the $(1, 3)$ conjugate of L is obtained by interchanging the roles of rows and symbols.

THEOREM 2.2. *Let L and Q be two Latin squares of order n both with entries from $(n]$. If L and Q are orthogonal, then their $(1, 3)$ conjugates form a family $\{(1, 3)L, (1, 3)Q\}$ C_4 -free.*

Proof. Suppose $(L)_{ab} = i$, $(L)_{cd} = i$, and $(Q)_{ab} = k$. Since L and Q are orthogonal, then the row x for which $(Q)_{xd} = k$ is never equal to c . The $(1, 3)$ conjugates are $((1, 3)L)_{ib} = a$, $((1, 3)L)_{id} = c$, $((1, 3)Q)_{kb} = a$, and $((1, 3)Q)_{kd} = x$, with $x \neq c$. Therefore, the position matrix of $\{(1, 3)L, (1, 3)Q\}$ is J_2 -free because

$$\begin{pmatrix} (P_a((1, 3)L))_{ib} & (P_c((1, 3)L))_{id} \\ (P_a((1, 3)Q))_{kb} & (P_c((1, 3)Q))_{kd} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}. \quad \square$$

Let q be a prime power and $\mathbb{F}_q$ the Galois field of order q . A set of $q - 1$ mutually orthogonal Latin squares (MOLS) of order q with entries $[[q - 1]]$ are obtained according to the following rule (see [14]):

$$(2.3) \quad (L^u)_{ij} = ui + j, \quad u, i, j \in \mathbb{F}_q, \quad u \neq 0.$$

Let L_q^u be the matrix as defined in (2.3), but the symbol zero is replaced with q . Clearly, the set $\{L_q^u : u \in \mathbb{F}_q \setminus \{0\}\}$ is a set of $q - 1$ MOLS of order q with entries in $[q]$. As a consequence of Theorem 2.2 we obtain the following result.

THEOREM 2.3. *Let q be a prime power and $\mathbb{F}_q$ the Galois field of order q . Then the following assertions hold:*

- (i) *For each $u \in \mathbb{F}_q$, $u \neq 0$ defines the matrix Σ_q^u by $(\Sigma_q^u)_{ij} = u(i + j)$ for $i, j \in \mathbb{F}_q$, $i \neq -j$, and q otherwise. Then $\{\Sigma_q^u : u \in \mathbb{F}_q, u \neq 0\}$ is C_4 -free.*
- (ii) *For each $u \in \mathbb{F}_q$, $u \neq 0$ defines the matrix Υ_q^u by $(\Upsilon_q^u)_{ij} = u(i + j)^{-1}$ for $i, j \in \mathbb{F}_q$, $i \neq -j$, and q otherwise. Then $\{\Upsilon_q^u : u \in \mathbb{F}_q, u \neq 0\}$ is C_4 -free.*
- (iii) *Let Σ^u, Υ^u be the Latin squares obtained from Σ_q^u and Υ_q^u , respectively, by replacing symbol q with zero. If M is the position matrix of $\{\Sigma^u : u \in \mathbb{F}_q, u \neq 0\}$, then the transpose of M is the position matrix of $\{\Upsilon^u : u \in \mathbb{F}_q, u \neq 0\}$.*

Proof. (i) Let $a = u(i + j)$ for $u, i, j \in \mathbb{F}_q$, $u \neq 0$, $i \neq -j$. Then the entry (a, j) of the $(1, 3)$ conjugate of the Latin square Σ^u is $((1, 3)\Sigma^u)_{aj} = i = u^{-1}a - j$; and for $i = -j$ we have $((1, 3)\Sigma^u)_{qj} = -j$ because $(\Sigma^u)_{-j, j} = q$. According to (2.3) the family $\{(1, 3)\Sigma^u : u \in \mathbb{F}_q, u \neq 0\}$ is a set of $q - 1$ MOLS of order q with entries from $[q]$. Hence by Theorem 2.2 we obtain that $\{\Sigma^u : u \in \mathbb{F}_q, u \neq 0\}$ is C_4 -free. Thus point (i) is valid.

(ii) Let $b = u(i + j)^{-1}$ for $u, i, j \in \mathbb{F}_q, u \neq 0, i \neq -j$. Then the entry (b, j) of the $(1, 3)$ conjugate of the Latin square Υ^u is $((1, 3)\Upsilon^u)_{bj} = i = ub^{-1} - j$; and for $i = -j$ we have $((1, 3)\Upsilon^u)_{qj} = -j$ because $(\Upsilon^u)_{-j,j} = q$. Then $\{\Upsilon_q^u : u \in \mathbb{F}_q, u \neq 0\}$ is a set of $q - 1$ MOLS of order q with entries from $(q]$ and again by Theorem 2.2 we conclude point (ii).

(iii) Let $P_x(\Sigma^u)$ be the position matrix of the element x in Σ^u , and observe that $P_x(\Sigma^u)$ is a symmetric matrix because $(P_x(\Sigma^u))_{ij} = 1$ if and only if $u(i + j) = x$, $i, j, x, u \in \mathbb{F}_q, u, x \neq 0$. Similarly let $P_x(\Upsilon^w)$ denote the position matrix of element x in Υ^w , which is also symmetric. According to (2.1), the position matrix of $\{\Sigma^u : u \in \mathbb{F}_q, u \neq 0\}$ is

$$M = \begin{pmatrix} P_1(\Sigma^1) & \cdots & P_{q-1}(\Sigma^1) \\ \vdots & \vdots & \vdots \\ P_1(\Sigma^{q-1}) & \cdots & P_{q-1}(\Sigma^{q-1}) \end{pmatrix},$$

and the position matrix of $\{\Upsilon^u : u \in \mathbb{F}_q, u \neq 0\}$ is

$$M^D = \begin{pmatrix} P_1(\Upsilon^1) & \cdots & P_{q-1}(\Upsilon^1) \\ \vdots & \vdots & \vdots \\ P_1(\Upsilon^{q-1}) & \cdots & P_{q-1}(\Upsilon^{q-1}) \end{pmatrix}.$$

To see that M^D is the transpose of M let us show that the column x of M is the row x of M^D ; i.e., let us show that $P_x(\Sigma^w) = P_w(\Upsilon^x)$. Since $(P_x(\Sigma^w))_{ij} = 1$ if and only if $w(i + j) = x$, then $i \neq -j$ because $x \neq 0$, hence $w = x(i + j)^{-1}$ yielding $(P_w(\Upsilon^x))_{ij} = 1$. Therefore the theorem is valid. $\square$

3. Method. Let us call *array of r symbols and n columns* the matrix of order $r \times n$

$$O_{r,n} = \begin{pmatrix} 1 & \cdots & 1 \\ 2 & \cdots & 2 \\ \vdots & \vdots & \vdots \\ r & \cdots & r \end{pmatrix}.$$

When $r = n$ the array is denoted by O_n . It is easy to see that $O_{r,n}$ is a C_4 -free matrix defined on $(r]$. Let $(k)]I_n$ denote the matrix obtained from the identity matrix I_n of order n by replacing each one with the set $(k]$.

The following theorem is a consequence of Theorem 2.3, and gives a general method for constructing the desired k -regular bipartite graphs of girth 6 when $k = q, q + 1$. The position matrix of $O_{q,n}$ will play an important role in what follows, and as in the above sections it will be denoted by $[\mathcal{P}(O_{q,n})]$.

THEOREM 3.1. *Let q be a prime power and $\mathbb{F}_q$ a finite Galois field of q elements. Let us consider the Latin squares $\Sigma^u = [u(i + j)]$ with $u, i, j \in \mathbb{F}_q, u \neq 0$, and the Latin squares Σ_q^u obtained from Σ^u by replacing 0 with q . Then the following assertions hold:*

- (i) *Assume $q \geq 3$. The position matrix of $\{\Sigma_q^1, \Sigma_q^2, \dots, \Sigma_q^{q-1}, (q)]I_q\}$ (given by (2.1)) is the incidence matrix of a q -regular bipartite graph of girth 6 and q^2 vertices in each partite set.*
- (ii) *The position matrix A of $\{\Sigma_q^1, \Sigma_q^2, \dots, \Sigma_q^{q-1}, (q)]I_q, O_q\}$ is the incidence matrix of an affine plane of order q .*

Let j_{q+1} be the row-matrix of order $1 \times (q+1)$ full of ones. The following $(0,1)$ -matrix

$$(3.1) \quad \begin{array}{|c|c|} \hline A & \mathcal{P}(O_{q+1,q})^\top \\ \hline 0 \cdots 0 & j_{q+1} \\ \hline \end{array}$$

is the incidence matrix of a projective plane of order q .

- (iii) Let M be the position matrix of $\{\Sigma^1, \Sigma^2, \dots, \Sigma^{q-2}, (q-1][I_q, O_{q-1,q}\}$. Then $[M | \mathcal{P}(O_{q-1,q})^\top]$ is the incidence matrix of a q -regular bipartite graph of girth 6 and $q^2 - 1$ vertices in each partite set.

Proof. (i) From Theorem 2.3 we know that $\{\Sigma_q^u : u \in \mathbb{F}_q \setminus \{0\}\}$ is C_4 -free, and moreover, by Remark 1, we know that $\{(q)[I_q, \Sigma_q^u]\}$ is C_4 -free. Thus it follows that $\mathcal{F} = \{\Sigma_q^1, \Sigma_q^2, \dots, \Sigma_q^{q-1}, (q)[I_q]\}$ is a C_4 -free family; that is, the position matrix N of $\mathcal{F}$ is free of J_2 . Since each matrix of $\mathcal{F}$ has q symbols in each one of its rows, then there are q ones in each row of N . On the other hand, the position matrix B of $\{\Sigma^1, \Sigma^2, \dots, \Sigma^{q-1}\}$ is a submatrix of N obtained by deleting the q last rows and the q last columns of N . From item (iii) of Theorem 2.3, it follows that the transpose $B^\top$ of B is the position matrix of $\{\Upsilon^1, \Upsilon^2, \dots, \Upsilon^{q-1}\}$. By the definition of the matrices Υ^u it is clear that there are $q-1$ ones in each row of $B^\top$. Therefore B also has $q-1$ ones in each column, and hence N has q ones in each column; thus N is the incidence matrix of a q -regular bipartite graph of q^2 vertices in each partite set of girth 6 (notice that the number of vertices of the constructed graphs is strictly less than the lower bound given in (1.1) for $g = 8$).

(ii) From Theorem 2.3 and by Remark 1, it follows that the position matrix A of $\{\Sigma_q^1, \Sigma_q^2, \dots, \Sigma_q^{q-1}, (q)[I_q, O_q]\}$ is free of J_2 , and clearly, there are q ones in each row of A . By item (i) we know that the position matrix of $\{\Sigma_q^1, \Sigma_q^2, \dots, \Sigma_q^{q-1}, (q)[I_q]\}$ has order $q^2 \times q^2$ and q ones in each row and column. Consequently, the matrix A has order $(q^2 + q) \times q^2$ and has q ones in each row and $q+1$ ones in each column. Hence, considering the rows of A as lines of a partial plane and the columns as points, the matrix A may be seen as the incidence matrix of an affine plane of order q . Thus a projective plane can be obtained by adding $q+1$ points (columns) and one new line (row) containing them. The $q+1$ points correspond to the last $q+1$ columns of (3.1) and the new line is the last row.

(iii) The position matrix M of $\{\Sigma^1, \Sigma^2, \dots, \Sigma^{q-2}, (q-1)[I_q, O_{q-1,q}]\}$ is a submatrix of the block A of (3.1). Indeed, Σ^u is obtained by replacing q with 0 in Σ_q^u . This is equivalent to deleting in A the last q columns. Moreover, now matrix Σ^{q-1} is not considered, which is equivalent to deleting in A the q rows corresponding to the position matrix of Σ^{q-1} . And take $O_{q-1,q}$ instead of O_q is the same as deleting in A the last row. Therefore $[M | \mathcal{P}(O_{q-1,q})^\top]$ is a submatrix of (3.1) and so is free of J_2 . This $(0,1)$ -matrix has $q^2 - 1$ rows and $(q^2 - q) + (q - 1) = q^2 - 1$ columns and, by construction, there are q ones in each row and in each column. Hence, it is the incidence matrix of a q -regular bipartite graph of $q^2 - 1$ vertices in each partite set and girth 6 (notice that the number of vertices of the constructed graphs is strictly less than the lower bound given in (1.1) for $g = 8$). $\square$

To illustrate the method contained in Theorem 3.1, let us consider the first cases $q = 2, 3, 4$. From now on, if there is no confusion a t -set $\{x_1, x_2, \dots, x_t\}$ will be denoted as a sequence of integers $x_1 x_2 \cdots x_t$.

Case $q = 2, 3$. The Latin squares provided by Theorem 2.3 and the incidence matrix of projective plane $PG(2, 2)$ (called Fano plane) and the incidence matrix of projective plane $PG(2, 3)$ given in (3.1) are the following:

Σ_2^1	2	1	0	1	0	1	0	0
	1	2	1	0	0	1	1	0
$(2)]I_2$	12	0	1	0	1	0	0	1
	0	12	0	1	0	1	0	1
O_2	1	1	1	1	0	0	0	0
	2	2	0	0	1	1	0	0
			0	0	0	0	1	1

Σ_3^1	3	1	2	0	1	0	0	0	1	1	0	0	0	0
	1	2	3	1	0	0	0	1	0	0	1	0	0	0
	2	3	1	0	0	1	1	0	0	0	1	0	0	0
Σ_3^2	3	2	1	0	0	1	0	1	0	1	0	0	0	0
	2	1	3	0	1	0	1	0	0	0	1	0	1	0
	1	3	2	1	0	0	0	1	0	1	0	0	1	0
$(3)]I_3$	123	0	0	1	0	0	1	0	0	1	0	0	0	1
	0	123	0	0	1	0	0	1	0	0	1	0	0	1
	0	0	123	0	0	1	0	0	1	0	0	1	0	0
O_3	1	1	1	1	1	1	0	0	0	0	0	0	0	1
	2	2	2	0	0	0	1	1	1	0	0	0	0	1
	3	3	3	0	0	0	0	0	0	1	1	1	0	0
				0	0	0	0	0	0	0	0	0	1	1

Note that the diagonal block of order 6×4 in the matrix on the left-hand side is the incidence matrix of the affine plane $AG(2, 2)$. The diagonal block of order 12×9 in the matrix on the right-hand side is the incidence matrix of the affine plane $AG(2, 3)$.

When $q = 3$, Latin square Σ^1 is obtained by changing 3 for 0 in Σ_3^1 , and $(2)]I_3$ is obtained from $(3)]I_3$ by deleting in the diagonal entries the symbol 3. The incidence matrix $[M | \mathcal{P}(O_{2,3})^\top]$ corresponding to item (iii) of Theorem 3.1 is shown next and gives the incidence matrix of a 3-regular bipartite graph of girth 6 with 8 vertices in each partite set.

Σ^1	0	1	2	0	1	0	0	0	1	1	0
	1	2	0	1	0	0	0	1	0	1	0
	2	0	1	0	0	1	1	0	0	1	0
$(2)]I_3$	12	0	0	1	0	0	1	0	0	0	1
	0	12	0	0	1	0	0	1	0	0	1
	0	0	12	0	0	1	0	0	1	0	1
$O_{2,3}$	1	1	1	1	1	1	0	0	0	0	0
	2	2	2	0	0	0	1	1	1	0	0

Case $q = 4$. The Latin squares provided by Theorem 2.3 and the incidence matrix of projective plane $PG(2, 4)$ given in (3.1) are the following:

Σ_4^1	4	1	2	3	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0
	1	4	3	2	1	0	0	0	0	0	1	0	0	1	0	1	0	0	0
	2	3	4	1	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0
	3	2	1	4	0	0	1	0	0	1	0	0	1	0	0	0	0	1	0
Σ_4^2	4	2	3	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0
	2	4	1	3	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0
	3	1	4	2	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
	1	3	2	4	1	0	0	0	0	0	1	0	1	0	0	0	0	1	0
Σ_4^3	4	3	1	2	0	0	1	0	0	0	1	0	1	0	0	1	0	0	0
	3	4	2	1	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0
	1	2	4	3	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0
	2	1	3	4	0	1	0	0	1	0	0	0	0	1	0	0	0	1	0
$(4)]I_4$	1234	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	1	0
	0	1234	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
	0	0	1234	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1	0
	0	0	0	1234	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1
O_4	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1
	2	2	2	2	0	0	0	0	1	1	1	1	0	0	0	0	0	0	1
	3	3	3	3	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0
	4	4	4	4	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

The diagonal block of order 20×16 is the incidence matrix of the affine plane $AG(2, 4)$.

Σ^1	0	1	2	3	0	1	0	0	0	0	1	0	0	0	0	0	1	1	0	0
	1	0	3	2	1	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0
	2	3	0	1		0	0	0	1	1	0	0	0	0	1	0	0	1	0	0
	3	2	1	0		0	0	1	0	0	1	0	0	0	1	0	0	0	1	0
Σ^2	0	2	3	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	1	0
	2	0	1	3		0	0	1	0	1	0	0	0	0	0	0	1	0	1	0
	3	1	0	2		0	1	0	0	0	0	0	1	1	0	0	0	0	1	0
	1	3	2	0		1	0	0	0	0	0	1	0	0	1	0	0	0	1	0
$(3)]I_4$	123	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1
	0	123	0	0		0	1	0	0	0	0	1	0	0	0	1	0	0	0	1
	0	0	123	0		0	0	1	0	0	0	0	1	0	0	1	0	0	0	1
	0	0	0	123		0	0	0	0	1	0	0	0	1	0	0	0	1	0	0
$O_{3,4}$	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
	2	2	2	2	2	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0
	3	3	3	3	3	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0

$$\Sigma_2^1 = \begin{bmatrix} 0 & 1 & 2 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 & 1 \\ 3 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 & 0 \end{bmatrix} \quad \Sigma_2^2 = \begin{bmatrix} 0 & 2 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 \\ 1 & 3 & 0 & 2 & 0 \\ 3 & 0 & 2 & 0 & 1 \end{bmatrix} \quad \Sigma_2^{2*} = \begin{bmatrix} 1 & 3 & 0 & 2 & 0 \\ 3 & 0 & 2 & 0 & 1 \end{bmatrix} \quad O_{2,5}^* = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 2 & 2 & 0 & 0 & 2 \end{bmatrix}$$

$$(3)_I I_5^* = \begin{bmatrix} 123 & 0 & 0 & 0 & 0 \\ 0 & 123 & 0 & 0 & 0 \\ 0 & 0 & 123 & 0 & 0 \\ 0 & 0 & 3 & 12 & 0 \\ 0 & 0 & 3 & 0 & 12 \end{bmatrix} \quad O_* = \boxed{1 \ 1 \ 0 \ 0 \ 1}$$

Proof. Note that $[M \mid \mathcal{P}(O^*)^\top]$ is a matrix free of J_2 because it is a submatrix of (3.1) by changing some ones for zeros. Besides, M has order $(q^2 - rq - 1) \times q(q - r)$,

and the last $q - r - 1$ columns of M are full of zeros, because in the last $q - r - 1$ columns of each Σ_r^u we have replaced the entry $q - r$ with zero. Thus, we consider the matrix M' once eliminating from M these columns full of zeros. Hence M' has order $(q^2 - rq - 1) \times (q(q - r) - (q - r - 1))$. Furthermore, notice that each of the $q^2 - rq - q + r + 1$ columns of M' has $q - r$ ones according to the definition of M' . On the other hand, the rows of the position matrix corresponding to $\{\Sigma_{r*}^{q-r-1}, (q - r)]I_q^*, O_{q-r-1,q}^*\}$ have $q - r$ ones each because of the definition of these matrices. Therefore, since the value $q - r$ remains in the first r columns of each Σ_r^u , $u = 1, \dots, q - r - 2$, then there are $(q - r - 2)(q - r)$ rows in M' having $q - r - 1$ ones, and the remaining rows have $q - r$ ones each. By joining $\mathcal{P}(O_*)^\top$ to M' we put $q - r - 2$ more columns, that due to the definition of O_* add an additional one just in the rows which had $q - r - 1$ ones in M' . Therefore $[M' | \mathcal{P}(O_*)^\top]$ is the incidence matrix of a $(q - r)$ -regular bipartite graph with $q^2 - rq - 1$ vertices in each partite set and girth at least 6. Since the number of vertices of the constructed graphs is strictly less than the lower bound given in (1.1) for $q = 8$, then the girth is 6. $\square$

To illustrate the method of Theorem 3.2 let us consider the first cases $q = 4, 5$.

Case $q = 4$ and $r = 1$. The matrices of Theorem 3.2 and the incidence matrix of a 3-regular bipartite graph of girth 6 with 11 vertices in each partite set are the following:

Σ_1^1	0	1	2	0	0	1	0	0	0	0	0	1	0	0	0	0	1
	1	0	0	2	1	0	0	0	0	0	0	1	0	0	0	0	1
	2	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	1
	3	2	1	0	0	0	1	0	0	1	0	0	1	0	0	0	0
Σ_{1*}^2	3	1	0	2	0	1	0	0	0	0	0	1	1	0	0	0	0
$(3)]I_4^*$	123	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
	0	123	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
	0	3	12	0	0	0	1	0	0	0	1	0	0	0	1	0	0
	0	3	0	12	0	0	0	1	0	0	0	1	0	1	0	0	0
$O_{2,4}^*$	1	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0
	2	2	2	0	0	0	0	0	1	1	1	0	0	0	0	0	0

The matrix M' of Theorem 3.2 is obtained by deleting columns 11 and 12 which are full of zeros.

Case $q = 5$ and $r = 2$. The matrices of Theorem 3.2 and the incidence matrix of a 3-regular bipartite graph of girth 6 with 14 vertices in each partite set are the following:

Σ_2^1	0	1	2	0	0	0	1	0	0	0	0	0	0	0	0	1
	1	2	0	0	0	0	1	0	0	0	0	0	1	0	0	0
	2	3	0	0	1	0	0	0	0	1	1	0	0	0	0	0
	3	0	0	1	2	0	0	0	1	0	0	0	0	1	0	0
	0	0	1	2	0	0	0	1	0	0	0	1	0	0	0	1
Σ_{2*}^2	1	3	0	2	0	0	1	0	0	0	0	0	1	0	0	0
	3	0	2	0	1	0	0	0	1	0	0	1	0	0	0	0
$(3)]I_5^*$	123	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0
	0	123	0	0	0	0	0	1	0	0	0	0	1	0	0	0
	0	0	123	0	0	0	0	0	1	0	0	0	0	1	0	0
	0	0	3	12	0	0	0	0	1	0	0	0	0	1	0	0
	0	0	3	0	12	0	0	0	0	1	0	0	0	1	0	0
$O_{2,5}^*$	0	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0
	2	2	0	0	2	0	0	0	0	0	1	1	0	0	1	0

The matrix M' of Theorem 3.2 is obtained by deleting columns 14 and 15 which are full of zeros.

4. Improvement for $(q-1)$ -regular bipartite graphs of girth 6. In this section we improve Theorem 3.2 for $r = 1$, finding incidence matrices of $(q-1)$ -regular bipartite graphs of girth 6 with $q^2 - q - 2$ vertices in each partite set. To this end we need to introduce the following definition.

DEFINITION 4.1. Two C_4 -free matrices A^1 and A^2 of the same number of columns whose elements are 1-sets of $[[n]]$ are said to be quasi row-disjoint if and only if the Cartesian product of any two rows $(A^1)_i, (A^2)_h$ contains at most one pair $(x, x) \in (A^1)_i \times (A^2)_h$ with $x \neq 0$.

An example of two quasi row-disjoint Latin squares of order 3×3 , one with entries in $\{3\}$ and the other with entries in $\{3\} \setminus \{1\}$, is the following.

$$(4.1) \quad L^1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} \quad L^2 = \begin{pmatrix} 0 & 3 & 2 \\ 3 & 2 & 0 \\ 2 & 0 & 3 \end{pmatrix}.$$

We can check by superimposing any two rows of L^1 and L^2 that there is at most one resulting pair (x, x) with $x \neq 0$. Then L^1 and L^2 are quasi row-disjoint.

THEOREM 4.2. Let Q^1 and Q^2 be two C_4 -free matrices of the same number of columns and whose elements are 1-sets of $[[n]]$. Then Q^1 and Q^2 are quasi row-disjoint if and only if the family $\{Q^1, Q^2\}$ is C_4 -free.

Proof. Let $(Q^\ell)_i$ denote the row i of Q^ℓ , $\ell = 1, 2$. Assume that $(x, x) \in (Q^1)_i \times (Q^2)_h$ with $x \in [[n]]$, $x \neq 0$. Then $x = (Q^1)_{ij} = (Q^2)_{hj}$ or, equivalently $(P_x(Q^1))_{ij} = (P_x(Q^2))_{hj} = 1$, for some column j , where $P_x(Q^\ell)$ is the position matrix of x in Q^ℓ .

If Q^1 and Q^2 are not quasi row-disjoint, then there exists $(y, y) \in (Q^1)_i \times (Q^2)_h$ with $y \in [[n]] \setminus \{x, 0\}$. This implies that $(P_y(Q^1))_{ik} = (P_y(Q^2))_{hk} = 1$ with $k \neq j$, and therefore the position matrix of $\{Q^1, Q^2\}$ contains a submatrix J_2 ; i.e., $\{Q^1, Q^2\}$ is not C_4 -free.

Conversely, if the position matrix of $\{Q^1, Q^2\}$ contains J_2 as a submatrix, then there exists $y \in [[n]] \setminus \{x, 0\}$ such that $(P_y(Q^1))_{ik} = (P_y(Q^2))_{hk} = 1$, or in other words, $y = (Q^1)_{ik} = (Q^2)_{hk}$, $k \neq j$. Hence Q^1, Q^2 are not quasi row-disjoint. $\square$

As a consequence of Theorems 2.3, Remark 1, and Theorem 4.2, the following corollary is now clear.

COROLLARY 4.3. Let q be a prime power and $\mathbb{F}_q$ the Galois field of order q . The family $\{\Sigma_q^u : u \in \mathbb{F}_q\} \cup \{(q)I_q, O_q\}$ is a set of $q+1$ quasi row-disjoint matrices of order q with entries in $[[q]]$.

Next, we construct a family of mutually quasi row-disjoint matrices of order $q-1$ by adding to each entry of the product table of the cyclic group $(\mathbb{F}_q \setminus \{0\}, \cdot)$ an element of the Galois field $\mathbb{F}_q$.

THEOREM 4.4. Let $q \geq 4$ be a prime power and $\mathbb{F}_q$ the Galois field of order q . Then the following assertions are true.

- (i) For every $u \in \mathbb{F}_q$ define the matrix Π^u by $(\Pi^u)_{ij} = u + ij$, $i, j \in \mathbb{F}_q, i, j \neq 0$. Then the family $\{\Pi^u : u \in \mathbb{F}_q\}$ is C_4 -free.
- (ii) For every $u \in \mathbb{F}_q$ define the matrix D^u by $(D^u)_{ij} = u - ij$, $i, j \in \mathbb{F}_q, i, j \neq 0$. Then the family $\{D^u : u \in \mathbb{F}_q\}$ is C_4 -free.
- (iii) If N is the position matrix of $\{\Pi^u : u \in \mathbb{F}_q, u \neq 0\}$, then the transpose of N is the position matrix of $\{D^u : u \in \mathbb{F}_q, u \neq 0\}$.
- (iv) Let N be the position matrix of $\{\Pi^3, \dots, \Pi^{q-1}, \Pi^0, (q-1)I_{q-1}, O_{q-1}^2\}$, where O_{q-1}^2 is the matrix obtained from O_{q-1} by deleting the first two rows. Then

the $(0, 1)$ -matrix $[N|\mathcal{P}(O_{q-3,q-1})^\top]$ is the incidence matrix of a $(q-1)$ -regular bipartite graph of girth 6 and $q^2 - q - 2$ vertices in each partite set.

Proof. (i) Clearly $\{\Pi^u = [ij + u] : u, i, j \in \mathbb{F}_q, i, j \neq 0\}$ is a set of q Latin squares of order $q - 1$, each one with entries from $[[q - 1]] \setminus \{u\}$, because $ij + u = u$ implies $i = 0$ or $j = 0$ which is impossible by definition of Π^u . Let us see that they are mutually quasi row-disjoint. Suppose $(\Pi^w)_{ij} = (\Pi^u)_{kj}$ for some $u, w \in \mathbb{F}_q$ with $u \neq w$ and $i, j, k \neq 0$. Therefore

$$ij + w = kj + u.$$

If the Latin squares are not quasi row-disjoint, then there exists $h \neq j$, $h \neq 0$, such that $(\Pi^w)_{ih} = (\Pi^u)_{kh}$, that is,

$$ih + w = kh + u.$$

Hence $i(j - h) = k(j - h)$. As $j \neq h$, then $j - h$ has an inverse respect to the product in the cyclic group $\mathbb{F}_q \setminus \{0\}$ yielding $i = k$, thus $u = w$ which is a contradiction. Therefore Π^u and Π^w are quasi row-disjoint, and hence by Theorem 4.2 we conclude that $\{\Pi^u : u \in \mathbb{F}_q\}$ is C_4 -free.

(ii) The proof of this item is the same as the proof of item (i).

(iii) Reasoning as in the proof of item (iii) of Theorem 2.3, it is enough to prove that $P_k(\Pi^w) = P_w(D^k)$ for all $k, w \in \mathbb{F}_q$, $k \neq 0$. Since $(P_k(\Pi^w))_{ij} = 1$ if and only if $w + ij = k$, hence $w = k - ij$ if and only if $(P_w(D^k))_{ij} = 1$. Then (iii) is valid.

(iv) By point (i) we know that $\{\Pi^u : u \in \mathbb{F}_q\}$ is C_4 -free. It is not difficult to see that O_{q-1} is quasi row-disjoint with each Π^u , and using Remark 1 we obtain that $\{\Pi^u : u \in \mathbb{F}_q\} \cup \{(q-1)I_{q-1}, O_{q-1}\}$ is a C_4 -free family of $q + 2$ mutually quasi-row-disjoint C_4 -free matrices of order $(q-1) \times (q-1)$ whose entries are subsets of $[[q-1]]$. Therefore the position matrix N of $\{\Pi^3, \dots, \Pi^{q-1}, \Pi^0, (q-1)I_{q-1}, O_{q-1}^2\}$ is free of J_2 .

Moreover, we have for all $u \in \mathbb{F}_q$ that the entries of Π^u belong to $[[q-1]] \setminus \{u\}$. Hence the rows of the position matrix of each Π^u have $q-2$ ones for $u \neq 0$, and $q-1$ ones for $u = 0$. Let A be the position matrix of $\{\Pi^1, \dots, \Pi^{q-1}\}$. From item (iii) it follows that $A^\top$ is the position matrix of $\{D^1, \dots, D^{q-1}\}$. This implies that the columns of matrix A have $q-2$ ones. Then deleting the matrices Π^1 and Π^2 results in a submatrix A' of A which has $q-3$ ones in $2(q-1)$ columns and $q-4$ ones in the $(q-3)(q-1)$ remaining columns. Let B be the position matrix of $\{\Pi^0, (q-1)I_{q-1}\}$, which is readily seen to have 2 ones in each column and $q-1$ ones in each row. Therefore joining to A' the rows of B we obtain a matrix C which has $q-1$ ones in $2(q-1)$ columns and $q-2$ ones in the $(q-3)(q-1)$ remaining columns. Finally, adding to C the rows of the position matrix of O_{q-1}^2 we obtain the matrix N having $q-1$ ones in each column. Thus, $[N|\mathcal{P}(O_{q-3,q-1})^\top]$ is the incidence matrix of a $(q-1)$ -regular bipartite graph of girth 6 and $q^2 - q - 2$ vertices in each partite set. $\square$

To illustrate the method contained in Theorem 4.4, let us consider the case $q = 4$.

Case $q = 4$. The C_4 -free quasi row-disjoint matrices provided by Theorem 4.4 and the incidence matrix of a 3-regular bipartite graph of girth 6 with 10 vertices in each partite set are the following:

TABLE 5.1
Order of some minimal regular graphs of girth 6.

k	known $(k, 6)$ -graph	Method
7	90 [16]	96
11	240 [1, 21]	240
13	336 [1]	336
15	462 [1]	476
16	504 [1]	510
q prime power		
$q + 1$	$2(q^2 + q + 1)$ Projective plane	$2(q^2 + q + 1)$
q	$2(q^2 - 1)$ [1]	$2(q^2 - 1)$
$q - 1$	$2(q^2 - q - 1)$ [1]	$2(q^2 - q - 2)$
$q - r, r \geq 2$	$2(q^2 - rq)$ [1, 2]	$2(q^2 - rq - 1)$

Π^3	2	1	0	0	1	0	1	0	0	0	0	1
	1	0	2	1	0	0	0	0	1	0	0	1
	0	2	1	0	0	1	0	1	0	0	0	1
Π^0	1	2	3	1	0	0	0	1	0	0	0	1
	2	3	1	0	0	1	1	0	0	0	1	0
	3	1	2	0	1	0	0	0	1	1	0	0
$(3)]I_3$	123	0	0	1	0	0	1	0	0	1	0	0
	0	123	0	0	1	0	0	1	0	0	1	0
	0	0	123	0	0	1	0	0	1	0	0	1
O_3^2	3	3	3	0	0	0	0	0	0	1	1	1

5. Conclusions. We have presented a method providing the incidence matrices of both affine planes and projective planes of order q , and q -regular bipartite graphs with $q^2 - 1$ vertices in each partite set when q is a prime power. From these incidence matrices, we obtain the incidence matrices of $(q - r)$ -regular bipartite graphs of girth 6 having $q^2 - rq - 1$ vertices in each partite set. Furthermore, using quasi row-disjoint Latin squares we improve this result for $r = 1$ finding $(q - 1)$ -regular bipartite graphs of girth 6 with $q^2 - q - 2$ vertices in each partite set. Table 5.1 shows a comparison between the number of vertices of some known smallest graphs [1, 19] and the graphs provided by the method proved in this work. We consider only degrees k in the interval $[q - r, q + 1]$ in which the only prime power is q , and $q - r$ is different from a prime power plus one.

Acknowledgments. I would like to express my thanks to the referees for their helpful comments and suggestions. Overall I want to thank to Kevin Phelps for suggesting to me that the Latin squares obtained by multiplying each entry of the addition table of the Galois field by an element distinct from zero are the $(1, 3)$ conjugates of Latin squares of a complete set of MOLS.

REFERENCES

- [1] M. ABREU, M. FUNK, D. LABBATE, AND V. NAPOLITANO, *On (minimal) regular graphs of girth 6*, Australas. J. Combin., 35 (2006), pp. 119–132.
- [2] G. ARAUJO, D. GONZÁLEZ-MORENO, J. J. MONTELLANO, AND O. SERRA, *Upper bounds and connectivity of cages*, Australas. J. Combin., 38 (2007), pp. 221–228.
- [3] E. BANNAI AND T. ITO, *On finite Moore graphs*, J. Fac. Sci., Univ. Tokio, Sect. I A Math., 20 (1973), pp. 191–208.
- [4] L. M. BATTEN, *Combinatorics of Finite geometries*, Cambridge University Press, Cambridge, UK, 1997.
- [5] C. T. BENSON, *Minimal regular graphs of girth eight and twelve*, Canad. J. Math., 18 (1966), pp. 1091–1094.
- [6] N. BIGGS, *Algebraic Graph Theory*, Cambridge University Press, New York, 1996.

- [7] N. BIGGS, *Construction for cubic graphs with large girth*, Electron. J. Combin., 5 (1998), p. A1.
- [8] G. BRINKMANN, B. D. MCKAY, AND C. SAAGER, *The smallest cubic graphs of girth nine*, Combin. Probab. Comput., 5 (1995), pp. 1–13.
- [9] P. ERDÖS AND H. SACHS, *Reguläre graphen gegebener Tailenweite mit minimaler Knotenzahl*, Wiss. Z. Uni. Halle (Math. Nat.), 12 (1963), pp. 251–257.
- [10] G. EXOO, *A simple method for constructing small cubic graphs of girths 14, 15 and 16*, Electron. J. Combin., 3 (1996).
- [11] W. FEIT AND G. HIGMAN, *The nonexistence of certain generalized polygons*, J. Algebra, 1 (1964), pp. 114–131.
- [12] C. GODSIL AND G. ROYLE, *Algebraic Graph Theory*, Springer, NY, 2000.
- [13] D. A. HOLTON AND J. SHEEHAN, *The Petersen Graph, Chapter 6: Cages*, Cambridge University, London, 1993.
- [14] J. H. VAN LINT AND R. M. WILSON, *A Course in Combinatorics*, Cambridge University Press, London, 1994.
- [15] M. O'KEEFE AND P. K. WONG, *A smallest graph of girth 10 and valency 3*, J. Combin. Theory (B), 29 (1980), pp. 91–105.
- [16] M. O'KEEFE AND P. K. WONG, *The smallest graph of girth 6 and valency 7*, J. Graph Theory, 5 (1981), pp. 79–85.
- [17] M. MERINGER, *Fast generation of regular graphs and construction of cages*, J. Graph Theory, 30 (1999), pp. 137–146.
- [18] T. PISANSKI, M. BOBEN, D. MARUSIC, A. ORBANIC, AND A. GRAOVAC, *The 10-cages and derived configurations*, Discrete Math., 275 (2004), pp. 265–276.
- [19] G. ROYLE, <http://www.cs.uwa.edu.au/~gordon/data.html>.
- [20] W. T. TUTTE, *A family of cubical graphs*, Proc. Cambridge Philos. Soc., (1947), pp. 459–474.
- [21] P. K. WONG, *A regular graph of girth 6 and valency 11*, Int. J. Math. Sci., 9 (1986), pp. 561–565.
- [22] P. K. WONG, *Cages—A survey*, J. Graph Theory, 6 (1982), pp. 1–22.

NEARLY OPTIMAL VISIBILITY REPRESENTATIONS OF PLANE GRAPHS*

XIN HE[†] AND HUAMING ZHANG[‡]

Abstract. The *visibility representation* (VR for short) is a classical representation of plane graphs. The VR has various applications and has been extensively studied in the literature. One of the main focuses of the study is to minimize the size of the VR. It is known that there exists a plane graph G with n vertices, where any VR of G requires a size at least $(\lfloor \frac{2n}{3} \rfloor) \times (\lfloor \frac{4n}{3} \rfloor - 3)$. For upper bounds, it is known that every plane graph has a VR with height at most $\lfloor \frac{4n-1}{5} \rfloor$, and a VR with width at most $\lfloor \frac{13n-24}{9} \rfloor$. In this paper, we prove that every plane graph has a VR with height at most $\frac{2n}{3} + 2\lceil \sqrt{n/2} \rceil$, and a VR with width at most $\frac{4n}{3} + 2\lceil \sqrt{n} \rceil$. These representations are nearly optimal in the sense that they differ from the lower bounds only by a lower order additive term. Both representations can be constructed in linear time. Our presentations use Schnyder's realizer to construct the *st*-orientations of plane graphs with special properties. As the *st*-orientation is a very useful concept in other applications, this result may be of independent interest.

Key words. plane graph, plane triangulation, visibility representation, *st*-orientation

AMS subject classifications. 05C62, 05C85, 05C90, 68R10

DOI. 10.1137/060651811

1. Introduction. Drawing plane graphs has emerged as a fast growing research area in recent years (see [7] for a survey). Several important concepts have been introduced for solving graph drawing problems, including the *canonical orderings* and *canonical ordering trees* of 3-connected plane graphs [10, 11, 12]; the *orderly spanning trees*, which generalize canonical ordering trees [5, 4, 6, 16]; the *well-orderly trees*, which are canonical ordering trees with some special properties [2], and *Schnyder's realizers* [23, 24]. These concepts and techniques have been successfully used in graph drawing, graph encoding, and graph generation [3, 1, 9, 18].

A *visibility representation* (VR for short) of a plane graph G is a drawing of G , where the vertices of G are represented by nonoverlapping horizontal line segments (called *vertex segments*), and each edge of G is represented by a vertical line segment touching the vertex segments of its end vertices. Figure 2.1 shows a VR of a plane graph G . The problem of computing a compact VR is important not only in algorithmic graph theory, but also in practical applications. A simple linear time VR algorithm was given in [22, 25] for 2-connected plane graphs. It uses an *st*-orientation of G and the corresponding *st*-orientation of its *st*-dual G^* to construct a VR. Using this approach, the height of the VR is bounded by $(n - 1)$ and the width of the VR is bounded by $(2n - 5)$ [22, 25].

As in many other graph drawing problems, one of the main concerns in the VR research is to minimize the size of the representation. For the lower bounds, it was shown in [27] that there exists a plane graph G with n vertices where any VR of G

*Received by the editors February 9, 2006; accepted for publication (in revised form) May 23, 2008; published electronically September 11, 2008.

<http://www.siam.org/journals/sidma/22-4/65181.html>

[†]Department of Computer Science and Engineering, State University of New York at Buffalo, Buffalo, NY 14260 (xinhe@cse.buffalo.edu). This author's research was supported in part by NSF grants CCR-0309953 and CCR-0635104.

[‡]Department of Computer Science, The University of Alabama in Huntsville, Huntsville, AL 35899 (hzhang@cs.uah.edu). This author's research was supported in part by NSF grant CCF-0728830.

requires a size at least $(\lfloor \frac{2n}{3} \rfloor) \times (\lfloor \frac{4n}{3} \rfloor - 3)$. Some work has been done to reduce the height and the width of the VR by carefully constructing special *st*-orientations. The following table summarizes related previous results.

Plane graph		4-connected plane graph	
Width	Height	Width	Height
$\leq (2n - 5)$ [22, 25]	$\leq (n - 1)$ [22, 25]		
$\leq \lfloor \frac{3n-6}{2} \rfloor$ [13]			
$\leq \lfloor \frac{22n-42}{15} \rfloor$ [17]		$\leq (n - 1)$ [14]	
	$\leq \lfloor \frac{5n}{6} \rfloor$ [27]		
$\leq \lfloor \frac{13n-24}{9} \rfloor$ [28]	$\leq \lfloor \frac{4n-1}{5} \rfloor$ [29]		$\leq \lceil \frac{3n}{4} \rceil$ [26]

All these results concentrated on one dimension of the VR (either the width or the height), with the other dimension bounded by the trivial upper bound (namely, $n - 1$ for the height and $2n - 5$ for the width). In [20, 21], heuristic algorithms were developed aiming at reducing the height and the width of VRs simultaneously.

In this paper, we prove that every plane graph of n vertices has a VR with height at most $\frac{2n}{3} + 2\lceil \sqrt{n/2} \rceil$ (and width at most $2n - 5$), and another VR with width at most $\frac{4n}{3} + 2\lceil \sqrt{n} \rceil$ (and height at most $n - 1$). These representations are nearly optimal in the sense that they differ from the lower bounds only by a lower order additive term. Both representations can be constructed in linear time. Our VRs use Schnyder's realizer to construct *st*-orientations of plane graphs with special properties. As the *st*-orientation is a very useful concept in other applications, this result may be of independent interest.

The present paper is organized as follows. Section 2 introduces preliminaries. Section 3 presents the construction of a VR with nearly optimal height. Section 4 presents the construction of a VR with nearly optimal width. Section 5 concludes the paper.

2. Preliminaries. In this section, we give definitions and preliminary results. Definitions not mentioned here are standard.

A *planar graph* is a graph G such that the vertices of G can be drawn in the plane and the edges of G can be drawn as nonintersecting curves. Such a drawing is called an *embedding*. The embedding divides the plane into a number of connected regions. Each region is called a *face*. The unbounded face is the *external face*. The other faces are *internal faces*. A *plane graph* is a planar graph with a fixed embedding. A *plane triangulation* is a plane graph where every face is a triangle (including the external face). We abbreviate the words “counterclockwise” and “clockwise” as ccw and cw, respectively.

When discussing VRs, we assume the input graph G is a plane triangulation. (If not, we get a plane triangulation G' by adding dummy edges into G . After constructing a VR for G' , we can get a VR of G by deleting the vertical line segments for the dummy edges.)

A *numbering* $\mathcal{O}$ of a set $S = \{a_1, \dots, a_k\}$ is a one-to-one mapping between S and the set $\{1, 2, \dots, k\}$. We write $\mathcal{O} = \langle a_{i_1}, a_{i_2}, \dots, a_{i_k} \rangle$ to indicate $\mathcal{O}(a_{i_1}) = 1$, $\mathcal{O}(a_{i_2}) = 2$, etc. A set S with a numbering written this way is called an *ordered list*. For two elements a_i and a_j , if a_i is assigned a smaller number than a_j in $\mathcal{O}$, we write $a_i \prec_{\mathcal{O}} a_j$. Let S_1 and S_2 be two disjoint sets. If $\mathcal{O}_1$ is a numbering of S_1 and $\mathcal{O}_2$ is a numbering of S_2 , their concatenation, written as $\mathcal{O} = \langle \mathcal{O}_1, \mathcal{O}_2 \rangle$, is the numbering of $S_1 \cup S_2$, where $\mathcal{O}(x) = \mathcal{O}_1(x)$ for all $x \in S_1$ and $\mathcal{O}(y) = \mathcal{O}_2(y) + |S_1|$ for all $y \in S_2$.

G is called a *directed graph* (digraph for short) if each edge of G is assigned a direction. An *orientation* of an (undirected) graph G is a digraph obtained from G by assigning a direction to each edge of G . We will use G to denote both the resulting digraph and the underlying undirected graph unless otherwise specified. (Its meaning will be clear from the context.)

Let $G = (V, E)$ be an undirected graph. A numbering $\mathcal{O}$ of V induces an orientation of G as follows: each edge of G is directed from its lower numbered end vertex to its higher numbered end vertex. The resulting digraph is called the *orientation derived from $\mathcal{O}$* which, obviously, is an acyclic digraph. We use $\text{length}_G(\mathcal{O})$ (or simply $\text{length}(\mathcal{O})$ if G is clear from the context) to denote the length of the longest path in the orientation of G derived from $\mathcal{O}$. (The length of a path is the number of edges in it.)

For a 2-connected plane graph G and an external edge (s, t) , an orientation of G is called an *st-orientation* if the resulting digraph is acyclic with s as the only source and t as the only sink. Such a digraph is also called an *st-graph*. Lempel, Even, and Cederbaum [15] showed that for every 2-connected plane graph G and an external edge (s, t) , there exists an *st-orientation*. For more properties of *st-orientation* and *st-graph*, we refer readers to [19].

Let G be a 2-connected plane graph and (s, t) an external edge. An *st-numbering* of G is a one-to-one mapping $\xi : V \rightarrow \{1, 2, \dots, n\}$, such that $\xi(s) = 1$, $\xi(t) = n$, and each vertex $v \neq s, t$ has two neighbors u, w with $\xi(u) < \xi(v) < \xi(w)$, where u (w , respectively) is called a *smaller neighbor* (*bigger neighbor*, respectively) of v . Given an *st-numbering* ξ of G , the orientation of G derived from ξ is obviously an *st-orientation* of G . On the other hand, if $G = (V, E)$ has an *st-orientation* $\mathcal{O}$, we can define a one-to-one mapping $\xi : V \rightarrow \{1, \dots, n\}$ by topological sort. It is easy to see that ξ is an *st-numbering* and the orientation derived from ξ is $\mathcal{O}$. From now on, we will interchangeably use the term “an *st-numbering*” of G and the term “an *st-orientation*” of G , where each edge of G is directed accordingly.

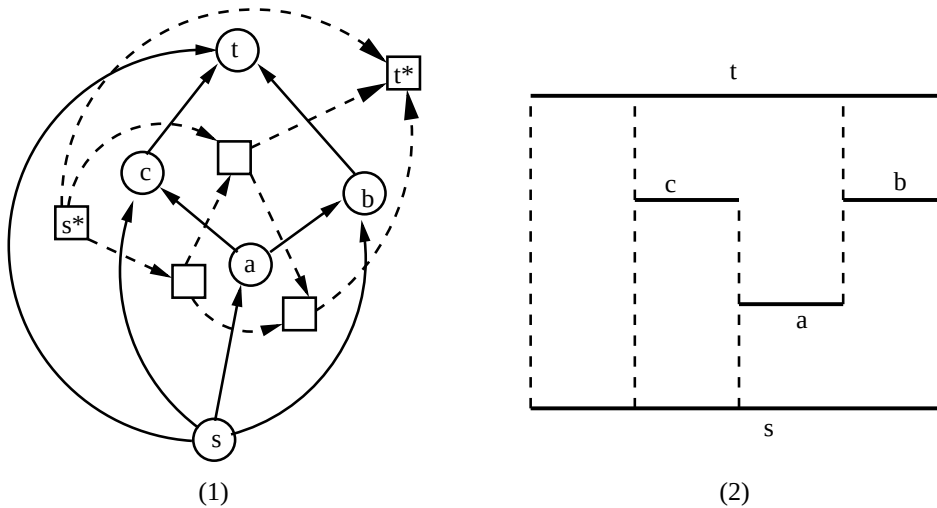
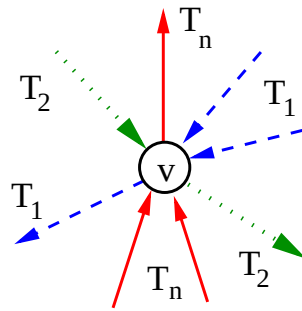
DEFINITION 2.1. Let G be a plane graph with an *st-orientation* $\mathcal{O}$, where (s, t) is an external edge drawn at the left on the external face of G . The *st-dual graph* G^* of G and the dual orientation $\mathcal{O}^*$ of $\mathcal{O}$ is defined as follows:

- Each face f of G corresponds to a node f^* of G^* . In particular, the unique internal face adjacent to the edge (s, t) corresponds to a node s^* in G^* ; the external face corresponds to a node t^* in G^* .
- For each edge $e \neq (s, t)$ of G separating a face f_1 on its left and a face f_2 on its right, there is a dual edge e^* in G^* from f_1^* to f_2^* .
- The dual edge of the external edge (s, t) is directed from s^* to t^* .

Figure 2.1(1) shows an *st-graph* G and its *st-dual graph* G^* (where circles and solid lines denote the vertices and the edges of G ; squares and dashed lines denote the nodes and the edges of G^*). It is well known that the *st-dual graph* G^* defined above is an *st-graph* with source s^* and sink t^* . The correspondence between an *st-orientation* $\mathcal{O}$ of G and the dual *st-orientation* $\mathcal{O}^*$ is a one-to-one correspondence. The following theorem was given in [22, 25].

THEOREM 2.2. Let G be a 2-connected plane graph with an *st-orientation* $\mathcal{O}$. Let $\mathcal{O}^*$ be the dual *st-orientation* of the *st-dual graph* G^* . A VR of G can be obtained from $\mathcal{O}$ in linear time. The height of the VR is $\text{length}(\mathcal{O}) \leq n - 1$. The width of the VR is $\text{length}(\mathcal{O}^*) \leq 2n - 5$.

Figure 2.1(2) shows a VR of the graph G shown in Figure 2.1(1). The width of the VR is $\text{length}(\mathcal{O}^*) = 4$. The height of the VR is $\text{length}(\mathcal{O}) = 3$.

FIG. 2.1. (1) An st -graph G and its st -dual graph G^* ; (2) a VR of G .FIG. 2.2. Edge pattern around an internal vertex v .

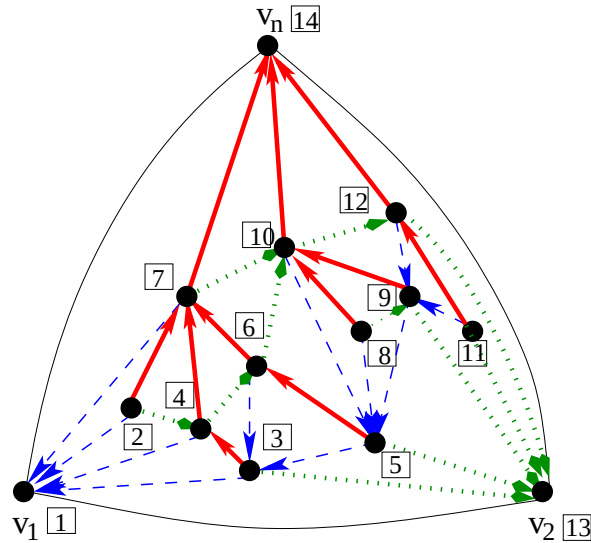
The following concept was introduced in [23] and is of central importance in our VR construction.

DEFINITION 2.3. Let G be a plane triangulation of n vertices with three external vertices v_1, v_2, v_n in ccw order. A realizer $\mathcal{R} = \{T_1, T_2, T_n\}$ of G is a partition of its internal edges into three sets T_1, T_2, T_n of directed edges such that the following hold:

- For each $i \in \{1, 2, n\}$, the internal edges incident to v_i are in T_i and directed towards v_i .
- For each internal vertex v of G , v has exactly one edge leaving v in each of T_1, T_2, T_n . The ccw order of the edges incident to v is leaving in T_1 , entering in T_n , leaving in T_2 , entering in T_1 , leaving in T_n , and entering in T_2 . Each entering block may be empty. (See Figure 2.2.)

Figure 2.3 shows a realizer of a plane triangulation G . The dashed lines (dotted lines and solid lines, respectively) are the edges in T_1 (T_2 and T_n , respectively).

In [23], Schnyder showed that every plane triangulation G has a realizer which can be constructed in linear time. It was also shown that each set T_i of a realizer is a tree rooted at v_i . For each T_i , we denote by $\bar{T}_i$ the tree composed of T_i augmented with the two edges of the external face incident to the root v_i of T_i . Obviously $\bar{T}_i$ is a spanning tree of G . For example, in Figure 2.3, $\bar{T}_n$ is T_n (the tree in solid lines)

FIG. 2.3. A plane triangulation G and a realizer $\mathcal{R}$ of G .

augmented with the edges (v_1, v_n) and (v_2, v_n) , which is a spanning tree of G .

For each internal vertex v of G and $i \in \{1, 2, n\}$, $p_i(v)$ denotes the path in $\bar{T}_i$ from v to the root v_i of $\bar{T}_i$. It was shown in [23] that $p_1(v)$, $p_2(v)$, and $p_n(v)$ have only the vertex v in common, and for two vertices $u \neq v$ and two indices $i \neq j$, $p_i(u)$ and $p_j(v)$ can have at most one common vertex. The following property was shown in [2, 5].

Property 1. Let v be an internal vertex of G .

1. All ancestors of v in $\bar{T}_1$ ($\bar{T}_2$, $\bar{T}_n$, respectively) constitute a nonempty set and they appear before v in the ccw postordering of the vertices of G with respect to $\bar{T}_n$ ($\bar{T}_1$, $\bar{T}_2$, respectively).
2. All ancestors of v in $\bar{T}_2$ ($\bar{T}_n$, $\bar{T}_1$, respectively) constitute a nonempty set and they appear before v in the cw postordering of the vertices of G with respect to $\bar{T}_n$ ($\bar{T}_1$, $\bar{T}_2$, respectively).

For example, in Figure 2.3, the numbering of the vertices of G is the ccw postordering with respect to $\bar{T}_n$. For the internal vertex 5, its ancestor set in $\bar{T}_1$ is $\{1, 3\}$. It is nonempty and all its elements appear before 5 in this numbering.

3. Visibility representation of nearly optimal height. By Theorem 2.2, we need to find an st -numbering $\mathcal{O}$ of G such that $\text{length}(\mathcal{O})$ is bounded. To do this, we use a realizer $\mathcal{R} = \{T_1, T_2, T_n\}$ of G . We will describe how to construct an st -numbering $\mathcal{O}_i$ of G from $\bar{T}_i$ (for each $i \in \{1, 2, n\}$) and show that, for one i , $\text{length}(\mathcal{O}_i)$ has the needed upper bound.

DEFINITION 3.1. Let $T = (V, E)$ be a tree drawn in the plane. The balanced partition of T is the partition of V into three ordered subsets A, B, C as follows. Let a_i be the i th vertex of T in ccw postordering and b_i the i th vertex of T in cw postordering. We mark the vertices of T in the order $a_1, b_1, a_2, b_2, \dots, a_i, b_i, \dots$. Continue this process and stop when one of the following two conditions occurs: (i) when trying to mark a_{k+1} , it has been marked already (as a vertex b_i for some $i \leq k$), or (ii) when trying to mark b_{k+1} , it has been marked (as a vertex a_i for some $i \leq k+1$). When the marking process stops, we define $A = \langle a_1, a_2, \dots, a_k \rangle$, $B = \langle b_1, b_2, \dots, b_k \rangle$, and $C = V - A - B$.

The vertex $a_{k+1} \in C$ is called the merge vertex of T .

Let $P = \langle c_1 (= a_{k+1}), c_2, \dots, c_p \rangle$ be the path in T ordered from the merge vertex to the root of T . We claim $C = P$ and call this path the *leftover* path of T . We prove this claim for case (ii). (The proof for case (i) is similar.) Let $P' = P - \{a_{k+1}\}$ and $R = V - A - P$. We need to show $B = R$. Note that in the ccw postordering of T , the vertices in A are numbered first, then the vertex a_{k+1} , then the vertices in $P' \cup R$. When the marking process stops, none of the vertices in B has been marked as an a_i vertex. Thus, we must have $B \subseteq P' \cup R$. Towards a contradiction, suppose there is a vertex $x \in B \cap P'$. Then in the cw postordering, x has been numbered as a b_i , while the merge vertex a_{k+1} has not been numbered. Since x is an ancestor of a_{k+1} in T , this is impossible. Thus, we must have $P' \cap B = \emptyset$ implying $B \subseteq R$. Towards another contradiction, suppose that there is a vertex $y \in R - B$. When the marking process stops, the vertex b_{k+1} has been marked as an a_i vertex for some $i \leq k+1$. So $b_{k+1} \in A \cup \{a_{k+1}\}$. On the other hand, y has not been marked. However, in the cw postordering of T , all vertices in R are numbered before the vertices in $A \cup \{a_{k+1}\}$. So this is impossible. Hence $R - B = \emptyset$, implying that $B = R$ is to be shown.

For example, the balanced partition of the tree $\bar{T}_n$ in Figure 2.3 is as follows: $A = \langle 1, 2, 3, 4, 5, 6 \rangle$, $B = \langle 13, 11, 12, 9, 8, 10 \rangle$, and $C = \langle 7, 14 \rangle$.

DEFINITION 3.2. Let $T = (V, E)$ be a tree with the balanced partition (A, B, C) . A numbering $\mathcal{O}$ of V is consistent with respect to T if the following hold.

1. For any $i < j$, $a_i \prec_{\mathcal{O}} a_j$, $b_i \prec_{\mathcal{O}} b_j$, and $c_i \prec_{\mathcal{O}} c_j$.
2. For any vertices $a_i \in A, b_j \in B$, and $c_l \in C$, $a_i \prec_{\mathcal{O}} c_l$ and $b_j \prec_{\mathcal{O}} c_l$.

By the definition of the consistent numbering, it is easy to see that the following property holds.

Property 2. Let T be a tree with balanced partition (A, B, C) and $\mathcal{O}$ a numbering consistent with respect to T . If u is a child of v in T , then $u \prec_{\mathcal{O}} v$.

LEMMA 3.3. Let $\mathcal{R} = \{T_1, T_2, T_n\}$ be a realizer of a plane triangulation G . For $i \in \{1, 2, n\}$, let $\mathcal{O}_i$ be a consistent numbering with respect to $\bar{T}_i$. Then $\mathcal{O}_i$ is an *st*-numbering of G .

Proof. We prove only the case $i = n$. The other two cases are symmetric. Let (A, B, C) be the balanced partition of $\bar{T}_n$.

For any vertex other than the root of $\bar{T}_n$, its parent is assigned a bigger number in $\mathcal{O}_n$. The root of $\bar{T}_n$ is assigned n .

For any nonleaf vertex of $\bar{T}_n$, their children are assigned smaller numbers in $\mathcal{O}_n$ by Property 2. Consider a leaf $u \neq v_1, v_2$ of $\bar{T}_n$. Clearly, u is an internal vertex of G . If $u \in A$ or $u \in C$ is the merge vertex, then according to Property 1(1), its ancestor set in $\bar{T}_1$ is nonempty and all its members appear before u in ccw postordering with respect to $\bar{T}_n$. Thus the parent of u is a smaller neighbor of u . Similarly, by Property 1(2), if $u \in B$, the parent of u in $\bar{T}_2$ is a smaller neighbor of u . For v_1 and v_2 , one of them is assigned 1, and it becomes a smaller neighbor of the other. Thus $\mathcal{O}_n$ is an *st*-numbering of G . $\square$

Our algorithm will use one *st*-numbering $\mathcal{O}_i$ mentioned in Lemma 3.3. However, if $\mathcal{O}_i$ is constructed arbitrarily, $\text{length}(\mathcal{O}_i)$ may be too large for our need. In the following, we explore how to construct $\mathcal{O}_i$ in order to bound $\text{length}(\mathcal{O}_i)$, by using a substructure defined below.

DEFINITION 3.4. A ladder graph of order k is a plane graph $L = (V_L, E_L)$. The vertex set V_L is partitioned into two ordered lists $A = \langle a_1, a_2, \dots, a_k \rangle$ and $B = \langle b_1, b_2, \dots, b_k \rangle$. $E_L = E_A \cup E_B \cup E_{\text{cross}}$, where $E_A = \{(a_i, a_{i+1}) | 1 \leq i < k\}$; $E_B = \{(b_i, b_{i+1}) | 1 \leq i < k\}$; E_{cross} consists of edges between a vertex $a_i \in A$ and a vertex

$b_j \in B$; and no two edges in E_{cross} cross each other. The edges in E_{cross} are called cross edges of L .

For a cross edge (a_i, b_j) , we define $\text{slope}(a_i, b_j) = i - j$. A cross edge (a_i, b_j) is called a level (or up or down, respectively) edge if $\text{slope}(a_i, b_j) = 0$ (or $\text{slope}(a_i, b_j) < 0$ or $\text{slope}(a_i, b_j) > 0$, respectively).

A numbering $\mathcal{O}$ of the vertices of a ladder graph $L = (A \cup B, E_L)$ is consistent with respect to L if for any $i < j$, $a_i \prec_{\mathcal{O}} a_j$ and $b_i \prec_{\mathcal{O}} b_j$.

LEMMA 3.5. Let $L = (A \cup B, E_L)$ be a ladder graph of order k . Let l be the number of level edges in L . Then L has a consistent numbering $\mathcal{O}$ such that $\text{length}(\mathcal{O}) \leq (k - 1) + l$. This numbering can be constructed in linear time.

Proof. Let $E_{\text{cross}} = \langle e_1, e_2, \dots, e_p \rangle$ be the ordered list of cross edges of L ordered from bottom up. (This ordering exists because the edges in E_{cross} do not cross each other.) We partition E_{cross} into blocks $E_1, E_2, \dots, E_q$. Each E_j is a maximal sublist of E_{cross} that contains only up edges, or level edges, or down edges. E_j is called an up, or a level, or a down block accordingly.

For each block E_j , let x_j be the largest index such that either a_{x_j} or b_{x_j} is incident to an edge in E_j . The span interval of E_j (for $1 \leq j \leq q$) is defined to be $I_1 = [1, x_1]$; $I_j = (x_{j-1}, x_j]$ (for $1 < j < q$); $I_q = (x_{q-1}, k]$. Note that $I_1, \dots, I_q$ partition the interval $[1, k]$. For each j ($1 \leq j \leq q$), we define a numbering $\mathcal{O}_j$ of the vertex set $\{a_t, b_t \mid t \in I_j\}$ as follows.

If E_j is an up or a level block, let $\mathcal{O}_j = \langle a_{x_{j-1}+1}, a_{x_{j-1}+2}, \dots, a_{x_j}, b_{x_{j-1}+1}, b_{x_{j-1}+2}, \dots, b_{x_j} \rangle$.

If E_j is a down block, let $\mathcal{O}_j = \langle b_{x_{j-1}+1}, b_{x_{j-1}+2}, \dots, b_{x_j}, a_{x_{j-1}+1}, a_{x_{j-1}+2}, \dots, a_{x_j} \rangle$.

Define a numbering of the vertices in L by $\mathcal{O} = \langle \mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_q \rangle$.

It is clear that $\mathcal{O}$ is a consistent numbering of L and the following hold: (i) the path consisting of the edges in E_A and the path consisting of the edges in E_B each has length $k - 1$; (ii) up edges and down edges do not increase the length of the longest path in the orientation of L derived from $\mathcal{O}$; and (iii) each level edge can increase the length of the longest path in the orientation of L derived from $\mathcal{O}$ by at most 1. Thus $\text{length}(\mathcal{O}) \leq k - 1 + l$. It is straightforward to construct $\mathcal{O}$ in linear time. $\square$

For an example, Figure 3.1(1) shows a ladder graph L of order 7. E_{cross} is partitioned into 6 blocks: $E_1 = \{e_1\}$, $E_2 = \{e_2, e_3\}$, $E_3 = \{e_4, e_5\}$, $E_4 = \{e_6\}$, $E_5 = \{e_7, e_8\}$, and $E_6 = \{e_9\}$. $I_1 = [1, 2]$, $I_2 = (2, 3]$, $I_3 = (3, 5]$, $I_4 = (5, 5] = \emptyset$, $I_5 = (5, 7]$, and $I_6 = (7, 7] = \emptyset$. The numbering constructed according to Lemma 3.5 is $\mathcal{O} = \langle b_1, b_2, a_1, a_2, a_3, b_3, a_4, a_5, b_4, b_5, a_6, a_7, b_6, b_7 \rangle$, where $\text{length}(\mathcal{O}) = 8$.

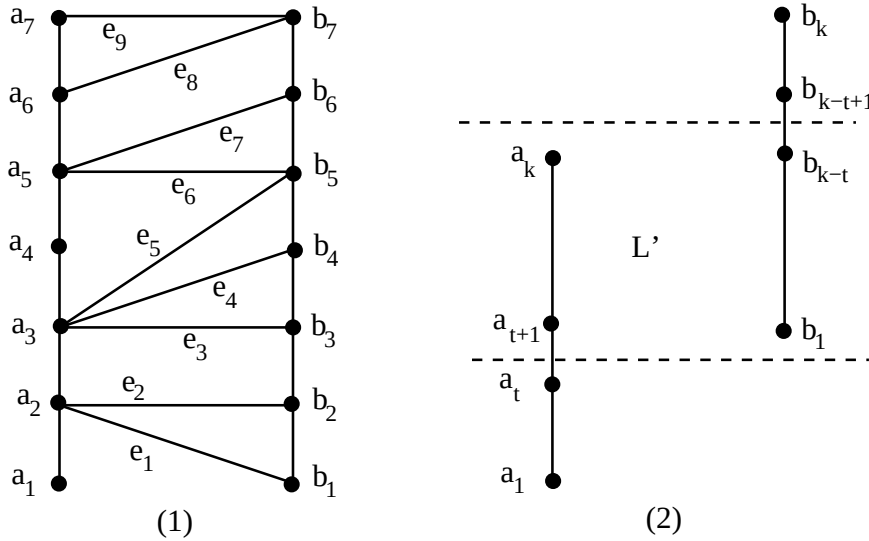
LEMMA 3.6. Let $L = (A \cup B, E_L)$ be a ladder graph of order k . Then L has a consistent numbering $\mathcal{O}$ such that $\text{length}(\mathcal{O}) \leq k + 2\lceil\sqrt{k}\rceil - 1$. $\mathcal{O}$ can be constructed in linear time.

Proof. Let E_A be the set of the edges in L between the vertices in A , E_B the set of the edges in L between the vertices in B , and E_{cross} the set of the cross edges in L . Clearly $|E_A| = |E_B| = k - 1$.

If $|E_{\text{cross}}| \leq 1$, then obviously $|E_{\text{cross}}| \leq 2k - 1$. Otherwise, the external face of L has at least $2k$ edges. Since L is a plane graph with $2k$ vertices, by Euler formula, $|E_L| \leq 3(2k) - (3 + 2k) = 4k - 3$. Thus $|E_{\text{cross}}| = |E_L| - (|E_A| + |E_B|) \leq (4k - 3) - (2k - 2) = 2k - 1$.

Let $p = \lceil\sqrt{k}\rceil$. For each i ($-p \leq i \leq p$), let C_i be the set of cross edges in L with slope i . Since $C_{-p}, \dots, C_{-1}, C_0, C_1, \dots, C_p$ are disjoint, there exists t such that $|C_t| \leq \frac{|E_{\text{cross}}|}{2p+1} \leq \frac{2k-1}{2p+1} \leq \lceil\sqrt{k}\rceil$.

We consider only the case $t \geq 0$. (If $t < 0$, the construction is symmetric.)

FIG. 3.1. (1) A ladder graph L ; (2) proof of Lemma 3.6.

Let $A' = \langle a_{t+1}, a_{t+2}, \dots, a_k \rangle$ and $B' = \langle b_1, b_2, \dots, b_{k-t} \rangle$. Let L' be the ladder graph consisting of the vertex set $A' \cup B'$ and all cross edges in L that are between the vertices in A' and B' . (See Figure 3.1(2).) L' is a ladder graph of order $k - t$ and all level edges of L' are in C_t . By Lemma 3.5, L' has a consistent numbering $\mathcal{O}'$ with $\text{length}(\mathcal{O}') \leq k - t - 1 + |C_t|$.

Let $\mathcal{O} = \langle a_1, \dots, a_t, \mathcal{O}', b_{k-t+1}, \dots, b_k \rangle$. Clearly $\mathcal{O}$ is a consistent numbering of L and $\text{length}(\mathcal{O}) = \text{length}(\mathcal{O}') + 2t \leq (k - t - 1 + |C_t|) + 2t = k - 1 + t + |C_t| \leq k + 2\lceil \sqrt{k} \rceil - 1$. It is straightforward to construct $\mathcal{O}$ in linear time. $\square$

LEMMA 3.7. Let $G = (V, E)$ be a plane triangulation with n vertices. Let $\bar{T}$ be a tree obtained from a realizer $\mathcal{R}$ of G . Let (A, B, C) be the balanced partition of $\bar{T}$. Then G has an st -numbering $\mathcal{O}$ such that $\text{length}(\mathcal{O}) \leq n/2 + |C|/2 + 2\lceil \sqrt{n/2} \rceil - 1$. $\mathcal{O}$ can be constructed in linear time.

Proof. Let $k = (n - |C|)/2$. Then $|A| = |B| = k$. Let $L = (A \cup B, E_L)$ be a graph constructed as follows. The vertex set of L is $A \cup B$. E_L contains the edges $\{(a_i, a_{i+1}), (b_i, b_{i+1}) \mid 1 \leq i < k\}$ and all edges of G between A and B (since G is planar, these edges do not cross). Thus L is a ladder graph of order k . By Lemma 3.6, L has a consistent numbering $\mathcal{O}_L$ with $\text{length}(\mathcal{O}_L) \leq k + 2\lceil \sqrt{k} \rceil - 1$.

Define $\mathcal{O} = \langle \mathcal{O}_L, C \rangle$. Clearly this is a numbering of G consistent with respect to $\bar{T}$. By Lemma 3.3, $\mathcal{O}$ is an st -numbering of G . We have $\text{length}(\mathcal{O}) \leq |C| + \text{length}(\mathcal{O}_L) \leq |C| + (n - |C|)/2 + 2\lceil \sqrt{k} \rceil - 1 \leq n/2 + |C|/2 + 2\lceil \sqrt{n/2} \rceil - 1$. It is straightforward to construct $\mathcal{O}$ in linear time. $\square$

LEMMA 3.8. Let G be a plane triangulation of n vertices with a realizer $\mathcal{R} = \{T_1, T_2, T_n\}$. For $i \in \{1, 2, n\}$, let (A_i, B_i, C_i) be the balanced partition of the tree $\bar{T}_i$. Then there exists $i \in \{1, 2, n\}$ such that $|C_i| \leq n/3 + 1$.

Proof. Each C_i is a path in $\bar{T}_i$. By the property of realizers, if $i \neq j$, C_i and C_j can intersect at at most one vertex. Thus there exists $i \in \{1, 2, n\}$ such that $|C_i| \leq (n + 3)/3 = n/3 + 1$. $\square$

We now can prove our main theorem in this section.

THEOREM 3.9. Let G be a plane triangulation with n vertices. Then G has a VR

with height $\leq 2n/3 + 2\lceil\sqrt{n/2}\rceil$ and width $\leq 2n - 5$, which can be constructed in linear time.

Proof. Let $\mathcal{R} = \{T_1, T_2, T_n\}$ be a realizer of G . For each $i \in \{1, 2, n\}$, let (A_i, B_i, C_i) be the balanced partition of $\overline{T}_i$. By Lemma 3.8, there exists $i \in \{1, 2, n\}$ such that $|C_i| \leq n/3 + 1$. By Lemma 3.7, the st -numbering $\mathcal{O}_i$ associated with this $\overline{T}_i$ satisfies $\text{length}(\mathcal{O}_i) \leq n/2 + |C_i|/2 + 2\lceil\sqrt{n/2}\rceil - 1 \leq n/2 + n/6 + 2\lceil\sqrt{n/2}\rceil - 1 + 1 = 2n/3 + 2\lceil\sqrt{n/2}\rceil$. By Theorem 2.2, the st -numbering $\mathcal{O}_i$ leads to a VR with height at most $2n/3 + 2\lceil\sqrt{n/2}\rceil$.

By Lemma 3.7 and Theorem 2.2, we can construct this VR in linear time. $\square$

4. Visibility representation of nearly optimal width. In order to find a VR of bounded width for a plane triangulation G , we construct an st -numbering $\mathcal{O}^*$ of the st -dual graph G^* with proper bound on $\text{length}(\mathcal{O}^*)$. By the one-to-one correspondence between st -orientations of G and the dual st -orientations of G^* , we can construct the st -orientation $\mathcal{O}$ for G that corresponds to $\mathcal{O}^*$. By Theorem 2.2, this will lead to a VR of G with the desired width.

In order to find such a numbering $\mathcal{O}^*$, our approach is parallel to that in section 3: starting with a realizer $\mathcal{R}^* = \{T_1^*, T_2^*, T_n^*\}$ of the dual graph G^* , we find an st -numbering $\mathcal{O}_i^*$ of G^* for each $i \in \{1, 2, n\}$ and show that one of them has the desired upper bound. However, the dual graph G^* is a 3-regular plane graph, not a plane triangulation. Therefore the concept of realizer defined in Definition 2.3 must be modified. In [8], the realizer concept is generalized from plane triangulation to 3-connected plane graph as follows.

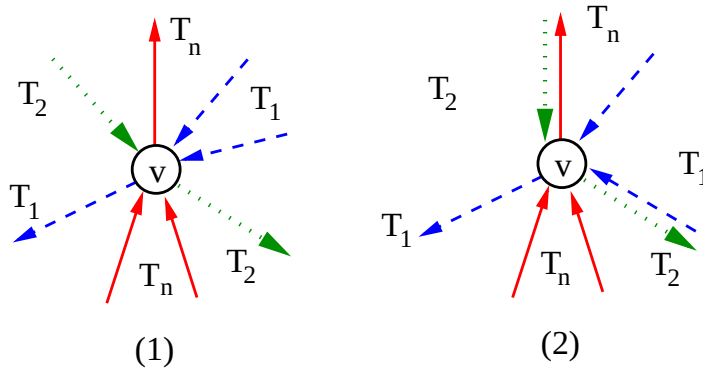
DEFINITION 4.1. *Let G be a 3-connected plane graph with three external vertices v_1, v_2, v_n in ccw order. A realizer of G is a triplet of rooted spanning trees $\{T_1, T_2, T_n\}$ of G with the following properties.*

1. *For $i \in \{1, 2, n\}$, the root of T_i is v_i ; the edges of G are directed from children to parent in T_i .*
2. *Each edge e of G is contained in at least one and at most two spanning trees. If e is contained in two spanning trees, then it has different directions in the two trees.*
3. *For each vertex $v \notin \{v_1, v_2, v_n\}$ of G , v has exactly one edge leaving v in each of T_1, T_2, T_n . The ccw order of the edges incident to v is leaving in T_1 , entering in T_n , leaving in T_2 , entering in T_1 , leaving in T_n , and entering in T_2 . Each entering block may be empty. An edge with two opposite directions is considered twice. The first and the last incoming edges are possibly coincident with the outgoing edges. (Figure 4.1 shows two examples of edge pattern around an interval vertex v . In the second example, the edge leaving v in T_2 and an edge entering v in T_2 are the same edge.)*
4. *For $i \in \{1, 2, n\}$, all the edges incident to v_i belong to T_i .*

We color the edges in T_1 blue, T_2 green, and T_n red. According to the definition, each edge of G is assigned one or two colors, and is said to be 1-colored or 2-colored, respectively.

Consider a realizer $\mathcal{R} = \{T_1, T_2, T_n\}$ (as defined in Definition 2.3) of a plane triangulation G . The triplet of the three trees $\{\overline{T}_1, \overline{T}_2, \overline{T}_n\}$ (each of the three external edges $(v_1, v_2), (v_2, v_n), (v_n, v_1)$ is in two trees) is a special case of the realizer for 3-connected plane graphs defined here, where the three external edges are the only 2-colored edges.

For each vertex v of G and $i \in \{1, 2, n\}$, $p_i(v)$ denotes the path in T_i from v to the root v_i of T_i . A subpath of $p_i(v)$ between the end vertex v and an ancestor u of v

FIG. 4.1. Two examples of edge pattern around an internal vertex v .

in T_i is denoted by $p_i(v, u)$. The subpath of the external face of G , with end vertices v_1 and v_2 , and not containing v_n , is denoted by $ext(v_1, v_2)$. The subpaths $ext(v_2, v_n)$ and $ext(v_n, v_1)$ are defined similarly.

It was shown in [8] that every 3-connected plane graph has a realizer, which can be computed in linear time. The properties of realizers have been studied extensively in [8] and are summarized in the following lemma.

LEMMA 4.2. *Let $G = (V, E)$ be a 3-connected plane graph with $|V| = n$ and $|E| = m$. Let $\mathcal{R} = (T_1, T_2, T_n)$ be a realizer of G , where T_i is rooted at the vertex v_i for $i \in \{1, 2, n\}$.*

1. *The number of 2-colored edges of G is $3n - m - 3$.*
2. *For each vertex v of G , $p_1(v)$, $p_2(v)$, and $p_n(v)$ have only the vertex v in common.*
3. *For $i, j \in \{1, 2, n\}$ ($i \neq j$) and two vertices u and v , the intersection of $p_i(u)$ and $p_j(v)$ is either empty or a common subpath.*
4. *For vertices v_1, v_2, v_n , the following hold: $p_1(v_2) = p_2(v_1) = ext(v_1, v_2)$; $p_2(v_n) = p_n(v_2) = ext(v_2, v_n)$; $p_n(v_1) = p_1(v_n) = ext(v_n, v_1)$.*

Similar to plane triangulations, we have the following version of Property 1 for 3-connected plane graph G with a realizer $\mathcal{R} = \{T_1, T_2, T_n\}$.

Property 3. Let v be an internal vertex of G .

1. All ancestors of v in T_1 (T_2, T_n , respectively) constitute a nonempty set and they appear before v in the ccw postordering of the vertices of G with respect to T_n (T_1, T_2 , respectively).
2. All ancestors of v in T_2 (T_n, T_1 , respectively) constitute a nonempty set and they appear before v in the cw postordering of the vertices of G with respect to T_n (T_1, T_2 , respectively).

DEFINITION 4.3 (see [8]). *Let $G = (V, E)$ be a 3-connected plane graph, with three external vertices v_1, v_2, v_n in ccw order. The extended dual graph G_e^* of G is defined as follows.*

1. *Each internal face f of G corresponds to a node f^* in G_e^* ; the external face of G corresponds to three nodes v_1^*, v_2^* , and v_n^* in G_e^* .*
2. *Each edge e of G corresponds to an edge e^* in G_e^* .*
3. *Two nodes f_1^* and f_2^* of G_e^* (different from v_1^*, v_2^* , and v_n^*) are connected by an edge $e^* = (f_1^*, f_2^*)$ in G_e^* if and only if the corresponding faces f_1 and f_2 in G have the edge e in common.*

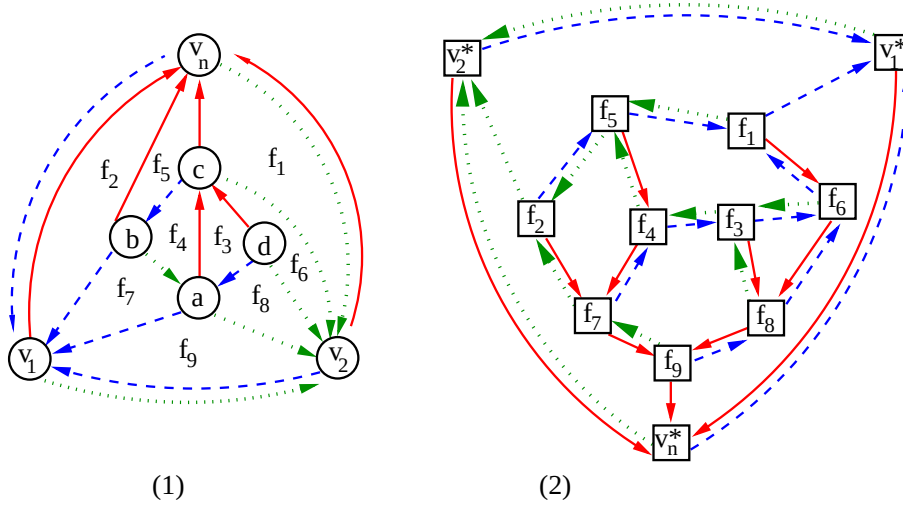


FIG. 4.2. (1) A plane triangulation G with a realizer $\mathcal{R}$; (2) its extended dual G_e^* with dual realizer $\mathcal{R}^*$.

4. v_1^* is adjacent to all the nodes of G_e^* corresponding to faces of G incident to an edge of $\text{ext}(v_2, v_n)$; v_2^* is adjacent to all the nodes of G_e^* corresponding to faces of G incident to an edge of $\text{ext}(v_n, v_1)$; v_n^* is adjacent to all the nodes of G_e^* corresponding to faces of G incident to an edge of $\text{ext}(v_1, v_2)$.

5. (v_1^*, v_2^*) , (v_2^*, v_n^*) , and (v_n^*, v_1^*) are three edges in G_e^* .

Figure 4.2(1) shows a plane triangulation G . Figure 4.2(2) shows its extended dual G_e^* . Note that the extended dual G_e^* of a 3-connected plane triangulation G is also a 3-connected plane graph.

Let $G = (V, E)$ be a 3-connected plane graph with a realizer $\mathcal{R} = \{T_1, T_2, T_n\}$. Let G_e^* be the extended dual of G . In [8], it was shown that $\mathcal{R}$ induces a dual realizer $\mathcal{R}^* = (T_1^*, T_2^*, T_n^*)$ of G_e^* . In our application, we need only the case where G is a plane triangulation. The following definition of $\mathcal{R}^*$ is restricted to this special case.

DEFINITION 4.4. Let $G = (V, E)$ be a plane triangulation with a realizer $\mathcal{R} = \{T_1, T_2, T_n\}$. Then $\mathcal{R}$ induces a dual realizer $\mathcal{R}^* = \{T_1^*, T_2^*, T_n^*\}$ of the extended dual G_e^* as follows.

1. Let $f^* \notin \{v_1^*, v_2^*, v_n^*\}$ be a node of G_e^* . Let f be the face of G corresponding to f^* and e an edge of G that is on the boundary of f . We color and direct the dual edge e^* in G_e^* according to the following rules.
 - (a) If e is blue (i.e., in T_1), then e^* is colored by green and red (i.e., e^* is in both T_2^* and T_n^*). If e is cw on f , then e^* is an outgoing red and incoming green edge for f^* . If e is ccw on f , then e^* is an outgoing green and incoming red edge for f^* .
 - (b) If e is green (i.e., in T_2), then e^* is colored by red and blue (i.e., e^* is in both T_n^* and T_1^*). If e is cw on f , then e^* is an outgoing blue and incoming red edge for f^* . If e is ccw on f , then e^* is an outgoing red and incoming blue edge for f^* .
 - (c) If e is red (i.e., in T_n), then e^* is colored by blue and green (i.e., e^* is in both T_1^* and T_2^*). If e is cw on f , then e^* is an outgoing green and incoming blue edge for f^* . If e is ccw on f , then e^* is an outgoing blue and incoming green edge for f^* .

2. The dual edge of (v_1, v_2) is red (i.e., in T_n^*) and directed to v_n^* . The dual edge of (v_2, v_n) is blue (i.e., in T_1^*) and directed to v_1^* . The dual edge of (v_n, v_1) is green (i.e., in T_2^*) and directed to v_2^* .
3. The edge (v_1^*, v_2^*) is blue and green (i.e., in both T_1^* and T_2^*). The edge (v_2^*, v_n^*) is green and red (i.e., in both T_2^* and T_n^*). The edge (v_n^*, v_1^*) is red and blue (i.e., in both T_n^* and T_1^*).

Figure 4.2(1) shows a plane triangulation G and a realizer $\mathcal{R}$. Figure 4.2(2) shows the dual realizer $\mathcal{R}^*$ of the extended dual graph G_e^* . The extended dual graph G^* and the dual realizer $\mathcal{R}^*$ of a plane triangulation G with a realizer $\mathcal{R}$ satisfy the following property.

Property 4. Let $\mathcal{R}$ be a realizer of a plane triangulation G with n vertices, and $\mathcal{R}^*$ the dual realizer of $\mathcal{R}$ of the extended dual graph G_e^* .

1. The number N of nodes in G_e^* is the number of faces of G plus 2 (since the external face of G corresponds to three nodes in G_e^*). Thus $N = 2n - 4 + 2 = 2n - 2$.
2. The three edges $(v_1, v_2)^*$, $(v_2, v_n)^*$, $(v_n, v_1)^*$ are the only 1-colored edges in $\mathcal{R}^*$.
3. Since $\mathcal{R}^*$ is just a special case of realizers defined for 3-connected plane graphs, Property 3 holds for $\mathcal{R}^*$.

For each tree T_i^* ($i \in \{1, 2, n\}$) of $\mathcal{R}^*$, we can define the balanced partition of T_i^* , and the consistent numbering $\mathcal{O}_i^*$ with respect to T_i^* , exactly the same way as in section 3.

Having covered the necessary concepts and properties, we now proceed to the construction of the VR with nearly optimal width. The approach is parallel to that in section 3. It turns out that almost all results in section 3 remain valid and with almost identical proof (with one major exception). In particular, the corresponding versions of Lemmas 3.3 and 3.7 are given below.

LEMMA 4.5. *Let G be a plane triangulation with a realizer $\mathcal{R} = \{T_1, T_2, T_n\}$. Let G_e^* be the extended dual of G and $\mathcal{R}^* = \{T_1^*, T_2^*, T_n^*\}$ the dual realizer of $\mathcal{R}$. For $i \in \{1, 2, n\}$, let $\mathcal{O}_i^*$ be a consistent numbering with respect to T_i^* . Then $\mathcal{O}_i^*$ induces an st -numbering, denoted by $\mathcal{O}_i^{*r}$, of the st -dual graph G^* of G .*

Proof. We prove only the case $i = n$. The other cases are symmetric. Note that $\mathcal{O}_n^*$ always assigns N to the node v_n^* and assigns the numbers $\{1, k\}$ to the nodes $\{v_1^*, v_2^*\}$ (where $1 < k < N$). First, we merge v_1^* , v_2^* , and v_n^* into a single node t^* . After the merging operation, G_e^* becomes the st -dual graph G^* of G as defined in Definition 2.1. (The merged node t^* corresponds to the external face of G .) The node t^* is still numbered by 1. Adjust $\mathcal{O}_n^*$ (closing the gap for the number k) so that it becomes a numbering of G^* . After the merging, the node with the largest number must be the child of v_n^* in T_n^* other than v_1^* and v_2^* , which corresponds to the internal face of G adjacent to the external edge (v_1, v_2) of G . Denote this node by s^* .

Let $\mathcal{O}_n^{*r}$ be the numbering of the st -dual graph G^* by reversing the numbering $\mathcal{O}_n^*$. Thus $\mathcal{O}_n^{*r}$ assigns 1 to s^* and $N - 2$ to t^* . By repeating the proof of Lemma 3.3, except that we use Property 3 here instead of Property 1, we can show that $\mathcal{O}_n^{*r}$ is an st -numbering of the st -dual graph G^* , in which s^* is the source and t^* is the sink. $\square$

LEMMA 4.6. *Let $G = (V, E)$ be a plane triangulation of n vertices with a realizer $\mathcal{R}$. Let G_e^* be the extended dual of G with $N = 2n - 2$ nodes. Let $\mathcal{R}^*$ be the dual realizer of $\mathcal{R}$. Let (A, B, C) be the balanced partition of a tree T^* in $\mathcal{R}^*$. Then the st -dual graph G^* has an st -numbering $\mathcal{O}^{*r}$ such that $\text{length}(\mathcal{O}^{*r}) \leq N/2 + |C|/2 + 2\lceil\sqrt{N/2}\rceil - 1$.*

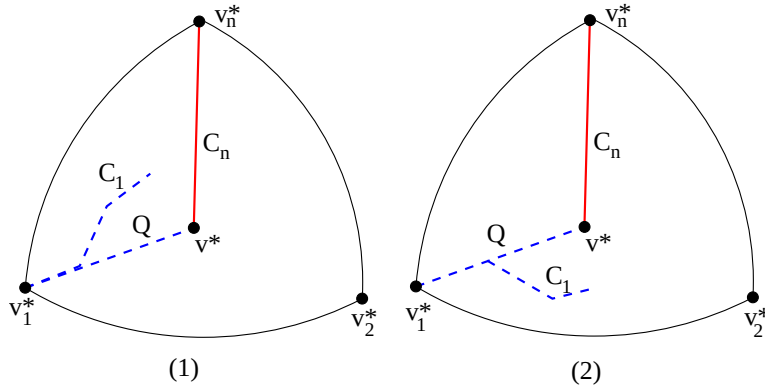


FIG. 4.3. (1) Case 1a; (2) Case 1b.

$\mathcal{O}^{*r}$ can be constructed in linear time.

Proof. By using the identical proof of Lemma 3.7, we can show that $\text{length}(\mathcal{O}^*)$ satisfies the given upper bound. When converting $\mathcal{O}^*$ to $\mathcal{O}^{*r}$, note that merging three nodes from G_e^* into one node in G^* and reversing the ordering of $\mathcal{O}^*$ does not increase the length of the longest path. $\square$

To complete our construction of the VR with nearly optimal width, we also need Lemma 3.8. However, although its conclusion holds, the proof of Lemma 3.8 is not valid here. Its proof relied on a simple fact: For a realizer $\mathcal{R} = \{T_1, T_2, T_n\}$ of a plane triangulation G , the leftover paths C_i and C_j ($i \neq j$) can intersect at at most one vertex. For the dual realizer $\mathcal{R}^* = \{T_1^*, T_2^*, T_n^*\}$, this is not true because the leftover paths C_i and C_j (for T_i^* and T_j^*) can have a common subpath of unbounded length. In the following we restate and prove the version of Lemma 3.8 needed here.

LEMMA 4.7. *Let G be a plane triangulation of n vertices with a realizer $\mathcal{R} = \{T_1, T_2, T_n\}$. Let G_e^* be the extended dual of G with $N = 2n - 2$ nodes, and $\mathcal{R}^* = \{T_1^*, T_2^*, T_n^*\}$ the dual realizer of $\mathcal{R}$. For each $i \in \{1, 2, n\}$, let (A_i, B_i, C_i) be the balanced partition of T_i^* . Then there exists $i \in \{1, 2, n\}$ such that $|C_i| \leq N/3 + 4$.*

Proof. The proof heavily depends on the properties of the dual realizer $\mathcal{R}^*$. We refer readers to [8] for the detailed discussions of these properties. There are several cases.

Case 1. There exist $i, j \in \{1, 2, n\}$ such that C_i and C_j do not have any common node. We assume C_1 and C_n have no common node. The other cases are similar. Let v^* be the starting node of C_n . Let Q be the path in T_1^* from v^* to the root v_1^* of T_1^* . Consider the path C_1 .

Case 1a. C_1 is on the right of Q (see Figure 4.3(1)). Then $C_1 \subseteq A_n$ and $|C_1| \leq |A_n| = (N - |C_n|)/2$. This implies $2|C_1| + |C_n| \leq N$. Hence, either $|C_1|$ or $|C_n|$ is $\leq N/3$.

Case 1b. C_1 is on the left of Q (see Figure 4.3(2)). Then $C_n \subseteq B_1$ and $|C_n| \leq |B_1| = (N - |C_1|)/2$. This implies $|C_1| + 2|C_n| \leq N$. Again, either $|C_1|$ or $|C_n|$ is $\leq N/3$.

Case 2. For every pair $i, j \in \{1, 2, n\}$, C_i and C_j intersect at at least one node, but no node is common to all three leftover paths C_1, C_2, C_n . By the property of the realizer, the intersection pattern of C_1, C_2 , and C_n must be one of the two cases shown in Figure 4.4. We consider only the first case. The other case is symmetric.

The three paths partition the internal region of G^* into several subregions: W is

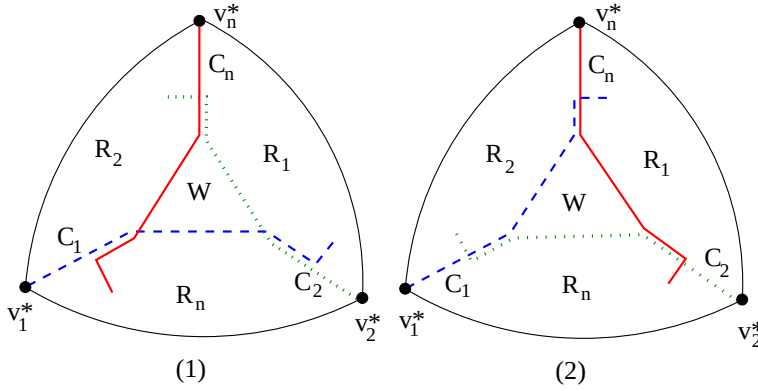


FIG. 4.4. Two possible subcases of Case 2.

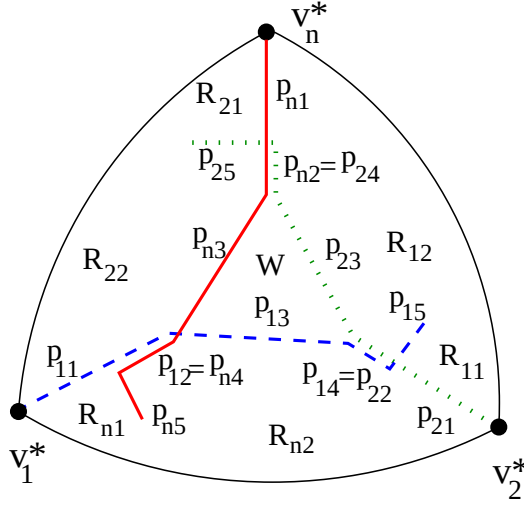


FIG. 4.5. Proof of Case 2.

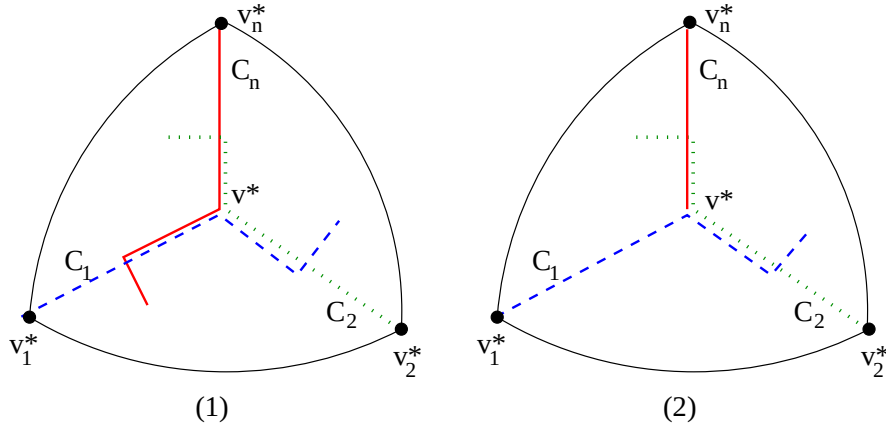
bounded by C_1, C_2 , and C_n ; R_1 is bounded by C_2 and C_n ; R_2 is bounded by C_n and C_1 ; R_n is bounded by C_1 and C_2 (see Figure 4.4(1)).

By the definition of balanced partition, we have

$$(4.1) \quad |A_1| = |B_1|, \quad |A_2| = |B_2|, \quad |A_n| = |B_n|.$$

Partition the path C_1 into 5 subpaths $p_{11}, p_{12}, \dots, p_{15}$ from root v_1^* to the merge node: where p_{11} is incident to v_1^* , p_{12} overlaps with C_n , p_{13} bounds the region W , p_{14} overlaps with C_2 , and p_{15} is in the left of C_2 (see Figure 4.5). Some of these subpaths may be empty. Partition C_2 and C_3 similarly. Thus $p_{12} = p_{n4}$, $p_{22} = p_{14}$, $p_{n2} = p_{24}$. For the region R_1 , let $R_{11} = R_1 \cap A_1$ and $R_{12} = R_1 \cap B_1$. Thus R_{11} , p_{15} , and R_{12} partition R_1 . Partition the regions R_2 and R_n similarly (see Figure 4.5). In the following, $|p_{ij}|$ and $|R_{ij}|$ denote the number of vertices in p_{ij} and R_{ij} , respectively. We have

$$\begin{aligned} |A_1| &= |R_{n1}| + |p_{n5}| + |R_{n2}| + |p_{21}| + |R_{11}| - 2, \\ |A_2| &= |R_{11}| + |p_{15}| + |R_{12}| + |p_{n1}| + |R_{21}| - 2, \\ |A_n| &= |R_{21}| + |p_{25}| + |R_{22}| + |p_{11}| + |R_{n1}| - 2. \end{aligned}$$

FIG. 4.6. *Proof of Case 3.*

(In the equation for $|A_1|$, the last term -2 is because one end node of p_{n5} and one end node p_{21} belong to C_1 , so they should not be counted in $|A_1|$. Similarly for $|A_2|$ and $|A_3|$.) Similarly,

$$\begin{aligned} |B_1| &= |R_{22}| + |p_{25}| + |R_{21}| + |p_{n3}| + |p_{n2}| + |p_{n1}| + |W| + |p_{23}| + |R_{12}| - 6, \\ |B_2| &= |R_{n2}| + |p_{n5}| + |R_{n1}| + |p_{13}| + |p_{12}| + |p_{11}| + |W| + |p_{n3}| + |R_{22}| - 6, \\ |B_n| &= |R_{12}| + |p_{15}| + |R_{11}| + |p_{23}| + |p_{22}| + |p_{21}| + |W| + |p_{13}| + |R_{n2}| - 6. \end{aligned}$$

(In the equation for $|B_1|$, the last term -6 is because one end node of p_{n3} and one end node of p_{23} belong to C_1 ; the common node of p_{n3} , p_{23} , and p_{n2} is counted three times, and the common node of p_{n2} , p_{25} , and p_{n1} is counted three times. Similarly for $|B_2|$ and $|B_3|$.)

Plug these values into (4.1), add them together, and simplify:

$$\begin{aligned} |R_{11}| + |R_{21}| + |R_{n1}| &= (|R_{12}| + |R_{22}| + |R_{n2}|) + 3|W| + 2(|p_{13}| + |p_{23}| + |p_{n3}|) \\ &\quad + (|p_{12}| + |p_{22}| + |p_{n2}|) - 12 \\ (4.2) \quad &\geq (|p_{12}| + |p_{22}| + |p_{n2}|) - 12. \end{aligned}$$

Let $Q_1 = (C_1 - p_{12}) \cup R_{n1}$, $Q_2 = (C_2 - p_{22}) \cup R_{11}$, and $Q_n = (C_n - p_{n2}) \cup R_{21}$. Since Q_1 , Q_2 , Q_n are disjoint, we have

$$(4.3) \quad |Q_1| + |Q_2| + |Q_n| \leq N.$$

By the definition of Q_i and inequalities (4.2) and (4.3), we get

$$(4.4) \quad |C_1| + |C_2| + |C_n| = |Q_1| + |Q_2| + |Q_n| + (|p_{12}| + |p_{22}| + |p_{n2}|) - (|R_{11}| + |R_{21}| + |R_{n1}|)$$

$$(4.5) \quad \leq |Q_1| + |Q_2| + |Q_n| + 12 \leq N + 12.$$

Thus, there exists $i \in \{1, 2, n\}$ such that $|C_i| \leq N/3 + 4$.

Case 3. C_1, C_2, C_n have a node v^* in common.

Case 3a. All three edges in G_e incident to v^* are 2-colored (see Figure 4.6(1)). This is a special case of Case 2 where $W = \emptyset$. The proof of Case 2 still holds.

Case 3b. Exactly two edges incident to v^* are 2-colored (see Figure 4.6(2)). Since $\mathcal{R}^*$ is the dual realizer of a realizer $\mathcal{R}$ of a plane triangulation, the only 1-colored edges of G_e^* are the three internal edges incident to v_1^*, v_2^* , and v_n^* . So there exists i such that C_i consists of a single edge and $|C_i| = 1$.

Case 3c. At most one edge incident to v^* is 2-colored. Similar to Case 3b, one C_i satisfies $|C_i| = 1$. $\square$

We now can prove our main theorem in this section.

THEOREM 4.8. *Let G be a plane triangulation of n vertices. Then G has a VR with width $\leq 4n/3 + 2\lceil\sqrt{n}\rceil$ and height $\leq n - 1$, which can be constructed in linear time.*

Proof. Let $\mathcal{R} = \{T_1, T_2, T_n\}$ be a realizer of G . Let G_e^* be the extended dual of G with $N = 2n - 2$ nodes, and $\mathcal{R}^* = \{T_1^*, T_2^*, T_n^*\}$ the dual realizer of $\mathcal{R}$. For $i \in \{1, 2, n\}$, let (A_i, B_i, C_i) be the balanced partition of T_i^* .

By Lemma 4.7, there exists $i \in \{1, 2, n\}$ such that $|C_i| \leq N/3 + 4$. Consider the numbering $\mathcal{O}_i^*$ of the extended dual graph G_e^* consistent with respect to this T_i^* . By Lemma 4.6, $\mathcal{O}_i^*$ induces an st -numbering $\mathcal{O}_i^{*r}$ of the st -dual G^* such that $\text{length}(\mathcal{O}_i^{*r}) \leq N/2 + |C_i|/2 + 2\lceil\sqrt{N/2}\rceil - 1 \leq N/2 + N/6 + 2\lceil\sqrt{N/2}\rceil - 1 + 2 = 2N/3 + 2\lceil\sqrt{N/2}\rceil + 1 < 4n/3 + 2\lceil\sqrt{n}\rceil$. By Theorem 2.2, this st -numbering leads to a VR of G with width at most $4n/3 + 2\lceil\sqrt{n}\rceil$, which can be constructed in linear time. $\square$

5. Conclusion. In this paper, we present two constructions for VR of plane graphs, one with nearly optimal height $2n/3 + 2\lceil\sqrt{n/2}\rceil$, and one with nearly optimal width $4n/3 + 2\lceil\sqrt{n}\rceil$. Recently, a new algorithm, based on the results in this paper, was polished that constructs a VR with optimal height $\leq 2n/3$ and width $\leq (2n - 5)$ [30]. We conclude the paper by mentioning two open problems:

1. Find a VR construction with optimal width $\leq 4n/3$.
2. Find a VR that simultaneously achieves optimal bounds on width and height.

REFERENCES

- [1] N. BONICHON, S. FELSNER, AND M. MOSBAH, *Convex drawings of 3-connected plane graphs*, Algorithmica, 47 (2007), pp. 399–420.
- [2] N. BONICHON, C. GAVOILLE, AND N. HANUSSE, *An information-theoretic upper bound of planar graphs using triangulation*, in Proceedings of STACS'03, Lecture Notes in Comput. Sci. 2607, Springer-Verlag, Berlin, 2003, pp. 499–510.
- [3] N. BONICHON, B. LE SAËC, AND M. MOSBAH, *Wagner's theorem on realizers*, in Proceedings of ICALP'02, Lecture Notes in Comput. Sci. 2380, Springer-Verlag, Berlin, 2002, pp. 1043–1053.
- [4] H.-L. CHEN, C.-C. LIAO, H.-I. LU, AND H.-C. YEN, *Some applications of orderly spanning trees in graph drawing*, in Proceedings of Graph Drawing'02, Lecture Notes in Comput. Sci. 2528, Springer-Verlag, Berlin, 2002, pp. 332–343.
- [5] Y.-T. CHIANG, C.-C. LIN, AND H.-I. LU, *Orderly spanning trees with applications to graph encoding and graph drawing*, in Proceedings of the Twelfth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, 2001, pp. 506–515.
- [6] Y.-T. CHIANG, C.-C. LIN, AND H.-I. LU, *Orderly spanning trees with applications*, SIAM J. Comput., 34 (2005), pp. 924–945.
- [7] G. DI BATTISTA, P. EADES, R. TAMASSIA, AND I. TOLLIS, *Graph Drawing: Algorithms for the Visualization of Graphs*, Prentice-Hall, Upper Saddle River, NJ, 1998.
- [8] G. DI BATTISTA, R. TAMASSIA, AND L. VISMARA, *Output-sensitive reporting of disjoint paths*, Algorithmica, 23 (1999), pp. 302–340.
- [9] S. FELSNER AND F. ZICKFELD, *Schneider woods and orthogonal surfaces*, in Proceedings of the 14th International Symposium on Graph Drawing, Lecture Notes in Comput. Sci. 4372, Springer-Verlag, Berlin, 2007, pp. 417–429.

- [10] U. FÖSSMEIER, G. KANT, AND M. KAUFMANN, *2-visibility drawings of planar graphs*, in Proceedings of the 4th International Symposium on Graph Drawing, Lecture Notes in Comput. Sci. 1190, Springer-Verlag, Berlin, 1996, pp. 155–168.
- [11] H. DE FRAYSSEIX, J. PACH, AND R. POLLACK, *How to draw a planar graph on a grid*, Combinatorica, 10 (1990), pp. 41–51.
- [12] G. KANT, *Drawing planar graphs using the canonical ordering*, Algorithmica, 16 (1996), pp. 4–32.
- [13] G. KANT, *A more compact visibility representation*, Internat. J. Comput. Geom. Appl., 7 (1997), pp. 197–210.
- [14] G. KANT AND X. HE, *Regular edge labeling of 4-connected plane graphs and its applications in graph drawing problems*, Theoret. Comput. Sci., 172 (1997), pp. 175–193.
- [15] A. LEMPEL, S. EVEN, AND I. CEDERBAUM, *An algorithm for planarity testing of graphs*, in Theory of Graphs (Internat. Sympos. Rome, 1966), Gordon and Breach, New York, Dunod, Paris, 1967, pp. 215–232.
- [16] C.-C. LIAO, H.-I. LU, AND H.-C. YEN, *Floor-planning via orderly spanning trees*, in Proceedings of the 9th International Symposium on Graph Drawing, Lecture Notes in Comput. Sci. 2265, Springer-Verlag, Berlin, 2002, pp. 367–377.
- [17] C.-C. LIN, H.-I. LU, AND I.-F. SUN, *Improved compact visibility representation of planar graph via Schnyder’s realizer*, SIAM J. Discrete Math., 18 (2004), pp. 19–29.
- [18] K. MIURA, M. AZUMA, AND T. NISHIZEKI, *Canonical decomposition, realizer, Schnyder labeling and orderly spanning trees of plane graphs*, Int. J. Found. Comput. Sci., 16 (2005), pp. 117–141.
- [19] P. OSSONA DE MENDEZ, *Orientations bipolaires*, Ph.D. thesis, Ecole des Hautes Etudes en Sciences Sociales, Paris, 1994.
- [20] C. PAPAMANTHOU AND I. G. TOLLIS, *Applications of parameterized st-orientations in graph drawing algorithms*, in Proceedings of the 13th International Symposium on Graph Drawing, Lecture Notes in Comput. Sci. 3843, Springer-Verlag, Berlin, 2006, pp. 355–367.
- [21] C. PAPAMANTHOU AND I. G. TOLLIS, *Parameterized st-orientations of graphs: Algorithms and experiments*, in Proceedings of the 14th International Symposium on Graph Drawing, Lecture Notes in Comput. Sci. 4372, Springer-Verlag, Berlin, 2007, pp. 220–233.
- [22] P. ROSENSTIEHL AND R. E. TARJAN, *Rectilinear planar layouts and bipolar orientations of planar graphs*, Discrete Comput. Geom., 1 (1986), pp. 343–353.
- [23] W. SCHNYDER, *Planar graphs and poset dimension*, Order, 5 (1989), pp. 323–343.
- [24] W. SCHNYDER, *Embedding planar graphs on the grid*, in Proceedings of the First Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, 1990, pp. 138–148.
- [25] R. TAMASSIA AND I. G. TOLLIS, *A unified approach to visibility representations of planar graphs*, Discrete Comput. Geom., 1 (1986), pp. 321–341.
- [26] H. ZHANG AND X. HE, *Canonical ordering trees and their applications in graph drawing*, Discrete Comput. Geom., 33 (2005), pp. 321–344.
- [27] H. ZHANG AND X. HE, *Visibility representation of plane graphs via canonical ordering tree*, Inform. Process. Lett., 96 (2005), pp. 41–48.
- [28] H. ZHANG AND X. HE, *Improved visibility representation of plane graphs*, Comput. Geom., 30 (2005), pp. 29–29.
- [29] H. ZHANG AND X. HE, *An application of well-orderly trees in graph drawing*, in Proceedings of the 13th International Symposium on Graph Drawing, Lecture Notes in Comput. Sci. 3843, Springer-Verlag, Berlin, 2006, pp. 458–467.
- [30] H. ZHANG AND X. HE, *Optimal st-orientations for plane triangulations*, J. Comb. Optim., to appear.

HAMILTON CYCLES IN PLANAR LOCALLY FINITE GRAPHS*

HENNING BRUHN[†] AND XINGXING YU[‡]

Abstract. A classical theorem by Tutte ensures the existence of a Hamilton cycle in every finite 4-connected planar graph. Extensions of this result to infinite graphs require a suitable concept of an infinite cycle. Such a concept was provided by Diestel and Kühn, who defined circles to be homeomorphic images of the unit circle in the Freudenthal compactification of the (locally finite) graph. With this definition we prove a partial extension of Tutte’s result to locally finite graphs.

Key words. Hamilton cycles, infinite graphs, Tutte

AMS subject classification. 05C45

DOI. 10.1137/050631458

1. Introduction. While Hamilton cycles have been investigated intensively in finite graphs, comparatively little attention has been paid to Hamilton cycles in infinite graphs. One reason for this is that it is not entirely clear what the infinite analogon of a cycle should be.

Adomaitis [1] avoided this question by defining a graph to be Hamiltonian if for every finite subset of the vertex set there is a spanning cycle. In contrast, Nash-Williams [8] addressed the problem and proposed spanning double rays as infinite analogons of Hamilton cycles. He noticed that for a spanning double ray to exist the graph needs to be 3-indivisible. (A graph is k -indivisible if the deletion of finitely many vertices leaves at most $k - 1$ infinite components.) Generalizing the following classical result by Tutte, Nash-Williams conjectured that a 3-indivisible 4-connected planar graph contains a spanning double ray.

THEOREM 1 (see Tutte [11]). *Every finite 4-connected planar graph has a Hamilton cycle.*

Recently, Yu [13, 14, 15, 16, 12] announced a proof of Nash-Williams’s conjecture.

The restriction to 3-indivisible graphs is a quite serious one that at first appears unavoidable. Yet, while double rays are the obvious first choice for an infinite analogon of cycles, there is a more subtle alternative which was introduced by Diestel and Kühn [5, 6]. They call the homeomorphic image C of the unit circle in the Freudenthal compactification of a locally finite graph G a *circle*; the subgraph $C \cap G$ is then a *cycle*—for finite graphs this definition coincides with the usual one. In a series of papers it has been shown that this notion is very successful and more suitable than double rays; see Diestel [3] for an introduction and a survey.

These circles overcome the restriction to 3-indivisible graphs. Indeed, in Figure 1 we see an infinite *Hamilton cycle* in a graph that is not k -indivisible for any k . The example is due to Diestel and Kühn [5]. Bruhn (see [3]) conjectured that, in this sense, Theorem 1 extends to locally finite graphs.

*Received by the editors May 12, 2005; accepted for publication (in revised form) May 17, 2008; published electronically September 11, 2008.

<http://www.siam.org/journals/sidma/22-4/63145.html>

[†]Mathematisches Seminar, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany (hbruhn@gmx.net).

[‡]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160 (yu@math.gatech.edu).

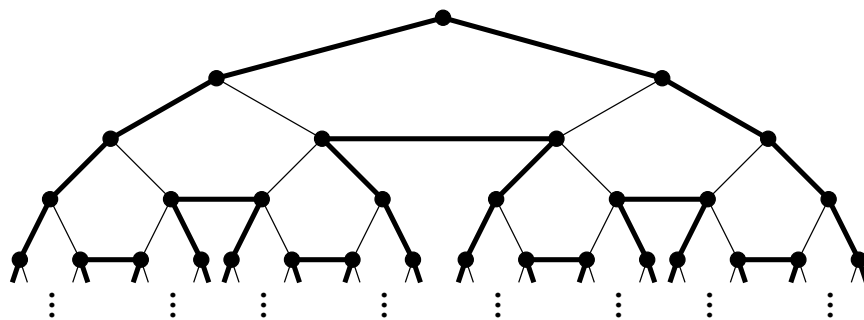


FIG. 1. A Hamilton circle (in bold).

CONJECTURE 2. Let G be a locally finite 4-connected planar graph. Then G has a Hamilton cycle.

In this paper we will present a partial result in this direction.

THEOREM 3. Let G be a locally finite 6-connected planar graph that is k -indivisible for some finite $k \in \mathbb{N}$. Then G has a Hamilton cycle.

Since submission of this paper, there has been some further work on infinite Hamilton cycles. Cui, Wang, and Yu [2] proved that a locally finite 4-connected planar graph has a Hamilton cycle, provided it has a *VAP-free* drawing, that is, a drawing without any vertex accumulation point. This is a substantial special case of Conjecture 2. Furthermore, Georgakopoulos [7] extended Fleischner's theorem on Hamilton cycles in squares of graphs to locally finite graphs.

2. Definitions. In general, our notation will be that of Diestel [4]. If not otherwise noted, all graphs will be simple. A 1-way infinite path is called a *ray*, a 2-way infinite path is a *double ray*, and the subrays of a ray or double ray are its *tails*. Let $G = (V, E)$ be any locally finite graph. Two rays in G are *equivalent* if no finite set of vertices separates them; the corresponding equivalence classes of rays are the *ends* of G . In several places we will implicitly use that every connected locally finite but infinite graph contains a ray, and thus an end.

We say that a finite vertex set S *separates* an end ω from a vertex set U if every ray $R \in \omega$ that starts in a vertex of U meets S . In a similar manner, S separates two ends ω and ω' if every double ray with one tail in ω and the other in ω' goes through S .

We define a topology on G together with its ends; i.e., our topological space consists of all vertices, all inner points of edges, and all ends of G . On G the topology will be that of a 1-complex. Thus, the basic open neighborhoods of an inner point on an edge are the open intervals on the edge containing that point, while the basic open neighborhoods of a vertex x are the unions of half-open intervals containing x , one from every edge at x . For every end ω and any finite set $S \subseteq V$ there is exactly one component $C = C(S, \omega)$ of $G - S$ which contains a tail of every ray in ω . We say that ω *belongs* to C and write $\overline{C}(S, \omega)$ for the component C together with all the ends belonging to it. Then the basic open neighborhoods of an end ω are all sets of the form

$$\hat{C}(S, \omega) := \overline{C}(S, \omega) \cup \mathring{E}(S, \omega).$$

$\overset{\circ}{E}(S, \omega)$ is the set of all inner points of edges between S and $C(S, \omega)$. This topological space will be denoted by $|G|$ and is also known as the *Freudenthal compactification* of G .

We will freely view G either as an abstract graph or as a subspace of $|G|$, i.e., the union of all vertices and edges of G with the usual topology of a 1-complex. It is not difficult to see that if G is connected and locally finite, then $|G|$ is compact. Note that in $|G|$ every ray converges to the end, of which it is an element.

A set $C \subseteq |G|$ is a *circle* if it is homeomorphic to the unit circle. Then C includes every edge of which it contains an inner point, and the graph consisting of these edges and their endvertices is the *cycle* defined by C . Conversely, $C \cap G$ is dense in C , so every circle is the closure in $|G|$ of its cycle and hence defined uniquely by it. Note that every finite cycle in G is also a cycle in this sense, but there can also be infinite cycles; see Diestel [3, 4] for examples and for more information. A *Hamilton cycle* is a cycle that contains every vertex of G ; the corresponding circle will then cover all ends.

G is called *planar* if there is an embedding of G (as a 1-complex) in the sphere S^2 . Since we work in $|G|$ rather than in G , the following fact is quite convenient.

THEOREM 4 (see Richter and Thomassen [9]). *Let G be a locally finite 2-connected planar graph. Then $|G|$ embeds in the sphere.*

It should be noted that for an embedding $\varphi : |G| \rightarrow S^2$, the topological space $|G|$ coincides with the closure of $\varphi(G)$ in the sphere. We call every connected component of $S^2 \setminus \varphi(|G|)$ a *face*. Its boundary is a *face boundary*. It is not difficult to see that each face boundary is (or, more precisely, corresponds to) the closure of a subgraph of G . The following lemma follows immediately from Proposition 3 in [9].

LEMMA 5. *Let G be a locally finite 2-connected graph with an embedding $\varphi : |G| \rightarrow S^2$. Then the face boundaries of $\varphi(|G|)$ are circles of $|G|$.*

For a subgraph H of G , an *H-bridge* is either a chord $e \notin E(H)$ together with its endvertices both of which lie in $V(H)$ or a component K of $G - H$ together with all edges between K and H , denoted by $E(K, H)$, and their incident vertices. We say that an *H-bridge* B is *chordal* if $E(B)$ consists of a single edge. All the vertices of an *H-bridge* B in H are *attachments* of B . For a subgraph F of G , we call a path (resp., cycle) P an *F-Tutte path* (resp., *cycle*) in G if every P -bridge of G has at most three attachments, and if every P -bridge containing an edge of F has at most two attachments.

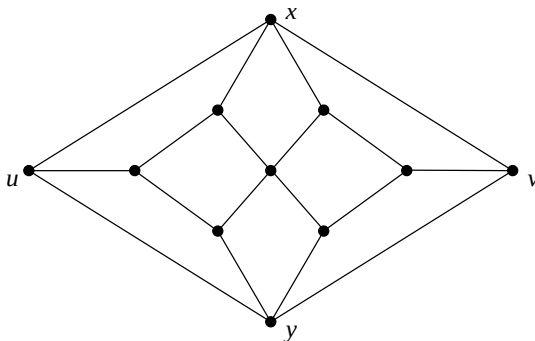


FIG. 2. The Herschel graph.

3. Discussion. The Herschel graph (see Figure 2) is a well-known example for a 3-connected planar graph without a Hamilton cycle which shows that in a sense

Theorem 1 is the best possible. In this section, let us briefly demonstrate that Conjecture 2 is also false for infinite graphs if we only assume 3-connectivity.

Consider Figure 3. There we have arranged copies of the Herschel graph (shaded) in a hexagonal grid. The copies are glued together in such a way that each copy of the vertex v in Figure 2 does not receive an extra edge and thus still has degree 3. Now assume that the resulting graph has a Hamilton cycle C , and consider a copy of the Herschel graph H in that graph. If the Hamilton cycle enters H in u and leaves H in either x or y , then C induces a Hamilton cycle of H , which is impossible. Thus, C enters H in x and leaves it in y , which implies that H has a Hamilton path between x and y . However, this is impossible since both x and y belong to the smaller partition class in the odd bipartite graph H .

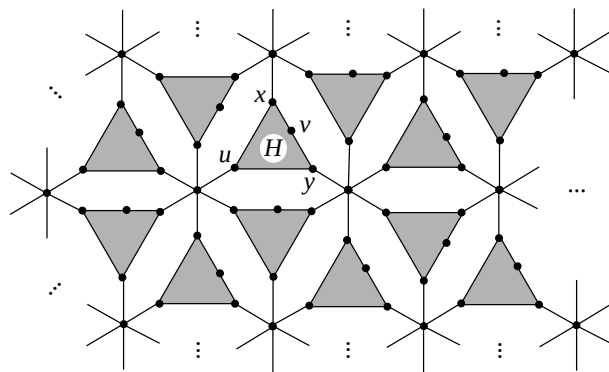


FIG. 3. An infinite 3-connected planar graph without a Hamilton cycle.

4. Proof of main result. Before we start proving Theorem 3, let us reformulate the theorem slightly. The notion of ends are central to the definition of infinite cycles, and we will therefore express the theorem in terms of ends. It is straightforward to see that a locally finite graph is k -indivisible if and only if it has at most $k - 1$ ends. Thus, we obtain the following alternative version of Theorem 3.

THEOREM 6. *Let G be a locally finite 6-connected planar graph with at most finitely many ends. Then G has a Hamilton cycle.*

Our main tool in the proof will be the following result of Thomassen, which itself implies Tutte's theorem.

THEOREM 7 (see Thomassen [10]). *Let G be a finite 2-connected plane graph with a face boundary C . Assume that $u \in V(C)$, $e \in E(C)$, and $v \in V(G) \setminus \{u\}$. Then G contains a C -Tutte path from u to v and through e .*

In order to make use of Thomassen's theorem, we need to chop all infinite parts off of the graph so that a finite part remains, in which we may apply Theorem 7. We then extend this finite Hamilton cycle to a finite part of the chopped off infinite components. For this to work, we need that we can separate 3-connected infinite components from some arbitrary finite vertex set. This is the task of our next lemma and, in particular, of its consequence, Lemma 10.

LEMMA 8. *Let G be a 4-connected locally finite planar graph with minimum degree at least 6, let $Y \subseteq V(G)$ be a finite vertex set, and let ω be an end of G . Then, there is a finite vertex set S such that for $C_\omega := C(S, \omega)$ the following hold:*

- (i) C_ω is disjoint from Y ; and
- (ii) if $X \subseteq V(C_\omega)$ with $|X| \leq 2$, then every component of $C_\omega - X$ is infinite.

Proof. We view G as a plane graph. Choose a finite vertex set S such that $C_\omega := C(S, \omega)$ is disjoint from Y and such that $|E(C_\omega, G - C_\omega)|$ is minimal with that property.

Suppose there is a set $X \subseteq V(C_\omega)$ with $|X| \leq 2$ such that there is a finite component K of $C_\omega - X$. Denote by K' the subgraph obtained by adding to K the vertices in X and the edges between X and K . In particular, note that any possibly existing edges between vertices in X are not included in K' . Put $r := |E(K, G - C_\omega)|$ and $s := |E(X, K)|$. We will show that $s < r$. Then $S' := S \cup V(K)$ leads to a smaller cut between $C(S', \omega)$ and the rest of the graph, which is a contradiction.

First, assume that K' is not a triangulation. Thus, putting $n := |V(K)|$, Euler's formula implies $|E(K')| < 3(n + |X|) - 6 \leq 3(n + 2) - 6 = 3n$. On the other hand,

$$2|E(K')| = \sum_{v \in V(K')} d_{K'}(v) = |E(X, K)| + \sum_{v \in V(K)} d_K(v) \geq s + 6n - r.$$

For the last inequality recall that the minimum degree of G is at least 6. Hence, $3n + (s - r)/2 \leq |E(K')| < 3n$, and thus $s < r$, as desired.

Second, let K' be a triangulation. We may assume that some vertices in S lie in the outer face of K' . Since G is 4-connected and since the face boundary of the outer face of K' is a triangle, T , say, no vertices of G can be contained in the interior face of T . Thus, $T = K'$, and consequently the set X consists of exactly one vertex as there are no edges between vertices of X in K' . This implies that there are exactly two edges between X and K , i.e., $s = 2$. Since the minimum degree is at least 6 in G , and as exactly two vertices in K can have at most one edge between them, it follows that $r \geq 3$. Again, we get $s < r$, as desired. $\square$

LEMMA 9. *Let G be a k -connected graph, and let $V(G) = A \cup B$ be a partition such that $G[B]$ is l -connected. Consider $X \subseteq A$ with $|X| \leq k - l$. Then, for every component K of $G[A] - X$ the graph $G[K \cup B]$ is still l -connected.*

Proof. Suppose there is a vertex set $Y \subseteq V(K) \cup B$ with $|Y| < l$ so that $G[K \cup B] - Y$ has (at least) two distinct components C and D . If both C and D contain vertices of B , then $Y \cap B$ separates two vertices of $G[B]$, contradicting that $G[B]$ is l -connected. So we may assume that $C \subseteq K$. Then, $X \cup Y$ separates C from B in G , but $|X \cup Y| < (k - l) + l = k$, which is a contradiction. $\square$

Lemma 10 will be used in each of the induction steps of the proof of Theorem 6.

LEMMA 10. *Let G be a locally finite planar 6-connected graph, let Y be a finite vertex set, and let ω be an end of G . Let $S \subseteq V(G)$ be a finite vertex set such that $C := C(S, \omega)$ is 2-connected and such that no end other than ω belongs to $C(S \cup Y, \omega)$. Then there is a finite vertex set S' such that $C(S', \omega) \subseteq C$ is 3-connected and disjoint from Y , and such that $C - C(S', \omega)$ is 2-connected.*

Proof. By adding finitely many vertices from C to the set $(Y \cup N(S)) \cap V(C)$, we can ensure that for the resulting finite set $Y' \subseteq V(C)$, the subgraph $G[Y']$ of C is 2-connected. Applying Lemma 8 to Y' and ω yields a finite vertex set $\tilde{S}$. Clearly, as $D := C(\tilde{S}, \omega)$ is disjoint from Y' it is also disjoint from Y . Since $Y' \supseteq N(S) \cap V(C)$, it follows that $D \subseteq C$. Moreover, D is 3-connected because of (ii) in Lemma 8 and the fact that there is only one end, namely, ω , belonging to D . Next, observe that since $G[Y'] \subseteq C$ is 2-connected, there is a block B of $C - D$ containing Y' .

We claim that $N(C - B) =: S'$ is finite and that $C(S', \omega) = C - B$ is 3-connected. As $C - B$ is disjoint from Y , because of $Y' \subseteq V(B)$, the statement of the lemma follows from this claim.

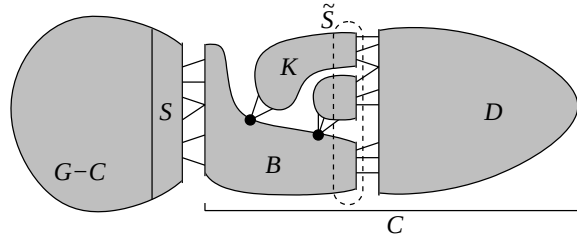


FIG. 4. Illustration of the proof of Lemma 10.

So, let us prove the claim. The vertex set of $C - B$ consists of vertices in D and of vertices from components of $C - D - B$. Let us consider such a component K of $C - D - B$ as illustrated in Figure 4. Observe that K may have at most one neighbor in B since any path between two vertices of a block lies entirely in the block. Also, K cannot have neighbors outside C as K is disjoint from $Y' \supseteq N(S) \cap V(C)$. Thus, all neighbors of K in G , except possibly one, lie in D . Consequently, Lemma 9 (set $X = N(K) \cap V(B)$) implies that $G[D \cup K]$ is 3-connected. Moreover, as G is 2-connected, K has a neighbor in D and thus meets the finite set $\tilde{S} \supseteq N(D)$. Therefore, $C - D - B$ has only finitely many components. Now, enlarging D , one by one, by the components of $C - D - B$, we see that $C - B$ is 3-connected.

Finally, it remains to show that $|S'| < \infty$. But this is easy: a vertex in S' either is contained in $N(D) \subseteq \tilde{S}$ (which is finite) or is the single neighbor in B of a component of $C - D - B$, of which, as shown above, there are only finitely many. $\square$

We will construct our Hamilton cycle in a piecewise manner. Slightly more precisely, we will construct finite nested subgraphs G_i of our locally finite graph G in which the application of Theorem 7 will yield subgraphs H_i that are finite approximations of the desired infinite Hamilton cycle. The following definition and lemma make sure that these subgraphs H_i indeed tend to a cycle, i.e., that $\bigcup_{i=1}^{\infty} H_i$ is a cycle.

If G' is a subgraph of G , then denote by G'^+ the subgraph obtained by adding to G' the edges in $E(V(G'), V(G - G'))$ together with their incident endvertices. We call a sequence $(G_1, H_1), \dots, (G_k, H_k)$ of finite induced subgraphs $G_1 \subseteq \dots \subseteq G_k \subseteq G$ and subgraphs $H_1 \subseteq \dots \subseteq H_k \subseteq G$ good if for $i = 1, \dots, k$ the following hold:

- (i) $G_{i-1}^+ \subseteq G_i$ for $i \geq 2$;
- (ii) $E(H_i \cap G_{i-1}^+) = E(H_{i-1})$ for $i \geq 2$;
- (iii) $V(G_i) \subseteq V(H_i) \subseteq V(G_i^+)$;
- (iv) if K is an infinite component of $G - G_i$, then $|E(K, G - K) \cap E(H_i)| = 2$; and
- (v) there is a finite cycle Z such that $Z \cap G_i = H_i \cap G_i$.

LEMMA 11. Let G be a locally finite connected graph with at most finitely many ends, and let every finite initial segment of $(G_1, H_1), (G_2, H_2), \dots$ be good. Then $H := \bigcup_{i=1}^{\infty} H_i$ is a Hamilton cycle of G .

Proof. By (i) it holds that $V(G) = \bigcup_{i=1}^{\infty} V(G_i)$, and hence (iii) implies $V(H) = V(G)$. Therefore, H is a Hamilton cycle if we can prove that it is a cycle at all. For this, we first show that no finite cut $F \subseteq E(G)$ avoiding $E(H)$ separates two vertices of H . Suppose that on the contrary there is such a cut F and that it separates $u, v \in V(H)$, say. Because of (i) we can choose N large enough so that $F \subseteq E(G_N)$ and $u, v \in V(G_N)$. Then a cycle Z as in (v) contains a $u-v$ path in G that is disjoint from F as $E(Z) \cap F = E(Z) \cap F \cap E(G_N) = E(H_N) \cap E(G_N) \cap F = \emptyset$, which is a contradiction.

Assume that one of the components of H , C , say, is a finite cycle. Since the edges between $V(C)$ and $V(G - C)$ then constitute a finite cut that avoids $E(H)$ but separates C from the rest of the graph, we obtain that $H = C$, and we are done.

So, as (v) implies that H is 2-regular, we may assume that H is the disjoint union of double rays. Let the set of these be $\mathcal{D}$. We define a multigraph H' with $\Omega(G)$ as vertex set and $\mathcal{D}$ as edge set: each $D \in \mathcal{D}$ will be incident with the ends it contains tails of. In particular, if all tails of D lie in the same end ω , then D becomes a loop at ω . Conditions (ii) and (iv) ensure that H' is 2-regular.

Suppose that H' is not connected. Let X be a component of H' , and set $Y := H' - X$. Choose N large enough so that no component of $G - G_N$ has more than one end belonging to it. Because of (iv) this prevents any infinite component of $G - G_N$ from meeting both a double ray from $E(X)$ and of $E(Y)$. Contracting each infinite component of $G - G_N$ (keeping parallel edges but deleting loops) yields a finite minor G' of G , in which no edge belonging to a double ray in $E(X)$ is adjacent with an edge lying in a double ray in $E(Y)$. Therefore, there is a cut F of G' separating these two sets of edges. This cut then also separates two vertices of H in G , contradicting what we have shown earlier. Consequently, H' is a cycle, and hence H is an infinite cycle in G . $\square$

Let us introduce one useful definition before we finally start with the proof. We call a vertex set U in a graph G *externally k -connected* if $|U| \geq k$ and if for every set $X \subseteq V(G)$ with $|X| < k$ there is a path between any two vertices of $U \setminus X$ in $G - X$.

Proof of Theorem 6. By Theorem 4, $|G|$ can be embedded in the sphere. We will identify $|G|$ with that embedding and thus view G as a plane graph.

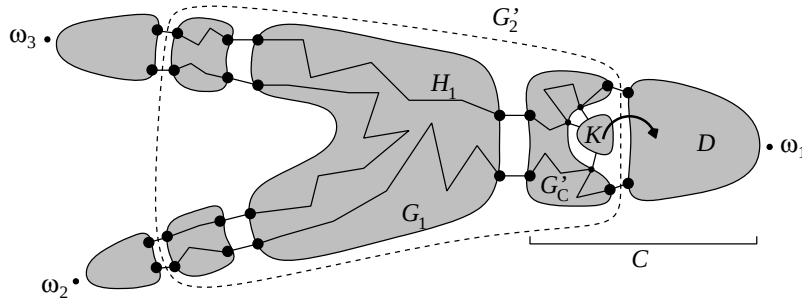
Inductively, we will construct connected finite induced subgraphs $G_1 \subseteq G_2 \subseteq \dots \subseteq G$ and subgraphs $H_1 \subseteq H_2 \subseteq \dots \subseteq G$ such that for every $i \geq 1$ it holds that

- (a) $(G_1, H_1), \dots, (G_i, H_i)$ is good;
- (b) if $\mathcal{C}_i$ is the set of components of $G - G_i$, then each $C \in \mathcal{C}_i$ is 3-connected and there is exactly one end belonging to C ; and
- (c) for each $C \in \mathcal{C}_i$ it follows that $|V(C) \cap V(H_i)| = 2$.

If this can be achieved, then, by Lemma 11, $\bigcup_{i=1}^{\infty} H_i$ is a Hamilton cycle.

Before we proceed, let us give a rough outline of the proof. Ideally, we would like to perform infinitely many steps as follows. In the first step, our foremost task is to separate all the ends. For this, we will find a finite subgraph G_1 so that all components of $G - G_1$ are 3-connected and have exactly one end belonging to them. By contracting these components we will obtain a minor of G , in which the application of Theorem 7 will deliver us the initial piece H_1 of the Hamilton cycle. Slightly more precisely, Theorem 7 yields a Tutte path in the minor—its edges, viewed as edges of G , together with the incident vertices in G form H_1 . In the second step, we need to extend H_1 into each of the components of $G - G_1$. We will treat one component after the other, starting with C , say. In C we will find a finite subgraph G'_C so that the unique component D of $C - G'_C$ is 3-connected. We will use Theorem 7 in the minor obtained from C by contracting D . As we will require the resulting Tutte path H'_C to start in one of the two vertices in $H_1 \cap C$ and end in the other (here, note (c)), H'_C and H_1 will fit together. Doing this for all components of $G - G_1$, we will construct H_2 ; G_2 will be defined to be the (induced subgraph of the) union of G_1 and all the G'_C . Then, we will repeat the second step infinitely often. See Figure 5 for an illustration.

We wrote “ideally” at the beginning of this outline, and that was for a reason: in one aspect the procedure has to be amended in order to work. It is crucial that the H_i cover all vertices of the G_i —otherwise $\bigcup H_i$ will never be a Hamilton cycle.

FIG. 5. Outline: Bridge K is not covered by H_2 .

However, H'_C will, in general, miss some vertices of G'_C and thus produce nonchordal H'_C -bridges in G'_C . Fortunately, and here the fact that the components of $G - G_i$ are 3-connected will come in handy, we will be able to force such a nonchordal bridge to have all its neighbors in $G'_C \cup D$. This will allow us to “put the bridge back”: it will be merged with D and, thus, covered in later steps.

Let us continue with the proof. As the construction of the base case, i.e., when $i = 1$, is quite similar to the general case, i.e., when $i \geq 2$, we will treat both at once. However, for some of the steps we will need to make case distinctions. Put $G_0 := \emptyset$, $H_0 := \emptyset$, and $\mathcal{C}_0 = \{G\}$, and assume (G_{i-1}, H_{i-1}) to be constructed for some $i \geq 1$. Consider a component $C \in \mathcal{C}_{i-1}$.

First, let $i = 1$, and note that as $C = G$ has only finitely many ends, we can choose a finite vertex set W such that no two ends belong to the same infinite component of $G - W$. Denote by F_C a face boundary of $|G|$, and pick an edge $x_C y_C \in E(F_C)$, a vertex $u_C \in V(F_C)$ with $u_C \notin \{x_C, y_C\}$, and a vertex $v_C \notin \{u_C, x_C, y_C\}$ that is adjacent to u_C . Put $U_C := W \cup \{u_C, v_C, x_C, y_C\}$.

For $i \geq 2$, observe that, by (c), H_{i-1} contains exactly two vertices of C , u_C and v_C , say. By (iii), u_C and v_C are neighbors of G_{i-1} . In the embedding of $\overline{C}$ induced by $|G|$, all the neighbors of the connected graph G_{i-1} that lie in C are contained in the same face boundary F_C of $\overline{C}$. Pick an edge $x_C y_C \in E(F_C)$ such that $x_C y_C \neq u_C v_C$, and let U_C be the union of $\{u_C, v_C, x_C, y_C\}$ together with all neighbors of G_{i-1} in C . Thus, $U_C \subseteq V(F_C)$ separates C from the rest of G .

For later use, we note that

$$(4.1) \quad \text{if } i = 1, \text{ then } U_C \text{ separates the ends of } G, \text{ and if } i \geq 2, \text{ then } U_C \subseteq V(F_C) \text{ separates } C \text{ from } G - C$$

and

$$(4.2) \quad u_C \neq v_C \text{ and } |\{u_C, v_C, x_C, y_C\}| \geq 3 \text{ (resp., } = 4 \text{ for } i = 1).$$

Since C is 3-connected (resp., 4-connected for $i = 1$) there are three (resp., four) internally disjoint paths in C between any two vertices in U_C . Denote by Y_C the vertex set of the union of three (resp., four) such paths for each pair of vertices in U_C . Clearly, U_C is externally 3-connected (resp., externally 4-connected for $i = 1$) in $C[Y_C]$.

Let $\omega_1, \dots, \omega_n$ be the ends belonging to C . (Note that if $n \geq 2$, then $i = 1$.) Setting $S_1 := N(C)$ we apply Lemma 10 to Y_C, ω_1, S_1 , which yields a finite vertex set S'_1 , so that $C(S'_1, \omega_1)$ is 3-connected and disjoint from Y_C and so that $C - C(S'_1, \omega_1)$

is 2-connected. Now, proceeding by induction, we set $S_j := S_{j-1} \cup (N(S'_{j-1}) \cap V(C(S'_{j-1}, \omega_{j-1})))$ and denote by S'_j the finite vertex set that the application of Lemma 10 to Y_C, ω_j, S_j yields. Note that $C(S'_j, \omega_j)$ is 3-connected and disjoint from Y_C and that $C(S_j, \omega_j) - C(S'_j, \omega_j) = C - \bigcup_{l=1}^j C(S'_l, \omega_l)$ is 2-connected. Having done this for $j = 1 \dots n$, we set $G'_C := C - \bigcup_{j=1}^n C(S'_j, \omega_j)$ and observe that G'_C is a finite graph: it is connected but cannot contain any ray, since each ray has a tail in one of the $C(S'_j, \omega_j)$. Since $C[Y_C] \subseteq G'_C$, it holds therefore that

$$(4.3) \quad G'_C \text{ is a finite 2-connected graph, and } U_C \text{ is externally 3-connected (resp., externally 4-connected for } i = 1) \text{ in } G'_C.$$

Moreover, as $U_C \subseteq V(G'_C)$ contains all neighbors of G_{i-1} in C , it follows that

$$(4.4) \quad G'_i := G[V(G_{i-1} \cup \bigcup_{B \in \mathcal{C}_{i-1}} G'_B)] \text{ is connected, and every component of } G - G'_i \text{ is 3-connected and has exactly one end of } G \text{ belonging to it.}$$

The graph G'_i will serve as a precursor to G_i . Note that each component of $G - G'_i$ is of the form $C(S'_j, \omega_j)$ for some j and $C \in \mathcal{C}_{i-1}$, and vice versa. To avoid several indices we will from now on denote the set of components of $G - G'_i$ that lie in a $C \in \mathcal{C}_{i-1}$ by $\mathcal{D}_C$. Observe that for $i > 1$ it holds that $|\mathcal{D}_C| = 1$.

As stated in the outline above, we now want to use Theorem 7 in the minor obtained from C by contracting all elements of $\mathcal{D}_C$. Working in this minor, instead of in G'_C , say, has the advantage that we can force the resulting Tutte path $\tilde{H}_C$ to go through each contraction vertex. If $E(\tilde{H}_C)$ is viewed as a subset of $E(G)$, then this means that $E(\tilde{H}_C)$ uses exactly two edges in the cut $E(G - D, D)$ for every $D \in \mathcal{D}_C$, which is (iv). However, we need slightly more: we need that these two edges are not incident with the same vertex in D —see condition (c). To achieve this, we will delete some edges from $E(G - D, D)$ before forming the minor. This is what we will do next; see Figure 6 for an illustration.

So, consider a $D \in \mathcal{D}_C$. Let M_D be obtained from $E(D, G - D)$ by deleting for each vertex in D all but one of the incident edges in $E(D, G - D)$. Thus,

$$(4.5) \quad \text{every neighbor of } G'_C \text{ in } D \text{ is incident with exactly one edge of } M_D.$$

Starting from G'_C define a 2-connected finite plane graph $\tilde{G}_C$ as follows: for every $D \in \mathcal{D}_C$ put a vertex z_D into the face of G'_i that contains D and link z_D to the vertices in G'_C that are incident with an edge in M_D . We will identify these linking edges with the edges in M_D . Note that the resulting graph, which is a minor of G , may have parallel edges. Then, since, for $i = 1$, F_C is a face boundary of $|G|$ such that $u_C, x_C, y_C \in V(F_C)$ and since, for $i \geq 2$, $U_C \subseteq V(F_C)$, this also holds for a face boundary $\tilde{F}_C$ of $\tilde{G}_C$, i.e.,

$$(4.6) \quad \text{for } i = 1 \text{ we have } u_C, x_C, y_C \in V(\tilde{F}_C), \text{ and for } i \geq 2 \text{ it holds that } U_C \subseteq V(\tilde{F}_C).$$

We apply Theorem 7 to $\tilde{G}_C$ and obtain an $\tilde{F}_C$ -Tutte path (resp., cycle for $i = 1$) $\tilde{H}_C$ from u_C to v_C (resp., through $u_C v_C$) and through $x_C y_C$. (More precisely, if $\tilde{G}_C$ has parallel edges, we first subdivide these before using Theorem 7; the obtained Tutte path then induces a Tutte path in $\tilde{G}_C$.) From (4.2) it follows that

$$(4.7) \quad |V(\tilde{H}_C) \cap U_C| \geq 3 \text{ (resp., } \geq 4 \text{ for } i = 1).$$

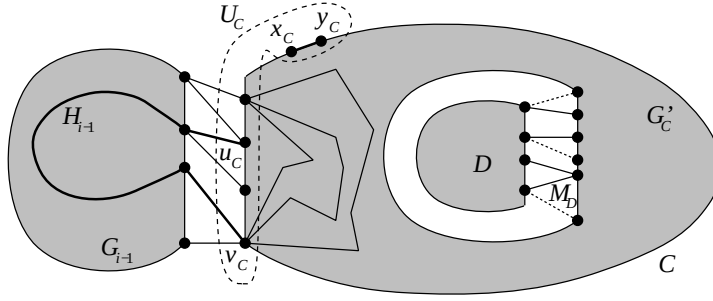


FIG. 6. Claims (4.1)–(4.5).

Next, we show that

- (4.8) *for every nonchordal $\tilde{H}_C$ -bridge K in $\tilde{G}_C$, $K - \tilde{H}_C$ is disjoint from U_C , and all its neighbors in G lie in $G'_C \cup D$ for some $D \in \mathcal{D}_C$.*

Suppose that $K - \tilde{H}_C$ meets U_C . If $i \geq 2$, then K contains an edge of $\tilde{F}_C$, as $U_C \subseteq V(\tilde{F}_C)$ by (4.6). Thus, K has at most two (resp., three for $i = 1$) attachments as $\tilde{H}_C$ is an $\tilde{F}_C$ -Tutte path (resp., cycle). Since U_C is externally 3-connected (resp., externally 4-connected) in G'_C by (4.3), this implies $U_C \subseteq V(K)$. Thus, by (4.7), $K - \tilde{H}_C$ contains a vertex in $V(\tilde{H}_C) \cap U_C$, which is a contradiction. Therefore, $K - \tilde{H}_C$ is disjoint from U_C .

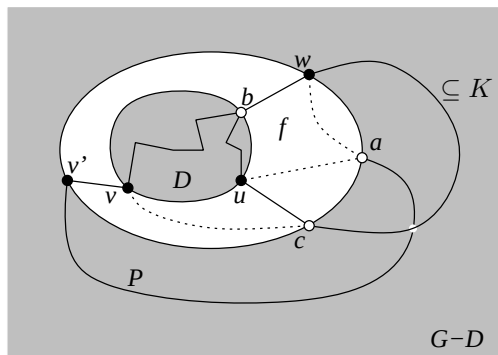
To prove the second assertion, first assume that $i = 1$. Then $C = G$ consists of G'_C and $\bigcup_{D \in \mathcal{D}_C} D$ (plus edges between these graphs). Let there be a neighbor of $K - \tilde{H}_C$ in $D \in \mathcal{D}$ (i.e., $z_D \in V(K)$). As $U_C \subseteq V(G'_C)$, all the $D' \in \mathcal{D}_C$ are disjoint from U_C . On the other hand, U_C separates the ends of G , by (4.1), and thus D from all $D' \in \mathcal{D}_C$, $D' \neq D$. Consequently, $K - \tilde{H}_C$, which is disjoint from U_C , cannot have neighbors in any $D' \in \mathcal{D}_C$ other than D . If, on the other hand, $i \geq 2$, then $\mathcal{D}_C$ consists of only one element, D , say. As $G'_C = C - D$ and as $K - \tilde{H}_C$ is disjoint from U_C , the assertion follows from (4.1).

It holds that

- (4.9) $z_D \in V(\tilde{H}_C)$ for every $D \in \mathcal{D}_C$.

Indeed, suppose there is a $D \in \mathcal{D}_C$ with $z_D \notin V(\tilde{H}_C)$. Thus, $z_D \in V(K - \tilde{H}_C)$ for some $\tilde{H}_C$ -bridge K . Denote by f_D the face of $G - D$ that contains D . Then there exists a vertex a in the face boundary of f_D that does not lie in $V(K)$. Indeed, otherwise the at most three attachments of K separate D from $U_C \setminus V(K) \neq \emptyset$ in G , contradicting that G is 6-connected. (Note here that because of (4.8) and as $z_D \in V(K - \tilde{H}_C)$, all vertices of attachment of K are indeed vertices of G .)

Next, observe that a is not incident with any edge in M_D as $z_D \in V(K - \tilde{H}_C)$. Therefore, there is a unique face $f \subseteq f_D$ of (the closure of) $(G - D) \cup D \cup M_D$ whose face boundary F contains a . Furthermore, F contains exactly two edges of M_D ; let these be uc and bw , where $u, b \in V(D)$; see Figure 7. As G is 6-connected, there is then a path P between a vertex $v \in V(D)$ and a that avoids u, b , and the at most three vertices of attachment of K (which are vertices of G , as above). We may assume that v is the only vertex on P in D . We note that $v \notin V(F)$: as the first edge on P lies in $E(D, G - D)$, v is incident with an edge in M_D , by (4.5), but the only two vertices in $V(F)$ incident with edges in M_D are u, b .

FIG. 7. Subdivision of $K_{3,3}$ in (4.9).

We claim now that by adding some (noncrossing) edges in the plane graph $(G - D) \cup D \cup \{uc, bw, vv'\}$, we obtain a subdivision of $K_{3,3}$ with partition sets $\{a, b, c\}$ and $\{u, v, w\}$, which clearly contradicts planarity. We already have the edges uc and bw , and the v - a path P . Since all of a, u, w lie in F , we can add new edges, drawn in f , between a and each of u, w . We add another new edge, between v and c , which will be drawn in $f_D \setminus f$; this is possible since $v \notin V(F)$. As D is 3-connected, there are two disjoint paths in D , one from b to u and the second from b to v ; these yield two more subdivided edges in the $K_{3,3}$. All that remains now is to find a path from c to w . As endvertices of edges in M_D , both c and w lie in K (because $z_D \in V(K - \tilde{H}_C)$). Thus, if $K - z_D$ is connected (in G), there will be c - w path in $K - z_D$ avoiding all the other subdivided edges—note that P is disjoint from $K - z_D$ as it could only leave K through one of its vertices of attachment.

So, suppose that $K - z_D$ has two components. Because of (4.8), each component must contain a vertex in the boundary of f_D . As K has at most three vertices of attachment, one of the components, K' , say, has at most one vertex in the boundary of f_D . But then this single vertex separates K' from the rest of G , which is a contradiction. This establishes (4.9).

Consider $E(\tilde{H}_C)$ as a subset of $E(G)$, and let H_C be the subgraph of G consisting of the edges $E(\tilde{H}_C)$ and the incident vertices. Put $H_i := H_{i-1} \cup \bigcup_{C \in \mathcal{C}_{i-1}} H_C$, and observe that the pair (G'_i, H_i) already satisfies almost all of the desired properties. Clearly, (b) is (4.4), and (c) holds because of (4.9) and the definition of M_D . Moreover, for (a), $(G_1, H_1), \dots, (G_{i-1}, H_{i-1}), (G'_i, H_i)$ should be a good sequence, and indeed, it at least satisfies (i), (ii), (iv), (v): (i) holds because of $U_C \subseteq V(G'_C)$ for every $C \in \mathcal{C}_{i-1}$ (for $i \geq 2$), (ii) because no H_C uses edges of G_{i-1}^+ , and (iv) because of (4.9). To see (v), add a path through D between the two vertices of H_i in D for every $D \in \mathcal{D}_C$, and denote the resulting subgraph by Z . Clearly, Z is connected, and every $w \in V(Z)$ has degree two: if $w \in V(G_{i-1})$, then because of (v) for $i - 1$, if $w \in \{u_C, v_C\}$, then because $u_C \neq v_C$, if $w \in V(H_i - H_{i-1})$ because H_i is a disjoint union of paths for every C , and finally if $w \in V(Z - H_i)$ because $Z - H_i$ is a disjoint union of paths. Thus, Z is a cycle.

However, (iii) may fail. To fix this, consider a nonchordal H_i -bridge K in G'_i . By (iii) for $i - 1$, we deduce that $K - H_i$ is disjoint from G_{i-1} . Observe that for each $C \in \mathcal{C}_{i-1}$, $K \cap G'_C$ is empty, a chord, or a union of H_C -bridges. Thus, $K - H_i$ is disjoint from U_C , by (4.8), which implies with Lemma 9 that $G'_i - (K - H_i)$ still satisfies

(i) and (b) and then also (i), (ii), (iv), (v), and (c). Thus, putting $G_i := G'_i - (G'_i - H_i)$, we see that for the pair (G_i, H_i) conditions (a)–(c) hold. $\square$

REFERENCES

- [1] R. ADOMAITIS, *Infinite Hamiltonian Graphs*, unpublished, 1995.
- [2] Q. CUI, J. WANG, AND X. YU, *Hamilton Circles in Infinite Planar Graphs*, preprint, Georgia Institute of Technology, Atlanta, GA, 2007.
- [3] R. DIESTEL, *The cycle space of an infinite graph*, *Combin. Probab. Comput.*, 14 (2005), pp. 59–79.
- [4] R. DIESTEL, *Graph Theory*, 3rd ed., Springer-Verlag, New York, 2005.
- [5] R. DIESTEL AND D. KÜHN, *On infinite cycles I*, *Combinatorica*, 24 (2004), pp. 69–89.
- [6] R. DIESTEL AND D. KÜHN, *On infinite cycles II*, *Combinatorica*, 24 (2004), pp. 91–116.
- [7] A. GEORGAKOPOULOS, *Infinite Hamilton Cycles in Squares of Locally Finite Graphs*, preprint, Universität Hamburg, Hamburg, Germany, 2006.
- [8] C. ST. J. A. NASH-WILLIAMS, *Hamiltonian lines in infinite graphs with few vertices of small valency*, *Aequationes Math.*, 7 (1971), pp. 59–81.
- [9] R. B. RICHTER AND C. THOMASSEN, *3-connected planar spaces uniquely embed in the sphere*, *Trans. Amer. Math. Soc.*, 354 (2002), pp. 4585–4595.
- [10] C. THOMASSEN, *A theorem on paths in planar graphs*, *J. Graph Theory*, 7 (1983), pp. 169–176.
- [11] W. T. TUTTE, *A theorem on planar graphs*, *Trans. Amer. Math. Soc.*, 82 (1956), pp. 99–116.
- [12] X. YU, *Infinite paths in planar graphs V, 3-indivisible graphs*, *J. Graph Theory*, to appear.
- [13] X. YU, *Infinite paths in planar graphs I, graphs with radial nets*, *J. Graph Theory*, 47 (2004), pp. 147–162.
- [14] X. YU, *Infinite paths in planar graphs II, structures and ladder nets*, *J. Graph Theory*, 48 (2004), pp. 247–266.
- [15] X. YU, *Infinite paths in planar graphs III, 1-way infinite paths*, *J. Graph Theory*, 51 (2006), pp. 175–198.
- [16] X. YU, *Infinite paths in planar graphs IV, dividing cycles*, *J. Graph Theory*, 53 (2006), pp. 173–195.

PROLIFIC CODES WITH THE IDENTIFIABLE PARENT PROPERTY*

SIMON R. BLACKBURN[†], TUVI ETZION[‡], AND SIAW-LYNN NG[†]

Abstract. Let $\mathcal{C}$ be a code of length n over an alphabet of size q . A word $\mathbf{d}$ is a *descendant* of a pair of codewords $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ if $d_i \in \{x_i, y_i\}$ for $1 \leq i \leq n$. A code $\mathcal{C}$ is an *identifiable parent property* (IPP) code if the following property holds. Whenever we are given $\mathcal{C}$ and a descendant $\mathbf{d}$ of a pair of codewords in $\mathcal{C}$, it is possible to determine at least one of these codewords. The paper introduces the notion of a prolific IPP code. An IPP code is *prolific* if all q^n words are descendants. It is shown that linear prolific IPP codes fall into three infinite (“trivial”) families, together with a single sporadic example which is ternary of length 4. There are no known examples of prolific IPP codes which are not equivalent to a linear example: the paper shows that for most parameters there are no prolific IPP codes, leaving a relatively small number of parameters unsolved. In the process the paper obtains upper bounds on the size of a (not necessarily prolific) IPP code which are better than previously known bounds.

Key words. error-correcting codes, identifiable parent property, linear codes, MDS codes, orthogonal arrays, copyright protection

AMS subject classifications. 94B60, 94A60, 94B65

DOI. 10.1137/070695551

1. Introduction. Codes with the identifiable parent property were first introduced by Hollmann et al. [9] in 1998, motivated by an application to prevent software piracy. IPP codes and various generalizations have since been intensively studied; see, for example, the papers of Alon et al. [1], Alon, Fischer, and Szegedy [2], Alon and Stav [3], Barg et al. [4], Barg and Kabatiansky [5], Blackburn [6], Lindkvist, Löfvenberg, and Svanström [10], Löfvenberg [11], Staddon, Stinson, and Wei [13], Tô and Safavi-Naini [14], van Trung and Martirosyan [15], and Yemane [16].

To define IPP codes we need the notion of a descendant, which is defined as follows. Let F be an alphabet of size q . Let $\mathbf{x} = x_1x_2 \dots x_n \in F^n$ and $\mathbf{y} = y_1y_2 \dots y_n \in F^n$ be q -ary words of length n . The set of *descendants* $\text{desc}(\mathbf{x}, \mathbf{y})$ of $\mathbf{x}$ and $\mathbf{y}$ is defined to be

$$\text{desc}(\mathbf{x}, \mathbf{y}) = \{d_1d_2 \dots d_n \in F^n : d_i \in \{x_i, y_i\} \text{ for } i = 1, 2, \dots, n\}.$$

If $\mathbf{d} \in \text{desc}(\mathbf{x}, \mathbf{y})$, we say that $\mathbf{d}$ is a *descendant* of $\mathbf{x}$ and $\mathbf{y}$, and we say that $\{\mathbf{x}, \mathbf{y}\}$ is a set of *parents* of $\mathbf{d}$. We say that the parent $\mathbf{x}$ *contributes to the i th component of $\mathbf{d}$* if $x_i = d_i$. Clearly $|\text{desc}(\mathbf{x}, \mathbf{y})| = 2^{d(\mathbf{x}, \mathbf{y})}$, where $d(\mathbf{x}, \mathbf{y})$ is the Hamming distance between $\mathbf{x}$ and $\mathbf{y}$.

Let $\mathcal{C}$ be an (n, q, M) -code (so $\mathcal{C}$ is a q -ary code of length n , containing M codewords). Informally, $\mathcal{C}$ has the identifiable parent property (we say $\mathcal{C}$ is an *IPP code* or an (n, q, M) -*IPP code*) if, whenever we are given a descendant $\mathbf{d}$ of two codewords,

*Received by the editors June 26, 2007; accepted for publication (in revised form) April 23, 2008; published electronically September 11, 2008. This research was supported in part by E.P.S.R.C. grant EP/E034632/1.

<http://www.siam.org/journals/sidma/22-4/69555.html>

[†]Department of Mathematics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom (s.blackburn@rhul.ac.uk, s.ng@rhul.ac.uk).

[‡]Technion — Israel Institute of Technology, Department of Computer Science, Technion City, Haifa 32000, Israel (etzion@cs.technion.ac.il).

we are able to identify one of the parents. More formally, $\mathcal{C}$ is an IPP code if the following holds. For $\mathbf{d} \in F^n$, define

$$P_{\mathbf{d}} = \{\{\mathbf{x}, \mathbf{y}\} \subseteq \mathcal{C} : \mathbf{d} \in \text{desc}(\mathbf{x}, \mathbf{y})\}.$$

Then $\mathcal{C}$ is an IPP code if for all $\mathbf{d} \in F^n$ which are descendants of one or more pairs of codewords

$$\bigcap_{\{\mathbf{x}, \mathbf{y}\} \in P_{\mathbf{d}}} \{\mathbf{x}, \mathbf{y}\} \neq \emptyset.$$

The following lemma, due to Hollmann et al. [9], gives simple criteria for a code to have the identifiable parent property.

LEMMA 1.1. *An (n, q, M) -code $\mathcal{C}$ is an IPP code if and only if the following hold:*

- IPP1. *For any three distinct codewords $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{C}$ there exists $i \in \{1, 2, \dots, n\}$ such that x_i, y_i , and z_i are distinct.*
- IPP2. *For any four codewords $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{v} \in \mathcal{C}$ such that $\{\mathbf{x}, \mathbf{y}\} \cap \{\mathbf{z}, \mathbf{v}\} = \emptyset$, there exists $i \in \{1, 2, \dots, n\}$ such that $\{x_i, y_i\} \cap \{z_i, v_i\} = \emptyset$.*

Hollmann et al. [9] observed that the ternary Hamming code of length 4 is an example of a $(4, 3, 9)$ -IPP code:

$$\mathcal{C} = \{0000, 0111, 0222, 1012, 1120, 1201, 2021, 2102, 2210\}.$$

To see why $\mathcal{C}$ is an IPP code, note that since all codewords are at distance 3, a descendant $\mathbf{d} \in \text{desc}(\mathbf{x}, \mathbf{y})$ is at distance at most 1 from exactly one of its parents $\mathbf{x}, \mathbf{y}$. But the minimum distance of the code shows that $\mathbf{d}$ cannot be of distance at most 1 from two distinct codewords. Thus $\mathcal{C}$ is an IPP code, with the identified parent of a descendant $\mathbf{d}$ being the unique codeword at distance at most 1 from $\mathbf{d}$.

This example has the beautiful property that every possible word is a descendant. We say that a code is *prolific* if every word is a descendant of some pair of codewords. The main question this paper asks is, What other examples of prolific IPP codes are there? This question is motivated by an attempt to draw parallels between error-correcting codes and IPP codes. There are clear connections between the two areas: at a most basic level, we observed above that the size of the set of descendants is related to Hamming distance; moreover, error-correcting codes of high minimum distance provide good explicit constructions of IPP codes (see Hollmann et al. [9, Theorem 4], for example). From this perspective, prolific IPP codes may be thought of as analogues of perfect error-correcting codes.

There are three *trivial* families of prolific IPP codes. First, the set F of all words of length 1 is a prolific $(1, q, q)$ -IPP code. Second, the repetition code of length 2 (with codewords of the form aa , where $a \in F$) is a prolific $(2, q, q)$ -IPP code. Third, any binary word and its complement form a prolific $(n, 2, 2)$ -IPP code. It is easy to see that a prolific IPP code which has length 1 or 2, or which is a binary code, must be equivalent to a member of one of these three families, and so from now on we assume that $n \geq 3$ and $q \geq 3$.

All the examples above are linear codes. One main goal of the paper is to show that the $(4, 3, 9)$ -code above is the only nontrivial example of a linear prolific IPP code (up to equivalence). In general, we conjecture that there are no more examples of prolific IPP codes (linear or not). We do not know how to show this, but we are able to prove that there are no more examples when $n \leq 5$ (due to space constraints, we just provide a sketch proof for the case $n = 5$). In addition, we rule out many

other parameters. As a side-benefit of our investigations, we are able to provide new upper bounds on the size of a (not necessarily prolific) IPP code.

For the rest of this paper, $\mathcal{C}$ will be a q -ary code of length n with M codewords. We write words and subwords in a bold font to distinguish them from components of codewords. We write $\ell(\mathbf{x})$ for the length of the (sub)word $\mathbf{x}$.

The paper is structured as follows. In section 2, we provide some simple upper and lower bounds for the size of a prolific IPP code. Section 3 shows that there are no nontrivial linear prolific IPP codes other than the $(4, 3, 9)$ -code above. We then turn our attention away from the linear case. We provide new upper bounds for (not necessarily prolific) IPP codes in section 4. Sections 5 and 6 show that there are no nontrivial examples of prolific IPP codes for lengths 3 and 4, respectively, other than the $(4, 3, 9)$ example. Section 7 contains a sketch proof that there are no nontrivial examples of length 5. Finally, section 8 summarizes the parameters where it is unknown whether a prolific IPP code exists and comments on possibilities for future work.

2. General bounds for prolific IPP codes. Since prolific IPP codes have many descendants, it seems intuitively plausible that they must be fairly large. This is indeed the case, and this section makes this precise by establishing lower bounds on the size of a prolific IPP code. For many parameters these lower bounds conflict with known upper bounds on the size of an IPP code, and so the bounds rule out the existence of a prolific IPP code for these parameters.

The simplest lower bound on the size of a prolific IPP code is stated in the following theorem.

THEOREM 2.1. *If $\mathcal{C}$ is an (n, q, M) prolific IPP code, then $\binom{M}{2}2^n \geq q^n$.*

Proof. There are at most $\binom{M}{2}$ pairs of codewords from $\mathcal{C}$. Each pair of codewords can produce at most 2^n descendants, since there are at most two possibilities for each component of a descendant once the pair of parents is fixed. So $\mathcal{C}$ has at most $\binom{M}{2}2^n$ descendants. The bound follows once we observe that all q^n words must be descendants since $\mathcal{C}$ is prolific. $\square$

The counting argument used above will tend to significantly overcount descendants which are close to a codeword (in terms of Hamming distance). We can overcome this problem by counting the descendants of the code in another way, giving us the following improvement on Theorem 2.1.

THEOREM 2.2. *Let $\mathcal{C}$ be an (n, q, M) prolific IPP code and let k be a positive integer. Then*

$$M \left(\sum_{i=0}^k \binom{n}{i} (q-1)^i + \frac{M-1}{2} \sum_{i=k+1}^{n-k-1} \binom{n}{i} \right) \geq q^n.$$

Proof. We count the descendants of $\mathcal{C}$ as follows. A sphere of radius k contains $\sum_{i=0}^k \binom{n}{i} (q-1)^i$ words, and so there are at most $M(\sum_{i=0}^k \binom{n}{i} (q-1)^i)$ descendants of $\mathcal{C}$ at distance at most k from the code. A descendant of a pair $\{\mathbf{c}_1, \mathbf{c}_2\} \subseteq \mathcal{C}$ of codewords is formed by choosing i components from $\mathbf{c}_1$ and the remaining $n-i$ components from $\mathbf{c}_2$ for some $i \in \{0, 1, \dots, n\}$. But when $0 \leq i \leq k$ or $n-k \leq i \leq n$ the resulting descendant is within distance k of the code. So each of the $\binom{M}{2}$ pairs of codewords gives rise to at most $\sum_{i=k+1}^{n-k-1} \binom{n}{i}$ descendants of distance greater than k from the code. Thus $\mathcal{C}$ has at most $M(\sum_{i=0}^k \binom{n}{i} (q-1)^i + \frac{M-1}{2} \sum_{i=k+1}^{n-k-1} \binom{n}{i})$ descendants, and the theorem follows by the same argument as in Theorem 2.1. $\square$

We finish this section by stating an upper bound on an IPP code due to Hollmann et al. [9]. For many parameters, this upper bound conflicts with the lower bounds above, showing that no prolific IPP codes exist for these parameters.

THEOREM 2.3. *Let $\mathcal{C}$ be an IPP code of length 3, where position i , $1 \leq i \leq 3$, of a codeword is taken from an alphabet Q_i . Then*

$$(2.1) \quad |\mathcal{C}| \leq |Q_1| + |Q_2| + |Q_3| - 1.$$

Hollmann et al. used (2.1) to obtain the following bounds on the size of IPP codes.

THEOREM 2.4. *Let $\mathcal{C}$ be an (n, q, M) -IPP code.*

(i) *If $n = 3\ell - 2$, then $M \leq q^\ell + 2q^{\ell-1} - 1$.*

(ii) *If $n = 3\ell - 1$, then $M \leq 2q^\ell + q^{\ell-1} - 1$.*

(iii) *If $n = 3\ell$, then $M \leq 3q^\ell - 1$.*

Proof. Write $n = k_1 + k_2 + k_3$, where $k_i \in \{\ell - 1, \ell\}$. Define $Q_i = F^{k_i}$. We may write any codeword in $\mathcal{C}$ in the form $\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3$, where $\ell(\mathbf{x}_i) = k_i$. So $\mathcal{C}$ can be regarded as a length 3 code, with symbols in position i taken from the alphabet Q_i . It is easy to verify that $\mathcal{C}$ is still an IPP code when thought of in this way, and so the bound follows by Theorem 2.3. $\square$

3. Linear codes. The goal of this section is to prove Theorem 3.2, that there are no nontrivial linear prolific IPP codes other than the $(4, 3, 9)$ -code from the introduction. The most difficult step is to prove this result in the special case when the code is maximal distance separable (MDS); see Lemma 3.1. (Recall that an *MDS code* is a linear- (n, q, q^k) code of minimum distance $n - k + 1$.)

Note that an (n, q, q^k) -code of minimum distance $n - k + 1$ meets the Singleton bound. In particular, whenever we restrict all codewords to a set of k positions, we find that every word of length k appears exactly once as a restriction. We refer to this property as the *MDS property* of the code.

We use the well-known result that an (n, q, q^k) -code $\mathcal{C}$ with minimum distance $n - k + 1$ can exist only when $q \geq n - k + 1$. To see this, note that we can construct an $(n - k + 2, q, q^2)$ -code $\mathcal{C}'$ of minimum distance $n - k + 1$ by taking all codewords in $\mathcal{C}$ ending in $k - 2$ zeros, and then removing these zeros to produce words of length $n - k + 2$. But then $\mathcal{C}'$ implies the existence of a set of $n - k$ mutually orthogonal Latin squares of order q (see Hill [8, Theorem 10.20]), and such a set can have size at most $q - 1$ (see Hill [8, Theorem 10.18]).

LEMMA 3.1. *Let $\mathcal{C}$ be a linear (n, q, q^k) -code of minimum distance $n - k + 1$ (in other words, $\mathcal{C}$ is an MDS code). Let $n \geq 3$ and $q \geq 3$. If $\mathcal{C}$ is a prolific IPP code, then $n = 4$, $q = 3$, and $k = 2$. In particular, no MDS code of length strictly greater than 4 is a prolific IPP code.*

Proof. Assume that we are not in the case when $n = 4$, $q = 3$, and $k = 2$. We need to show that $\mathcal{C}$ is not a prolific IPP code. We deal with the length 3 and 4 cases first, and then go on to consider the remaining cases.

Suppose that $n = 3$. When $k = 0$ or $k = 1$, we see that $\mathcal{C}$ is too small to be a prolific code, by Theorem 2.1. When $k = 2$ or $k = 3$ we see that $\mathcal{C}$ is too large to be an IPP code, by Theorem 2.4. So we get a contradiction when $n = 3$, as required.

Suppose that $n = 4$. Theorem 2.1 implies that $k \geq 2$, and Theorem 2.4 implies that $k \leq 2$. So we may assume that $k = 2$, and thus $q > 3$ and $\mathcal{C}$ is a $(4, q, q^2)$ code of minimum distance 3. The union of spheres of radius 1 about codewords contains $q^2(1 + 4(q - 1))$ words, and since $q > 3$ we have that $q^2(1 + 4(q - 1)) < q^4$. So there exists a word $\mathbf{d} = d_1d_2d_3d_4$ of distance at least 2 from any codeword. By the

MDS property of $\mathcal{C}$, there exist codewords $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4 \in \mathcal{C}$ of the form $\mathbf{c}_1 = d_1 d_2 **$, $\mathbf{c}_2 = ** d_3 d_4$, $\mathbf{c}_3 = d_1 ** d_4$, and $\mathbf{c}_4 = * d_2 d_3 *$. These codewords are distinct, since $\mathbf{d}$ is at distance 2 from $\mathcal{C}$. But then the sets $\{\mathbf{c}_1, \mathbf{c}_2\}$ and $\{\mathbf{c}_3, \mathbf{c}_4\}$ violate IPP2. (Recall the definition of IPP2 from Lemma 1.1.) So we have a contradiction in this case.

It remains to consider the situation when $n > 4$. We distinguish between six cases. In the first two cases we show that $\mathcal{C}$ is not an IPP code; in the remaining cases we show that $\mathcal{C}$ cannot be prolific.

Case 1. $n \leq 3k - 3$. Define integers n_1, n_2 , and n_3 by $n_1 = \lceil \frac{n}{3} \rceil$, $n_2 = \lfloor \frac{n}{3} \rfloor$, and $n_3 = n - \lceil \frac{n}{3} \rceil - \lfloor \frac{n}{3} \rfloor$. We can write each codeword in the form $\mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_3$, where $\ell(\mathbf{x}_i) = n_i$. Let $\mathbf{c}_0 = \mathbf{000}$ be the all-zero codeword. Since $n_1 \leq k$, there are exactly q^{k-n_1} codewords of the form $\mathbf{0}**$. Since $n_1 < k$, we have that $q^{k-n_1} > 1$, and so there exists a codeword $\mathbf{c}_1 \in \mathcal{C} \setminus \{\mathbf{c}_0\}$ of the form $\mathbf{c}_1 = \mathbf{0}*\mathbf{y}$ for some word $\mathbf{y}$ of length n_3 . Similarly, since $n_2 < k$ there exists a codeword $\mathbf{c}_2$ distinct from $\mathbf{c}_0$ of the form $\mathbf{c}_2 = *\mathbf{0}*$, and since $n_3 < k$ there is a codeword $\mathbf{c}_3$ distinct from $\mathbf{c}_1$ of the form $**\mathbf{y}$. If $\mathbf{y} = \mathbf{0}$, then the codewords $\mathbf{c}_0, \mathbf{c}_1$, and $\mathbf{c}_2$ violate IPP1; if $\mathbf{y} \neq \mathbf{0}$, then the sets $\{\mathbf{c}_0, \mathbf{c}_3\}$ and $\{\mathbf{c}_1, \mathbf{c}_2\}$ violate IPP2. So $\mathcal{C}$ is not an IPP code.

Case 2. $n = 3k - 2$. Note that since $n > 4$, we have that $k \geq 3$. Define $n_1 = n_2 = k - 1$ and $n_3 = k$. As before, we can write any codeword in the form $\mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_3$, where $\ell(\mathbf{x}_i) = n_i$. Let $\mathbf{c}_0 \in \mathcal{C}$ be the all-zero codeword. Since $n_1 < k$, there exists a codeword $\mathbf{c}_1 \in \mathcal{C} \setminus \{\mathbf{c}_0\}$ of the form $\mathbf{0}*\mathbf{y}$ for some word $\mathbf{y}$ of length n_3 . Since $n_2 = k - 1$, there are $q - 1$ nonzero codewords of the form $*\mathbf{0}*$; moreover, no two distinct words of this form can agree anywhere in their last n_3 positions, as this would contradict the MDS property of the code. This implies (since $q \geq 3$ and $n_3 = k \geq 3$) that we may choose a codeword $\mathbf{c}_2$ distinct from $\mathbf{c}_0$ of the form $\mathbf{c}_2 = *\mathbf{0}\mathbf{z}$, where $d(\mathbf{y}, \mathbf{z}) \geq 2$. Let $\mathbf{w}$ be a word of length n_3 such that $\mathbf{w} \in \text{desc}(\mathbf{y}, \mathbf{z}) \setminus \{\mathbf{y}, \mathbf{z}\}$. Such a word exists since $d(\mathbf{y}, \mathbf{z}) \geq 2$. Let $\mathbf{c}_3$ be the (unique) codeword of the form $**\mathbf{w}$. The sets $\{\mathbf{c}_0, \mathbf{c}_3\}$, $\{\mathbf{c}_1, \mathbf{c}_2\}$ violate IPP2. Hence the code is not an IPP code.

Case 3. $n = 3k - 1$ and $k \geq 4$. We can write any word in the form $\mathbf{x}\mathbf{y}$, where $\ell(\mathbf{x}) = 2k - 2$ and $\ell(\mathbf{y}) = k + 1$. Consider the set $\mathcal{D}$ of all words of the form $\mathbf{0}\mathbf{y}$, where $\mathbf{y}$ has length $k + 1$, all entries in $\mathbf{y}$ are nonzero, and $\mathbf{y}$ does not occur as a suffix of a codeword. There are exactly $(q - 1)^k$ codewords that end in k nonzero symbols, and so $|\mathcal{D}| \geq (q - 1)^{k+1} - (q - 1)^k = (q - 1)^k(q - 2)$. We aim to show that $\mathcal{C}$ cannot be prolific since it cannot have all the words in $\mathcal{D}$ as descendants.

Note that the all-zero word cannot be a parent of any word $\mathbf{d} \in \mathcal{D}$. To see this, note that the all-zero word cannot contribute to any of the last $k + 1$ components of $\mathbf{d}$, and so these $k + 1$ components must come from the other parent. But this would mean that the last $k + 1$ entries of $\mathbf{d}$ would be a suffix of a codeword, contradicting the definition of $\mathcal{D}$.

The MDS property of the code shows that any nonzero codeword in $\mathcal{C}$ has at most $k - 1$ zero entries (for otherwise the codeword would be too close to the all-zero codeword). Since any word $\mathbf{d} \in \mathcal{D}$ begins with $2(k - 1)$ zeros, any pair of parents $\mathbf{c}_1$ and $\mathbf{c}_2$ for $\mathbf{d}$ must each have $k - 1$ zeros in their first $2(k - 1)$ positions, and the positions where the zeros of $\mathbf{c}_1$ occur must be disjoint from the positions where the zeros of $\mathbf{c}_2$ occur. Without loss of generality, we may assume that $\mathbf{c}_1$ has a zero in its first position. There are $\frac{1}{2} \binom{2k-2}{k-1}$ choices for the positions where $\mathbf{c}_1$ is zero; the positions where $\mathbf{c}_2$ is zero are determined by this choice. By the MDS property, there are exactly $q - 1$ choices for a nonzero codeword $\mathbf{c}_1$ with zeros in the specified positions; similarly, there are $q - 1$ choices for $\mathbf{c}_2$. Each pair of codewords $\{\mathbf{c}_1, \mathbf{c}_2\}$ gives rise to at most 2^{k+1} descendants which start with $2k - 2$ zeros. So the pair $\{\mathbf{c}_1, \mathbf{c}_2\}$ can have at most $2^{k+1} - 2$ descendants in $\mathcal{D}$, since no element of $\mathcal{D}$ ends with the suffix

of a codeword. Moreover, the MDS property shows that for every choice of $\mathbf{c}_1$, there is a unique choice for $\mathbf{c}_2$ that agrees with $\mathbf{c}_1$ in its last position. When $\mathbf{c}_2$ is of this form, the pair gives rise to at most $2^k - 2$ descendants in $\mathcal{D}$. Thus $\mathcal{C}$ can have at most

$$\frac{1}{2} \binom{2k-2}{k-1} (q-1) ((q-2)(2^{k+1}-2) + (2^k-2))$$

descendants in $\mathcal{D}$. For $\mathcal{C}$ to be prolific, all words in $\mathcal{D}$ must be descendants, and so

$$\binom{2k-2}{k-1} (q-1) ((q-2)(2^k-1) + (2^{k-1}-1)) \geq |\mathcal{D}| \geq (q-1)^k (q-2).$$

Using the fact that $q \geq n - k + 1 = 2k$, we find that there are no solutions k and q to this inequality, since we are assuming that $k \geq 4$. So there are no prolific codes in this case, as required.

Case 4. $n = 3k - 1$ with $k = 2$. Note that $(5, q, q^2)$ -codes of minimum distance $5 - 1 = 4$ do not exist when $q = 3$, and so we may assume that $q \geq 4$.

A descendant of $\mathcal{C}$ must agree with one of its parents in at least three positions, and so is at distance 2 from this parent. So the set of descendants is contained in the union of the spheres of radius 2 about codewords. We show that these spheres cannot cover all words, and so $\mathcal{C}$ cannot be prolific.

We begin by counting the number of words $\mathbf{d}$ that are in spheres of radius 2 about two codewords $\mathbf{c}_1$ and $\mathbf{c}_2$. Note that since all codewords are at distance at least 4, the word $\mathbf{d}$ has distance exactly 2 from both $\mathbf{c}_1$ and $\mathbf{c}_2$, and the distance from $\mathbf{c}_1$ to $\mathbf{c}_2$ is exactly 4. This implies that the positions where $\mathbf{c}_1$ and $\mathbf{d}$ differ must be disjoint from the positions where $\mathbf{c}_2$ and $\mathbf{d}$ differ, and so no codeword $\mathbf{d}$ lies in more than two spheres of radius 2 about codewords, since we cannot have three pairwise disjoint 2-subsets of a 5-set.

There are $5(q-1)$ codewords at distance 4 from a fixed codeword, and so the number of pairs of codewords at distance 4 is $5q^2(q-1)/2$. Each such pair gives rise to exactly $\binom{4}{2} = 6$ words that lie in spheres of radius 2 about both codewords. So the number of words that are in two spheres of radius 2 is exactly $15q^2(q-1)$ (and no words lie in three or more spheres). Hence the number of words in spheres of radius 2 about $\mathcal{C}$ is

$$q^2(1 + 5(q-1) + 10(q-1)^2) - 15q^2(q-1) = q^2(1 - 10(q-1) + 10(q-1)^2).$$

Since $q \geq 4$, this expression is less than q^5 , and so there are words that are not descendants of the code. Thus $\mathcal{C}$ is not a prolific IPP code, as required.

Case 5. $n = 3k - 1$ and $k = 3$. We may assume that $q \geq n - k + 1 = 6$. But q must be a prime power as $\mathcal{C}$ is linear, and so we may assume that $q \geq 7$. We show the code cannot be prolific by showing that the number of descendants of the code is less than q^8 .

There are $q^3(1 + 8(q-1) + \binom{8}{2}(q-1)^2)$ words within spheres of radius 2 about codewords. All these words are descendants, by the MDS property of the code. It remains to count descendants of distance at least 3 from every codeword.

Let p_i be the number of (unordered) pairs of codewords at distance i . So $p_i = 0$ when $1 \leq i \leq 5$. An upper bound for the number of descendants at distance at least 3 from the code is

$$(3.1) \quad p_6 \binom{6}{3} + p_7 \left(\binom{7}{3} + \binom{7}{4} \right) + p_8 \left(\binom{8}{3} + \binom{8}{4} + \binom{8}{5} \right).$$

Now $p_i = q^3 s_i / 2$, where s_i is the number of codewords of weight i . The weight distribution of any MDS code is known; indeed, we have that

$$\begin{aligned} s_6 &= 28q - 28, \\ s_7 &= 8q^2 - 56q + 48, \\ s_8 &= q^3 - 8q^2 + 28q - 21 \end{aligned}$$

by [12, Theorem 11.6], for example. Substituting these values into (3.1) above, and adding the term which counts descendants at distance at most 2 from the code, we can easily check that the number of descendants is less than q^8 whenever $q \geq 7$, and so $\mathcal{C}$ cannot be prolific in this case.

Case 6. $n \geq 3k$. Consider the set $\mathcal{D}$ of words $\mathbf{0y}$, where $\ell(\mathbf{y}) = k + 1$ and $\mathbf{y}$ is not a suffix of a codeword and contains no zero entries. Note that $|\mathcal{D}| \geq (q - 1)^{k+1} - (q - 1)^k = (q - 1)^k(q - 2) > 0$, and so $\mathcal{D}$ is nonempty. We show that no word in $\mathcal{D}$ can be a descendant of $\mathcal{C}$, and so $\mathcal{C}$ cannot be prolific.

Suppose, for a contradiction, that $\mathbf{d} \in \mathcal{D}$ is a descendant of $\mathcal{C}$. A descendant must agree with one of its parents in at least k of its first $2k - 1$ positions. Since $\mathbf{d}$ begins with $2k - 1$ zeros, one of its parents must have k zero entries and so must be the all-zero codeword. But the final $k + 1$ entries of $\mathbf{d}$ are nonzero, and so the other parent must have contributed them. But this implies that the last $k + 1$ entries of $\mathbf{d}$ are a suffix of this parent, contradicting the definition of $\mathcal{D}$ as required.

From these six cases we deduce that there are no prolific IPP (n, q, q^k) -codes of minimum distance $n - k + 1$ when $n \geq 3$, unless $n = 4$, $q = 3$, and $k = 2$. So the theorem is established. $\square$

We remark that Lemma 3.1 remains true when the assumption that $\mathcal{C}$ is linear is removed; see [7, Theorem 3.1] for details.

THEOREM 3.2. *The only linear nonbinary prolific IPP code of length 3 or more is the $(4, 3, 9)$ -IPP code.*

Proof. Let $\mathcal{C}$ be a prolific $[n, k]$ linear IPP code, where $n \geq 3$. By Lemma 3.1 there are no prolific MDS IPP codes other than the $(4, 3, 9)$ example, so we may assume that $\mathcal{C}$ is not an MDS code. The theorem follows if we can derive a contradiction from this assumption.

Since $\mathcal{C}$ is not MDS, we may permute the columns of the code so that the last k columns of the generator matrix $\mathcal{G}$ form a $k \times k$ matrix with rank $k - 1$.

We can write each codeword in the form $\mathbf{xy}$, where $\ell(\mathbf{x}) = n - k$ and $\ell(\mathbf{y}) = k$. Consider a word $\mathbf{0y}$ where $\mathbf{y}$ has no zeros and is not a suffix of any codeword. A choice for $\mathbf{y}$ certainly exists, since there are at most $(q - 1)^{k-1}$ suffixes of length k of codewords with no zeros, by our condition on the generator matrix $\mathcal{G}$, and so there are at least $(q - 1)^k - (q - 1)^{k-1}$ choices for $\mathbf{y}$.

Since $\mathcal{C}$ is prolific it follows that $\mathbf{0y}$ is a descendant, so there exist codewords $\mathbf{x}_1\mathbf{y}_1, \mathbf{x}_2\mathbf{y}_2 \in \mathcal{C}$ such that $\mathbf{0y} \in \text{desc}(\mathbf{x}_1\mathbf{y}_1, \mathbf{x}_2\mathbf{y}_2)$. Since $\mathbf{y}$ is not a suffix of a codeword, these parents are distinct; moreover, neither parent can be the all-zero codeword. Clearly, $\mathbf{0y}_1 \in \text{desc}(\mathbf{x}_1\mathbf{y}_1, \mathbf{x}_2\mathbf{y}_2)$. The suffix $\mathbf{y}_1$ appears in q codewords of $\mathcal{C}$ by our condition on $\mathcal{G}$, and hence there is a codeword $\mathbf{x}_3\mathbf{y}_1$, where $\mathbf{x}_3 \notin \{\mathbf{x}_1, \mathbf{x}_2\}$. But then we have that $\mathbf{0y}_1 \in \text{desc}(\mathbf{00}, \mathbf{x}_3\mathbf{y}_1)$, which implies that $\mathcal{C}$ is not an IPP code. This contradiction establishes the theorem, as required. $\square$

4. New upper bound on the size of IPP codes. This section establishes a new upper bound on the size of an IPP code, which improves the leading coefficient

in the bound of Theorem 2.4. When the code has length 5, our techniques yield especially good results (and we will need a good bound in section 7). We begin by considering this special case.

LEMMA 4.1. *Let $\mathcal{C}$ be a $(5, q, M)$ -IPP code, where $M > q^2$. Then the minimum distance $d(\mathcal{C})$ of $\mathcal{C}$ is at least 3.*

Proof. Assume, for a contradiction, that $d(\mathcal{C}) \leq 2$. Suppose $\mathbf{c}_1$ and $\mathbf{c}_2$ are codewords at distance 1 or 2. Without loss of generality, assume that $\mathbf{c}_1$ and $\mathbf{c}_2$ agree in their first three positions. Since $M > q^2$, there exist distinct codewords $\mathbf{c}_3$ and $\mathbf{c}_4$ that agree in their final two positions. If the sets $\{\mathbf{c}_1, \mathbf{c}_3\}$ and $\{\mathbf{c}_2, \mathbf{c}_4\}$ are disjoint, then IPP2 is violated. Otherwise, the set $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4\}$ has size 3 and IPP1 is violated. In either case, we have a contradiction, as required. $\square$

Theorem 2.4 implies that for a $(5, q, M)$ -IPP code we have $M \leq 2q^2 + q - 1$. The theorem below significantly improves this bound.

THEOREM 4.2. *If $\mathcal{C}$ is a $(5, q, M)$ -IPP code, then $M < \frac{5}{4}q^2 + 5q$.*

Proof. If there is a symbol $x \in F$ and a position $i \in \{1, 2, 3, 4, 5\}$ such that x occurs just once as the i th position of a codeword, we remove this codeword to produce a smaller code. Repeating this process as often as is necessary, we eventually obtain a code $\mathcal{C}'$ in which no symbol appears exactly once in any fixed position. Note that we have removed at most $5q$ codewords to obtain $\mathcal{C}'$, and so $\mathcal{C}'$ has M' codewords where $M - M' \leq 5q$. To prove the theorem, it suffices to show that $M' < \frac{5}{4}q^2$. The theorem follows trivially when $M' \leq q^2$, and so we may assume that $M' = q^2 + \mu$ for some positive integer μ .

For integers i and j such that $1 \leq i < j \leq 5$, define the subset $S_{ij} \subseteq \mathcal{C}'$ by

$$S_{ij} = \{\mathbf{x} \in \mathcal{C}' : \exists \mathbf{y} \in \mathcal{C}' \setminus \{\mathbf{x}\} \text{ such that } x_i = y_i \text{ and } x_j = y_j\}.$$

Note that $|S_{ij}| > \mu$.

We claim that $S_{ij} \cap S_{i'j'} = \emptyset$ whenever $\{i, j\}$ and $\{i', j'\}$ are disjoint pairs of positions. Without loss of generality, it is sufficient to show that $S_{12} \cap S_{34} = \emptyset$. Suppose, for a contradiction, that $\mathbf{c}_1 \in S_{12} \cap S_{34}$. Writing $\mathbf{c}_1 = x_1x_2x_3x_4x_5$, there exist codewords $\mathbf{c}_2, \mathbf{c}_3 \in \mathcal{C}' \setminus \{\mathbf{c}_1\}$ of the form $\mathbf{c}_2 = x_1x_2**y$ and $\mathbf{c}_3 = **x_3x_4z$. Note that $\mathbf{c}_2 \neq \mathbf{c}_3$, by Lemma 4.1. If $|\{x_5, y, z\}| < 3$, then $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ violates IPP1. If x_5, y and z are distinct, let $\mathbf{c}_4 \in \mathcal{C}'$ be another codeword ending in y (which exists by our choice of $\mathcal{C}'$). Then the sets $\{\mathbf{c}_1, \mathbf{c}_4\}$, $\{\mathbf{c}_2, \mathbf{c}_3\}$ violate IPP2. This contradiction establishes our claim.

For any given disjoint pairs $\{i_1, j_1\}, \{i_2, j_2\} \subseteq \{1, 2, 3, 4, 5\}$ define $Q_{i_1j_1i_2j_2}$ to be the set of symbols that occur in the ℓ th positions of codewords in $S_{i_1j_1}$, where ℓ is the unique position not equal to any of i_1, j_1, i_2, j_2 .

We claim that $Q_{i_1j_1i_2j_2} \cap Q_{i_2j_2i_1j_1} = \emptyset$ for any disjoint pairs $\{i_1, j_1\}$ and $\{i_2, j_2\}$. (In particular, since there are q symbols in total, this claim implies that $|Q_{i_1j_1i_2j_2}| + |Q_{i_2j_2i_1j_1}| \leq q$.) To see why our claim holds, we show (without loss of generality) that $Q_{1234} \cap Q_{3412} = \emptyset$. Assume, for a contradiction, that $x \in Q_{1234} \cap Q_{3412}$. Then the following four codewords lie in $\mathcal{C}'$: $\mathbf{c}_1 = x_1x_2**x$, $\mathbf{c}_2 = x_1x_2***$, $\mathbf{c}_3 = **x_3x_4x$, and $\mathbf{c}_4 = **x_3x_4*$. (These codewords are distinct, since $S_{12} \cap S_{34} = \emptyset$.) But then the pairs $\{\mathbf{c}_1, \mathbf{c}_4\}$ and $\{\mathbf{c}_2, \mathbf{c}_3\}$ violate IPP2, and so our claim follows.

There are $\binom{5}{2}\binom{3}{2} = 30$ subsets $Q_{i_1j_1i_2j_2} \subseteq F$, and the previous paragraph shows that at least half of them are “small” in the sense of satisfying $|Q_{i_1j_1i_2j_2}| \leq \frac{1}{2}q$.

The map from S_{12} to $Q_{1235} \times Q_{1234}$, where $x_1x_2x_3x_4x_5 \mapsto x_4x_5$, is injective, since $S_{12} \cap S_{45} = \emptyset$. Thus $|S_{12}| \leq |Q_{1235}||Q_{1234}|$. Arguing similarly, we find that

$$|S_{12}| \leq \min\{|Q_{1234}||Q_{1235}|, |Q_{1234}||Q_{1245}|, |Q_{1235}||Q_{1245}|\}.$$

A similar inequality exists for any subset S_{ij} . There are in total $\binom{5}{2} = 10$ such inequalities, each involving three subsets $Q_{iji'j'}$. Averaging over all pairs $\{i, j\}$, the expected number of the subsets $Q_{iji'j'}$ in such an inequality which satisfy $|Q_{iji'j'}| \leq \frac{1}{2}q$ is at least $\frac{3}{2}$. So we may find a pair $\{i, j\}$ such that the inequality involves at least two sets $Q_{iji'j'}$ of size at most $\frac{1}{2}q$. But then the inequality implies that $|S_{ij}| \leq (\frac{1}{2}q)^2 = \frac{1}{4}q^2$. Since $\frac{1}{4}q^2 \geq |S_{ij}| > \mu$, we find that $M' = q^2 + \mu < \frac{5}{4}q^2$, as required. $\square$

We comment that, arguing more carefully, it is possible to reduce the term $5q$ in the bound above. Indeed, we have the outline of a proof (with many special cases) that the term can be eliminated. For the sake of simplicity, we content ourselves with proving a bound that eliminates this term in the case of prolific IPP codes.

THEOREM 4.3. *If $\mathcal{C}$ is a $(5, q, M)$ prolific IPP code, then $M < \frac{5}{4}q^2$.*

Proof. Suppose that there is no symbol x that appears just once as the i th position of a codeword. The argument of Theorem 4.2 (where $M = M'$ in our situation) now shows that $M < \frac{5}{4}q^2$. So we may assume that there is a symbol x and a position i such that x appears exactly once as the i th position of a codeword. Replacing $\mathcal{C}$ by an equivalent code if necessary, we may assume (without loss of generality) that $x = 0$, $i = 1$, and $\mathcal{C}$ contains the all-zero word $\mathbf{0}$. So no codeword starts with 0, other than the all-zero codeword.

The word 01111 is a descendant, since $\mathcal{C}$ is prolific. Now, $\mathbf{0}$ is a parent, since no other codeword can contribute to the first position of the descendant. But $\mathbf{0}$ cannot contribute to any of the remaining positions, and so there exists a codeword $\mathbf{c}_1$ of the form $\mathbf{c}_1 = *1111$. Similarly, considering the descendant 01112, there exists a codeword of the form $\mathbf{c}_2 = *1112$. But then $d(\mathbf{c}_1, \mathbf{c}_2) \leq 2$, and so Lemma 4.1 implies that $M \leq q^2 < \frac{5}{4}q^2$, as required. $\square$

We do not see how to generalize the bound of Theorem 4.2, as it is not clear what the analogue of the final paragraph of the proof should be. However, we are able to establish the following theorem.

THEOREM 4.4. *Let $\mathcal{C}$ be an (n, q, M) -IPP code, where $n = 3k - 1$. Then $M < \frac{3}{2}q^k + 3q^{k-1}$.*

Proof. We begin by proving the weaker bound

$$(4.1) \quad M < \frac{3}{2}q^k + \binom{n}{k-1}q^{k-1},$$

and we will then show how our argument can be modified to give the bound of the theorem.

If there are any codewords $\mathbf{c}$ that are uniquely defined by a set of $k - 1$ positions (so there exists a $(k - 1)$ -set X of positions such that $\{\mathbf{u} \in \mathcal{C} : c_i = u_i \text{ for all } i \in X\} = \{\mathbf{u}\}$), we remove them. Repeating this process as often as is necessary, we obtain a code $\mathcal{C}'$ with the property that for all $\mathbf{c} \in \mathcal{C}'$ and for all $(k - 1)$ -sets $X \subseteq \{1, 2, \dots, n\}$ we have that

$$|\{\mathbf{u} \in \mathcal{C}' : c_i = u_i \forall i \in X\}| \geq 2.$$

Note that $\mathcal{C}'$ is an (n, q, M') -code with $M - M' < \binom{n}{k-1}q^{k-1}$. If $M' \leq q^k$, then bound (4.1) holds trivially, and so we may assume that $M' > q^k$. Let μ be the positive integer such that $M' = q^k + \mu$. To show (4.1) holds, it suffices to show that $\mu \leq \frac{1}{2}q^2$.

For a subset $T \subseteq \{1, 2, \dots, n\}$ of k positions, define a subset $S_T \subseteq \mathcal{C}'$ by

$$S_T = \{\mathbf{x} \in \mathcal{C}' : \exists \mathbf{y} \in \mathcal{C}' \setminus \{\mathbf{x}\} \text{ such that } x_i = y_i \text{ for } i \in T\}.$$

Note that $|S_T| > \mu$. We claim that $S_{T_1} \cap S_{T_2} = \emptyset$ whenever T_1 and T_2 are disjoint. To see this, assume (without loss of generality) that $T_1 = \{1, 2, \dots, k\}$ and $T_2 = \{k+1, k+2, \dots, 2k\}$ and suppose (for a contradiction) that $\mathbf{c} \in S_{T_1} \cap S_{T_2}$. We may write $\mathbf{c} = \mathbf{x}\mathbf{y}\mathbf{z}$, where $\ell(\mathbf{x}) = \ell(\mathbf{y}) = k$ and $\ell(\mathbf{z}) = k-1$. Since $\mathbf{c} \in S_{T_1} \cap S_{T_2}$, there exist codewords $\mathbf{c}_2, \mathbf{c}_3 \in \mathcal{C}' \setminus \{\mathbf{c}_1\}$ of the form $\mathbf{c}_2 = \mathbf{x}\mathbf{w}$ and $\mathbf{c}_3 = \mathbf{y}\mathbf{u}$. Suppose that $|\{\mathbf{z}, \mathbf{w}, \mathbf{u}\}| < 3$. Then we find that IPP1 is violated. For if $\mathbf{c}_2 \neq \mathbf{c}_3$, then $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ violates IPP1. Moreover, if $\mathbf{c}_2 = \mathbf{c}_3$ (which implies $\mathbf{z} \neq \mathbf{w}$), then $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_4\}$ violates IPP1, where $\mathbf{c}_4 \in \mathcal{C}$ has $\mathbf{w}$ as a suffix and is distinct from $\mathbf{c}_2$. (Our construction of $\mathcal{C}'$ guarantees that $\mathbf{c}_4$ exists.) Now suppose that $|\{\mathbf{z}, \mathbf{w}, \mathbf{u}\}| = 3$. We find that IPP2 is violated by the sets $\{\mathbf{c}_1, \mathbf{c}_4\}$ and $\{\mathbf{c}_2, \mathbf{c}_3\}$, where $\mathbf{c}_4$ is defined as before. So we have a contradiction, and therefore our claim follows.

Define $S_1 = S_{\{1, 2, \dots, k\}}$ and $S_2 = S_{\{k+1, k+2, \dots, 2k\}}$. For $i = 1, 2$, let P_i be the set of length $k-1$ suffixes of codewords in S_i . We claim that $P_1 \cap P_2 = \emptyset$. To see this, suppose (for a contradiction) that there exists a suffix $\mathbf{z} \in P_1 \cap P_2$. Since $\mathbf{z} \in P_1$, there exist distinct codewords of the form $\mathbf{c}_1 = \mathbf{x}\mathbf{z}$ and $\mathbf{c}_2 = \mathbf{x}\mathbf{**}$. Since $\mathbf{z} \in P_2$, there exist distinct codewords of the form $\mathbf{c}_3 = \mathbf{y}\mathbf{z}$ and $\mathbf{c}_4 = \mathbf{y}\mathbf{*}$. Since $\{\mathbf{c}_1, \mathbf{c}_2\} \subseteq S_1$ and $\{\mathbf{c}_3, \mathbf{c}_4\} \subseteq S_2$, and since $S_1 \cap S_2 = \emptyset$, we find that the codewords $\mathbf{c}_i$ are pairwise distinct. But then the sets $\{\mathbf{c}_1, \mathbf{c}_4\}$ and $\{\mathbf{c}_2, \mathbf{c}_3\}$ violate IPP2. This contradiction shows that $P_1 \cap P_2 = \emptyset$, as required.

Since P_1 and P_2 are disjoint subsets of a set of the q^{k-1} possible suffixes of length $k-1$, we find that $|P_i| \leq \frac{1}{2}q^{k-1}$ for some i . Suppose that $|P_1| \leq \frac{1}{2}q^{k-1}$, so the number of length $k-1$ suffixes of codewords in S_1 is at most $\frac{1}{2}q^{k-1}$. The number of suffixes of length k of codewords in S_1 is therefore at most $\frac{1}{2}q^k$, and the length k suffixes of any two codewords in S_1 are distinct, since $S_1 \cap S_{\{2k, 2k+1, \dots, 3k-1\}} = \emptyset$. Thus $\frac{1}{2}q^k \geq |S_1| > \mu$, and so (4.1) holds in this case. In the case when $|P_2| \leq \frac{1}{2}q^{k-1}$, a similar argument establishes (4.1): instead of suffixes, we consider subwords consisting of the first component and the last $k-1$ components of a word, and we use the fact that $S_2 \cap S_{\{1, 2k+1, 2k+2, \dots, 3k-1\}} = \emptyset$.

It remains to show that the above argument can be tightened in order to establish the theorem.

The argument above uses the fact that $S_{T_1} \cap S_{T_2} = \emptyset$ for a limited range of sets T_1 and T_2 . (Indeed, it uses this equality when $T_1 = \{1, 2, \dots, k\}$ and $T_2 = \{k+1, k+2, \dots, 2k\}$, when $T_1 = \{1, 2, \dots, k\}$ and $T_2 = \{2k, 2k+1, \dots, 3k-1\}$, and when $T_1 = \{k+1, k+2, \dots, 2k\}$ and $T_2 = \{1, 2k+1, 2k+2, \dots, 3k-1\}$.) Because of this, we see that the argument still works when $\mathcal{C}'$ is defined to be a larger subcode, where less than $3q^{k-1}$ codewords have been removed: we remove codewords that are uniquely defined by their positions in X , where $X = \{2k+1, 2k+2, \dots, 3k-1\}$, $X = \{k+1, k+2, \dots, 2k-1\}$, or $X = \{2, 3, \dots, k\}$. This modification establishes the theorem, as required. $\square$

5. Prolific IPP codes of length 3. The goal of this section is to prove Theorem 5.7, which states that there are no nontrivial prolific IPP codes of length 3.

LEMMA 5.1. *Let $\mathcal{C}$ be a nonbinary prolific IPP code of length 3. Then $|\mathcal{C}| > q$, and the minimum distance $d(\mathcal{C})$ of $\mathcal{C}$ is at least 2.*

Proof. If $\mathcal{C}$ contains q or fewer codewords, then the bound of Theorem 2.1 is violated. So $|\mathcal{C}| > q$.

Suppose $\mathcal{C}$ contains codewords $\mathbf{x}$ and $\mathbf{y}$ at distance 1. There exist distinct codewords $\mathbf{u}, \mathbf{v} \in \mathcal{C}$ that agree at the position where $\mathbf{x}$ and $\mathbf{y}$ disagree, since $|\mathcal{C}| > q$. If the codewords $\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v}$ are not distinct (and so form a set of size 3), then IPP1 is violated; if the codewords are distinct, then IPP2 is violated. This contradiction shows that $d(\mathcal{C}) \geq 2$. $\square$

LEMMA 5.2. *There are no nontrivial prolific IPP codes of length 3 when $q \geq 6$.*

Proof. Let $\mathcal{C}$ be a $(3, q, M)$ prolific IPP code. Every symbol must occur at least once at the start of a codeword, since if there are no codewords of the form $x**$, then there are no descendants of the form $x**$, contradicting the fact that $\mathcal{C}$ is prolific. Since $M \leq 3q - 1$ by Theorem 2.4, there is a symbol that occurs (A) exactly once, or (B) exactly twice as the first position of a codeword. Without loss of generality, suppose this value is 0.

Case (A). Without loss of generality, we may assume that $000 \in \mathcal{C}$ and that no other codeword starts with 0. There are $(q - 1)^2$ words $\mathbf{d}$ of the form $0ab$ where $a, b \neq 0$, and all of these must occur as descendants. Now 000 must be a parent of $\mathbf{d}$ (since no other codeword starts with 0) but cannot contribute to the remaining two positions of $\mathbf{d}$ (since $a, b \neq 0$). So the other parent must be a codeword of the form $*ab$. Thus there are at least $(q - 1)^2$ codewords not equal to 000 . But $|\mathcal{C}| \geq (q - 1)^2 + 1$ contradicts Theorem 2.4 since $q \notin \{3, 4\}$.

Case (B). We may assume that $000, 0xy \in \mathcal{C}$ for some $x, y \in \{0, 1, \dots, q - 1\}$ that are not both zero. A similar argument to (A), with the codewords d of the form $0ab$ where $a \notin \{0, x\}$ and $b \notin \{0, y\}$, shows that there are at least $(q - 2)^2$ other codewords, and so $|\mathcal{C}| \geq (q - 2)^2 + 2$ in this case. This contradicts Theorem 2.4 when $q \geq 6$.

Thus, no $(3, q, M)$ prolific IPP codes exist if $q \geq 6$. $\square$

The following lemma is a special case of a result of Tô and Safavi-Naini [14, Theorem 34]; see [7, Lemma 6.3] for an elementary proof.

LEMMA 5.3. *A 3-ary IPP code $\mathcal{C}$ of length 3 must have $|\mathcal{C}| \leq 4$.*

COROLLARY 5.4. *There is no 3-ary prolific IPP code of length 3.*

Proof. Suppose a 3-ary prolific IPP code $\mathcal{C}$ of length 3 exists. All symbols must occur as the start of a codeword, since $\mathcal{C}$ is prolific. Since $|\mathcal{C}| \leq 4$, there is a symbol that starts a unique codeword. So we are in Case (A) of the proof of Lemma 5.2. The argument there shows that $|\mathcal{C}| \geq 1 + (q - 1)^2 = 5$, and this contradicts Lemma 5.3, as required. $\square$

LEMMA 5.5. *There is no 4-ary prolific IPP code of length 3.*

Proof. Suppose, for a contradiction, that a symbol occurs exactly once as the start of a codeword. Without loss of generality, $000 \in \mathcal{C}$ and no other codeword starts with 0. Considering parents of the nine descendants of the form $0xy$, where $x, y \neq 0$, we see that there are codewords of the form $*xy$ for all x, y ; moreover, these codewords cannot start with 0, since 0 starts a unique codeword. But these nine codewords now form a 3-ary IPP code, and this contradicts Lemma 5.3.

Using the argument above on the second and third positions of $\mathcal{C}$, we may assume all symbols occur at least twice in every position in the code.

Choose a codeword $\mathbf{x}$. Choose a codeword $\mathbf{y} \neq \mathbf{x}$ such that $x_1 = y_1$. Choose a codeword $\mathbf{z} \neq \mathbf{y}$ such that $y_2 = z_2$. Choose a codeword $\mathbf{w} \neq \mathbf{z}$ such that $z_3 = w_3$. Since $\mathcal{C}$ has minimum distance 2, we find that $\mathbf{x} \neq \mathbf{z}$ and $\mathbf{y} \neq \mathbf{w}$. There are two cases as follows: If $\mathbf{x} \neq \mathbf{w}$, then the pairs $\{\mathbf{x}, \mathbf{z}\}$, and $\{\mathbf{y}, \mathbf{w}\}$ violate IPP2. If $\mathbf{x} = \mathbf{w}$, then $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ violates IPP1. We have produced a contradiction, and so the lemma follows. $\square$

LEMMA 5.6. *There is no 5-ary prolific IPP code of length 3.*

Proof. If such code $\mathcal{C}$ exists, then we cannot be in Case (A) in the proof of Lemma 5.2 (as $1 + (q - 1)^2 \geq 3q$, contradicting Theorem 2.4). So we may assume that for all positions i and symbols a there are at least two codewords equal to a in position i . But now the argument in the final paragraph of the proof of Lemma 5.5 shows that $\mathcal{C}$ cannot be an IPP code. This contradiction establishes the lemma. $\square$

The above results together show the following.

THEOREM 5.7. *There are no nonbinary prolific codes of length 3.*

6. Prolific IPP codes of length 4. This section aims to prove Theorem 6.5, which states that there are no nontrivial examples of prolific codes of length 4.

LEMMA 6.1. *Let $\mathcal{C}$ be a nonbinary prolific IPP code of length 4. Then the minimum distance $d(\mathcal{C})$ of $\mathcal{C}$ is at least 3.*

Proof. Suppose $\mathcal{C}$ has M codewords. Theorem 2.1 shows that $\binom{M}{2}2^4 \geq q^4$. If $M \leq q$, we have that $q^22^3 \geq q^4$, which implies that $q \leq 2^{3/2} < 3$, a contradiction. So we may assume that $M > q$.

Suppose, for a contradiction, that $d(\mathcal{C}) = 1$. Without loss of generality, we may assume that $0000, 0001 \in \mathcal{C}$. If there is another codeword whose final symbol is 0 or 1, then IPP1 is violated. If this is not the case, choose distinct codewords $\mathbf{c}_1$ and $\mathbf{c}_2$ that agree in their final position. (Such codewords exist since $|\mathcal{C}| > q$.) Then $\{0000, \mathbf{c}_1\}$ and $\{0001, \mathbf{c}_2\}$ violate IPP2. This contradiction shows that $d(\mathcal{C}) \geq 2$.

We claim that every symbol must occur at least twice in any position of the code. (The prolific property shows that every symbol must occur at least once.) For assume that a symbol, 0 say, occurs exactly once as the start of a codeword. Without loss of generality, we may assume that $0000 \in \mathcal{C}$. For $x, y, z \in F \setminus \{0\}$, the descendant $0xyz$ must have 0000 and $*xyz$ as parents. So there are at least $(q-1)^3$ codewords of the form $*xyz$, where x, y , and z are nonzero. None of these codewords can start with 0, and so we have a collection $\mathcal{C}'$ of $(q-1)^3$ codewords over an alphabet of size $q-1$. Since $(q-1)^3 > (q-1)^2$, we can find distinct codewords $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}'$ that agree in their first two positions. There are $q-2$ codewords in $\mathcal{C}' \setminus \{\mathbf{c}_2\}$ that agree with $\mathbf{c}_2$ in its last two positions: pick $\mathbf{c}_3$ of this form. Then $\mathbf{c}_1$ and $\mathbf{c}_3$ have $\mathbf{c}_2$ as their descendant. This contradiction establishes our claim.

Now suppose, for a contradiction, that we can find two distinct codewords $\mathbf{c}_1, \mathbf{c}_2$ that are at distance 2. Without loss of generality, we may take $\mathbf{c}_1 = 0000, \mathbf{c}_2 = 0011$. Let $\mathbf{c}_3 \in \mathcal{C} \setminus \{\mathbf{c}_1\}$ be of the form $**0*$. Similarly let $\mathbf{c}_4 \in \mathcal{C} \setminus \{\mathbf{c}_2\}$ be a codeword of the form $***1$. If $\mathbf{c}_3 = \mathbf{c}_4$, then the set $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ violates IPP1. But if $\mathbf{c}_3 \neq \mathbf{c}_4$, then $\{\mathbf{c}_1, \mathbf{c}_4\}$ and $\{\mathbf{c}_2, \mathbf{c}_3\}$ violate IPP2. This contradiction shows that there are no pairs of codewords at distance less than 3, and so the lemma follows. $\square$

LEMMA 6.2. *Let $\mathcal{C}$ be a q -ary prolific IPP code of length 4, and let $i \in \{1, 2, 3, 4\}$. Then every symbol occurs in the i th position of either $q-1$ or q codewords.*

Proof. Without loss of generality, we may assume that $i = 1$.

Since $d(\mathcal{C}) \geq 3$, the first two positions of a codeword uniquely determine that codeword. So each symbol occurs at most q times as the start of a codeword (as there are q pairs starting with this symbol).

Suppose a symbol, 0 say, occurs less than q times as the start of a codeword. Then there exist symbols $x, y, z \in F$ such that there are no codewords of the form $0x**$, $0*y*$, or $0**z$. Without loss of generality, assume that $x = y = z = 0$.

Consider the descendant 0000 . One parent must start with 0, but this codeword cannot contribute to any of the remaining positions in the descendant. So there is a codeword of the form $w000$ for some w (and w is clearly nonzero). Since $d(\mathcal{C}) = 3$, we see that $w000$ is the unique codeword of the form $*00*$.

Now consider the descendant $000a$ where $a \neq 0$. One parent starts with 0 and this parent cannot contribute to the middle two positions. Hence $w000$ is the other parent. But $w000$ cannot contribute to the last position, and so there must be a codeword of the form $0**a$. Since we have $q-1$ choices for a , the symbol 0 occurs at least $q-1$ times as the start of a codeword. This proves the lemma. $\square$

LEMMA 6.3. *There is no q -ary prolific IPP code of length 4 when $q > 4$.*

Proof. Let C be a q -ary prolific IPP code of length 4. Consider the words starting with 0. By Lemma 6.2 there are at least $q - 1$ such words and no two of these words can agree in any position other than the first (since $d(C) = 3$). So, without loss of generality, we may assume the codewords starting with 0 are $0111, 0222, \dots, 0(q - 1)(q - 1)(q - 1)$, and possibly 0000 .

Fix a symbol $\ell \in F$. Choose $m \in F$ such that $\ell \neq m$ and there exists a codeword of the form $*\ell m*$. Such a choice for m exists; indeed, by Lemma 6.2 there are at least $(q - 1) - 1$ choices for m . Choose $n \in F$ such that for $\ell \neq n, m \neq n$, there exists a codeword of the form $**mn$ but there does not exist a codeword of the form $*\ell mn$. There are at least $q - 1$ choices for n such that there is a codeword of the form $**mn$, and at most three of these choices are ruled out by the other conditions we place on n . Since $q - 1 - 3 \geq 1$, there exists at least one choice for n .

Consider the descendant $0\ell mn$. This is a descendant of $0\ell\ell\ell$ and the codeword $**mn$, and of $0nnn$ and $*\ell m*$. Our choice of ℓ, m, n means that these sets of parents are disjoint. So we get a contradiction to IPP2, as required. $\square$

LEMMA 6.4. *There is no 4-ary prolific IPP code of length 4.*

Proof. The proof of Lemma 6.3 works (with $q - 1$ replaced by q at various points) when every symbol occurs four times in every position of a codeword. So we know that if a 4-ary prolific IPP code C of length 4 exists, then $12 = q(q - 1) \leq |C| < q^2 = 16$. We now argue that no such code can exist.

Let $M = |C|$. There are $M(1 + 4 \times 3) = 13M$ words at distance at most 1 from C . All the remaining words must be descendants, and indeed they must be descendants of a unique pair of codewords at distance 4. (Codewords at distance 3 can never produce descendants of distance more than 1 from the code.)

A pair of codewords at distance 4 produce exactly six descendants at distance 2 from the code. So $4^4 = 6M_4 + 13M$, where M_4 is the number of pairs of codewords at distance 4. Therefore

$$M_4 = \frac{256 - 13M}{6},$$

and so $256 - 13M \equiv 0 \pmod{6}$ and so $M \equiv 4 \pmod{6}$. But this cannot happen, as no number M such that $12 \leq M \leq 15$ is such that $M \equiv 4 \pmod{6}$. This proves the theorem. $\square$

THEOREM 6.5. *If C is a nonbinary prolific IPP code of length 4, then C is equivalent to the $(4, 3, 9)$ -code given in the introduction.*

Proof. Let C be a q -ary prolific IPP code of length 4, where $q > 2$. Lemmas 6.3 and 6.4 show that we must, in fact, have that $q = 3$.

Lemma 6.1 shows that $d(C) \geq 3$, and so the Singleton bound shows that $|C| \leq 9$. Moreover, we know from Lemma 6.2 that every symbol occurs at least twice in each position of the code.

We claim that $|C| = 9$. Suppose, for a contradiction, that $|C| \leq 8$ and so some symbol occurs twice at the start of a codeword. Without loss of generality, we may assume that 0 occurs just twice at the start of a codeword, and that $0000, 0111 \in C$. Since 0222 is a descendant of the code, we must have a codeword of the form $x222$ where $x \in \{1, 2\}$.

The descendant 0012 must either have parents 0000 and $**12$, or 0111 and $*0*2$. Since $d(C) \geq 3$, any word of the form $**12$ or $*0*2$ must in fact be of the form $*012$ (for otherwise there would be a codeword too close to 0000 , 0111 , or $x222$), and so we may deduce that C contains a codeword of the form $*012$. Indeed, if $a, b, c \in F$ are

distinct, we may argue in the same way (using the descendant $0abc$) that there exists a codeword of the form $*abc$. These six codewords, together with the codewords 0000, 0111, and $x222$ show that $|\mathcal{C}| \geq 9$, contradicting our assumption that $|\mathcal{C}| < 9$.

So we may assume that $|\mathcal{C}| = 9$. But this implies that $\mathcal{C}$ is an MDS code, and so is equivalent to the $(4, 3, 9)$ -code from the introduction. $\square$

7. Prolific IPP codes of length 5. This section provides a sketch proof of the following theorem. Full details can be found in [7, section 8].

THEOREM 7.1. *There are no nonbinary prolific IPP codes of length 5.*

Let $\mathcal{C}$ be a nonbinary prolific IPP code of length 5. We sketch a method for deriving a contradiction by breaking our argument up into a proof of various statements as follows:

1. *Each symbol in F appears at least twice in every coordinate of $\mathcal{C}$.* If a symbol occurs exactly once in some position, the argument in Case (A) of the proof of Lemma 5.2 shows that $|\mathcal{C}| \geq (q-1)^4 + 1$. This contradicts the upper bound of Theorem 4.3.

2. *The minimum distance of $\mathcal{C}$ is 3.* The difficult case is showing that $d(\mathcal{C}) \neq 4$. When $d(\mathcal{C}) = 4$, the Singleton bound shows that $|\mathcal{C}| \leq q^2$; the prolific property shows that all q^2 pairs of symbols must occur as prefixes of codewords. But then $\mathcal{C}$ is a $(5, q, q^2)$ -code, which cannot be a prolific IPP code by the nonlinear version of Lemma 3.1 (see the remark following the proof of the lemma).

3. *Let $\{i, j, k, \ell, m\} = \{1, 2, 3, 4, 5\}$. Let distinct codewords $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}$ agree in positions i and j . Then $\mathbf{c}_1$ does not agree with any other codeword $\mathbf{c}_3$ in positions k and ℓ .* Statement 1 implies that there exists another codeword $\mathbf{c}_4$ that agrees with $\mathbf{c}_1$ in position m . But then these four codewords provide a counterexample to the IPP property.

4. *Suppose that there exist codewords $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4$ of the form $\mathbf{c}_1 = x_1x_2a**$, $\mathbf{c}_2 = x_1x_2***$, $\mathbf{c}_3 = **bx_4x_5$, and $\mathbf{c}_4 = ***x_4x_5$. Suppose that $\mathbf{c}_1 \neq \mathbf{c}_2$ and $\mathbf{c}_3 \neq \mathbf{c}_4$. Then $a \neq b$.* This follows easily from statement 3.

5. *No 3-ary prolific IPP code of length 5 exists.* Without loss of generality, we may assume that $00000, 00111 \in \mathcal{C}$. By statement 4, no pair of codewords can agree on two of the last three of their coordinates. In particular, $|\mathcal{C}| \leq 9$. But this also implies that any descendant $ab000$ with a and b nonzero must have a parent of the form $ab***$. This argument may be extended to show that all nine possible prefixes of length 2 must occur at the start of codewords. Thus $|\mathcal{C}| \geq 10$ and we have our contradiction. This argument is also useful in the case when $q > 3$ and can be used to establish the following two statements:

6. *Suppose (without loss of generality) that $00000, 00111 \in \mathcal{C}$. Then there exists a pair of codewords which agree in positions i and j , where $\{i, j\} \subseteq \{3, 4, 5\}$.*

7. *Suppose there are two codewords $\mathbf{c}, \mathbf{c}' \in \mathcal{C}$ that agree in positions i and j . Then for all $a, b \in F$ with $a \neq c_i$, $b \neq c_j$ there exists a codeword whose i th and j th positions are equal to a and b , respectively.*

8. *No set of $q-1$ codewords can pairwise agree in any fixed set of two positions.* Suppose a set X of $q-1$ codewords exists that agree in their (say) first two positions. By statement 2, all their remaining positions differ. But then we get a contradiction by statement 4 if there are two codewords that agree in two of their last three positions, and we get a contradiction by statement 6 if not.

9. *Every symbol occurs at least q times in each coordinate of $\mathcal{C}$.* Suppose, for example, there are less than q codewords starting with 0. Let $a, b, c, d \in F$ be such that there are no codewords of the form $0a***$, $0*b**$, $0**c*$, or $0***d$. For any $z \neq d$,

there must be a codeword of the form $0***z$ (consider the parents of $0abcz$). So there are exactly $q - 1$ codewords starting with 0, and all their last coordinates differ. By statement 8, there exist $u, v, w \in F$ such that there are no codewords of the form $*abu*$, $*abv*$, or $*abw*$. Consider parents of $0**uw$ and $0**vw$ to show that there exist distinct codewords starting with 0 and agreeing in their last position, giving us the contradiction we need.

10. *Assume that $00000, 00111 \in \mathcal{C}$ and that there are two codewords that agree in their last two positions. Then the nonzero codewords of the form $**0**$ must be of the form $**01a$ or $**0a1$ for some $a \in F \setminus \{0, 1\}$.* Using statement 3, it is not difficult to reduce this statement to proving that the word $\mathbf{c} = **022$ is not a codeword. If $\mathbf{c} \in \mathcal{C}$, then by statement 4, $\mathbf{c}$ is the only codeword of the form $***22$; by statement 3, 00111 is the only codeword of the form $***11$. So by statement 7, there exists $\mathbf{c}' \in \mathcal{C}$ of the form $***21$ or $***12$. But then the IPP property is violated, giving us the contradiction we need.

11. *Let $i, j \in \{1, 2, \dots, 5\}$ be a pair of positions. Then there cannot exist a set of three codewords that pairwise agree in position i and position j .* Without loss of generality, assume that $00000, 00111, 00222 \in \mathcal{C}$. By statement 6, we may assume that there is a pair of codewords that agree in their last two positions. Statement 10 implies that codewords of the form $**0**$ have the form $**01a$ or $**0a1$ where $a \geq 2$. Applying statement 10 with 1 replaced by 2, these codewords also have the form $**02b$ or $**0b2$ where $b \notin \{0, 2\}$. Since $d(\mathcal{C}) = 3$ and $q > 3$, this implies there are at most three codewords of the form $**0**$, contradicting statement 9.

12. *No q -ary length 5 prolific IPP code exists when $q \geq 5$.* Statements 10 and 11 imply that there are at most five codewords of the form $**0**$, and without loss of generality these have the forms $00000, **012, **013, **021$, and $**031$. Statement 11 and the fact that $\mathcal{C}$ is prolific imply that there are either one or two codewords of the form $***44$. So there are at least six choices for $a, b \in F$ such that there are no codewords of the form $a***44$ or $*b***44$. But then at least one of these choices has the property that there are no codewords of the form $ab0**$. This is a contradiction, as $ab044$ is not a descendant of $\mathcal{C}$.

13. *A 4-ary prolific IPP code $\mathcal{C}$ of length 5 has exactly 16 codewords.* Statement 9 implies that $|\mathcal{C}| \geq 16$. If no two codewords agree in a fixed pair of positions $|\mathcal{C}| \leq 16$ trivially, so we may assume that this does not happen. This implies that two pairs of symbols cannot appear more than once as the start of a codeword; for either we get a contradiction to the IPP property or we violate statement 4. So, by statement 11 we find $|\mathcal{C}| \leq 17$. When $|\mathcal{C}| = 17$, a short argument using statement 10 shows that two pairs of codewords must agree in their third and fourth positions, giving the contradiction we seek.

14. *No 4-ary prolific IPP code of length 5 exists.* By statement 13, $|\mathcal{C}| = 16$, and by statement 9 every symbol occurs exactly four times in any coordinate of the code. Assume $00000, 00111 \in \mathcal{C}$. Using statement 10 several times, we see that $\mathcal{C}$ must (without loss of generality) consist of codewords of the forms given in Table 7.1, where the last three positions of the final five codewords do not involve 0 or 1. There are $x, y \in F$ such that there are no codewords of the form $0x***$ or $y0***$. A codeword $\mathbf{c}$ exists of the form $**021, **210$, or $**102$ which is not of the form $xy***$. By statement 11, $\mathbf{c}$ is not of the form $00***$. But then $0xc_3c_4c_5$ and $y0c_3c_4c_5$ cannot both be descendants of the code, and we have a contradiction, as required.

8. Conclusion and open problems. As we stated in the introduction, we conjecture that there are no examples of prolific IPP codes other than those listed in

TABLE 7.1
Structure of the 4-ary code.

0	0	0	0	0
0	0	1	1	1
*	*	0	1	2
*	*	0	2	1
*	*	0	1	3
*	*	2	1	0
*	*	1	2	0
*	*	1	3	0
*	*	1	0	2
*	*	2	0	1
*	*	3	0	1
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*
*	*	*	*	*

the introduction. We begin this section by discussing how close we are to proving this conjecture.

When $n \geq 3$ but $n \neq 4$, the lower bound of Theorem 2.2 and the upper bound due to Hollmann et al. (Theorem 2.4) together imply that for any fixed length n , there are no q -ary prolific IPP codes of length n provided that q is sufficiently large. When $n = 4$, the same is true if an improved, but less explicit, bound due to Alon, Fischer, and Szegedy [2] is used; or we may deduce this from Theorem 6.5. Sadly, we are not able to bring the number of open parameters n and q (where it is not known whether a prolific q -ary IPP code of length n exists) down to a finite number. In particular, when q is fixed and $3 \leq q \leq 8$, all the parameters n and q are open for $n \geq 6$. Table 8.1 lists the parameters for which the existence of a prolific (n, q, M) -IPP code is as yet undetermined. For the values of n listed in the table, a nonbinary prolific (n, q, M) -IPP code might exist for $3 \leq q \leq q_{max}$, and for $n > 68$, a nonbinary prolific (n, q, M) -IPP code might exist for $3 \leq q \leq 8$. We used the lower bound of Theorem 2.2 with $k = \lceil \frac{n}{3} \rceil$. We used the upper bound of Theorem 4.4 when $n \equiv 2 \pmod 3$, and the upper bound of Theorem 2.4 otherwise.

As a step towards proving the conjecture, is it possible to show that no q -ary prolific IPP code of length n exists for all sufficiently large n , where q is a fixed integer with $3 \leq q \leq 8$?

We remark that the notion of a prolific code makes sense for any class of code which has a natural notion of a descendant. In particular, are there non-trivial examples of prolific k -IPP codes where $k > 2$? (See Staddon, Stinson, and Wei [13] or Blackburn [6] for the definition of a k -IPP code.)

Is it possible to improve our upper bounds on IPP codes (Theorems 4.2 and 4.4) significantly? We believe that the constants in the leading terms of these upper bounds can be reduced. Indeed, it might be possible to prove more than this. Let n be fixed, and suppose that 3 does not divide n . Let ϵ be a positive constant. Is it the case that a q -ary IPP code $\mathcal{C}$ of length n must satisfy $|\mathcal{C}| \leq \epsilon q^{\lceil n/3 \rceil}$ when q is sufficiently large? This is true when $n = 4$, by a bound of Alon, Fischer, and Szegedy [2, Theorem 2.5].

Is it possible to generalize the techniques of Theorems 4.2 and 4.4 to provide better upper bounds for k -IPP codes when $k > 2$? We have established new bounds on 3-IPP codes using these techniques. We hope that these bounds will form the subject of a future paper, but we also hope that our techniques can be stretched further.

TABLE 8.1
Open parameters for prolific IPP codes.

n	q_{max}	n	q_{max}	n	q_{max}
6	11	7	34	8	13
9	10	10	18	11	12
12	10	13	14	14	11
15	10	16	13	17	10
18	9	19	12	20	10
21	9	22	11	23	10
24	9	25	10	26	9
27	9	28	10	29	9
30	9	31	10	32	9
33	9	34	10	35	9
36	9	37	9	38	9
39	8	40	9	41	9
42	8	43	9	44	9
45	8	46	9	47	9
48	8	49	9	50	8
51	8	52	9	53	8
54	8	55	9	56	8
57	8	58	9	59	8
60	8	61	9	62	8
63	8	64	9	65	8
66	8	67	8	68	8

Acknowledgment. The authors would like to thank Noga Alon for some helpful remarks.

REFERENCES

- [1] N. ALON, G. COHEN, M. KRIVELEVICH, AND S. LITSYN, *Generalised hashing and parent identifying codes*, J. Combin. Theory Ser. A, 104 (2003), pp. 207–115.
- [2] N. ALON, E. FISCHER, AND M. SZEGEDY, *Parent-identifying codes*, J. Combin. Theory Ser. A, 95 (2001), pp. 349–359.
- [3] N. ALON AND U. STAV, *New bounds on parent-identifying codes: The case of multiple parents*, Combin. Probab. Comput., 13 (2004), pp. 795–807.
- [4] A. BARG, G. COHEN, S. ENCHEVA, G. KABATIANSKY, AND G. ZEMOR, *A hypergraph approach to the identifying parent property: The case of multiple parents*, SIAM J. Discrete Math., 14 (2001), pp. 423–431.
- [5] A. BARG AND G. KABATIANSKY, *A class of IPP codes with efficient identification*, J. Complexity, 20 (2004), pp. 137–147.
- [6] S. R. BLACKBURN, *An upper bound on the size of a code with the k -identifiable parent property*, J. Combin. Theory Ser. A, 102 (2003), pp. 179–185.
- [7] S. R. BLACKBURN, T. ETZION, AND S.-L. NG, *Prolific Codes with the Identifiable Parent Property*, Report 2007/276, Cryptology ePrint Archive, available online at <http://eprint.iacr.org/2007/276.pdf> (2007).
- [8] R. HILL, *A First Course in Coding Theory*, Oxford University Press, Oxford, UK, 1986.
- [9] H. D. L. HOLLMANN, J. H. VAN LINT, J.-P. LINNARTZ, AND L. M. G. M. TOLHUIZEN, *On codes with the identifiable parent property*, J. Combin. Theory Ser. A, 82 (1998), pp. 121–133.
- [10] T. LINDKVIST, J. LÖFVENBERG, AND M. SVANSTRÖM, *A class of traceability codes*, IEEE Trans. Inform. Theory, 48 (2002), pp. 2094–2096.
- [11] J. LÖFVENBERG, *Binary fingerprinting codes*, Des. Codes Cryptogr., 36 (2005), pp. 69–81.
- [12] F. J. MACWILLIAMS AND N. J. A. SLOANE, *The Theory of Error Correcting Codes*, North-Holland, Amsterdam, 1977.
- [13] J. N. STADDON, D. R. STINSON, AND R. WEI, *Combinatorial properties of frameproof and traceability codes*, IEEE Trans. Inform. Theory, 47 (2001), pp. 1042–1049.

- [14] V. D. TÔ AND R. SAFAVI-NAINI, *On the maximal codes of length 3 with the 2-identifiable parent property*, SIAM J. Discrete Math., 17 (2004), pp. 548–570.
- [15] T. VAN TRUNG AND S. MARTIROSYAN, *New constructions for IPP codes*, Des. Codes Cryptog., 35 (2005), pp. 227–239.
- [16] Y. YEMANE, *Codes with the k -Identifiable Parent Property*, Ph.D. Thesis, Department of Mathematics, Royal Holloway, University of London, 2002.

THE INDEPENDENT EVEN FACTOR PROBLEM*

SATORU IWATA[†] AND KENJIRO TAKAZAWA[‡]

Abstract. This paper deals with the independent even factor problem. For odd-cycle-symmetric digraphs, in which each arc in any odd dicycle has the reverse arc, a min-max formula is established as a common generalization of the Tutte–Berge formula for matchings and the min-max formula of Edmonds [*Submodular functions, matroids, and certain polyhedra*, in *Combinatorial Structures and Their Applications*, R. Guy et al., eds., Gordon and Breach, New York, 1970, pp. 69–87] for matroid intersection. We devise a combinatorial efficient algorithm to find a maximum independent even factor in an odd-cycle-symmetric digraph accompanied by general matroids, which commonly extends two of the alternating-path-type algorithms, the even factor algorithm of Pap [*Math. Program.*, 110 (2007), pp. 57–69], and the matroid intersection algorithms. This algorithm gives a proof of the min-max formula and contains a new operation on matroids, which corresponds to shrinking factor-critical components in the matching algorithm of Edmonds [*Canad. J. Math.*, 17 (1965), pp. 449–467]. The running time of the algorithm is $O(n^4Q)$, where n is the number of vertices and Q is the time for an independence test. The algorithm also gives a common generalization of the Edmonds–Gallai decomposition for matchings and the principal partition for matroid intersection.

Key words. matching, even factor, matroid intersection, polynomial-time algorithm

AMS subject classifications. 05C70, 05C85, 52B40

DOI. 10.1137/060657066

1. Introduction. One of the most fundamental problems in combinatorial optimization is the matching problem. Tutte [24] characterized graphs that have a perfect matching. Then, Berge [1] showed that Tutte’s characterization implies a min-max formula, now called the Tutte–Berge formula. Edmonds [5] gave a combinatorial algorithm to find a maximum matching. This algorithm implies a structural result, which was also found independently by Gallai [9], and is now called the Edmonds–Gallai decomposition.

The matroid intersection problem is also of importance. Edmonds [6] presented a min-max formula for that problem. Polynomial-time algorithms for finding a maximum common independent set are known, such as those of Edmonds [7] and Lawler [15].

As a common generalization of the matching and matroid intersection problems, Cunningham and Geelen [3] introduced the notion of *independent path-matchings*. They showed a min-max formula, a totally dual integral polyhedral description, and polynomial-time solvability of the independent path-matching problem, which generalize those in the matching and matroid intersection problems. The independent path-matching problem with free matroids is called the *path-matching problem*. For the path-matching problem, Frank and Szegő [8] simplified the min-max formula of Cunningham and Geelen [3] and provided its combinatorial proof. Spille and Szegő [21] showed that the path-matching problem has a property which generalizes the Edmonds–Gallai structure for matchings.

*Received by the editors April 12, 2006; accepted for publication (in revised form) April 20, 2008; published electronically September 11, 2008.

<http://www.siam.org/journals/sidma/22-4/65706.html>

[†]Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan (iwata@kurims.kyoto-u.ac.jp). This author is supported by Grant-in-Aid for Scientific Research for Japan Society for the Promotion of Science (JSPS).

[‡]Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan, and Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan (takazawa@misojiro.t.u-tokyo.ac.jp). This author is supported by Grant-in-Aid for JSPS Fellows.

The proof of polynomial-time solvability in Cunningham and Geelen [3] relies on the ellipsoid method. A combinatorial polynomial-time algorithm for finding an optimal path-matching had been of interest. Spille and Weismantel [22] proposed generalizing the combinatorial algorithms for the matching problem [5] and the matroid intersection problem [7, 15] to independent path-matchings.

For the sake of a combinatorial approach to path-matchings, Cunningham and Geelen [4] introduced the *even factor problem* in digraphs, which is yet a generalization of the path-matching problem. While finding a maximum even factor is NP-hard in general digraphs, they showed that the problem is solvable in polynomial time in *weakly symmetric* digraphs with an algebraic approach. A digraph is said to be weakly symmetric if every arc in any dicycle has the reverse arc. In addition, they presented a min-max formula and generalized the Edmonds–Gallai decomposition to the even factor problem in weakly symmetric digraphs, followed by simplifications of Pap and Szegő [18]. These properties are known to hold in a broader class of digraphs, called *odd-cycle-symmetric*. A digraph is odd-cycle-symmetric if each arc in any odd dicycle has the reverse arc. Pap [17] devised a combinatorial algorithm to find a maximum even factor in odd-cycle-symmetric digraphs, which provides an algorithmic proof of the min-max formula and the Edmonds–Gallai-type structure. This algorithm has some properties similar to Edmonds’ matching algorithm [5].

As for the weighted even factor problem, Király and Makai [14] considered odd-cycle-symmetric weighted digraphs, which are odd-cycle-symmetric digraphs accompanied by a weight vector such that the total weight of edges in an odd cycle is equal to that of the reverse cycle. For odd-cycle-symmetric weighted digraphs, a combinatorial algorithm to find a maximum weight even factor is given by Takazawa [23].

In fact, as mentioned above, the even factor problem leaves matroids out of consideration, or in other words, it deals only with free matroids. For a common generalization of matchings and matroid intersection, Cunningham and Geelen [4] proposed the notion of *basic even factors*. They showed a polynomial reduction of the basic even factor problem in weakly symmetric digraphs to the matroid intersection problem. In the resulting algorithm, one needs to compute the cardinality of a maximum even factor in a certain digraph for each independence test.

In this paper, we deal with the *independent even factor problem*, which is a variant of the basic even factor problem. Extending the approach of Cunningham and Geelen [4], we obtain a min-max formula for the independent even factor problem in odd-cycle-symmetric digraphs, which commonly extends the Tutte–Berge formula for matchings and the min-max formula for matroid intersection of Edmonds [6]. For the independent even factor problem in odd-cycle-symmetric digraphs accompanied by general matroids, we devise a combinatorial algorithm to find a maximum independent even factor. The algorithm commonly extends the even factor algorithm [17] and the matroid intersection algorithms [7, 15], in both of which the notion of *alternating-paths* plays an essential part. The algorithm runs in $O(n^4Q)$ time, where n is the cardinality of the vertex set of the digraph and Q is the time needed to test if a given set is independent. By this algorithm, we obtain a different proof of the min-max formula.

The algorithm also gives a common extension of the Edmonds–Gallai decomposition for matchings and the principal partition for matroid intersection (cf. [12, 13]). Unlike the existing theorems on the Edmonds–Gallai-type structure for path-matchings or even factors, we treat the two matroids equally to achieve a unified decomposition principle including the principal partition for matroid intersection.

The maximum independent even factor algorithm includes a new operation on matroids, called *shrinking*. For a matroid $\mathbf{M}$ and its independent set I , eliminate I and create a new element i to obtain a new matroid. We call this procedure *shrinking* I . Note that the resulting matroid differs from the one derived by contraction or deletion. This procedure corresponds to shrinking a factor-critical component in the matching algorithm of Edmonds [5] and the even factor algorithm of Pap [17].

Recently, Harvey [10, 11] gave a new algebraic characterization of basic path-matchings, which achieved a maximum basic path-matching algorithm. His approach needs the matroids to be linearly represented over the same field. The algorithm runs in $O(n^\omega)$ time, where n is the number of vertices and ω is the exponent for matrix multiplication, bounded by $\omega < 2.38$ (due to Coppersmith and Winograd [2]). One can easily extend this approach to obtain an $O(n^\omega)$ algorithm for the independent even factor problem in which the matroids are linearly represented over the same field. Our algorithm, however, does not need the assumption, that is, it can deal with general matroids.

Before closing this section, let us give some notations and definitions used in the following sections.

A digraph G with the vertex set V and the arc set A is denoted by $G = (V, A)$. For vertices u and v , let (u, v) denote an arc which starts in u and ends in v . For an arc $a = (u, v)$, let $\partial^+ a$ (resp., $\partial^- a$) denote the initial (resp., terminal) vertex of a , that is, $\partial^+ a = u$ and $\partial^- a = v$. For a vertex v , define $\delta^+ v = \{a \mid a \in A, \partial^+ a = v\}$ and $\delta^- v = \{a \mid a \in A, \partial^- a = v\}$. For an arc set B , define $\partial^+ B = \{v \mid v \in V, \exists a \in B, v = \partial^+ a\}$ and $\partial^- B = \{v \mid v \in V, \exists a \in B, v = \partial^- a\}$. For a vertex set U , denote $\delta^+ U = \{a \mid a \in A, \partial^+ a \in U, \partial^- a \notin U\}$ and $\delta^- U = \{a \mid a \in A, \partial^- a \in U, \partial^+ a \notin U\}$.

We denote the reverse arc of a by $\bar{a}$, i.e., $\bar{a} = (v, u)$ for $a = (u, v)$. An arc $a \in A$ is called *symmetric* if $\bar{a} \in A$. A digraph is said to be *symmetric* if every arc is symmetric.

For a subgraph H of G , the vertex set and the arc set of H are denoted by $V(H)$ and $A(H)$, respectively. We say H is *odd* (resp., *even*) if $|V(H)|$ is odd (resp., even). A strongly connected component C with $\delta^- V(C) = \emptyset$ (resp., $\delta^+ V(C) = \emptyset$) is called a *source-component* (resp., *sink-component*). For two vertex sets X^+ and X^- , let $G[X^+, X^-]$ denote the subgraph whose vertex set is $X^+ \cup X^-$ and arc set is $A[X^+, X^-] = \{a \mid a \in A, \partial^+ a \in X^+, \partial^- a \in X^-\}$. For a vertex set X , we abbreviate $G[X, X]$ and $A[X, X]$ as $G[X]$ and $A[X]$, respectively. We denote by $\text{odd}^+(X)$ (resp., $\text{odd}^-(X)$) the number of odd source-components (resp., odd sink-components) in $G[X]$.

For $v_0, v_1, \dots, v_k \in V$ and $a_1, \dots, a_k \in A$, a sequence $W = (v_0, a_1, v_1, a_2, v_2, \dots, v_{k-1}, a_k, v_k)$ is called a *walk* if $a_i = (v_{i-1}, v_i)$ for all $i = 1, \dots, k$. We denote the vertex set $\{v_0, \dots, v_k\} = V(W)$ and the arc set $\{a_1, \dots, a_k\} = A(W)$. For a walk $W = (v_0, a_1, v_1, \dots, v_{k-1}, a_k, v_k)$, the length k of W is denoted by $|W|$. A walk W is said to be *odd* (resp., *even*) if $|W|$ is odd (resp., even). A walk $W = (v_0, a_1, v_1, \dots, v_{k-1}, a_k, v_k)$ is called a *path* if $v_0, v_1, \dots, v_k$ are pairwise distinct and called a *cycle* if $v_0, v_1, \dots, v_{k-1}$ are pairwise distinct and $v_0 = v_k$. For a walk $W = (v_0, a_1, v_1, \dots, v_{k-1}, a_k, v_k)$, let $\bar{W}$ denote the reverse walk of W (if it exists), that is, $\bar{W} = (v_k, \bar{a}_k, v_{k-1}, \dots, v_1, \bar{a}_1, v_0)$. Recall that a digraph is *odd-cycle-symmetric* if any odd cycle has the reverse cycle.

For two vertex sets X^+ and X^- , we call (X^+, X^-) a *stable pair* if there is neither an arc $a \in A$ with $\partial^+ a \in X^+ \setminus X^-$ and $\partial^- a \in X^-$ nor $a \in A$ with $\partial^+ a \in X^+$ and $\partial^- a \in X^- \setminus X^+$. This definition generalizes that of Cunningham and Geelen [3], who first introduced the notion of stable pairs for the independent path-matching problem.

The outline of this paper is as follows. Section 2 introduces even factors and matroid intersection and describes some important results which are to be extended in independent even factors. Section 3 provides the operation of shrinking on matroids. Section 4 introduces independent even factors and proves a min-max formula structurally. Section 5 presents a new algorithm of ours. Finally, section 6 reveals the Edmonds–Gallai-type structure for independent even factors.

2. Even factors and matroid intersection.

2.1. Even factors. Let $G = (V, A)$ be a digraph. An *even factor* is the arc set of a vertex-disjoint collection of some paths and even cycles. For a digraph G , let $\nu(G)$ denote the cardinality of a maximum even factor in G .

One approach to computing $\nu(G)$ is extending the notion of the Tutte matrix in the matching problem [24]. Let $\{x_{uv} \mid (u, v) \in A\}$ be the indeterminates such that $x_{uv} = -x_{vu}$ for $(u, v), (v, u) \in A$. The *Tutte matrix* $T = (t_{uv})$ is defined by

$$t_{uv} = \begin{cases} x_{uv} & ((u, v) \in A), \\ 0 & (\text{otherwise}), \end{cases}$$

where rows and columns are both indexed by V . Cunningham and Geelen [4] asserted that

$$(2.1) \quad \nu(G) = \text{rank } T$$

if G is weakly symmetric. Based on this fact, the algorithm of Cunningham and Geelen [4] figures out $\nu(G)$ by computing $\text{rank } T$ for a weakly symmetric digraph G .

In fact, the proof of (2.1) in Cunningham and Geelen [4] applies to odd-cycle-symmetric digraphs. That is, (2.1) and their algorithm are valid for odd-cycle-symmetric digraphs.

THEOREM 2.1. *Let G be an odd-cycle-symmetric digraph and let T be the Tutte matrix of G . Then $\nu(G) = \text{rank } T$ holds.*

Another approach to $\nu(G)$, proposed by Pap [17], is extending the matching algorithm of Edmonds [5]. Pap's algorithm finds a maximum even factor in an odd-cycle-symmetric digraph and brings about the following min-max formula.

THEOREM 2.2. *For an odd-cycle-symmetric digraph $G = (V, A)$, it holds that*

$$\nu(G) = \min_{(X^+, X^-)} \{ |V \setminus X^+| + |V \setminus X^-| + |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-) \},$$

where (X^+, X^-) runs over all stable pairs.

This formula modestly extends that of Cunningham and Geelen [4, Theorem 1.3], in which the minimum is taken over all (X^+, X^-) such that $G[X^+, X^-]$ is symmetric. One easily sees that such (X^+, X^-) forms a stable pair.

2.2. Matroid intersection. Let V be a finite set and its subset family $\mathcal{I} \subseteq 2^V$ satisfy the following:

(I0) $\emptyset \in \mathcal{I}$;

(I1) $I \subseteq J \in \mathcal{I} \Rightarrow I \in \mathcal{I}$;

(I2) $I, J \in \mathcal{I}$, $|I| < |J| \Rightarrow \exists j \in J \setminus I$, $I \cup \{j\} \in \mathcal{I}$.

A set function $\rho : 2^V \rightarrow \mathbb{Z}_+$ is defined by

$$\rho(X) = \max\{|J| \mid J \subseteq X, J \in \mathcal{I}\} \quad (X \subseteq V).$$

For V , $\mathcal{I}$, and ρ defined above, we say $\mathbf{M} = (V, \mathcal{I}, \rho)$ forms a *matroid*. We call V the *ground set*, $\mathcal{I}$ the *independent set family*, $I \in \mathcal{I}$ an *independent set*, and ρ the *rank function*. A subset of V is said to be *dependent* if it is not an independent set.

Given a matroid $(V, \mathcal{I}, \rho)$, a subset $B \subseteq V$ is called a *base* if B is an inclusion-wise maximal independent set, and a subset $C \subseteq V$ is called a *circuit* if C is an inclusion-wise minimal dependent set.

The *closure function* $\text{cl} : 2^V \rightarrow 2^V$ is defined by

$$\text{cl}(X) = \{j \mid \rho(X \cup \{j\}) = \rho(X)\} \quad (X \subseteq V).$$

For any $I \in \mathcal{I}$ and $j \in \text{cl}(I) \setminus I$, the union $I \cup \{j\}$ contains a unique circuit, called the *fundamental circuit* denoted by $C(I \mid j)$.

Let us be given two matroids on the same ground set, say, $\mathbf{M}^+ = (V, \mathcal{I}^+, \rho^+)$ and $\mathbf{M}^- = (V, \mathcal{I}^-, \rho^-)$. The *matroid intersection problem* is a problem for finding a maximum common independent set $I \in \mathcal{I}^+ \cap \mathcal{I}^-$.

A min-max formula for the matroid intersection problem is given by Edmonds [6].

THEOREM 2.3 (Edmonds [6]). *For two matroids $\mathbf{M}_1 = (V, \mathcal{I}_1, \rho_1)$ and $\mathbf{M}_2 = (V, \mathcal{I}_2, \rho_2)$, it holds that*

$$\max\{|I| \mid I \subseteq V, I \in \mathcal{I}_1 \cap \mathcal{I}_2\} = \min_{X \subseteq V} \{\rho_1(V \setminus X) + \rho_2(X)\}.$$

There are some combinatorial polynomial-time algorithms for finding a maximum common independent set, such as Lawler's [15] and Edmonds' [7]. We remark that the notion of *alternating-paths* plays an essential part in those algorithms, as is the case with the matching algorithm of Edmonds [5] and the even factor algorithm of Pap [17].

3. Shrinking of matroids. This section is devoted to the notion of *shrinking* of matroids. Note that it is different from the well-known notion of contraction of matroids. We remark that a similar notion is independently introduced by Pap in [16], where it is named *projection*.

3.1. Definition. Let $\mathbf{M} = (V, \mathcal{I}, \rho)$ be a matroid and $W \subseteq V$ be an independent set in $\mathbf{M}$. Then, consider $\tilde{V} = (V \setminus W) \cup \{w\}$, where w is a new element. Define a subset family $\tilde{\mathcal{I}}$ of $\tilde{V}$ and a set function $\tilde{\rho} : 2^{\tilde{V}} \rightarrow \mathbb{Z}_+$ by

$$\tilde{\mathcal{I}} = \{I \mid I \subseteq V \setminus W, \exists J \subseteq W, |J| = |W| - 1, I \cup J \in \mathcal{I}\}$$

$$\cup \{I \cup \{w\} \mid I \subseteq V \setminus W, I \cup W \in \mathcal{I}\},$$

$$\tilde{\rho}(X) = \max\{|J| \mid J \subseteq X, J \in \tilde{\mathcal{I}}\} \quad (X \subseteq \tilde{V}).$$

LEMMA 3.1. *The tuple $(\tilde{V}, \tilde{\mathcal{I}}, \tilde{\rho})$ forms a matroid.*

Proof. It is easily seen that $\tilde{\mathcal{I}}$ satisfies (I0) and (I1). We prove (I2) by showing that for any $\tilde{I}, \tilde{J} \in \tilde{\mathcal{I}}$ with $|\tilde{I}| < |\tilde{J}|$, there exists an element v such that

$$(3.1) \quad v \in \tilde{J} \setminus \tilde{I}, \quad \tilde{I} \cup \{v\} \in \tilde{\mathcal{I}}.$$

Suppose $w \in \tilde{I}$, i.e., $\tilde{I} = \tilde{I}' \cup \{w\}$ with $\tilde{I}' \subseteq V \setminus W$. As $\tilde{I}, \tilde{J} \in \tilde{\mathcal{I}}$, it follows that $I = \tilde{I}' \cup W \in \mathcal{I}$, and there exists $J \in \mathcal{I}$ with $|J| = |\tilde{J}| + |W| - 1$ and $J \setminus \tilde{J} \subseteq W$.

Then, $|I| < |J|$ holds, which implies that there exists $v \in J \setminus I$ such that $I \cup \{v\} \in \mathcal{I}$. Hence, it holds that $\tilde{I} \cup \{v\} \in \mathcal{I}$ and $v \in \tilde{J} \setminus \tilde{I}$. Thus, v satisfies (3.1).

Suppose $w \notin \tilde{I}$. Then, there exists $W_I \subseteq W$ such that $|W_I| = |W| - 1$ and $I = \tilde{I} \cup W_I \in \mathcal{I}$. Define J and v as mentioned above. If $v \notin W$, we have that v satisfies (3.1). So, we may assume $v \in W$, which implies $W = W_I \cup \{v\}$ and $\tilde{I} \cup \{w\} \in \tilde{\mathcal{I}}$. If $w \in \tilde{J}$, then w satisfies (3.1). Otherwise, let $W_J = J \setminus \tilde{J}$. Note that $|W_J| = |W| - 1$ and $W_J \subseteq W$. Then, it holds that $I' = \tilde{I} \cup W_J \subseteq \tilde{I} \cup W$. Hence, by (I1) we have $I' \in \mathcal{I}$, and $|I'| = |\tilde{I}| < |J|$. Therefore, there exists an element $u \in J \setminus I' \subseteq \tilde{J} \setminus \tilde{I}$ such that $I' \cup \{u\} \in \mathcal{I}$, which implies (3.1) holds for u . $\square$

We denote the new matroid $(\tilde{V}, \tilde{\mathcal{I}}, \tilde{\rho})$ defined above by $\mathbf{M}_W$. The procedure to obtain $\mathbf{M}_W$ is called *shrinking* W , and we say that W is *shrunk into* w . Conversely, if I_W is an independent set in $\mathbf{M}_W$, there exists at least one independent set $I \in \mathcal{I}$ such that $I_W \setminus \{w\} \subseteq I$ and $|I| = |I_W| + |W| - 1$. The procedure of obtaining I is said to be *expanding* W .

Lemma 3.1 can also be confirmed through combining well-known operations in matroid theory, induction through a bipartite graph, and contraction.

Let $\mathbf{M} = (V, \mathcal{I}, \rho)$ be a matroid and $W \subseteq V$ be its independent set. Make a copy $\mathbf{M}' = (V', \mathcal{I}', \rho')$ of $\mathbf{M}$. We denote the copy of $v \in V$ (resp., $U \subseteq V$) by v' (resp., U'). Then, construct a bipartite graph $G = (V', V; A)$, where

$$A = \{(u', v) \mid u, v \in W\} \cup \{(v', v) \mid v \in V \setminus W\}.$$

Define a subset family $\mathcal{I}^\circ$ of V associated with $\mathcal{I}'$ and G as follows: a subset $I \subseteq V$ belongs to $\mathcal{I}^\circ$ if and only if there exists a matching $M \subseteq A$ in G such that $\partial^- M = I$ and $\partial^+ M \in \mathcal{I}'$. Perfect [19] observed that $\mathcal{I}^\circ$ satisfies (I0)–(I2). We denote the resulting matroid by $\mathbf{M}^\circ = (V, \mathcal{I}^\circ, \rho^\circ)$.

Next, choose an arbitrary subset S of W with $|S| = |W| - 1$ and contract S in $\mathbf{M}^\circ$ to obtain a new matroid $\tilde{\mathbf{M}}$. By denoting the unique element in $W \setminus S$ by w , one easily checks that $\tilde{\mathbf{M}}$ is the same as $\mathbf{M}_W$.

From this construction, we obtain an explicit description of $\tilde{\rho}$.

THEOREM 3.2. *Let $\mathbf{M} = (V, \mathcal{I}, \rho)$ be a matroid, $W \subseteq V$ be an independent set in $\mathbf{M}$, and $\mathbf{M}_W = (\tilde{V}, \tilde{\mathcal{I}}, \tilde{\rho})$ be the matroid obtained by shrinking W . Then, for $X \subseteq \tilde{V}$, it holds that*

$$\tilde{\rho}(X) = \begin{cases} \rho(X \cup W) - |W| + 1 & (w \in X), \\ \min\{\rho(X), \rho(X \cup W) - |W| + 1\} & (w \notin X). \end{cases}$$

Proof. Given $\mathbf{M} = (V, \mathcal{I}, \rho)$ and $W \in \mathcal{I}$, define $G = (V', V; A)$ and $\mathbf{M}^\circ = (V, \mathcal{I}^\circ, \rho^\circ)$ as above. Let S be a subset of W with $|S| = |W| - 1$ and w be the unique element in $W \setminus S$. Then, for $X \subseteq \tilde{V}$, it holds that

$$\tilde{\rho}(X) = \rho^\circ(X \cup S) - \rho^\circ(S) = \min_{Y \subseteq X \cup S} \{\rho(\Gamma(Y)) + |X \cup S \setminus Y|\} - |W| + 1,$$

where $\Gamma(Y) = \{u \mid u \in V, \exists(u', v) \in A, v \in Y\}$ for $Y \subseteq V$.

Suppose $w \in X$. In this case, we have

$$\begin{aligned}
& \min_{Y \subseteq X \cup S} \{\rho(\Gamma(Y)) + |X \cup S \setminus Y|\} \\
&= \min \left\{ \min_{Y \subseteq X \setminus \{w\}} \{\rho(Y) + |X \cup S \setminus Y|\}, \min_{\substack{Y \subseteq X \cup S, \\ Y \cap W \neq \emptyset}} \{\rho(Y \cup W) + |X \cup S \setminus Y|\} \right\} \\
&= \min \{\rho(X \setminus \{w\}) + |W|, \rho(X \cup W)\} \\
&= \rho(X \cup W),
\end{aligned}$$

where the second equality can be seen through putting $Y = X \setminus \{w\}, X \cup S$.

Suppose $w \notin X$. In this case,

$$\begin{aligned}
& \min_{Y \subseteq X \cup S} \{\rho(\Gamma(Y)) + |X \cup S \setminus Y|\} \\
&= \min \left\{ \min_{Y \subseteq X} \{\rho(Y) + |X \cup S \setminus Y|\}, \min_{\substack{Y \subseteq X \cup S, \\ Y \cap S \neq \emptyset}} \{\rho(Y \cup W) + |X \cup S \setminus Y|\} \right\} \\
&= \min \{\rho(X) + |W| - 1, \rho(X \cup W)\},
\end{aligned}$$

where the second equality can be seen through putting $Y = X, X \cup S$. $\square$

3.2. Independence test. Here we describe how to test independence in matroids obtained by repeated shrinking. Suppose we have an independence oracle for a matroid $\mathbf{M} = (V, \mathcal{I}, \rho)$, and we obtained a matroid $\tilde{\mathbf{M}} = (\tilde{V}, \tilde{\mathcal{I}}, \tilde{\rho})$ by applying to $\mathbf{M}$ the procedure shrinking repeatedly. Let $w_1, \dots, w_k$ be the newly created elements in $\tilde{V}$, that is, $\{w_1, \dots, w_k\} = \tilde{V} \setminus V$, and denote the set of vertices shrunk into the pseudovortex w_i by $W_i \subseteq V$ ($i = 1, \dots, k$). If the vertices shrunk into w_i are nested (that is, w_i is obtained by shrinking vertices that includes pseudovertrices created in preceding shrinking), it suffices to deal with w_i as if we executed a single procedure of shrinking W_i into w_i .

Assume that we have an independent set $\tilde{I} \in \tilde{\mathcal{I}}$, and we know how to expand W_i ($i = 1, \dots, k$) to obtain an independent set $I \in \mathcal{I}$. We focus on testing whether $\tilde{I} \cup \{v\}$ is independent for all $v \in \tilde{V} \setminus \tilde{I}$. Moreover, if $\tilde{I} \cup \{v\} \notin \tilde{\mathcal{I}}$, we enumerate all $u \in \tilde{I}$ such that $(\tilde{I} \setminus \{u\}) \cup \{v\} \in \tilde{\mathcal{I}}$. These independence tests can be performed efficiently through constructing an auxiliary graph.

Construct an auxiliary graph $G^\natural = (V^\natural, A^\natural; S^1, S^2)$ as follows. The vertex set is $V^\natural = V^1 \cup V^2$, where V^1 and V^2 are copies of V . The copy of $v \in V$ (resp., $W \subseteq V$) in V^i is denoted by v^i (resp., W^i) for $i = 1, 2$. The arc set $A^\natural$ consists of $A^\natural = A_V \cup A_M \cup J$, where

$$\begin{aligned}
A_V &= \left(\bigcup_{i=1}^k \{(u^1, v^2) \mid u, v \in W_i\} \right) \cup \{(v^1, v^2) \mid v \in V \setminus (W_1 \cup \dots \cup W_k)\}, \\
A_M &= \{(v^2, v^1) \mid v \in I\}, \\
J &= \{(u^2, v^2) \mid u \in \text{cl}(I) \setminus I, v \in I, v \in C(I \mid u)\}.
\end{aligned}$$

The two subsets $S^1 \subseteq V^1$ and $S^2 \subseteq V^2$ are defined by

$$S^1 = \{v^1 \mid v \in V \setminus I\}, \quad S^2 = \{v^2 \mid v \in V \setminus I, I \cup \{v\} \in \mathcal{I}\}.$$

We refer to S^1 and S^2 as the set of *source-vertices* and *sink-vertices*, respectively, and refer to a path which starts from S^1 and ends in S^2 as a *source-sink path*.

Fix $v \in V \setminus I$. Assume v is not a pseudovortex. If there exists a source-sink path in $G^\natural$ which starts from v^1 , then $\tilde{I} \cup \{v\} \in \tilde{\mathcal{I}}$. If not, one can find all elements $u \in \tilde{I}$ which satisfy $(\tilde{I} \setminus \{u\}) \cup \{v\} \in \tilde{\mathcal{I}}$ as follows: For a vertex $u \in \tilde{I} \setminus \{w_1, \dots, w_k\}$, it holds that $(\tilde{I} \setminus \{u\}) \cup \{v\} \in \tilde{\mathcal{I}}$ if and only if there exists a path in $G^\natural$ which starts from v^1 and ends in u^2 ; for a pseudovortex $w_i \in \tilde{I}$, it holds that $(\tilde{I} \setminus \{w_i\}) \cup \{v\} \in \tilde{\mathcal{I}}$ if and only if there exists a path in $G^\natural$ which starts from v^1 and ends in w^2 with $w \in W_i$.

Assume v is a pseudovortex, say, $v = w_1$, and let $w_* \in V$ be the unique vertex in $W_1 \setminus I$. In that case, if there exists a source-sink path in $G^\natural$ which starts from w_*^1 , then $\tilde{I} \cup \{v\} \in \tilde{\mathcal{I}}$. If not, one can find all elements $u \in \tilde{I}$ which satisfy $(\tilde{I} \setminus \{u\}) \cup \{v\} \in \tilde{\mathcal{I}}$ as follows: For a vertex $u \in \tilde{I} \setminus \{w_1, \dots, w_k\}$, it holds that $(\tilde{I} \setminus \{u\}) \cup \{v\} \in \tilde{\mathcal{I}}$ if and only if there exists a path in $G^\natural$ which starts from w_*^1 and ends in u^2 ; for a pseudovortex $w_i \in \tilde{I}$, it holds that $(\tilde{I} \setminus \{w_i\}) \cup \{v\} \in \tilde{\mathcal{I}}$ if and only if there exists a path in $G^\natural$ which starts from w_*^1 and ends in w^2 with $w \in W_i$.

4. The independent even factor problem. From now on, we deal with the *independent even factor problem*, which is a variant of the *basic even factor problem* [4].

DEFINITION 4.1. Let $G = (V, A)$ be a digraph and $\mathbf{M}^+ = (V, \mathcal{I}^+, \rho^+)$, $\mathbf{M}^- = (V, \mathcal{I}^-, \rho^-)$ be matroids. An even factor M in G is called an *independent even factor* in $(G, \mathbf{M}^+, \mathbf{M}^-)$ if $\partial^+ M \in \mathcal{I}^+$ and $\partial^- M \in \mathcal{I}^-$. If $\partial^+ M$ and $\partial^- M$ are bases of $\mathbf{M}^+$ and $\mathbf{M}^-$, respectively, M is called a *basic even factor*.

For a digraph $G = (V, A)$ and matroids $\mathbf{M}^+ = (V, \mathcal{I}^+)$, $\mathbf{M}^- = (V, \mathcal{I}^-)$, let $\nu(G, \mathbf{M}^+, \mathbf{M}^-)$ denote the cardinality of a maximum independent even factor in $(G, \mathbf{M}^+, \mathbf{M}^-)$. The independent even factor problem contains the even factor problem and the matroid intersection problem as special cases.

Even factor problem. If $\mathbf{M}^+$ and $\mathbf{M}^-$ are free, an arc set is an independent even factor in $(G, \mathbf{M}^+, \mathbf{M}^-)$ if and only if it is an even factor in G .

Matroid intersection problem. Let $\mathbf{M}_1 = (V, \mathcal{I}_1, \rho_1)$ and $\mathbf{M}_2 = (V, \mathcal{I}_2, \rho_2)$ be an instance of the matroid intersection problem. Then, construct an associated bipartite digraph $G = (V_1, V_2; A)$ as follows. Both of the vertex sets V_1 and V_2 are copies of V , respectively, and the arc set A is given by

$$A = \{(v_1, v_2) \mid v \in V\},$$

where $v_1 \in V_1$ (resp., $v_2 \in V_2$) denotes the copy of $v \in V$. On the vertex set $V_1 \cup V_2$, define matroids $\mathbf{M}^+ = (V_1 \cup V_2, \mathcal{I}^+, \rho^+)$ and $\mathbf{M}^- = (V_1 \cup V_2, \mathcal{I}^-, \rho^-)$ by the following rank functions $\rho^+, \rho^- : 2^{V_1 \cup V_2} \rightarrow \mathbb{Z}_+$:

$$\rho^+(X) = \rho_1(X \cap V_1), \quad \rho^-(X) = \rho_2(X \cap V_2).$$

Then, a subset $I \subseteq V$ belongs to $\mathcal{I}_1 \cap \mathcal{I}_2$ if and only if the arc set $\{(v_1, v_2) \mid v \in I\}$ is an independent even factor in $(G, \mathbf{M}^+, \mathbf{M}^-)$.

For any independent even factor in $(G, \mathbf{M}^+, \mathbf{M}^-)$ and any stable pair in G , the following inequality holds.

LEMMA 4.2. Let $G = (V, A)$ be a digraph and $\mathbf{M}^+ = (V, \mathcal{I}^+, \rho^+)$, $\mathbf{M}^- = (V, \mathcal{I}^-, \rho^-)$ be matroids. Then, for any independent even factor M in $(G, \mathbf{M}^+, \mathbf{M}^-)$ and any stable pair (X^+, X^-) in G , it holds that

$$(4.1) \quad |M| \leq \rho^+(V \setminus X^+) + \rho^-(V \setminus X^-) + |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-).$$

Proof. As (X^+, X^-) is a stable pair, it holds that $M = M_1 \cup M_2 \cup M_3$, where

$$(4.2) \quad M_1 = \{a \mid a \in M, \partial^+ a \in V \setminus X^+\},$$

$$(4.3) \quad M_2 = \{a \mid a \in M, \partial^- a \in V \setminus X^-\},$$

$$(4.4) \quad M_3 = \{a \mid a \in M, a \in A[X^+ \cap X^-]\}.$$

Note that M_1 and M_2 are not necessarily disjoint. We accomplish the proof by showing

$$(4.5) \quad |M_1| \leq \rho^+(V \setminus X^+), \quad |M_2| \leq \rho^-(V \setminus X^-), \quad |M_3| \leq |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-).$$

First, as M is an independent even factor, it follows that $\partial^+ M \in \mathcal{I}^+$, especially $\partial^+ M_1 \in \mathcal{I}^+$. By the definition of M_1 , we have $\partial^+ M_1 \subseteq V \setminus X^+$, which implies $|M_1| = |\partial^+ M_1| \leq \rho^+(V \setminus X^+)$.

Next, a similar argument shows that $|M_2| \leq \rho^-(V \setminus X^-)$.

Finally, let us estimate $|M_3|$ by counting $|\partial^- M_3|$. For a source-component C in $G[X^+ \cap X^-]$, it holds that $|V(C) \cap \partial^- M_3| = |A(C) \cap M_3|$. Hence, if C is odd and $|V(C) \cap \partial^- M_3| = |C|$, then $A(C) \cap M_3$ would contain an odd cycle, which contradicts that M is an even factor in G . Thus, for every odd source-component C in $X^+ \cap X^-$, it holds that $|V(C) \cap \partial^- M_3| \leq |C| - 1$, which implies $|M_3| = |\partial^- M_3| \leq |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-)$. $\square$

A stable pair (X^+, X^-) is called *minimizing* if (X^+, X^-) minimizes the right-hand side of (4.1). In fact, for a maximum independent even factor M and a minimizing stable pair (X^+, X^-) , (4.1) holds with equality if G is odd-cycle-symmetric.

THEOREM 4.3. For an odd-cycle-symmetric digraph $G = (V, A)$ and matroids $\mathbf{M}^+ = (V, \mathcal{I}^+, \rho^+)$, $\mathbf{M}^- = (V, \mathcal{I}^-, \rho^-)$, it holds that

$$(4.6) \quad \nu(G, \mathbf{M}^+, \mathbf{M}^-) = \min_{(X^+, X^-)} \{ \rho^+(V \setminus X^+) + \rho^-(V \setminus X^-) + |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-) \},$$

where (X^+, X^-) runs over all stable pairs.

One easily sees that Theorem 4.3 generalizes Theorems 2.2 and 2.3.

Theorem 4.3 can be proved by extending Cunningham and Geelen's approach [4]. Let $G = (V, A)$ be an odd-cycle-symmetric digraph, $\mathbf{M}^+ = (V, \mathcal{I}^+, \rho^+)$ and $\mathbf{M}^- = (V, \mathcal{I}^-, \rho^-)$ be matroids, and T be the Tutte matrix of G . Let us denote by $T[I^+, I^-]$ the submatrix of T whose row set is $I^+ \subseteq V$ and column set $I^- \subseteq V$. Note that the submatrix $T[I^+, I^-]$ is the Tutte matrix of the subgraph $G[I^+, I^-]$. By Theorem 2.1, we have that if a submatrix $T[I^+, I^-]$ such that $I^+ \in \mathcal{I}^+$ and $I^- \in \mathcal{I}^-$ is nonsingular, there exists an independent even factor M with $\partial^+ M = I^+$ and $\partial^- M = I^-$. As for such submatrices, there exists a min-max formula, which is a special case of Schrijver's one for linking systems [20, Theorem 3.6].

THEOREM 4.4. *Let T be a matrix whose row set and column set are both indexed by V , and let $\mathbf{M}^+ = (V, \mathcal{I}^+, \rho^+)$ and $\mathbf{M}^- = (V, \mathcal{I}^-, \rho^-)$ be matroids. Then, it holds that*

$$\begin{aligned} & \max \{ |I^+| \mid I^+ \in \mathcal{I}^+, I^- \in \mathcal{I}^-, T[I^+, I^-] \text{ is nonsingular} \} \\ &= \min_{U^+, U^- \subseteq V} \{ \rho^+(V \setminus U^+) + \rho^-(V \setminus U^-) + \text{rank } T[U^+, U^-] \}. \end{aligned}$$

We remark here that Schrijver [20] proved the min-max formula using Theorem 2.3. Combining Theorems 2.2 and 4.4 procures us a proof of Theorem 4.3.

Proof of Theorem 4.3. Let T be the Tutte matrix of G . By Theorem 4.4, we have

$$\nu(G, \mathbf{M}^+, \mathbf{M}^-) = \min_{U^+, U^- \subseteq V} \{ \rho^+(V \setminus U^+) + \rho^-(V \setminus U^-) + \text{rank } T[U^+, U^-] \}.$$

Moreover, we know by Theorems 2.1 and 2.2 that

$$\begin{aligned} \text{rank } T[U^+, U^-] &= \nu(G[U^+, U^-]) \\ &= \min_{(X^+, X^-)} \{ |U^+ \setminus X^+| + |U^- \setminus X^-| + |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-) \}, \end{aligned}$$

where (X^+, X^-) runs over all stable pairs in $G[U^+, U^-]$. Note that the second equality follows from a variant of Theorem 2.2, which can also be proved by Pap's even factor algorithm [17]. Therefore, we have that $\nu(G, \mathbf{M}^+, \mathbf{M}^-)$ is equal to

$$\begin{aligned} & \min_{U^+, U^- \subseteq V} \left\{ \rho^+(V \setminus U^+) + \rho^-(V \setminus U^-) \right. \\ & \quad \left. + \min_{(X^+, X^-)} \{ |U^+ \setminus X^+| + |U^- \setminus X^-| + |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-) \} \right\} \\ &= \min_{(X^+, X^-)} \left\{ \min_{U^+, U^-} \{ \rho^+(V \setminus U^+) + \rho^-(V \setminus U^-) + |U^+ \setminus X^+| + |U^- \setminus X^-| \right. \\ & \quad \left. + |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-) \} \right\}, \end{aligned}$$

where U^+ and U^- satisfy that (X^+, X^-) is a stable pair in $G[U^+, U^-]$. The inner minimum in the right-hand side is achieved when $U^+ = X^+$ and $U^- = X^-$. Hence, it follows that $\nu(G, \mathbf{M}^+, \mathbf{M}^-)$ is equal to

$$\min_{(X^+, X^-)} \{ \rho^+(V \setminus X^+) + \rho^-(V \setminus X^-) + |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-) \},$$

where (X^+, X^-) runs over all stable pairs. $\square$

5. An alternating-path algorithm. In this section, we describe a combinatorial algorithm to find a maximum independent even factor in $(G, \mathbf{M}^+, \mathbf{M}^-)$, where G is odd-cycle-symmetric. This algorithm is a natural generalization of the maximum even factor algorithm [17] and the matroid intersection algorithms [7, 15], in both of which the notion of *alternating-path* plays an essential part. The algorithm also finds a minimizing stable pair which verifies (4.6).

5.1. The algorithm. Let $G = (V, A)$ be an odd-cycle-symmetric digraph, and let $\mathbf{M}^+ = (V, \mathcal{I}^+, \rho^+)$ and $\mathbf{M}^- = (V, \mathcal{I}^-, \rho^-)$ be matroids, whose closure functions and fundamental circuits are denoted by $\text{cl}^+(\cdot), \text{cl}^-(\cdot)$ and $C^+(\cdot | \cdot), C^-(\cdot | \cdot)$, respectively. Assume we have an independent even factor M in $(G, \mathbf{M}^+, \mathbf{M}^-)$. We show a procedure for finding an independent even factor M' with $|M'| = |M| + 1$ if it exists, and if not, determining that M is a maximum independent even factor.

For M , construct an auxiliary graph $G_M = (V^*, A^*; S^+, S^-)$ as follows. The vertex set V^* is given by $V^* = V^+ \cup V^-$, where V^+ and V^- are two copies of V . For a vertex $v \in V$, we denote the copy of v in V^+ (resp., V^-) by v^+ (resp., v^-).

The arc set A^* is composed of

$$A^* = A^\circ \cup M^\circ \cup J^+ \cup J^-,$$

where

$$\begin{aligned} A^\circ &= \{(u^+, v^-) \mid (u, v) \in A\} && \text{(copy of } A), \\ M^\circ &= \{(v^-, u^+) \mid (u, v) \in M\} && \text{(reverse copy of } M), \\ J^+ &= \{(u^+, v^+) \mid u \in \partial^+ M, v \in \text{cl}^+(\partial^+ M) \setminus \partial^+ M, u \in C^+(\partial^+ M | v)\}, \\ J^- &= \{(u^-, v^-) \mid u \in \text{cl}^-(\partial^- M) \setminus \partial^- M, v \in \partial^- M, v \in C^-(\partial^- M | u)\}. \end{aligned}$$

We call the arcs in J^+ and J^- the *jumping arcs*.

The two subsets $S^+ \subseteq V^+$ and $S^- \subseteq V^-$ are defined by

$$S^+ = \{v^+ \mid v \in V \setminus \text{cl}^+(\partial^+ M)\}, \quad S^- = \{v^- \mid v \in V \setminus \text{cl}^-(\partial^- M)\}.$$

We refer to S^+ and S^- as the set of *source-vertices* and *sink-vertices*, respectively.

A vertex $v \in V^*$ is called *source-reachable* if there exists a path which starts in S^+ and ends in v , and is called *sink-reachable* if there exists a path which starts in v and ends in S^- . A path P is called a *source-sink path* if P starts in S^+ and ends in S^- . For a path P in G_M , we denote the master copy of $A(P) \cap (A^\circ \cup M^\circ)$ by $A^\circ(P)$ (that is, $A^\circ(P) \subseteq A$).

Suppose that G_M has source-sink paths. Let P be the shortest source-sink path and $A^\circ(P) = \{a_1, m_1, a_2, \dots, a_k, m_k, a_{k+1}\}$, which is in order of appearance in P . Note that $a_1, \dots, a_{k+1} \in A \setminus M$ and $m_1, \dots, m_k \in M$. Then, we obtain an independent even factor M' with $|M'| = |M| + 1$ by putting $M' = M \triangle A^\circ(P) = (M \setminus \{m_1, \dots, m_k\}) \cup \{a_1, \dots, a_{k+1}\}$, if M' does not contain odd cycles. We refer to this procedure as $\text{Augment}(M, P)$.

If there exist odd cycles in M' , we avoid the creation of the cycles as follows. Denote

$$A^\circ(P_i) = \begin{cases} \emptyset & (i = 0), \\ \{a_1, m_1, a_2, \dots, a_i, m_i\} & (i = 1, \dots, k), \\ A^\circ(P) & (i = k + 1). \end{cases}$$

Let j be the smallest number such that $M \triangle A^\circ(P_j)$ contains odd cycles. Note that $M \triangle A^\circ(P_j)$ contains a unique odd cycle, denoted by C , and observe that $V(C) \in \mathcal{I}^+ \cap \mathcal{I}^-$. Now, take the symmetric difference of M and $A^\circ(P_{j-1})$, say, $N = M \triangle A^\circ(P_{j-1})$. Let us remark that N is an independent even factor with $|N| = |M|$ and $|N \cap A(C)| =$

$|C| - 1$. Then, shrink $V(C)$ into a new vertex v_C to obtain a new digraph $\tilde{G} = (\tilde{V}, \tilde{A})$ and matroids $\tilde{\mathbf{M}}^+, \tilde{\mathbf{M}}^-$. In need of an explicit representation of the shrunk odd cycle, we shall denote $\tilde{G}$, $\tilde{V}$, and $\tilde{A}$ by G_C , V_C , and A_C , respectively. Note that the arc set $\tilde{N} = N \cap \tilde{A}$ is an independent even factor in $(\tilde{G}, \tilde{\mathbf{M}}^+, \tilde{\mathbf{M}}^-)$ with $|\tilde{N}| = |N| - |C| + 1$. We call this procedure $\text{Shrink}(M, P, C)$.

We show the validity of the procedure $\text{Shrink}(M, P, C)$ by the propositions below. First, the following proposition ensures the odd-cycle-symmetry of $\tilde{G}$.

PROPOSITION 5.1 (Pap [17]). *Let $G = (V, A)$ be an odd-cycle-symmetric digraph. For an odd cycle C , it holds that $G_C = (V_C, A_C)$ is also odd-cycle-symmetric.* $\square$

By Lemma 3.1 and Proposition 5.1, it is ensured that the algorithm also runs in $(\tilde{G}, \tilde{\mathbf{M}}^+, \tilde{\mathbf{M}}^-)$.

Next, for an independent even factor $\tilde{M}$ in $(\tilde{G}, \tilde{\mathbf{M}}^+, \tilde{\mathbf{M}}^-)$, we describe how to obtain an independent even factor M in $(G, \mathbf{M}^+, \mathbf{M}^-)$ with $|M| = |\tilde{M}| + |C| - 1$.

Suppose $v_C \in \partial^+ \tilde{M} \cap \partial^- \tilde{M}$. Then, we have that $(\partial^+ \tilde{M} \setminus \{v_C\}) \cup V(C) \in \mathcal{I}^+$ and $(\partial^- \tilde{M} \setminus \{v_C\}) \cup V(C) \in \mathcal{I}^-$. In this case, it suffices to add $|C| - 1$ arcs in $A[V(C)]$ to $\tilde{M}$ so that the resulting arc set forms an even factor in G . Let us denote by a^+ (resp., a^-) the unique arc in $\delta^+ V(C) \cap \tilde{M}$ (resp., $\delta^- V(C) \cap \tilde{M}$). As C is odd and G is odd-cycle-symmetric, there exists an even path P from $\partial^- a^-$ to $\partial^+ a^+$ in C or $\bar{C}$. In addition, there exist cycles of length two that cover every vertex $v \in V(C)$ exposed by P exactly once. The union of $\tilde{M}$, $A(P)$, and the arcs in these cycles is an even factor in G of size $|\tilde{M}| + |C| - 1$, $\partial^+ M = (\partial^+ \tilde{M} \setminus \{v_C\}) \cup V(C)$, and $\partial^- M = (\partial^- \tilde{M} \setminus \{v_C\}) \cup V(C)$.

In the case where $v_C \notin \partial^+ \tilde{M}$, we know which $|C| - 1$ vertices in $V(C)$ are in $\partial^+ M$ by $G^\natural$, the auxiliary graph for the independence test in $\tilde{\mathbf{M}}^+$. The same applies to $\partial^- M$, and so an argument similar to the one above shows which $|C| - 1$ arcs in $A(C) \cup A(\bar{C})$ are to be added. We call this procedure $\text{Expand}(\tilde{M}, C)$.

The procedure $\text{Expand}(\tilde{M}, C)$ ensures that for every independent even factor $\tilde{M}$ in $(\tilde{G}, \tilde{\mathbf{M}}^+, \tilde{\mathbf{M}}^-)$, there exists an independent even factor M in $(G, \mathbf{M}^+, \mathbf{M}^-)$ with $|M| = |\tilde{M}| + |C| - 1$. Hence, the following proposition holds.

PROPOSITION 5.2. *Let $G = (V, A)$ be an odd-cycle-symmetric digraph, and let $\mathbf{M}^+, \mathbf{M}^-$ be two matroids on V . For an odd cycle C with $V(C) \in \mathcal{I}^+ \cup \mathcal{I}^-$, it follows that*

$$\nu(G, \mathbf{M}^+, \mathbf{M}^-) \geq \nu(G_C, \mathbf{M}_{V(C)}^+, \mathbf{M}_{V(C)}^-) + |C| - 1.$$

We are now ready to describe the steps for finding an independent even factor of cardinality larger than M by one.

Step 1. Construct the auxiliary graph G_M . If there exists a source-sink path in G_M , then go to Step 2. Otherwise, go to Step 4.

Step 2. Let P be the shortest source-sink path. If $M \triangle A^\circ(P)$ does not contain any odd cycle, apply $\text{Augment}(M, P)$, and then go to Step 4. Otherwise, go to Step 3.

Step 3. Apply $\text{Shrink}(M, P, C)$, update $(G, \mathbf{M}^+, \mathbf{M}^-) := (\tilde{G}, \tilde{\mathbf{M}}^+, \tilde{\mathbf{M}}^-)$, and $M := \tilde{N}$. Then, go to Step 1.

Step 4. Apply $\text{Expand}(M, C)$ while there exists a pseudovortex v_C .

In the case that no source-sink path is found in Step 1, the algorithm determines that M is a maximum independent even factor. We certify this by showing that there exists a stable pair (X^+, X^-) such that M and (X^+, X^-) certify (4.6).

Algorithmic proof of Theorem 4.3. Let $(G, \mathbf{M}^+, \mathbf{M}^-)$ be the tuple in which no source-sink path was found in Step 1. Let R be the set of the source-reachable vertices in G_M , $X^+ = \{v \mid v^+ \in R\}$, and $X^- = \{v \mid v^- \notin R\}$. Observe that (X^+, X^-) is a

stable pair, and define $M_1, M_2, M_3 \subseteq M$ by (4.2)–(4.4). By the definition of R , we have that $M_1 \cap M_2 = \emptyset$, and $M_3 = \emptyset$. Moreover, $A[X^+ \cap X^-] = \emptyset$, which implies that $|X^+ \cap X^-| = \text{odd}^+(X^+ \cap X^-)$.

Pick up $v \in (V \setminus X^+) \setminus \partial^+ M$. As v^+ is not source-reachable, v^+ is not a source-vertex, that is, $\partial^+ M \cup \{v\} \notin \mathcal{I}^+$. Then, consider the fundamental circuit $C^+(\partial^+ M \mid v)$. If $C^+(\partial^+ M \mid v)$ and X^+ were to intersect at some vertex u , there would be a jumping arc from u^+ to v^+ , which contradicts that v^+ is not source-reachable. Hence, we have that $C^+(\partial^+ M \mid v) \subseteq V \setminus X^+$ for all $v \in (V \setminus X^+) \setminus \partial^+ M$, which implies that $|M_1| = \rho^+(V \setminus X^+)$. A similar argument shows that $|M_2| = \rho^-(V \setminus X^-)$. Therefore, it holds that $|M| = \rho^+(V \setminus X^+) + \rho^-(V \setminus X^-) + |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-)$.

The argument above just ensures that (4.6) holds in $(G, \mathbf{M}^+, \mathbf{M}^-)$, which may be achieved as a result of applying $\text{Shrink}(\cdot, \cdot, \cdot)$ repeatedly. We show below that (4.6) holds before the repeated $\text{Shrink}(\cdot, \cdot, \cdot)$.

Suppose that $(G, \mathbf{M}^+, \mathbf{M}^-)$ is the tuple obtained by applying $\text{Shrink}(\hat{M}, P, C)$ to $(\hat{G}, \hat{\mathbf{M}}^+, \hat{\mathbf{M}}^-)$, where $\hat{G} = (\hat{V}, \hat{A})$, $\hat{\mathbf{M}}^+ = (\hat{V}, \hat{\mathcal{I}}^+, \hat{\rho}^+)$, and $\hat{\mathbf{M}}^- = (\hat{V}, \hat{\mathcal{I}}^-, \hat{\rho}^-)$. Associated with X^+ and X^- mentioned above, define $\hat{X}^+, \hat{X}^- \subseteq \hat{V}$ as the inverse-image of X^+, X^- , respectively, that is,

$$\hat{X}^+ = \begin{cases} X^+ & (v_C \notin X^+), \\ X^+ \cup V(C) & (v_C \in X^+), \end{cases}$$

as is $\hat{X}^-$. Now we prove that $(\hat{X}^+, \hat{X}^-)$ is a stable pair which certifies (4.6).

First, one can easily see that $(\hat{X}^+, \hat{X}^-)$ forms a stable pair in $\hat{G}$, which follows from the fact that (X^+, X^-) is a stable pair in G .

Second, we estimate the value $\hat{\rho}^+(\hat{V} \setminus \hat{X}^+) + \hat{\rho}^-(\hat{V} \setminus \hat{X}^-) + |\hat{X}^+ \cap \hat{X}^-| - \text{odd}^+(\hat{X}^+ \cap \hat{X}^-)$.

One can easily see that v_C is a source-vertex in G_M , which implies that $v_C \in X^+$. Then, by Theorem 3.2, we have $\hat{\rho}^+(\hat{V} \setminus \hat{X}^+) = \rho(V \setminus X^+)$, since $\hat{\rho}^+(\hat{V} \setminus \hat{X}^+) \leq \hat{\rho}^+((\hat{V} \setminus \hat{X}^+) \cup V(C)) - |C| + 1$ follows that we can extend M with $|C| - 1$ arcs in $A(C) \cup A(\bar{C})$ to obtain an independent even factor in $(\hat{G}, \hat{\mathbf{M}}^+, \hat{\mathbf{M}}^-)$.

Assume $v_C \in X^+ \cap X^-$. Then, we have $\hat{\rho}^-(\hat{V} \setminus \hat{X}^-) = \rho^-(V \setminus X^-)$ from an argument similar to the one above. Moreover, $|\hat{X}^+ \cap \hat{X}^-| - \text{odd}^+(\hat{X}^+ \cap \hat{X}^-) = |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-) + |C| - 1$ follows from the following equations: $|\hat{X}^+ \cap \hat{X}^-| = |X^+ \cap X^-| + |C| - 1$; $\text{odd}^+(\hat{X}^+ \cap \hat{X}^-) = \text{odd}^+(X^+ \cap X^-)$.

Assume $v_C \in X^+ \setminus X^-$. In this case, Theorem 3.2 implies that $\hat{\rho}^-(\hat{V} \setminus \hat{X}^-) = \rho^-(V \setminus X^-) + |C| - 1$, and it is obvious that $|\hat{X}^+ \cap \hat{X}^-| - \text{odd}^+(\hat{X}^+ \cap \hat{X}^-) = |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-)$.

Therefore, in any case we have

$$(5.1) \quad \hat{\rho}^+(\hat{V} \setminus \hat{X}^+) + \hat{\rho}^-(\hat{V} \setminus \hat{X}^-) + |\hat{X}^+ \cap \hat{X}^-| - \text{odd}^+(\hat{X}^+ \cap \hat{X}^-) \\ = \rho^+(V \setminus X^+) + \rho^-(V \setminus X^-) + |X^+ \cap X^-| - \text{odd}^+(X^+ \cap X^-) + |C| - 1.$$

Hence, combining Lemma 4.2, Proposition 5.2, and (5.1), we have that (4.6) holds for $(\hat{G}, \hat{\mathbf{M}}^+, \hat{\mathbf{M}}^-)$. $\square$

5.2. Complexity. We discuss the complexity of the algorithm described in section 5.1. Let $(G, \mathbf{M}^+, \mathbf{M}^-)$, where $G = (V, A)$, $\mathbf{M}^+ = (V, \mathcal{I}^+, \rho^+)$, and $\mathbf{M}^- = (V, \mathcal{I}^-, \rho^-)$, be the original instance. Recall that $n = |V|$ and Q is the time needed to test if a given set $U \subseteq V$ is independent in $\mathbf{M}^+$ or $\mathbf{M}^-$.

The most time-consuming part of the algorithm is the construction of the auxiliary graphs $G_M = (V', A'; S^+, S^-)$. We can identify the source-vertices S^+ , sink-vertices S^- , and jumping arcs J^+, J^- by searching paths in the auxiliary graphs $G^\natural$ described in section 3.2. The construction of the auxiliary graph $G^\natural$ takes $O(n^2Q)$ time. In the procedure $\text{Shrink}(\cdot, \cdot, \cdot)$, the number of vertices decreases by at least two, which implies that we have to construct $O(n)$ auxiliary graphs G_M between augmentations. Since there are at most n augmentations, the total time complexity of the algorithm is $O(n^4Q)$.

6. An Edmonds–Gallai-type structure. In this section, we show that independent even factors have a structure which generalizes both the Edmonds–Gallai decomposition for matchings or even factors, and the principal partition for matroid intersection.

Let $G = (V, A)$ be an odd-cycle-symmetric digraph, and let $\mathbf{M}^+ = (V, \mathcal{I}^+, \rho^+)$, $\mathbf{M}^- = (V, \mathcal{I}^-, \rho^-)$ be matroids. In the independent even factor algorithm, we obtain a stable pair (X^+, X^-) in the shrunk digraph when no source-sink path is found. Then, expand pseudovertrices to obtain a stable pair $(\hat{X}^+, \hat{X}^-)$ in the original graph G , which we have seen is minimizing in $(G, \mathbf{M}^+, \mathbf{M}^-)$; i.e., $(\hat{X}^+, \hat{X}^-)$ minimizes the right-hand side of (4.1) among all stable pairs. In addition, exchange the roles of $\mathbf{M}^+$ and $\mathbf{M}^-$ in the algorithm to obtain another stable pair $(\hat{Y}^+, \hat{Y}^-)$, which minimizes

$$(6.1) \quad \rho^+(V \setminus Y^+) + \rho^-(V \setminus Y^-) + |Y^+ \cap Y^-| - \text{odd}^-(Y^+ \cap Y^-)$$

among all stable pairs (Y^+, Y^-) . For these stable pairs, the following theorem holds.

THEOREM 6.1. *For the stable pairs $(\hat{X}^+, \hat{X}^-)$ and $(\hat{Y}^+, \hat{Y}^-)$ defined above, set $V_D^+ = \hat{X}^+$, $V_A^+ = V \setminus \hat{X}^-$, $V_D^- = \hat{Y}^-$, and $V_A^- = V \setminus \hat{Y}^+$. Then, the following four items hold:*

1. $V_D^+ = \{v \mid v \in V, \exists \text{ maximum independent even factor } M, v \notin \text{cl}^+(\partial^+ M)\}$, $V_D^- = \{v \mid v \in V, \exists \text{ maximum independent even factor } M, v \notin \text{cl}^-(\partial^- M)\}$.
2. Every strongly connected component C in $G[V_D^+ \cap (V \setminus V_A^+)]$ or in $G[V_D^- \cap (V \setminus V_A^-)]$ is an odd component with $V(C) \in \mathcal{I}^+ \cap \mathcal{I}^-$.
3. If M is a maximum independent even factor in $(G, \mathbf{M}^+, \mathbf{M}^-)$, then the following three items hold:
 - (i) For every strongly connected component C in $G[V_D^+ \cap (V \setminus V_A^+)]$ or in $G[V_D^- \cap (V \setminus V_A^-)]$, $|M \cap A(C)| = |C| - 1$.
 - (ii) $|(V \setminus V_D^+) \cap \partial^+ M| = \rho^+(V \setminus V_D^+)$, $|V_A^+ \cap \partial^- M| = \rho^-(V_A^+)$, $|V_A^- \cap \partial^+ M| = \rho^+(V_A^-)$, and $|(V \setminus V_D^-) \cap \partial^- M| = \rho^-(V \setminus V_D^-)$.
 - (iii) For a vertex $v \in V_A^+ \cap \partial^- M$, there exists a vertex $u \in V_D^+$ with $(u, v) \in M$, and for a vertex $v \in V_A^- \cap \partial^+ M$, there exists a vertex $u \in V_D^-$ with $(v, u) \in M$.
4. Any minimizing stable pair (X^+, X^-) satisfies that $V_D^+ \subseteq X^+$ and $V_D^- \subseteq X^-$. Similarly, any stable pair (Y^+, Y^-) that minimizes (6.1) satisfies that $V_D^+ \subseteq Y^+$ and $V_D^- \subseteq Y^-$.

Proof. Here we show only the statements on the stable pair $(V_D^+, V \setminus V_A^+)$. Those on $(V \setminus V_A^-, V_D^-)$ can be proved similarly by exchanging the roles of $\mathbf{M}^+$ and $\mathbf{M}^-$ in the maximum independent even factor algorithm.

Assertion 2 follows from the fact that $(V_D^+, V \setminus V_A^+)$ is the minimizing stable pair which the algorithm finds. The fact also implies that

$$V_D^+ \subseteq \{v \mid v \in V, \exists \text{ maximum independent even factor } M, v \notin \text{cl}^+(\partial^+ M)\}.$$

For a maximum independent even factor M , define M_1 , M_2 , and M_3 by (4.2)–(4.4), where $X^+ = V_D^+$ and $X^- = V \setminus V_A^+$. As M is a maximum independent

even factor and (X^+, X^-) is a minimizing stable pair, (4.5) hold with equalities and $M_1 \cap M_2 = \emptyset$. One checks that these conditions induce

$$V_D^+ \supseteq \{v \mid v \in V, \exists \text{ maximum independent even factor } M, v \notin \text{cl}^+(\partial^+ M)\}$$

and assertion 3 (cf. the proof of Lemma 4.2).

Now what is left is assertion 4. Suppose there exists a minimizing stable pair (X^+, X^-) such that there is a vertex u with $u \in V_D^+ \setminus X^+$. As $u \in V_D^+$, there is a maximum independent even factor M with $u \notin \text{cl}^+(\partial^+ M)$. On the other hand, the stable pair (X^+, X^-) and M satisfy (4.5) with equality. The equation $|M_1| = \rho^+(V \setminus X^+)$, however, contradicts that $u \in V \setminus X^+$ and $u \notin \text{cl}^+(\partial^+ M)$. Therefore, for any minimizing stable pair (X^+, X^-) , it holds that $V_D^+ \subseteq X^+$. $\square$

Let us describe how Theorem 6.1 generalizes the Edmonds–Gallai decomposition for matchings and the principal partition for matroid intersection.

Edmonds–Gallai decomposition for matchings. Let $G = (V, E)$ be an undirected graph with the vertex set V and the edge set E , and let $\vec{G} = (V, A)$ be a symmetric digraph whose arc set is

$$A = \{(u, v), (v, u) \mid \exists \text{ edge between } u \text{ and } v \text{ in } E\}.$$

Consider an instance of the matching problem G and an instance of the independent even factor problem $(\vec{G}, \mathbf{M}^+, \mathbf{M}^-)$, where $\mathbf{M}^+$ and $\mathbf{M}^-$ are free matroids. Then, it follows that $V_D^+ = V_D^- = V_D$, $V_A^+ = V_A^- = V_A$, where

$$V_D = \{v \mid v \in V, \exists \text{ maximum matching which does not cover } v\},$$

$$V_A = \{v \mid v \in V \setminus V_D, v \text{ is adjacent to some vertex in } V_D\}.$$

By Theorem 6.1, the following statements on the undirected graph G hold:

1. The cardinality of a maximum matching in G is equal to

$$\begin{aligned} \frac{1}{2} \nu(\vec{G}, \mathbf{M}^+, \mathbf{M}^-) &= \frac{1}{2} (|V \setminus V_D^+| + |V_A^+| + |V_D^+ \cap (V \setminus V_A^+)| - \text{odd}^+(V_D^+ \cap (V \setminus V_A^+))) \\ &= \frac{1}{2} (|V| + |V_A| - \text{odd}(V_D)), \end{aligned}$$

where $\text{odd}(X)$ denotes the number of odd component in $G[X]$ ($X \subseteq V$).

2. Every connected component C in $G[V_D]$ is odd and factor-critical.

3. If M is a maximum matching, then the following hold:

- (i) For every connected component C in $G[V_D]$, C has $(|C| - 1)/2$ edges in M ;
- (ii) every vertex in $V \setminus V_D$ is covered by M ; and
- (iii) for a vertex $v \in V_A$, there exists a vertex $u \in V_D$ such that an edge in M connects u and v .

4. For $V_D' \subseteq V$, define

$$V_A(V_D') = \{v \mid v \in V \setminus V_D', v \text{ is adjacent to some vertex in } V_D'\}.$$

If V_D' minimizes $|V_A(V_D')| - \text{odd}(V_D')$, it follows that $V_D \subseteq V_D'$.

These statements are exactly those of the Edmonds–Gallai decomposition.

Principal partition for matroid intersection. Let $\mathbf{M}_1 = (V, \mathcal{I}_1, \rho_1)$ and $\mathbf{M}_2 = (V, \mathcal{I}_2, \rho_2)$ be an instance of the matroid intersection problem, and define

$$V_D^1 = \{v \mid v \in V, \exists \text{ maximum common independent set } I, v \notin \text{cl}_1(I)\},$$

$$V_D^2 = \{v \mid v \in V, \exists \text{ maximum common independent set } I, v \notin \text{cl}_2(I)\}.$$

Recalling the instance of the independent even factor problem $(G, \mathbf{M}^+, \mathbf{M}^-)$ to which we reduced the matroid intersection instance in section 4, we have that

$$V_D^+ = V_A^+ = \bigcup_{v \in V_D^1} \{v_1, v_2\}, \quad V_D^- = V_A^- = \bigcup_{v \in V_D^2} \{v_1, v_2\}.$$

Applying Theorem 6.1 to $(G, \mathbf{M}^+, \mathbf{M}^-)$, we obtain the following statements.

1. The cardinality of a maximum common independent set in $\mathcal{I}_1 \cap \mathcal{I}_2$ is equal to

$$\nu(G, \mathbf{M}^+, \mathbf{M}^-) = \rho_1(V \setminus V_D^1) + \rho_2(V_D^1) = \rho_1(V_D^2) + \rho_2(V \setminus V_D^2).$$

2. If I is a maximum common independent set, then it holds that $|(V \setminus V_D^1) \cap I| = \rho_1(V \setminus V_D^1)$, $|V_D^1 \cap I| = \rho_2(V_D^1)$, $|V_D^2 \cap I| = \rho_1(V_D^2)$, and $|(V \setminus V_D^2) \cap I| = \rho_2(V \setminus V_D^2)$.

3. Any subset $X \subseteq V$ that minimizes $\rho_1(V \setminus X) + \rho_2(X)$ satisfies that $V_D^1 \subseteq X$ and $V_D^2 \subseteq V \setminus X$. That is, V_D^1 (resp., V_D^2) is the minimal (resp., maximal) subset minimizing the submodular function $\rho_1(V \setminus X) + \rho_2(X)$.

Acknowledgments. The authors are thankful to Kazuo Murota for his helpful comments.

REFERENCES

- [1] C. BERGE, *Sur le couplage maximum d'un graphe*, C. R. Acad. Sci. Paris, 247 (1958), pp. 258–259.
- [2] D. COPPERSMITH AND S. WINOGRAD, *Matrix multiplication via arithmetic progressions*, J. Symbolic Comput., 9 (1990), pp. 251–280.
- [3] W. H. CUNNINGHAM AND J. F. GEELEN, *The optimal path-matching problem*, Combinatorica, 17 (1997), pp. 315–337.
- [4] W. H. CUNNINGHAM AND J. F. GEELEN, *Vertex-Disjoint Dipaths and Even Dicircuits*, manuscript.
- [5] J. EDMONDS, *Paths, trees, and flowers*, Canad. J. Math., 17 (1965), pp. 449–467.
- [6] J. EDMONDS, *Submodular functions, matroids, and certain polyhedra*, in Combinatorial Structures and Their Applications, R. Guy, H. Hanani, N. Sauer, and J. Schönheim, eds., Gordon and Breach, New York, 1970, pp. 69–87.
- [7] J. EDMONDS, *Matroid intersection*, Ann. Discrete Math., 4 (1979), pp. 39–49.
- [8] A. FRANK AND L. SZEGŐ, *Note on the path-matching formula*, J. Graph Theory, 41 (2002), pp. 110–119.
- [9] T. GALLAI, *Maximale Systeme unabhängiger Kanten*, Magyar Tud. Akad. Mat. Kutató Int. Közl., 9 (1964), pp. 401–413.
- [10] N. J. A. HARVEY, *Algebraic structures and algorithms for matching and matroid problems*, in Proceedings of the 47th Annual Symposium on Foundations of Computer Science, IEEE, Washington, DC, 2006, pp. 531–542.
- [11] N. J. A. HARVEY, *Algebraic algorithms for matching and matroid problems*, SIAM J. Comput., to appear.
- [12] M. IRI, *A review of recent work in Japan on principal partitions of matroids and their applications*, Ann. New York Acad., 319 (1979), pp. 306–319.
- [13] M. IRI AND S. FUJISHIGE, *Use of matroid theory in operations research, circuits and systems theory*, Internat. J. Systems Sci., 12 (1981), pp. 27–54.

- [14] T. KIRÁLY AND M. MAKAI, *On polyhedra related to even factors*, in Integer Programming and Combinatorial Optimization, Lecture Notes in Comput. Sci. 3064, D. Bienstock and G. Nemhauser, eds., Springer-Verlag, Berlin, 2004, pp. 416–430.
- [15] E. L. LAWLER, *Matroid intersection algorithms*, Math. Programming, 9 (1975), pp. 31–56.
- [16] G. PAP, *A Constructive Approach to Matching and Its Generalizations*, Ph.D. thesis, Eötvös Loránd University, Budapest, 2006.
- [17] G. PAP, *Combinatorial algorithms for matchings, even factors and square-free 2-factors*, Math. Program., 110 (2007), pp. 57–69.
- [18] G. PAP AND L. SZEGŐ, *On the maximum even factor in weakly symmetric graphs*, J. Combin. Theory Ser. B, 91 (2004), pp. 201–213.
- [19] H. PERFECT, *Independence spaces and combinatorial problems*, Proc. London Math. Soc. (3), 19 (1969), pp. 17–30.
- [20] A. SCHRIJVER, *Matroids and linking systems*, J. Combin. Theory Ser. B, 26 (1979), pp. 349–369.
- [21] B. SPILLE AND L. SZEGŐ, *A Gallai-Edmonds-type structure theorem for path-matchings*, J. Graph Theory, 46 (2004), pp. 93–102.
- [22] B. SPILLE AND R. WEISMANTEL, *A generalization of Edmonds’ matching and matroid intersection algorithms*, in Integer Programming and Combinatorial Optimization, Lecture Notes in Comput. Sci. 2337, W. J. Cook and A. S. Schulz, eds., Springer-Verlag, Berlin, 2002, pp. 9–20.
- [23] K. TAKAZAWA, *A weighted even factor algorithm*, Math. Program., 115 (2008), pp. 223–237.
- [24] W. T. TUTTE, *The factorization of linear graphs*, J. London Math. Soc., 22 (1947), pp. 107–111.

MULTIDIMENSIONAL MANHATTAN STREET NETWORKS*

F. COMELLAS[†], C. DALFÓ[†], AND M. A. FIOL[†]

Abstract. We formally define the n -dimensional Manhattan street network M_n —a special case of an n -regular digraph—and we study some of its structural properties. In particular, we show that M_n is a Cayley digraph, which can be seen as a subgroup of the n -dimensional version of the wallpaper group pgg . These results induce a useful new representation of M_n , which can be applied to design a local (shortest-path) routing algorithm and to study some other metric properties, such as the diameter. We also show that the n -dimensional Manhattan street networks are Hamiltonian and, in the standard case (that is, in dimension two), we give sufficient conditions for a 2-dimensional Manhattan street network to be decomposable into two arc-disjoint Hamiltonian cycles.

Key words. Manhattan street network, line digraph, Cayley digraph, diameter, Hamiltonian cycle

AMS subject classifications. 05C20, 05C25, 05C12, 05C45, 90B10

DOI. 10.1137/07068446X

1. Introduction. The study of a class of toroidal directed networks, commonly known in the literature as Manhattan street networks, has received significant attention since they were introduced independently (in different contexts) by Morillo, Fiol, and Fàbrega [18] and Maxemchuk [16], as a unidirectional regular mesh structure resembling locally the topology of the avenues and streets of Manhattan (or *l'Exemple* in downtown Barcelona); see Figure 2.1. More precisely, given two even integers N_i , $i = 1, 2$, the (2-dimensional) Manhattan street network $M_2 = M(N_1, N_2)$ is a 2-regular digraph with vertices $\mathbf{u} = (u_1, u_2)$, $u_i \in \mathbb{Z}_{N_i}$, and vertex $\mathbf{u}$ has two outgoing arcs: one horizontal $(u_1 + (-1)^{u_2}, u_2)$, and the other vertical $(u_1, u_2 + (-1)^{u_1})$.

A Manhattan street network must not be confused with the so-called “Manhattan network,” which, for a given set of points in the plane, consists of a set of axis-parallel line segments whose union contains an x - and y -monotone path for each pair of points (see, for instance, the paper of Benkert et al. [5]).

Back to the Manhattan street networks, Morillo, Fiol, and Fàbrega [18] related these networks to plane tessellations, and this association facilitates the study of some metric properties. Most of the work on Manhattan street networks has been devoted to the 2-dimensional case. Khasnabish [15] and Chung and Agrawal [7] gave the average distance of $M(N_1, N_2)$, which has the following expression when N_1 and N_2 are both divisible by 4:

$$(1.1) \quad \overline{D} = \frac{N_1 + N_2}{4} + 1 - \frac{4}{N_1 N_2}.$$

Maxemchuk [16] studied the generation of routing schemes of the Manhattan street network.

*Received by the editors March 6, 2007; accepted for publication (in revised form) May 29, 2008; published electronically September 19, 2008. This research was supported by the Ministry of Science and Technology (Spain) and the European Regional Development Fund (ERR) under projects MTM2005-08990-C02-01 and TEC2005-03575.

<http://www.siam.org/journals/sidma/22-4/68446.html>

[†]Departament de Matemàtica Aplicada IV, Universitat Politècnica de Catalunya, Jordi Girona 1–3, Mòdul C3, Campus Nord, 08034 Barcelona, Catalonia, Spain (comellas@ma4.upc.edu, cdalfo@ma4.upc.edu, fiol@ma4.upc.edu).

These results have usually been inspired by conjectures supported by computer simulations. Also, the study of spanning trees (by Chung and Agrawal [7]) in the Manhattan street network $M(N_1, N_2)$ has allowed the computation of its diameter:

$$(1.2) \quad D = \begin{cases} N_1/2 + N_2/2 + 1 & \text{if } N_1 \equiv N_2 \equiv 0 \pmod{4}, \\ N_1/2 + N_2/2 & \text{otherwise,} \end{cases}$$

and the design of a multiport broadcasting algorithm. More recently, Varvarigos [21] evaluated again the mean internodal distance of such a network, and also provided a shortest path routing algorithm and two edge-disjoint Hamiltonian cycles in the 2-dimensional case $N_1 = N_2 = N$.

The multidimensional natural extension of the Manhattan street networks was considered by Banerjee, Jain, and Shah [2] and Banerjee, Mukherjee, and Ramamurthy [3] with the determination of the average distance of a 3-dimensional Manhattan street network $M(N_1, N_2, N_3)$, which, in the case $N_i \equiv 0 \pmod{4}$, $i = 1, 2, 3$, turns out to be

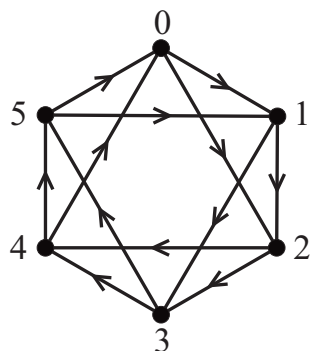
$$(1.3) \quad \overline{D} = \frac{N_1 + N_2 + N_3}{4} + 1 - \frac{4}{N_1 N_2 N_3},$$

a result to be compared with (1.1). Moreover, in the same paper the authors gave a conjecture for higher dimensions although, neither in the 3-dimensional nor in the n -dimensional case, no formal algebraic definition was provided, but only an idea of how the adjacencies should be. Namely, quoting Banerjee, Mukherjee, and Ramamurthy [3], “...in a torus network with unidirectional links, the number of nodes in each dimension is even and all adjacent links point in opposite directions.”

Finally, Chung and Agrawal [8] studied the diameter and provided routing schemes for a 3-dimensional construction based on 2-dimensional Manhattan street networks, although the proposed resulting network is not strictly a 3-dimensional Manhattan street network.

In this paper, we give a formal definition of an n -dimensional Manhattan street network M_n and its main properties (section 2). We show that M_n is a Cayley digraph (section 3). This fact allows us to introduce a new representation of M_n (section 3.1), which can be applied to prove a number of results, such as the design of a local (shortest-path) routing algorithm (section 4.2). We also analyze some of its distance-related parameters, such as the diameter in the cases 2-dimensional (section 4.1) and n -dimensional (section 4.3). Finally, we provide some details on the cycle structure and a proof of the Hamiltonicity of these digraphs, and we give sufficient conditions for a 2-dimensional Manhattan street network to be decomposable into two arc-disjoint Hamiltonian cycles (section 5). Consequently, according to the survey of Banerjee, Jain, and Shah [2], such n -dimensional Manhattan street networks would be good candidates for modeling regular logical (multihoop) topologies for lightwave networks.

Some notation on digraphs. Recall that a digraph $G = (V, A)$ consists of a set of vertices V , together with a set of arcs A , which can be seen as ordered pairs of vertices $A \subset V \times V = \{(u, v) : u, v \in V\}$. An arc (u, v) is usually depicted by an arrow with *initial vertex* u and *terminal vertex* v ; that is, $u \rightarrow v$. In this case we say that v is *adjacent from* u or that u is *adjacent to* v . The *out-neighborhood* $\Gamma^+(u)$ (respectively, *in-neighborhood* $\Gamma^-(u)$) of a vertex u is the set of vertices adjacent from (respectively, to) u . The *outdegree* of vertex u is $\delta^+(u) = |\Gamma^+(u)|$, whereas its *indegree* is $\delta^-(u) = |\Gamma^-(u)|$. (Equivalently $\delta^+(u)$ and $\delta^-(u)$ are the numbers of arcs with initial and terminal vertex u , respectively.) Then G is δ -regular when $\delta^-(u) = \delta^+(u) = \delta$ for

FIG. 1.1. The Cayley digraph $\text{Cay}(\mathbb{Z}_6, \{1, 2\})$.

every vertex $u \in V$. Given a digraph $G = (V, A)$, its *converse* digraph $\overline{G} = (V, \overline{A})$ is obtained from G by reversing all of the orientations of the arcs in A ; that is, $(u, v) \in \overline{A}$ if and only if $(v, u) \in A$.

A (directed) *path* $\mathbf{p}$ is a subdigraph of the form $u_0 \rightarrow u_1 \rightarrow u_2 \rightarrow \cdots \rightarrow u_\ell$ with *initial* vertex u_0 , *terminal* vertex u_ℓ , and *interior* vertices $u_1, u_2, \dots, u_{\ell-1}$. The *length* of $\mathbf{p}$ is defined as the number of arcs: $\|\mathbf{p}\| = \ell$. If $\mathbf{p}$ has the minimum number of arcs, then its length is the *distance* $\text{dist}(u_1, u_\ell)$ from the initial vertex u_1 to the terminal vertex u_ℓ . The *eccentricity* $\text{ecc}(u)$ of vertex u is then defined as the maximum distance from u to all of the other vertices. Then, the *diameter* of a digraph G , denoted by $D = D(G)$, is the maximum of all of the eccentricities of its vertices. A digraph is *strongly connected* when, for any pair of vertices u, v , there is a directed path from u to v (in other words, $D(G)$ is finite). Note that G is strongly connected if and only if its converse digraph $\overline{G}$ is. More precisely, $D(G) = D(\overline{G})$.

Given a group Γ with (finite) generating set Δ , the *Cayley digraph* $\text{Cay}(\Gamma, \Delta)$ has vertices representing the elements of Γ and arcs of the form $(g, h\partial)$ where $g, h \in \Gamma$ and $\partial \in \Delta$. The Cayley digraph $\text{Cay}(\Gamma, \Delta)$ is a vertex-transitive strongly connected regular digraph. As an example, Figure 1.1 shows the Cayley digraph of the cyclic group $\Gamma = \mathbb{Z}_6$ with generating set $\Delta = \{1, 2\}$.

A *homomorphism* Ψ from a digraph G to a digraph H is a mapping from the vertex set of G to the vertex set of H preserving adjacencies; that is, if (u, v) is an arc of G , then $(\Psi(u), \Psi(v))$ is an arc of H . Moreover, if both digraphs are arc-colored (all arcs with the same initial or terminal vertex receive different colors) and Ψ preserves the colors, then we say that Ψ is a *colored homomorphism* or simply that it is *color-preserving*. A (colored) *automorphism* of a digraph G is simply a (colored) homomorphism from G to itself (note that, in this case, the above mapping corresponds to a permutation of the vertices of G). The set of (colored) automorphisms of G forms a group called the (colored) *automorphism group* of the digraph. The well-known characterization result of Sabidussi [19] states that a digraph is a Cayley digraph (for some pair Γ, Δ) if and only if its automorphism group contains a regular subgroup (that is, a subgroup which, for every pair of vertices u, v , contains exactly one permutation mapping u into v).

Other standard definitions and basic results about graphs and digraphs not recalled here can be found, for instance, in the textbooks of Bang-Jensen and Gutin [4] or Chartrand and Lesniak [6].

2. Definition and structure. Recall that the standard Manhattan street network $M(N_1, N_2)$ is defined as a 2-regular digraph in the following way. Every vertex is represented by a pair of integers $\mathbf{u} = (u_1, u_2)$, with $0 \leq u_i \leq N_i - 1$, for some even integers N_i , $i = 1, 2$, and vertex $\mathbf{u}$ has two outgoing arcs: one horizontal ($u_1 \pm 1, u_2$), and the other vertical ($u_1, u_2 \pm 1$) (where the sign depends on the parity of the other component, and arithmetic must be understood to be modulo N_i). More precisely, a horizontal arc points *east* (respectively, *west*) when it is on an *even* (respectively, *odd*) row. Similarly, a vertical arc points *north* (respectively, *south*) if it is on an *even* (respectively, *odd*) column.

Locally the structure is as shown in Figure 2.1, and corresponds to a standard pattern for the allowed traffic directions in some neighborhoods of our modern cities, like New York or Barcelona, with their system of straight orthogonal streets. In many papers, such as those of Maxemchuk [16], Chung and Agrawal [7], and Comellas and Dalfo [9], the above mentioned toroidal version of M_2 was considered, whereas Morillo, Fiol, and Fàbrega [18] aimed to construct locally Manhattan street networks with the maximum number of vertices for a given diameter.

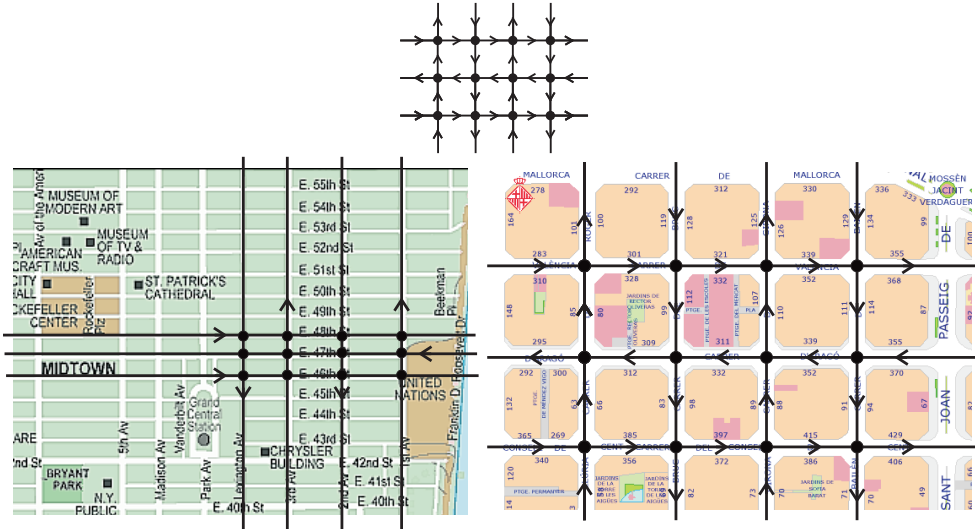


FIG. 2.1. The local pattern of a Manhattan street network and two real-life examples: Orthogonal streets of Manhattan and l'Eixample in Barcelona.

A formal definition of the toroidal version for the general n -dimensional case is the following.

DEFINITION 2.1. Given n even positive integers $N_1, N_2, \dots, N_n$, the n -dimensional Manhattan street network $M_n = M(N_1, N_2, \dots, N_n)$ is a digraph with vertex set $V(M_n) = \mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2} \times \dots \times \mathbb{Z}_{N_n}$. Thus, each of its vertices is represented by an n -vector $\mathbf{u} = (u_1, u_2, \dots, u_n)$, with $0 \leq u_i \leq N_i - 1$, $i = 1, 2, \dots, n$. The arc set $A(M_n)$ is defined by the following adjacencies (here called i -arcs):

$$(2.1) \quad (u_1, \dots, u_i, \dots, u_n) \rightarrow (u_1, \dots, u_i + (-1)^{\sum_{j \neq i} u_j}, \dots, u_n) \quad (1 \leq i \leq n).$$

Therefore, M_n is an n -regular digraph on $N = \prod_{i=1}^n N_i$ vertices.

In particular, notice that when $N_i = 2$, $1 \leq i \leq n$, we always have $(-1)^{\sum_{j \neq i} u_j} = 1$. Hence, in this case the n -dimensional Manhattan street network is isomorphic to the symmetric digraph Q_n^* , with Q_n being the hypercube of dimension n (the n -cube).

Some other simple consequences of the definition of M_n are presented in the following lemma.

LEMMA 2.2. *Every n -dimensional Manhattan street network $M_n = M(N_1, N_2, \dots, N_n)$ satisfies the following properties:*

- (a) *Given any permutation σ of the numbers $N_1, N_2, \dots, N_n$, say $P_1, P_2, \dots, P_n$, the Manhattan street networks M_n and $M_n^\sigma = M(P_1, P_2, \dots, P_n)$ are isomorphic digraphs.*
- (b) *M_n is isomorphic to its converse $\overline{M}_n$.*
- (c) *For any $n - k$ fixed integers $x_i \in \mathbb{Z}_{N_i}$, $i = k + 1, k + 2, \dots, n$, the subdigraph of M_n induced by the vertices of the form $(u_1, u_2, \dots, u_k, x_{k+1}, \dots, x_n)$ is either the k -dimensional Manhattan street network $M_k = M(N_1, N_2, \dots, N_k)$ or its converse $\overline{M}_k$, depending on whether $\alpha := \sum_{i=k+1}^n x_n$ is even or odd, respectively.*
- (d) *M_n is both a 2^n -partite and bipartite digraph.*
- (e) *There exists a homomorphism from M_n to the symmetric digraph of the hypercube Q_n^* .*

Proof. The result in (a) is clear with σ acting on the (components of the) vertices of M_n . To prove (b), note that in the converse digraph $\overline{M}_n$, the adjacencies are just

$$(2.2) \quad (u_1, \dots, u_i, \dots, u_n) \rightarrow (u_1, \dots, u_i - (-1)^{\sum_{j \neq i} u_j}, \dots, u_n) \quad (1 \leq i \leq n).$$

Hence, it is readily checked that the mapping $\varphi : V(M_n) \rightarrow V(\overline{M}_n)$ defined by $\varphi(\mathbf{u}) = -\mathbf{u}$ is the required isomorphism. The result in (c) follows from the “converse adjacencies” in (2.2) and the fact that $(-1)^{\sum_{j=1, j \neq i}^k u_i + \alpha} = \pm(-1)^{\sum_{j=1, j \neq i}^k u_i}$ depending on the parity of α . Moreover, (d) holds since M_n is a 2^n -partite digraph with independent sets $V_{\mathbf{b}}$, where $\mathbf{b} = (b_1, b_2, \dots, b_n)$ is an n -binary string. A vertex $\mathbf{u} = (u_1, u_2, \dots, u_n)$ belongs to $V_{\mathbf{b}}$ when the parities of u_i and b_i coincide for every $1 \leq i \leq n$. In particular, M_n is bipartite with stable vertex sets V_0 and V_1 constituted by the vertices whose corresponding binary string represents an even or odd number, respectively. Finally, the claimed homomorphism in (e) is simply

$$(u_1, u_2, \dots, u_n) \mapsto (\pi(u_1), \pi(u_2), \dots, \pi(u_n)),$$

where the parity function π takes the expected values in $\{0, 1\}$. $\square$

Concerning property (d), let us mention that Morillo, Fiol, and Fàbrega [18] introduced the local structure of a standard (2-dimensional) Manhattan street network as a type of 4-partite digraph in the following way: Let a digraph $G = (V, A)$ have order $N = |V|$, a multiple of 4, with $V = V_0 \cup V_1 \cup V_2 \cup V_3$, where

$$(2.3) \quad V_j = \{i : 0 \leq i \leq N - 1; i \equiv j \pmod{4}\} \quad (0 \leq j \leq 3),$$

with every vertex i being adjacent to the vertices $i + a_j$, $i + b_j \pmod{N}$, for some given integers $a_j \equiv 3$, and $b_j \equiv 1 \pmod{4}$ satisfying

$$a_0 + a_2 \equiv -a_1 - a_3 \equiv b_0 + b_2 \equiv -b_1 - b_3 \pmod{N}.$$

See Figure 2.2 to check whether the above conditions impose a local Manhattan street structure.

The line digraph structure. Here we show that the standard Manhattan street network, which is the 2-dimensional case M_2 , has the structure of a line digraph. To the knowledge of the authors, this relevant fact has not been published before,

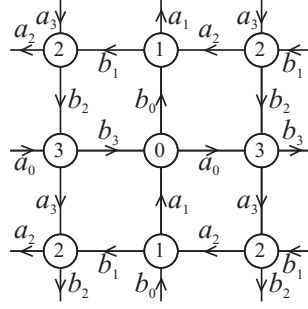


FIG. 2.2. An alternative definition of the local structure of a 2-dimensional Manhattan street network seen as a 4-partite digraph. (All vertices in V_j are denoted by j .)

although it was used in the thesis by Villar [22]. As a consequence, M_2 can be seen as the line digraph of a digraph M'_2 whose order is one half of the order of M_2 and, what is more important, some properties of M_2 can be derived from those of M'_2 .

First, recall that given a digraph $G = (V, A)$ with n vertices and m arcs, its line digraph $L(G) = (V_L, A_L)$ has vertices representing the arcs of G ; so that we identify each vertex $ij \in V_L$ with the arc $(i, j) \in A$, and its adjacencies are naturally induced by the arc adjacencies in G . More precisely, vertex $ij \in V_L$ is adjacent to vertex jk since the arc $(i, j) \in A$ has the same terminal vertex as the initial vertex of (j, k) . Thus, the order of $L(G)$ equals the size m of G and, if G is δ -regular, so is $L(G)$ and it has δn vertices. Also, it is known that if G is different from a (directed) cycle and has diameter D , then its line digraph $L(G)$ has diameter $D + 1$; see Fiol, Yebra, and Alegre [13]. Some other interesting metric properties of line digraphs were given by Comellas et al. [10] and Fiol, Yebra, and Alegre [13]. With respect to the spectral properties, Balbuena et al. [1] showed that the spectrum of the line digraph $L(G)$ has the same nonzero eigenvalues as G , including (algebraic) multiplicities. In fact, the eigenvalue sets differ only in the number of zeros since their corresponding characteristic polynomials, $p_{L(G)}$ and p_G , satisfy $p_{L(G)}(x) = x^{m-n}p_G(x)$.

As stated before, the next result shows that the line digraph structure is inherent to the 2-dimensional case.

LEMMA 2.3. *For any N_1, N_2 , the 2-dimensional Manhattan street network M_2 is a line digraph.*

Proof. It suffices to check Heuchenne's condition [14], which says that a digraph is a line digraph if and only if it has no multiple arc and the out-neighbor (in-neighbor) sets of any two of its vertices are either identical or disjoint. With this aim, assume that two different vertices, $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$, have a common out-neighbor $\mathbf{w}$. Then we claim that the arcs $\mathbf{u} \rightarrow \mathbf{w}$ and $\mathbf{v} \rightarrow \mathbf{w}$ must be of different type; that is, one 1-arc and the other a 2-arc. Otherwise, if both were 1-arcs, say, we would have

$$\mathbf{w} = (u_1 + (-1)^{u_2}, u_2) = (v_1 + (-1)^{v_2}, v_2),$$

which leads to $u_2 = v_2$ and $u_1 = v_1$, so that $\mathbf{u} = \mathbf{v}$ against the hypothesis. The same contradiction is reached if we suppose that both adjacencies are 2-arcs. Thus assume, without loss of generality, that $\mathbf{u} \rightarrow \mathbf{w}$ is a 1-arc and $\mathbf{v} \rightarrow \mathbf{w}$ is a 2-arc. Then,

$$\mathbf{w} = (u_1 + (-1)^{u_2}, u_2) = (v_1, v_2 + (-1)^{v_1}),$$

whence

$$\begin{aligned} u_1 &= v_1 - (-1)^{u_2} = v_1 - (-1)^{v_2+(-1)^{v_1}} = v_1 + (-1)^{v_2}, \\ v_2 &= u_2 - (-1)^{v_1} = u_2 - (-1)^{u_1+(-1)^{u_2}} = u_2 + (-1)^{u_1}, \end{aligned}$$

which imply the existence of another common out-neighbor w' of u and v , such that $u \rightarrow w'$ is a 2-arc and $v \rightarrow w'$ is a 1-arc, namely

$$w' = (u_1, u_2 + (-1)^{u_1}) = (v_1 + (-1)^{v_2}, v_2).$$

We then get the claimed result. $\square$

Summarizing, we have seen that two different vertices u, v have the same out-neighborhood if and only if they are of the form

$$u = (a, b), \quad v = (a + (-1)^b, b + (-1)^a),$$

for some integers $a \in \mathbb{Z}_{N_1}$, $b \in \mathbb{Z}_{N_2}$. Then, according to the parity (equal “●” or distinct “■” in Figure 2.3) of a, b , we have the two possible situations shown on the left of the figure. From this, notice that the digraph G which M_2 comes from (that is, $M_2 = L(G)$) is also bipartite, with independent sets $\{\bullet\}$ and $\{\blacksquare\}$. In fact, the infinite pattern corresponds to another planar crystallographic group; namely, the one denoted by $p4$.

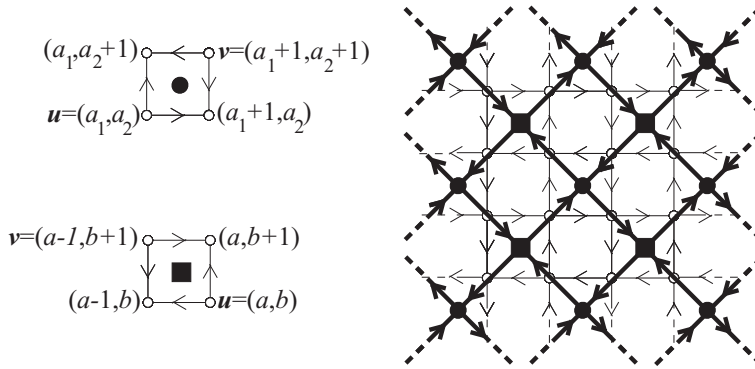


FIG. 2.3. The local structure of a 2-dimensional Manhattan street network M_2 (thin lines and white vertices) and the digraph (thick lines and black vertices) whose line digraph is M_2 .

It is worth noting that the property of being line digraphs is not shared, in general, by the Manhattan street networks with dimension greater than two. For instance, $M(8, 6, 10)$ does not satisfy Heuchenne’s condition since the out-neighborhoods

$$\Gamma^+((1, 1, 5)) = \{(2, 1, 5), (1, 2, 5), (1, 1, 6)\} \text{ and}$$

$$\Gamma^+((6, 1, 4)) = \{(1, 1, 6), (2, 2, 6), (2, 1, 5)\}$$

are neither equal nor disjoint.

Two simple consequences of Lemma 2.3 are the following: First, M_2 is Hamiltonian, as it is the line digraph of a 2-regular, and thus is a Eulerian digraph (see, for instance, Chartrand and Lesniak [6]). In fact, in section 5 we show that the Hamiltonian property is shared by all n -dimensional Manhattan street networks. Second, from the commented results of Balbuena et al. [1] (see also Fiol and Mitjana [12]) it follows that the spectrum of $M_2 = M(N_1, N_2)$ has the eigenvalue 0 with (geometric) multiplicity at least $N_1 N_2 / 2$.

3. The colored automorphism group. Here we investigate the symmetries of the Manhattan street networks.

THEOREM 3.1. *The n -dimensional Manhattan street network M_n is the Cayley digraph of the group*

$$(3.1) \quad \Gamma = \langle a_1, a_2, \dots, a_n \mid a_i^{N_i} = (a_i a_j)^2 = (a_i a_j^{-1})^2 = 1, \ i, j = 1, \dots, n \rangle.$$

Proof. Let us show that the mappings ϕ_j , $1 \leq j \leq n$, defined by

$$(3.2) \quad \phi_j(u_1, \dots, u_j, \dots, u_n) = (-u_1, \dots, -u_{j-1}, u_j + 1, -u_{j+1}, \dots, -u_n),$$

are all isomorphisms of M_n mapping i -arcs into i -arcs. Indeed, let $\gamma_i^+ \mathbf{u}$ denote the vertex adjacent from vertex $\mathbf{u} = (u_1, \dots, u_n)$ through the i -arc. Then, assuming first that $j \neq i$, say $j < i$,

$$\begin{aligned} \phi_j(\gamma_i^+ \mathbf{u}) &= \phi_j(u_1, \dots, u_j, \dots, u_i + (-1)^{\sum_{k \neq i} u_k}, \dots, u_n) \\ &= (-u_1, \dots, u_j + 1, \dots, -u_i + (-1)^{1 + \sum_{k \neq i} u_k}, \dots, -u_n) \\ &= \gamma_i^+(-u_1, \dots, u_j + 1, \dots, -u_i, \dots, -u_n) \\ &= \gamma_i^+ \phi_j(\mathbf{u}). \end{aligned}$$

Otherwise, if $j = i$, then we have

$$\begin{aligned} \phi_i(\gamma_i^+ \mathbf{u}) &= \phi_i(u_1, \dots, u_i + (-1)^{\sum_{k \neq i} u_k}, \dots, u_n) \\ &= (-u_1, \dots, u_i + 1 + (-1)^{\sum_{k \neq i} u_k}, \dots, -u_n) \\ &= \gamma_i^+(-u_1, \dots, u_i + 1, \dots, -u_n) \\ &= \gamma_i^+ \phi_i(\mathbf{u}). \end{aligned}$$

Therefore, the mappings ϕ_j , $1 \leq j \leq n$, are all color-preserving automorphisms of M_n . Let us now show that the permutation group $\langle \phi_i \mid 1 \leq i \leq n \rangle$ acts transitively on the set $\mathbb{Z} \times \mathbb{Z} \times \dots \times \mathbb{Z}$ (and, hence, also on the vertex set of $M_n = M(N_1, N_2, \dots, N_n)$). With this aim, it is enough to show that any vertex $\mathbf{u} = (u_1, u_2, \dots, u_n)$ can be mapped into vertex $\mathbf{0} = (0, 0, \dots, 0)$. Let us first distinguish two cases depending on the sign of u_n (the superscripts of the isomorphisms indicate how many times they are applied):

- $u_n < 0$:

$$(u_1, u_2, \dots, u_n) \xrightarrow{\phi_n^{|u_n|}} (\pm u_1, \pm u_2, \dots, 0);$$

- $u_n > 0$:

$$\begin{aligned} (u_1, \dots, u_i, \dots, u_n) &\xrightarrow{\phi_i} (-u_1, \dots, u_i + 1, \dots, -u_n) \\ &\xrightarrow{\phi_n^{u_n}} (\pm u_1, \dots, \pm(u_i + 1), \dots, 0), \end{aligned}$$

where $i < n$, and, in both cases, the sign in $\pm$ depends on the parity of u_n .

Then, by applying the same procedure $n - 1$ times, we obtain a vertex of the form $(v_1, 0, \dots, 0)$. From this vertex, the desired path is obtained by taking into

consideration the following cases. Let k be a nonnegative integer:

- $v_1 = -k$:

$$(-k, 0, \dots) \xrightarrow{\phi_1^k} (0, 0, \dots).$$

- $v_1 = 2k + 1$:

$$(2k + 1, 0, \dots) \xrightarrow{\phi_2} (-2k - 1, 1, \dots) \xrightarrow{\phi_1^{2k+1}} (0, -1, \dots) \xrightarrow{\phi_2} (0, 0, \dots).$$

- $v_1 = 2k$:

$$(2k, 0, \dots) \xrightarrow{\phi_2} (-2k, 1, \dots) \xrightarrow{\phi_1^{2k}} (0, 1, \dots) \xrightarrow{\phi_1} (1, -1, \dots) \\ \xrightarrow{\phi_2} (-1, 0, \dots) \xrightarrow{\phi_1} (0, 0, \dots).$$

Thus, the group $\Gamma = \langle \phi_1, \dots, \phi_n \rangle$ is a regular subgroup of the automorphism group $\text{Aut } M_n$ and M_n is a Cayley digraph. Concerning the structure of Γ , let us check only the second defining relation in (3.1), as the others are proved similarly,

$$\begin{aligned} (\phi_i \phi_j)^2(\mathbf{u}) &= \phi_i \phi_j \phi_i \phi_j(u_1, \dots, u_i, \dots, u_j, \dots, u_n) \\ &= \phi_i \phi_j \phi_i(-u_1, \dots, -u_i, \dots, u_j + 1, \dots, -u_n) \\ &= \phi_i \phi_j(u_1, \dots, -u_i + 1, \dots, -u_j - 1, \dots, u_n) \\ &= \phi_i(-u_1, \dots, u_i - 1, \dots, -u_j, \dots, -u_n) \\ &= (u_1, \dots, u_i, \dots, u_j, \dots, u_n) = \mathbf{u}. \quad \square \end{aligned}$$

This structural result has some appealing consequences, the most evident being the following corollary.

COROLLARY 3.2. *The n -dimensional Manhattan street network M_n is a vertex-symmetric (but not necessarily arc-symmetric) digraph.*

The reader familiar with group theory will have already noted that, in the 2-dimensional case, the presentation in (3.1) without the first generating relations $a_1^{N_1} = a_2^{N_2} = 1$ corresponds to the (plane) crystallographic group $p\bar{g}g$; see, for instance, Coxeter and Moser [11]. Consequently, we have the following result.

COROLLARY 3.3. *The underlying Cayley digraph of the 2-dimensional Manhattan street network M_2 , with respect to the arc-coloring defined in (2.1), is a (normal) subgroup of the crystallographic group $p\bar{g}g$.*

In particular, for $N_1 = n$, $N_2 = 2$, we obtain the dihedral group D_n (the symmetry group in two dimensions of a regular n -gon). See Figure 3.1 for the standard drawing of $M(8, 2)$ and the Cayley digraph of D_8 .

3.1. An alternative definition. The previous results imply an alternative definition of the Manhattan street networks.

DEFINITION 3.4. *The vertex set of $M_n = M(N_1, N_2, \dots, N_n)$ is, as stated previously, $\mathbb{Z}_{N_1} \times \dots \times \mathbb{Z}_{N_n}$ and the (i -)arcs ($1 \leq i \leq n$) are now*

$$(3.3) \quad (u_1, \dots, u_i, \dots, u_n) \rightarrow (-u_1, \dots, -u_{i-1}, u_i + 1, -u_{i+1}, \dots, -u_n).$$

LEMMA 3.5. *The graphs defined by (2.1) and (3.3) are isomorphic.*

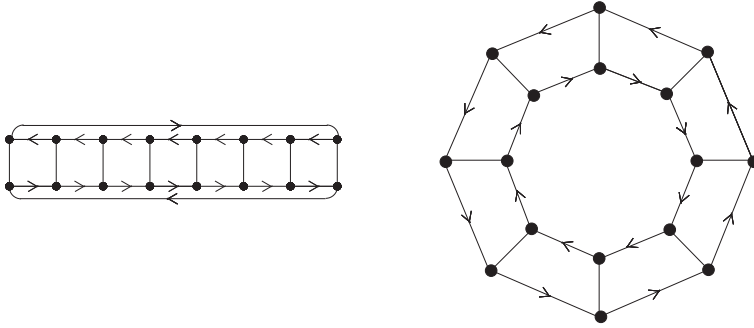


FIG. 3.1. Standard representations of the Manhattan street network $M(8,2)$ and the Cayley digraph of D_8 (each undirected edge stands for two opposite arcs).

Proof. We introduce the isomorphism from the standard definition to the new representation as follows:

$$(3.4) \quad \Psi(u_1, \dots, u_i, \dots, u_n) \\ = ((-1)^{\sum_{j \neq 1} u_j} u_1, \dots, (-1)^{\sum_{j \neq i} u_j} u_i, \dots, (-1)^{\sum_{j \neq n} u_j} u_n).$$

This application preserves the adjacencies and their “colors.” Indeed,

$$\begin{aligned} \Psi(u_1, \dots, u_i, \dots, u_n) &= ((-1)^{\sum_{j \neq 1} u_j} u_1, \dots, (-1)^{\sum_{j \neq i} u_j} u_i, \dots, (-1)^{\sum_{j \neq n} u_j} u_n) \\ &\rightarrow (-(-1)^{\sum_{j \neq 1} u_j} u_1, \dots, (-1)^{\sum_{j \neq i} u_j} u_i + 1, \dots, -(-1)^{\sum_{j \neq n} u_j} u_n) \\ &= ((-1)^{\sum_{j \neq 1} u_j} + (-1)^{\sum_{j \neq i} u_j} u_1, \dots, (-1)^{\sum_{j \neq i} u_j} (u_i + (-1)^{\sum_{j \neq i} u_j}), \dots, \\ &(-1)^{\sum_{j \neq n} u_j} + (-1)^{\sum_{j \neq i} u_j} u_n) = \Psi(u_1, \dots, u_i + (-1)^{\sum_{j \neq i} u_j}, \dots, u_n). \quad \square \end{aligned}$$

As an example, Figure 3.2 shows both the standard definition and the new representation of $M(6,4)$. (The torus surface is drawn as usual, where the directed dashed lines represent the identification of parallel sides of the rectangle.)

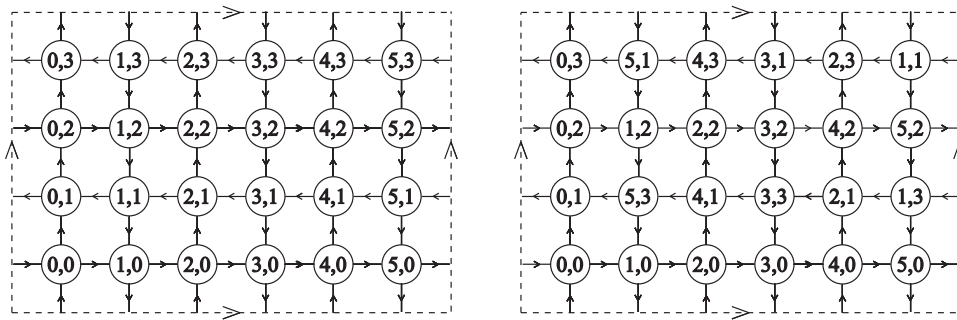


FIG. 3.2. The vertices with their standard labels and with labels induced by applying Ψ to the Manhattan street network $M(6,4)$.

As suggested by this example, it can be readily checked that Ψ is involutive, and hence the mapping from the alternative definition to the standard one is simply $\Psi^{-1} = \Psi$.

4. The diameter of Manhattan street networks. In the 2-dimensional case, the diameter of the Manhattan street network M_2 was first explicitly given by Chung and Agrawal [7] by using spanning trees, although it follows easily from the results of Varvarigos [21], where the mean distance was computed. The same result was proved by Comellas and Dalfó [9] from the comparison of the distance distribution in M_2 and the corresponding undirected toroidal mesh. In [21, 9], the distribution of vertices at each distance was also given, which allows closed formulas for the mean distance. In particular, for large values of N_1 and N_2 , the number of vertices at distance $k \geq 4$ from a given vertex, say $\mathbf{0}$, of $M_2(N_1, N_2)$, is $4(k-1)$ (see Figure 4.1 for the cases $k = 7, 8$). This was also noted by Morillo, Fiol, and Fàbrega [18] for the (not necessarily toroidal) 2-dimensional Manhattan street network with vertex set as in (2.3). Besides, in the same paper it was shown that, considering the digraph as bipartite, if it has diameter $D(> 4)$, then its order has the following Moore-like upper bound (see Miller and Širáň [17]):

$$N(2, D) = \begin{cases} 2(D-1)^2, & D \text{ odd,} \\ 2[(D-1)^2 + 1], & D \text{ even.} \end{cases}$$

In fact, if we do not impose the “toroidal closure” of the network, the above values can be attained since the corresponding tiles (that is, the sets of unit squares associated to the vertices which are at distance at most D from $\mathbf{0}$) tessellate the plane periodically (see Figure 4.1). More precisely, when D is odd, the “steps” (see Figure 2.2) for attaining the maximum order are, for instance,

$$\begin{aligned} a_0 &= 3, & a_1 &= 2D-3, & a_2 &= -2D+1, & a_3 &= -1; \\ b_0 &= 1, & b_1 &= -3, & b_2 &= -2D+3, & b_3 &= 2D-1; \end{aligned}$$

whereas, for D even, we have

$$\begin{aligned} a_0 &= -3, & a_1 &= 2D+1, & a_2 &= -2D+1, & a_3 &= 1; \\ b_0 &= -1, & b_1 &= 3, & b_2 &= -2D-1, & b_3 &= 2D-1. \end{aligned}$$

4.1. Dimension 2. As the reader will have already noted, in the toroidal case studied here, and for a given number of vertices $N = N_1 N_2$, the diameter is much greater than the expression obtained above. In this case we have the following result, which we use afterwards to study the n -dimensional case. As mentioned above, to the best of our knowledge this result was first proved by Chung and Agrawal [7] by using spanning trees. Here we give a constructive proof by explicitly deriving a shortest path from any vertex to vertex $(0, 0)$. Thus our proof leads to a local (shortest-path) routing scheme.

THEOREM 4.1. *The diameter of the Manhattan street network $M(N_1, N_2)$ is*

- (a) $D = \frac{N_1}{2} + \frac{N_2}{2} + 1$ if $N_1 \equiv N_2 \equiv 0 \pmod{4}$,
- (b) $D = \frac{N_1}{2} + \frac{N_2}{2}$ otherwise.

Proof. The proof is based on our Definition 3.4. Let α, β be some even and odd integers, respectively, in $[0, N_i/2 - 1]$, $i = 1, 2$. Similarly, let γ be any integer in the interval $[0, N_i/2 - 1]$. Because of the symmetry of the digraph, it suffices to consider a path from a generic vertex $\mathbf{u} = (u_1, u_2)$ to vertex $\mathbf{0} = (0, 0)$, whose length never exceeds the above values for D . With this aim, we first consider the cases where

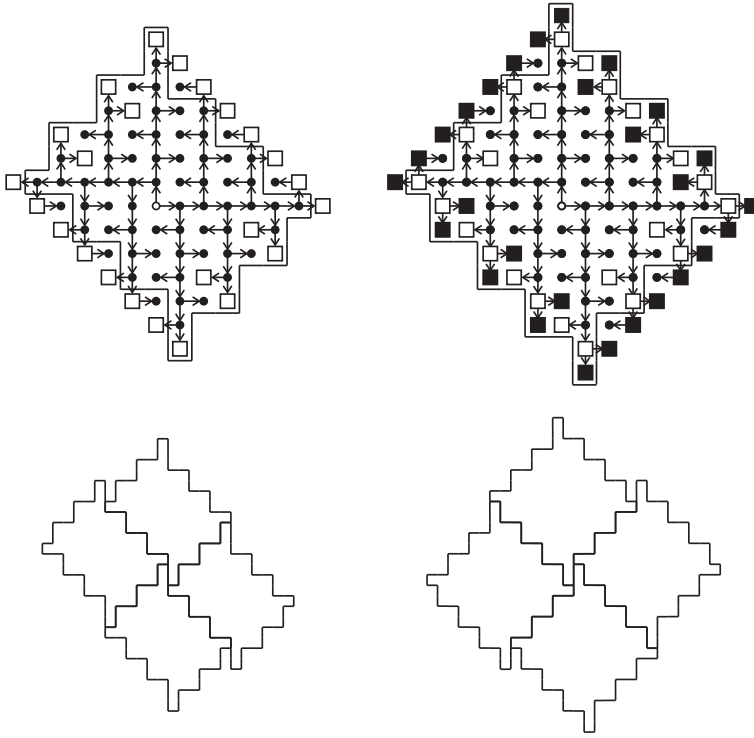


FIG. 4.1. The vertices at distance 7 (“□”) and at distance 8 (“■”) from vertex $(0,0)$ (“○”), and the corresponding tessellations.

some u_i equals $\pm N_i/2$ (the sign is irrelevant since we are in $\mathbb{Z}_{N_i}$ with N_i even). Here the arrows and the numbers above them represent the paths and their lengths, respectively,

(i) $u_1 = N_1/2, u_2 = N_2/2$:

$$(N_1/2, N_2/2) \xrightarrow{+N_1/2} (0, N_2/2) \xrightarrow{+N_2/2} (0, 0).$$

(ii) $u_1 = N_1/2$ (even), $u_2 = -\gamma$:

$$(N_1/2, -\gamma) \xrightarrow{+N_1/2} (0, -\gamma) \xrightarrow{+\gamma} (0, 0).$$

(iii) $u_1 = N_1/2$ (even), $u_2 = \alpha$:

$$(N_1/2, \alpha) \xrightarrow{+1} (N_1/2 + 1, -\alpha) \xrightarrow{+\alpha} (N_1/2 + 1, 0) \xrightarrow{+N_1/2-1} (0, 0).$$

(iv) $u_1 = N_1/2$ (even), $u_2 = \beta$:

$$(N_1/2, \beta) \xrightarrow{+N_1/2} (0, \beta) \xrightarrow{+1} (1, -\beta) \xrightarrow{+\beta} (-1, 0) \xrightarrow{+1} (0, 0).$$

(v) $u_1 = N_1/2$ (odd), $u_2 = \gamma$:

$$(N_1/2, \gamma) \xrightarrow{+N_1/2} (0, -\gamma) \xrightarrow{+\gamma} (0, 0).$$

(vi) $u_1 = N_1/2$ (odd), $u_2 = -\gamma$:

$$(N_1/2, -\gamma) \xrightarrow{+\gamma} (N_1/2, 0) \xrightarrow{+N_1/2} (0, 0).$$

Note that all of the above paths have length $\text{dist}(\mathbf{u}, \mathbf{0}) \leq N_1/2 + N_2/2$, except in case (iv) where we have

$$(4.1) \quad \text{dist}(\mathbf{u}, \mathbf{0}) = \frac{N_1}{2} + \beta + 2 \leq \frac{N_1}{2} + \frac{N_2}{2} + 1,$$

and equality is attained when $\beta = N_2/2 - 1$ or, in the symmetric case, $u_1 = \beta = N_1/2 - 1$ and $u_2 = N_2/2$. Note that, in both cases, $N_1/2$ and $N_2/2$ must be even, so that we are in case (a) of the theorem. Besides, when $\beta = N_2/2 - 2$, (4.1) gives $\text{dist}(\mathbf{u}, \mathbf{0}) = N_1/2 + N_2/2$ (for $N_1/2$ even and $N_2/2$ odd); and the same holds in the symmetric case $u_1 = \beta = N_1/2 - 2$, $u_2 = N_2/2$ (for $N_1/2$ odd and $N_2/2$ even). Such cases correspond to case (b) in Theorem 4.1.

Moreover, when neither of the entries u_i equals $N_i/2$, we need to consider the following cases (as before, α_i and β_i represent even and odd integers, respectively, in the interval $[0, N_i - 1]$, $i = 1, 2$):

$$(1) \quad u_1 = -\alpha_1, u_2 = -\alpha_2:$$

$$(-\alpha_1, -\alpha_2) \xrightarrow{+\alpha_1} (0, -\alpha_2) \xrightarrow{+\alpha_2} (0, 0).$$

$$(2) \quad u_1 = -\alpha, u_2 = -\beta:$$

$$(-\alpha, -\beta) \xrightarrow{+\alpha} (0, -\beta) \xrightarrow{+\beta} (0, 0).$$

$$(3) \quad u_1 = -\beta_1, u_2 = -\beta_2:$$

$$(-\beta_1, -\beta_2) \xrightarrow{+\beta_1} (0, \beta_2) \xrightarrow{+1} (1, -\beta_2) \xrightarrow{+\beta_2} (-1, 0) \xrightarrow{+1} (0, 0).$$

$$(4) \quad u_1 = -\alpha, u_2 = \beta:$$

$$(-\alpha, \beta) \xrightarrow{+\alpha} (0, \beta) \xrightarrow{+1} (1, -\beta) \xrightarrow{+\beta} (-1, 0) \xrightarrow{+1} (0, 0).$$

$$(5) \quad u_1 = -\beta, u_2 = \alpha:$$

$$(-\beta, \alpha) \xrightarrow{+\beta} (0, -\alpha) \xrightarrow{+\alpha} (0, 0).$$

$$(6) \quad u_1 = -\alpha_1, u_2 = \alpha_2:$$

$$(-\alpha_1, \alpha_2) \xrightarrow{+1} (-\alpha_1 + 1, -\alpha_2) \xrightarrow{+\alpha_2} (-\alpha_1 + 1, 0) \xrightarrow{+\alpha_1 - 1} (0, 0).$$

$$(7) \quad u_1 = -\beta_1, u_2 = \beta_2:$$

$$(-\beta_1, \beta_2) \xrightarrow{+\beta_1} (0, -\beta_2) \xrightarrow{+\beta_2} (0, 0).$$

$$(8a) \quad u_1 = \alpha_1 < N_1/2 - 1, u_2 = \alpha_2:$$

$$\begin{aligned} &(\alpha_1, \alpha_2) \xrightarrow{+1} (\alpha_1 + 1, -\alpha_2) \xrightarrow{+\alpha_2} (\alpha_1 + 1, 0) \\ &\xrightarrow{+1} (-\alpha_1 - 1, 1) \xrightarrow{+\alpha_1 + 1} (0, -1) \xrightarrow{+1} (0, 0). \end{aligned}$$

$$(8b) \quad u_1 = \alpha_1 = N_1/2 - 1, u_2 = \alpha_2:$$

$$(\alpha_1, \alpha_2) \xrightarrow{+1} (\alpha_1 + 1, -\alpha_2) = (N_1/2, -\alpha_2) \xrightarrow{+\alpha_2} (N_1/2, 0) \xrightarrow{+N_1/2} (0, 0).$$

$$(9) \quad u_1 = \alpha, u_2 = \beta:$$

$$(\alpha, \beta) \xrightarrow{+1} (\alpha + 1, -\beta) \xrightarrow{+\beta} (-\alpha - 1, 0) \xrightarrow{+\alpha+1} (0, 0).$$

$$(10) \quad u_1 = \beta_1, u_2 = \beta_2:$$

$$(\beta_1, \beta_2) \xrightarrow{+1} (\beta_1 + 1, -\beta_2) \xrightarrow{+\beta_2} (-\beta_1 - 1, 0) \xrightarrow{+\beta_1+1} (0, 0).$$

Observe again that all of the above paths have length $\text{dist}(\mathbf{u}, \mathbf{0}) < N_1/2 + N_2/2$, except in the cases (4), (8b), (9) where we have

$$\text{dist}(\mathbf{u}, \mathbf{0}) = \frac{N_1}{2} + \frac{N_2}{2},$$

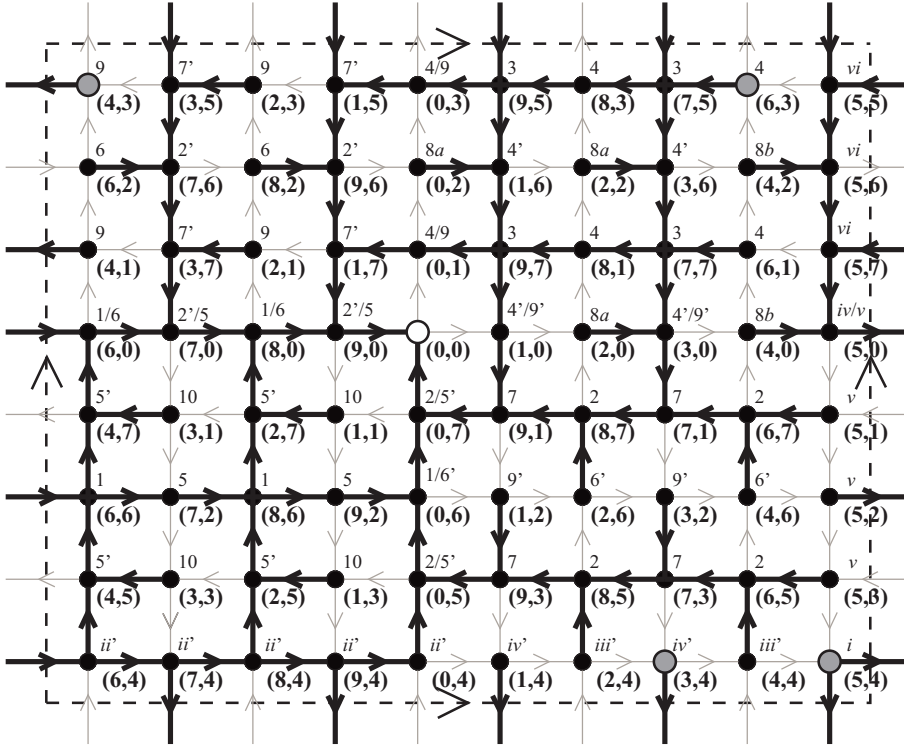
for different parities of $N_1/2$ and $N_2/2$. For instance, in case (8b) the maximum is attained when $\alpha_2 = N_1/2 - 1$; that is, for a vertex of the form $(N_1/2 - 1, N_2/2 - 1)$, where $N_1 \equiv N_2 \equiv 2 \pmod{4}$. (This and all of the other vertices at maximum distance are listed below.) This completes the proof. $\square$

Notice that, from the above proof, the vertices at maximum distance from $\mathbf{0}$ are, depending on the case (recall that we are using the alternative definition):

- (a) $N_1 \equiv N_2 \equiv 0 \pmod{4}$:
 - (iv): $(N_1/2, N_2/2 - 1), (N_1/2 - 1, N_2/2)$;
- (b1) $N_1 \equiv 0, N_2 \equiv 2 \pmod{4}$:
 - (i): $(N_1/2, N_2/2)$,
 - (iv): $(N_1/2, N_2/2 - 2)$,
 - (4): $(N_1/2 - 1, N_2/2 + 1)$,
 - (9): $(N_1/2 - 1, N_2/2 - 1)$;
- (b2) $N_1 \equiv 2, N_2 \equiv 0 \pmod{4}$:
 - (i): $(N_1/2, N_2/2)$,
 - (iv): $(N_1/2 - 2, N_2/2)$,
 - (4): $(N_1/2 + 1, N_2/2 - 1)$,
 - (9): $(N_1/2 - 1, N_2/2 - 1)$;
- (b3) $N_1 \equiv N_2 \equiv 2 \pmod{4}$:
 - (i): $(N_1/2, N_2/2)$,
 - (8b): $(N_1/2 - 1, N_2/2 - 1)$.

4.2. A routing scheme. As we have already mentioned, the interest of the above proof is that it leads to a shortest-path routing scheme. Indeed, by the symmetry of the network, it suffices to exhibit a path from each vertex to the vertex $(0, 0)$. More precisely, for every node $\mathbf{u} = (u_1, u_2)$ and from the cases (i)–(vi) and (1)–(10) considered in the proof, we can determine which is the shortest path from $\mathbf{u}$ to $(0, 0)$. This is illustrated in Figure 4.2 for the Manhattan street network $M(10, 8)$ where the small number above each vertex indicates the corresponding case. Note that sometimes, when one of the entries u_i is zero, there are two equivalent cases. A label with an apostrophe refers to the corresponding symmetric case obtained by interchanging the roles of u_1 and u_2 . The whole path from $\mathbf{u}$ to $(0, 0)$ is obtained from the case where $\mathbf{u}$ belongs to. For instance, vertex $(6, 1)$ belongs to case (4), which gives the following path (with thick lines in the figure):

$$(6, 1) \rightarrow (7, 7) \rightarrow (8, 1) \rightarrow (9, 7) \rightarrow (0, 1) \rightarrow (1, 7) \rightarrow (9, 0) \rightarrow (0, 0).$$

FIG. 4.2. The shortest paths to $(0,0)$ in $M(10,8)$, according to the proof of Theorem 4.1.

If in some step there is more than one option, then we always take the horizontal link. For instance, in the previous example, as node $(6,1)$ belongs to case (4), which gives the previous path, the next step must be horizontal. This leads us to node $(7,7)$, which belongs to case (3) yielding the subpath $(7,7) \rightarrow (8,1) \rightarrow \dots \rightarrow (0,0)$. Thus, the next step must be again horizontal to reach vertex $(8,1)$ belonging to case (4), and so on. We proceed in this way until we reach $(0,0)$.

Just for illustration, the four vertices $(5,4)$, $(3,4)$, $(6,3)$, and $(4,3)$ at maximum distance 9 to $(0,0)$ —case (b2) in the proof—are indicated by grey in Figure 4.2.

4.3. Dimension n . To derive the diameter of the n -dimensional Manhattan street network, we introduce the following notation: Given N (even) and $0 \leq u \leq N$, let $\|u\|_N$ be the distance between 0 and u in the undirected cycle C_N ; that is, $\|u\|_N = \min\{u \bmod N, -u \bmod N\}$ (so that $0 \leq \|u\|_N \leq N/2$).

LEMMA 4.2. Consider the vertices $\mathbf{0} = (0, 0, \dots, 0)$, $\mathbf{u} = (u_1, \dots, u_{n-1}, u_n)$ in $M_n = (N_1, \dots, N_{n-1}, N_n)$, $n > 2$; and $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{u}' = (u_1, \dots, u_{n-1})$ in $M_{n-1} = (N_1, \dots, N_{n-1})$. Let $\alpha = (-1)^{u_n}$. Then,

- (a) $\text{dist}_{M_n}(\mathbf{u}, \mathbf{0}) \geq \sum_{i=1}^n \|u_i\|_{N_i}$;
- (b) $\text{dist}_{M_n}(\mathbf{u}, \mathbf{0}) \leq \text{dist}_{M_{n-1}}(\alpha \mathbf{u}', \mathbf{0}) + \|u_n\|_{N_n}$, if $u_n \in [N_n/2, N_n]$;
- (c) $\text{dist}_{M_n}(\mathbf{u}, \mathbf{0}) \leq \text{dist}_{M_{n-1}}(\alpha(\mathbf{u}' + \mathbf{e}_1), \mathbf{0}) + \|u_n\|_{N_n} + 1$, if $u_n \in [0, N_n/2 - 1]$.

Proof. The inequality in (a) is a direct consequence of the fact that the underlying undirected graph of M_n is just the direct product of the cycles C_{N_i} , $1 \leq i \leq n_i$.

If $N_n/2 \leq u_n \leq N_n$, then there is a path of length $N_n - u_n = \|u_n\|_{N_n}$ from $\mathbf{u}$ to $(\alpha u_1, \dots, \alpha u_{n-1}, 0)$ in M_n . From this vertex, we will need at most $\text{dist}_{M_{n-1}}(\alpha \mathbf{u}', \mathbf{0})$ steps to reach $\mathbf{0}$. This proves (b).

Otherwise, when $0 \leq u_n \leq N_n/2 - 1$, we go first, in one step, from $\mathbf{u}$ to $(u_1 + 1, -u_2, \dots, -u_{n-1}, -u_n)$. Then, from this vertex we have a path of length $u_n = \|u_n\|_{N_n}$ to $(\alpha(u_1 + 1), -\alpha u_2, \dots, -\alpha u_{n-1}, 0)$, and reasoning as before we get (c). $\square$

As a consequence, from the vertex symmetry of the involved digraphs, we have

$$\sum_{i=1}^n \frac{N_i}{2} \leq \text{ecc}_{M_n}(\mathbf{0}) \leq \text{ecc}_{M_{n-1}}(\mathbf{0}) + \frac{N_n}{2}$$

since, in case (c), $\|u_n\|_{N_n} \leq N_n/2 - 1$, and the same formula applies for the respective diameters:

$$(4.2) \quad \sum_{i=1}^n \frac{N_i}{2} \leq D(M_n) \leq D(M_{n-1}) + \frac{N_n}{2}.$$

THEOREM 4.3. *The diameter of the n -dimensional Manhattan street network $M_n = M(N_1, \dots, N_n)$ is*

$$(a) \quad D(M_n) = \sum_{i=1}^n \frac{N_i}{2} + 1 \quad \text{if } N_i \equiv 0 \pmod{4} \text{ for every } 1 \leq i \leq n;$$

$$(b) \quad D(M_n) = \sum_{i=1}^n \frac{N_i}{2} \quad \text{otherwise.}$$

Proof. Since Lemma 4.2 clearly applies to any two components of the vertex $(u_1, u_2, \dots, u_n)$, we can apply recursively the above results to get

$$(4.3) \quad D(M_n) \leq D(M(N_j, N_k)) + \frac{1}{2} \sum_{i \neq j, k} N_i \quad (1 \leq j < k \leq n).$$

Then, under the hypothesis in (b), there is some $1 \leq j \leq n$ such that $N_j \not\equiv 0 \pmod{4}$ and $D(M(N_j, N_k)) = N_j/2 + N_k/2$ by Theorem 4.1. This, together with the lower bound in (4.2), proves the equality in (b).

Otherwise, if $N_i \equiv 0 \pmod{4}$ for any $1 \leq i \leq n$, then Theorem 4.1 and (4.3) yield

$$(4.4) \quad D(M_n) \leq \frac{1}{2} \sum_{i=1}^n N_i + 1.$$

Thus, to prove (a) we need only to show that equality is attained for some vertex. In fact, the vertices at maximum distance to $\mathbf{0}$ are, in this case,

$$\mathbf{u}_i = (N_1/2, N_2/2, \dots, N_i/2 - 1, \dots, N_n/2) \quad (1 \leq i \leq n)$$

(compare with the 2-dimensional case (a), after the proof of Theorem 4.1). For instance, let us check that $\text{dist}(\mathbf{u}_1, \mathbf{0})$ attains the upper bound in (4.4). With this aim, note that the first entry of $\mathbf{u}_1$, $N_1/2 - 1$, is odd, whereas the others, $N_i/2$, $2 \leq i \leq n$, are even. Moreover, for all $1 \leq i \leq n$, we need at least $N_i/2$ steps to bring the i th entry to 0, which gives $\text{dist}(\mathbf{u}_1, \mathbf{0}) \geq \sum_{i=1}^n N_i/2$. In particular, this requires changing, in some step, the first component from $N_1/2 - 1$ to $-N_1/2 + 1 = N_1/2 + 1$, which is accomplished when the number, say r , of previous steps going through j -arcs,

$$(u_1, \dots, u_j, \dots, u_n) \rightarrow (-u_1, \dots, u_j + 1, \dots, -u_n) \quad (2 \leq j \leq n),$$

is odd. Then, although we have taken r steps in the “right direction” to $\mathbf{0}$, there is some j such that the first and j th entries of the reached vertex $\mathbf{u}'$, $u'_1 = N_1/2 + 1$, and $u'_j \in [N_j/2 + 1, N_j - 1]$, are odd. But now, as in case (3) in the proof of Theorem 4.1,

it is impossible to bring these components to 0 without spending at least $\|u'_1\| + \|u'_j\| + 2$ steps (no matter what the other entries are). Hence, we have $\text{dist}(\mathbf{u}_1, \mathbf{0}) = \sum_{i=1}^n N_i/2 + 1$, which completes the proof. $\square$

From the previous results (Theorems 4.1 and 4.3, and Lemma 4.2) and their proofs, it is now apparent how a message can be routed along a path whose length is bounded by the diameter of M_n . Indeed, to go from vertex $(u_1, u_2, \dots, u_n)$ to $\mathbf{0}$ we first go to vertex $(0, \dots, u_i, \dots, u_j, \dots, 0)$ for some appropriately chosen i, j (proof of Lemma 4.2) and, to deal with this two remaining components, we follow the routing induced by the proof of Theorem 4.1—as we were in $M(N_i, N_j)$ —to get to the destination vertex $\mathbf{0}$.

5. Hamiltonian cycles. In this section we first show that Manhattan street networks are Hamiltonian digraphs. Moreover, in the 2-dimensional case, some sufficient conditions for M_2 to be decomposable into two (arc-disjoint) Hamiltonian cycles are derived.

THEOREM 5.1. *The Manhattan street network M_n is Hamiltonian.*

Proof. We proceed by induction over n . For $n = 1$, M_1 can be seen as a directed cycle and hence is trivially Hamiltonian. (We can also start from $N = 2$ since we already know that M_2 is Hamiltonian by Lemma 2.3.) Now suppose that there exists a Hamiltonian cycle for M_{n-1} . Then, we build a Hamiltonian cycle for M_n by appropriately joining N_n Hamiltonian cycles (without some arcs) of its N_n subdigraphs isomorphic to M_{n-1} (see Lemma 2.2(c)). More precisely, the cycle begins, say, in $(u_1, u_2, \dots, u_{n-2}, 0, 0)$ and goes from this vertex to the vertices $(u_1, u_2, \dots, u_{n-2}, 0, 1), \dots, (u_1, u_2, \dots, u_{n-2}, 0, N_n - 1)$. From this last vertex, we follow a clockwise Hamiltonian cycle in a subdigraph M_{n-1} (without the last step) until vertex $(u_1, u_2, \dots, u_{n-2}, N_{n-1} - 1, N_n - 1)$. Then, we go to vertex $(u_1, u_2, \dots, u_{n-2}, N_{n-1} - 1, N_n - 2)$ and, from it, we follow in a counterclockwise cycle until $(u_1, u_2, \dots, u_{n-2}, 1, N_n - 2)$. From here, we go to $(u_1, u_2, \dots, u_{n-2}, 1, N_n - 3)$. Now we repeat this pattern several times until we reach the counterclockwise Hamiltonian cycle of a M_{n-1} (without the last step), which finishes in $(u_1, u_2, \dots, u_{n-2}, 0, 0)$ and thus closes the Hamiltonian cycle of M_n (see Figure 5.1). $\square$

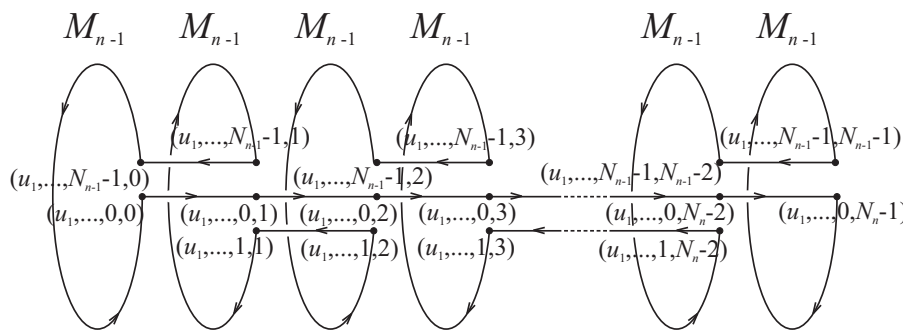
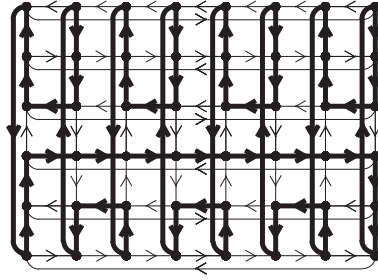


FIG. 5.1. A Hamiltonian cycle in M_n .

As an example of a Hamiltonian cycle in $M(N_1, N_2)$, see Figure 5.2.

Besides the Hamiltonian property, in some applications it is interesting to have a decomposition of the network into (two or more) edge-disjoint Hamiltonian cycles. In this context, Varvarigos [21] showed the presence of two edge-disjoint Hamiltonian cycles in the 2-dimensional Manhattan street network $M(N, N)$. Generalizing


 FIG. 5.2. A Hamiltonian cycle in $M(8,6)$.

his result, we give here some sufficient conditions for $M(N_1, N_2)$ to be Hamiltonian decomposable in such a way.

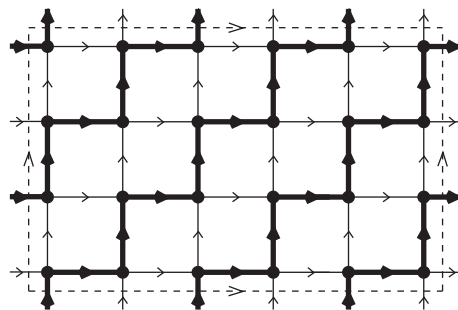
PROPOSITION 5.2. *Let $M_2 = M(N_1, N_2)$ be a 2-dimensional Manhattan street network. If the two following conditions hold:*

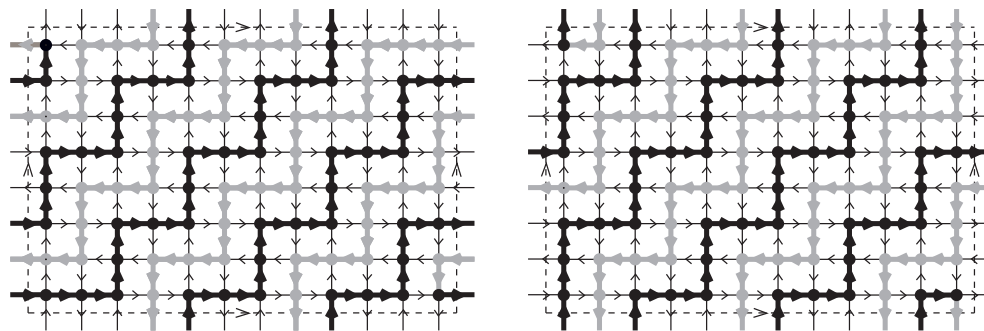
- (a) $\gcd\{N_1/2, N_2/2\} = d \geq 2$;
- (b) *there exist $d_1, d_2 > 0$, $d_1 + d_2 = d$, such that*

$$\gcd\{d_1, N_1/2\} = \gcd\{d_2, N_2/2\} = 1;$$

then M_2 contains two edge-disjoint Hamiltonian cycles.

Proof. The proof is based on a result of Trotter and Erdős [20], who showed that (a) and (b) are sufficient conditions for the Cartesian product of the directed cycles $C_{N_1/2} \times C_{N_2/2}$ to be Hamiltonian. For instance, such conditions are fulfilled with $N_1/2 = 6$ and $N_2/2 = 4$, since $d = 2$ and we can simply take $d_1 = d_2 = 1$, yielding the Hamiltonian cycle shown in Figure 5.3. Note that to obtain such a cycle we simply repeat periodically a subpath of d_1 horizontal steps and d_2 vertical steps. Then, Figure 5.4 illustrates how this Hamiltonian cycle (with arcs going east and north and constructed by taking into account only the even vertices, that is with both entries u_1, u_2 even) induces a Hamiltonian cycle in M_2 . The key idea is that, before closing, this east-north quasi-cycle—in black in the figure—changes its last step from vertical to horizontal and turns into the symmetric west-south quasi-cycle—in grey (now by considering only the odd vertices). Then, at the end, it changes again (last step from vertical to horizontal) to connect with the first quasi-cycle and completing the Hamiltonian cycle in M_2 . With respect to the other (edge-disjoint) cycle of $M(N_1, N_2)$, it is simply the complement of the first one. Alternatively, it can be seen as a Hamiltonian cycle constructed in the same way as the first one, but on the


 FIG. 5.3. A Hamiltonian cycle in $C_6 \times C_4$.

FIG. 5.4. A decomposition of $M(12, 8)$ into two edge-disjoint Hamiltonian cycles.

converse digraph of the Manhattan street network $M(N_2, N_1)$; see again Figure 5.4 for the details. $\square$

Acknowledgment. The authors sincerely acknowledge the anonymous referee for the valuable comments and suggestions that led to a significant improvement of this paper.

REFERENCES

- [1] C. BALBUENA, D. FERRERO, X. MARCOTE, AND I. PELAYO, *Algebraic properties of a digraph and its line digraph*, J. Interconnection Networks, 4 (2003), pp. 377–393.
- [2] S. BANERJEE, V. JAIN, AND S. SHAH, *Regular multihop logical topologies for lightwave networks*, IEEE Comm. Surveys and Tutorials, First Quarter, 2 (1999) (electronic).
- [3] D. BANERJEE, B. MUKHERJEE, AND S. RAMAMURTHY, *The multidimensional torus: Analysis of average hop distance and application as a multihop lightwave network*, in Proceedings of the IEEE International Conference on Communications, Vol. 3, New Orleans, LA, 1994, pp. 1675–1680.
- [4] J. BANG-JENSEN AND G. GUTIN, *Digraphs: Theory, Algorithms and Applications*, Springer Monographs in Mathematics, Springer, London, 2000.
- [5] M. BENKERT, A. WOLFF, F. WIDMANN, AND T. SHIRABE, *The minimum Manhattan network problem: Approximations and exact solutions*, Comput. Geom., 35 (2006), pp. 188–208.
- [6] G. CHARTRAND AND L. LESNIAK, *Graphs & Digraphs*, 3rd ed., Chapman and Hall, London, 1996.
- [7] T. Y. CHUNG AND D. P. AGRAWAL, *On network characterization of and optimal broadcasting in the Manhattan Street Network*, in Proceedings of the IEEE INFOCOM'90, Vol. 2, San Francisco, CA, 1990, pp. 465–472.
- [8] T. Y. CHUNG AND D. P. AGRAWAL, *Design and analysis of multidimensional Manhattan Street Networks*, IEEE Trans. Commun., 41 (1993), pp. 295–298.
- [9] F. COMELLAS AND C. DALFÓ, *Optimal broadcasting in 2-dimensional Manhattan street networks*, Parallel Distrib. Comput. Networks, 246 (2005), pp. 135–140.
- [10] F. COMELLAS, M. A. FIOLO, J. GIMBERT, AND M. MITJANA, *Weakly distance-regular digraphs*, J. Combin. Theory Ser. B, 90 (2004), pp. 233–255.
- [11] H. S. M. COXETER AND W. O. J. MOSER, *Generators and Relations for Discrete Groups*, 4th ed., Springer-Verlag, Berlin, 1980.
- [12] M. A. FIOLO AND M. MITJANA, *The spectra of some families of digraphs*, Linear Algebra Appl., 423 (2007), pp. 109–118.
- [13] M. A. FIOLO, J. L. A. YEBRA, AND I. ALEGRE, *Line digraph iterations and the (d, k) digraph problem*, IEEE Trans. Comput., C-33 (1984), pp. 400–403.
- [14] C. HEUCHENNE, *Sur une certaine correspondance entre graphes*, Bull. Soc. Roy. Sci. Liège, 33 (1964) pp. 743–753.
- [15] B. KHASNABISH, *Topological properties of Manhattan street networks*, Electronics Lett., 25 (1989), pp. 1388–1389.
- [16] N. F. MAXEMCHUK, *Routing in the Manhattan street network*, IEEE Trans. Commun., 35 (1987), pp. 503–512.

- [17] M. MILLER AND J. ŠIRÁŇ, *Moore graphs and beyond: A survey of the degree/diameter problem*, Electronic J. Combin., DS14 (2005); available online at www.combinatorics.org/Surveys/ds14.pdf.
- [18] P. MORILLO, M. A. FIOL, AND J. FÀBREGA, *The diameter of directed graphs associated to plane tessellations*, Ars Comb., 20 (1985), pp. 17–27.
- [19] G. SABIDUSSI, *On a class of fixed-point-free graphs*, Proc. Amer. Math. Soc., 9 (1958), pp. 800–804.
- [20] W. T. TROTTER AND P. ERDŐS, *When the Cartesian product of directed cycles is Hamiltonian*, J. Graph Theory, 2 (1978), pp. 137–142.
- [21] E. A. VARVARIGOS, *Optimal communication algorithms for Manhattan street networks*, Discrete Appl. Math. 83 (1998), pp. 303–326.
- [22] J. L. VILLAR, *Contribución a la Teoría de Redes con Enlaces Bidireccionales Derivadas de los Digrafos Línea*, Ph.D. thesis, Universitat Politècnica de Catalunya, Catalonia, Spain, 1994.

PARTITIONS OF FAULTY HYPERCUBES INTO PATHS WITH PRESCRIBED ENDVERTICES*

TOMÁŠ DVOŘÁK† AND PETR GREGOR†

Abstract. Given a set $\mathcal{P} = \{a_i, b_i\}_{i=1}^m$ of pairs of vertices in a graph G , is there a collection of paths $\{P_i\}_{i=1}^m$ such that P_i connects a_i with b_i and $\{V(P_i)\}_{i=1}^m$ partitions $V(G)$? We study this problem for the graph $Q_n - \mathcal{F}$ obtained from the n -dimensional hypercube Q_n by removing a set $\mathcal{F}$ of faulty vertices. We show that an obvious necessary condition for the existence of such a partition is also sufficient provided $2|\mathcal{P}| + 3|\mathcal{F}| \leq n - 3$. As a corollary, we obtain a similar characterization for the existence of a hamiltonian cycle and a hamiltonian path of $Q_n - \mathcal{F}$ provided $|\mathcal{F}| \leq (n - 5)/3$. On the other hand, if the size of $\mathcal{F}$ is not limited, the problems are NP-complete.

Key words. path partition, hamiltonian cycle, hamiltonian path, hypercube, faulty vertex, fault tolerance, NP-complete

AMS subject classifications. 05C70, 05C45, 68Q17, 68M07, 68M15, 68R10

DOI. 10.1137/060678476

1. Introduction. The n -dimensional hypercube Q_n is a graph whose structural properties have been a subject of interest to graph theorists [12], but it has been also studied by computer scientists as a prospective architecture for interconnection networks [14]. This application also inspired questions related to the reliability of hypercubic networks: if some vertices or edges, corresponding to busy or faulty processors or links, are removed from Q_n , under what conditions does the resulting subgraph preserve the functionality of the original network?

Hamiltonian cycles in Q_n exist for all $n \geq 2$ and their properties and applications have been thoroughly explored [19]. The problem of existence of a hamiltonian cycle of Q_n avoiding an arbitrary number of faulty edges is known to be NP-complete, while in case that this number does not exceed $2n - 5$, a hamiltonian cycle exists provided the natural condition that every vertex is incident with at least two nonfaulty edges is satisfied [4]. The upper bound is optimal in the sense that for $n \geq 3$ there exists Q_n with $2n - 4$ faulty edges satisfying the natural condition but without a hamiltonian cycle avoiding the faulty edges. There is a similar result for hamiltonian paths with given endvertices [20] as well as a generalization to k -ary hypercubes Q_n^k for $k \geq 3$ [1].

The status of the problem for faulty vertices is, however, less satisfactory. It is known that for any distinct vertices u, v, w of Q_n there is a hamiltonian path between u and v in $Q_n - \{w\}$ iff u, v belong to the same partite set, which is different from that containing w [15]. Moreover, this path can avoid a given set of at most $n - 3$ faulty edges [21]. The other results say that Q_n with f faulty vertices contains a cycle of length at least $2^n - 2f$ if $f \leq 2n - 4$ [9] and a path of length at least $2^n - 2f - 1$ ($2^n - 2f - 2$) between two arbitrary vertices at odd (even) distance from each other if $f \leq n - 2$ [10].

The motivation of this article is to obtain results on hamiltonicity of hypercubes with faulty vertices forming a natural counterpart of the results for faulty edges.

*Received by the editors December 22, 2006; accepted for publication (in revised form) June 4, 2008; published electronically September 25, 2008. The second author is partially supported by Czech Science Foundation grant 201/08/P298.

<http://www.siam.org/journals/sidma/22-4/67847.html>

†Faculty of Mathematics and Physics, Charles University, Malostranské nám. 25, 118 00 Prague, Czech Republic (tomas.dvorak@mff.cuni.cz, petr.gregor@mff.cuni.cz).

In order to accomplish this task, we study a more general problem: Given a set $\mathcal{P} = \{a_i, b_i\}_{i=1}^m$ of pairs of vertices in a graph G , is there a collection of paths $\{P_i\}_{i=1}^m$ such that P_i connects a_i with b_i and $\{V(P_i)\}_{i=1}^m$ partitions $V(G)$? This question was investigated previously for general graphs [8], hypercubes [3, 7], and hypercube-like interconnection networks with faulty edges and vertices [16]. We study this problem for the graph $Q_n - \mathcal{F}$ obtained from the n -dimensional hypercube Q_n by removing a set $\mathcal{F}$ of faulty vertices. We show that an obvious necessary condition for the existence of such a partition is also sufficient provided $2|\mathcal{P}| + 3|\mathcal{F}| \leq n - 3$. As a corollary, we obtain a similar characterization for the existence of a hamiltonian cycle of $Q_n - \mathcal{F}$ provided $|\mathcal{F}| \leq (n - 5)/3$. On the other hand, if $|\mathcal{F}|$ is not limited, both problems are NP-complete.

2. Preliminaries. The graph-theoretic terminology and notation used in this paper but not defined explicitly can be found, e.g., in [2]. A graph G with vertices $V(G)$ and edges $E(G)$ is a *spanning subgraph* of a graph H if $V(G) = V(H)$ and $E(G) \subseteq E(H)$. For a set $V \subseteq V(G)$, $\langle V \rangle$ denotes the subgraph of G induced by V , while $G - V$ denotes the subgraph $\langle V(G) \setminus V \rangle$. The distance of vertices u and v in G is denoted by $d_G(u, v)$, the subscript being omitted if no ambiguity may arise.

A path with endvertices a and b is denoted by P_{ab} . In particular, P_{aa} denotes the path consisting of a single vertex a . If paths P_{ab} and P_{cd} are vertex disjoint and vertices b and c are adjacent, $P_{ab} + P_{cd}$ denotes the path with endvertices a and d , obtained as a concatenation of P_{ab} with P_{cd} . Path P_1 is called a *subpath* of a path P_2 if P_1 forms a subgraph of P_2 . A collection $\{P_i\}_{i=1}^m$ of paths in a graph G forms a *path partition* of G if $\{V(P_i)\}_{i=1}^m$ partitions $V(G)$.

In this text n always denotes a positive integer, while $[n]$ stands for the set $\{1, 2, \dots, n\}$. Since it is sometimes convenient to deal with subsets of $[n]$ instead of binary vectors of $\{0, 1\}^n$, define a mapping $\pi : \{0, 1\}^n \rightarrow [n]$ which assigns to each $v = (v_1, v_2, \dots, v_n)$ the set $\pi(v) = \{i \in [n] \mid v_i = 1\}$. For binary vectors $u, v \in \{0, 1\}^n$ let $u \oplus v$ denote coordinatewise addition modulo 2. For any $x \in [n]$ let e_x^n denote the binary vector $\pi^{-1}(\{x\})$. In order to simplify the notation, write a^x instead of $a \oplus e_x^n$.

The n -dimensional hypercube Q_n is the graph with vertex set $V(Q_n) = \{0, 1\}^n$ and edge set $E(Q_n) = \{\{u, v\} \mid |\pi(u \oplus v)| = 1\}$. We define the *balance* $\beta(V)$ of a set $V \subseteq V(Q_n)$ by $\beta(V) = \sum_{v \in V} \beta(v)$, where $\beta(v) = (-1)^{|\pi(v)|}$ for any $v \in V(Q_n)$. A set V with $\beta(V) = 0$ is called *balanced*.

A set $\mathcal{P}$ of pairs $\{a_i, b_i\}_{i=1}^m$ of vertices of Q_n (of $V \subseteq V(Q_n)$) is called a *pair set* in Q_n (in V) if $\{a_i, b_i\} \cap \{a_j, b_j\} = \emptyset$ for all distinct $i, j \in [m]$. The union $\bigcup_{i=1}^m \{a_i, b_i\}$ is denoted by $\bigcup \mathcal{P}$. We extend the definition of balance to pair sets by setting $\beta(\mathcal{P}) = \sum_{\{a,b\} \in \mathcal{P}} \beta(a, b)$, where $\beta(a, b) = (\beta(a) + \beta(b))/2$ for any $a, b \in V(Q_n)$. A pair $\{a, b\} \in \mathcal{P}$ is called a *singleton* (*nonsingleton*) if $a = b$ ($a \neq b$). The set of all singletons of $\mathcal{P}$ is denoted by $s(\mathcal{P})$.

We say that a pair set $\mathcal{P}$ is

regular if $\mathcal{P}$ contains at least one nonsingleton,

strongly regular if $\mathcal{P}$ is nonempty and contains only nonsingletons,

balanced if $\beta(\mathcal{P}) = 0$,

connectable in G for some subgraph G of Q_n if there exists a path partition $\{P_{ab}\}_{\{a,b\} \in \mathcal{P}}$ of G , called a *connector* of $\mathcal{P}$ in G . The reference to G may be omitted if $G = Q_n$ and no ambiguity may arise.

Employing the terms introduced above, the main problems studied in this paper can be formulated in the following way.

Hamiltonian cycle in hypercube with faulty vertices ($HCQ_n - V$).

INSTANCE: A subset S of vertices of Q_n for some n .

QUESTION: Is the induced subgraph $\langle S \rangle$ of Q_n hamiltonian?

Hamiltonian path in hypercube with faulty vertices ($\text{HP}Q_n - V$).

INSTANCE: A subset S of vertices of Q_n for some n .

QUESTION: Does the induced subgraph $\langle S \rangle$ of Q_n contain a hamiltonian path?

Hamiltonian path with prescribed endvertices in hypercube with faulty vertices ($\text{HPP}Q_n - V$).

INSTANCE: A subset S of vertices of Q_n for some n and two distinct vertices $a, b \in S$.

QUESTION: Does the induced subgraph $\langle S \rangle$ of Q_n contain a hamiltonian path between a and b ?

Partition of hypercube with faulty vertices into paths with prescribed endvertices ($\text{PP}Q_n - V$).

INSTANCE: A subset S of vertices of Q_n for some n and a strongly regular pair set $\mathcal{P}$ in S .

QUESTION: Is $\mathcal{P}$ connectable in $\langle S \rangle$?

Recall that as $S \subseteq V(Q_n)$ is a set of binary vectors of length n , the instance size of any of the first three problems is $\Theta(n|S|)$, while the instance size of $\text{PP}Q_n - V$ is $\Theta(n(|S| + |\mathcal{P}|))$. It should also be noted that if $\{a, a\}$ is a singleton, then the only path connecting a with a is the path P_{aa} . Singletons can thus play the role of faulty vertices, and therefore it suffices to deal only with the connectability of regular pair sets in Q_n .

The main strategy used in this paper to attack the above problems is based on transformations of pair sets. In order to accomplish this task, we need a few more definitions. Let $x \in [n]$ and $\mathcal{P}$ be a pair set in Q_n . A vertex $v \in V(Q_n)$ is called x -free with respect to $\mathcal{P}$ if $\{v, v^x\} \cap \bigcup \mathcal{P} = \emptyset$. The reference to $\mathcal{P}$ may be omitted if no ambiguity may arise. For any $i \in \{0, 1\}$ let $V_{x,i}$ denote the set $\{(v_1, v_2, \dots, v_n) \in V(Q_n) \mid v_x = i\}$, and let $\mathcal{P}_{x,i}$ denote the set $\{\{a, b\} \in \mathcal{P} \mid \{a, b\} \subseteq V_{x,i}\}$, while $\mathcal{P}_{x,2} = \mathcal{P} \setminus (\mathcal{P}_{x,0} \cup \mathcal{P}_{x,1})$. The pair set $\mathcal{P}$ is called x -separate if $\mathcal{P}_{x,2} = \emptyset$; that is, we can deal with $\mathcal{P}_{x,0}$ in $V_{x,0}$ and $\mathcal{P}_{x,1}$ in $V_{x,1}$ separately.

DEFINITION 2.1. Let $\mathcal{P}, \bar{\mathcal{P}}$ be pair sets in Q_n and $x \in [n]$. We say that $\bar{\mathcal{P}}$ is obtained from $\mathcal{P}$ by

- (i) x -separation if $\bar{\mathcal{P}} = (\mathcal{P} \setminus \{\{a, b\}\}) \cup \{\{a, c\}, \{c^x, b\}\}$ for some $\{a, b\} \in \mathcal{P}_{x,2}$ and an x -free vertex c such that $a, c \in V_{x,i}$ for some $i \in \{0, 1\}$,
- (ii) x -detour if $\bar{\mathcal{P}} = (\mathcal{P} \setminus \{\{a, b\}\}) \cup \{\{a, c\}, \{c^x, d^x\}, \{d, b\}\}$ for some nonsingleton $\{a, b\} \in \mathcal{P}_{x,i}$ and distinct x -free vertices $c, d \in V_{x,i}$ for some $i \in \{0, 1\}$.

$\bar{\mathcal{P}}$ is called an x -extension of $\mathcal{P}$ if there is a sequence $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_m$ of pair sets such that $\mathcal{P}_1 = \mathcal{P}$, $\mathcal{P}_m = \bar{\mathcal{P}}$, and $\mathcal{P}_{i+1}$ is obtained from $\mathcal{P}_i$ by an x -separation or x -detour for each $i \in [m-1]$.

To simplify the notation, throughout the rest of the paper we omit parentheses in $(A \setminus B) \cup C$ and write simply $A \setminus B \cup C$. We conclude this section with a list of properties preserved by x -extensions that shall be needed later.

PROPOSITION 2.2. Let $x \in [n]$ and $\bar{\mathcal{P}}$ be an x -extension of a pair set $\mathcal{P}$. Then

- (i) $s(\bar{\mathcal{P}}) = s(\mathcal{P})$,
- (ii) $\beta(\bar{\mathcal{P}}) = \beta(\mathcal{P})$,
- (iii) $v \in \bigcup \bar{\mathcal{P}} \setminus \bigcup \mathcal{P}$ iff $v^x \in \bigcup \bar{\mathcal{P}} \setminus \bigcup \mathcal{P}$,
- (iv) if $\bar{\mathcal{P}}_{x,i} = s(\bar{\mathcal{P}}_{x,i})$ for some $i \in \{0, 1\}$, then $\bar{\mathcal{P}} = \mathcal{P}$, and
- (v) if $\bar{\mathcal{P}}$ is connectable, so is $\mathcal{P}$.

Proof. Parts (i)–(iii) follow from the observation that both x -separation and x -detour preserve singletons, balance, and the condition in (iii).

To see why part (iv) is true, note that if $\bar{\mathcal{P}}$ was formed by x -separation or x -detour, then both $\bar{\mathcal{P}}_{x,0}$ and $\bar{\mathcal{P}}_{x,1}$ contain at least one nonsingleton.

Finally, to verify part (v), observe that if $\mathcal{C}$ is a connector of $\bar{\mathcal{P}}$ and $\bar{\mathcal{P}}$ is obtained from $\mathcal{P}$ by x -separation (x -detour), then $\mathcal{C} \setminus \{P_{ac}, P_{c^x b}\} \cup \{P_{ac} + P_{c^x b}\}$ ($\mathcal{C} \setminus \{P_{ac}, P_{c^x d^x}, P_{db}\} \cup \{P_{ac} + P_{c^x d^x} + P_{db}\}$) is a connector of $\mathcal{P}$. $\square$

3. History. This section provides a survey of known results on connectability in Q_n . We start with a simple observation saying that each connectable pair set satisfies an obvious necessary condition, which can be expressed in terms of balance.

PROPOSITION 3.1. *If pair set $\mathcal{P}$ is connectable in $\langle V \rangle$ for some $V \subseteq V(Q_n)$, then $\beta(\mathcal{P}) = \beta(V)$.*

Proof. If $\mathcal{P}$ is connectable in $\langle V \rangle$, then $\sum_{\{a,b\} \in \mathcal{P}} \beta(V(P_{ab})) = \sum_{v \in V} \beta(v) = \beta(V)$. Since $\beta(u) \neq \beta(v)$ for any $\{u, v\} \in E(Q_n)$, it follows that

$$\beta(V(P_{ab})) = \begin{cases} 0 = \beta(a, b) & \text{if } \beta(a) \neq \beta(b), \\ \beta(a) = \beta(b) = \beta(a, b) & \text{otherwise.} \end{cases}$$

Consequently, $\beta(\mathcal{P}) = \sum_{\{a,b\} \in \mathcal{P}} \beta(a, b) = \sum_{\{a,b\} \in \mathcal{P}} \beta(V(P_{ab})) = \beta(V)$. $\square$

In particular, this means that any connectable pair set in Q_n must be balanced. On the other hand, there are previous results showing that balanced pair sets are connectable if their size is limited. A classical theorem on hamiltonian paths in hypercubes may be in our terms formulated in the following way.

THEOREM 3.2 (see [13]). *A balanced pair set of size one is connectable in Q_n .*

There is a generalization to pair sets of size two.

THEOREM 3.3 (see [6]). *A pair set $\mathcal{P} = \{\{a_1, b_1\}, \{a_2, b_2\}\}$ is connectable in Q_n provided $\beta(a_1, b_1) = \beta(a_2, b_2) = 0$.*

A further generalization to a nonconstant number of pairs has been obtained by Caha and Koubek.

THEOREM 3.4 (see [3]). *A strongly regular balanced pair set $\mathcal{P}$ is connectable in Q_n provided $|\mathcal{P}| \leq n/6$.*

The bound on the size of $|\mathcal{P}|$ has been recently improved.

THEOREM 3.5 (see [11]). *A strongly regular balanced pair set $\mathcal{P}$ is connectable in Q_n provided $|\mathcal{P}| \leq (n-1)/2$.*

4. Construction. The aim of this section is to transform the problem of connectability of regular pair sets to the strongly regular case so that we can apply Theorem 3.5. We start with a rather technical lemma that provides a sufficient condition which guarantees that there are always enough x -free vertices needed for our constructions. Put $l_x(\mathcal{P}) = |\{\{v, v^x\} \mid \{v, v^x\} \subseteq \bigcup \mathcal{P}\}|$.

LEMMA 4.1. *Let $\mathcal{P}$ be a pair set in Q_n , $x \in [n]$. If $n \geq 6$ and $|\mathcal{P}| \leq n + l_x(\mathcal{P})/2 - 1$, then for any $i \in \{0, 1\}$, $j \in \{1, -1\}$,*

$$|\{v \in V_{x,i} \mid v \text{ is } x\text{-free and } \beta(v) = j\}| \geq n.$$

Proof. Denote the left-hand side of the last inequality by $l_{i,j}$ and put $l := l_x(\mathcal{P})$. Since there are exactly $|\bigcup \mathcal{P}| - l$ vertices in $V_{x,i}$ that are not x -free and

$$\left| \bigcup \mathcal{P} \right| \leq 2|\mathcal{P}| \leq 2(n + l/2 - 1) = 2n + l - 2,$$

it follows that

$$l_{i,j} \geq |\{v \in V_{x,i} \mid \beta(v) = j\}| - \left(\left| \bigcup \mathcal{P} \right| - l \right) \geq 2^{n-2} - (2n - 2) \geq n$$

for $n \geq 6$. $\square$

Now we are ready to transform regular pair sets to strongly regular ones so that the balance is preserved. In the first step we describe a transformation of a pair set to an x -separate one.

LEMMA 4.2. *Let $n \geq 6$, $x \in [n]$, and $\mathcal{P}$ be a balanced pair set in Q_n such that $|\mathcal{P}| + |\mathcal{P}_{x,2}| \leq n$. Then there is an x -separate x -extension $\bar{\mathcal{P}}$ of $\mathcal{P}$ such that for both $i \in \{0, 1\}$*

- (i) $|\bar{\mathcal{P}}_{x,i}| = |\mathcal{P}_{x,i}| + |\mathcal{P}_{x,2}|$,
- (ii) $|\beta(\bar{\mathcal{P}}_{x,i})| \leq \min\{|\beta(\mathcal{P}_{x,0})|, |\beta(\mathcal{P}_{x,1})|\}$.

Proof. First note that since $\mathcal{P}$ is balanced, we have

$$(1) \quad \beta(\mathcal{P}_{x,0}) + \beta(\mathcal{P}_{x,1}) + \beta(\mathcal{P}_{x,2}) = 0.$$

If $|\mathcal{P}_{x,2}| = 0$, put $\bar{\mathcal{P}} = \mathcal{P}$, and note that condition (i) is trivial, while (ii) follows from (1). If $|\mathcal{P}_{x,2}| > 0$, let $i \in \{0, 1\}$ be such that $|\beta(\mathcal{P}_{x,i})| \leq |\beta(\mathcal{P}_{x,1-i})|$. We can assume that $\mathcal{P}_{x,2} = \{a_k, b_k\}_{k=1}^m$ such that $a_k \in V_{x,i}$ and $b_k \in V_{x,1-i}$ for each $k \in [m]$. Since $|\mathcal{P}| \leq n - |\mathcal{P}_{x,2}| \leq n - 1 \leq n + l_x(\mathcal{P})/2 - 1$, Lemma 4.1 guarantees that for each $k \in [m]$ there is a distinct x -free vertex $c_k \in V_{x,i}$ such that

$$(2) \quad \beta(a_k, c_k) = 0,$$

which means that

$$(3) \quad \beta(c_k^x, b_k) = \beta(a_k, b_k).$$

Observe that

$$\bar{\mathcal{P}} := \mathcal{P} \setminus \{\{a_k, b_k\} \mid k \in [m]\} \cup \{\{a_k, c_k\}, \{c_k^x, b_k\} \mid k \in [m]\}$$

is an x -separate x -extension of $\mathcal{P}$ satisfying (i). From (2) we have $\beta(\bar{\mathcal{P}}_{x,i}) = \beta(\mathcal{P}_{x,i})$, while from (3) we obtain $\beta(\bar{\mathcal{P}}_{x,1-i}) = \beta(\mathcal{P}_{x,1-i}) + \beta(\mathcal{P}_{x,2})$, which equals $-\beta(\mathcal{P}_{x,i})$ by (1). Hence (ii) holds as well. $\square$

The next lemma shows how to perform a transformation so that the final x -separate pair set $\bar{\mathcal{P}}$ has both parts $\bar{\mathcal{P}}_{x,0}$ and $\bar{\mathcal{P}}_{x,1}$ balanced.

LEMMA 4.3. *Let $n \geq 6$, $x \in [n]$, and $\mathcal{P}$ be an x -separate regular balanced pair set in Q_n such that $|\mathcal{P}| + |\beta(\mathcal{P}_{x,0})| \leq n + l_x(\mathcal{P})/2$. Then there is an x -separate x -extension $\bar{\mathcal{P}}$ of $\mathcal{P}$ such that for both $i \in \{0, 1\}$,*

- (i) $\bar{\mathcal{P}}_{x,i}$ is balanced,
- (ii) $|\bar{\mathcal{P}}_{x,i}| = |\mathcal{P}_{x,i}| + |\beta(\mathcal{P}_{x,i})|$.

Proof. First note that since $\mathcal{P}$ is balanced and x -separate, $\beta(\mathcal{P}_{x,0}) = -\beta(\mathcal{P}_{x,1})$. We argue by induction on $|\beta(\mathcal{P}_{x,0})|$. If $|\beta(\mathcal{P}_{x,0})| = 0$, put $\bar{\mathcal{P}} = \mathcal{P}$. If $|\beta(\mathcal{P}_{x,0})| > 0$, note that since $\mathcal{P}$ is regular, there is a nonsingleton $\{a, b\} \in \mathcal{P}_{x,j}$ for some $j \in \{0, 1\}$. As $|\mathcal{P}| \leq |\mathcal{P}| + |\beta(\mathcal{P}_{x,0})| - 1 \leq n + l_x(\mathcal{P})/2 - 1$, by Lemma 4.1 we can select x -free vertices $c, d \in V_{x,j}$ satisfying $\beta(c) = \beta(d) = -\text{sgn}(\beta(\mathcal{P}_{x,j}))$. Then $\mathcal{P}^* := \mathcal{P} \setminus \{\{a, b\}\} \cup \{\{a, c\}, \{b, d\}, \{c^x, d^x\}\}$ is an x -extension of $\mathcal{P}$. Since $\beta(a, c) + \beta(b, d) - \beta(a, b) = \beta(c, d) = -\text{sgn}(\beta(\mathcal{P}_{x,j}))$ while $\beta(c^x, d^x) = -\text{sgn}(\beta(\mathcal{P}_{x,1-j}))$, it follows that $|\beta(\mathcal{P}_{x,j}^*)| = |\beta(\mathcal{P}_{x,1-j}^*)| = |\beta(\mathcal{P}_{x,0})| - 1$ and that $\mathcal{P}^*$ is balanced. Moreover, since $l_x(\mathcal{P}^*) = l_x(\mathcal{P}) + 2$, we have $|\mathcal{P}^*| + |\beta(\mathcal{P}_{x,0}^*)| = (|\mathcal{P}| + 2) + (|\beta(\mathcal{P}_{x,0})| - 1) \leq n + l_x(\mathcal{P})/2 + 1 = n + l_x(\mathcal{P}^*)/2$. Therefore, by the induction hypothesis, there exists an x -separate x -extension $\bar{\mathcal{P}}$ of $\mathcal{P}^*$ such that for both $i \in \{0, 1\}$

- (i) $\bar{\mathcal{P}}_{x,i}$ is balanced, and
 - (ii) $|\bar{\mathcal{P}}_{x,i}| = |\mathcal{P}_{x,i}^*| + |\beta(\mathcal{P}_{x,i}^*)| = (|\mathcal{P}_{x,i}| + 1) + (|\beta(\mathcal{P}_{x,i})| - 1) = |\mathcal{P}_{x,i}| + |\beta(\mathcal{P}_{x,i})|$,
- as desired. $\square$

Now we can put the previous two lemmas together.

LEMMA 4.4. *Let $x \in [n]$ and $\mathcal{P}$ be a regular balanced pair set in Q_n such that $2 \leq |\mathcal{P}| \leq (n - |s(\mathcal{P})| - 3)/2$. Then there is an x -separate x -extension $\bar{\mathcal{P}}$ of $\mathcal{P}$ such that for both $i \in \{0, 1\}$,*

- (i) $\bar{\mathcal{P}}_{x,i}$ is balanced,
- (ii) $|\bar{\mathcal{P}}_{x,i}| \leq |\mathcal{P}|$, and
- (iii) $\bar{\mathcal{P}}_{x,i}$ is either regular or empty.

Proof. First note that we can assume that $n \geq 7$, for otherwise $|\mathcal{P}| \leq (n - 3)/2 < 2$, contrary to our assumption. Since $|\mathcal{P}| + |\mathcal{P}_{x,2}| \leq 2|\mathcal{P}| \leq n - 3 < n$, we can use Lemma 4.2 to obtain an x -separate x -extension $\mathcal{P}^*$ of $\mathcal{P}$ such that for both $i \in \{0, 1\}$

$$(4) \quad |\mathcal{P}_{x,i}^*| + |\beta(\mathcal{P}_{x,i}^*)| \leq |\mathcal{P}_{x,i}| + |\mathcal{P}_{x,2}| + \min_{j \in \{0,1\}} |\mathcal{P}_{x,j}| \leq |\mathcal{P}|.$$

Consequently, as $|\mathcal{P}_{x,1}^*| = |\mathcal{P}_{x,1}| + |\mathcal{P}_{x,2}|$ by part (i) of Lemma 4.2, we obtain that

$$\begin{aligned} |\mathcal{P}^*| + |\beta(\mathcal{P}_{x,0}^*)| &= |\mathcal{P}_{x,1}^*| + |\mathcal{P}_{x,0}^*| + |\beta(\mathcal{P}_{x,0}^*)| \\ &= |\mathcal{P}_{x,1}| + |\mathcal{P}_{x,2}| + |\mathcal{P}_{x,0}^*| + |\beta(\mathcal{P}_{x,0}^*)| \leq 2|\mathcal{P}| \leq n - 3 < n. \end{aligned}$$

Since $\mathcal{P}^*$ is also a regular balanced pair set by Proposition 2.2, we can use Lemma 4.3 to obtain an x -separate x -extension $\bar{\mathcal{P}}$ of $\mathcal{P}^*$ such that for both $i \in \{0, 1\}$, $\bar{\mathcal{P}}_{x,i}$ is balanced and

$$|\bar{\mathcal{P}}_{x,i}| = |\mathcal{P}_{x,i}^*| + |\beta(\mathcal{P}_{x,i}^*)| \leq |\mathcal{P}|,$$

using (4) for the last inequality. This verifies the validity of parts (i) and (ii) for $\bar{\mathcal{P}}$.

If part (iii) fails for $\bar{\mathcal{P}}$, then $\bar{\mathcal{P}}_{x,j} = s(\bar{\mathcal{P}}_{x,j}) \neq \emptyset$ for some $j \in \{0, 1\}$. Consequently, part (iv) of Proposition 2.2 implies that $\bar{\mathcal{P}} = \mathcal{P}$. In this case, we construct the desired x -separate x -extension of $\mathcal{P}$ in a different way.

Since $\mathcal{P}$ is regular and x -separate, there exists a nonsingleton $\{a, b\} \in \mathcal{P}_{x,1-j}$. As $|\mathcal{P}| < n$, by Lemma 4.1 there are x -free vertices $c, d \in V_{x,1-j}$ with $\beta(c) \neq \beta(d)$. Then $\tilde{\mathcal{P}} := \mathcal{P} \setminus \{\{a, b\}\} \cup \{\{a, c\}, \{b, d\}, \{c^x, d^x\}\}$ is an x -separate x -extension of $\mathcal{P}$ satisfying

- (i) $\beta(\tilde{\mathcal{P}}_{x,i}) = \beta(\mathcal{P}_{x,i}) = 0$ as $\beta(a, c) + \beta(b, d) = \beta(a, b)$ and $\beta(c^x, d^x) = 0$,
- (ii) $|\tilde{\mathcal{P}}_{x,i}| = |\mathcal{P}_{x,i}| + 1 \leq |\mathcal{P}_{x,i}| + |\mathcal{P}_{x,1-i}| = |\mathcal{P}|$ as $\mathcal{P}_{x,0} \neq \emptyset \neq \mathcal{P}_{x,1}$, and
- (iii) $\tilde{\mathcal{P}}_{x,i}$ is regular as $\{c^x, d^x\} \in \tilde{\mathcal{P}}_{x,j}$ while $\{a, c\} \in \tilde{\mathcal{P}}_{x,1-j}$

for both $i \in \{0, 1\}$. $\square$

The last lemma of this section settles the connectability of pair sets with exactly one singleton.

LEMMA 4.5. *A regular balanced pair set $\mathcal{P}$ with exactly one singleton is connectable in Q_n provided $|\mathcal{P}| \leq (n - 4)/2$.*

Proof. Let $s(\mathcal{P}) = \{\{a, a\}\}$. We claim that there exists $x \in [n]$ such that

$$(5) \quad \text{the only neighbor of } a^x \text{ in } \bigcup \mathcal{P} \text{ is } a.$$

To prove the claim, we say that a vertex $v \in \bigcup \mathcal{P}$ kills the neighbor a^x of a if $v = a \oplus e_x^n \oplus e_y^n$ for some $y \neq x$. Note that if v kills a neighbor of a , then $d(v, a) = 2$, and therefore

- (i) $\beta(a) = \beta(v)$, and
- (ii) v kills exactly two neighbors of a .

Suppose that the claim does not hold. That is, every neighbor of a is killed. Then, there are at least $n/2$ killing vertices in $\bigcup \mathcal{P} \setminus \{a\}$ by (ii). Moreover, they have the

same balance $\beta(a)$ by (i). Since $\mathcal{P}$ is balanced, the set $\bigcup \mathcal{P}$ contains another at least $n/2$ vertices of the balance $-\beta(a)$. But then, $\mathcal{P}$ contains at least $n/2$ pairs apart from the singleton $\{a, a\}$, which yields a contradiction with $|\mathcal{P}| \leq (n-4)/2$. Hence there exists $x \in [n]$ satisfying (5).

Next, observe that as $|s(\mathcal{P})| = 1$ and the regularity of $\mathcal{P}$ implies $|\mathcal{P}| > |s(\mathcal{P})|$, we have

$$2 \leq |\mathcal{P}| \leq (n - |s(\mathcal{P})| - 3)/2.$$

Now Lemma 4.4 yields that there exists an x -separate x -extension $\bar{\mathcal{P}}$ of $\mathcal{P}$ satisfying the conclusions of Lemma 4.4. Fix $i \in \{0, 1\}$ so that $a \in V_{x,i}$ and use Lemma 4.1 to choose $b \in V_{x,i}$ which is x -free with respect to $\bar{\mathcal{P}}$ and for which $\beta(a) = \beta(b)$. Since $\bar{\mathcal{P}}_{x,i}$ is balanced, the same holds for $\mathcal{P}_1 := \bar{\mathcal{P}}_{x,i} \setminus \{\{a, a\}\} \cup \{\{a, b\}\}$. Moreover, as $\{a, a\}$ is, by part (i) of Proposition 2.2, the sole singleton of $\bar{\mathcal{P}}$, $\mathcal{P}_1$ is strongly regular and $|\mathcal{P}_1| = |\bar{\mathcal{P}}_{x,i}| \leq |\mathcal{P}| \leq (n-4)/2 < ((n-1)-1)/2$. Hence by Theorem 3.5 there exists a connector $\mathcal{C}_1$ of $\mathcal{P}_1$ in $\langle V_{x,i} \rangle$. Denote by c the neighbor of a on the path $P_{ab} \in \mathcal{C}_1$ and note that $c \notin \bigcup \bar{\mathcal{P}}$ (as $\bar{\mathcal{P}}$ is x -separate). Since $c^x \notin \bigcup \mathcal{P}$ by (5), part (iii) of Proposition 2.2 implies that $c^x \notin \bigcup \bar{\mathcal{P}}$. Furthermore, as $\beta(c) = -\beta(a) = -\beta(b)$, we have $\beta(c^x, b^x) = \beta(c, b) = 0$. It remains to show that $\bar{\mathcal{P}}$ is connectable, which by part (v) of Proposition 2.2 implies the desired connectability of $\mathcal{P}$. We consider two cases (cf. Figure 1).

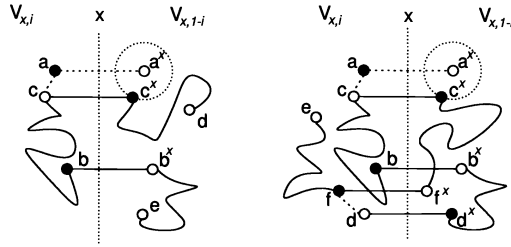


FIG. 1. Two cases in the proof of Lemma 4.5.

1. $\bar{\mathcal{P}}_{x,1-i} \neq \emptyset$. Then $\bar{\mathcal{P}}_{x,1-i}$ is regular by part (iii) of Lemma 4.4 and therefore there is a nonsingleton $\{d, e\} \in \bar{\mathcal{P}}_{x,1-i}$. Since $\bar{\mathcal{P}}_{x,1-i}$ is balanced, the same holds for $\mathcal{P}_2 := \bar{\mathcal{P}}_{x,1-i} \setminus \{\{d, e\}\} \cup \{\{d, c^x\}, \{b^x, e\}\}$ as $\beta(d, c^x) + \beta(b^x, e) = \beta(d, e) + \beta(c^x, b^x) = \beta(d, e)$ (recall that $\beta(c^x, b^x) = 0$). Moreover, as $s(\mathcal{P}_2) = s(\bar{\mathcal{P}}_{x,1-i}) = s(\mathcal{P}_{x,1-i}) = \emptyset$, $\mathcal{P}_2$ is strongly regular and $|\mathcal{P}_2| = |\bar{\mathcal{P}}_{x,1-i}| + 1 \leq |\mathcal{P}| + 1 \leq ((n-1)-1)/2$. Hence by Theorem 3.5 there exists a connector $\mathcal{C}_2$ of $\mathcal{P}_2$ in $\langle V_{x,1-i} \rangle$. The connector of $\bar{\mathcal{P}}$ in Q_n can be formed by $(\mathcal{C}_1 \setminus \{P_{ab}\}) \cup (\mathcal{C}_2 \setminus \{P_{dc^x}, P_{b^x e}\}) \cup \{P_{dc^x} + P_{cb} + P_{b^x e}\}$, where P_{cb} is the subpath of P_{ab} .

2. $\bar{\mathcal{P}}_{x,1-i} = \emptyset$. Since $\bar{\mathcal{P}}_{x,i}$ is balanced and $s(\bar{\mathcal{P}}_{x,i}) = s(\mathcal{P}_{x,i}) = \{\{a, a\}\}$, there exists $\{d, e\} \in \bar{\mathcal{P}}_{x,i}$ with $\beta(d) = -\beta(a)$. It follows that $\beta(d) = \beta(c)$. Let f be the neighbor of d on the path $P_{de} \in \mathcal{C}_1$. Observe that $\beta(d^x, b^x) = \beta(c^x, f^x) = 0$, and hence by Theorem 3.3 there exists a connector $\{P_{d^x b^x}, P_{c^x f^x}\}$ of $\{\{d^x, b^x\}, \{c^x, f^x\}\}$ in $\langle V_{x,1-i} \rangle$. The connector of $\bar{\mathcal{P}}$ can be formed by $\mathcal{C}_1 \setminus \{P_{ab}, P_{de}\} \cup \{P_{dd} + P_{d^x b^x} + P_{bc} + P_{c^x f^x} + P_{fe}\}$, where P_{bc} and P_{fe} are subpaths of P_{ba} and P_{de} , respectively. $\square$

5. Results. Now we can employ the constructions of the previous section to prove the main positive results.

THEOREM 5.1. *Let $\mathcal{P}$ be a regular pair set in Q_n such that $2|\mathcal{P}| + |s(\mathcal{P})| \leq n - 3$. Then $\mathcal{P}$ is connectable iff it is balanced.*

Proof. The necessity of the condition was already demonstrated by Proposition 3.1. To verify the sufficiency, assume that $\mathcal{P}$ is balanced. We argue by induction on $|s(\mathcal{P})|$.

1. If $|s(\mathcal{P})| = 0$, then $\mathcal{P}$ is strongly regular and therefore connectable by Theorem 3.5.

2. If $|s(\mathcal{P})| = 1$, then $\mathcal{P}$ is connectable by Lemma 4.5.

3. If $s(\mathcal{P})$ contains distinct singletons $\{a, a\}, \{b, b\}$, choose $x \in \pi(a \oplus b)$ and let $\bar{\mathcal{P}}$ be the x -separate x -extension of $\mathcal{P}$ obtained by Lemma 4.4. Since $s(\bar{\mathcal{P}}) = s(\mathcal{P})$ by Proposition 2.2, it follows that $|s(\bar{\mathcal{P}}_{x,i})| \leq |s(\mathcal{P})| - 1$ for both $i \in \{0, 1\}$. Recall that $\bar{\mathcal{P}}_{x,i}$ is balanced and its size does not exceed $|\mathcal{P}|$, which means that

$$2|\bar{\mathcal{P}}_{x,i}| + |s(\bar{\mathcal{P}}_{x,i})| \leq 2|\mathcal{P}| + |s(\mathcal{P})| - 1 \leq (n-1) - 3$$

for both $i \in \{0, 1\}$. Moreover, since $\bar{\mathcal{P}}_{x,i}$ contains either $\{a, a\}$ or $\{b, b\}$, it is nonempty and therefore regular by part (iii) of the same lemma. Hence by the induction hypothesis there exist connectors $\mathcal{C}_i$ of $\bar{\mathcal{P}}_{x,i}$ in $\langle V_{x,i} \rangle$ for both $i \in \{0, 1\}$. Since $\bar{\mathcal{P}}$ is x -separate, $\mathcal{C}_0 \cup \mathcal{C}_1$ forms a connector of $\bar{\mathcal{P}}$ in Q_n . It remains to apply part (v) of Proposition 2.2 to obtain a connector of $\mathcal{P}$. $\square$

The following corollaries show how the above result provides a partial solution to the problems formulated in section 2.

COROLLARY 5.2. *Let $\mathcal{F} \subseteq V(Q_n)$ and $\mathcal{P}$ be a strongly regular pair set in $Q_n - \mathcal{F}$ such that $2|\mathcal{P}| + 3|\mathcal{F}| \leq n - 3$. Then $\mathcal{P}$ is connectable in $Q_n - \mathcal{F}$ iff $\beta(\mathcal{P}) + \beta(\mathcal{F}) = 0$.*

Proof. Put $\bar{\mathcal{P}} = \mathcal{P} \cup \{\{a, a\} \mid a \in \mathcal{F}\}$. Note that $\bar{\mathcal{P}}$ is connectable in Q_n iff $\mathcal{P}$ is connectable in $Q_n - \mathcal{F}$, and $\beta(\bar{\mathcal{P}}) = \beta(\mathcal{P}) + \beta(\mathcal{F})$. Since

$$2|\mathcal{P}| + 3|\mathcal{F}| = 2|\bar{\mathcal{P}}| + |s(\bar{\mathcal{P}})|,$$

the statement of the corollary follows from Theorem 5.1. $\square$

COROLLARY 5.3. *Let $\mathcal{F} \subseteq V(Q_n)$ and u, v be distinct vertices of $Q_n - \mathcal{F}$. If $|\mathcal{F}| \leq (n-5)/3$, then $Q_n - \mathcal{F}$ contains a hamiltonian path between u and v iff $\beta(u, v) + \beta(\mathcal{F}) = 0$.*

Proof. Put $\mathcal{P} = \{\{u, v\}\}$ and apply Corollary 5.2. $\square$

COROLLARY 5.4. *Let $\mathcal{F} \subseteq V(Q_n)$. If $|\mathcal{F}| \leq (n-5)/3$, then $Q_n - \mathcal{F}$ contains a hamiltonian path iff $|\beta(\mathcal{F})| \leq 1$.*

Proof. Observe that there are distinct vertices u and v in $Q_n - \mathcal{F}$ such that $\beta(u, v) + \beta(\mathcal{F}) = 0$ iff $\beta(\mathcal{F}) \in \{-1, 0, 1\}$, and apply Corollary 5.3. $\square$

COROLLARY 5.5. *Let $\mathcal{F} \subseteq V(Q_n)$. If $|\mathcal{F}| \leq (n-5)/3$, then $Q_n - \mathcal{F}$ is hamiltonian iff $\mathcal{F}$ is balanced.*

Proof. Choose an arbitrary $\{u, v\} \in E(Q_n - \mathcal{F})$ and apply Corollary 5.3. $\square$

6. NP-completeness. The results of the previous section imply that there are bounds on the number $|\mathcal{F}|$ of faulty vertices and the size $|\mathcal{P}|$ of a pair set such that $\mathbf{HC}Q_n - V$, $\mathbf{HP}Q_n - V$, $\mathbf{HPP}Q_n - V$, and $\mathbf{PP}Q_n - V$ are decidable in polynomial time. This section is, on the other hand, devoted to negative results. We show that if $|\mathcal{F}|$ is not limited, all problems are NP-complete. We start with a similar problem for faulty edges whose NP-completeness is already known [4].

Hamiltonian cycle in hypercube with faulty edges ($\mathbf{HC}Q_n - E$).

INSTANCE: A spanning subgraph G of Q_n for some n .

QUESTION: Is G hamiltonian?

For an instance of the problem $\mathbf{HC}Q_n - E$, we can assume that G has minimal degree at least two; otherwise it cannot contain a hamiltonian cycle. We will transform it to an instance of the problem $\mathbf{HC}Q_n - V$ in Q_{5n+2} . Each vertex $u \in V(G)$

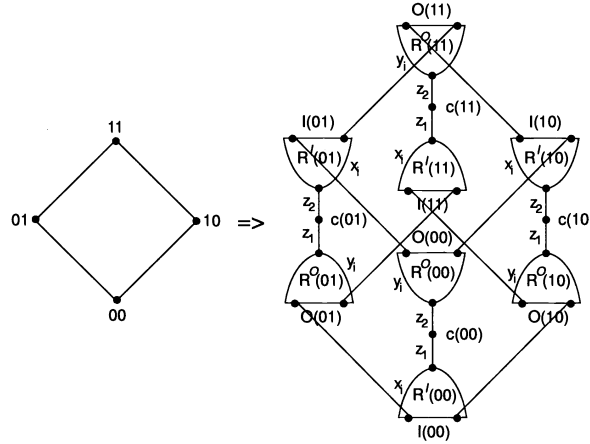


FIG. 2. A scheme of routers.

will be replaced with a set $R(u)$ of vertices of Q_{5n+2} , called a router of u ; see Figure 2. It consists of a center $c(u)$, an in-gate $R^I(u)$, and an out-gate $R^O(u)$. The in-gate contains a set $I(u) \subset R^I(u)$ of special vertices called in-ports of u ; similarly, we define a set $O(u) \subset R^O(u)$ of out-ports of u . Our aim is to guarantee that for every edge $\{u, v\} \in E(G)$, routers $R(u)$ and $R(v)$ are joined with exactly two edges $\{u^I, v^O\}, \{u^O, v^I\}$ of Q_{5n+2} , where $u^I \in I(u)$, $u^O \in O(u)$, $v^I \in I(v)$, and $v^O \in O(v)$, while no other edge joins different routers.

There are two similar kinds of routers $R(u)$, depending on $\beta(u)$. A design of the out-gate $R^O(u)$ depends on the edges of G incident with u , while the design of the in-gate $R^I(u)$ is the same for all vertices u . The position of $R(u)$ in Q_{5n+2} is uniquely determined by its center $c(u)$, which will be specified later.

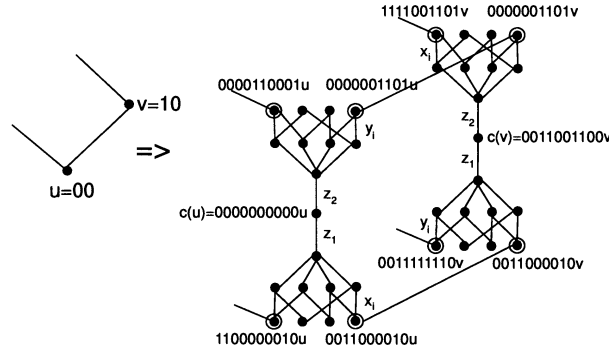


FIG. 3. Router details. The vertices in circles are ports.

Let $z_1 = e_{4n+1}^m$, $z_2 = e_{4n+2}^m$ and $x_i = e_i^m$, $y_i = e_{2n+i}^m$, where $i \in [2n]$ and $m = 5n+1$. For our convenience, put $x_0 = x_{2n}$, $y_0 = y_{2n}$ and $x_{2n+1} = x_1$, $y_{2n+1} = y_1$. First we describe $R(u)$ for $u \in V(G)$ with $\beta(u) = 1$. Recall that as G is a subgraph of Q_n , any edge $\{u, v\} \in E(G)$ satisfies $u \oplus v = e_i$ for some $i \in [n]$. Let $i_1 < \dots < i_k$ be all $i_j \in [n]$ with $\{u, u \oplus e_{i_j}\} \in E(G)$ for $j \in [k]$. Note that we have $k \geq 2$ since u

is incident with at least two edges. For our convenience, put $i_0 = i_k$ and $i_{k+1} = i_1$. Suppose we already have a center $c(u) \in V(Q_{5n+2})$, and define

$$\begin{aligned}
 (6) \quad & p = c(u) \oplus z_1, \\
 & I(u) = \{p \oplus x_{2l-1} \oplus x_{2l} \mid l \in [n]\}, \\
 & R^I(u) = \{p, p \oplus x_{2l-1}, p \oplus x_{2l}, p \oplus x_{2l} \oplus x_{2l+1} \mid l \in [n]\} \cup I(u), \\
 & o = c(u) \oplus z_2, \\
 & O(u) = \{o \oplus y_{2i_j-1} \oplus y_{2i_j} \mid j \in [k]\}, \\
 & R^O(u) = \{o, o \oplus y_{2i_j-1}, o \oplus y_{2i_j}, o \oplus y_{2i_j} \oplus y_{2i_{j+1}-1} \mid j \in [k]\} \cup O(u), \\
 & R(u) = R^O(u) \cup R^I(u) \cup \{c(u)\}.
 \end{aligned}$$

The router $R(u)$ for u with $\beta(u) = -1$ is constructed analogously; merely switch z_1 and z_2 . See Figure 3 for an illustration.

LEMMA 6.1. *For any spanning subgraph G of Q_n and $u \in V(G)$ the following properties hold:*

- (i) *For every $a \in I(u)$, $b \in O(u)$ there is a hamiltonian path of $\langle R(u) \rangle$ between a and b .*
- (ii) *Let $\mathcal{P} = \{P_{a_i b_i}\}_{i=1}^m$ be a path partition of $\langle R(u) \rangle$ such that $a_i, b_i \in I(u) \cup O(u)$ for all $i \in [m]$. Then $\mathcal{P} = \{P_{ab}\}$ for some $a \in I(u)$ and $b \in O(u)$.*

Proof. Let $l \in [n]$ and $j \in [k]$ be such that $a = p \oplus x_{2l-1} \oplus x_{2l}$ and $b = o \oplus y_{2i_j-1} \oplus y_{2i_j}$. Observe that

$$p \oplus x_1, p \oplus x_1 \oplus x_2, p \oplus x_2, p \oplus x_2 \oplus x_3, \dots, p \oplus x_{2n}, p \oplus x_{2n} \oplus x_1$$

is a hamiltonian cycle of $\langle R^I(u) \setminus \{p\} \rangle$. Similarly,

$$o \oplus y_{2i_1-1}, o \oplus y_{2i_1-1} \oplus y_{2i_1}, o \oplus y_{2i_1}, o \oplus y_{2i_1} \oplus y_{2i_2-1}, \dots, o \oplus y_{2i_k}, o \oplus y_{2i_k} \oplus y_{2i_1-1}$$

is a hamiltonian cycle of $\langle R^O(u) \setminus \{o\} \rangle$. Thus

$$\begin{aligned}
 & a, p \oplus x_{2l}, p \oplus x_{2l} \oplus x_{2l+1}, p \oplus x_{2l+1}, \dots, p \oplus x_{2l-2} \oplus x_{2l-1}, p \oplus x_{2l-1}, p, c(u), \\
 & o, o \oplus y_{2i_j-1}, o \oplus y_{2i_j-1} \oplus y_{2i_j}, \dots, o \oplus y_{2i_{j+1}-1}, o \oplus y_{2i_j} \oplus y_{2i_{j+1}-1}, o \oplus y_{2i_j}, b
 \end{aligned}$$

is a hamiltonian path of $\langle R(u) \rangle$ between a and b and part (i) holds.

If the assumption of part (ii) is satisfied, by Proposition 3.1 we have $\beta(R(u)) = \beta(\{a_i, b_i\}_{i=1}^m)$. Observe that $\beta(v)$ is the same for all $v \in I(u) \cup O(u)$, and therefore $\beta(\{a_i, b_i\}_{i=1}^m) \in \{m, -m\}$. Moreover, Proposition 3.1 applied to part (i) of this lemma reveals that $\beta(R(u)) \in \{1, -1\}$. It follows that $m = 1$. Moreover, (6) implies that $c(u)$ is a cutvertex of $\langle R(u) \rangle$ such that $I(u)$ and $O(u)$ belong to different components of $\langle R(u) - \{c(u)\} \rangle$. As $P_{a_1 b_1}$ must pass through $c(u)$, it cannot have both endvertices in $I(u)$ or both in $O(u)$. $\square$

At this point we are ready to describe the center of a router by defining the mapping $c : V(Q_n) \rightarrow V(Q_{5n+2})$ such that

$$(7) \quad c(u_1 u_2 \dots u_n) = u_1 u_1 u_2 u_2 \dots u_n u_n u_1 u_1 u_2 u_2 \dots u_n u_n 0 u_1 u_2 \dots u_n.$$

Note that this together with (6) uniquely determines the router $R(u)$ for any vertex $u \in V(G)$.

LEMMA 6.2. *For a spanning subgraph G of Q_n and two distinct vertices $u, v \in V(G)$, the set of all edges of Q_{5n+2} joining a vertex of $R(u)$ with a vertex of $R(v)$*

- (i) contains exactly two edges $\{u^I, v^O\}, \{u^O, v^I\}$ such that $u^I \in I(u)$, $u^O \in O(u)$, $v^I \in I(v)$, and $v^O \in O(v)$ if $\{u, v\} \in E(G)$, and
- (ii) is empty if $\{u, v\} \notin E(G)$.

Proof. Put $m = 5n + 2$ and let $a \in R(u)$, $b \in R(v)$ be arbitrary vertices. First note that as $d_{Q_m}(c(u), c(v)) = 5d_{Q_n}(u, v)$ by (7), the triangle inequality implies

$$(8) \quad d_{Q_m}(a, b) \geq 5d(u, v) - d_{Q_m}(c(u), a) - d_{Q_m}(c(v), b).$$

Moreover, $d_{Q_m}(a, c(u)), d_{Q_m}(b, c(v)) \leq 3$ by (6). Hence if $\{u, v\} \notin E(Q_n)$, then (8) implies $d_{Q_m}(a, b) \geq 4$, and therefore $R(u)$ and $R(v)$ are not joined by any edge in this case.

If $\{u, v\} \in E(Q_n)$, then $u \oplus v = e_i$ for some $i \in [n]$. We can assume without loss of generality that $\beta(u) = 1 = -\beta(v)$. Assume that $\{a, b\} \in E(Q_m)$. Since $\pi(a \oplus c(u)) \subseteq [4n + 2]$, $\pi(b \oplus c(v)) \subseteq [4n + 2]$, and $4n + 2 + i \in \pi(c(u), c(v))$ by (6) and (7), it follows that $a \oplus b = e_{4n+2+i}^m$. Moreover,

$$c(u) \oplus c(v) = x_{2i-1} \oplus x_{2i} \oplus y_{2i-1} \oplus y_{2i} \oplus e_{4n+2+i}^m$$

by (7), so

$$a \oplus b = c(u) \oplus c(v) \oplus x_{2i-1} \oplus x_{2i} \oplus y_{2i-1} \oplus y_{2i}.$$

From (6) observe that for every $r, s \in \{2l, 2l - 1 \mid l \in [n]\}$, the set $\pi(a \oplus c(u))$ cannot contain both x_r and y_s . Moreover, if $x_r \in \pi(a \oplus c(u))$ or $y_s \in \pi(a \oplus c(u))$, then $z_j \in \pi(a \oplus c(u))$ for some $j \in \{1, 2\}$. The same also holds for the set $\pi(b \oplus c(v))$. Hence we obtain that

$$a = c(u) \oplus z_j \oplus a', \quad b = c(v) \oplus z_j \oplus b',$$

where $j \in \{1, 2\}$ and

$$\{a', b'\} = \{x_{2i-1} \oplus x_{2i}, y_{2i-1} \oplus y_{2i}\}.$$

Therefore, we have either

$$a = c(u) \oplus z_1 \oplus x_{2i-1} \oplus x_{2i}, \quad b = c(v) \oplus z_1 \oplus y_{2i-1} \oplus y_{2i},$$

which means that $a \in I(u)$, $b \in O(v)$, or

$$a = c(u) \oplus z_2 \oplus y_{2i-1} \oplus y_{2i}, \quad b = c(v) \oplus z_2 \oplus x_{2i-1} \oplus x_{2i},$$

which means that $a \in O(u)$, $b \in I(v)$. It remains only to observe that if $\{u, v\} \in E(Q_n) \setminus E(G)$, then by (6) we have $b \notin R(v)$ in the former case or $a \notin R(u)$ in the latter case, which contradicts the assumption that $\{a, b\} \in E(Q_m)$. $\square$

THEOREM 6.3. **HCQ_n - V** and **HPPQ_n - V** are NP-complete problems.

Proof. Both problems are clearly in NP. To prove that they are also NP-hard, we describe polynomial-time transformations from **HCQ_n - E** to both **HCQ_n - V** and **HPPQ_n - V**.

Let G be a spanning subgraph of Q_n of minimum degree at least two, representing an instance of **HCQ_n - E**. We construct a set $S \subseteq V(Q_{5n+2})$ of size at most $(8n+3)2^n$ and a pair a, b of distinct vertices of S such that S is an instance of **HCQ_n - V**, while S, a, b is an instance of **HPPQ_n - V** satisfying

$$(9) \quad G \text{ is hamiltonian} \begin{cases} \text{(a) iff } \langle S \rangle \text{ is hamiltonian,} \\ \text{(b) iff } \langle S \rangle \text{ contains a hamiltonian path between } a \text{ and } b. \end{cases}$$

To that end, put $S = \bigcup_{u \in V(G)} R(u)$, where $R(u)$ is the router of u described by (6) and (7). Then choose an arbitrary $u \in V(Q_n)$ and put $a = c(u)$, $b = c(u) \oplus z_1$, where $c(u)$ and z_1 are defined by (7) and (6), respectively. This completes the construction. We leave it to the reader to verify that it really describes an algorithm whose time complexity is polynomial in the input size of G , which is $\Theta(n2^n)$.

To verify part (a) of (9), first suppose that there exists a hamiltonian cycle $u_1, u_2, \dots, u_{2^n}$ of G . Put $u_{2^n+1} = u_1$ and observe that Lemma 6.2 guarantees the existence of a pair of adjacent vertices $u_i^O \in O(u_i)$, $u_{i+1}^I \in I(u_{i+1})$ for all $i \in [2^n]$. Using part (i) of Lemma 6.1 we can connect u_i^I and u_i^O with a hamiltonian path P_i of $\langle R(u_i) \rangle$. Then the path $P_1 + P_2 + \dots + P_{2^n}$ with the edge $\{u_{2^n}^O, u_1^I\}$ appended forms a hamiltonian cycle of $\langle S \rangle$.

On the other hand, suppose there is a hamiltonian cycle H of $\langle S \rangle$. Then for any $u \in V(G)$, the subgraph of H induced by $R(u)$ forms a path partition $\mathcal{P}_u = \{P_{a_i b_i}\}_{i=1}^m$ of $\langle R(u) \rangle$. Moreover, Lemma 6.2 implies that $a_i, b_i \in I(u) \cup O(u)$ for all $i \in [m]$, while part (ii) of Lemma 6.1 says that then $\mathcal{P}_u = \{P_{a_u b_u}\}$ for some $a_u \in I(u)$ and $b_u \in O(u)$. It remains to apply Lemma 6.2 again to see that $\{b_u, a_v\} \in E(H)$ implies $\{u, v\} \in E(G)$, and therefore the cycle C with vertices $V(C) = V(G)$ and edges $E(C) = \{\{u, v\} \mid \{b_u, a_v\} \in E(H)\}$ forms a hamiltonian cycle of G .

Finally, to verify part (b) of (9), note that by (6), $c(u) \oplus z_1$ is one of the two neighbors of $c(u)$ in $\langle R(u) \rangle$. Moreover, Lemma 6.2 guarantees that there are no other edges incident with $c(u)$ in $\langle S \rangle$. Hence any hamiltonian cycle of $\langle S \rangle$ must pass through the edge $\{c(u), c(u) \oplus z_1\}$. It follows that $\langle S \rangle$ is hamiltonian iff $\langle S \rangle$ contains a hamiltonian path between a and b , and therefore the validity of part (a) implies that part (b) holds as well. $\square$

Since each instance of $\mathbf{HPPQ}_n - \mathbf{V}$ is also an instance of $\mathbf{PPQ}_n - \mathbf{V}$, we immediately obtain the following.

COROLLARY 6.4. $\mathbf{PPQ}_n - \mathbf{V}$ is an NP-complete problem.

The variant without prescribed vertices is also NP-complete.

COROLLARY 6.5. $\mathbf{HPQ}_n - \mathbf{V}$ is an NP-complete problem.

Proof. Transform $\mathbf{HPPQ}_n - \mathbf{V}$ to $\mathbf{HPQ}_n - \mathbf{V}$ using the observation that if G is an induced subgraph of Q_n and a, b distinct vertices of G , then graph G' obtained from G by adding new vertices a' and b' and edges $\{a, a'\}$, $\{b, b'\}$ is an induced subgraph of Q_{n+2} such that G contains a hamiltonian path between a and b iff G' contains a hamiltonian path. $\square$

The following problem is studied in [17, 18] for the purpose of data compression.

Traveling salesman in faulty hypercube ($\mathbf{TSQ}_n - \mathbf{V}$).

INSTANCE: A sequence $v_1, v_2, \dots, v_m$ of distinct vertices of Q_n for some n and a positive integer k .

QUESTION: Is there a permutation $p : [m] \rightarrow [m]$ such that

$$\sum_{i=1}^{m-1} d(v_{p(i)}, v_{p(i+1)}) \leq k?$$

Since each instance of $\mathbf{HPQ}_n - \mathbf{V}$ also forms an instance of $\mathbf{TSQ}_n - \mathbf{V}$ with $k = m - 1$, we obtain the following corollary.

COROLLARY 6.6. $\mathbf{TSQ}_n - \mathbf{V}$ is an NP-complete problem.

7. Concluding remarks. The results of section 5 provide bounds on the number of faulty vertices or on the size of a pair set, which guarantees that the four problems considered in this paper are polynomial. However, the values we obtained

are probably not optimal. In particular, let $f(n)$ denote the maximum integer such that the statement of Corollary 5.5 holds for every $\mathcal{F} \subseteq V(Q_n)$ with $|\mathcal{F}| \leq f(n)$. We showed that $f(n) \geq (n-5)/3$. On the other hand, consider a set $\mathcal{F}$ containing $n-1$ neighbors of some vertex u of Q_n and another $n-1$ vertices different from u but of the same balance as u . Then $\mathcal{F}$ is a balanced set of $2n-2$ vertices of Q_n , but $Q_n - \mathcal{F}$ is not hamiltonian since the vertex u has only one nonfaulty neighbor. It follows that $f(n) < 2n-2$. It is natural to ask, What is the exact value of $f(n)$?

In section 6 we showed that the existence of a hamiltonian cycle in an induced subgraph of Q_n is an NP-complete problem. We can proceed one step further and address the complexity of the same problem for distance-preserving subgraphs of a hypercube, also known as partial cubes. Note that the induced subgraph of Q_{5n+2} , constructed in the proof of Theorem 6.3 to form an instance of $\mathbf{HC}Q_n - \mathbf{V}$, is not a partial cube, so our method does not resolve this problem. It should be also noted that recognizing partial cubes is known to be polynomial [5].

Acknowledgments. We are very grateful to the anonymous referees for their insightful comments that essentially improved the exposition of the paper. We are also obliged to Václav Koubek for fruitful discussions on this topic.

REFERENCES

- [1] Y. A. ASHIR AND I. A. STEWART, *Fault-tolerant embeddings of hamiltonian circuits in k -ary n -cubes*, SIAM J. Discrete Math., 15 (2002), pp. 317–328.
- [2] B. BOLLOBÁS, *Modern Graph Theory*, 2nd ed., Springer-Verlag, New York, 2002.
- [3] R. CAHA AND V. KUBEK, *Spanning multi-paths in hypercubes*, Discrete Math., 307 (2007), pp. 2053–2066.
- [4] M. Y. CHAN AND S.-J. LEE, *On the existence of Hamiltonian circuits in faulty hypercubes*, SIAM J. Discrete Math., 4 (1991), pp. 511–527.
- [5] D. Z. DJOKOVIĆ, *Distance-preserving subgraphs of hypercubes*, J. Combin. Theory Ser. B, 14 (1973), pp. 263–267.
- [6] T. DVOŘÁK, *Hamiltonian cycles with prescribed edges in hypercubes*, SIAM J. Discrete Math., 19 (2005), pp. 135–144.
- [7] T. DVOŘÁK, P. GREGOR, AND V. KUBEK, *Spanning paths in hypercubes*, Discrete Math. Theor. Comput. Sci., AE (2005), pp. 363–368.
- [8] H. ENOMOTO AND K. OTA, *Partitions of a graph into paths with prescribed endvertices and lengths*, J. Graph Theory, 34 (2000), pp. 163–169.
- [9] J.-S. FU, *Fault-tolerant cycle embedding in the hypercube*, Parallel Comput., 29 (2003), pp. 821–832.
- [10] J.-S. FU, *Longest fault-free paths in hypercubes with vertex faults*, Inform. Sci., 176 (2006), pp. 759–771.
- [11] P. GREGOR AND T. DVOŘÁK, *Path partitions of hypercubes*, Inform. Process. Lett., to appear.
- [12] F. HARARY, J. P. HAYES, AND H.-J. WU, *A survey of the theory of hypercube graphs*, Comput. Math. Appl., 15 (1988), pp. 277–289.
- [13] I. HAVEL, *On hamiltonian circuits and spanning trees of hypercubes*, Čas. Pěst. Mat., 109 (1984), pp. 135–152.
- [14] F. T. LEIGHTON, *Introduction to Parallel Algorithms and Architectures: Arrays, Trees, Hypercubes*. Morgan Kaufmann, San Mateo, CA, 1992.
- [15] M. LEWINTER AND W. WIDULSKI, *Hyper-Hamilton laceable and caterpillar-spannable product graphs*, Comput. Math. Appl., 34 (1997), pp. 99–104.
- [16] J.-H. PARK, H.-C. KIM, AND H.-S. LIM, *Many-to-many disjoint path covers in hypercube-like interconnection networks with faulty elements*, IEEE Trans. Parallel and Distributed Systems, 17 (2006), pp. 227–240.
- [17] A. PINAR, T. TAO, AND H. FERHATOSMANOGLU, *Compressing bitmap indices by data reorganization*, in Proceedings of the 21st International Conference on Data Engineering (ICDE 2005), IEEE Computer Society, Washington, DC, 2005, pp. 310–321.
- [18] D. RICHARDS, *Data compression and Gray-code sorting*, Inform. Process. Lett., 22 (1986), pp. 201–205.

- [19] C. SAVAGE, *A survey of combinatorial Gray codes*, SIAM Rev., 39 (1997), pp. 605–629.
- [20] C.-H. TSAI, *Linear array and ring embeddings in conditional faulty hypercubes*, Theoret. Comput. Sci., 314 (2004), pp. 431–443.
- [21] C.-H. TSAI, J. J. M. TAN, T. LIANG, AND L.-H. HSU, *Fault-tolerant hamiltonian laceability of hypercubes*, Inform. Process. Lett., 83 (2002), pp. 301–306.

TOTAL-COLORING OF PLANE GRAPHS WITH MAXIMUM DEGREE NINE*

ŁUKASZ KOWALIK[†], JEAN-SÉBASTIEN SERENI[‡], AND RISTE ŠKREKOVSKI[§]

Abstract. The central problem of the total-colorings is the total-coloring conjecture, which asserts that every graph of maximum degree Δ admits a $(\Delta + 2)$ -total-coloring. Similar to edge-colorings—with Vizing’s edge-coloring conjecture—this bound can be decreased by 1 for plane graphs of higher maximum degree. More precisely, it is known that if $\Delta \geq 10$, then every plane graph of maximum degree Δ is $(\Delta + 1)$ -totally-colorable. On the other hand, such a statement does not hold if $\Delta \leq 3$. We prove that every plane graph of maximum degree 9 can be 10-totally-colored.

Key words. total-coloring, planar graph, discharging method

AMS subject classification. 05C15

DOI. 10.1137/070688389

1. Introduction. Given a graph $G = (V, E)$ and a positive integer k , a k -total-coloring of G is a mapping $\lambda : V \cup E \rightarrow \{1, 2, \dots, k\}$ such that

- (i) $\lambda(u) \neq \lambda(v)$ for every pair u, v of adjacent vertices,
- (ii) $\lambda(v) \neq \lambda(e)$ for every vertex v and every edge e incident to v ,
- (iii) $\lambda(e) \neq \lambda(e')$ for every pair e, e' of incident edges.

This notion was independently introduced by Behzad [3] in his doctoral thesis and Vizing [15]. It is now a prominent notion in graph coloring, to which a whole book is devoted [17]. Both Behzad and Vizing made the celebrated *total-coloring conjecture*, stating that every graph of maximum degree Δ admits a $(\Delta + 2)$ -total-coloring. Notice that every such graph cannot be totally-colored with less than $\Delta + 1$ colors, and that a cycle of length 5 cannot be 3-totally-colored. The best general bound so far has been obtained by Molloy and Reed [10], who established that every graph of maximum degree Δ can be $(\Delta + 10^{26})$ -totally-colored. Moreover, the conjecture has been shown to be true for several special cases, namely for $\Delta = 3$ by Rosenfeld [11] and Vijayaditya [14], and then for $\Delta \in \{4, 5\}$ by Kostochka [9].

Another natural subclass to consider is the one of planar graphs, which has attracted a considerable amount of attention and for which several results were obtained. First, Borodin [5] proved that if $\Delta \geq 9$, then every plane graph of maximum degree Δ fulfills the conjecture. This result can be extended to the case where $\Delta = 8$ by the use of the four color theorem [1, 2], combined with Vizing’s theorem about edge coloring—the reader can consult the book by Jensen and Toft [8] for more details.

*Received by the editors April 16, 2007; accepted for publication April 25, 2008; published electronically October 1, 2008.

<http://www.siam.org/journals/sidma/22-4/68838.html>

[†]Institute of Informatics, Warsaw University, Banacha 2, 02-097 Warsaw, Poland (kowalik@mimuw.edu.pl). This author’s work was supported by a grant from the Ministry of Science and Higher Education, project N206 005 32/0807.

[‡]Institute for Theoretical Computer Science (ITI) and Department of Applied Mathematics (KAM), Faculty of Mathematics and Physics, Charles University, Malostranské náměstí 25, 118 00 Prague 1, Czech Republic (sereni@kam.mff.cuni.cz). This author’s work was supported by the European project IST FET AEOLUS.

[§]Department of Mathematics, University of Ljubljana, Jadranska 19, 1111 Ljubljana, Slovenia (bluesky2high@yahoo.com). This author’s work was partially supported by the Ministry of Science and Technology of Slovenia, Research Program P1-0297.

Elsewhere, Sanders and Zhao [12] solved the case $\Delta = 7$ of the total-coloring conjecture for plane graphs. So the only open case regarding plane graphs is $\Delta = 6$. Interestingly, $\Delta = 6$ is also the only remaining open case for Vizing's edge-coloring conjecture, after Sanders and Zhao [13] resolved the case $\Delta = 7$.

However, plane graphs with high maximum degree allow a stronger assertion. More precisely, Borodin [5] showed that if $\Delta \geq 14$, then every plane graph with maximum degree Δ is $(\Delta + 1)$ -totally-colorable, and asked whether 14 could be decreased. Borodin, Kostochka, and Woodall extended this result to the case where $\Delta \geq 12$ [6], and later to $\Delta = 11$ [7]. Recently, Wang [16] established the result for $\Delta = 10$. On the other hand, this bound is not true if $\Delta \leq 3$. The complete graphs K_2 , K_4 and the cycles of length $3k + 2$ with $k \geq 1$ are examples of plane graphs that cannot be $(\Delta + 1)$ -totally-colored. We continue along those lines and establish the following theorem.

THEOREM 1. *Every plane graph of maximum degree 9 is 10-totally-colorable.*

So, the values of Δ , for which it is not known whether all plane graphs of maximum degree Δ are $(\Delta + 1)$ -totally-colorable, are now 4, 5, 6, 7, and 8. Recall that the case where $\Delta = 6$ is even open for the total-coloring conjecture. We also note that if $\Delta \geq 3$, then every outerplane graph with maximum degree Δ can be $(\Delta + 1)$ -totally-colored [19]. Another result of the same type is that every Halin graph of maximum degree 4 admits a 5-total-coloring [18]. Note also that the complete r -partite balanced graph K_{r*n} , whose maximum degree Δ is $n(r - 1)$, admits a $(\Delta + 2)$ -total-coloring, and the cases where this bound can be decreased by 1 have been characterized [4].

We prove Theorem 1 by contradiction. From now on, we let $G = (V, E)$ be a minimum counterexample to the statement of Theorem 1, in the sense that the quantity $|V| + |E|$ is minimum. In particular, every proper subgraph of G is 10-totally-colorable. First, we establish various structural properties of G in section 2. Then, relying on these properties, we use the discharging method in section 3 to obtain a contradiction.

In what follows, a vertex of degree d is called a d -vertex. A vertex is a $(\leq d)$ -vertex if its degree is at most d ; it is a $(\geq d)$ -vertex if its degree is at least d . If f is a face of G , then the *degree of f* is its length, i.e., the number of its incident vertices. The notions of d -face, $(\leq d)$ -face, and $(\geq d)$ -face are defined analogously as for the vertices. Moreover, if a vertex v is adjacent to a d -vertex u , then we say that u is a d -neighbor of v . A cycle of length 3 is called a *triangle*. For integers a, b, c , an $(\leq a, \leq b, \leq c)$ -triangle is a triangle xyz of G with $\deg(x) \leq a$, $\deg(y) \leq b$, and $\deg(z) \leq c$. The notions of $(a, \leq b, \leq c)$ -triangles, $(a, b, \geq c)$ -triangles, and so on are defined analogously.

2. Reducible configurations. In this section, we establish some structural properties of the graph G . We prove that some planar graphs are *reducible configurations*; i.e., they cannot be subgraphs of G .

For convenience, we sometimes define configurations by depicting them in figures. In all of the figures of this paper, 2-vertices are represented by small black bullets, 3-vertices by black triangles, 4-vertices by black squares, and white bullets represent vertices whose degree is at least the one shown on the figure.

Let λ be a (partial) 10-total-coloring of G . For each element $x \in V \cup E$, we define $\mathcal{C}(x)$ to be the set of colors (with respect to λ) of vertices and edges incident or adjacent to x . Also, we set $\mathcal{F}(x) := \{1, 2, \dots, 10\} \setminus \mathcal{C}(x)$. If $x \in V$, then we define $\mathcal{E}(x)$ to be the set of colors of the edges incident to x . Moreover, λ is *nice* if only some (≤ 4) -vertices are not colored. Observe that every nice coloring can be greedily extended to a 10-total-coloring of G , since $|\mathcal{C}(v)| \leq 8$ for each (≤ 4) -vertex v ; i.e., v

has at most 8 forbidden colors. Therefore, in the rest of this paper, we shall always suppose that such vertices are colored at the very end. More precisely, every time we consider a partial coloring of G , we uncolor all (≤ 4)-vertices, and implicitly color them at the very end of the coloring procedure of G . We make the following observation about nice colorings and use it implicitly throughout this paper.

Observation. Let uv be an edge with $\deg(v) \leq 4$. There exists a nice coloring λ of $G - e$, in which u is colored and v is uncolored. Moreover, it then suffices to properly color the edge e with a color from $\{1, 2, \dots, 10\}$ to extend λ to a nice coloring of G .

We now study the structural properties of G in a series of lemmas.

LEMMA 2. *The graph G has the following properties:*

- (i) *the minimum degree is at least 2;*
- (ii) *if vu is an edge with $\deg(v) \leq 4$, then $\deg(u) \geq 11 - \deg(v)$;*
- (iii) *a 9-vertex is adjacent to at most one 2-vertex;*
- (iv) *a triangle incident to a 3-vertex must also contain a 9-vertex;*
- (v) *there is no $(4, \leq 7, \leq 8)$ -triangle;*
- (vi) *a triangle contains at most one (≤ 5) -vertex.*

Proof. (i) Suppose that v is a 1-vertex, and let u be its neighbor. By the minimality of G , the graph $G - v$ admits a nice coloring in which u is colored. Since the degree of u in $G - v$ is at most 8, we obtain $|\mathcal{C}(vu)| \leq 9$. Thus, the edge vu can be properly colored, which yields a nice coloring of G .

(ii) Suppose that $vu \in E$ with $\deg(v) \leq 4$ and $\deg(u) \leq 10 - \deg(v)$. There exists a nice coloring of $G' := G - vu$, in which u is colored and v is uncolored. Therefore, $|\mathcal{C}(vu)| \leq \deg(v) - 1 + \deg(u) - 1 + 1 \leq 9$. Hence we can color properly the edge vu , thereby obtaining a nice coloring of G .

(iii) Suppose that v is a 9-vertex adjacent to two 2-vertices x and y . Let x' be the neighbor of x different from v , and let y' be the neighbor of y different from v . Notice that we may have $x' = y'$. By the previous assertion, x' and y' are 9-vertices. It is enough to consider the following two possibilities.

v is adjacent to neither x' nor y' . Then, we construct the graph G' by first removing x and y , and then adding the edge vx' . If $y' \neq x'$, then we additionally add the edge vy' . Note that G' is a simple plane graph of maximum degree 9 with fewer vertices and edges than G . Therefore, it admits a nice coloring λ by the minimality of G . We easily modify λ to obtain a nice coloring of G . First, put $\lambda(xx') := \lambda(vy) := \lambda'(vx')$. Now, if $x' \neq y'$, then we put $\lambda(vx) := \lambda(yy') := \lambda'(vy')$. See Figure 1(a) for an illustration. And, if $x' = y'$, then we note that each of the edges yy' and vx has at most 9 forbidden colors. Thus, both of them can be colored and the obtained 10-total-coloring of G is nice.

v is adjacent to x' . Thus $vx x'$ is a triangle. Consider a nice coloring of $G - vy$. To extend it to G , it suffices to properly color the edge vy . If this cannot be done greedily, then $|\mathcal{C}(vy)| = 10$, and up to a permutation of the colors, we can assume that the coloring is the one shown in Figure 1(b). If $a \neq 10$, then recolor vx with 10 and color vy with 5 to obtain a nice coloring of G . And if $a = 10$, then we interchange the colors of vx' and xx' , and afterwards color vy with 4.

(iv) By (ii), a 3-vertex has only (≥ 8)-neighbors. Thus we may suppose that $v w u$ is a $(3, 8, 8)$ -triangle, with u being the 3-vertex. Consider a nice coloring of $G - vu$. To extend it to G , again it suffices to properly color the edge vu . If we cannot do this

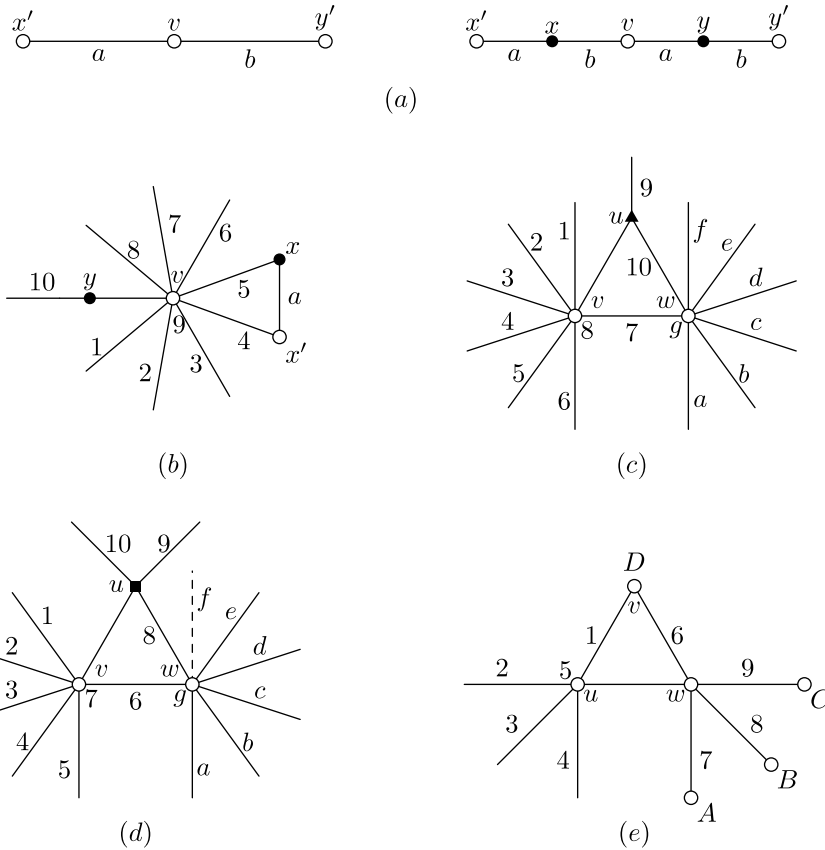


FIG. 1. Configurations for the proof of Lemma 2.

greedily, then it means that $|\mathcal{C}(vu)| = 10$. Thus, up to a permutation of the colors, the coloring is the one shown in Figure 1(c). If the edge wu can be properly recolored, then we do so, and afterwards color the edge vu with 10, which gives a nice coloring of G . So we deduce that $|\mathcal{C}(wu)| = 9$. Consequently, $\{a, b, c, d, e, f, g\} = \{1, 2, 3, 4, 5, 6, 8\}$. Thus we obtain $9 \notin \mathcal{C}(vw)$. So, we can recolor vw with 9 and color vu with 7 to conclude the proof.

(v) By (ii), it is enough to prove that there is no $(4, 7, \delta)$ -triangle in G for $\delta \in \{7, 8\}$. Suppose that vwu is such a triangle with w having degree δ and u degree 4. Consider a nice coloring of $G - vu$. It is sufficient to properly color the edge vu to obtain a nice coloring of G . Again, $|\mathcal{C}(vu)| = 10$, so up to a permutation of the colors, we assume that the coloring is the one of Figure 1(d). If the edge wu can be properly recolored, then do so, and color vu with 8 to obtain a nice coloring of G . Thus, we deduce that $|\mathcal{C}(wu)| = 9$. Therefore, $\{1, 2, 3, 4, 5, 7\} \subset \{a, b, c, d, e, f, g\}$. From this we infer that $|\mathcal{C}(vw)| \leq 6 + \delta - 6 = \delta \leq 8$. Thus, the edge vw can be properly recolored, and so the edge vu can be colored with 6, yielding a nice coloring of G .

(vi) Let vu be a triangle with $\deg(u) = \deg(w) = 5$. Consider a total-coloring of $G - uw$, and uncolor the vertex w . Observe that $|\mathcal{F}(uw)| \geq 1$ and $|\mathcal{F}(w)| \geq 1$. Furthermore, these two sets must actually be equal and of size 1; otherwise we can extend the coloring to G . Up to a permutation of the colors, the coloring is the one shown in Figure 1(e), with $\{A, B, C, D\} = \{1, 2, 3, 4\}$. Notice that the colors of the

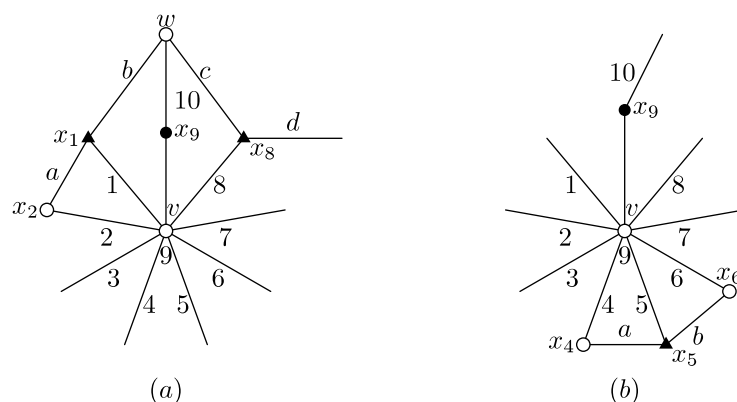


FIG. 2. Reducible configurations of Lemma 3(vi) and (vii).

edges vu and wu can be safely interchanged. Now, the vertex w can be properly colored with 6, and the edge uw with 10. $\square$

LEMMA 3. For the graph G , the following assertions hold.

- (i) There is no $(5, 6, 6)$ -triangle.
- (ii) A 6-vertex has at most two 5-neighbors.
- (iii) Suppose that v is a 7-vertex, and let x_1 be one of its neighbors. If v and x_1 have at least two common neighbors, then at most one of them has degree 4.
- (iv) Suppose that vwu and vwu' are two triangles with $\deg(u) = 2$. Then, $\deg(u') \geq 4$.
- (v) Suppose that v is a 9-vertex incident to a $(2, 9, 9)$ -triangle. Then it is not incident to a $(\leq 3, \geq 8, 9)$ -triangle.
- (vi) The configuration of Figure 2(a) is reducible.
- (vii) The configuration of Figure 2(b) is reducible.

Proof. (i) Suppose on the contrary that G contains a $(5, 6, 6)$ -triangle uvw with u being of degree 5. The proof is in two steps. In the first step, we prove the existence of a 10-total-coloring of G in which only u is uncolored. And in the second step, we establish that such a coloring can be extended to G . Consider a nice coloring of $G - vu$, and uncolor the vertex u . Our only goal in the first step is to properly color the edge vu . If we cannot do this greedily, then $|\mathcal{C}(vu)| = 10$, and thus we can assume that the coloring is the one of Figure 3(a). We infer that $\{6, 7, 8, 9, 10\} = \{a, b, c, d, e\}$; otherwise we can choose a color $\alpha \in \{6, 7, 8, 9, 10\} \setminus \{a, b, c, d, e\}$, recolor uw with α , and color vu with 4. Consequently, we have $\mathcal{C}(vw) = \{4, 6, 7, 8, 9, 10\}$. Thus, we can recolor vw with 1, and color vu with 5.

For the second step, consider a partial 10-total-coloring of G such that only u is not colored. If we cannot greedily extend it to G , then without loss of generality the coloring is the one of Figure 3(b). Note that if $|\mathcal{C}(vu)| \leq 8$, then we can recolor vu , and color u with 5. Thus, we infer that $\{a, b, c, d, e\} \supset \{7, 8, 9, 10\}$. Similarly, $\{e, f, g, h, i\} \supset \{6, 8, 9, 10\}$. Observe that $|\mathcal{C}(v)| = 9$; otherwise we just properly recolor v , and color u with 6.

We assert that we can assume that $e \in \{1, 2, 3\}$. If it is not the case, then $e \in \{8, 9, 10\}$, say $e = 10$. By what precedes, $|\mathcal{C}(vw)| \leq 12 - 4 = 8$ and $\{4, 5, 6, 7, 8, 9\} \subset \mathcal{C}(vw)$. Thus at least one color among 1, 2, 3 can be used to recolor vw , which proves the assertion. Therefore, $\{a, b, c, d\} = \{7, 8, 9, 10\}$ and $\{f, g, h, i\} = \{6, 8, 9, 10\}$. Thus vw can be recolored by every color of $\{1, 2, 3\}$. So, if there exists a color $\alpha \in \{1, 2, 3\} \setminus$

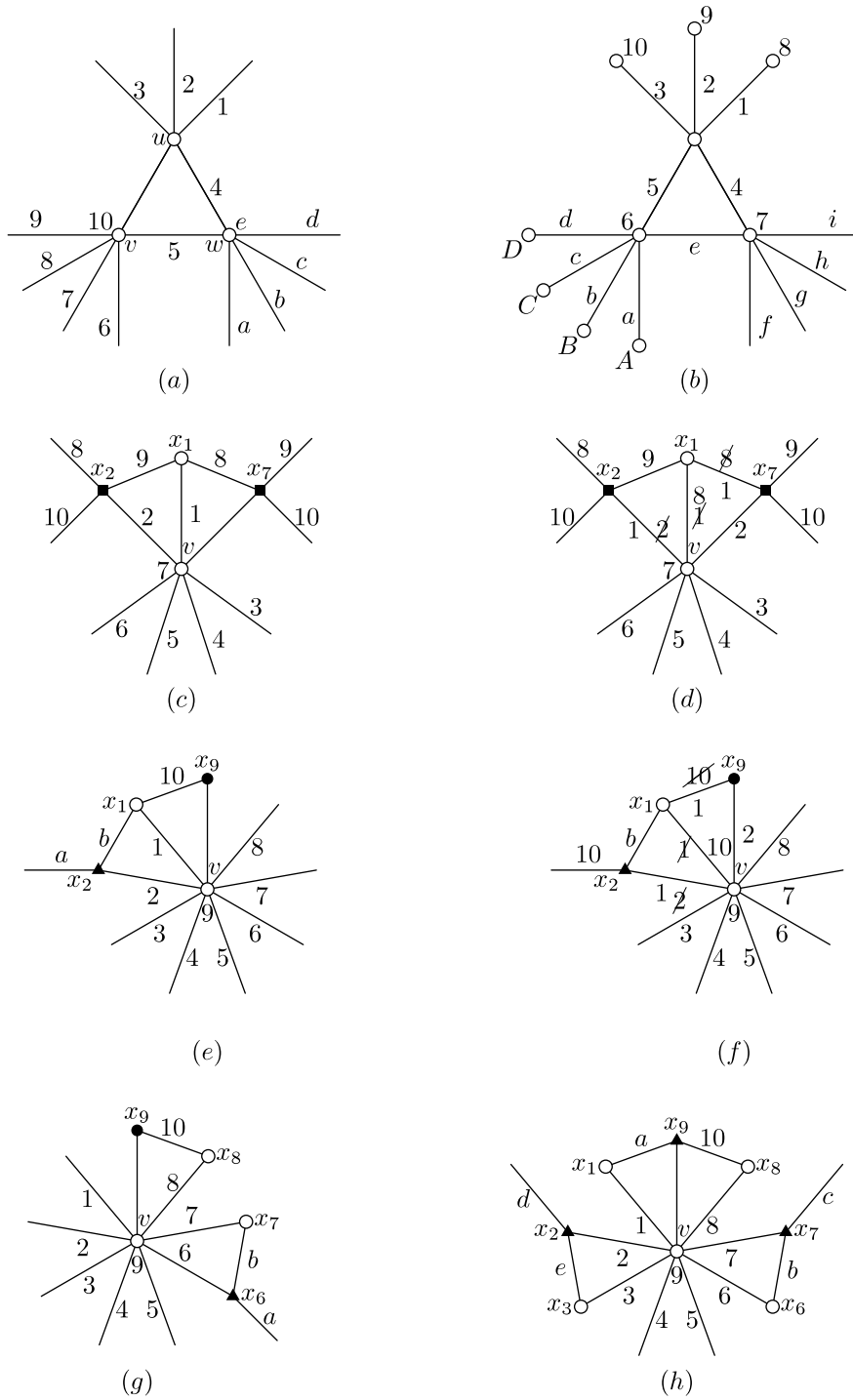


FIG. 3. Configurations for the proofs of Lemmas 3 and 4.

$\{A, B, C, D\}$, we can recolor vw with a color of $\{1, 2, 3\}$ different from α , recolor v with α , and color u with 6. Hence $\{1, 2, 3\} \subseteq \{A, B, C, D\}$. Now, recall that $|\mathcal{C}(v)| = 9$, thus $4 \in \{A, B, C, D\}$. Consequently, we can interchange safely the colors of vu and wu , recolor v with 5, and finally color u with 6.

(ii) Suppose that v is a 6-vertex with three 5-neighbors x_1, x_2, x_3 . By Lemma 2(vi), these three vertices are pairwise nonadjacent. Let λ be a nice coloring of $G - vx_1$, and uncolor the edges vx_2 and vx_3 as well as the vertices v, x_1, x_2 , and x_3 . Notice that $|\mathcal{C}(x_i)| \leq 8$ and $|\mathcal{C}(vx_i)| \leq 7$ for each $i \in \{1, 2, 3\}$. Moreover, $|\mathcal{C}(v)| \leq 6$. Recall that $\mathcal{F}(x) := \{1, 2, \dots, 10\} \setminus \mathcal{C}(x)$ for every $x \in V \cup E$. Observe that for each $i \in \{1, 2, 3\}$, we have $\mathcal{F}(v) \cap \mathcal{F}(x_i) \subseteq \mathcal{F}(vx_i)$. Hence, we infer that $\mathcal{F}(v) \cap (\mathcal{F}(vx_i) \cup \mathcal{F}(x_i)) = \mathcal{F}(v) \cap \mathcal{F}(vx_i)$. Consequently, there exists a color $\alpha \in \mathcal{F}(v)$ such that, after setting $\lambda(v) := \alpha$, it holds that $|\mathcal{F}(x_3)| \geq 2$ and $|\mathcal{F}(vx_3)| \geq 3$. If we color properly x_1, vx_1, x_2 , and vx_2 , then we will be able to color greedily x_3 and vx_3 , and hence the proof would be complete. Observe that if α does not belong to $\mathcal{F}(x_1)$ or to $\mathcal{F}(vx_2)$, then the coloring can be extended greedily to x_1, x_2, vx_1, vx_2 —just color x_1 or vx_2 last, respectively. Therefore we assume that α belongs to these two lists. Uncolor v and color x_1 and vx_2 with α . With respect to this coloring, note that $|\mathcal{F}(vx_1)| \geq 2$, $|\mathcal{F}(v)| \geq 3$, $|\mathcal{F}(x_2)| \geq 1$, $|\mathcal{F}(vx_3)| \geq 3$, and $|\mathcal{F}(x_3)| \geq 2$. Hence, we can color x_2 . Now, if there exists $\beta \in \mathcal{F}(vx_1) \cap \mathcal{F}(x_3)$, then we let $\lambda(vx_1) := \lambda(x_3) := \beta$, and afterwards greedily color v and vx_3 .

So, $\mathcal{F}(vx_1) \cap \mathcal{F}(x_3) = \emptyset$. If there exists $\kappa \in \mathcal{F}(v) \cap \mathcal{F}(x_3)$, then we set $\lambda(v) := \kappa$, and afterwards we greedily color x_3, vx_3 , and vx_1 in this order. This is possible since $\kappa \notin \mathcal{F}(vx_1)$. Otherwise, greedily coloring vx_1, v, vx_3 , and x_3 in this order yields a nice coloring of G .

(iii) Suppose that the statement is false, so the graph G contains the configuration of Figure 3(c). Consider a nice coloring λ of $G - vx_7$. If it cannot be extended to G , then $|\mathcal{C}(vx_7)| = 10$. Furthermore, $|\mathcal{C}(vx_2)| = 9$; otherwise we can color the edge vx_7 with $\lambda(vx_2)$ and greedily recolor the edge vx_2 , thereby obtaining a nice coloring of G . Therefore, we can assume that the coloring is the one shown in Figure 3(c). Then a nice coloring of G is obtained by interchanging the colors of the edges x_7x_1 and vx_1 , recoloring vx_2 with 1, and coloring vx_7 with 2, as shown in Figure 3(d).

(iv) Suppose on the contrary that G contains the configuration of Figure 3(e). Consider a nice coloring of $G - vx_9$. If the edge vx_9 cannot be greedily colored, then $|\mathcal{C}(vx_9)| = 10$. Thus we may assume that the coloring is the one shown in Figure 3(e). Notice that $a = 10$; otherwise we recolor vx_2 with 10 and color vx_9 with 2. So, the recoloring in Figure 3(f) is nice.

(v) Suppose that G contains the configuration of Figure 3(g), and consider a nice coloring λ of $G - vx_9$. Without loss of generality, we may assume that it is the one of Figure 3(g). Observe that $10 \in \{a, b\}$; otherwise we obtain a nice coloring of G by setting $\lambda(vx_6) := 10$ and $\lambda(vx_9) := 6$. Now, we consider two cases regarding b .

$b = 10$. If $a \neq 7$, then we can interchange the colors of the edges x_6x_7 and vx_7 , and color vx_9 with 7 to obtain a nice coloring of G . And if $a = 7$, then we interchange the colors of the edges x_9x_8 and vx_8 , and then we let $\lambda(vx_6) := 8$ and $\lambda(vx_9) := 6$.

$b \neq 10$. In this case, $a = 10$. We interchange the colors of x_9x_8 and vx_8 . Similar to before, we deduce that $b = 8$. Now, the previous case applies with 8 playing the role of color 10.

(vi) Suppose on the contrary that G contains the configuration of Figure 2(a). Up to a permutation of the colors, every nice coloring of $G - vx_9$ is as the one of Figure

2(a). Note that $d = 10$; otherwise recolor vx_8 with 10 and color vx_9 with 8. Similarly, $a = 10$. Now, interchange the colors of the edges x_1x_2 and vx_2 . If $b \neq 2$, then the obtained coloring extends to G by coloring vx_9 with 2. If $b = 2$, then interchange the colors of the edges x_9w and x_1w , thereby obtaining a nice coloring of $G - vx_9$. Since $d = 10 \neq 2$, observe that we can extend it to G as before; i.e., we recolor vx_8 with 2 and color vx_9 with 8.

(vii) Suppose that G contains the configuration of Figure 2(b). Consider a nice coloring of $G - vx_9$. Without loss of generality, we may assume that it is the one of Figure 2(b). Note that $10 \in \{a, b\}$; otherwise recolor vx_5 with 10 and color vx_9 with 5. By symmetry, we can assume that $a = 10$. Interchange the colors of the edges x_5x_4 and vx_4 . If $b \neq 4$, then we have a nice coloring of $G - vx_9$, and we extend it to G by coloring vx_9 with 4. Otherwise, $b = 4$, we interchange the colors of the edges x_5x_6 and vx_6 , and color vx_9 with 6, which yields a nice coloring of G . $\square$

LEMMA 4. *The configuration of Figure 3(h) is reducible.*

Proof. Consider a nice coloring of $G - vx_9$. If it cannot be greedily extended to G , then $|\mathcal{C}(vx_9)| = 10$, and so we can assume that the coloring is the one of Figure 3(h). First, we note that if $a \neq 7$, then $10 \in \{b, c\}$; otherwise, we recolor vx_7 by 10, and color vx_9 with 7. Similarly, if $a \neq 2$, then $10 \in \{d, e\}$. We now split the proof into three cases.

$a \notin \{6, 8\}$. Since a is different from either 2 or 7, we may assume that $a \neq 7$. As mentioned above, we must have $10 \in \{b, c\}$. Moreover, if we interchange the colors of the edges x_9x_8 and vx_8 , then we deduce as before that $8 \in \{b, c\}$, the color 8 playing the role of color 10. Hence $\{b, c\} = \{8, 10\}$. Now, interchange the colors of the edges x_7x_6 and vx_6 , and color vx_9 with 6. If $b = 10$, then the obtained coloring is proper, and if $b = 8$, then we additionally interchange the colors of the edges x_9x_8 and vx_8 to obtain the desired coloring.

$a = 8$. In this case $10 \in \{b, c\}$. By interchanging the colors of the edges x_9x_8 and vx_8 , and also of x_9x_1 and vx_1 , we infer that $1 \in \{b, c\}$. Hence $\{b, c\} = \{1, 10\}$. Similar to the previous case, interchange the colors of x_7x_6 and vx_6 , and afterwards color vx_9 with 6. If $b = 10$, then the obtained coloring of G is proper, and if $b = 1$, then it suffices to additionally interchange the colors of the edges x_9x_1 and vx_1 , and also of x_9x_8 and vx_8 , to obtain a nice coloring of G .

$a = 6$. Then, $10 \in \{d, e\}$. Note that the colors of the edges x_9x_8 and vx_8 can be interchanged safely, because $a \neq 8$. Therefore, as $a \neq 2$, we infer that $8 \in \{d, e\}$, and hence $\{d, e\} = \{8, 10\}$. We interchange now the colors of the edges x_2x_3 and vx_3 , and color vx_9 with 3. If $e = 10$, then the obtained coloring of G is proper. And, if $e = 8$, then it suffices to interchange the colors of the edges x_9x_8 and vx_8 to obtain the desired coloring. $\square$

LEMMA 5. *If uvw is a triangle with an 8-vertex v and a 3-vertex u , then v has no 3-neighbor distinct from u .*

Proof. Suppose that v is an 8-vertex that contradicts the lemma. Let u and w be two 3-neighbors of v , and assume that vuz is a triangle. We consider a nice coloring of $G - vu$. If we cannot extend it to G , then, without loss of generality, we may assume that the coloring is the one shown on Figure 4(a). Observe that $\{a, b\} = \{9, 10\}$; otherwise we obtain the desired coloring by recoloring vw with either 9 or 10, and coloring vu with 2. Now, as depicted in Figure 4(b), we interchange the colors of the edges uz and vz , recolor vw with 1, and color vu with 2 to obtain the sought coloring. $\square$

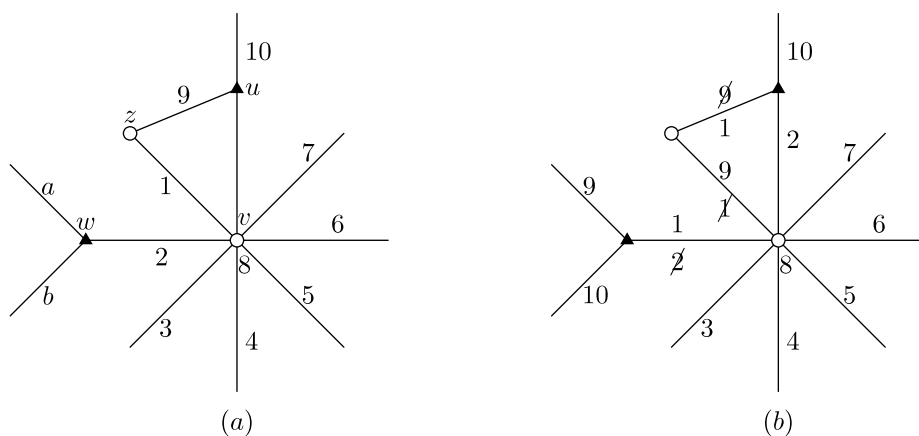


FIG. 4. Coloring and recoloring for the proof of Lemma 5.

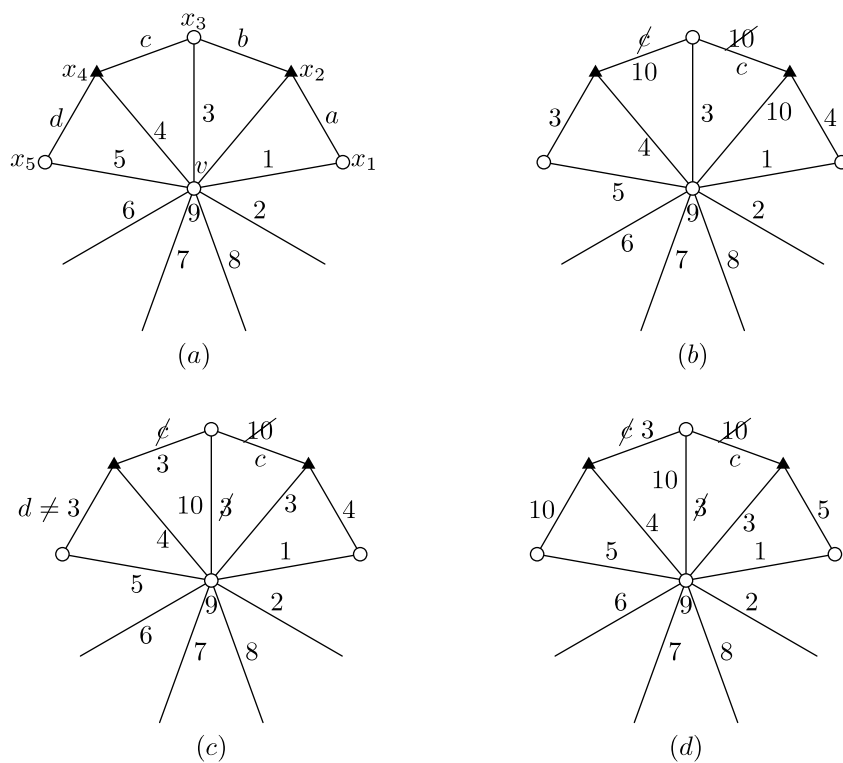


FIG. 5. Configurations for Lemma 6.

LEMMA 6. *The configuration of Figure 5(a) is reducible.*

Proof. Consider a nice coloring of $G - vx_2$. Up to a permutation of the colors, it is the one of Figure 5(a). Note that $10 \in \{a, b\}$; otherwise we obtain a nice coloring of G by coloring vx_2 with 10. We split the proof into two cases, regarding the value of b .

Case 1. $b = 10$. If $a = 4$, then apply the recolorings of Figures 5(b) and (c),

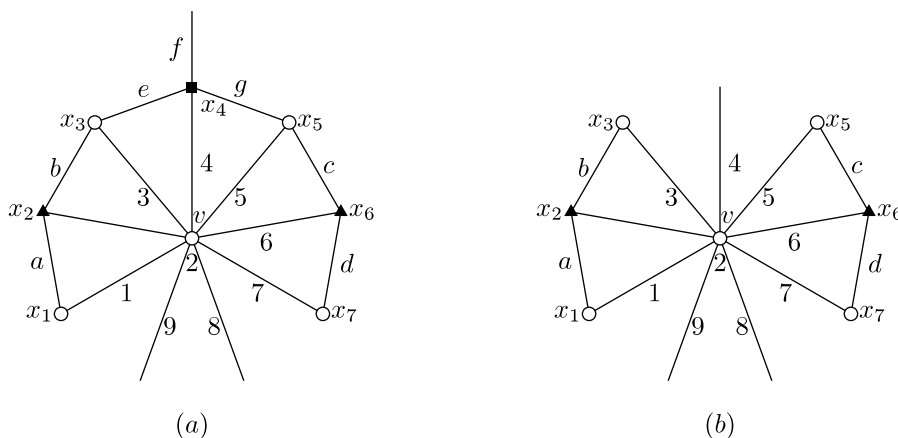


FIG. 6. Configurations for Lemmas 7 and 8.

regarding whether d is 3.

Suppose now that $a \neq 4$. In this case, we deduce that $d = 10$; otherwise, we can recolor vx_4 with 10, and color vx_2 with 4. If $c \neq 5$, then the desired coloring can be obtained as follows. If $a \neq 5$, then interchange the colors of the edges x_4x_5 and vx_5 , and color vx_2 with 5, and if $a = 5$, then the recoloring of Figure 5(d) is nice.

We may assume now that $c = 5$. Interchange the colors of the edges x_4x_5 and vx_5 , and also of the edges x_4x_3 and vx_3 . If $a \neq 3$, then it suffices to color vx_2 with 3. And, if $a = 3$, then additionally interchange the colors of the edges x_2x_1 and vx_1 , recolor vx_4 with 1, and color vx_2 with 4 to obtain the sought coloring.

Case 2. $b \neq 10$. Therefore, $a = 10$. First, note that $10 \in \{c, d\}$; otherwise, we recolor vx_4 with 10, and color vx_2 with 4. Either the obtained coloring of G is nice, or $b = 4$. In the latter case, we additionally interchange the colors of x_2x_3 and x_4x_3 to obtain the desired coloring.

Suppose now that $c = 10$. Then, $b = 4$; otherwise, we uncolor vx_4 , color vx_2 with 4, and apply Case 1 to the obtained coloring with x_4 playing the role of the vertex x_2 . Now, interchange the colors of x_4x_3 and vx_3 . The obtained coloring is nice if $d \neq 3$, and we extend it to G by coloring vx_2 with 3. And, if $d = 3$, then we additionally interchange the colors of x_4x_5 and vx_5 , and color vx_2 with 5.

Finally, assume that $c \neq 10$, and hence $d = 10$. Up to interchanging the colors of x_2x_3 and x_4x_3 , we may assume that $b \neq 5$. Interchange the colors of x_4x_5 and vx_5 . If $c \neq 5$, the obtained coloring is nice, and we extend it to G by coloring vx_2 with 5. And, if $c = 5$, then we additionally interchange the colors of x_4x_3 and vx_3 , and color vx_2 with 3. $\square$

LEMMA 7. *The configuration of Figure 6(a) is reducible.*

Our proof of Lemma 7 uses the following result. Given a coloring λ and a vertex v , recall that $\mathcal{E}(v)$ is the set of colors assigned to the edges incident to v . Let $\mathcal{E}'(v) := \{1, 2, \dots, 10\} \setminus (\mathcal{E}(v) \cup \{\lambda(v)\})$.

LEMMA 8. *Suppose that G contains the configuration of Figure 6(b). Then, for every nice coloring λ of $G - vx_2$, it holds that $\mathcal{E}'(v) \cup \{\lambda(vx_6)\} \subseteq \mathcal{E}(x_2)$.*

Proof. Up to a permutation of the colors, the coloring λ is the one of Figure 6(b). Notice that $\mathcal{E}'(v) = \{10\}$, $\lambda(vx_6) = 6$, and $\mathcal{E}(x_2) = \{a, b\}$. First, $10 \in \{a, b\}$; otherwise, we just color vx_2 with 10. By symmetry, we may assume that $a = 10$. Thus, to finish the proof, it remains only to prove that $b = 6$. Suppose on the

contrary that $b \neq 6$. Note that $10 \in \{c, d\}$; otherwise, we recolor vx_6 with 10 and color vx_2 with 6. By symmetry, we may assume that $d = 10$. We consider two possibilities regarding the value of b .

$b = 1$. Interchange the colors of the edges x_6x_7 and vx_7 . The obtained coloring of G is nice if $c \neq 7$, and if $c = 7$ we additionally interchange the colors of x_6x_5 and vx_5 . Now, coloring vx_2 with 7 or 5 yields a nice coloring of G , a contradiction.

$b \neq 1$. In this case, $c = 1$. Indeed, if $c \neq 1$, then we recolor vx_6 with 1, interchange the colors of x_2x_1 and vx_1 , and color vx_2 with 6 to obtain a nice coloring of G . Now, if $b \neq 7$, then interchange the colors of x_6x_7 and vx_7 and color vx_2 with 7. And, if $b = 7$, then interchange the colors of x_6x_5 and vx_5 , and also of x_2x_1 and vx_1 , and color vx_2 with 5. $\square$

Proof of Lemma 7. Consider a nice coloring λ of $G - vx_2$. Up to a permutation of the colors, we assume that the coloring is the one of Figure 6(a). By Lemma 8, we have $\{a, b\} = \{6, 10\}$. We consider two cases.

$a = 10$ and $b = 6$. If there exists a color $\alpha \in \{1, 10\} \setminus \{e, f, g\}$, then recolor vx_4 with α , and color vx_2 with 4. The obtained coloring is nice if $\alpha = 10$. And, if $\alpha = 1$, then it suffices to additionally interchange the colors of x_2x_1 and vx_1 . Thus, $\{1, 10\} \subset \{e, f, g\}$.

Suppose that $6 \notin \{e, f, g\}$. We start by interchanging the colors of the edges x_2x_3 and x_4x_3 . If $e = 10$, then we additionally interchange the colors of x_2x_1 and vx_1 . Observe that the obtained coloring does not fulfill the conclusion of Lemma 8, a contradiction. Hence, $\{e, f, g\} = \{1, 6, 10\}$ and so $e \in \{1, 10\}$. We interchange the colors of x_4x_3 and vx_3 , and color vx_2 with 3. Either this coloring of G is nice, or $e = 1$ and hence additionally interchanging the colors of x_2x_1 and vx_1 yields a nice coloring of G .

$a = 6$ and $b = 10$. If there exists $\alpha \in \{3, 10\} \setminus \{f, g\}$, then recolor vx_4 with α , and color vx_2 with 4. If the obtained coloring is not nice, then $\alpha = 3$ and hence interchanging the colors of x_2x_3 and vx_3 yields a nice coloring of G , a contradiction. Observe that we may assume that $f = 3$ and $g = 10$. Indeed, if it is not the case, then we interchange the colors of x_2x_3 and vx_3 and obtain the desired condition, with 3 playing the role of color 10.

Furthermore, $e = 5$; otherwise, we interchange the colors of x_4x_5 and vx_5 , and color vx_2 with 5. Now, observe that $d = 10$; otherwise, we recolor vx_6 with 10, vx_4 with 6, and color vx_2 with 4 to obtain a nice coloring of G . Finally, we interchange the colors of x_6x_7 and vx_7 . If $c = 7$, then we additionally interchange the colors of x_6x_5 and vx_5 . Now, coloring vx_2 with 7 or 5 yields a nice coloring of G , a contradiction. $\square$

LEMMA 9. *The configurations of Figure 7 are reducible.*

Proof. Consider a nice coloring of $G - vu$. We may assume that the coloring is the one of Figure 7. Let $\alpha \in \{1, 7, 9, 10\} \setminus \{a, b, c\}$. We recolor vx_3 with α and color vu with 3. The obtained coloring of G is nice unless $\alpha \in \{1, 7\}$. If $\alpha = 1$, then we additionally interchange the colors of uw and vw . And if $\alpha = 7$, then we interchange the colors of ut and vt . $\square$

LEMMA 10. *A 6-vertex incident to 6 triangles is not adjacent to two 5-vertices.*

Proof. Suppose that v is a 6-vertex. We let $x_1, x_2, \dots, x_6$ be its neighbors, such that x_i is adjacent to x_{i+1} if $i \in \{1, 2, \dots, 5\}$ and x_6 is adjacent to x_1 . We also assume that x_6 is a 5-vertex, and we let w be the other 5-vertex. By symmetry and Lemma 2(vi), we may assume that $w \in \{x_2, x_3\}$. The proof is in two steps. In the

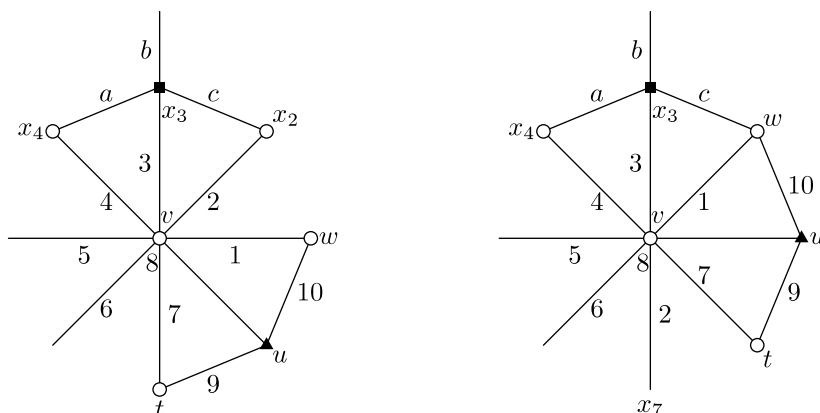


FIG. 7. Reducible configurations of Lemma 9. We assume that the degree of v in G is 8.

first step, we show that there exists a partial 10-total-coloring of G in which only x_6 is uncolored. In the second step, we show how to extend it to a 10-total-coloring of G .

Given a total-coloring and an element $x \in V \cup E$, recall that $\mathcal{C}(x)$ is the set of colors of all the elements of $V \cup E$ incident or adjacent to x . Recall also that if $x \in V$, then $\mathcal{E}(x)$ is the set of colors of all the edges incident to x .

Let λ be a total-coloring of $G - vx_6$, in which, furthermore, we uncolor the vertex x_6 . Our goal is to properly color the edge vx_6 . Note that $|\mathcal{C}(vx_6)| = 10$; otherwise, the edge vx_6 can be greedily colored. Without loss of generality, we may assume that the coloring is the one shown in Figure 8(a).

We want to color vx_6 with $\lambda(vw)$. Recall that w is either x_2 or x_3 . We set $\mathcal{E} := \mathcal{E}(w) \cup \{\lambda(w)\}$. If there exists a color $\alpha \in \{7, 8, 9, 10\} \setminus \mathcal{E}$, then we set $\lambda(vx_6) := \lambda(vw)$ and $\lambda(vw) := \alpha$. Furthermore, if $1 \notin \mathcal{E}$, then we interchange the colors of x_6x_1 and vx_1 , color vx_6 with $\lambda(vw)$, and recolor vw with 1. Thus, $1 \in \mathcal{E}$. Similarly, we deduce that $5 \in \mathcal{E}$. Finally, note that either 2 or 3 belongs to $\mathcal{E}$, according to whether w is x_2 or x_3 . Consequently, this shows that $|\mathcal{E}| \geq 7$. But w has degree 5, thus $|\mathcal{E}| = 6$, a contradiction. This concludes the first step.

Suppose now that we are given a partial 10-total-coloring of G in which only x_6 is not colored. If we cannot extend it to G , then, without loss of generality, we may assume that the coloring is the one shown in Figure 8(b). If there exists a color $\alpha \in \{2, 4, 6, 10\} \setminus \{a, b, c, d, e\}$, then recolor vx_6 with α and color x_6 with 7 to obtain a 10-total-coloring of G . Hence, $\{2, 4, 6, 10\} \subset \{a, b, c, d, e\}$. Suppose that $a \notin \{2, 4, 6\}$. In this case, $\{b, c, d, e\} = \{2, 4, 6, 10\}$, and thus $e \in \{2, 4, 10\}$. Interchange the colors of the edges x_6x_5 and vx_5 . Now, if $a \neq 5$, then the obtained coloring is proper, and we extend it to G by coloring x_6 with 5. And, if $a = 5$, then we additionally interchange the colors of x_6x_1 and vx_1 , and color v with 9. Consequently, we obtain $a \in \{2, 4, 6\}$.

If $9 \notin \{b, c, d, e\}$, then we can apply a similar recoloring. More precisely, we can interchange the colors of the edges x_6x_1 and vx_1 . The obtained coloring is proper and can be extended to G by coloring x_6 with 9. So $9 \in \{b, c, d, e\}$, and hence $5 \notin \mathcal{E}(v)$. We interchange the colors of x_6x_5 and vx_5 , and color x_6 with 5. Either the obtained 10-total-coloring of G is proper, or $e = 9$. In the latter case, we additionally interchange the colors of x_6x_1 and vx_1 to obtain the sought contradiction. $\square$

LEMMA 11. *The configuration of Figure 9(a) is reducible.*

Proof. Consider a nice coloring of $G - vx_9$. Without loss of generality, it is the one

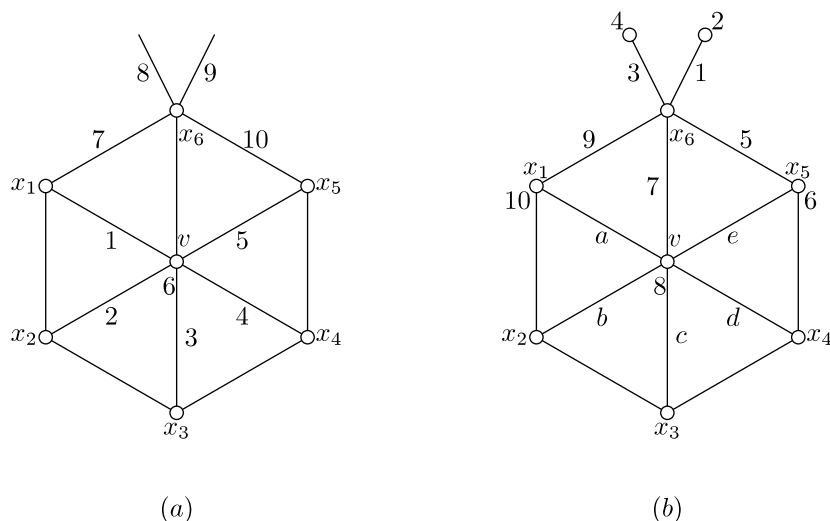


FIG. 8. Proof of Lemma 10: (a) coloring of $G - vx_6$, (b) partial coloring of G in which x_6 is not colored.

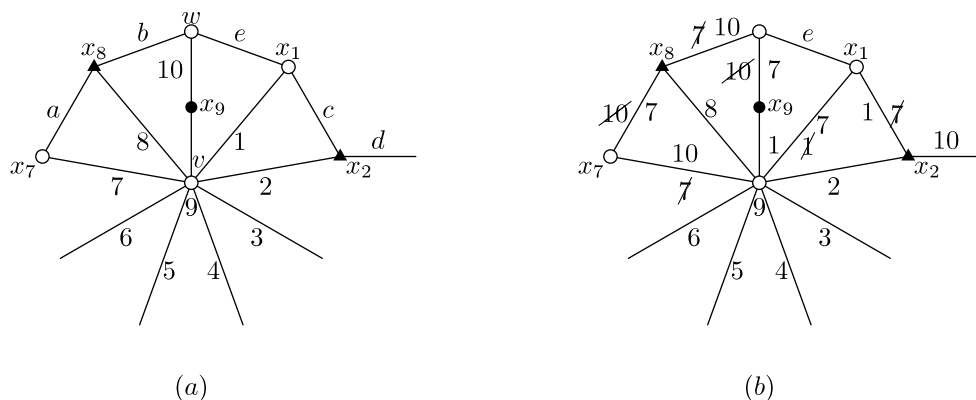


FIG. 9. Precoloring and recoloring for the proof of Lemma 11.

of Figure 9(a). First, note that $a = 10$; otherwise, we can recolor the edge vx_8 with 10, and color vx_9 with 8. Next, we infer that $b = 7$; otherwise, we can interchange the colors of x_8x_7 and vx_7 , and color vx_9 with 7. Now, observe that $10 \in \{c, d\}$; otherwise, we recolor vx_2 with 10, and color vx_9 with 2. Furthermore, $7 \in \{c, d\}$; otherwise, we interchange the colors of x_8w and x_9w , and also of x_8x_7 and vx_7 , recolor vx_2 with 7, and color vx_9 with 10. Thus, $\{c, d\} = \{7, 10\}$. If $d = 7$ and $c = 10$, then we just interchange the colors of the edges x_2x_1 and vx_1 , and color vx_9 with 1. And, if $d = 10$ and $c = 7$, then the recoloring shown in Figure 9(b) is a nice coloring of G . $\square$

3. Discharging part. Recall that $G = (V, E)$ is a minimum counterexample to the statement of Theorem 1, in the sense that $|V| + |E|$ is minimum. We obtain a contradiction by using the discharging method. Here is an overview of the proof. We fix a planar embedding of G . Each vertex and face of G is assigned an initial charge. The total sum of the charges is negative by Euler's formula. Then, some redistribution rules are applied, and vertices and faces send or receive some charge

according to these rules. The total sum of the charges is not changed during this step, but at the end we infer that the charge of each vertex and face is nonnegative, a contradiction.

Initial charge. We assign a charge to each vertex and face. For every $x \in V \cup F$, we define the initial charge $\text{ch}(x)$ to be $\deg(x) - 4$, where $\deg(x)$ is the degree of x in G . By Euler's formula the total sum is

$$\sum_{v \in V} \text{ch}(v) + \sum_{f \in F} \text{ch}(f) = -8.$$

Rules. We need the following definitions to state the discharging rules. A 2-vertex is *bad* if it is not incident to a (≥ 5) -face. A triangle is *bad* if it contains a vertex of degree at most 4. Recall that a triangle with vertices x, y , and z is a $(\deg(x), \deg(y), \deg(z))$ -triangle.

Rule R0. A (≥ 5) -face sends 1 to each incident 2-vertex.

Rule R1. A 5-vertex v sends $1/5$ to each incident triangle.

Rule R2. A 6-vertex sends $13/35$ to each incident $(5, 6, \geq 7)$ -triangle, $1/3$ to each incident $(6, 6, 6)$ -triangle, and $2/7$ to each incident $(6, \geq 6, \geq 7)$ -triangle.

Rule R3. A 7-vertex sends $1/2$ to each incident bad triangle, $3/7$ to each incident nonbad $(\leq 7, \leq 7, 7)$ -triangle, and $1/3$ to each incident nonbad triangle containing a (≥ 8) -vertex.

Rule R4. A 8-vertex sends

- (i) $1/3$ to each adjacent 3-vertex,
- (ii) $1/2$ to each incident bad triangle,
- (iii) $7/15$ to each incident $(5, \leq 7, 8)$ -triangle and each incident $(6, 6, 8)$ -triangle,
- (iv) $2/5$ to each incident $(5, \geq 8, 8)$ -triangle, each incident $(6, 7, 8)$ -triangle, and each incident $(6, 8, 8)$ -triangle,
- (v) $1/3$ to each incident $(6, 8, 9)$ -triangle and each incident $(\geq 7, \geq 7, 8)$ -triangle.

Rule R5. A 9-vertex sends

- (i) 1 to each adjacent bad 2-vertex and $1/2$ to each adjacent nonbad 2-vertex,
- (ii) $1/3$ to each adjacent 3-vertex,
- (iii) $1/2$ to each incident bad triangle and each incident $(5, \leq 7, 9)$ -triangle,
- (iv) $3/7$ to each incident $(6, 6, 9)$ -triangle,
- (v) $2/5$ to each incident $(5, \geq 8, 9)$ -triangle and each incident $(6, \geq 7, 9)$ -triangle,
- (vi) $1/3$ to each incident $(\geq 7, \geq 7, 9)$ -triangle.

In what follows, we prove that the final charge $\text{ch}^*(x)$ of every $x \in V \cup F$ is nonnegative. Hence, we obtain

$$-8 = \sum_{v \in V} \text{ch}(v) + \sum_{f \in F} \text{ch}(f) = \sum_{v \in V} \text{ch}^*(v) + \sum_{f \in F} \text{ch}^*(f) \geq 0,$$

a contradiction. This contradiction establishes the theorem.

Final charge of faces. Let f be a d -face. Our goal is to show that $\text{ch}^*(f) \geq 0$. By Lemma 2(ii) and (iii), f is incident to at most $\lfloor \frac{d}{3} \rfloor$ vertices of degree 2. Therefore, if $d \geq 5$, then by Rule R0 we obtain $\text{ch}^*(f) \geq d - 4 - \lfloor \frac{d}{3} \rfloor = \lceil \frac{2d}{3} \rceil - 4 \geq 0$. A 4-face neither sends nor receives any charge, so its charge stays 0.

Finally, let $f = xyz$ be a triangle with $\deg(x) \leq \deg(y) \leq \deg(z)$. The initial charge of f is -1 , and we assert that its final charge $\text{ch}^*(f)$ is at least 0. We consider several cases and subcases according to the degrees of x, y , and z .

$\deg(x) = 2$. Then both y and z have degree 9 by Lemma 2(ii), and hence f receives $1/2$ from each of y and z by Rule R5(iii).

- $\deg(x) = 3$. In this case, by Lemma 2(ii) and (iv), we infer that $\deg(y) \geq 8$ and $\deg(z) = 9$. Thus, f receives $\frac{1}{2} + \frac{1}{2} = 1$ by Rules R4(ii) and R5(iii).
- $\deg(x) = 4$. Then, by Lemma 2(ii) and (v), $\deg(y) \geq 7$ and $\deg(z) \geq 8$. Hence, by Rules R3, R4(ii), and R5(iii), f receives $\frac{1}{2} + \frac{1}{2} = 1$ from y and z .
- $\deg(x) = 5$. According to Lemma 2(vi), $\deg(y) \geq 6$, and by Lemma 3(i), $\deg(z) \geq 7$. By Rule R1, f receives $1/5$ from x , so we need only to show that it receives at least $4/5$ from y and z together. Consider the following subcases.
- $\deg(z) = 7$. By Rule R3, z sends $3/7$ to f , and by Rules R2 and R3, y sends at least $13/35$. Thus, f receives at least $\frac{13}{35} + \frac{3}{7} = \frac{4}{5}$ from y and z , as needed.
- $\deg(z) = 8$. If $\deg(y) \leq 7$, then z sends $7/15$ to f by Rule R4(iii), and y sends at least $1/3$ by Rules R2 and R3. And, if $\deg(y) = 8$, then both y and z send $2/5$ to f by Rule R4(iv). So, in both cases f receives $4/5$ from y and z together.
- $\deg(z) = 9$. Suppose first that $\deg(y) \leq 7$. Then, by Rule R5(iii), z sends $1/2$ to f . Moreover, by Rules R2 and R3, y sends at least $1/3$ to f , which proves the assertion. Now, if $\deg(y) \geq 8$, then according to Rules R4(iv) and R5(v) f receives $2/5$ from each of y and z , as needed.
- $\deg(x) = 6$. First, if $\deg(z) = 6$, then f receives $1/3$ from each of its vertices by Rule R2. So we assume that $\deg(z) \geq 7$. In this case, f receives $2/7$ from x by Rule R2. Hence, we need only to show that y and z send at least $5/7$ to f in total. We consider several cases, regarding the degree of z .
- $\deg(z) = 7$. Then f receives $3/7$ from z by Rule R3, and at least $2/7$ from y by Rules R2 and R3, as desired.
- $\deg(z) = 8$. If $\deg(y) = 6$, then z sends $7/15$ by Rule R4(iii) and y sends $2/7$ by Rule R2. And, if $\deg(y) \geq 7$, then y sends at least $1/3$ by Rules R3 and R4(iv), and z sends at least $2/5$ by Rule R4(iv).
- $\deg(z) = 9$. If $\deg(y) = 6$, then f receives $2/7$ from y by Rule R2 and $3/7$ from z by Rule R5(iv). And, if $\deg(y) \geq 7$, then f receives at least $1/3$ from y by Rules R3, R4(v), and R5(v), and at least $2/5$ from z by Rule R5(v), which yields the result.
- $\deg(x) \geq 7$. The assertion follows from Rules R3, R4(v), and R5(vi).

Final charge of vertices. Let v be an arbitrary vertex of G . We have $\deg(v) \geq 2$ by Lemma 2(i). For every positive integer d , we define v_d to be the number of d -neighbors of v , and f_d to be the number of d -faces incident to v . Let $x_1, x_2, \dots, x_{\deg(v)}$ be the neighbors of v in clockwise order. We prove that the final charge of v is nonnegative. To do so, we consider several cases, regarding the degree of v .

If $\deg(v) = 2$, then its two neighbors are 9-vertices by Lemma 2(ii). If v is bad, then it receives 1 from each of its two 9-neighbors by Rule R5(i), while otherwise it receives at least 1 from its incident faces by Rule R0, and $1/2$ from each of its two 9-neighbors by Rule R5(i). Thus, in both cases its final charge is at least 0.

If $\deg(v) = 3$, then all of its neighbors have degree at least 8, so by Rules R4(i) and R5(ii) it receives $1/3$ from each of its neighbors, setting its final charge to 0. If $\deg(v) = 4$, then it neither sends nor receives anything, so its charge stays 0. If v is a 5-vertex, then by Rule R1 it sends $1/5$ to each of its at most five incident triangles; therefore, its final charge is nonnegative.

Suppose now that v is a 6-vertex. All of its neighbors have degree at least 5 by Lemma 2(ii). Note that if $f_3 \leq 5$, then, according to Rule R2, $\text{ch}^*(v) \geq 2 - 5 \cdot \frac{13}{35} > 0$. So, we assume now that $f_3 = 6$; i.e., v is incident to 6 triangles. Thus, we infer

from Lemma 10 that $v_5 \leq 1$. If $v_5 = 0$, then following Rule R2, the vertex v sends at most $6 \cdot \frac{1}{3} = 2$; therefore, its final charge is at least 0. And, if $v_5 = 1$, then let x_1 be the unique 5-neighbor of v . By Lemma 3(i), it holds that $\deg(x_2) \geq 7$ and $\deg(x_6) \geq 7$. Consequently, vx_3x_2 and vx_5x_6 are two $(6, \geq 6, \geq 7)$ -triangles. Thus, $\text{ch}^*(v) \geq 2 - 2 \cdot \frac{13}{35} - 2 \cdot \frac{1}{3} - 2 \cdot \frac{2}{7} = \frac{2}{105} > 0$ by Rule R2.

Suppose that v is a 7-vertex. If $f_3 \leq 6$, then $\text{ch}^*(v) \geq 3 - 6 \cdot \frac{1}{2} = 0$ by Rule R3. So, we assume now that $f_3 = 7$. We consider several cases, according to the number of 4-neighbors of v . Note that, by Lemmas 2(ii) and 3(iii), the vertex v has at most two such neighbors; i.e., $v_4 \leq 2$.

$v_4 = 0$. According to Rule R3, we have $\text{ch}^*(v) \geq 3 - 7 \cdot \frac{3}{7} = 0$.

$v_4 = 1$. Let x_1 be this 4-neighbor. So, x_2 and x_7 both are 9-vertices by Lemma 2(v).

According to Rule R3, the vertex v sends $1/3$ to each of vx_2x_3 and vx_5x_6 .

Furthermore, v is incident to exactly two bad triangles and sends at most $3/7$ to each nonbad triangle. Therefore, we obtain $\text{ch}^*(v) \geq 3 - 2 \cdot \frac{1}{2} - 3 \cdot \frac{3}{7} - 2 \cdot \frac{1}{3} = \frac{1}{21} > 0$.

$v_4 = 2$. Without loss of generality, we assume that x_1 has degree 4. According to Lemmas 2(ii) and 3(iii), the other 4-neighbor of v must be x_4 or x_5 , say x_4 by symmetry. By Lemma 2(v), the vertices x_2, x_7, x_3 , and x_5 all have degree 9.

Note that x_6 has degree at least 5. Consequently, $\text{ch}^*(v) = 3 - 4 \cdot \frac{1}{2} - 3 \cdot \frac{1}{3} = 0$.

Suppose now that v is an 8-vertex. If $v_3 = 0$, then $\text{ch}^*(v) \geq 4 - 8 \cdot \frac{1}{2} = 0$ by Rule R4. Thus, we assume now that x_1 is a 3-vertex. Notice that, by Lemma 5, if a 3-neighbor of v is on a triangle, then $v_3 = 1$. Therefore, $v_3 + f_3 \leq 9$. If $f_3 \leq 6$, then we obtain $\text{ch}^*(v) \geq 4 - 6 \cdot \frac{1}{2} - 3 \cdot \frac{1}{3} = 0$. If $f_3 = 7$, then we infer from Lemma 5 that $v_3 \leq 1$, and so $\text{ch}^*(v) \geq 4 - 7 \cdot \frac{1}{2} - \frac{1}{3} = \frac{1}{6} > 0$. Now we suppose that $f_3 = 8$, and thus $v_3 = 1$. According to Lemma 2(iv), it holds that $\deg(x_2) = 9$ and $\deg(x_9) = 9$. Moreover, by Lemma 9, all the vertices but x_1 have degree at least 5. Thus, by Rule R4, we infer that $\text{ch}^*(v) \geq 4 - \frac{1}{3} - 2 \cdot \frac{1}{2} - 4 \cdot \frac{7}{15} - 2 \cdot \frac{2}{5} = 0$.

Finally, suppose that v is a 9-vertex. By Lemma 2(iii), v is adjacent to at most one 2-vertex. We consider two cases.

Case 1. $v_2 = 0$. Suppose first that v is incident to a (≥ 4) -face; i.e., $f_3 \leq 8$. If $f_3 = 8$, then $v_3 \leq 3$ by Lemmas 4 and 6, and hence $\text{ch}^*(v) \geq 5 - 8 \cdot \frac{1}{2} - 3 \cdot \frac{1}{3} = 0$. If $f_3 \leq 7$, then we assert that $f_3 + v_3 \leq 12$. Indeed, if $v_3 \geq 6$, then, as two 3-vertices are not adjacent, we infer that $f_3 \leq 2(9 - v_3)$, which yields the assertion. So if $f_3 \leq 6$, then we obtain $\text{ch}^*(v) \geq 5 - 6 \cdot \frac{1}{2} - 6 \cdot \frac{1}{3} = 0$. If $f_3 = 7$, then we can see that $v_3 \leq 4$ by Lemmas 4 and 6. Consequently, $\text{ch}^*(v) \geq 5 - 7 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} = \frac{1}{6} > 0$. Now assume that $f_3 = 9$. Then, $v_3 \leq 2$ according to Lemma 4. If $v_3 \leq 1$, then $\text{ch}^*(v) \geq 5 - 9 \cdot \frac{1}{2} - \frac{1}{3} = \frac{1}{6} > 0$. Assume now that $v_3 = 2$. Without loss of generality, say that x_1 is a 3-vertex; thus both x_2 and x_9 are (≥ 8) -vertices. Note that by Lemma 6, both x_3 and x_8 are (≥ 4) -vertices. Therefore, up to symmetry, it suffices to consider the following two cases.

x_4 is the second 3-neighbor. Then $\deg(x_3) \geq 8$, so by Rule R5(vi) the vertex v sends $\frac{1}{3}$ to vx_2x_3 . Hence, we infer that $\text{ch}^*(v) \geq 5 - 2 \cdot \frac{1}{3} - 8 \cdot \frac{1}{2} - \frac{1}{3} = 0$.

x_5 is the second 3-neighbor. In this case, $\deg(x_4) \geq 8$ and $\deg(x_6) \geq 8$. Furthermore, $\deg(x_3) \geq 5$ by Lemma 7. Consequently, x_3x_2v and x_3x_4v are both $(\geq 5, \geq 8, 9)$ -triangles. Hence, v sends at most $\frac{2}{5}$ to each of them by Rule R5(v). So, $\text{ch}^*(v) \geq 5 - 2 \cdot \frac{1}{3} - 7 \cdot \frac{1}{2} - 2 \cdot \frac{2}{5} = \frac{1}{30} > 0$.

Case 2. $v_2 = 1$. Let x_1 be the 2-neighbor. Observe that by Lemma 3(vii), v cannot have a 3-neighbor on two triangles. Moreover, x_1 cannot be incident to two triangles, so $f_3 \leq 8$. We consider the following possibilities.

x_1 is on a triangle. Let this triangle be vx_1x_2 . From Lemma 3(iv) and (v), we infer that $f_3 + v_3 \leq 8$. So, $\text{ch}^*(v) \geq 5 - 1 - 8 \cdot \frac{1}{2} = 0$.

x_1 is bad but not on a triangle. In this case, x_1 is on two 4-faces; therefore, in particular, $f_3 \leq 7$. Note that by Lemma 3(vi), either one vertex among x_2, x_9 has degree at least 4, or $f_3 \leq 5$. Besides, according to Lemma 3(vii) there is no 3-neighbor on two triangles. Observe also that if both vx_2x_3 and vx_8x_9 are triangles, then Lemma 11 implies that $v_3 \leq 6$. Let us consider several cases regarding the value of f_3 .

$f_3 \leq 4$. Then $f_3 + v_3 \leq 10$; otherwise, we obtain a contradiction by Lemma 3(vi) and (vii). Thus, $\text{ch}^*(v) \geq 5 - 1 - 4 \cdot \frac{1}{2} - 6 \cdot \frac{1}{3} = 0$.

$f_3 = 5$. Using Lemma 3(vi) and (vii), a small case-analysis shows that $v_3 \leq 5$. Moreover, if $v_3 = 5$, then the obtained configuration is the one of Lemma 11, which is reducible. And, if $v_3 \leq 4$, then we obtain $\text{ch}^*(v) \geq 5 - 1 - 5 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} = \frac{1}{6} > 0$.

$f_3 = 6$. In this case, $v_3 \leq 3$ by Lemma 3(vi) and (vii). Thus, $\text{ch}^*(v) \geq 5 - 1 - 6 \cdot \frac{1}{2} - 3 \cdot \frac{1}{3} = 0$.

$f_3 = 7$. By Lemma 3(vi) and (vii), v has at most one 3-neighbor, namely x_2 or x_9 . Thus, $\text{ch}^*(v) \geq 5 - 1 - 7 \cdot \frac{1}{2} - \frac{1}{3} = \frac{1}{6} > 0$.

x_1 is neither bad nor on a triangle. Notice that $f_3 \leq 7$. Let us again consider several cases regarding the value of f_3 .

$f_3 \leq 5$. In this case, $f_3 + v_3 \leq 11$ by Lemma 3(vii). So, $\text{ch}^*(v) \geq 5 - \frac{1}{2} - 5 \cdot \frac{1}{2} - 6 \cdot \frac{1}{3} = 0$.

$f_3 = 6$. Similarly as before, we infer that $v_3 \leq 4$, and hence $\text{ch}^*(v) \geq 5 - \frac{1}{2} - 6 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} = \frac{1}{6} > 0$.

$f_3 = 7$. By Lemma 3(vii) the vertex v has at most two 3-neighbors, namely x_2 and x_9 . Thus, $\text{ch}^*(v) \geq 5 - \frac{1}{2} - 7 \cdot \frac{1}{2} - 2 \cdot \frac{1}{3} = \frac{1}{3} > 0$.

This establishes that the final charge of every vertex is nonnegative; therefore, the proof of Theorem 1 is now complete. $\square$

Acknowledgment. The third author thanks the Department of Applied Mathematics of Charles University for hospitality during his stay in Prague in December 2006, which was supported by Research Project MSM 0021620838, when part of this research was conducted.

REFERENCES

- [1] K. APPEL AND W. HAKEN, *Every planar map is four colorable. I. Discharging*, Illinois J. Math., 21 (1977), pp. 429–490.
- [2] K. APPEL, W. HAKEN, AND J. KOCH, *Every planar map is four colorable. II. Reducibility*, Illinois J. Math., 21 (1977), pp. 491–567.
- [3] M. BEHZAD, *Graphs and Their Chromatic Numbers*, Doctoral thesis, Michigan State University, East Lansing, MI, 1965.
- [4] J.-C. BERMOND, *Nombre chromatique total du graphe r -parti complet*, J. London Math. Soc. (2), 9 (1974/75), pp. 279–285.
- [5] O. V. BORODIN, *On the total coloring of planar graphs*, J. Reine Angew. Math., 394 (1989), pp. 180–185.
- [6] O. V. BORODIN, A. V. KOSTOCHKA, AND D. R. WOODALL, *List edge and list total colorings of multigraphs*, J. Combin. Theory Ser. B, 71 (1997), pp. 184–204.
- [7] O. V. BORODIN, A. V. KOSTOCHKA, AND D. R. WOODALL, *Total colorings of planar graphs with large maximum degree*, J. Graph Theory, 26 (1997), pp. 53–59.
- [8] T. R. JENSEN AND B. TOFT, *Graph Coloring Problems*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, New York, 1995.
- [9] A. V. KOSTOCHKA, *The total chromatic number of any multigraph with maximum degree five is at most seven*, Discrete Math., 162 (1996), pp. 199–214.

- [10] M. MOLLOY AND B. REED, *A bound on the total chromatic number*, *Combinatorica*, 18 (1998), pp. 241–280.
- [11] M. ROSENFELD, *On the total coloring of certain graphs*, *Israel J. Math.*, 9 (1971), pp. 396–402.
- [12] D. P. SANDERS AND Y. ZHAO, *On total 9-coloring planar graphs of maximum degree seven*, *J. Graph Theory*, 31 (1999), pp. 67–73.
- [13] D. P. SANDERS AND Y. ZHAO, *Planar graphs of maximum degree seven are class I*, *J. Combin. Theory Ser. B*, 83 (2001), pp. 201–212.
- [14] N. VIJAYADITYA, *On total chromatic number of a graph*, *J. London Math. Soc. (2)*, 3 (1971), pp. 405–408.
- [15] V. G. VIZING, *Some unsolved problems in graph theory*, *Uspehi Mat. Nauk*, 23 (1968), pp. 117–134.
- [16] W. WANG, *Total chromatic number of planar graphs with maximum degree ten*, *J. Graph Theory*, 54 (2007), pp. 91–102.
- [17] H. P. YAP, *Total Colorings of Graphs*, *Lecture Notes in Math.* 1623, Springer-Verlag, Berlin, 1996.
- [18] Z. ZHANG, L. LIU, J. WANG, AND H. LI, *A note on the total chromatic number of Halin graphs with maximum degree 4*, *Appl. Math. Lett.*, 11 (1998), pp. 23–27.
- [19] Z. F. ZHANG, J. X. ZHANG, AND J. F. WANG, *The total chromatic number of some graphs*, *Sci. Sinica Ser. A*, 31 (1998), pp. 1434–1441.

ODD MINIMUM CUT SETS AND b -MATCHINGS REVISITED*

ADAM N. LETCHFORD[†], GERHARD REINELT[‡], AND DIRK OLIVER THEIS[§]

Abstract. The famous Padberg–Rao separation algorithm for b -matching polyhedra can be implemented to run in $\mathcal{O}(|V|^2|E|\log(|V|^2/|E|))$ time in the uncapacitated case, and in $\mathcal{O}(|V||E|^2\log(|V|^2/|E|))$ time in the capacitated case. We give a new and simple algorithm for the capacitated case which can be implemented to run in $\mathcal{O}(|V|^2|E|\log(|V|^2/|E|))$ time.

Key words. matching, polyhedra, separation

AMS subject classifications. Primary, 90C27; Secondary, 05C70, 90C57, 90C10

DOI. 10.1137/060664793

1. Introduction. Let $G = (V, E)$ be an undirected graph, let $b \in \mathbb{Z}_+^V$ be a vector of vertex capacities, and let $u \in \mathbb{Z}_+^{|E|}$ be a vector of edge capacities. A u -capacitated b -matching is a family of edges, possibly containing multiple copies, such that

- for each $i \in V$, there are at most b_i edges in the family incident on i ;
- at most, u_e copies of edge e are used.

If we define for each edge e the integer variable x_e , representing the number of times e appears in the matching, then the incidence vectors of u -capacitated b -matchings are the solutions to

$$(1) \quad \sum_{e \in \delta(i)} x_e \leq b_i \quad \text{for all } i \in V,$$

$$(2) \quad 0 \leq x_e \leq u_e \quad \text{for all } e \in E,$$

$$(3) \quad x_e \in \mathbb{Z} \quad \text{for all } e \in E.$$

Here, as usual, $\delta(i)$ represents the set of edges incident on i .

The convex hull in $\mathbb{R}^{|E|}$ of solutions to (1)–(3) is called the u -capacitated b -matching polytope. Edmonds and Pulleyblank (see [Edm65] and [Pul73]) gave a complete linear description of this polytope. It is described by the *degree inequalities* (1), the *bounds* (2), and the following *blossom inequalities*:

$$(4) \quad \sum_{e \in E(W)} x_e + \sum_{f \in F} x_f \leq \left\lfloor \frac{b(W) + \sum_{f \in F} u_f}{2} \right\rfloor$$

for all $W \subset V, F \subset \delta(W)$ with $b(W) + \sum_{f \in F} u_f$ odd.

*Received by the editors July 11, 2006; accepted for publication (in revised form) May 13, 2008; published electronically October 1, 2008. An extended abstract of this paper appeared in *Integer Programming and Combinatorial Optimization 10*, Lecture Notes in Comput. Sci. 3064, Springer-Verlag, Berlin, 2004, pp. 196–205.

<http://www.siam.org/journals/sidma/22-4/66479.html>

[†]Department of Management Science, Lancaster University, Lancaster LA1 4YW, UK (A.N. Letchford@lancaster.ac.uk).

[‡]Institute of Computer Science, University of Heidelberg, Heidelberg, Germany (Gerhard.Reinelt@informatik.uni-heidelberg.de).

[§]Service de Géométrie, Combinatoire et Théorie des Groupes, Département de Mathématique, Université Libre de Bruxelles, Brussels, Belgium (Dirk.Theis@ulb.ac.be). This author's research was supported by Deutsche Forschungsgemeinschaft (DFG, RE 776/9-1).

Here, $E(W)$ (respectively, $\delta(W)$) represents the set of edges with both endvertices (respectively, exactly one endvertex) in W , and $b(W)$ denotes $\sum_{i \in W} b_i$.

An important special case is where the upper bounds u_e are not present (or, equivalently, $u_{ij} \geq \max\{b_i, b_j\}$ for all $\{i, j\} \in E$). The associated *uncapacitated* b -matching polytope is described by the degree inequalities, the nonnegativity inequalities $x_e \geq 0$ for all $e \in E$, and the *simplified blossom inequalities*

$$(5) \quad \sum_{e \in E(W)} x_e \leq \left\lfloor \frac{b(W)}{2} \right\rfloor \quad \text{for all } W \subset V \text{ with } b(W) \text{ odd.}$$

In their seminal paper, Padberg and Rao [PR82] devised an efficient combinatorial *separation algorithm* for b -matching polytopes. A separation algorithm is (see [GLS88]) a procedure which, given a point $x^* \in \mathbb{Q}^{|E|}$ lying outside of the polytope, finds a linear inequality which is valid for the polytope yet violated by x^* . Clearly, testing if a degree inequality or bound is violated can be performed in linear time; therefore, the main contribution of [PR82] is to identify violated blossom inequalities.

For *uncapacitated* b -matching, the Padberg–Rao algorithm involves the solution of up to $|V| - 1$ max-flow problems on graphs with $\mathcal{O}(|V|)$ vertices and $\mathcal{O}(|E|)$ edges. Using the well-known *preflow push* algorithm [GT88] to solve the max-flow problems, this leads to an overall running time of $\mathcal{O}(|V|^2 |E| \log(|V|^2/|E|))$.

The Padberg–Rao separation algorithm for *capacitated* b -matching, however, is substantially more time consuming. It involves the solution of up to $|V| + |E| - 1$ max-flow problems on graphs with $\mathcal{O}(|E|)$ vertices and $\mathcal{O}(|E|)$ edges. Using the preflow push algorithm, this leads to a worst-case running time of $\mathcal{O}(|E|^3 \log|V|)$.

In 1987, Grötschel and Holland [GH87] observed that the max-flow computations needed in the capacitated case can in fact be carried out on graphs with only $\mathcal{O}(|V|)$ vertices and $\mathcal{O}(|E|)$ edges. Although the idea behind this is simple, it reduces the overall running time for the capacitated case to $\mathcal{O}(|V||E|^2 \log(|V|^2/|E|))$.

In this paper, we propose a new separation algorithm for the capacitated case, whose running time is the same as that for the uncapacitated case. As well as being faster than the previous approaches, the new algorithm is simpler and easier to implement. It also has a surprisingly simple proof of correctness. Like the previous algorithms, it can also be easily adapted to the case of *perfect* capacitated b -matchings, in which the inequalities (1) are changed to equations, and to certain other problems in the matching family.

The importance of blossom separation extends, of course, far beyond matching problems. This is because blossom-like inequalities can be defined for many important $\mathcal{NP}$ -hard combinatorial optimization problems involving constraints on vertex degrees and bounds on the variables. The most famous problem of this type is the *traveling salesman problem*, for which certain blossom inequalities are facet-inducing and make very useful cutting planes (see [GP79a], [GP79b], [PR90]). Other problems for which blossom-like inequalities can be defined and separated more quickly using the new algorithm include, for example, the *Capacitated Vehicle Routing Problem* [LLE04], the *Rural Postman Problem* [GL00], the *Capacitated Arc Routing Problem* [BB98], the *k-edge Connected Spanning Subgraph Problem* [Mah94], and the *Network Connectivity Problem* [GMS92].

Another very interesting application of blossom separation is described in [CF96]. In that paper, a family of cutting planes for general integer linear programs is defined, called $\{0, \frac{1}{2}\}$ -*Chvátal–Gomory cuts*, and it is shown that the associated separation problem can be solved in polynomial time if the constraint matrix satisfies certain

conditions. Under one of these conditions, the separation problem essentially reduces to blossom separation, and our algorithm can be used. (We omit details for the sake of brevity.)

Note: An extended abstract of this paper appeared in the proceedings of IPCO X [LRT04]. However, the proof of correctness of the algorithm has been simplified substantially.

2. Review of key concepts. To make the paper self-contained, and also for the sake of clarity, we review in this section some key concepts from the literature.

2.1. Cut-trees. The notion of a *cut-tree* was introduced by Gomory and Hu [GH61]. Let $G = (V, E)$ be an undirected graph, $w \in \mathbb{Q}_+^{|E|}$ a vector of edge weights, and $X \subset V$ a set of *terminal vertices*. A cut-tree is an edge-weighted tree, spanning X , that compactly represents the minimum cut in G between every pair of vertices in X . More formally, the cut-tree consists of a mapping $\pi: V \rightarrow X$ with $\pi(x) = x$ for all $x \in X$, and an adjacency relation $\sim$ on the set X . The adjacency relation shall make the set of terminal vertices into a tree, i.e., $x \sim y$ means that x and y are connected by an edge of the cut-tree. Deleting an edge $x \sim y$ of the cut-tree partitions the set X into two sets X_x and X_y , and thus defines a cut $(U, \bar{U})$ in G by letting $U := \pi^{-1}(X_x)$ and $\bar{U} := \pi^{-1}(X_y)$. We call this the cut *induced* by the edge $x \sim y$. Now, the following condition must hold:

- (6) for all pairs $x, y \in X$ with $x \sim y$, the cut induced by the edge $x \sim y$ is a minimum (x, y) -cut in G with respect to the weights w .

The Gomory–Hu algorithm for producing a cut-tree involves solving $|X| - 1$ max-flow/min-cut problems on a series of graphs obtained from G by suitable contractions. Using the preflow push algorithm as a subroutine, the Gomory–Hu algorithm runs in $\mathcal{O}(|X||V||E| \log(|V|^2/|E|))$ time.

2.2. Minimum odd cuts. We now move on to consider *minimum odd cuts*. As before, let $G = (V, E)$ be an undirected graph with edge weights $w \in \mathbb{Q}_+^{|E|}$. Let $T \subset V$, with $|T|$ even, be a set of vertices that are labelled *odd*. A cut $\delta(U)$ is called *T-odd*, or just *odd* for short, if $|T \cap U|$ is an odd number. The *minimum odd cut problem* asks for an odd cut $\delta(U)$ with minimum weight $w(\delta(U))$.

In [PR82], Padberg and Rao gave a surprisingly simple algorithm to compute minimum odd cuts. It involves simply invoking the Gomory–Hu algorithm with terminal set T and then checking each of the $|T| - 1$ cuts produced; see Algorithm 1. It is easy to show that the bottleneck of the algorithm is the Gomory–Hu call, and therefore the algorithm can be implemented to run in $\mathcal{O}(|T||V||E| \log(|V|^2/|E|))$ time.

2.3. The Padberg–Rao separation algorithms. Padberg and Rao then showed how to reduce the blossom separation problem to the minimum odd cut problem.

First, consider the uncapacitated case. We introduce, for each $i \in V$, the term $s_i := b_i - \sum_{e \in \delta(i)} x_e$, which can be viewed as the *slack* of the corresponding degree inequality (1). The simplified blossom inequality (5) can then easily be rewritten to take the following form:

$$(7) \quad \sum_{i \in W} s_i + \sum_{e \in \delta(W)} x_e \geq 1.$$

Now, let $x^* \in \mathbb{R}_+^{|E|}$ be a given vector satisfying the degree inequalities. Take the graph G and assign the weight x_e^* to each $e \in E$. Add a new *dummy* vertex v and, for each

Algorithm 1. Minimum odd cut [PR82]

Input:

Graph G , set $T \subset V$ of odd vertices, weight vector $w \in \mathbb{Q}_+^{|E|}$.

Output:

A minimum odd cut.

1: Let $r = \infty$.

2: Compute a cut-tree for the graph G with weights w and terminal vertex set T .

3: **For** each of the $|T| - 1$ edges of the cut-tree

4: Let $\delta(U)$ denote the cut induced by the cut-tree edge.

5: *Check the cut:*

 If $|T \cap U|$ is odd and $w(\delta(U)) < r$, then store U and set $r := w(\delta(U))$.

6: **End for**

7: Output the cut $\delta(U)$.

$i \in V$, an edge $\{i, v\}$ with weight $s_i^* := b_i - \sum_{e \in \delta(i)} x_e^*$. Denote the resulting graph by $G^+ = (V^+, E^+)$. Now, label the vertices $i \in V$ that have an odd value of $b(i)$ *odd*. Also label the dummy vertex v *odd* if $b(V)$ is odd.

Clearly, there is a violated blossom inequality in the form (7) if and only if there exists an odd cut in G^+ of weight less than 1. Thus, one can use the minimum odd cut algorithm described in the previous subsection. As a result, blossom separation can be performed in $\mathcal{O}(|V|^2|E| \log(|V|^2/|E|))$ time in the uncapacitated case. This is $\mathcal{O}(|V|^4)$ in the worst case.

The capacitated case is a little more complicated. First, the blossom inequality (4) needs to be rewritten in the following form:

$$(8) \quad \sum_{i \in W} s_i + \sum_{e \in \delta(W) \setminus F} x_e + \sum_{e \in F} (u_e - x_e) \geq 1.$$

Now, let x^* be a given vector satisfying the degree inequalities and bounds. Padberg and Rao begin by constructing the so-called split graph, which is obtained from the graph G^+ mentioned above by subdividing each edge $\{i, j\} \in E$ into two halves. (That is, each edge $\{i, j\} \in E$ is replaced with two edges, say $\{i, k\}$ and $\{j, k\}$.) One of the two halves (which we shall call the *normal* half) is given a weight of x_e^* , whereas the other half (which we shall call the *complemented* half) is given a weight of $u_e - x_e^*$.

Parities are then assigned to the edges and vertices of the split graph. The normal halves are labelled *even*, the complemented halves are labelled *even* or *odd* according to whether the corresponding upper bound u_e is even or odd, and the edge $\{i, v\}$ is labelled *even* or *odd* according to whether the upper bound b_i is even or odd. Finally, a vertex is labelled *even* or *odd* according to whether it is incident on an even or odd number of odd edges.

It can then be shown that there is a violated blossom inequality in the form (8) if and only if there exists an odd cut in the split graph of weight less than 1. Unfortunately, the split graph is substantially larger than G , with $|V| + |E| + 1$ vertices and $2|E| + |V|$ edges. Thus, the minimum odd cut computation now involves $\mathcal{O}(|E|)$ max-flow computations in graphs with $\mathcal{O}(|E|)$ vertices and edges. Using the preflow push algorithm, this leads to an overall running time $\mathcal{O}(|E|^3 \log |V|)$, or $\mathcal{O}(|V|^6 \log |V|)$ in the worst case.

2.4. The Grötschel–Holland modification. As mentioned in the introduction, a simple but significant improvement to the algorithm for the capacitated case

was made by Grötschel and Holland [GH87]. (In fact, they were only interested in 2-matchings, but the same argument applies in the general case.) Consider an arbitrary edge $e = \{i, j\} \in E$. It is obvious that the maximum amount of flow which can be directly sent from i to j in the split graph is equal to the minimum of x_e^* and $u_e - x_e^*$. Consequently, at each iteration of the Gomory–Hu algorithm, for each $e \in E$, one of the two halves can be contracted before computing a max-flow. As a result, the max-flow problems can be conducted on graphs which are of a similar size to G . This reduces the overall running time to $\mathcal{O}(|V||E|^2 \log(|V|^2/|E|))$, or $\mathcal{O}(|V|^5)$ in the worst case.

3. The new algorithm. We now present a new, faster algorithm for blossom separation in the capacitated case. In fact, our algorithm solves a slightly more general problem. Given a graph $G = (V, E)$ and a set $T \subset V$ of odd vertices, let us say that a *blossom* is a pair (U, F) consisting of a set of vertices $U \subset V$ and a set of edges $F \subset \delta(U)$ with the property that $|T \cap U| + |F|$ is an odd number. Now suppose that we are given, in addition, two nonnegative weight vectors $c, c' \in \mathbb{Q}_+^{|E|}$. We define the *blossom minimization problem* as the problem of finding a blossom whose *value*

$$\beta(U, F) := \sum_{e \in \delta(U) \setminus F} c_e + \sum_{f \in F} c'_f$$

is as small as possible. To cast the blossom separation problem in this form, it suffices to construct the graph G^+ used by Padberg and Rao for the uncapacitated case (see subsection 2.3), and set

$$(c_e, c'_e) := \begin{cases} (x_e^*, u_e - x_e^*) & \text{if } u_e \text{ is odd,} \\ (\min(x_e^*, u_e - x_e^*), \infty) & \text{if } u_e \text{ is even} \end{cases}$$

for all $e \in E$. For the edges $\{i, v\}$ of G^+ , we set $c_{iv} := s_i^*$ and $c'_{iv} := \infty$ (v is the dummy vertex in G^+). With this construction, it is an easy exercise to verify the following proposition.

PROPOSITION 3.1. *Let (U, F^+) be a blossom in G^+ with the weight vectors (c, c') just described. W.l.o.g. we may assume that the dummy vertex v is not in U . Let*

$$W := U \quad \text{and} \quad F := F^+ \cup \{e \in \delta(W) \mid u_e \text{ even and } u_e - x_e^* x_e^* < x_e^*\}.$$

Then W and F satisfy the conditions required in the definition of a blossom inequality, and therefore we have

$$(9) \quad \sum_{i \in W} s_i + \sum_{e \in \delta(W) \setminus F} x_e + \sum_{e \in F} (u_e - x_e) = \beta(U, F^+).$$

(The left-hand side of the equation is the left-hand side of inequality (8).) Conversely, every blossom inequality for which $F \subset \delta(W)$ is chosen such that the left-hand side in (8) is minimized defines a blossom (U, F^+) by letting $U := W$ and $F^+ := \{e \in F \mid u_e \text{ odd}\}$, for which (9) holds.

Our algorithm for the blossom minimization problem is displayed as Algorithm 2.

Remark 3.2. For fixed $U \subset V$, it has been observed by Padberg and Rinaldi [PR90] that

$$(10) \quad \beta(U) := \min \left\{ \beta(U, F) \mid F \subset \delta(U), |T \cap U| + |F| \text{ odd} \right\}$$

Algorithm 2. Blossom minimization**Input:**

Graph G , set $T \subset V$ of odd vertices, weight vectors $c, c' \in \mathbb{Q}_+^{|E|}$.

Output:

A minimum blossom.

- 1: Let $r = \infty$.
- 2: Compute a cut-tree for G with weights $\min(c, c')$ and terminal vertex set V .
- 3: **For** each of the $|V| - 1$ edges of the cut-tree
- 4: Let $\delta(U)$ denote the cut induced by the cut-tree edge.
- 5: *Check the cut:*
 Compute $\beta(U)$ as in (10).
- 6: If $\beta(U) < r$, then store U along with the arg-min F and set $r := \beta(U)$.
- 7: **End for**
- 8: Output the best blossom (U, F) .

can be computed in time $\mathcal{O}(|\delta(U)|)$ by first tentatively taking $F := \{e \in \delta(U) \mid c'_e < c_e\}$. Now if $|T \cap U| + |F|$ is odd, then we have found a minimizing F . Otherwise, find $f \in \delta(U)$ minimizing $|c_f - c'_f|$ over $f \in \delta(U)$, because then the symmetric difference of F and $\{f\}$ minimizes $\beta(U, \cdot)$.

This implies that the loop in steps 2–6 in Algorithm 2 runs in time $\mathcal{O}(|V|^2)$, and that the running time of Algorithm 2 is dominated by the computation of the cut-tree in step 1, which amounts to $\mathcal{O}(|V|^2|E| \log(|V|^2/|E|))$.

The similarity between the Padberg–Rao minimum odd cut Algorithm 1 and our blossom minimization Algorithm 2 is striking. Moreover, in the next section, we give a short and elegant proof of correctness of Algorithm 2, which is similar to a proof of correctness of Algorithm 1 given by Rizzi [Riz02].

We remark that, just like the original Padberg–Rao blossom separation algorithm, our algorithm can be modified so that it outputs every violated blossom inequality found, rather than only the most violated one. We also remark that the function $\beta(U)$ defined in (10), unlike the cut function $w(\delta(U))$, is not, in general, submodular.

4. A simple proof of the correctness of Algorithm 2. The proof makes use of a well-known fact concerned with T -odd cuts and so-called T -joins, which we briefly review. Let $G = (V, E)$ be a graph, and let $T \subset V$ (with $|T|$ even) be a set of odd vertices, as before. A set $F \subset E$ is called a T -join if it has the following property: a vertex of G is odd if and only if there is an odd number of edges in F incident on it. The fact that we need is the following: if F is a T -join and U is an arbitrary subset of V , then $|U \cap T|$ is odd if and only if $|\delta(U) \cap F|$ is odd. That is, a cut is T -odd if and only if it meets every T -join an odd number of times.

Consider once again a graph $G = (V, E)$, a weight vector $w \in \mathbb{Q}^{|E|}$, and a set $T \subset V$ of odd vertices. Suppose that we have computed a cut-tree with terminal vertex set V . We say that an edge $x \sim y$ of the cut-tree is T -odd if the vertex sets on each shore of the cut defined by $x \sim y$ are T -odd. This defines a T -join in the cut-tree. The next theorem, due to Rizzi [Riz02], is an extension of a key theorem in [PR82]. For the sake of clarity, we repeat the proof of Rizzi.

THEOREM 4.1 (see [Riz02]). *One of the T -odd edges of the cut-tree induces a minimum odd cut in G .*

Proof. Let $U \subset V$ be an odd set such that $\delta(U)$ is a minimum odd cut in G . Since U is an odd set, there exists an odd number of T -odd cut-tree edges leaving U . Let

$x \sim y$ be one of them, and let $\delta(U')$ be the minimum (x, y) -cut that it induces in G (by property (6)). Since $\delta(U)$ is also an (x, y) -cut in G , we have $w(\delta(U')) \leq w(\delta(U))$. Since $x \sim y$ is a T -odd cut-tree edge, $\delta(U')$ is a minimum odd cut in G . $\square$

Now we come to the proof of correctness of Algorithm 2.

THEOREM 4.2. *One of the edges of the cut-tree computed in Algorithm 2 induces a set $U \subset V$ that minimizes $\beta(\cdot)$.*

Proof. Let $E' := \{e \in E \mid c'_e < c_e\}$ and T' denote the set of vertices that are incident on an odd number of edges in E' . This way, E' is a T' -join. Define the set T^Δ as the symmetric difference of T with T' . Furthermore, let $w := \min(c, c')$ denote the weight function used in Algorithm 2. We note the following facts:

- (a) A set U is T^Δ -odd if and only if $|T \cap U| + |E' \cap \delta(U)|$ is odd.
- (b) For all sets U , we have $\beta(U) \geq w(\delta(U))$. If U is T^Δ -odd, then the equality holds. If U is not T^Δ -odd, then, by Remark 3.2,

$$\beta(U) = w(\delta(U)) + \min_{e \in \delta(U)} |c'_e - c_e|.$$

Item (4) follows from the elementary fact about T -cuts and T -joins mentioned previously by taking a T -join J and noting that the symmetric difference of J and E' is a T^Δ -join.

Now let $U \subset V$ be a set that minimizes $\beta(\cdot)$. We distinguish two cases.

Case 1: U is T^Δ -odd. The proof of Theorem 4.1, applied to T^Δ and w , shows that there exists a T^Δ -odd edge of the cut-tree which, by (4), determines a set $U' \subset V$ minimizing $\beta(\cdot)$.

Case 2: U is not T^Δ -odd. Let $f = \{i, j\} \in \delta(U)$ have the minimal value of $|c_f - c'_f|$ among all edges in $\delta(U)$. On the path from i to j in the cut-tree, at least one edge $x \sim y$ has one end in U and the other not in U . Let $\delta(U')$ be the minimum (x, y) -cut in G corresponding to this edge. Notice that $f \in \delta(U')$. We then have

$$\beta(U) = w(\delta(U)) + |c_f - c'_f| \geq w(\delta(U')) + |c_f - c'_f| \geq \beta(U').$$

The first inequality holds because $\delta(U)$ is an (x, y) -cut in G . The second follows from the definition of the function $\beta(\cdot)$ and the fact that $f \in \delta(U')$. $\square$

REFERENCES

- [BB98] J. M. BELENGUER AND E. BENAVENT, *The capacitated arc routing problem: Valid inequalities and facets*, Comput. Optim. Appl., 10 (1998), pp. 165–187.
- [CF96] A. CAPRARA AND M. FISCHETTI, $\{0, \frac{1}{2}\}$ -Chvátal-Gomory cuts, Math. Programming, 74 (1996), pp. 221–235.
- [Edm65] J. EDMONDS, *Maximum matching and a polyhedron with 0-1 vertices*, J. Res. Nat. Bur. Standards Sect. B, 69B (1965), pp. 125–130.
- [GL00] G. GHIANI AND G. LAPORTE, *A branch-and-cut algorithm for the undirected rural postman problem*, Math. Program., 87 (2000), pp. 467–481.
- [GH61] R. E. GOMORY AND T. C. HU, *Multi-terminal network flows*, J. Soc. Indust. Appl. Math., 9 (1961), pp. 551–570.
- [GH87] M. GRÖTSCHEL AND O. HOLLAND, *A cutting plane algorithm for minimum perfect 2-matching*, Computing, 39 (1987), pp. 327–344.
- [GLS88] M. GRÖTSCHEL, L. LOVÁSZ, AND A. J. SCHRIJVER, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, Berlin, 1988.
- [GMS92] M. GRÖTSCHEL, C. L. MONMA, AND M. STOER, *Computational results with a cutting plane algorithm for designing communication networks with low-connectivity constraints*, Oper. Res., 40 (1992), pp. 309–330.

- [GP79a] M. GRÖTSCHEL AND M. W. PADBERG, *On the symmetric travelling salesman problem I: Inequalities*, Math. Programming, 16 (1979), pp. 265–280.
- [GP79b] M. GRÖTSCHEL AND M. W. PADBERG, *On the symmetric travelling salesman problem II: Lifting theorems and facets*, Math. Programming, 16 (1979), pp. 281–302.
- [GT88] A. V. GOLDBERG AND R. E. TARJAN, *A new approach to the maximum flow problem*, J. Assoc. Comput. Math., 35 (1988), pp. 921–940.
- [LRT04] A. N. LETCHFORD, G. REINELT, AND D. O. THEIS, *A faster exact separation algorithm for blossom inequalities*, in Integer Programming and Combinatorial Optimization 10, D. Bienstock and G. Nemhauser, eds., Lecture Notes in Comput. Sci. 3064, Springer-Verlag, Berlin, 2004, pp. 196–205.
- [LLE04] J. LYSGAARD, A. N. LETCHFORD, AND R. W. EGGLESE, *A new branch-and-cut algorithm for the capacitated vehicle routing problem*, Math. Programming, 100 (2004), pp. 423–445.
- [Mah94] A. R. MAHJOUB, *Two-edge connected spanning subgraphs and polyhedra*, Math. Programming, 64 (1994), pp. 199–208.
- [PR82] M. W. PADBERG AND M. R. RAO, *Odd minimum cut-sets and b -matchings*, Math. Oper. Res., 7 (1982), pp. 67–80.
- [PR90] M. PADBERG AND G. RINALDI, *Facet identification for the symmetric traveling salesman polytope*, Math. Program., 47 (1990), pp. 219–257.
- [Pul73] W. R. PULLEYBLANK, *Faces of Matching Polyhedra*, Ph.D. thesis, University of Waterloo, Waterloo, Ontario, 1973.
- [Riz02] R. RIZZI, *Minimum T -cuts and optimal T -pairings*, Discrete Math., 257 (2002), pp. 177–181.

ANALYSIS OF SERIAL TURBO CODES OVER ABELIAN GROUPS FOR SYMMETRIC CHANNELS*

FEDERICA GARIN[†] AND FABIO FAGNANI[†]

Abstract. In this paper we study serial turbo interconnections of Abelian group codes to be used on symmetric channels. Particular attention is devoted to AWGN channel with input restricted to m -PSK constellation with corresponding group structure $\mathbb{Z}_m$. We establish the exact asymptotic decay of the average symbol and word error probabilities when the interleaver length goes to infinity (interleaver gain). Moreover, we give a detailed characterization of the distance parameter characterizing the behavior for the signal-to-noise ratio going to infinity (effective free distance). Some of our results are new also in the binary context; in particular, the lower bound to the error probability decay.

Key words. serial interconnected codes, turbo codes, trellis-coded modulation, average error probability, codes over groups, effective distance

AMS subject classifications. 94B10, 94B12, 94B70, 90C27

DOI. 10.1137/07068802X

1. Introduction. Turbo codes made their appearance in 1993 with the pioneering work [6] (see also [5]). Since then, there has been a huge effort to understand their great performance by a significant number of researchers in the field, e.g., [3, 4, 9, 10, 12, 14, 28, 35, 37, 38, 39, 40, 41]. In the meanwhile, many variations and extensions have been proposed with respect to the original parallel concatenated scheme in [5]. A particularly relevant scheme is the serially concatenated version, first proposed in [2] and studied further in [28]. Other significant variations concern the possibility to work with a more restricted class of coupling interleavers [7, 11] in order to achieve a higher convergence abscissa in the iterative decoding algorithm.

In the theoretical analysis of these coding schemes we can find two main lines: on the one hand, they have been studied in combination with the suboptimal iterative decoding algorithm (see, e.g., [35, 10, 14]), and on the other hand, they have been studied in a more classical setting, considering optimal maximum likelihood (ML) decoding. This second approach separates the analysis of the coding schemes from the use of the suboptimal iterative decoding (see, e.g., [3, 4, 2, 9, 12, 28, 37, 39, 40]). The results in [4, 2, 28] provide upper bounds for the decay of the averaged bit error rate $\overline{P_b(e)}$, of type $CN^{-\alpha}$ where α is a parameter, called interleaver gain, depending only on the constituent codes of the turbo scheme. Further, a more refined analysis has been done for the situation when the Bhattacharyya parameter γ of the channel goes to zero, showing that the above constant C could be estimated as $C = K\gamma^{d^*}$ where d^* is a sort of generalized minimum distance of the scheme and K is a universal constant. The main contribution of this line of research has been to discover these two parameters, α and d^* , and their relation to the constituent encoders, which makes them interesting from a design point of view.

*Received by the editors April 11, 2007; accepted for publication (in revised form) May 25, 2008; published electronically October 1, 2008.

<http://www.siam.org/journals/sidma/22-4/68802.html>

[†]Dipartimento di Matematica, Politecnico di Torino, C.so Duca degli Abruzzi 24, 10129 Torino, Italy (garin@dei.unipd.it, fabio.fagnani@polito.it). The first author is now with the Department of Information Engineering (DEI), Università di Padova, Via Gradenigo 6/a, 35131 Padova, Italy.

Another significant extension of the theory of turbo codes has been the study of schemes for nonbinary alphabets, in order to use them over channels with a nonbinary input alphabet.

Mostly, the construction of such nonbinary turbo codes has followed the so-called *pragmatic design* approach: turbo codes are designed for the binary case and optimized independently from the external high-order input constellation. Joint optimization is very little, usually regarding the mapping of the coded bits into the constellation signals. Performance evaluation is typically obtained by simulation. As a matter of fact, almost all schemes presented in the literature belong to this class (see, e.g., [31, 36, 21]).

On the contrary, in the so-called *analytical design* turbo codes are designed and optimized by taking into account the chosen nonbinary channel. This is the same philosophy of classical trellis coded modulation (TCM) schemes. Pioneering works in this direction are [23, 34, 16]. This approach exploits the symmetries of some important families of codes, such as additive Gaussian channels with geometrically uniform input constellation, by using a suitable algebraic structure matched to the channel. There are two fundamental difficulties in the theoretical developments of such turbo schemes, new with respect to the binary case:

- in general, the relevant distance affecting error probability and its estimations is not Hamming distance, e.g., for Gaussian channels Euclidean distance matters;
- in order to obtain bit or word error rate to be independent from the transmitted information word, one needs to carefully choose the algebraic structure of the constituent encoders. For example, the so-called bit-geometrically-uniform encoders [23, 16] ensure a sort of isometry in the scheme between the input Hamming distance and the output Euclidean distance; these encoders give a bit error rate independent from the information word, but this independence is only reached in the average sense and not for a single realization of the interleaver and, moreover, it forces the use of encoders on non-Abelian groups, creating many algebraic technical problems in the treatment of the convolutional codes.

This paper also follows the *analytical design* approach, but taking a different road with respect to previous works. Here, we study serial turbo interconnections on an Abelian group (typically $\mathbb{Z}_m$). The constituent encoders are chosen to be homomorphic, which does not guarantee the bit error rate $P_b(e)$ (or its average version $\overline{P_b(e)}$) to be independent from the transmitted information word. However, if we consider the symbol error rate $P_s(e)$, where the symbol is to be intended as a suitable aggregation of the input bits in order to form the input group of the encoder (typically $\mathbb{Z}_m$ again), then this independence is ensured. Of course $P_s(e)$ and $P_b(e)$ are strictly linked to each other, so that estimations on the former lead to estimations on the latter. Moreover, the choice to work with Abelian groups allows us to use a much richer and flexible theory at the level of convolutional codes, widely studied in the past [29, 17, 18, 19].

In this paper, we study the average symbol and word error rate $\overline{P_s(e)}$, $\overline{P_w(e)}$ for such schemes on symmetric channels and under ML decoding. We determine the asymptotic decay $N^{-\alpha}$ when the interleaver length N goes to infinity. Our results extend and complete the bounds given in [2, 28]. The extension is two-fold: we pass from binary codes to schemes over generic Abelian groups, and we move from the uniform interleaver assumption by also considering structured families of interleavers. Moreover, we provide a lower bound which was missing even for the binary case.

Technical proofs are inspired by [15], which analyzes parallel interconnection schemes for binary codes.

Finally, we carry on a detailed investigation of the asymptotic analysis when the signal-to-noise ratio (SNR) goes to infinity. Similar to results in [2] for the binary case, we find that the leading term of both word and symbol error probabilities (for $N \rightarrow +\infty$) depends on the channel's SNR through a sort of distance parameter q^* , which is a function of the constituent encoders. However, different from the binary case, the exact characterization of q^* is given as the solution of a combinatorial optimization problem involving both constituent encoders.

Here, we give a brief outline of this paper. Sections 2 and 3 contain some basic notation and results on block and convolutional encoders over Abelian groups: special attention is devoted to homomorphic state-space realizations, which play a fundamental role in all of our estimations. In section 4 we introduce the serially concatenated schemes and the specific examples we want to focus on. Section 5 contains the fundamental original theoretical results of this paper: Theorem 5.4 describes the asymptotic estimations for word and symbol error probabilities when $N \rightarrow +\infty$ and, further, when the SNR goes to infinity. In particular, the expression for the interleaver gain α is derived and the combinatorial optimization problem describing the effective distance q^* is formulated. All proofs are given in section 6. Section 7 gives a detailed account of the computation of α and d^* for all of our examples. An appendix containing some results on the algebraic properties of encoders over $\mathbb{Z}_m$ completes the paper.

2. Homomorphic block encoders for symmetric channels.

2.1. Notation. Given a set Ω and $A \subseteq \Omega$, the symbol $\chi_A : \Omega \rightarrow \{0, 1\}$ will denote the indicator function of A , i.e., $\chi_A(x) = 1$ if and only if $x \in A$. $|A|$ will denote the cardinality of A . We will denote by $\mathbb{N}$ the set of nonnegative integers and by $\mathbb{N}^*$ the set of positive integers. Throughout this paper, vectors will always be column vectors and be denoted by boldface letters. Given a vector $\mathbf{w} \in \mathbb{N}^k$, we let $|\mathbf{w}| = \sum_j w_j$. We will denote by $\mathbf{e}_j$ a vector of the appropriate length (clear by the context or explicitly stated) made by all zeros except a one in position j . Given two sets, A and B , B^A will denote vectors with entries in B , having length $|A|$, and components indexed by elements of A instead of integers $1, \dots, |A|$. By $\log$ and $\exp$ we will denote logarithm and exponential with respect to the same base $b > 1$. Given groups G and H , $\text{Hom}(G, H)$ will denote the group of all homomorphisms from G to H , while $\text{Aut}(G)$ will be the group of automorphisms of G .

2.2. Symmetric channels. A memoryless channel is described by:

- an input alphabet $\mathcal{X}$ (which we will always assume is finite),
- an output alphabet $\mathcal{Y}$, endowed with a σ -algebra $\mathcal{B} \subseteq 2^{\mathcal{Y}}$ and a probability measure μ ,
- and a family of transition probability densities $W(\cdot|x)$ on $\mathcal{Y}$, indexed by the inputs $x \in \mathcal{X}$.

Such a channel will be denoted by $(\mathcal{X}, \mathcal{Y}, W)$. In most applications, either $\mathcal{Y}$ is finite and μ is the counting measure so $W(\cdot|x)$ are simply probability vectors, or $\mathcal{Y} = \mathbb{R}^n$ and μ is the Lebesgue measure.

To give a formal definition of symmetric memoryless channels, we need to recall some definitions of group actions. Given a group $(G, +)$ with neutral element 0 and given a set A , G acts on A if for every $g \in G$ there exists a map $a \mapsto ga$ from A to A , such that $(h + g)a = h(ga)$ for all $h, g \in G$, $a \in A$, and $0a = a$ for all $a \in A$. For finite A , the group action of G on A is said to be (simply) transitive if for every

$a, b \in A$, there exists a (unique) element $g \in G$ such that $ga = b$. If G acts simply transitively on A , then G and A are in bijection through the map $\theta : G \rightarrow A$ defined by $\theta(g) = ga_0$ for an arbitrary but fixed $a_0 \in A$.

Given a probability space $\mathcal{Y}$ with σ -algebra $\mathcal{B}$ and probability measure μ , we say that a group G acts isometrically on $\mathcal{Y}$ if there exists an action of G on $\mathcal{Y}$ consisting of measurable bijections such that $\mu(gA) = \mu(A)$ for all $A \in \mathcal{B}$, for all $g \in G$. If $\mathcal{Y}$ is finite, then all group actions on $\mathcal{Y}$ are isometric. If $\mathcal{Y} = \mathbb{R}^n$, then an action is isometric when all maps $y \mapsto gy$ are isometries of $\mathbb{R}^n$.

Given a group G , a memoryless channel $(\mathcal{X}, \mathcal{Y}, W)$ is called G -symmetric if

1. G acts simply transitively on $\mathcal{X}$,
2. G acts isometrically on $\mathcal{Y}$,
3. $W(y|x) = W(gy|gx)$ for every $g \in G$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$.

In this case, the bijection $\theta : G \rightarrow \mathcal{X}$ defined by $\theta(g) = gx_0$ for some fixed $x_0 \in \mathcal{X}$ is called an isometric labelling.

The most common examples of G -symmetric channels are the following.

- **Binary-input output-symmetric channels.** $\mathbb{Z}_2$ -symmetric channels are known in the coding literature as binary-input output-symmetric (BIOS) channels. Well-known examples are binary symmetric channel (BSC), binary erasure channel (BEC), and binary-input AWGN (BIAWGN) channel.
- **Geometrically uniform AWGN channels.** An n -dimensional constellation is a finite subset $S \subset \mathbb{R}^n$ that spans $\mathbb{R}^n$; we denote with $\Gamma(S)$ its symmetry group, i.e., the group of Euclidean isometries of $\mathbb{R}^n$ mapping S into S itself. A constellation S is said to be geometrically uniform (GU) with generating group G if G is a subgroup of $\Gamma(S)$ whose action on S is simply transitive. The simplest example of GU constellation is the 1-dimensional antipodal constellation $\{-1, 1\}$ (a.k.a. 2-points pulse amplitude modulation (2-PAM)). A 2-dimensional example is the m -PSK constellation

$$S = \{e^{\frac{2\pi il}{m}} : l = 0, \dots, m-1\} \subseteq \mathbb{C} \simeq \mathbb{R}^2$$

which always has the generating group $\mathbb{Z}_m$ (seen as rotations of angles multiple of $2\pi/m$) and, if m is even, then also has the non-Abelian generating group $D_{m/2}$. For a complete theory of GU constellations and generating groups, see [20] and [32].

Given a GU constellation $S \subset \mathbb{R}^q$ with generating group G , define the S -AWGN channel as the memoryless channel $(S, \mathbb{R}^n, W)$ where the family W of n -dimensional transition densities is given by $W(y|x) = N(y-x)$ for any $x \in S$; $N(\cdot)$ is the density of an n -dimensional diagonal Gaussian random variable (r.v.) $N(y) = (2\pi\sigma^2)^{-n/2} e^{-\|y\|^2/(2\sigma^2)}$.

Other examples of G -symmetric channels can be obtained from the S -AWGN by suitable symmetric quantizations of the channel output, e.g., quantizing with respect to the Voronoi regions of the same constellation S .

- **m -ary symmetric channels.** This is a simple generalization of the BSC: $\mathcal{X} = \mathcal{Y} = \{0, 1, \dots, m-1\}$ and $W(y|x) = \epsilon/(m-1)$ if $y \neq x$, $W(y|x) = 1 - \epsilon$ if $y = x$. This channel is G -symmetric for any group G with $|G| = m$; in particular, for $G = \mathbb{Z}_m$.

Now fix a G -symmetric channel $(\mathcal{X}, \mathcal{Y}, W)$, and denote by W_n , μ_n , and θ the natural extensions of W , μ , and θ to multiple uses of the channel. Define the pairwise equivocation probability of a word $\mathbf{c} \in G^n$, $P(\mathbf{0} \rightarrow \mathbf{c})$, to be the probability that, for some fixed decoding rule, the decoder will prefer $\mathbf{c}$ to $\mathbf{0}$, given that $\theta(\mathbf{0})$ was transmitted. In this paper, we consider maximum likelihood decoding, with the choice

to break ties uniformly at random, so that

$$\begin{aligned} P(\mathbf{0} \rightarrow \mathbf{c}) &= \int_{\mathcal{Y}^n} W_n(\cdot | \boldsymbol{\theta}(\mathbf{0})) \chi_{\{W_n(\cdot | \boldsymbol{\theta}(\mathbf{c})) > W_n(\cdot | \boldsymbol{\theta}(\mathbf{0}))\}} d\mu_n \\ &\quad + \frac{1}{2} \int_{\mathcal{Y}^n} W_n(\cdot | \boldsymbol{\theta}(\mathbf{0})) \chi_{\{W_n(\cdot | \boldsymbol{\theta}(\mathbf{c})) = W_n(\cdot | \boldsymbol{\theta}(\mathbf{0}))\}} d\mu_n. \end{aligned}$$

Note that $P(\mathbf{0} \rightarrow \mathbf{c})$ depends only on the number of occurrences in $\mathbf{c}$ of each nonzero element of G . We will call the type of $\mathbf{c}$, denoted $\mathbf{w}_T(\mathbf{c}) \in \mathbb{N}^{G \setminus \{0\}}$, a counter of such occurrences: $(\mathbf{w}_T(\mathbf{c}))_g = \sum_{i=1}^n \chi_{\{g\}}(\mathbf{c}_i)$. Given $\mathbf{w} \in \mathbb{N}^{G \setminus \{0\}}$, we will use the notation $Q(\mathbf{w})$ to denote $P(\mathbf{0} \rightarrow \mathbf{c})$ for any $\mathbf{c}$ with $\mathbf{w}_T(\mathbf{c}) = \mathbf{w}$.

In what follows, we will use the following two estimations of $P(\mathbf{0} \rightarrow \mathbf{c})$, both involving $w_H(\mathbf{c})$, the Hamming weight of $\mathbf{c}$ (number of nonzero elements). The well-known Bhattacharyya upper bound is

$$\begin{aligned} P(\mathbf{0} \rightarrow \mathbf{c}) &\leq \int_{\mathcal{Y}^n} W_n(\cdot | \boldsymbol{\theta}(\mathbf{0})) \chi_{\{W_n(\cdot | \boldsymbol{\theta}(\mathbf{c})) \geq W_n(\cdot | \boldsymbol{\theta}(\mathbf{0}))\}} d\mu_n \\ &\leq \prod_{i=1}^n \int_{\mathcal{Y}} \sqrt{W(\cdot | \boldsymbol{\theta}(0)) W(\cdot | \boldsymbol{\theta}(c_i))} d\mu \\ &\leq \gamma^{w_H(\mathbf{c})}, \end{aligned}$$

where γ is the (worse) Bhattacharyya noise parameter of the channel defined as

$$\gamma = \max_{g \neq 0} \int_{\mathcal{Y}} \sqrt{W(\cdot | \boldsymbol{\theta}(0)) W(\cdot | \boldsymbol{\theta}(g))} d\mu.$$

A lower bound for pairwise equivocation probability is easily obtained,

$$\begin{aligned} P(\mathbf{0} \rightarrow \mathbf{c}) &\geq \int_{\mathcal{Y}^n} W_n(\cdot | \boldsymbol{\theta}(\mathbf{0})) \chi_{\{W_n(\cdot | \boldsymbol{\theta}(\mathbf{c})) > W_n(\cdot | \boldsymbol{\theta}(\mathbf{0}))\}} d\mu_n \\ &\geq \prod_{i=1}^n \int_{\mathcal{Y}} W(\cdot | \boldsymbol{\theta}(0)) \chi_{\{W(\cdot | \boldsymbol{\theta}(c_i)) > W(\cdot | \boldsymbol{\theta}(0))\}} d\mu \\ &\geq p^{w_H(\mathbf{c})}, \end{aligned}$$

where p is the (worse) equivocation probability of the channel defined as

$$p = \min_{g \neq 0} \int_{\mathcal{Y}} W(\cdot | \boldsymbol{\theta}(0)) \chi_{\{W(\cdot | \boldsymbol{\theta}(c_i)) > W(\cdot | \boldsymbol{\theta}(0))\}} d\mu.$$

Let's see what these definitions give in the examples of G -symmetric channels we have presented.

- **BIOS channels.** The names Bhattacharyya parameter and equivocation probability for γ and p are mostly used only in this context, where there is only one nonzero $g \in G$ and so there is no maximization (resp., minimization) in the definition of γ (resp., p).

For BSC with crossover probability ϵ , if $w_H(\mathbf{c}) = w$, then $P(\mathbf{0} \rightarrow \mathbf{c}) = Q(w)$ with

$$Q(w) = \begin{cases} \sum_{r=\lceil w/2 \rceil}^w \binom{w}{r} \epsilon^r (1-\epsilon)^{w-r} & \text{if } w \text{ is odd} \\ \sum_{r=1+w/2}^w \binom{w}{r} \epsilon^r (1-\epsilon)^{w-r} + \frac{1}{2} \binom{w}{w/2} \epsilon^{w/2} (1-\epsilon)^{w/2} & \text{if } w \text{ is even.} \end{cases}$$

The Bhattacharyya parameter is $\gamma = 2\sqrt{\epsilon(1-\epsilon)}$, while the equivocation probability is $p = \epsilon$.

For BEC with erasure probability ϵ , the only terms in $P(\mathbf{0} \rightarrow \mathbf{c})$ come from breaking ties: if $w_H(\mathbf{c}) = w$, then $P(\mathbf{0} \rightarrow \mathbf{c}) = Q(w) = \frac{1}{2}\epsilon^w$. Here $\gamma = \epsilon$, while $p = 0$.

For BIAWGN channel, see below.

- **S-AWGN channels.** Here, with $S \subset \mathbb{R}^d$ and codewords of length n ,

$$P(\mathbf{0} \rightarrow \mathbf{c}) = \int_{\mathbb{R}^{dn}} W_n(\cdot | \boldsymbol{\theta}(\mathbf{0})) \chi_{\{W_n(\cdot | \boldsymbol{\theta}(\mathbf{c})) > W_n(\cdot | \boldsymbol{\theta}(\mathbf{0}))\}} d\mu_n = \frac{1}{2} \operatorname{erfc} \left(\frac{\|\boldsymbol{\theta}(\mathbf{c}) - \boldsymbol{\theta}(\mathbf{0})\|}{2\sqrt{2\sigma^2}} \right),$$

where $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-t^2} dt$ and $\|\cdot\|$ denotes Euclidean norm.

It is well known that all points of a GU constellation lie on a sphere, and it is usually assumed that constellations have barycenter in the origin, so the radius of the sphere, squared, is the signal energy per transmitted symbol E_s . This remark allows us to find explicit dependence of $P(\mathbf{0} \rightarrow \mathbf{c})$ on the SNR:

$$P(\mathbf{0} \rightarrow \mathbf{c}) = \frac{1}{2} \operatorname{erfc} \left(\frac{\|\boldsymbol{\theta}(\mathbf{c}) - \boldsymbol{\theta}(\mathbf{0})\|}{2\sqrt{E_s}} \sqrt{\frac{E_s}{N_0}} \right),$$

where E_s/N_0 is the SNR per transmitted symbol.

Clearly, having fixed S and θ , $P(\mathbf{0} \rightarrow \mathbf{c})$ depends only on $\mathbf{w}_T(\mathbf{c})$: with the notation $\mathbf{h} = \mathbf{w}_T(\mathbf{c})$, one can write $\|\boldsymbol{\theta}(\mathbf{c}) - \boldsymbol{\theta}(\mathbf{0})\| = \sqrt{\sum_g \mathbf{h}_g \|\theta(g) - \theta(0)\|^2}$.

However, the expression for $P(\mathbf{0} \rightarrow \mathbf{c})$ suggests defining a different weight, which better captures the geometry of S ; it is a rescaled squared Euclidean distance from zero, which we will call Euclidean weight: $w_E(\mathbf{c}) = \|\boldsymbol{\theta}(\mathbf{c}) - \boldsymbol{\theta}(\mathbf{0})\|^2 / (4E_s)$.

One can also compute $\gamma = \max_{g \neq 0} e^{-\|\theta(g) - \theta(0)\|^2 / (8\sigma^2)} = (e^{-E_s/N_0})^{\min_{g \neq 0} w_E(g)}$.

Finally, $p = \min_{g \neq 0} \frac{1}{2} \operatorname{erfc} \left(\frac{\|\theta(g) - \theta(0)\|}{2\sqrt{2\sigma^2}} \right) = \frac{1}{2} \operatorname{erfc} \sqrt{\max_{g \neq 0} w_E(g) \frac{E_s}{N_0}}$.

- **m-ary symmetric channels.** A symmetric channel with alphabet size m and error probability ϵ has

$$P(\mathbf{0} \rightarrow \mathbf{c}) = \sum_{s=1}^w \binom{w}{s} \left[\left(1 - \frac{1}{m-1}\right) \epsilon \right]^{w-s} \sum_{r=\lfloor s/2 \rfloor + 1}^s \binom{s}{r} \left(\frac{\epsilon}{m-1} \right)^r (1-\epsilon)^{s-r} \\ + \frac{1}{2} \sum_{s=0}^{\lfloor w/2 \rfloor} \binom{w}{2s} \left[\left(1 - \frac{1}{m-1}\right) \epsilon \right]^{w-2s} \binom{2s}{s} \left(\frac{\epsilon(1-\epsilon)}{m-1} \right)^s.$$

In this case, $\gamma = \sqrt{\frac{\epsilon}{m-1}} (2\sqrt{1-\epsilon} + (m-2)\sqrt{\frac{\epsilon}{m-1}})$ and $p = \frac{\epsilon}{m-1}$.

2.3. Weights. In this paper, we will deal with different kinds of weight: we have already seen Hamming, type, and Euclidean weight, in relation to the channel's noise, but we will need other weights related to the coding schemes. In this section, we present a general definition and establish some useful properties.

DEFINITION 2.1. A weight on an Abelian group Z consists of a positive integer p and a map $\mathbf{w} : Z \rightarrow \mathbb{N}^p$ satisfying the following properties:

1. $\mathbf{w}(0) = \mathbf{0}$,
2. $|\mathbf{w}(z^1 + z^2)| \leq |\mathbf{w}(z^1)| + |\mathbf{w}(z^2)|$ for every $z^1, z^2 \in Z$,
3. $\{\mathbf{e}_1, \dots, \mathbf{e}_\rho\} \subseteq \mathbf{w}(Z)$ (here $\mathbf{e}_j \in \mathbb{N}^\rho$).

A few considerations on the above definition:

- Item 2 says that the summation in Z cannot create any extra weight.
- Item 3 is a simple minimality assumption which ensures that the full semi-group structure of $\mathbb{N}^\rho$ is used.

Whenever we have a weight $\mathbf{w}$ we will consider its natural extension to vectors by componentwise sum $\mathbf{w} : Z^N \rightarrow \mathbb{N}^\rho$, defined by $\mathbf{w}(\mathbf{z}) = \sum_j \mathbf{w}(z_j)$.

Given $\mathbf{h} \in \mathbb{N}^\rho$, we will use the following notation: $Z_{\mathbf{h}}^N = \{\mathbf{z} \in Z^N : \mathbf{w}(\mathbf{z}) = \mathbf{h}\}$. Moreover, if $\mathbf{h} \in \mathbb{N}^\rho$ we will use the notation

$$\binom{N}{\mathbf{h}} = \begin{cases} \frac{N!}{\mathbf{h}_1! \cdots \mathbf{h}_\rho! (N - |\mathbf{h}|)!} & \text{if } |\mathbf{h}| \leq N \\ 0 & \text{otherwise.} \end{cases}$$

The following result will be useful later.

LEMMA 2.2. *Suppose $\mathbf{w}$ is a weight on Z . For every $\mathbf{h} \in \mathbb{N}^\rho$ we have that*

$$\binom{N}{\mathbf{h}} \leq |Z_{\mathbf{h}}^N| \leq (|Z| N)^{|\mathbf{h}|}.$$

Proof. For $i = 1, \dots, \rho$, let $\eta_i \in Z$ be such that $\mathbf{w}(\eta_i) = \mathbf{e}_i \in \mathbb{N}^\rho$ (which surely exist by point 3 of the definition of weight). The lower bound is trivially true if $|\mathbf{h}| > N$. Otherwise, consider the words in Z^N with support cardinality $|\mathbf{h}|$ made by exactly $\mathbf{h}_j$ times η_j , for $j = 1, \dots, \rho$: there are $\binom{N}{\mathbf{h}}$ such words, and all of them have invariants weight vector $\mathbf{h}$.

The upper bound is clearly true if $\mathbf{h} = \mathbf{0}$. Therefore assume that $\mathbf{h} \neq \mathbf{0}$. For any $\mathbf{z} \in Z_{\mathbf{h}}^N$ consider the subset $\mathcal{J}$ of indices $j \in \{1, \dots, N\}$ for which $z_j \neq 0$. Clearly, $1 \leq |\mathcal{J}| \leq |\mathbf{h}|$. It thus follows that the number of elements in $Z_{\mathbf{h}}^N$ can be upper bound considering all possible subsets $\mathcal{J}$ of cardinality $1 \leq |\mathcal{J}| \leq |\mathbf{h}|$ and all possible elements of Z in the positions in $\mathcal{J}$. In other words

$$|Z_{\mathbf{h}}^N| \leq \sum_{j=1}^{|\mathbf{h}|} \binom{N}{j} |Z|^j \leq |Z|^{|\mathbf{h}|} \sum_{j=1}^{|\mathbf{h}|} \binom{N}{j}.$$

It is now sufficient to use the inequality $\sum_{j=1}^{|\mathbf{h}|} \binom{N}{j} \leq N^{|\mathbf{h}|}$ to obtain the result. $\square$

Hamming weight and type weight are examples of weights always available on any set Z :

- **Hamming weight:** $\rho = 1$, $w_H(z) = 1 - \chi_{\{0\}}(z)$.
- **Type weight:** $\rho = |Z| - 1$, or better $\mathbf{w}_T(z) \in \mathbb{N}^{Z \setminus \{0\}}$, because we prefer indexing the components of $\mathbf{w}_T(z)$ directly by the elements in $Z \setminus \{0\}$ instead of by integers $1, \dots, |Z| - 1$; define $\mathbf{w}_T(z)_a = \chi_{\{a\}}(z)$ for every $a \in Z \setminus \{0\}$.

Clearly, $|\mathbf{w}_T(z)| = w_H(z)$ and, moreover, for any weight $\mathbf{w}$ on Z , it necessarily holds that

$$w_H(\mathbf{z}) \leq |\mathbf{w}(\mathbf{z})| \leq w_{\max} w_H(\mathbf{z}),$$

where $w_{\max} = \max_{z \in Z} |\mathbf{w}(z)|$.

On Abelian groups, it will be particularly important to consider the weights compatible with the algebraic structure, as defined below.

DEFINITION 2.3. *Given an Abelian group U , a distance d on U is called compatible with the group structure of U if $d(u, v) = d(u + w, v + w)$ for all $u, v, w \in U$.*

A weight $\mathbf{w} : U \rightarrow \mathbb{N}$ is called compatible with the group U if there exists a distance d compatible with U such that, for all $u \in U$, $\mathbf{w}(u) = d(u, 0)$.

Notice that if d is a compatible distance on U , the natural extension (by componentwise summation) on U^k remains compatible and it will be denoted with the same symbol d , as well as the associated weight w . Also notice that Hamming and type weights are always compatible with any fixed group U .

2.4. Homomorphic block encoders. Word and symbol error rate. We fix a finite Abelian group Γ and consider transmission on a memoryless Γ -symmetric channel. Given another finite Abelian group U , we define a block encoder of rate k/n over Γ with inputs in U to be any injective group homomorphism $\phi : U^k \rightarrow \Gamma^n$; we define the corresponding code to be the image of the encoder.

We let ξ be a r.v. uniformly distributed on U^k (ξ is the word to be sent) and independent from the channel noise. We let $\hat{\xi}$ be the ML estimate of ξ from the received word y . In this setting, we can clearly define the word error probability of our code in the usual way:

$$P_w(e|\mathbf{u}) = \mathbb{P}(\hat{\xi} \neq \mathbf{u} | \xi = \mathbf{u}), \quad P_w(e) = \mathbb{P}(\hat{\xi} \neq \xi) = \frac{1}{|U|^k} \sum_{\mathbf{u} \in U^k} P_w(e|\mathbf{u}).$$

Our assumptions ensure that the uniform error property holds, i.e., the word error probability does not depend on which word has been sent and, in particular, we can assume that the all-zero word has been sent: $P_w(e) = P_w(e|\mathbf{0})$.

Another interesting property of a code (or, more precisely, of an encoder) is its bit error rate. In our abstract setting, it is more convenient to consider a symbol error rate, where symbols can be, for example, the elements of U or, as we will see, also something smaller. We propose the following definition.

Given a distance d compatible with U and such that $d(u, 0) \neq 0$ for all $u \neq 0$, we define a *symbol error rate* with respect to d as

$$P_s(e|\mathbf{u}) = \sum_{\hat{\mathbf{u}} \in U^k} \frac{d(\hat{\mathbf{u}}, \mathbf{u})}{k\rho_U} \mathbb{P}(\hat{\xi} = \hat{\mathbf{u}} | \xi = \mathbf{u}),$$

where ρ_U is the diameter of U with respect to d . Moreover, we put

$$P_s(e) = \frac{1}{|U|^k} \sum_{\mathbf{u} \in U^k} P_s(e|\mathbf{u}).$$

The compatibility of the distance with U , together with the previous assumptions, ensures that also for $P_s(e)$ the uniform error property holds true:

$$P_s(e) = P_s(e|\mathbf{0}) = \sum_{\hat{\mathbf{u}} \in U^k} \frac{\mathbf{w}(\hat{\mathbf{u}})}{k\rho_U} \mathbb{P}(\hat{\xi} = \hat{\mathbf{u}} | \xi = \mathbf{0}),$$

where $\mathbf{w}$ is the weight associated with the distance d . In this case, $\rho_U = \max_{u \in U} \mathbf{w}(u)$, and we have the inequality

$$P_s(e) \geq \frac{1}{N} \frac{\min_{u \in U, u \neq 0} \mathbf{w}(u)}{\max_{u \in U} \mathbf{w}(u)} P_w(e).$$

When d and w are Hamming distance and weight, respectively, the above definition simply gives the usual symbol error rate, where symbols are elements in U , and if $U = \mathbb{Z}_2$, then this is the classical bit error rate. When $U = \mathbb{Z}_2^a$, in addition to the symbol error rate, we can find also the bit error rate taking as distance the number of different bits (Hamming weight in $(\mathbb{Z}_2^a)^k$ identified with $\mathbb{Z}_2^{ak}$) instead of the number of different symbols.

3. Convolutional encoders over Abelian groups. In this section we will recall some basic facts of the theory of convolutional codes over Abelian groups which will be needed for what follows. Further details can be found in [29, 17, 18, 19] and the reference therein.

3.1. State maps and error events. Let U and Y be two finite Abelian groups. Consider the spaces of sequences $U^{\mathbb{N}}$ and $Y^{\mathbb{N}}$, respectively, both equipped with the componentwise group structure. Convolutional codes for us will be homomorphic maps $\phi : U^{\mathbb{N}} \rightarrow Y^{\mathbb{N}}$ satisfying some properties which are introduced below. In coding theory the only maps between sequence spaces which are really relevant are those which admit a realization through finite state maps.

A (homomorphic) state map η from $U^{\mathbb{N}}$ to $Y^{\mathbb{N}}$ consists of another Abelian group X and of four homomorphisms

$$\begin{aligned} F : X &\rightarrow X, & L : U &\rightarrow X, \\ R : X &\rightarrow Y, & S : U &\rightarrow Y. \end{aligned}$$

X is called the *state space* of the state map, and if X is finite, then the state map is said to be a finite state map. A state map is formally denoted by the quadruple $\eta = (F, L, R, S)$. A finite state map can be pictorially described by a trellis in the usual way: at each time step, we draw vertices corresponding to the elements of X ; then we draw an edge from vertex x at time t to vertex x' at time $t+1$, with input tag u and output label y if and only if $x' = Fx + Lu$ and $y = Rx + Su$.

Given a homomorphic state map η and a state $x \in X$, we can define a map $\eta_x : U^{\mathbb{N}} \rightarrow Y^{\mathbb{N}}$ mapping $\mathbf{u} \in U^{\mathbb{N}}$ into $\mathbf{y} = \eta_x(\mathbf{u})$ computed recursively starting from the initial condition $x_0 = x$, as follows:

$$(3.1) \quad \begin{cases} x_{t+1} &= Fx_t + Lu_t, \\ y_t &= Rx_t + Su_t \end{cases} \quad \forall t \in \mathbb{N}.$$

Explicitly, we can write

$$(3.2) \quad y_t = RF^t x + R \sum_{j=1}^t F^j Lu_{t-j} + Su_t.$$

Notice that η_0 is a homomorphism.

A homomorphic map $\phi : U^{\mathbb{N}} \rightarrow Y^{\mathbb{N}}$ is said to be a *convolutional encoder* if there exists a homomorphic finite state map $\eta = (F, L, R, S)$ such that $\phi = \eta_0$. In this case η is said to be a *state space realization* of ϕ . Given a convolutional encoder $\phi : U^{\mathbb{N}} \rightarrow Y^{\mathbb{N}}$, there may exist many homomorphic finite state maps realizing ϕ . A state map η is said to be a *minimal realization* of ϕ if it has the state space with the minimal number of states among the possible realizations of ϕ . An important consequence of minimality ([13, p. 48] or [30, p. 192]) is the following properties:

- **Observability:** Let $\mathbf{u}', \mathbf{u}'' \in U^{\mathbb{N}}$ and $\mathbf{x}', \mathbf{x}'' \in X^{\mathbb{N}}$ be such that both pairs $(\mathbf{u}', \mathbf{x}')$ and $(\mathbf{u}'', \mathbf{x}'')$ satisfy the first relation of (3.1). Let $\mathbf{y}', \mathbf{y}''$ be the corresponding output sequences. If $\mathbf{u}'_t = \mathbf{u}''_t$ and $\mathbf{y}'_t = \mathbf{y}''_t$ for all $t = 0, \dots, |X| - 1$, then $\mathbf{x}'_0 = \mathbf{x}''_0$ must hold.
- **Reachability:** For any $x, x' \in X$ there exist $t \leq |X| - 1$, $\mathbf{u} \in U^{\mathbb{N}}$, and $\mathbf{x} \in X^{\mathbb{N}}$ satisfying the first relation in (3.1) such that $\mathbf{x}_0 = x$ and $\mathbf{x}_t = x'$. The smallest t for which this condition holds for any $x, x' \in X$ is called the *reachability index* of η which is denoted by ν .

From now on, whenever we are considering a convolutional encoder $\phi : U^{\mathbb{N}} \rightarrow Y^{\mathbb{N}}$, we will assume that an underlying minimal state space representation η has been fixed once and for all: in particular, to any given $\mathbf{u} \in U^{\mathbb{N}}$ and initial state x we can unambiguously associate a state sequence $\mathbf{x} \in X^{\mathbb{N}}$. Whenever the initial state x is not explicitly mentioned, we assume that $x = 0$. Notice, moreover, that x_t depends only on $\mathbf{u}$ up to time $t - 1$.

We now define the key concept of error event.

DEFINITION 3.1. Let $\mathbf{u} \in U^{\mathbb{N}}$ be an input sequence with associated state sequence $\mathbf{x}$. $\mathbf{u}$ is said to be an input error event for ϕ if there exist $t_1 \leq t_2$ such that

- (i) $u_t = 0 \ \forall t < t_1 \text{ and } t > t_2$,
- (ii) $x_t = 0 \ \forall t \leq t_1 \text{ and } t > t_2$,
- (iii) $x_t \neq 0 \ \forall t \in]t_1, t_2]$.

The corresponding codeword $\mathbf{y} = \phi(\mathbf{u})$ is said to be an error event. We call $[t_1, t_2]$ the active window and $t_2 - t_1$ the length of the (input) error event, which we denote by $l(\mathbf{u})$ or by $l(\mathbf{y})$.

The following property shows that the length of an error event cannot grow unbounded. We omit the proof since it is a straightforward generalization of Lemma 20 in [15] (binary case) using the observability property of the minimal realization.

PROPOSITION 3.2. Given a convolutional encoder $\phi : U^{\mathbb{N}} \rightarrow Y^{\mathbb{N}}$, there exists a constant $L > 0$ such that any error event $\mathbf{u}$ has length $l(\mathbf{u}) \leq L (\mathbf{w}_H(\mathbf{u}) + \mathbf{w}_H(\phi(\mathbf{u})))$.

The support of a sequence $\mathbf{u} \in U^{\mathbb{N}}$ is defined by

$$\text{supp}(\mathbf{u}) = \{t \in \mathbb{N} : u_t \neq 0\}.$$

$\mathbf{u}$ is said to have *finite support* if its support has finite cardinality. Notice that the cardinality of the support of a sequence coincides with Hamming weight.

3.2. Laurent series formalism. In many situations the description of a convolutional encoder through a state representation or the corresponding trellis is sufficient and quite appropriate. As in the classical binary case there are also more algebraic but equivalent ways to describe convolutional codes which, on the other hand, turn out to be quite useful in investigating concepts like recursiveness, noncatastrophicity, etc. This is what we are going to do next.

Given a group U , we consider the group of *Laurent series*

$$U((D)) = \left\{ \sum u_k D^k : u_k \in U, \exists k_0 \in \mathbb{Z} u_k = 0 \ \forall k < k_0 \right\}.$$

Inside $U((D))$ there are two relevant subgroups: the polynomials $U[D]$ and the usual formal power series $U[[D]]$.

Relation (3.2), for $x = 0$, can be interpreted as a multiplicative operator (product being defined in the Cauchy style) from $U((D))$ to $G((D))$ with the multiplicative

symbol given by

$$(3.3) \quad \phi(D) = \sum_{j=1}^{\infty} (RF^j L) D^j + S \in \text{Hom}(U, Y)[[D]].$$

$\phi(D)$ is called the *transfer function* associated with ϕ . Conversely, given a generic $\phi(D) \in \text{Hom}(U, Y)[[D]]$, we can ask whether it is the transfer function of a convolutional encoder. The answer is that this is true if and only if $\phi(D)$ is rational. Rationality is defined similarly to the field case. Consider the ring $\mathbb{Z}((D))$ of Laurent series with coefficients in $\mathbb{Z}$. The invertible elements in $\mathbb{Z}((D))$ are those Laurent series whose trailing coefficient is equal to 1 or -1 ; we denote this subset with the symbol $\mathbb{Z}((D))^*$. Given any Abelian group U , $U((D))$ is naturally a $\mathbb{Z}((D))$ -module. We define the submodule of *rational elements* of $U((D))$ as

$$U(D) = \{u(D) \in U((D)) : \exists p(D) \in \mathbb{Z}[D] \cap \mathbb{Z}((D))^*, p(D)u(D) \in U[D]\}.$$

Notice that rational Laurent series can always be represented as a fraction $u(D) = p(D)^{-1}v(D)$ for some suitable polynomials $p(D) \in \mathbb{Z}[D] \cap \mathbb{Z}((D))^*$ and $v(D) \in U[D]$. The assumption on $p(D)$ is exactly to make sure that $1/p(D)$ is a meaningful element of $\mathbb{Z}((D))$. It can be proven that $\phi(D) = \sum_{k=0}^{\infty} \phi_k D^k \in \text{Hom}(U, Y)[[D]]$ is the transfer function of a convolutional encoder if and only if it is rational (see Proposition 5.2 in [17]). Rationality has a useful characterization at the level of the underlying sequence ϕ_k : $\phi(D)$ is rational if and only if ϕ_k is periodic for sufficiently large k . Special types of convolutional encoders are the polynomial ones, namely those for which $\phi(D) \in \text{Hom}(U, Y)[D]$.

In what follows we will often confuse the group sequence $U^{\mathbb{N}}$ with the formal power series $U[[D]]$ through the one-to-one correspondence

$$(u_t)_{t \in \mathbb{N}} \leftrightarrow \sum_t u_t D^t.$$

In particular, $u_0 D^{t_0}$ will often be used to denote the sequence $\mathbf{u}$ which is equal to u_0 at time t_0 and equal to 0, otherwise. Notice that finite support sequences are represented in this way by polynomials in D and polynomial encoders transform polynomials into polynomials.

3.3. Properties of convolutional encoders. In this section we describe how some classical properties can be generalized to our setting; we will need them to analyze our concatenated schemes. Some further properties, specific for the case when the input and output groups are free $\mathbb{Z}_m$ -modules, will be given in the appendix.

3.3.1. Noncatastrophicity. The classical definition of noncatastrophic encoders is the following.

DEFINITION 3.3. *A convolutional encoder $\phi : U^{\mathbb{N}} \rightarrow V^{\mathbb{N}}$ is noncatastrophic if, for all $\mathbf{u} \in U^{\mathbb{N}}$, $w_H(\phi(\mathbf{u})) < \infty$ implies that $w_H(\mathbf{u}) < \infty$.*

An useful remark is that systematic encoders are surely noncatastrophic. Also, noncatastrophic encoders have the following nice characterization (direct consequence of [19, Corollary 1, p. 41]).

PROPOSITION 3.4. *Let $\phi : U^{\mathbb{N}} \rightarrow V^{\mathbb{N}}$ be a convolutional encoder. The following conditions are equivalent:*

1. ϕ is noncatastrophic,
2. ϕ admits a polynomial left inverse,
3. there exists $\mu > 0$ such that, for all $\mathbf{u} \in U^{\mathbb{N}}$, $w_H(\mathbf{u}) \leq \mu w_H(\phi(\mathbf{u}))$ holds.

Notice that condition 2 gives a practical tool for testing if an encoder is noncatastrophic, and it also shows that noncatastrophicity is a property stronger than injectivity. Instead condition 3 is a sort of continuity reformulation.

3.3.2. Recursiveness. Binary convolutional encoders are defined to be recursive when no input word with Hamming weight 1 can give a finite-weight output; this property can be easily generalized to our setting.

DEFINITION 3.5. *Given a weight $\mathbf{w} : U \rightarrow \mathbb{N}^\rho$, a convolutional encoder $\phi : U^\mathbb{N} \rightarrow Y^\mathbb{N}$ is $\mathbf{w}$ -recursive if, for all $\mathbf{u} \in U^\mathbb{N}$ such that $|\mathbf{w}(\mathbf{u})| = 1$, $w_H(\phi(\mathbf{u})) = +\infty$ holds.*

When $\mathbf{w}$ is the Hamming weight, this is the usual definition of recursiveness.

See the appendix for a characterization of recursive encoders on free $\mathbb{Z}_m$ -modules which allows one to easily test for recursiveness.

3.3.3. Small input-weight codewords. All convolutional encoders, including the recursive ones, admit nonzero finite support input sequences whose image also has finite support. This fact is obvious from the rationality property. Indeed, if the transfer function $\phi(D)$ is of type $\phi(D) = p(D)^{-1}\phi'(D)$, where $p(D) \in \mathbb{Z}((D))^* \cap \mathbb{Z}[D]$ and $\phi'(D) \in \text{Hom}(U, Y)[D]$, we can observe that any polynomial input of type $u(D) = p(D)v(D)$ for some $v(D) \in U[D]$ is transformed into another polynomial $\phi(D)u(D) = \phi'(D)v(D)$.

We now present a sharper result which shows how to construct input sequences with support of cardinality 2, whose image has finite support: this will be useful later on.

PROPOSITION 3.6. *Let $\phi : U^\mathbb{N} \rightarrow Y^\mathbb{N}$ be a convolutional encoder, and let $u_1, \dots, u_r \in U$ be such that $\sum u_i = 0$. We can find time instants $t_1, \dots, t_r$ such that given $\mathbf{u} = \sum u_j D^{t_j}$ we have that $\phi(\mathbf{u})$ has finite support.*

Proof. Consider the transfer function $\phi(D) = \sum_k \phi_k D^k$. By rationality we know there exist $k_0 \in \mathbb{N}$ and $T \in \mathbb{N}$ such that $\phi_k = \phi_{k+T}$ for every $k \geq k_0$. Consider now the input sequence $\mathbf{u} = \sum_j u_j D^{(j-1)T}$. We have that

$$\phi(\mathbf{u})_t = \sum_{j=1}^r \phi_{t-(j-1)T} u_j.$$

Notice that if we choose $t \geq k_0 + (r-1)T$, then we can easily obtain that $\phi_{t-(j-1)T} = \phi_t$ for every j , and therefore $\phi(\mathbf{u})_t = 0$. This proves the result. $\square$

From the above result, we obtain as an immediate corollary the following property—well known at least for the binary case.

PROPOSITION 3.7. *Given a recursive convolutional encoder $\phi : U^\mathbb{N} \rightarrow Y^\mathbb{N}$, there exists $\delta \in \mathbb{N}$ such that for any $u \in U$ the input sequence $\mathbf{u} = u - uD^\delta$ is an error event.*

3.3.4. Free distance. In the classical analysis by Benedetto et al. [2], an essential design parameter is the free distance of the outer encoder. When the concatenating group is not the group of all permutations (the classical uniform interleaver), we have to consider a slightly different parameter; instead of taking the minimum Hamming weight among non-zero outer codewords, we minimize some other suitable weight.

DEFINITION 3.8. *Given a convolutional encoder $\phi : U^\mathbb{N} \rightarrow Y^\mathbb{N}$ and a weight $\mathbf{w} : Y^\mathbb{N} \rightarrow \mathbb{N}^\rho$, we define the $\mathbf{w}$ -free distance of ϕ to be*

$$d_f(\phi, \mathbf{w}) = \min\{|\mathbf{w}(\mathbf{c})| : \mathbf{c} = \phi(\mathbf{u}), \mathbf{u} \in U^\mathbb{N}, \mathbf{u} \neq \mathbf{0}\}.$$

The classical free distance is the w_H -free distance.

3.4. Terminated convolutional encoders. Suppose $\phi : U^{\mathbb{N}} \rightarrow Y^{\mathbb{N}}$ is a convolutional encoder with minimal state space X . We now define the terminated block codes associated with ϕ as follows.

Fix $N \in \mathbb{N}^*$. Given a vector $\mathbf{u} = (u_0, \dots, u_{N-1}) \in U^N$, let x_N be the corresponding state at time N . Because of the reachability condition it is possible to find input elements $\tilde{u}_N, \dots, \tilde{u}_{N+\nu-1}$ such that the state at time $N + \nu - 1$ is equal to 0. This input string may not be unique and we assume we have fixed a specific one as a function of the terminal state x_N we had reached in such a way that the mapping

$$x_N \mapsto (\tilde{u}_N, \dots, \tilde{u}_{N+\nu-1})$$

is a homomorphism. It is a straightforward algebraic verification that this is always indeed possible. Given $\mathbf{u} = (u_0, \dots, u_{N-1}) \in U^N$, we now consider the associated input sequence $\tilde{\mathbf{u}} = (u_0, \dots, u_{N-1}, \tilde{u}_N, \dots, \tilde{u}_{N+\nu-1}, 0, 0, \dots)$. We then define the *N-terminated block encoder* as

$$\phi^N : U^N \rightarrow Y^{N+\nu}, \quad \phi^N(\mathbf{u}) = \phi(\tilde{\mathbf{u}})|_{[0, N+\nu-1]}.$$

For the assumptions made, ϕ^N is also a homomorphism. $\mathcal{C}^N = \text{Im } \phi^N$ is called the *N-block code* associated to ϕ^N .

An input vector $\mathbf{u} \in U^N$ is an input error event for ϕ^N if $\tilde{\mathbf{u}} \in U^{\mathbb{N}}$ is an input error event for ϕ . In this case $\mathbf{c} = \phi^N(\mathbf{u})$ is called an error event for ϕ^N . Suppose the active window of $\tilde{\mathbf{u}}$ is equal to $[t_1, t_2]$. Then, the (input) error event is said to be *regular* if $t_2 \leq N$, otherwise it is called *terminated*. For a terminated error event, we define its length to be $N - t_1$.

Notice that any codeword $\mathbf{c} \in \mathcal{C}^N$ can be written as $\mathbf{c} = \sum_{j=1}^{n+1} \mathbf{c}_j$, where $\mathbf{c}_j$ are regular error events for $j = 1, \dots, n$ and $\mathbf{c}_{n+1}$ is either zero or a terminating error event and the active windows of all of these events are disjoint. We will use the notation $n(\mathbf{c})$ to denote the number of regular error events in the above decomposition of $\mathbf{c}$. Also notice that the above decomposition is unique, up to a permutation of the regular error events.

Some codewords have a decomposition in error events which is the same up to shifts of their error events, and for this reason they share many important properties. More formally, we propose the following definition:

- Two error events $\mathbf{c} = \phi^N(\mathbf{u})$ and $\mathbf{c}' = \phi^{N'}(\mathbf{u}')$ (notice that possibly $N \neq N'$) are said to be *shift equivalent* if the corresponding extended inputs $\tilde{\mathbf{u}}, \tilde{\mathbf{u}}' \in U^{\mathbb{N}}$ differ only by a shift.
- Two codewords $\mathbf{c} = \phi^N(\mathbf{u})$ and $\mathbf{c}' = \phi^{N'}(\mathbf{u}')$ are said to be *shift equivalent* if there exist error event decompositions $\mathbf{c} = \sum_{j=1}^{n+1} \mathbf{c}_j$ and $\mathbf{c}' = \sum_{i=1}^{n'+1} \mathbf{c}'_i$ such that $\mathbf{c}_i$ and $\mathbf{c}'_i$ are shift equivalent for all i .

Notice that, given two shift equivalent codewords $\mathbf{c}$ and $\mathbf{c}'$, it is clear that $n(\mathbf{c}) = n(\mathbf{c}')$ and, moreover, given a weight $\mathbf{w}$ on the alphabet Y , $\mathbf{w}(\mathbf{c}) = \mathbf{w}(\mathbf{c}')$.

Remark 3.9. Now we want to underline a property which is somehow similar to an inclusion of $\mathcal{C}^N$ in $\mathcal{C}^{N'}$ for $N \leq N'$ (while strictly speaking an inclusion cannot occur, as the two codes are subsets of different spaces). If $N \leq N'$, then for all $\mathbf{c} \in \mathcal{C}^N$ we can construct $\mathbf{c}' \in \mathcal{C}^{N'}$ such that $\mathbf{c}$ and $\mathbf{c}'$ are shift equivalent, by properly adding zeros.

3.5. Enumerating functions and growth estimates. A fundamental concept for all encoders is the so-called weight enumerating function, since it is well known to play a basic role in all performance evaluations. While in the binary case,

there is only one possible weight to be considered, namely the Hamming one, in our setting many choices are possible, and we will need to consider different possibilities in later sections. We start by defining the basic possibility based on the Hamming weights in the input and output groups.

DEFINITION 3.10. *Given a convolutional encoder $\phi : U^{\mathbb{N}} \rightarrow Y^{\mathbb{N}}$, consider its associated N -terminated block encoder $\phi^N : U^N \rightarrow Y^{N+\nu}$. Define its input/output support enumerating coefficients as*

$$\Lambda_{w,d,n}^N = |\{\mathbf{u} \in U^N : w_H(\mathbf{u}) = w, w_H(\phi^N(\mathbf{u})) = d, n(\phi^N(\mathbf{u})) = n\}|.$$

In some cases, we will need to replace the Hamming weight with other possible weights in the input and the output. We will use the notation $\Lambda_{\mathbf{w},\mathbf{d},n}^N$ to denote enumerating coefficients relative to some specified input weight $\mathbf{w}$ and output weight $\mathbf{d}$.

The following proposition gives a growth estimation for input/output support enumerating coefficients, which will allow us to have general bounds (even if quite loose) on all of the different weight enumerators. We omit the proof, which is a straightforward generalization of Proposition 10 in [15].

PROPOSITION 3.11. *There exist two positive constants a and b such that*

$$\Lambda_{w,d,n}^N \leq \binom{N+n}{n} a^w b^d.$$

4. Serial ensembles.

4.1. Serial interconnections. We now precisely define the serial interconnected schemes we are going to consider. We start with a Γ -symmetric channel. We also fix the *input* Abelian group U . All encoders we will consider will be driven by words on U and will output symbols in Γ . The interconnection will take place through a third Abelian group, say Y , called the *interconnection* group. We now fix two convolutional encoders denoted, respectively, the *outer* and *inner* encoder:

$$\phi_o : U^{\mathbb{N}} \rightarrow (Y^r)^{\mathbb{N}}, \quad \phi_i : (Y^s)^{\mathbb{N}} \rightarrow (\Gamma^l)^{\mathbb{N}}.$$

Denote by ν_o and ν_i the reachability indices of ϕ_o and ϕ_i , respectively, and define the set $\mathcal{N} = \{N \in \mathbb{N}^* : s|r(N + \nu_o)\}$. Consider now the terminations for $N \in \mathcal{N}$:

$$\phi_o^N : U^N \rightarrow Y^{r(N+\nu_o)}, \quad \phi_i^N : Y^{sM_N} \rightarrow \Gamma^{l(M_N+\nu_i)},$$

where $sM_N = r(N + \nu_o)$. We now fix, for every $N \in \mathbb{N}^*$, a subgroup $G_N \subseteq \text{Aut}(Y^{r(N+\nu_o)})$. The triple $(\phi_o, \phi_i, (G_N)_{N \in \mathcal{N}})$ is said to be a *serial interconnected ensemble*. The asymptotic rate of the serial interconnected ensemble above is defined by the product

$$R = \frac{\log |U|}{r} \frac{s}{l} \text{ bits per channel use.}$$

To the serial interconnected ensemble above we can associate a random sequence of encoders and codes as follows. Define Π_N to be a r.v. uniformly distributed over G_N , and consider the corresponding homomorphic encoder $\Phi^N = \phi_i^N \circ \Pi_N \circ \phi_o^N$ and group code $\mathcal{C}^N = \text{Im}(\Phi^N)$, which are called, respectively, the *random encoder* and the *random code* associated with the given ensemble.

The following picture describes the above construction:

$$U^N \xrightarrow{\quad} \boxed{\phi_o^N} \xrightarrow{Y^{r(N+\nu_o)}} \boxed{\Pi_N} \xrightarrow{Y^{sM_N}} \boxed{\phi_i^N} \xrightarrow{\Gamma^{l(M_N+\nu_i)}}.$$

In what follows we will denote by $\mathbb{P}$ and $\mathbb{E}$ probability and expected value, respectively, made with respect to the probabilistic space underlying the sequence Π_N . We will also use the notation $\overline{P_w(e)}$ and $\overline{P_s(e)}$, respectively, for the average word and symbol error probabilities.

4.2. Regular ensembles. Our aim is to give asymptotic results for $\overline{P_w(e)}$ and $\overline{P_s(e)}$ when $N \rightarrow \infty$, keeping fixed the constituent encoders. To do so, we need to make further assumptions on the groups G_N : roughly, we need to enforce some compatibility among the groups as N varies, and to avoid that the number of invariants of the group action grows with N . Following [15] we propose the following definition.

DEFINITION 4.1. *The sequence of groups G_N (and the corresponding ensemble) is said to be regular if there exists a weight $\mathbf{w}_G : Y \rightarrow \mathbb{N}^\rho$ such that, for every N and for all $\mathbf{y}, \mathbf{z} \in Y^{r(N+\nu_0)}$, it holds that*

$$\mathbf{w}_G(\mathbf{y}) = \mathbf{w}_G(\mathbf{z}) \Leftrightarrow \exists \sigma \in G_N : \sigma \mathbf{y} = \mathbf{z}.$$

$\mathbf{w}_G(\mathbf{y})$ will be called the invariants weight vector of $\mathbf{y} \in Y^{r(N+\nu_0)}$.

Property of regularity simply says that all actions of the groups G_N on the sets $Y^{r(N+\nu_0)}$ can be described through a finite (constant) family of invariants: the ρ components of the weight $\mathbf{w}_G$. We will use the notation $Y_{\mathbf{h}}^L = \{\mathbf{x} \in Y^L : \mathbf{w}_G(\mathbf{x}) = \mathbf{h}\}$. Moreover, we denote by $G_N(\mathbf{y}, \mathbf{z})$ the subset of elements in G_N mapping $\mathbf{y}$ to $\mathbf{z}$. Using standard results on group actions (the class formula) [26], we can show the following remark.

Remark 4.2.

$$\frac{|G_N(\mathbf{u}, \mathbf{v})|}{|G_N|} = \begin{cases} 0 & \text{if } \mathbf{w}_G(\mathbf{u}) \neq \mathbf{w}_G(\mathbf{v}) \\ 1/|Y_{\mathbf{h}}^{r(N+\nu_0)}| & \text{if } \mathbf{w}_G(\mathbf{u}) = \mathbf{w}_G(\mathbf{v}) = \mathbf{h}. \end{cases}$$

This technical result will be needed later.

LEMMA 4.3. *Assume that $\mathbf{y}, \mathbf{z} \in Y^{r(N+\nu_0)}$ are such that for every index $i \in \{0, \dots, N + \nu_0 - 1\}$, $\mathbf{y}_i \neq \mathbf{0}$ yields $\mathbf{z}_i = \mathbf{0}$. Then, given any $\sigma \in G_N$ and given any i we have that $(\sigma \mathbf{y})_i \neq \mathbf{0} \Rightarrow (\sigma \mathbf{z})_i \neq (\sigma \mathbf{y})_i$.*

Proof. Notice that $|\mathbf{w}_G(\sigma \mathbf{y} + \sigma(-\mathbf{z}))| = |\mathbf{w}_G(\mathbf{y} - \mathbf{z})| = |\mathbf{w}_G(\mathbf{y})| + |\mathbf{w}_G(-\mathbf{z})|$. On the other hand, if $\sigma \mathbf{y}$ and $\sigma \mathbf{z}$ were equal in an index where they are not equal to $\mathbf{0}$, then by point 2 of Definition 2.1 we would have

$$|\mathbf{w}_G(\sigma \mathbf{y} + \sigma(-\mathbf{z}))| < |\mathbf{w}_G(\sigma \mathbf{y})| + |\mathbf{w}_G(\sigma(-\mathbf{z}))| = |\mathbf{w}_G(\mathbf{y})| + |\mathbf{w}_G(-\mathbf{z})|.$$

This ends the proof. $\square$

We now present two fundamental examples of regular actions.

1. *Symbol permutation.* In this case we simply take $G_N = S_{r(N+\nu_0)}$, i.e., the full symmetric group acting on $Y^{r(N+\nu_0)}$ by standard permutation. In this case the invariant weight is the type weight: $\rho = |Y| - 1$ and $\mathbf{w}_G(y) \in \mathbb{N}^\rho$ by $(\mathbf{w}_G(y))_a = \chi_a(y)$ as a varies in $Y \setminus \{0\}$. Notice that

$$|Y_{\mathbf{h}}^{r(N+\nu_0)}| = \binom{r(N+\nu_0)}{\mathbf{h}}.$$

2. *Separate channels symbol permutation.* Assume $Y = Y_1 \times Y_2$, and assume $G_{N,1}$ and $G_{N,2}$ are sequences of groups acting regularly on $Y_1^{r(N+\nu_0)}$ and $Y_2^{r(N+\nu_0)}$, respectively, with invariant weights $\mathbf{w}_G^1 : Y_1 \rightarrow \mathbb{N}^{\rho_1}$ and $\mathbf{w}_G^2 : Y_2 \rightarrow \mathbb{N}^{\rho_2}$, respectively. Then, we can consider a regular action given by $G_N = G_{N,1} \times G_{N,2}$ acting

componentwise on $Y_1 \times Y_2$. Its invariant weight is given by $\mathbf{w}_G : Y_1 \times Y_2 \rightarrow \mathbb{N}^{\rho_1 + \rho_2}$, $\mathbf{w}_G(y_1, y_2) = (\mathbf{w}_G^1(y_1), \mathbf{w}_G^2(y_2))$. Notice that in this case

$$\left| Y_{\mathbf{h}_1, \mathbf{h}_2}^{r(N+\nu_o)} \right| = \binom{r(N+\nu_o)}{\mathbf{h}_1} \binom{r(N+\nu_o)}{\mathbf{h}_2}.$$

4.3. Examples of serial ensembles. Below we assume that $\Gamma = \mathbb{Z}_m$.

- *Repeat-convolute codes.* We choose $U = Y = \mathbb{Z}_m$ and $\phi_o = \text{Rep}_r : \mathbb{Z}_m^{\mathbb{N}} \rightarrow (\mathbb{Z}_m^r)^{\mathbb{N}}$ to be the r -repetition encoder $\text{Rep}_r(\mathbf{u})_t = (u_t, \dots, u_t)$. We let $\phi_i : (\mathbb{Z}_m^s)^{\mathbb{N}} \rightarrow (\mathbb{Z}_m^s)^{\mathbb{N}}$ be a rate-1 noncatastrophic convolutional encoder. Finally, we choose for the coupling interleavers the symbol permutation groups $G_N = S_{rN}$. The corresponding invariant weight is thus the type weight $\mathbf{w}_T : \mathbb{Z}_m^r \rightarrow \mathbb{N}^{m-1}$, where $(\mathbf{w}_T(\mathbf{y}))_j$ is the number of elements equal to j in the vector $\mathbf{y} \in \mathbb{Z}_m^r$. The rate of the scheme is

$$R = \frac{\log m}{r} \text{ bits per channel use.}$$

For this ensemble, we will need the assumption that ϕ_i is $\mathbf{w}$ -recursive, which is the same as requiring that it is $\mathbf{w}_H$ -recursive.

- *Structured LDPC codes.* We choose $U = \mathbb{Z}_m$, $Y = \mathbb{Z}_m^c \times \mathbb{Z}_m$, and the systematic encoder $\phi_o : \mathbb{Z}_m^{\mathbb{N}} \rightarrow (\mathbb{Z}_m^c \times \mathbb{Z}_m)^{\mathbb{N}}$, $\phi_o(\mathbf{u}) = (\text{Rep}_c(\mathbf{u}), \mathbf{u})$. Instead ϕ_i is itself the serial interconnection of two encoders. We consider $\text{Sum}_d : (\mathbb{Z}_m^d)^{\mathbb{N}} \rightarrow \mathbb{Z}_m^{\mathbb{N}}$ defined by $\text{Sum}_d(\mathbf{y}) = (y_1 + \dots + y_d, y_{d+1} + \dots + y_{2d}, \dots)$, and we take a $\mathbf{w}_H$ -recursive noncatastrophic rate-1 convolutional encoder $\psi : \mathbb{Z}_m^{\mathbb{N}} \rightarrow \mathbb{Z}_m^{\mathbb{N}}$. Finally, we take $\phi_i : (\mathbb{Z}_m^d \times \mathbb{Z}_m)^{\mathbb{N}} \rightarrow (\mathbb{Z}_m \times \mathbb{Z}_m)^{\mathbb{N}}$, defined by

$$\phi_i(\mathbf{y}^1, \mathbf{y}^2) = ((\psi \circ \text{Sum}_d)(\mathbf{y}^1), \mathbf{y}^2).$$

When taking the truncated versions of these encoders, we must make sure to have suitable lengths, so we take $\phi_o^N : \mathbb{Z}_m^{dN} \rightarrow \mathbb{Z}_m^{cdN} \times \mathbb{Z}_m^{dN}$ and $\phi_i^N : \mathbb{Z}_m^{cdN} \times \mathbb{Z}_m^{dN} \rightarrow \mathbb{Z}_m^{cN+\nu_\psi} \times \mathbb{Z}_m^{dN} = \Gamma^{(c+d)N+\nu_\psi}$. Therefore, the design rate of the serial encoder $\phi^N = \phi_i^N \circ \Pi_N \circ \phi_o^N$ is $R = \log m \frac{d}{c+d}$.

As an interconnection group, we choose the separated channels symbol permutation $G_N = S_{cdN} \times S_{dN}$.

This family of codes is a generalization of the repeat-convolute codes: the additional summator Sum_d is the same as the grouping factor introduced in irregular repeat accumulate codes [27].

We can easily construct a parity-check matrix for the code $\mathcal{C}^N = \text{Im}(\Phi^N) \subseteq \mathbb{Z}_m^{(c+d)N+\nu_\psi}$, which is sparse and has a structured and a random part, so that we have a structured LDPC ensemble, generalizing staircase LDPC codes. In fact, notice that

$$\begin{aligned} (\mathbf{c}^1, \mathbf{c}^2) \in \mathcal{C}^N &\Leftrightarrow \mathbf{c}^1 = \psi^N \circ \text{Sum}_d \circ \pi_N^1 \circ \text{Rep}_c \circ (\pi_N^2)^{-1}(\mathbf{c}^2) \\ &\Leftrightarrow (\psi^N)^{-1}(\mathbf{c}^1) = \text{Sum}_d \circ \pi_N^1 \circ \text{Rep}_c \circ (\pi_N^2)^{-1}(\mathbf{c}^2). \end{aligned}$$

It is clear that the permutation π_N^2 does not play any essential role: we needed it only to fit this scheme in our assumptions, but we can take it out without changing the performance of the scheme.

Note that the noncatastrophicity of ϕ_i is needed to make the syndrome matrix low density, i.e., with a number of nonzero elements per row and per column

which is small and does not grow with N . More precisely, the matrix $H_2 = \text{Sum}_d \pi_N^1 \text{Rep}_d$ is a random low density matrix with entries in $\{0, 1\}$, depending only on c , d , and π_N , with at most c elements equal to 1 in each column and at most d on any row. Instead $H_1 = (\psi^N)^{-1}$ depends on the choice of ψ and is also low density, having a number of nonzero elements per row and per column at most equal to the degree of the polynomial $\psi^{-1}(D)$.

5. Main result: Interleaver gain. The well-known analysis by Benedetto et al. [2] showed an interleaver gain, in the sense that average error probability is asymptotically vanishing when the interleaver length grows. Their result holds true under the assumption that both constituent encoders are systematic recursive convolutional encoders and that the free distance of the outer encoder was $d_f^o \geq 2$ to ensure $\overline{P_b(e)} \rightarrow 0$ and $d_f^o \geq 3$ to have also $\overline{P_w(e)} \rightarrow 0$.

In this section, we will comment on how the classical assumptions on the constituent encoders can be adapted to our setting, and we will state our results on the interleaver gain; all of the proofs will be given in section 6. Proofs' techniques are inspired by [15], but they cannot be seen as consequences of the results in [15]: the fact of working with serial interconnections and in a nonbinary situation makes the derivation quite different and the combinatorics more involved.

From now on, we will be always considering a regular serial ensemble (see Definition 4.1) with outer encoder $\phi_o : U^{\mathbb{N}} \rightarrow (Y^r)^{\mathbb{N}}$ and inner encoder $\phi_i : (Y^s)^{\mathbb{N}} \rightarrow (\Gamma^l)^{\mathbb{N}}$, and with a family of interconnection groups (G_N) with invariants weight vector $\mathbf{w}_G$. The symbol error probability will be with respect to a fixed weight on the input group U , denoted $\mathbf{w}_{\text{in}}$, with the requirement that $\mathbf{w}_{\text{in}}(u) \neq 0$ for all $u \neq 0$.

First of all, we have to generalize the assumptions about the constituent encoders introduced in [2].

When considering one single convolutional encoder, noncatastrophicity is usually needed to ensure good asymptotic properties. However, when dealing with a concatenated scheme the assumption that all constituent encoders are noncatastrophic can be slightly weakened, as it was already recognized, for example in [15] and [25] (in the latter, the authors consider serial schemes where the inner encoder is heavily punctured and becomes not injective). The essential assumption is that the overall scheme is noncatastrophic, and this can be obtained by asking classical noncatastrophicity of the outer encoder and a weaker property of the inner encoder: ϕ_i must be noncatastrophic when restricted to the inputs he will actually receive, i.e., the permuted outer codewords.

When we are dealing with ensembles of concatenated codes, each code of the ensemble must be noncatastrophic in the sense specified above. This leads to the following definition.

DEFINITION 5.1. *A regular serial ensemble with constituent encoders $\phi_o : U^{\mathbb{N}} \rightarrow (Y^r)^{\mathbb{N}}$ and $\phi_i : (Y^s)^{\mathbb{N}} \rightarrow (\Gamma^l)^{\mathbb{N}}$ and regular group family (G_N) is concatenatedly noncatastrophic if there exist two positive constants μ_o and μ_i such that for all $N \in \mathbb{N}^*$ and for all $\mathbf{u} \in U^{\mathbb{N}}$:*

1. $w_H(\mathbf{u}) \leq \mu_o |\mathbf{w}_G(\phi_o^N(\mathbf{u}))|$,
2. for all $\pi \in G_N$, $|\mathbf{w}_G(\phi_o^N(\mathbf{u}))| \leq \mu_i w_H(\phi_i^N \circ \pi \circ \phi_o^N(\mathbf{u}))$.

Notice that 1. is equivalent to the noncatastrophicity of ϕ_o (see Proposition 3.4). In the examples introduced in section 4.3, we have an example where both encoders are noncatastrophic (repeat-convolute), and an example where only concatenated noncatastrophicity is true (structured LDPC). In fact, in this second example, noncatastrophicity of ψ ensures sparsity of the parity-check matrix, but due to noninjectivity

of Sum_d the inner encoder is indeed catastrophic; overall noncatastrophicity of the concatenated scheme is ensured by the systematic branch.

Concerning the other classical assumptions on the constituent encoders ($d_f^o \geq 3$ and recursiveness of ϕ_i), clearly they must be restated considering the suitable connecting weight $\mathbf{w}_G$ instead of Hamming weight, using the definitions introduced in section 3. However, we will comment later in this section why these assumptions are sufficient and not necessary to obtain some interleaver gain, and how they can be weakened.

Now we will introduce some useful definitions, and then state the interleaver gain result, which will answer the questions: Is the average error probability asymptotically vanishing when the interleaver length grows to infinity? If so, how fast is the decay? From now on, we will always assume that we are considering a concatenatedly noncatastrophic ensemble.

Let $\mathcal{C}_o^N = \phi_o^N(U^N) \subseteq Y^{r(N+\nu_o)}$ be the outer block code, and let

$$H = \{\mathbf{w}_G(\mathbf{c}) : \mathbf{c} \in \mathcal{C}_o^N \text{ for some } N, \mathbf{c} \neq \mathbf{0}\}.$$

Notice that, with this notation, (2) in Definition 5.1 is equivalent to

$$\forall N \in \mathbb{N}^*, \forall \mathbf{h} \in H, \forall \mathbf{c} \in Y^{r(N+\nu_o)} \text{ such that } \mathbf{w}_G(\mathbf{c}) = \mathbf{h}, |\mathbf{h}| \leq \mu_i |\mathbf{w}_T(\phi_i^N(\mathbf{c}))|.$$

Given $\mathbf{h} \in H$, we look at the decomposition of codewords in error events, as defined in sections 3.1 and 3.4, and we define

- $n_o(\mathbf{h}) = \max\{n(\mathbf{c}) \text{ such that } \exists N, \exists \mathbf{c} \in \mathcal{C}_o^N : \mathbf{w}_G(\mathbf{c}) = \mathbf{h}\},$
- $n_i(\mathbf{h}) = \max\{n(\mathbf{x}) \text{ s.t. } \exists N, \exists \mathbf{u} \in Y^{r(N+\nu_o)} : \mathbf{x} = \phi_i^N(\mathbf{u}), \mathbf{w}_G(\mathbf{u}) = \mathbf{h}\},$
- $f(\mathbf{h}) = 1 + |\mathbf{h}| - n_o(\mathbf{h}) - n_i(\mathbf{h}).$

Remark 5.2. Both maxima in the above definition are well defined, since we clearly have $n(\mathbf{c}) \leq |\mathbf{h}|$, $n(\mathbf{x}) \leq |\mathbf{h}|$. Moreover, notice that, because of Remark 3.9 in section 3.4, the sequence of sets $\{n(\mathbf{c}) \text{ such that } \exists \mathbf{c} \in \mathcal{C}_o^N : \mathbf{w}_G(\mathbf{c}) = \mathbf{h}\}$ is increasing in N . Therefore there exists $\bar{N}(\mathbf{h})$ such that

$$n_o(\mathbf{h}) = \max\{n(\mathbf{c}) \text{ such that } \exists \mathbf{c} \in \mathcal{C}_o^{\bar{N}(\mathbf{h})} : \mathbf{w}_G(\mathbf{c}) = \mathbf{h}\}.$$

An analogous statement holds true for $n_i(\mathbf{h})$.

It is also clear that the maximum defining $n_o(\mathbf{h})$ can always be obtained with a codeword which admits only regular error events, while this is not necessarily true for $n_i(\mathbf{h})$.

Finally, we define

$$(5.1) \quad \alpha = \inf\{f(\mathbf{h}), \mathbf{h} \in H\}.$$

Notice that the function f takes values in $\mathbb{Z}$ and H is nonempty, so either $\alpha = -\infty$ or $\alpha = \min\{f(\mathbf{h}), \mathbf{h} \in H\}$.

Our main result (formally stated in Theorem 5.4) is that, for sufficiently good channels, if $\alpha \geq 1$, then

$$\overline{P_s(e)} \asymp N^{-\alpha} \quad \text{and} \quad \overline{P_w(e)} \asymp N^{-\alpha+1} \quad \text{for } N \rightarrow \infty.$$

We will give sufficient conditions ensuring that $\alpha \geq 1$, enforcing some properties on the constituent encoders (Propositions 5.5 and 5.7).

In addition to the decay of $\overline{P_s(e)}$ and $\overline{P_w(e)}$ asymptotically in N , we want to underline the dependence of the error probability on the channel's SNR, following

the steps of Benedetto et al. [2] and looking for a design parameter analogous of the classical effective free distance.

We define the set of the vectors $\mathbf{h}$ minimizing $f(\mathbf{h})$:

$$\mathcal{H} = \{\mathbf{h} \in H : f(\mathbf{h}) = \alpha\},$$

and we define the effective distance

$$q^* = \max_{\mathbf{h} \in \mathcal{H}} \{q^*(\mathbf{h})\},$$

where

$$q^*(\mathbf{h}) = \max\{P(\mathbf{0} \rightarrow \mathbf{x}) : \exists N, \exists \mathbf{u} \in Y^{r(N+\nu_o)} : \mathbf{x} = \phi_i^N(\mathbf{u}), \mathbf{w}_G(\mathbf{u}) = \mathbf{h}, n(\mathbf{x}) = n_i(\mathbf{h})\}.$$

Remark 5.3. We can prove that maxima in the definitions of $q^*(\mathbf{h})$ and q^* are well defined. In principle the number of $\mathbf{x}$ involved in the maximum defining $q^*(\mathbf{h})$ is infinite. However, we can always restrict the search to a finite set in the following way. As a first step, we can find an upper bound on the values of $P(\mathbf{0} \rightarrow \mathbf{x})$ to consider trivially by computing $\bar{q} = P(\mathbf{0} \rightarrow \bar{\mathbf{x}})$ for one admissible $\bar{\mathbf{x}}$. Then we restrict our search to the set

$$\bar{X} = \{\mathbf{x} : P(\mathbf{0} \rightarrow \mathbf{x}) \geq \bar{q} \text{ and } \exists N, \exists \mathbf{u} \in Y^{r(N+\nu_o)} : \mathbf{x} = \phi_i^N(\mathbf{u}), \mathbf{w}_G(\mathbf{u}) = \mathbf{h}, n(\mathbf{x}) = n_i(\mathbf{h})\}.$$

Now note that $P(\mathbf{0} \rightarrow \mathbf{x}) \geq \bar{q}$ implies $\gamma^{\mathbf{w}_H(\mathbf{x})} \geq \bar{q}$; i.e., $\mathbf{w}_H(\mathbf{x}) \leq \log \bar{q} / \log \gamma$. Now, by Proposition 3.2, we can bound the length of all error events in the decomposition of $\mathbf{x} \in \bar{X}$. This implies that, up to shift equivalence, the family of all possible error events appearing in $\mathbf{x} \in \bar{X}$ is finite. Therefore, $\bar{X}$, up to shift equivalence, is finite. The same argument also applies to q^* .

Later, we will also see that under suitable assumptions $\mathcal{H}$ is a finite set (Proposition 5.6).

Using the definition of q^* , we can state the interleaver gain result in a stronger way that underlines, additionally to the decay with N , also the dependence on the channel.

THEOREM 5.4. *Consider a regular and concatenatedly noncatastrophic serial ensemble $(\phi_o, \phi_i, (G_N)_{N \in \mathcal{N}})$ corresponding to the encoding scheme*

$$\xrightarrow{U^N} \boxed{\phi_o^N} \xrightarrow{Y^{r(N+\nu_o)}} \boxed{\pi_N} \xrightarrow{Y^{sM_N}} \boxed{\phi_i^N} \xrightarrow{\Gamma^{l(M_N+\nu_i)}}.$$

If $\alpha \geq 1$, then there exist positive constants c, c_1, c_2 , and γ_0 , depending only on ϕ_o, ϕ_i , and (G_N) , such that for all Γ -symmetric channels with Bhattacharyya parameter $\gamma < \gamma_0$

$$c q^* N^{-\alpha} \leq \overline{P_w(e)} \leq c_1 q^* c_2^{(\log q^* / \log \gamma)} N^{-\alpha} + O(N^{-\alpha-1}).$$

Moreover, for a given input weight $\mathbf{w}_{\text{in}}$ (compatible with U and satisfying $\mathbf{w}_{\text{in}}(u) \neq 0$ for all $u \neq 0$),

$$c \frac{\mathbf{w}_{\text{in}}^{\max}}{\mathbf{w}_{\text{in}}^{\min}} q^* N^{-\alpha+1} \leq \overline{P_s(e)} \leq c_1 \frac{\mathbf{w}_{\text{in}}^{\max}}{\mathbf{w}_{\text{in}}^{\min}} q^* c_2^{(\log q^* / \log \gamma)} N^{-\alpha+1} + O(N^{-\alpha}),$$

where $\mathbf{w}_{\text{in}}^{\max} = \max_{u \in U} \mathbf{w}_{\text{in}}(u)$ and $\mathbf{w}_{\text{in}}^{\min} = \min_{u \in U \setminus \{0\}} \mathbf{w}_{\text{in}}(u)$.

The terms q^* in the lower bound and $q^* c_2^{(\log q^* / \log \gamma)}$ in the upper bound describe the behavior of $\overline{P_s(e)}$ and $\overline{P_w(e)}$ with respect to the channel's noise. Note that $q^* =$

$P(\mathbf{0} \rightarrow \mathbf{c})$ for some word $\mathbf{c}$, so that, with the notation $w^* = w_H(\mathbf{c})$, we have $q^* \leq \gamma^{w^*}$, and $\frac{\log q^*}{\log \gamma} \leq w^*$. Theorem 5.4 holds true for any Γ -symmetric channel with $\gamma \leq \gamma_0$. However, it is particularly relevant for families of channels where $\gamma \rightarrow 0$ by preserving the geometry of the channel, thus having constant w^* for all the family: in this case, q^* fully describes the decay of error probability when $\gamma \rightarrow 0$ (i.e., SNR grows to infinity). For example, one can consider BSC with crossover probability ϵ and let $\epsilon \rightarrow 0$, or an S -AWGN channel, fix the shape of the constellation and let $E_s/N_0 \rightarrow \infty$.

Now we will show how the free distance of ϕ_o and the recursiveness of ϕ_i come into the picture. First of all, we generalize the classical assumptions in the most natural way, simply replacing Hamming weight with the interconnection weight $\mathbf{w}_G$, which will ensure that there is an interleaving gain (namely that $\alpha \geq 1$). From now on, let's denote by d_f^o the $\mathbf{w}_G$ -free distance of ϕ_o .

PROPOSITION 5.5. *Assume that $d_f^o \geq 2$ and ϕ_i is $\mathbf{w}_G$ -recursive. Then*

$$\lfloor (d_f^o + 1)/2 \rfloor \leq \alpha \leq d_f^o.$$

In particular, $\alpha \geq 1$, and if $d_f^o \geq 3$, then $\alpha \geq 2$.

In some particular cases we will give tighter upper bounds on α (see section 7).

The strong assumptions used in Proposition 5.5 have another interesting consequence.

PROPOSITION 5.6. *If $d_f^o \geq 3$ and ϕ_i is $\mathbf{w}_G$ -recursive, then $\mathcal{H}$ is a finite set.*

However, these assumptions are not necessary to obtain an interleaver gain. For example, in the case of parallel concatenations with multiple branches, these assumptions would mean that all constituent encoders are recursive, while it is known that there is an interleaver gain, even if smaller, when only some of them are recursive [15]. Also, a relaxation of the classical assumptions will allow us to give results about very interesting examples, such as the heavily punctured serial schemes considered in [25], or the class of structured LDPC interpreted as serial schemes that we introduced in section 4.3. Thus, we are interested in a generalization of Proposition 5.5.

PROPOSITION 5.7. *Assume that the interconnection weight has the structure $\mathbf{w}_G = (\mathbf{w}_1, \mathbf{w}_2) : Y^{sM} \rightarrow \mathbb{N}^{\rho_1} \times \mathbb{N}^{\rho_2}$ (possibly $\rho_2 = 0$, but $\rho_1 \geq 1$); denote by $d_{f,1}^o$ the $\mathbf{w}_1$ -free distance of ϕ_o . Assume that $d_{f,1}^o \geq 2$ and ϕ_i is $\mathbf{w}_1$ -recursive. Then,*

$$\lfloor (d_{f,1}^o + 1)/2 \rfloor \leq \alpha \leq d_{f,1}^o.$$

In particular, $\alpha \geq 1$, and if $d_{f,1}^o \geq 3$, then $\alpha \geq 2$.

Notice that Proposition 5.5 is a particular case of Proposition 5.7, where $\rho_2 = 0$ and therefore $d_f^o = d_{f,1}^o$.

6. Proofs of the main results. In this section, we prove our main results; i.e., Theorem 5.4 and Proposition 5.7. We prove the upper bound for $\overline{P_s(e)}$ and the lower bound for $\overline{P_w(e)}$; the whole result stated in Theorem 5.4 is then obtained by the simple remark

$$\overline{P_s(e)} \geq \frac{1}{N} \frac{\mathbf{w}_{\text{in}}^{\min}}{\mathbf{w}_{\text{in}}^{\max}} \overline{P_w(e)}.$$

6.1. Upper bound. This proof is based on the union-Bhattacharyya bound (see, e.g., [28]) and on estimations of the weight enumerating coefficients of the constituent encoders.

We will consider only the case when the symbol error rate $P_s(e)$ is defined with respect to Hamming input weight ($\mathbf{w}_{\text{in}} = \mathbf{w}_H$); however, this will give results true for every other compatible weight, up to a positive constant factor.

The well-known union bound gives

$$(6.1) \quad \overline{P_s(e)} \leq \sum_w \sum_{\mathbf{d}} \frac{w}{N} \overline{A_{w,\mathbf{d}}}^N Q(\mathbf{d}),$$

where $\overline{A_{w,\mathbf{d}}}^N$ is the average number of codewords of a serial ensemble with input Hamming weight w and output type weight $\mathbf{d}$.

The standard technique (see [28, 2]) is to express $\overline{A_{w,\mathbf{d}}}^N$ as a function of suitable enumerating coefficients of the constituent encoders. Here, we need the following:

- $A_{w,\mathbf{h}}^{o,N}$, the number of codewords of ϕ_o^N with input Hamming weight w , and output invariants weight vector $\mathbf{h}$,
- $A_{\mathbf{h},\mathbf{d}}^{i,N}$, the number of codewords of ϕ_i^N with input invariants weight vector $\mathbf{h}$, and output type weight $\mathbf{d}$.

PROPOSITION 6.1. $\overline{A_{w,\mathbf{d}}}^N = \sum_{\mathbf{h} \in H} \frac{A_{w,\mathbf{h}}^{o,N} A_{\mathbf{h},\mathbf{d}}^{i,N}}{|Y_{\mathbf{h}}^{r(N+\nu_o)}|}$.

Proof.

$$\overline{A_{w,\mathbf{d}}}^N = \sum_{\mathbf{u}: \mathbf{w}_H(\mathbf{u})=w} \sum_{\mathbf{v}: \mathbf{w}_T(\phi_i^N(\mathbf{v}))=\mathbf{d}} \mathbb{P}(\Pi_N(\phi_o^N(\mathbf{u})) = \mathbf{v}).$$

By Remark 4.2,

$$\mathbb{P}(\Pi_N(\phi_o^N(\mathbf{u})) = \mathbf{v}) = \frac{|G_N(\phi_o^N(\mathbf{u}), \mathbf{v})|}{|G_N|} = \begin{cases} 0 & \text{if } \mathbf{w}_G(\phi_o^N(\mathbf{u})) \neq \mathbf{w}_G(\mathbf{v}) \\ \frac{1}{|Y_{\mathbf{h}}^{r(N+\nu_o)}|} & \text{if } \mathbf{w}_G(\phi_o^N(\mathbf{u})) = \mathbf{w}_G(\mathbf{v}) = \mathbf{h}. \end{cases}$$

Substituting in the expression above, we obtain the thesis. $\square$

By Lemma 2.2, we know that $|Y_{\mathbf{h}}^{r(N+\nu_o)}| \geq \binom{r(N+\nu_o)}{\mathbf{h}}$. Thus, by the inequality (6.1) and Proposition 6.1 we have

$$(6.2) \quad \overline{P_s(e)} \leq \sum_{w,\mathbf{h},\mathbf{d}} \frac{w}{N} \frac{1}{\binom{r(N+\nu_o)}{\mathbf{h}}} A_{w,\mathbf{h}}^{o,N} A_{\mathbf{h},\mathbf{d}}^{i,N} Q(\mathbf{d}).$$

We have some inequalities involving the indexes $w, \mathbf{h}, \mathbf{d}$ which are necessary conditions to have nonzero $A_{w,\mathbf{h}}^{o,N} A_{\mathbf{h},\mathbf{d}}^{i,N}$. They are listed in Definition 6.2 and Proposition 6.3.

DEFINITION 6.2. Let $I \subseteq \mathbb{N}^* \times H \times \mathbb{N}^{\Gamma \setminus \{0\}}$ be the set of triples $(w, \mathbf{h}, \mathbf{d})$ satisfying the following conditions:

- $1 \leq w \leq N$,
- $|\mathbf{d}| \leq l(M_N + \nu_i)$,
- $w \leq \mu_o |\mathbf{h}|$ and $|\mathbf{h}| \leq \mu_i |\mathbf{d}|$ (μ_o and μ_i as in Definition 5.1).

PROPOSITION 6.3. If $(w, \mathbf{h}, \mathbf{d}) \notin I$, then $A_{w,\mathbf{h}}^{o,N} A_{\mathbf{h},\mathbf{d}}^{i,N} = 0$.

Proof. The first two inequalities are trivial remarks about the length of the input and codewords and the definition of free distance; the last one is the concatenated noncatastrophicity of the ensemble (see Definition 5.1). $\square$

Now, we need to estimate the product $A_{w,\mathbf{h}}^{o,N} A_{\mathbf{h},\mathbf{d}}^{i,N}$ when it is nonzero. We start with the following inequalities derived from Proposition 3.11.

PROPOSITION 6.4. There exist some positive constants a_o, a_i, b_o, b_i such that, for every $(w, \mathbf{h}, \mathbf{d}) \in I$,

$$\begin{aligned}
 1. \quad A_{w, \mathbf{h}}^{o, N} &\leq \sum_{n_o=1}^{n_o(\mathbf{h})} \binom{N+n_o}{n_o} a_o^w b_o^{|\mathbf{h}|}. \\
 2. \quad A_{\mathbf{h}, \mathbf{d}}^{i, N} &\leq \sum_{n_i=0}^{n_i^{\max}} \binom{N+n_i}{n_i} a_i^{|\mathbf{h}|} b_i^{|\mathbf{d}|}, \quad \text{where } n_i^{\max} = \begin{cases} n_i(\mathbf{h}) & \text{if } Q(\mathbf{d}) \leq q^*(\mathbf{h}) \\ n_i(\mathbf{h}) - 1 & \text{if } Q(\mathbf{d}) > q^*(\mathbf{h}). \end{cases}
 \end{aligned}$$

Proof. Let $w_{\max} = \max\{|\mathbf{w}_G(v)| : v \in Y^r\}$, so that $|\mathbf{w}_G(\mathbf{v})|/w_{\max} \leq w_H(\mathbf{v}) \leq |\mathbf{w}_G(\mathbf{v})|$ for all $\mathbf{v} \in Y^{r(N+\nu_o)}$. Then

$$A_{w, \mathbf{h}}^{o, N} \leq \sum_{n_o=1}^{n_o(\mathbf{h})} \sum_{h'=\lfloor |\mathbf{h}|/w_{\max} \rfloor}^{|\mathbf{h}|} \Lambda_{w, h', n_o}^{o, N},$$

where $\Lambda_{w, h', n_o}^{o, N}$ is the input/output support enumerating coefficient of ϕ_o^N , as defined in section 3.5. The conclusion now follows in a straightforward way from Proposition 3.11.

The proof for the inner encoder is similar, but we want to exploit the fact that, by definition of $q^*(\mathbf{h})$, there are no codewords of input weight $\mathbf{h}$ and output weight $\mathbf{d}$ such that $Q(\mathbf{d}) > q^*(\mathbf{h})$ having $n_i(\mathbf{h})$ error events in their decomposition. To do so, we need to define $A_{\mathbf{h}, \mathbf{d}, n}^{i, N}$ to be the number of codewords of ϕ_i^N with input invariants weight vector $\mathbf{h}$, output type weight $\mathbf{d}$, and n error events in the decomposition, so that

$$A_{\mathbf{h}, \mathbf{d}}^{i, N} = \sum_{n_i=0}^{n_i(\mathbf{h})} A_{\mathbf{h}, \mathbf{d}, n_i}^{i, N} = \sum_{n_i=0}^{n_i^{\max}} A_{\mathbf{h}, \mathbf{d}, n_i}^{i, N}$$

because $A_{\mathbf{h}, \mathbf{d}, n_i(\mathbf{h})}^{i, N} = 0$ if $Q(\mathbf{d}) > q^*(\mathbf{h})$. Then we conclude the proof as for the outer encoder, with $\hat{w}_{\max} = \max\{|\mathbf{w}_T(g)| : g \in \Gamma^l\}$:

$$A_{\mathbf{h}, \mathbf{d}}^{i, N} \leq \sum_{n_i=1}^{n_i^{\max}} \sum_{h'=\lfloor \frac{|\mathbf{h}|}{w_{\max}} \rfloor}^{|\mathbf{h}|} \sum_{d'=\lfloor \frac{|\mathbf{d}|}{\hat{w}_{\max}} \rfloor}^{|\mathbf{d}|} \Lambda_{h', d', n_i}^{i, N}. \quad \square$$

We now prove the following combinatorial inequality.

PROPOSITION 6.5. *There exists a constant $C > 0$ such that for all $\mathbf{h} \in H$ with $|\mathbf{h}| \leq w_{\max} r(N + \nu_o)$ (where $w_{\max} = \max\{|\mathbf{w}_G(v)| : v \in Y^r\}$),*

$$\frac{1}{\binom{r(N+\nu_o)}{\mathbf{h}}} \sum_{n_o=1}^{n_o(\mathbf{h})} \sum_{n_i=0}^{n_i^{\max}} \binom{N+n_o}{n_o} \binom{N+n_i}{n_i} \leq \begin{cases} C^{|\mathbf{h}|} \frac{|\mathbf{h}|^{f(\mathbf{h})-1}}{N^{f(\mathbf{h})-1}} & \text{if } n_i^{\max} = n_i(\mathbf{h}) \\ C^{|\mathbf{h}|} \frac{|\mathbf{h}|^{f(\mathbf{h})}}{N^{f(\mathbf{h})}} & \text{if } n_i^{\max} = n_i(\mathbf{h}) - 1. \end{cases}$$

Proof. First, we have $\binom{r(N+\nu_o)}{\mathbf{h}} \geq \left\lceil \frac{r(N+\nu_o)}{e|\mathbf{h}|} \right\rceil^{|\mathbf{h}|}$. This gives $\frac{1}{\binom{r(N+\nu_o)}{\mathbf{h}}} \leq C^{|\mathbf{h}|} \left[\frac{|\mathbf{h}|}{N} \right]^{|\mathbf{h}|}$ for some constant $C > 0$.

For the other terms, we use the following simple combinatorial inequalities:

- $\binom{N+n}{n} \leq \left\lceil \frac{s}{N} \right\rceil^{s-n} \binom{N+s}{s} \quad \forall n \geq 0, s, N \geq 1$ satisfying $s \geq n$,
- there exists a constant $c > 0$ such that $\binom{N+n}{n} \leq c \left[\frac{N+n}{N} \right]^N \left[\frac{N+n}{n} \right]^n \quad \forall n, N \geq 1$,
- $\left[\frac{N+n}{N} \right]^N \leq e^n \quad \forall n \geq 0, N \geq 1$.

As $n_o(\mathbf{h}) \leq |\mathbf{h}|$, these inequalities give

$$\begin{aligned} \sum_{n_o=1}^{n_o(\mathbf{h})} \binom{N+n_o}{n_o} &\leq \sum_{n_o=1}^{n_o(\mathbf{h})} \left[\frac{|\mathbf{h}|}{N} \right]^{|\mathbf{h}|-n_o} c e^{|\mathbf{h}|} \left[\frac{N+|\mathbf{h}|}{|\mathbf{h}|} \right]^{|\mathbf{h}|} \\ &\leq c_o^{|\mathbf{h}|} \sum_{n_o=1}^{n_o(\mathbf{h})} |\mathbf{h}|^{-n_o} N^{n_o} \\ &\leq C_o^{|\mathbf{h}|} |\mathbf{h}|^{-n_o(\mathbf{h})} N^{n_o(\mathbf{h})} \end{aligned}$$

(for some positive constants c_o and C_o). The second inequality is true thanks to the assumption that $|\mathbf{h}| \leq w_{\max} r(N + \nu_o)$. Similar estimation can be obtained for the summation concerning the inner encoder which yields the result. $\square$

If we substitute the estimations given by Propositions 6.4 and 6.5 into the expression (6.2) and use Proposition 6.3, we get, for some positive constants C_1, C_2, C_3 ,

$$(6.3) \quad \overline{P_s(e)} \leq \sum_{\substack{(w, \mathbf{h}, \mathbf{d}) \in I: \\ Q(\mathbf{d}) \leq q^*(\mathbf{h})}} \frac{|\mathbf{h}|^{f(\mathbf{h})-1}}{N^{f(\mathbf{h})}} C_1^w C_2^{|\mathbf{h}|} C_3^{|\mathbf{d}|} Q(\mathbf{d}) + \sum_{\substack{(w, \mathbf{h}, \mathbf{d}) \in I: \\ Q(\mathbf{d}) > q^*(\mathbf{h})}} \frac{|\mathbf{h}|^{f(\mathbf{h})}}{N^{f(\mathbf{h})+1}} C_1^w C_2^{|\mathbf{h}|} C_3^{|\mathbf{d}|} Q(\mathbf{d}).$$

Now, we split the first summation into two terms, separating $\mathbf{h} \in \mathcal{H}$ from $\mathbf{h} \notin \mathcal{H}$. Define

- $I_\alpha = \{(w, \mathbf{h}, \mathbf{d}) \in I : f(\mathbf{h}) = \alpha, Q(\mathbf{d}) \leq q^*(\mathbf{h})\}$,
- $I_> = \{(w, \mathbf{h}, \mathbf{d}) \in I : f(\mathbf{h}) > \alpha, Q(\mathbf{d}) \leq q^*(\mathbf{h})\}$,
- $I_* = \{(w, \mathbf{h}, \mathbf{d}) \in I : Q(\mathbf{d}) > q^*(\mathbf{h})\}$.

Equation (6.3) can be rewritten as follows:

$$\begin{aligned} \overline{P_s(e)} &\leq \frac{1}{N^\alpha} \sum_{(w, \mathbf{h}, \mathbf{d}) \in I_\alpha} |\mathbf{h}|^{\alpha-1} C_1^w C_2^{|\mathbf{h}|} C_3^{|\mathbf{d}|} Q(\mathbf{d}) \\ &\quad + \frac{1}{N^{\alpha+1}} \sum_{(w, \mathbf{h}, \mathbf{d}) \in I_>} C_1^w \left(\frac{|\mathbf{h}|}{N} \right)^{f(\mathbf{h})-\alpha-1} |\mathbf{h}|^\alpha C_2^{|\mathbf{h}|} C_3^{|\mathbf{d}|} Q(\mathbf{d}) \\ &\quad + \frac{1}{N^{\alpha+1}} \sum_{(w, \mathbf{h}, \mathbf{d}) \in I_*} C_1^w \left(\frac{|\mathbf{h}|}{N} \right)^{f(\mathbf{h})-\alpha} |\mathbf{h}|^\alpha C_2^{|\mathbf{h}|} C_3^{|\mathbf{d}|} Q(\mathbf{d}). \end{aligned}$$

In the following, we will show that the first summation is bounded by the expression $c q^* \exp(\log q^* / \log \gamma)$ (Proposition 6.6), and that the second and third summations are bounded by a constant $c'(\gamma)$ (Proposition 6.7), thus ending the proof of the upper bound.

PROPOSITION 6.6. *There exist some positive constants γ_0 and c such that, for any BIOS channel with $\gamma < \gamma_0$,*

$$\sum_{(w, \mathbf{h}, \mathbf{d}) \in I_\alpha} |\mathbf{h}|^{\alpha-1} C_1^w C_2^{|\mathbf{h}|} C_3^{|\mathbf{d}|} Q(\mathbf{d}) \leq c q^* \exp(\log q^* / \log \gamma).$$

Proof. Recall that $(w, \mathbf{h}, \mathbf{d}) \in I_\alpha$ implies that $w \leq \mu_o |\mathbf{h}|$, $|\mathbf{h}| \leq \mu_i |\mathbf{d}|$, and $Q(\mathbf{d}) \leq q^*(\mathbf{h}) \leq q^*$. So therefore,

$$\begin{aligned} & \sum_{(w, \mathbf{h}, \mathbf{d}) \in I_\alpha} |\mathbf{h}|^{\alpha-1} C_1^w C_2^{|\mathbf{h}|} C_3^{|\mathbf{d}|} Q(\mathbf{d}) \\ & \leq \sum_{\substack{\mathbf{d} \in \mathbb{N}^{\Gamma \setminus \{0\}} \\ Q(\mathbf{d}) \leq q^*}} \sum_{\substack{\mathbf{h} \in \mathbb{N}^\rho \\ |\mathbf{h}| \leq \mu_i |\mathbf{d}|}} |\mathbf{h}|^{\alpha-1} C_2^{|\mathbf{h}|} \left(\sum_{w \leq \mu_o |\mathbf{h}|} C_1^w \right) C_3^{|\mathbf{d}|} Q(\mathbf{d}) \\ & \leq \sum_{\substack{\mathbf{d} \in \mathbb{N}^{\Gamma \setminus \{0\}} \\ Q(\mathbf{d}) \leq q^*}} \sum_{\substack{\mathbf{h} \in \mathbb{N}^\rho \\ |\mathbf{h}| \leq \mu_i |\mathbf{d}|}} |\mathbf{h}|^{\alpha-1} C_2^{|\mathbf{h}|} \mu_o |\mathbf{h}| C_1^{|\mathbf{h}|} C_3^{|\mathbf{d}|} Q(\mathbf{d}) \\ & \leq \sum_{\substack{\mathbf{d} \in \mathbb{N}^{\Gamma \setminus \{0\}} \\ Q(\mathbf{d}) \leq q^*}} K^{|\mathbf{d}|} Q(\mathbf{d}) \quad (\text{for some suitable } K > 0). \end{aligned}$$

Now, we split the summation, recalling the Bhattacharyya bound $Q(\mathbf{d}) \leq \gamma^{|\mathbf{d}|}$:

$$\sum_{\substack{\mathbf{d} \in \mathbb{N}^{\Gamma \setminus \{0\}} \\ Q(\mathbf{d}) \leq q^*}} K^{|\mathbf{d}|} Q(\mathbf{d}) = \sum_{\substack{\mathbf{d} \in \mathbb{N}^{\Gamma \setminus \{0\}} \\ Q(\mathbf{d}) \leq q^*, \gamma^{|\mathbf{d}|} > q^*}} K^{|\mathbf{d}|} q^* + \sum_{\substack{\mathbf{d} \in \mathbb{N}^{\Gamma \setminus \{0\}} \\ \gamma^{|\mathbf{d}|} \leq q^*}} K^{|\mathbf{d}|} \gamma^{|\mathbf{d}|}.$$

Now let's find a bound for the number of $\mathbf{d}$'s involved in the first summation: $\gamma^{|\mathbf{d}|} > q^*$ means $|\mathbf{d}| < \log q^* / \log \gamma$ and so there are less than $(|\Gamma| - 1)^{\log q^* / \log \gamma}$ type weights satisfying this inequality:

$$\sum_{\substack{\mathbf{d} \in \mathbb{N}^{\Gamma \setminus \{0\}} \\ Q(\mathbf{d}) \leq q^*, \gamma^{|\mathbf{d}|} \geq q^*}} K^{|\mathbf{d}|} q^* \leq ((|\Gamma| - 1)K)^{\log q^* / \log \gamma} q^*.$$

For the second term, note that for $\gamma < 1/K$, the series is convergent and bounded by a constant times its first term, which has $|\mathbf{d}| = \log q^* / \log \gamma$; i.e.,

$$\sum_{\substack{\mathbf{d} \in \mathbb{N}^{\Gamma \setminus \{0\}} \\ \gamma^{|\mathbf{d}|} \leq q^*}} K^{|\mathbf{d}|} \gamma^{|\mathbf{d}|} \leq CK^{(\log q^* / \log \gamma)} q^*. \quad \square$$

PROPOSITION 6.7. *There exists a constant $\gamma_0 > 0$ depending on ϕ_o, ϕ_i and (G_N) , and there exists $c'(\gamma) > 0$ depending only on γ such that for all $\gamma < \gamma_0$*

$$\begin{aligned} & \sum_{(w, \mathbf{h}, \mathbf{d}) \in I_>} C_1^w \left(\frac{|\mathbf{h}|}{N} \right)^{f(\mathbf{h}) - \alpha - 1} |\mathbf{h}|^\alpha C_2^{|\mathbf{h}|} C_3^{|\mathbf{d}|} Q(\mathbf{d}) \\ & + \sum_{(w, \mathbf{h}, \mathbf{d}) \in I_*} C_1^w \left(\frac{|\mathbf{h}|}{N} \right)^{f(\mathbf{h}) - \alpha} |\mathbf{h}|^\alpha C_2^{|\mathbf{h}|} C_3^{|\mathbf{d}|} Q(\mathbf{d}) \leq c'(\gamma). \end{aligned}$$

Proof. Notice that for $(w, \mathbf{h}, \mathbf{d}) \in I_>$, $0 \leq f(\mathbf{h}) - \alpha - 1 \leq |\mathbf{h}| \leq cN$, where the first inequality holds true because $\mathbf{h} \in H \setminus \mathcal{H}$, the second immediately follows from the definitions of $f(\mathbf{h})$ and α , and the third is true, for a suitable $c > 1$, because $|\mathbf{h}| \leq r(N + \nu_o)$

for all $\mathbf{h} \in H$. These inequalities imply that $(|\mathbf{h}|/N)^{f(\mathbf{h})-\alpha-1} \leq c^{|\mathbf{h}|}$. Analogously, for all $(w, \mathbf{h}, d) \in I_*$, $0 \leq f(\mathbf{h}) - \alpha \leq |\mathbf{h}| \leq cN$ and so $(|\mathbf{h}|/N)^{f(\mathbf{h})-\alpha} \leq c^{|\mathbf{h}|}$.

This gives

$$\begin{aligned} & \sum_{(w, \mathbf{h}, d) \in I_>} C_1^w \left(\frac{|\mathbf{h}|}{N} \right)^{f(\mathbf{h})-\alpha-1} |\mathbf{h}|^\alpha C_2^{|\mathbf{h}|} C_3^{|d|} Q(d) \\ & + \sum_{(w, \mathbf{h}, d) \in I_*} C_1^w \left(\frac{|\mathbf{h}|}{N} \right)^{f(\mathbf{h})-\alpha} |\mathbf{h}|^\alpha C_2^{|\mathbf{h}|} C_3^{|d|} Q(d) \\ & \leq \sum_{(w, \mathbf{h}, d) \in I} C_1^w c^{|\mathbf{h}|} C_2^{|\mathbf{h}|} C_3^{|d|} Q(d). \end{aligned}$$

Noticing also that $\sum_{w \leq \mu_o |\mathbf{h}|} C_1^w \leq \mu_o |\mathbf{h}| C_1^{|\mathbf{h}|}$, we have

$$\sum_{(w, \mathbf{h}, d) \in I} C_1^w c^{|\mathbf{h}|} C_2^{|\mathbf{h}|} C_3^{|d|} Q(d) \leq \sum_{d \in \mathbb{N}^{\Gamma \setminus \{0\}}} \sum_{\mathbf{h}: |\mathbf{h}| \leq \mu_i |d|} \mu_o |\mathbf{h}| C_1^{|\mathbf{h}|} c^{|\mathbf{h}|} |\mathbf{h}|^\alpha C_2^{|\mathbf{h}|} C_3^{|d|} Q(d).$$

Now notice that, for some $K > 1$,

$$\sum_{\mathbf{h}: |\mathbf{h}| \leq \mu_i |d|} \mu_o |\mathbf{h}| C_1^{|\mathbf{h}|} c^{|\mathbf{h}|} |\mathbf{h}|^\alpha C_2^{|\mathbf{h}|} \leq K^{|d|}.$$

Finally, we use the Bhattacharyya bound.

$$\sum_{d \in \mathbb{N}^{\Gamma \setminus \{0\}}} (KC)_3^{|d|} Q(d) \leq \sum_{d \in \mathbb{N}} \sum_{d \in \mathbb{N}^{\Gamma \setminus \{0\}}: |d|=d} (KC_3 \gamma)^d = c'(\gamma) < \infty$$

if γ is sufficiently small to ensure convergence. $\square$

6.2. Lower bound. The lower bound is based on the following simple remark involving the equivocation probability.

Remark 6.8. If $\mathbf{c} \in \mathcal{C}^N$, then $P_w(e) \geq \mathbb{P}(\mathbf{0} \rightarrow \mathbf{c})$. Therefore, if one defines $Q_{\max}(\pi_N) = \max\{\mathbb{P}(\mathbf{0} \rightarrow \mathbf{c}), \mathbf{c} \in \phi_i^N \circ \pi_N \circ \phi_o^N(U^N)\}$, then, for any q ,

$$\overline{P_w(e)} \geq q \mathbb{P}(Q_{\max}(\Pi_N) \geq q).$$

We focus our attention on the value $q = q^*$, and we find the following lower bound to $\mathbb{P}(Q_{\max}(\Pi_N) \geq q)$, thus ending the proof of the lower bound in Theorem 5.4.

PROPOSITION 6.9. *If $\alpha \geq 1$, there exists a constant $C > 0$ such that*

$$\mathbb{P}(Q_{\max}(\Pi_N) \geq q) \geq CN^{-\alpha+1}.$$

In the remainder of this section, we will prove Proposition 6.9. To do so, we need to define some particular codewords which are essential for the bound. We start fixing once and for all the following objects:

1. A weight vector $\mathbf{h} \in H$ such that $q^*(\mathbf{h}) = q^*$.
2. An outer codeword $\mathbf{c}^* \in \phi_o^N(U^N)$ such that $\mathbf{w}_G(\mathbf{c}^*) = \mathbf{h}$ and $n(\mathbf{c}) = n_o(\mathbf{h})$. Let $n_o = n_o(\mathbf{h})$, and let $\mathbf{c}^* = \mathbf{c}_1^* + \dots + \mathbf{c}_{n_o+1}^*$ be an error event decomposition of $\mathbf{c}^*$ (see section 3.4). Denote by l_k the length of $\mathbf{c}_k^*$, and define $l_{\max} = \max\{l_1, \dots, l_{n_o}\}$. If $\alpha = 1$, we need a slightly different definition: $l_{\max} = \max\{l_1, \dots, l_{n_o}, L\}$ for a suitable constant L , which will be chosen in Lemma 6.12.

3. An input word $\mathbf{u}^*$ for the inner encoder, such that $\mathbf{w}_G(\mathbf{u}^*) = \mathbf{h}$ and such that $\mathbf{x}^* = \phi_i^N(\mathbf{u}^*)$, has equivocation $P(\mathbf{0} \rightarrow \mathbf{x}^*) = q^*(\mathbf{h})$ and $n(\mathbf{x}^*) = n_i(\mathbf{h})$. Let $n_i = n_i(\mathbf{h})$, and let $\mathbf{x}^* = \mathbf{x}_1^* + \dots + \mathbf{x}_{n_i+1}^*$ be an error event decomposition of $\mathbf{x}^*$. Denote by $\mathbf{u}_k^*$ the input error event corresponding to $\mathbf{x}_k^*$ and by λ_k its length (with $\lambda_{n_i+1} = 0$ if there is no terminating event). Define $\lambda_{\max} = \max\{\lambda_1, \dots, \lambda_{n_i}\}$. Again, for $\alpha = 1$ we need a different definition $\lambda_{\max} = \max\{\lambda_1, \dots, \lambda_{n_i}, \Lambda\}$ for a suitable constant Λ described in Lemma 6.12.

Notice that $\mathbf{c}^*$ can be chosen in such a way that it doesn't have any terminating event and does not depend on N , while this may not be possible for $\mathbf{u}^*$. However, we can assume that the error events $\mathbf{x}_k^*$ and their inputs $\mathbf{u}_k^*$ remain the same apart from some possible translations (see Remark 3.9). Also remember that $n_o \geq 1$, $n_i \geq 0$.

Now, we select a sufficiently big set of shift equivalent words for both $\mathbf{c}^*$ and $\mathbf{x}^*$, choosing many positions for the error events of $\mathbf{c}^*$ and for the input error events of $\mathbf{u}^*$, across all the time axis $[0, N + \nu_o - 1]$ for $\mathbf{c}^*$ and $[0, M_N - 1]$ for $\mathbf{u}^*$.

Let's start with $\mathbf{c}^*$. Define $\mathcal{A} = [0, \lfloor \frac{N}{n_o \lambda_{\max}} \rfloor - 1]$. Given $\mathbf{a} \in \mathcal{A}^{n_o}$, we define $\mathbf{c}_\mathbf{a}^*$ to be the outer codeword which, for every $k = 1, \dots, n_o$, contains exactly one shifted copy of the error event $\mathbf{c}_k^*$ starting at time $a_k \lambda_{\max} + (k-1)|\mathcal{A}| \lambda_{\max}$. Clearly by construction all error events in $\mathbf{c}_\mathbf{a}^*$ have disjoint support.

In the same way, we consider the inner input word $\mathbf{u}^*$. For $n_i \geq 1$, define $\mathcal{B} = [0, \lfloor \frac{M_N - \lambda_{n_i+1}}{n_i \lambda_{\max}} \rfloor - 1]$. Given $\mathbf{b} \in \mathcal{B}^{n_i}$ we define $\mathbf{u}_\mathbf{b}^*$ to be the inner input word which, for every $k = 1, \dots, n_i$, contains exactly one translated copy of the input error event $\mathbf{u}_k^*$ starting at time $b_k \lambda_{\max} + (k-1)|\mathcal{B}| \lambda_{\max}$, while the terminating event $\mathbf{u}_{n_i+1}^*$ (if there is one) remains fixed in its position in the interval $[M_N - \lambda_{n_i+1} - 1, M_N - 1]$. Let $\mathbf{x}_\mathbf{b}^*$ be the output $\mathbf{x}_\mathbf{b}^* = \phi_i^N(\mathbf{u}_\mathbf{b}^*)$.

Given $\mathbf{a} \in \mathcal{A}^{n_o}$ and $\mathbf{b} \in \mathcal{B}^{n_i}$, if $n_i \geq 1$ we define the event

$$E_{\mathbf{a}, \mathbf{b}} = \{\Pi_N(\mathbf{c}_\mathbf{a}^*) = \mathbf{u}_\mathbf{b}^*\} = G_N(\mathbf{c}_\mathbf{a}^*, \mathbf{u}_\mathbf{b}^*),$$

and we also define $E_\mathbf{a} = \bigcup_{\mathbf{b} \in \mathcal{B}} E_{\mathbf{a}, \mathbf{b}}$ (notice that this is a union of disjoint events). If $n_i = 0$, then we simply let $E_\mathbf{a} = \{\Pi_N(\mathbf{c}_\mathbf{a}^*) = \mathbf{u}^*\} = G_N(\mathbf{c}_\mathbf{a}^*, \mathbf{u}^*)$.

Remark 6.10. Clearly $\pi_N \in E_{\mathbf{a}, \mathbf{b}}$ implies $Q_{\max}(\pi_N) \geq P(\mathbf{0} \rightarrow \mathbf{x}_\mathbf{b}^*) = q^*$. Hence,

$$\mathbb{P}(Q_{\max}(\Pi_N) \geq q^*) \geq \mathbb{P}\left(\bigcup_{\mathbf{a} \in \mathcal{A}^{n_o}} E_\mathbf{a}\right).$$

Our aim is now to estimate this last probability, using

$$\mathbb{P}\left(\bigcup_{\mathbf{a} \in \mathcal{A}^{n_o}} E_\mathbf{a}\right) \geq \sum_{\mathbf{a} \in \mathcal{A}^{n_o}} \mathbb{P}(E_\mathbf{a}) - \sum_{\substack{\mathbf{a}, \mathbf{a}' \in \mathcal{A}^{n_o} \\ \mathbf{a} \neq \mathbf{a}'}} \mathbb{P}(E_\mathbf{a} \cap E_{\mathbf{a}'}).$$

We will prove a lower bound for the first term (Lemma 6.11) and an upper bound for the second term (Lemma 6.12).

LEMMA 6.11.

$$\sum_{\mathbf{a} \in \mathcal{A}^{n_o}} \mathbb{P}(E_\mathbf{a}) \geq |\mathcal{A}|^{n_o} |\mathcal{B}|^{n_i} \frac{1}{\lceil |Y| r(N + \nu_o) \rceil |\mathbf{h}|},$$

where in the case $n_i = 0$ (for which $\mathcal{B}$ is not defined), we define $|\mathcal{B}| = 1$.

Proof. Clearly

$$\mathbb{P}(E_{\mathbf{a}}) = \frac{|E_{\mathbf{a}}|}{|G_N|} = \frac{|\mathcal{B}|^{n_i} |G_N(\mathbf{c}^*, \mathbf{u}^*)|}{|G_N|}.$$

By Remark 4.2 and Lemma 2.2,

$$\frac{|G_N(\mathbf{c}^*, \mathbf{u}^*)|}{|G_N|} = \frac{1}{|Y_{\mathbf{h}}^{r(N+\nu_o)}|} \geq \frac{1}{[|Y|r(N+\nu_o)]^{|\mathbf{h}|}}. \quad \square$$

LEMMA 6.12. *If $\alpha \geq 2$, then*

$$\sum_{\substack{\mathbf{a}, \mathbf{a}' \in \mathcal{A}^{n_o} \\ \mathbf{a} \neq \mathbf{a}'}} \mathbb{P}(E_{\mathbf{a}} \cap E_{\mathbf{a}'}) \leq |\mathcal{A}|^{2n_o} |\mathcal{B}|^{2n_i} \left(\frac{2e|\mathbf{h}|}{r(N+\nu_o)} \right)^{2|\mathbf{h}|},$$

while if $\alpha = 1$, then

$$\sum_{\substack{\mathbf{a}, \mathbf{a}' \in \mathcal{A}^{n_o} \\ \mathbf{a} \neq \mathbf{a}'}} \mathbb{P}(E_{\mathbf{a}} \cap E_{\mathbf{a}'}) \leq \frac{|\mathcal{A}|^{n_o} |\mathcal{B}|^{n_i}}{[r(N+\nu_o)]^{|\mathbf{h}|}} C$$

for some constant $0 < C < \frac{1}{|Y|^{|\mathbf{h}|}}$.

Proof. We have

$$E_{\mathbf{a}} \cap E_{\mathbf{a}'} = \bigcup_{\mathbf{b}, \mathbf{b}' \in \mathcal{B}^{n_i}} (E_{\mathbf{a}, \mathbf{b}} \cap E_{\mathbf{a}', \mathbf{b}'}).$$

Consequently, by the union bound,

$$\sum_{\substack{\mathbf{a}, \mathbf{a}' \in \mathcal{A}^{n_o} \\ \mathbf{a} \neq \mathbf{a}'}} \mathbb{P}(E_{\mathbf{a}} \cap E_{\mathbf{a}'}) \leq \sum_{\substack{\mathbf{a}, \mathbf{a}' \in \mathcal{A}^{n_o} \\ \mathbf{a} \neq \mathbf{a}'}} \sum_{\mathbf{b}, \mathbf{b}' \in \mathcal{B}^{n_i}} \mathbb{P}(E_{\mathbf{a}, \mathbf{b}} \cap E_{\mathbf{a}', \mathbf{b}'}).$$

Now we need to deal with $\mathbb{P}(E_{\mathbf{a}, \mathbf{b}} \cap E_{\mathbf{a}', \mathbf{b}'})$. First of all, note that if there is an incomplete error event in $\mathbf{u}^*$, surely $\mathbb{P}(E_{\mathbf{a}, \mathbf{b}} \cap E_{\mathbf{a}', \mathbf{b}'}) = 0$ for all $\mathbf{a} \neq \mathbf{a}'$ and for all $\mathbf{b}, \mathbf{b}'$. Now consider the case of only regular events. Fix any $\mathbf{a} \neq \mathbf{a}'$ and $\mathbf{b}, \mathbf{b}'$ such that there exists $\pi \in E_{\mathbf{a}, \mathbf{b}} \cap E_{\mathbf{a}', \mathbf{b}'}$. By the definition of $\mathbf{c}_{\mathbf{a}}^*$ and $\mathbf{c}_{\mathbf{a}'}^*$, we can find outer codewords $\tilde{\mathbf{c}}^*$, $\tilde{\mathbf{c}}_{\mathbf{a}}^*$, $\tilde{\mathbf{c}}_{\mathbf{a}'}^*$ (possibly $\tilde{\mathbf{c}}^* = \mathbf{0}$) having disjoint supports, each consisting of some of the error events $\mathbf{c}_k^*$, such that $\mathbf{c}_{\mathbf{a}}^* = \tilde{\mathbf{c}}^* + \tilde{\mathbf{c}}_{\mathbf{a}}^*$ and $\mathbf{c}_{\mathbf{a}'}^* = \tilde{\mathbf{c}}^* + \tilde{\mathbf{c}}_{\mathbf{a}'}^*$. More precisely, letting $\tilde{n}_o = d_H(\mathbf{a}, \mathbf{a}')$, i.e., the number of i 's such that $\mathbf{a}_i \neq \mathbf{a}'_i$, $\tilde{\mathbf{c}}^*$ consists of $n_o - \tilde{n}_o$ error events, and $\tilde{\mathbf{c}}_{\mathbf{a}}^*$ consists of $\tilde{n}_o$ error events and is shift equivalent to $\tilde{\mathbf{c}}_{\mathbf{a}'}^*$. Clearly, $\mathbf{w}_G(\tilde{\mathbf{c}}_{\mathbf{a}}^*) = \mathbf{w}_G(\tilde{\mathbf{c}}_{\mathbf{a}'}^*) = \mathbf{h} - \mathbf{w}_G(\tilde{\mathbf{c}}^*)$.

Similarly, we can find inner input words $\tilde{\mathbf{u}}^*$, $\tilde{\mathbf{u}}_{\mathbf{b}}^*$, $\tilde{\mathbf{u}}_{\mathbf{b}'}^*$ (possibly $\tilde{\mathbf{u}}^* = \mathbf{0}$ or $\tilde{\mathbf{u}}_{\mathbf{b}}^* = \tilde{\mathbf{u}}_{\mathbf{b}'}^* = \mathbf{0}$) having disjoint supports, each consisting of some of the input error events $\mathbf{u}_k^*$, such that $\mathbf{u}_{\mathbf{b}}^* = \tilde{\mathbf{u}}^* + \tilde{\mathbf{u}}_{\mathbf{b}}^*$ and $\mathbf{u}_{\mathbf{b}'}^* = \tilde{\mathbf{u}}^* + \tilde{\mathbf{u}}_{\mathbf{b}'}^*$. Letting $\tilde{n}_i = d_H(\mathbf{b}, \mathbf{b}')$, $\tilde{\mathbf{u}}^*$ has $n_i - \tilde{n}_i$ error events and $\tilde{\mathbf{u}}_{\mathbf{b}}^*$ has $\tilde{n}_i$ error events and is shift equivalent to $\tilde{\mathbf{u}}_{\mathbf{b}'}^*$. Clearly, $\mathbf{w}_G(\tilde{\mathbf{u}}_{\mathbf{b}}^*) = \mathbf{w}_G(\tilde{\mathbf{u}}_{\mathbf{b}'}^*) = \mathbf{h} - \mathbf{w}_G(\tilde{\mathbf{u}}^*)$.

As a consequence of Lemma 4.3, if $\pi \in E_{\mathbf{a}, \mathbf{b}} \cap E_{\mathbf{a}', \mathbf{b}'}$, then $\pi(\tilde{\mathbf{c}}^*) = \tilde{\mathbf{u}}^*$, $\pi(\tilde{\mathbf{c}}_{\mathbf{a}}^*) = \tilde{\mathbf{u}}_{\mathbf{b}}^*$ and $\pi(\tilde{\mathbf{c}}_{\mathbf{a}'}^*) = \tilde{\mathbf{u}}_{\mathbf{b}'}^*$. This implies that $\mathbf{w}_G(\tilde{\mathbf{u}}^*) = \mathbf{w}_G(\tilde{\mathbf{c}}^*)$ and that $\mathbf{w}_G(\tilde{\mathbf{u}}_{\mathbf{b}}^*) = \mathbf{w}_G(\tilde{\mathbf{u}}_{\mathbf{b}'}^*) = \mathbf{w}_G(\tilde{\mathbf{c}}_{\mathbf{a}}^*) = \mathbf{w}_G(\tilde{\mathbf{c}}_{\mathbf{a}'}^*) = \mathbf{h} - \mathbf{w}_G(\tilde{\mathbf{c}}^*)$. We will use the notation $\tilde{\mathbf{h}} = \mathbf{w}_G(\tilde{\mathbf{u}}_{\mathbf{b}}^*)$. Note that

if $\mathbb{P}(E_{\mathbf{a},\mathbf{b}} \cap E_{\mathbf{a}',\mathbf{b}'}) \neq 0$ and $(\mathbf{a}, \mathbf{b}) \neq (\mathbf{a}', \mathbf{b}')$, then surely both $\mathbf{a} \neq \mathbf{a}'$ and $\mathbf{b} \neq \mathbf{b}'$. Also note that

$$\mathbb{P}(E_{\mathbf{a},\mathbf{b}} \cap E_{\mathbf{a}',\mathbf{b}'}) \leq \mathbb{P}(\Pi_N(\tilde{\mathbf{c}}^* + \tilde{\mathbf{c}}_{\mathbf{a}}^* + \tilde{\mathbf{c}}_{\mathbf{a}'}^*) = \tilde{\mathbf{u}}^* + \tilde{\mathbf{u}}_{\mathbf{b}}^* + \tilde{\mathbf{u}}_{\mathbf{b}'}^*) = \frac{1}{Y_{\mathbf{h}+\tilde{\mathbf{h}}}^{r(N+\nu_o)}}.$$

In the simple case when $\tilde{\mathbf{h}} = \mathbf{h}$, this gives

$$\mathbb{P}(E_{\mathbf{a},\mathbf{b}} \cap E_{\mathbf{a}',\mathbf{b}'}) \leq \frac{1}{|Y_{2\mathbf{h}}^{r(N+\nu_o)}|} \leq \left(\frac{2e|\mathbf{h}|}{r(N+\nu_o)} \right)^{2|\mathbf{h}|},$$

where the last line comes from Lemma 2.2.

Now notice that $\tilde{\mathbf{h}} \in H$ and $\mathbf{h} - \tilde{\mathbf{h}} \in H \cup \{\mathbf{0}\}$, so that

$$(6.4) \quad 1 + |\tilde{\mathbf{h}}| - \tilde{n}_o - \tilde{n}_i \geq f(\tilde{\mathbf{h}}) \geq \alpha = 1 + |\mathbf{h}| - n_o - n_i$$

and, if $\tilde{\mathbf{h}} \neq \mathbf{h}$, then

$$(6.5) \quad 1 + |\mathbf{h} - \tilde{\mathbf{h}}| - (n_o - \tilde{n}_o) - (n_i - \tilde{n}_i) \geq f(\mathbf{h} - \tilde{\mathbf{h}}) \geq \alpha = 1 + |\mathbf{h}| - n_o - n_i.$$

Equations (6.4) and (6.5) together are possible only in the case when $\alpha = 1$. So, for $\alpha \geq 2$, surely $\mathbf{h} = \tilde{\mathbf{h}}$, which ends the proof, as follows:

$$\sum_{\mathbf{a}, \mathbf{a}' \in \mathcal{A}^{n_o}, \mathbf{a} \neq \mathbf{a}'} \sum_{\mathbf{b}, \mathbf{b}' \in \mathcal{B}^{n_i}, \mathbf{b} \neq \mathbf{b}'} \mathbb{P}(E_{\mathbf{a},\mathbf{b}} \cap E_{\mathbf{a}',\mathbf{b}'}) \leq |\mathcal{A}|^{2n_o} |\mathcal{B}|^{2n_i} \left(\frac{2e|\mathbf{h}|}{r(N+\nu_o)} \right)^{2|\mathbf{h}|}.$$

For $\alpha = 1$, instead, note that in this case $\tilde{n}_0 + \tilde{n}_i = |\tilde{\mathbf{h}}|$ and $n_0 + n_i = |\mathbf{h}|$. We can estimate the following:

$$\begin{aligned} \sum_{\substack{\mathbf{a}, \mathbf{a}' \in \mathcal{A}^{n_o} \\ \mathbf{a} \neq \mathbf{a}'}} \mathbb{P}(E_{\mathbf{a}} \cap E_{\mathbf{a}'}) &= \sum_{1 \leq \tilde{n}_o \leq n_o} \sum_{\substack{\mathbf{a}, \mathbf{a}' \in \mathcal{A}^{n_o} \\ 1 \leq \text{d}_H(\mathbf{a}, \mathbf{a}') \leq \tilde{n}_o}} \sum_{1 \leq \tilde{n}_i \leq n_i} \sum_{\substack{\mathbf{b}, \mathbf{b}' \in \mathcal{B}^{n_i} \\ 1 \leq \text{d}_H(\mathbf{b}, \mathbf{b}') \leq \tilde{n}_i}} \mathbb{P}(E_{\mathbf{a},\mathbf{b}} \cap E_{\mathbf{a}',\mathbf{b}'}) \\ &\leq \sum_{1 \leq \tilde{n}_o \leq n_o} \binom{n_o}{\tilde{n}_o} |\mathcal{A}|^{n_o + \tilde{n}_o} \sum_{1 \leq \tilde{n}_i \leq n_i} \binom{n_i}{\tilde{n}_i} |\mathcal{B}|^{n_i + \tilde{n}_i} \left(\frac{e(|\mathbf{h}| + \tilde{n}_o + \tilde{n}_i)}{r(N+\nu_o)} \right)^{|\mathbf{h}| + \tilde{n}_o + \tilde{n}_i} \\ &\leq \frac{|\mathcal{A}|^{n_o} |\mathcal{B}|^{n_i}}{[r(N+\nu_o)]^{|\mathbf{h}|}} 2e|\mathbf{h}| \left[\left(1 + \frac{|\mathcal{A}|2e|\mathbf{h}|}{r(N+\nu_o)} \right)^{n_o} - 1 \right] \left[\left(1 + \frac{|\mathcal{B}|2e|\mathbf{h}|}{r(N+\nu_o)} \right)^{n_i} - 1 \right] \\ &\leq \frac{|\mathcal{A}|^{n_o} |\mathcal{B}|^{n_i}}{[r(N+\nu_o)]^{|\mathbf{h}|}} C \end{aligned}$$

for some constant $C > 0$. Choosing suitably large constants L and Λ in the definition of $l_{\max}$ and $\lambda_{\max}$ ensures that $|\mathcal{A}|$ and $|\mathcal{B}|$ are small enough to have $C < \frac{1}{|Y|^{|\mathbf{h}|}}$, which will be useful later. $\square$

Now we can conclude the proof of Proposition 6.9. Using Lemmas 6.11 and 6.12, for $\alpha \geq 2$ we get

$$\sum_{\mathbf{a} \in \mathcal{A}^{n_o}} \mathbb{P}(E_{\mathbf{a}}) - \sum_{\substack{\mathbf{a}, \mathbf{a}' \in \mathcal{A}^{n_o} \\ \mathbf{a} \neq \mathbf{a}'}} \mathbb{P}(E_{\mathbf{a}} \cap E_{\mathbf{a}'}) \geq \frac{|\mathcal{A}|^{n_o} |\mathcal{B}|^{n_i}}{[r(N+\nu_o)]^{|\mathbf{h}|}} \left(\frac{1}{|Y|^{|\mathbf{h}|}} - |\mathcal{A}|^{n_o} |\mathcal{B}|^{n_i} \frac{(2e|\mathbf{h}|)^{2|\mathbf{h}|}}{[r(N+\nu_o)]^{|\mathbf{h}|}} \right).$$

For $N \rightarrow \infty$, as $|\mathcal{A}| \asymp N$ and $|\mathcal{B}| \asymp N$, we have $\frac{|\mathcal{A}|^{n_o} |\mathcal{B}|^{n_i}}{[r(N+\nu_o)]^{|\mathbf{h}|}} \asymp N^{-\alpha+1}$. We conclude by noticing that $|\mathcal{A}|^{n_o} |\mathcal{B}|^{n_i} \frac{(2e|\mathbf{h}|)^{2|\mathbf{h}|}}{[r(N+\nu_o)]^{|\mathbf{h}|}} \asymp N^{-\alpha+1} \rightarrow 0$.

For $\alpha = 1$, Lemmas 6.11 and 6.12 give

$$\sum_{\mathbf{a} \in \mathcal{A}^{n_o}} \mathbb{P}(E_{\mathbf{a}}) - \sum_{\substack{\mathbf{a}, \mathbf{a}' \in \mathcal{A}^{n_o} \\ \mathbf{a} \neq \mathbf{a}'}} \mathbb{P}(E_{\mathbf{a}} \cap E_{\mathbf{a}'}) \geq \frac{|\mathcal{A}|^{n_o} |\mathcal{B}|^{n_i}}{[r(N+\nu_o)]^{|\mathbf{h}|}} \left(\frac{1}{|Y|^{|\mathbf{h}|}} - C \right)$$

for some $0 < C < \frac{1}{|Y|^{|\mathbf{h}|}}$. Then the equality $n_o + n_i = |\mathbf{h}|$ ensures that the right-hand side is bounded from below by a strictly positive constant.

6.3. Proof of Propositions 5.5, 5.6, and 5.7. We prove here Proposition 5.6 and Proposition 5.7; clearly the latter one also implies the weaker Proposition 5.5, which can be obtained as a special case just taking $\rho_2 = 0$ so that $\mathbf{w}_G = \mathbf{w}_1$ and $d_f^o = d_{f,1}^o$.

LEMMA 6.13. *Under the same assumptions as in Proposition 5.7 for all $\mathbf{h} = (\mathbf{h}_1, \mathbf{h}_2) \in H \subseteq \mathbb{N}^{\rho_1} \times \mathbb{N}^{\rho_2}$ we have the following:*

- $1 \leq n_o(\mathbf{h}) \leq \lfloor |\mathbf{h}_1|/d_{f,1}^o \rfloor$,
- $0 \leq n_i(\mathbf{h}) \leq \lfloor |\mathbf{h}_1|/2 \rfloor + |\mathbf{h}_2|$,
- $1 + |\mathbf{h}_1| - \lfloor |\mathbf{h}_1|/d_{f,1}^o \rfloor - \lfloor |\mathbf{h}_1|/2 \rfloor \leq f(\mathbf{h}) \leq |\mathbf{h}|$.

Proof. The upper bounds for $n_o(\mathbf{h})$ and $n_i(\mathbf{h})$ are an immediate consequence of the definition of $d_{f,1}^o$ and of the $\mathbf{w}_1$ -recursiveness of ϕ_i . For the lower bounds, see Remark 5.2. Then the estimations for $f(\mathbf{h})$ directly follow.

The definitions of d_f^o and H now clearly imply that $\alpha \leq d_f^o$, while the lower bound for α comes from the following property. We omit its easy proof, already sketched in [2].

PROPOSITION 6.14. *Given any constant $c \geq 2$, the function $g : \mathbb{N} \rightarrow \mathbb{N}$ defined by $g_c(h) = 1 + h - \lfloor h/c \rfloor - \lfloor h/2 \rfloor$ has $\min_{h \geq c} g_c(h) = \lfloor (c+1)/2 \rfloor$ and*

$$\arg \min_{h \geq c} g_c(h) = \begin{cases} 2\mathbb{N}^* & \text{if } c = 2 \\ \{c, c+1, 2c\} & \text{if } c = 3 \\ \{c, c+1\} & \text{if } c \text{ is odd, } c \geq 5 \\ \{c\} & \text{if } c \text{ is even, } c \geq 4. \end{cases}$$

The lower bound for $f(\mathbf{h})$ in Lemma 6.13 can be rewritten as $f(\mathbf{h}) \geq g_{d_{f,1}^o}(|\mathbf{h}_1|)$. For all $\mathbf{h} \in H$, clearly $|\mathbf{h}_1| \geq d_{f,1}^o$ and so, by Proposition 6.14,

$$\alpha = \min_{\mathbf{h} \in H} f(\mathbf{h}) \geq \lfloor (d_{f,1}^o + 1)/2 \rfloor.$$

Clearly, this lower bound for α immediately means that $d_{f,1}^o \geq 2$ gives $\alpha \geq 1$, and $d_{f,1}^o \geq 3$ gives $\alpha \geq 2$.

Finally, we prove that $\mathcal{H}$ is a finite set under the assumption that $\rho_2 = 0$ and $d_f^o \geq 3$. For any $\mathbf{h} \in \mathcal{H}$, i.e., such that $f(\mathbf{h}) = \alpha$, by Lemma 6.13 we get

$$\alpha = f(\mathbf{h}) \geq 1 + |\mathbf{h}| - \lfloor |\mathbf{h}|/d_f^o \rfloor - \lfloor |\mathbf{h}|/2 \rfloor \geq 1 + |\mathbf{h}| \left(\frac{1}{2} - \frac{1}{d_f^o} \right)$$

which gives $|\mathbf{h}| \leq (\alpha - 1) \frac{2d_f^o}{d_f^o - 2}$, thus proving that $\mathcal{H}$ is finite.

7. Examples. In this section we consider particular cases where we can characterize α and q^* exactly, or can give tighter bounds than the general ones. We will particularly focus on the relevant examples introduced in section 4.3. Throughout this section, we will consider $\Gamma = \mathbb{Z}_m$; in some cases we will restrict our attention to m -PSK-AWGN channels.

7.1. Classical free $\mathbb{Z}_m$ serial scheme. We call this scheme classical, because it is the simplest and most natural generalization of the classical binary serial concatenations introduced in [2].

In the general scheme, take $U = \mathbb{Z}_m^k$, $Y = \mathbb{Z}_m$, $\Gamma = \mathbb{Z}_m$, $G_N = S_{r(N+\nu_o)}$, and consider constituent encoders which are rational matrices $\phi_o \in \mathbb{Z}_m(D)^{k \times r}$ and $\phi_i \in \mathbb{Z}_m(D)^{s \times l}$. See the appendix for properties of convolutional encoders in this particular setting.

Consider symbol error probability with respect to Hamming weight on $\mathbb{Z}_m$ (extended componentwise). Take as interconnection group $G_N = S_{r(N+\nu_o)}$; i.e., all of the permutations moving around the elements of $\mathbb{Z}_m$. Clearly, the invariant weight w_G will be the type weight w_T on $\mathbb{Z}_m$ (extended componentwise). Notice that, in this scheme we can think of symbols in the most intuitive way; i.e., to be the elements of $\mathbb{Z}_m$, both in input and in the interconnection and at the output. Clearly, if one takes $m = 2$, then symbols are just bits, and type weight and Euclidean weight are equal to Hamming weight; therefore we get the classical binary schemes introduced in [2].

For this ensemble, we have an explicit expression for α if m is a power of 2, and tight bounds for α for general m ; we also have simple examples showing that, without more information about the constituent encoders, nothing tighter than these bounds can be found.

Let $m = p_1^{\alpha_1} \cdots p_l^{\alpha_l}$ be the prime factors decomposition of m , and let $\phi_{j,o} : \mathbb{Z}_{p_j}^{kN} \rightarrow \mathbb{Z}_{p_j}^{rN}$ be obtained by taking the restriction of ϕ_o to inputs in $\frac{m}{p_j} \mathbb{Z}_m^k$ and then identifying $\frac{m}{p_j} \mathbb{Z}_m$ with $\mathbb{Z}_{p_j}$ through the natural fields isomorphism. With this notation, the following bounds for α hold true.

PROPOSITION 7.1. *For the classical free $\mathbb{Z}_m$ ensemble,*

$$\lfloor (d_f^o + 1)/2 \rfloor \leq \alpha \leq d_f^o - \lfloor d_f^o/p_{\min} \rfloor,$$

where $p_{\min} = \min\{p_j : d_f^o = d_f(\phi_{j,o})\}$.

Proof. We already have the bound $\lfloor (d_f^o + 1)/2 \rfloor \leq \alpha \leq d_f^o$ (Proposition 5.5), so we just need to prove the tighter upper bound.

Notice that, by Proposition A.2, $\mathcal{P} := \{p_j : d_f^o = d_f(\phi_{j,o})\} \neq \emptyset$. We want to prove that $\alpha \leq d_f^o - \lfloor d_f^o/p \rfloor$ for all $p \in \mathcal{P}$. Therefore, we must fix any $p \in \mathcal{P}$; consider a word $\mathbf{c} \in \frac{m}{p} \phi_o^N(\mathbb{Z}_m^{kN})$ such that $w_H(\mathbf{c}) = d_f^o$; let $\mathbf{h} = \mathbf{w}_T(\mathbf{c})$.

Let $a_1, \dots, a_{d_f^o} \in \frac{m}{p} \mathbb{Z}_m \setminus \{0\}$ be the nonzero symbols of the word $\mathbf{c}$ (possibly with the same symbol repeated many times). Consider $a_1, \dots, a_p$. By Lemma A.3 (applied to $\mathbb{Z}_p$), there exist indexes $\{j_1, \dots, j_n\} \subseteq \{1, \dots, p\}$ such that $a_{j_1} + \dots + a_{j_n} = 0 \pmod{m}$. Then, by Proposition 3.6, there exist distinct times $t_1, \dots, t_n$ such that $\phi_i(a_{j_1} D^{t_1} + \dots + a_{j_n} D^{t_n})$ has finite support; i.e., is formed by some (at least one) error events. By applying the same argument to $a_{p+1}, \dots, a_{2p}$ and so on up to $a_{(\lfloor d_f^o/p \rfloor - 1)p+1}, \dots, a_{\lfloor d_f^o/p \rfloor p}$, we obtain that $n_i(\mathbf{h}) \geq \lfloor d_f^o/p \rfloor$. Clearly, $n_o(\mathbf{h}) = 1$ and so we can conclude that $\alpha \leq f(\mathbf{h}) \leq 1 + d_f^o - 1 - \lfloor d_f^o/p \rfloor$. $\square$

From Proposition 7.1, together with Proposition A.2 (see the appendix), we get the exact value of α for the case when m is a power of 2.

COROLLARY 7.2. *For the classical free $\mathbb{Z}_m$ ensemble, if m is a power of 2, then*

$$\alpha = \lfloor (d_f^o + 1)/2 \rfloor.$$

It is more difficult to get an explicit formula for q^* . We can just notice that whenever $\alpha = \lfloor (d_f^o + 1)/2 \rfloor$ (in particular for the classical free $\mathbb{Z}_m$ ensemble when m is a power of 2), the inequalities $n_o(\mathbf{h}) \leq \lfloor |\mathbf{h}|/d_f^o \rfloor$, $n_i(\mathbf{h}) \leq \lfloor |\mathbf{h}|/2 \rfloor$ (Lemma 6.13), and Proposition 6.14 make simpler the description of $\mathcal{H}$. When $m = 2$, the description of $\mathcal{H}$ gets even simpler, because the type weight is a scalar, equal to the Hamming weight, and for all $h \in H$, $n_i(h) = \lfloor h/2 \rfloor$. This allows us to find the following explicit formula $q^* = Q(d^*)$ in the binary classical ensemble (d^* was already described by Benedetto et al. [2], but here our result is more precise for odd values of d_f^o). Define $d_{f,2}^i$ and $d_{f,3}^i$ to be the minimum output Hamming weight of a regular error event of the inner encoder constrained to input Hamming weight 2 and 3, respectively ($d_{f,3} = +\infty$ if such an event does not exist). Also define $d_{1,\text{term}}^i$ to be the minimum output Hamming weight of a terminated error event with input Hamming weight 1.

PROPOSITION 7.3. *For the binary classical ensemble, $q^* = Q(d^*)$, where*

- if d_f^o is even, then $d^* = \frac{1}{2}d_f^o d_{f,2}^i$;
- if d_f^o is odd ($d_f^o \geq 5$), then

$$d^* = \begin{cases} \frac{d_f^o - 3}{2} d_{f,2}^i + \min \left\{ d_{f,2}^i + d_{1,\text{term}}^i, d_{f,3}^i, 2d_{f,2}^i \right\} & \text{if } d_f^o + 1 \in H, \\ \frac{d_f^o - 3}{2} d_{f,2}^i + \min \left\{ d_{f,2}^i + d_{1,\text{term}}^i, d_{f,3}^i \right\} & \text{if } d_f^o + 1 \notin H, \end{cases}$$

- if $d_f^o = 3$, then

$$d^* = \begin{cases} \min \left\{ d_{f,2}^i + d_{1,\text{term}}^i, d_{f,3}^i, 2d_{f,2}^i \right\} & \text{if } 4 \in H \\ \min \left\{ d_{f,2}^i + d_{1,\text{term}}^i, d_{f,3}^i, 3d_{f,2}^i \right\} & \text{if } 4 \notin H. \end{cases}$$

Now we give three examples with $\Gamma = \mathbb{Z}_m$ and with $G_N = S_{r(N+\nu_o)}$. For the computation of q^* , we consider the specific case of the m -PSK-AWGN channel, for which we can find an explicit expression. Remember that, for the S -AWGN channel, given a type $\mathbf{d}$, $Q(\mathbf{d}) = \frac{1}{2} \operatorname{erfc} \sqrt{(\sum_g d_g w_E(g)) E_s / N_0}$. Note that when S is m -PSK, $w_E(g) = \sin^2(g\pi/m)$.

Example 7.4 (repeat-accumulate codes). The encoders are $\phi_o = \operatorname{Rep}_r$ (with $r \geq 2$) and $\phi_i = \frac{1}{1-D}$. We assume that the termination rule for the accumulator is the one that always brings to the zero state in one trellis step (using the input $-a$ if we are in state a).

We obtain $\alpha = \min\{r-1, r - \lfloor r/p \rfloor\}$, where p is the smallest prime divisor of m , and $q^* = \frac{1}{2} \operatorname{erfc} \sqrt{d^* E_s / N_0}$, where we have the following:

- For $m = 2$, $d^* = \lfloor (r+1)/2 \rfloor$.
- For even $m \geq 4$,

$$d^* = \begin{cases} r w_E(1) & \text{if } r = 2 \text{ or } r = 3 \\ \lfloor (r+1)/2 \rfloor & \text{if } r \geq 4. \end{cases}$$

- For odd $m \geq 3$, let p be the smallest prime divisor of m , and let $n = m/p$. Define

$$d_*(r, m) = \left\lfloor \frac{r}{p} \right\rfloor \sum_{i=1}^{p-1} w_E(in) + \min_{1 \leq j \leq p-1} \sum_{i=1}^{r \bmod p} w_E(ijn \bmod p).$$

Then

$$d^* = \begin{cases} r\mathbf{w}_E(1) & \text{if } r < p \\ d_*(r, m) & \text{if } r \geq 2p \\ \min\{r\mathbf{w}_E(1), d_*(r, m)\} & \text{if } p \leq r < 2p. \end{cases}$$

A sketch of how to get this result follows:

- $m = 2$:

We can use the explicit expressions we have for α and d^* in the binary case. For Rep_r , $d_f^o = r$ and $d_f^o + 1 \notin \mathcal{H}$. For the accumulator, $d_{f,2}^i = d_{1,\text{term}}^i = 1$ and $d_{f,3}^i = +\infty$.

- even $m \geq 4$:

Notice that $d_f(\frac{m}{p_i}\text{Rep}_r) = r$ for all prime $p_i|m$, so that, if $2|m$ by Proposition 7.1, then $\alpha = \lfloor (r+1)/2 \rfloor$. Then compute the following:

- $H = (r\mathbb{N})^{m-1} \setminus \{\mathbf{0}\}$.
- If $r \geq 4$, then $\mathcal{H} = \{\mathbf{h} \in H : |\mathbf{h}| = r, n_o(\mathbf{h}) = 1, n_i(\mathbf{h}) = \lfloor |\mathbf{h}|/2 \rfloor\} = \{\mathbf{k}\}$, where $\mathbf{k}_{m/2} = r$, and $\mathbf{k}_i = 0$ for all $i \neq m/2$. This gives $d^* = \lfloor (r+1)/2 \rfloor \mathbf{w}_E(m/2)$. Notice that $\mathbf{w}_E(m/2) = 1$.
- If $r = 3$, then we have $\mathcal{H} = \{\mathbf{k}, 2\mathbf{k}, \mathbf{k}^{(1)}, \dots, \mathbf{k}^{(m/2-1)}\}$ with $\mathbf{k}$ as above and $\mathbf{k}^{(j)}$ defined by $\mathbf{k}_j^{(j)} = \mathbf{k}_{m-j}^{(j)} = r$ and $\mathbf{k}_i^{(j)} = 0$ for all $i \notin \{j, m-j\}$. Then

$$d^* = \min\{\lfloor \frac{r+1}{2} \rfloor \mathbf{w}_E(\frac{m}{2}), r\mathbf{w}_E(1), \dots, r\mathbf{w}_E(m)\} = \min\{2\mathbf{w}_E(\frac{m}{2}), 3\mathbf{w}_E(1)\}.$$

Then $m \geq 4$ implies $\mathbf{w}_E(1) \leq 1/2 = \mathbf{w}_E(m/2)/2$, so $d^* = 3\mathbf{w}_E(1)$.

- If $r = 2$, then $\mathcal{H} = \{\mathbf{h} \in H : \mathbf{h}_i = \mathbf{h}_{m-i} \forall i = 1, \dots, m/2 - 1\}$, and, with the same reasonings as above, we find again that $d^* = \lfloor (r+1)/2 \rfloor$ for $m = 2$ and $d^* = d^*((2, 0, \dots, 0, 2)) = 2\mathbf{w}_E(\phi_i(1-D)) = 2\mathbf{w}_E(1)$ for $m \geq 4$.

- odd $m \geq 3$:

- $H = (r\mathbb{N})^{m-1} \setminus \{\mathbf{0}\}$.
- Compute

$$\min_{\mathbf{h} \in H : |\mathbf{h}| = kr} f(\mathbf{h}) = \begin{cases} 1 + kr - k - \frac{k}{2}r & \text{if } k \text{ even,} \\ 1 + kr - k - \frac{k-1}{2}r - \lfloor r/p \rfloor & \text{if } k \text{ odd.} \end{cases}$$

Notice that both expressions are nondecreasing in k and increasing in k if $r \geq 3$, so that $\alpha = \min\{r-1, r - \lfloor r/p \rfloor\}$.

- If $r = 2$, then $\alpha = 1$ and $\mathcal{H} = \{\mathbf{h} \in H : \mathbf{h}_i = \mathbf{h}_{m-i} \forall i = 1, \dots, m/2 - 1\}$, so that $d^* = r\mathbf{w}_E(1)$, obtained for $\mathbf{h} = r\mathbf{e}_1 + r\mathbf{e}_{-1}$.
- If $2 < r < p$, then $\alpha = r-1$ and $\mathcal{H} = \{\mathbf{k}^{(1)}, \dots, \mathbf{k}^{((m-1)/2)}\}$, the $\mathbf{k}^j$'s defined as for even m . So again $d^* = r\mathbf{w}_E(1)$.
- If $r \geq 2p$, then $\alpha = r - \lfloor r/p \rfloor$ and $\mathcal{H} = \{r\mathbf{e}_{m/p}, r\mathbf{e}_{2m/p}, \dots, r\mathbf{e}_{(p-1)m/p}\}$, from which the expression for d^* easily follows.
- If $p \leq r < 2p$, then $\alpha = r-1 = r - \lfloor r/p \rfloor$, and $\mathcal{H}$ is the union of the set $\mathcal{H}$ computed for $r < p$ and the one computed for $r \geq 2p$; thus, d^* is the minimum of the two values obtained before.

The repeat-accumulate code on $\mathbb{Z}_3$ ($r \geq 3$) is an example where the upper bound $\alpha \leq d_f^o - \lfloor d_f^o/3 \rfloor$ is reached with equality. Now, we show another simple repeat-convolute code on $\mathbb{Z}_3$ such that the lower bound $\alpha \geq \lfloor (d_f^o + 1)/2 \rfloor$ is reached with

equality, showing that the bounds in Proposition 7.1 are the best possible for general m .

Example 7.5. Consider $m = 3$, $\phi_o = \text{Rep}_r$ ($r \geq 2$), and $\phi_i = 1/(1+D) = \sum_{t \geq 0} D^{2t} + 2D^{2t+1}$, with the termination rule that always brings the zero state back to zero in one trellis step (i.e., if at time t the codeword has $c_t = a$, we terminate using the input $u_{t+1} = -a$ if t is even, or a if t is odd).

Then, as for the repeat-accumulate, $H = r\mathbb{N}^2 \setminus \{(0,0)\}$, and given $\mathbf{h} = (rh_1, rh_2)$ we have $n_o(\mathbf{h}) = h_1 + h_2$. But now, when we look at the inner encoder to compute $n_i(\mathbf{h})$, we find $n_i(\mathbf{h}) = \lfloor |\mathbf{h}|/2 \rfloor$, because all of the following inputs produce a complete error event of ϕ_i : $\mathbf{u} = D^t + D^{t+1}$, $\mathbf{u} = 2D^t + 2D^{t+1}$, $\mathbf{u} = D^t + 2D^{t+2}$, and $\mathbf{u} = 2D^t + D^{t+2}$.

As a consequence,

$$\alpha = \lfloor (r+1)/2 \rfloor.$$

Let's compute $q^* = \frac{1}{2} \text{erfc} \sqrt{d^* E_s / N_0}$ for this example. First of all we need $\mathcal{H}$:

$$\mathcal{H} = \begin{cases} H & \text{if } r = 2 \\ \{(r,0), (0,r), (r,r), (2r,0), (0,2r)\} & \text{if } r = 3 \\ \{(r,0), (0,r)\} & \text{if } r > 3. \end{cases}$$

Now consider that $w_E(\phi_i(D^t + D^{t+1})) = w_E(D^t) = w_E(1) = 3/4$; analogously $w_E(\phi_i(2D^t + 2D^{t+1})) = w_E(2) = 3/4$; while $w_E(\phi_i(D^t + 2D^{t+2})) = w_E(D^t + 2D^{t+1}) = w_E(1) + w_E(2) = 3/2$ and $w_E(\phi_i(2D^t + D^{t+2})) = 3/2$. Assuming that termination is always done in one single step, we also have that $w_E(\phi_i^N(D^{N-1})) = w_E(\phi_i^N(2D^{N-1})) = 3/4$. Finally, we get $d^* = \frac{3}{4} \lfloor \frac{r+1}{2} \rfloor$.

Notice that for repeat-accumulate codes (Example 7.4), $\alpha = \lfloor (r+1)/2 \rfloor$ for all even m . This is true for all repeat-convolute codes by Proposition 7.1 together with the remark that $d_f(\frac{m}{p_i} \text{Rep}_r) = r$ for all prime $p_i | m$. However, for a general outer encoder ϕ_o this is not true: the assumption that m is a power of two is essential in Corollary 7.2, as shown by the following example.

Example 7.6. Let $m = 6$. Consider ϕ_o , which is the following slight variation of a repeat code: $\phi_o = [1, 1, 1, 1, 3]^T$. Let the inner encoder be the accumulator $\phi_i = \frac{1}{1-D}$. For $p_1 = 2$ we have $\phi_{1,o} = [1, 1, 1, 1, 1]^T$, which has $d_f(\phi_{1,o}) = 5$, while for $p_2 = 3$ we have $\phi_{2,o} = [1, 1, 1, 1, 0]^T$, which has $d_f(\phi_{2,o}) = 4$, and so $d_f^o = d_f(\phi_{2,o}) = 4$. The bounds given in Proposition 7.1 give us $2 \leq \alpha \leq 3$, and now we will show that $\alpha = 3$. Notice that $f(\mathbf{h}) = \alpha$ implies that $1 + |\mathbf{h}| - \frac{|\mathbf{h}|}{d_f^o} - \frac{|\mathbf{h}|}{2} \leq \alpha \leq 3$ and then $|\mathbf{h}| \leq 8$, so $\mathcal{H} \subseteq \{\mathbf{h} \in H : |\mathbf{h}| \leq 8\}$. There are seven elements in H with $|\mathbf{h}| \leq 8$. By computing $f(\mathbf{h})$ for all of them, we get $\alpha = 3$, $\mathcal{H} = \{(0,4,0,0,0), (0,0,5,0,0), (0,0,0,4,0)\}$, and, finally, for 6-PSK-AWGN channel, $q^* = \frac{1}{2} \text{erfc} \sqrt{d^* E_s / N_0}$ with $d^* = 2w_E(2) + w_E(4) = 9/4$, reached when $\mathbf{h} = (0,4,0,0,0)$ and $\mathbf{h} = (0,0,0,4,0)$.

7.2. Subgroups of permutations for the $\mathbb{Z}_m$ scheme. In the previous section, we have considered $\mathbb{Z}_m$ -schemes

$$\xrightarrow{\mathbb{Z}_m^{kN}} \boxed{\phi_o^N} \xrightarrow{\mathbb{Z}_m^{r(N+\nu_o)}} \boxed{\pi_N} \xrightarrow{\mathbb{Z}_m^{sM_N}} \boxed{\phi_i^N} \xrightarrow{\mathbb{Z}_m^{l(M_N+\nu_i)}}$$

by taking $U = \mathbb{Z}_m^k$, $Y = \mathbb{Z}_m$, $\Gamma = \mathbb{Z}_m$ in the general serial scheme. However, we can also obtain some $\mathbb{Z}_m$ schemes by taking $Y = \mathbb{Z}_m^a$. Then, if we consider on $\mathbb{Z}_m^a$ a weight

given by the componentwise extension of the type weight on $\mathbb{Z}_m$, then we again get the same scheme as above. However, in this case we can also consider permutations moving around not single elements of $\mathbb{Z}_m$, but only the vectors in $\mathbb{Z}_m^a$, so that the invariant weight is the type weight on $\mathbb{Z}_m^a$. Or, on the contrary, we can consider a *separate channels permutation*: the invariant weight is $\mathbf{w} \in (\mathbb{N}^{m-1})^a$ given by the type weight on each separate component of $\mathbb{Z}_m^a$.

Even though these schemes are quite similar to the classical one, differing only for a restriction of the permutations to a subgroup of $S_{r(N+\nu_o)}$, Proposition A.2 and Corollary 7.2 do not hold true. We give here a simple example, for the binary case $m = 2$, and for the *separate channels permutation*, where $\alpha > \lfloor (d_f^o + 1)/2 \rfloor$.

Example 7.7. Consider the following outer and inner binary encoders:

$$\phi_o = \left[1, \frac{1}{1+D+D^3} \right]^T, \quad \phi_i = \begin{bmatrix} \frac{1}{1+D} & 0 \\ 0 & \frac{1}{1+D} \end{bmatrix},$$

and consider the *separate channels permutation* ensemble (here $m = 2$ and $a = 2$, and therefore $\mathbf{w} \in \mathbb{N}^2$ is the Hamming weight of the two streams). The outer encoder has free distance $d_f^o = 4$ and all of the words $\mathbf{c}$ of the outer code such that $d_H(\mathbf{c}) = d_f^o$ are obtained when input is $1 + D + D^3$ or its shifts and have $\mathbf{w}(\mathbf{c}) = (3, 1)$. The inner encoder is simply the rate-1 accumulator, but acting separately on the two input streams.

We claim that for this scheme $\alpha = 3 > \lfloor (d_f^o + 1)/2 \rfloor = 2$. In fact, we know that $\alpha \geq \lfloor (d_f^o + 1)/2 \rfloor = 2$, where equality could be reached only if there was $\mathbf{h} \in H$ such that $|\mathbf{h}| = 4$, $n_o(|\mathbf{h}|) = 1$, $n_i(\mathbf{h}) = 2$, but this is not possible as the only $\mathbf{h} \in H$ such that $|\mathbf{h}| = 4$ is $\mathbf{h} = (3, 1)$, which has $n_o(\mathbf{h}) = 1$ and $n_i(\mathbf{h}) = 1$, giving $f(\mathbf{h}) = 3$ so $\alpha = 3$.

By an exhaustive listing of all small-weight codewords, we can also find $\mathcal{H}$, noting that $\mathbf{h} \in \mathcal{H}$ implies $|\mathbf{h}| \leq 8$. Then we can find $q^* = Q(3)$.

Remark 7.8. The separate channels ensemble is particularly interesting because it allows us to also include in our generalized serial concatenations traditional parallel turbo codes (as it was already noticed, e.g., in [1]): a turbo code with b parallel branches, each with an encoder ψ_j of rate k_j/n_j , can always be seen as a repeat-convolute scheme, where $\phi_o = \text{Rep}_r$ and $r = \sum k_j$ as the interleaver acts separately on the b streams of $k_j \times N$ bits and ϕ_i is a block diagonal matrix where the blocks are the ψ_j 's.

7.3. Structured LDPC ensemble. For a description of these schemes, see section 4.3. Here we give some statements about the parameters α and d^* .

First of all, we have the following tight bounds for α .

PROPOSITION 7.9. *For the structured LDPC ensemble,*

$$\lfloor (c + 1)/2 \rfloor \leq \alpha \leq c - \lfloor c/p_{\min} \rfloor,$$

where $p_{\min} = \min\{p_j \geq 2 : p_j | m\}$.

Proof. By Proposition 5.7, we have $\alpha \geq \lfloor (d_{f,1}^o + 1)/2 \rfloor = \lfloor (c + 1)/2 \rfloor$. The proof of the upper bound is similar to the proof of Proposition 7.1. Notice that here $p_{\min}$ is computed considering all prime factors of m because the outer encoder is a simple repetition code. $\square$

In particular, this proposition implies that for all even m the interleaver gain is

$$\alpha = \lfloor (c + 1)/2 \rfloor.$$

In the binary case ($m = 2$), we can also characterize q^* . In fact, we can easily describe $\mathcal{H}$ as

$$\mathcal{H} = \begin{cases} \{(2w, w) : w \in \mathbb{N}^*\} & \text{if } c = 2 \\ \{(3, 1), (6, 2)\} & \text{if } c = 3 \\ \{(c, 1)\} & \text{if } c \geq 4. \end{cases}$$

Then compute $q^* = Q(d^*)$,

$$d^* = \begin{cases} 1 & \text{if } c \text{ is even} \\ 2 & \text{if } c = 3 \\ 1 + \min \{d_{1,\text{term}}(\psi), d_{f,3}(\psi)\} & \text{if } c \text{ is odd, } c \geq 5, \end{cases}$$

where $d_{1,\text{term}}(\psi), d_{f,3}(\psi)$ are defined as $d_{1,\text{term}}^i, d_{f,3}^i$ in Proposition 7.3 but referring here to ψ instead of ϕ_i . If the inner encoder is truncated instead of terminated, then $d^* = 2$ for all odd c .

Notice that the choice of ψ has almost no influence on d^* . This happens because pairs of bits which are repetition of a same information bit can be permuted by some interleaver in such a way that they are summed up by Sum_d , producing a zero output. The value of d^* is given by this worst case scenario. This remark suggests considering interleavers with a better spread, enforcing the fact that 1's coming from the same error event of Rep_c cannot end up in positions where they would be summed up by Sum_d . However, the analysis of such a smaller ensemble, with a set of interleavers which is not a group, is beyond the scope of this paper. See [24] for the results that can be obtained and a sketch of the techniques used; more details on structured LDPC codes will be given in a paper now in preparation.

For general m there is no explicit simple characterization of q^* , nor is there one for all even m . We can just notice that, on m -PSK-AWGN channels, by the same argument used for $m = 2$, if both m and c are even, then $q^* = \frac{1}{2} \text{erfc} \sqrt{d^* E_s / N_0}$ with $d^* \leq w_E(m/2) = 1$. This upper bound is achieved, for example, by simply taking $\psi = 1/(1-D)$. To see that this upper bound is not always achieved, take, for example, $m = 6$ and $\psi = 1/(1+D)$: we have that $\psi(1+D) = 1$, which is an error event of Euclidean weight $w_E(1) = 1/4$, so that for $c = 2$ or $c = 4$ we have $d^* = \frac{1}{4} + \frac{c}{2} \frac{1}{4} < 1$.

8. Conclusion. In this paper we have studied the average ML performance of a wide class of generalized serial turbo schemes, coupling two convolutional encoders over groups through an interleaver respecting the group structure; these codes are designed to be used on symmetric channels. A particularly relevant example is the case when the convolutional codes are modules on $\mathbb{Z}_m$, the interleaver is a permutation, and the channel is AWGN with m -PSK input constellation.

We have obtained the exact asymptotic decay of the symbol and word error probability when the interleaver length goes to infinity and also the behavior when the SNR goes to infinity. The performance is characterized by the interleaver gain α and the effective distance q^* , which are defined as the solution of an optimization problem and, in general, jointly depend on both constituent encoders, differently from the binary case. To make clear the meaning of these parameters, we have explicitly computed them in some examples encompassing most of the relevant scenarios.

This work is a first attempt to give a rigorous analysis of generalized serial turbo schemes for nonbinary modulations. It leaves many open questions among which the most natural ones are, in our opinion, the following:

- Our analysis provides design parameters. The next research step would be to extensively search for pairs of constituent encoders being good with respect to these parameters, and to confirm the validity of the approach by a significant simulation analysis, as it has already been done for the binary turbo codes. However, here the optimization of the constituent encoders under some fixed complexity bound (e.g., the size of the state spaces) seems a quite challenging problem, due to the complexity of the combinatorial optimization problem.
- The ensemble analysis in this paper is limited to the average behavior. The study of the average properties is usually only the first step in the understanding of a family of codes. The next step is to find the typical performance. For the binary classical serial turbo ensemble, the typical error probability turns out to be much better than the average, with subexponential vs. polynomial decay, but depending on the same design parameters [8]. This is analogous of the well-known behavior of ML-decoded LDPC codes (see [22, 33]): for the (c, d) -regular LDPC ensemble, with $c \geq 3$, the average error probability is known to decrease to zero as $N^{1-c/2}$ for even c and N^{2-c} for odd c , while the error probability of a typical code goes to zero exponentially fast. In both cases, the average is strongly affected by a small fraction of very bad codes, which can be expurgated. The proofs proposed in [8] heavily rely on properties of binary codes, and so the generalization to our setting is not straightforward.

Appendix. Properties of free $\mathbb{Z}_m$ convolutional encoders. In this section, we consider convolutional encoders $\phi : \mathbb{Z}_m^{k\mathbb{N}} \rightarrow \mathbb{Z}_m^{n\mathbb{N}}$ which can be represented as matrices $\phi \in \mathbb{Z}_m^{k \times n}(D) \simeq \mathbb{Z}_m(D)^{k \times n}$. We will call them free $\mathbb{Z}_m$ convolutional encoders. They are the most straightforward generalization of classical binary convolutional encoders, and they have some interesting properties. Let us start with a simple algebraic remark: we know that ϕ can be represented as $\phi = p(D)^{-1}q(D)$ for some $p(D) \in \mathbb{Z}[D] \cap \mathbb{Z}((D))^*$ and $q(D) \in \mathbb{Z}_m[D]^{k \times n}$. Since all of the algebraic structures involved are also $\mathbb{Z}_m$ -modules, it turns out that we can as well assume that $p(D) \in \mathbb{Z}_m[D] \cap \mathbb{Z}_m((D))^*$ which, in practice, means that $p(D)$ has all coefficients in $\mathbb{Z}_m$ and the trailing coefficient is in $\mathbb{Z}_m^*$.

In section 3.3.2, we gave a general definition of recursiveness. In the binary case (for simplicity consider scalar input, i.e., $\phi : \mathbb{Z}_2^{\mathbb{N}} \rightarrow \mathbb{Z}_2^{n\mathbb{N}}$), there are well-known characterizations of w_H -recursive encoders: ϕ is recursive when its shift-register state representation has a feedback, or, equivalently, when $\phi = \frac{1}{q(D)}[p_1(D), \dots, p_n(D)]$ (with $\gcd\{q, p_1, \dots, p_n\} = 1$) has nontrivial denominator, i.e., $q(D) \neq D^h$. This latter characterization allows us to check very easily if an encoder is recursive, which we will now generalize to recursiveness of free $\mathbb{Z}_m$ encoders with respect to Hamming or, equivalently, to type weight in $\mathbb{Z}_m$ (not the Hamming weight in $\mathbb{Z}_m^k$).

First of all, without loss of generality we can restrict ourselves to considering scalar encoders $\phi : \mathbb{Z}_m^{\mathbb{N}} \rightarrow \mathbb{Z}_m^{\mathbb{N}}$: if not so, notice that $\phi : \mathbb{Z}_m^{k\mathbb{N}} \rightarrow \mathbb{Z}_m^{n\mathbb{N}}$ is w -recursive (w being the Hamming or the type weight in $\mathbb{Z}_m$) if and only if each column of its matrix has at least one entry which is a scalar w -recursive encoder.

Then, if m is a prime (so that $\mathbb{Z}_m$ is a field), then $\phi = p(D)/q(D)$ with $p(D), q(D) \in \mathbb{Z}_m[D]$ and $\gcd(p, q) = 1$ is w -recursive if and only if $q(D) \neq D^t$: as in the binary case, we can identify recursive encoders at a glance, just looking at their denominator.

If m is not a prime, then let $m = p_1^{\alpha_1} \dots p_l^{\alpha_l}$ be its prime factors decomposition, and let $\phi_i : \mathbb{Z}_{p_i}^{\mathbb{N}} \rightarrow \mathbb{Z}_{p_i}^{\mathbb{N}}$ be obtained by taking the restriction of ϕ to inputs in $\frac{m}{p_i}\mathbb{Z}_m$ and then identifying $\frac{m}{p_i}\mathbb{Z}_m$ with $\mathbb{Z}_{p_i}$ through the natural fields isomorphism mapping.

PROPOSITION A.1. ϕ is $\mathbf{w}$ -recursive if and only if $\phi_1, \dots, \phi_l$ are $\mathbf{w}$ -recursive.

Proof. The first implication is trivial. Conversely, knowing that $\phi_1, \dots, \phi_l$ is recursive, but we want to show that $w_H(\phi(D)u) = +\infty$ for any $u \in \mathbb{Z}_m \setminus \{0\}$. Since $u < m$, there exist $i \in \{1, \dots, l\}$ and $r \in \mathbb{N}$ such that $p_i^r | u$, $p_i^{r+1} \nmid u$, and $p_i^{r+1} | m$. Consider $\tilde{u} = (p_i^{-1}m)(p_i^{-r}u) = (p_i^{-r-1}m)u$. Clearly, $\tilde{u} \neq 0$ and, by the assumptions made, $w_H(\phi(D)\tilde{u}) = \infty$. This clearly implies that $w_H(\phi(D)u) = \infty$ also. $\square$

The characterization given by Proposition A.1 is helpful because the encoders $\phi_1, \dots, \phi_l$ can be obtained very easily from ϕ : writing $\phi(D) = p(D)/q(D)$ with $p(D)$ and $q(D) \in \mathbb{Z}_m[D]$, we have $\phi_j(D) = \tilde{p}(D)/\tilde{q}(D)$, where $\tilde{p}(D), \tilde{q}(D)$ are polynomials in $\mathbb{Z}_{p_j}$ obtained multiplying each coefficient of $p(D)$ (resp., $q(D)$) by m/p_j (modulo m) and then identifying the corresponding element in $\mathbb{Z}_{p_j}$.

For example, the encoder $\phi : \mathbb{Z}_8^\mathbb{N} \rightarrow \mathbb{Z}_8^\mathbb{N}$ defined by $\phi(D) = \frac{1+3D^2}{1+7D}$ is not recursive, but we cannot tell by simply looking at the denominator, which is nontrivial. However, we can tell by using the definition given in section 3.3.2: notice that $\phi(D) = (1 + 3D^2) \sum_{t \geq 0} D^t$, and then input $u(D) = 2$ produces output $2\phi(D) = 2 + 2D \in \mathbb{Z}_8[D]$. We can also check the recursiveness of ϕ using Proposition A.1: as $m = 8$ has only one prime divisor $p_1 = 2$, we need to check only one encoder $\phi_1 = \frac{1+D^2}{1+D} = 1 + D$ which clearly isn't recursive.

By the same technique of looking at the encoders $\phi_1, \dots, \phi_l$ defined previously, we can obtain a characterization of the free distance of ϕ with respect to Hamming or type weight in $\mathbb{Z}_m$. This characterization is not interesting under a computational point of view, as the computation of the free distance of encoders over fields or rings does not have a different complexity, but it is essential to find tight bounds for the interleaver gain of free $\mathbb{Z}_m$ serial schemes (Proposition 7.1 and Corollary 7.2).

PROPOSITION A.2. Let d_f be the $\mathbf{w}$ -free distance of ϕ , and $d_f(\phi_j)$ be the $\mathbf{w}$ -free distance of ϕ_j , where $\mathbf{w}$ is Hamming or type weight in $\mathbb{Z}_m$ and $\mathbb{Z}_{p_j}$, respectively. Then

$$d_f = \min_{j=1, \dots, l} \{d_f(\phi_j)\}.$$

Proof. Clearly, for all $j = 1, \dots, l$, $d_f(\phi_j) = d_f(\frac{m}{p_j}\phi) \geq d_f$. Now we will prove that there exists at least one j such that $d_f(\phi_j) = d_f$.

Let $\mathcal{C} = \phi(\mathbb{Z}_m^k((D)))$, and let $\mathbf{x} \in \mathcal{C}$ be a codeword such that $w_H(\mathbf{x}) = d_f$. The key remark is that $w_H(\mathbf{x}) = d_f$ implies that all nonzero symbols (i.e., elements of $\mathbb{Z}_m$) of $\mathbf{x}$ have the same annihilator. In fact, $w_H(\mathbf{x}) = d_f$ means that for all $\mathbf{y} \in \mathcal{C} \setminus \{\mathbf{0}\}$, $w_H(\mathbf{y}) \geq w_H(\mathbf{x})$, which implies that

$$\nexists a \in \mathbb{Z}_m \text{ such that } 0 < w_H(a\mathbf{x}) < w_H(\mathbf{x}).$$

Therefore for all $a \in \mathbb{Z}_m$, either $a\mathbf{x} = \mathbf{0}$ or $w_H(a\mathbf{x}) = d_f$; i.e., $a\mathbf{x}_i \neq 0$ for all $\mathbf{x}_i \neq 0$.

This remark implies that there exists $d|m$ (possibly $d = 1$) and there exists p_j , a prime factor of m such that $w_H(d\mathbf{x}) = w_H(\mathbf{x})$ and $p_j d\mathbf{x} = \mathbf{0}$. Now, choosing $\mathbf{c} = d\mathbf{x}$ we have a codeword $\mathbf{c} \in \frac{m}{p_j}\mathcal{C}$ such that $w_H(\mathbf{c}) = d_f$, so that we can conclude that $d_f(\phi_j) = d_f(\frac{m}{p_j}\phi) = d_f$. $\square$

Finally, when proving Proposition 7.1 we also need the following simple lemma, even though it is just a property of $\mathbb{Z}_m$ and not of convolutional codes.

LEMMA A.3. Given $a_1, \dots, a_m \in \mathbb{Z}_m \setminus \{0\}$, there exist indexes $\{i_1, \dots, i_n\} \subseteq \{1, \dots, m\}$ such that $a_{i_1} + \dots + a_{i_n} = 0 \pmod m$.

Proof. By contradiction, assume that $\sum_{i \in \mathcal{I}} a_i \neq 0 \pmod m$ for all nonempty $\mathcal{I} \subseteq \{1, \dots, m\}$. Then, in particular, $\sum_{i=1}^n a_i \neq 0 \pmod m$ for all $n = 2, \dots, m$ and so

$a_1 \notin \{-a_2, -a_2 - a_3, \dots, -\sum_{j=2}^m a_j\}$, which, being a set of $m - 1$ distinct nonzero elements of $\mathbb{Z}_m$ is $\mathbb{Z}_m \setminus \{0\}$ itself, contradicts $a_1 \in \mathbb{Z}_m \setminus \{0\}$. $\square$

REFERENCES

- [1] L. BAZZI, M. MAHDIAN, AND D. SPIELMAN, *The minimum distance of turbo-like codes*, IEEE Trans. Inform. Theory, 2003, submitted; available online at <http://math.mit.edu/~spielman/PAPERS/mindist.pdf>.
- [2] S. BENEDETTO, D. DIVSALAR, G. MONTORSI, AND F. POLLARA, *Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding*, IEEE Trans. Inform. Theory, 44 (1998), pp. 909–926.
- [3] S. BENEDETTO AND G. MONTORSI, *Unveiling turbo codes: Some results on parallel concatenated coding schemes*, IEEE Trans. Inform. Theory, 42 (1996), pp. 409–428.
- [4] S. BENEDETTO AND G. MONTORSI, *Design of parallel concatenated convolutional codes*, IEEE Trans. Commun., 44 (1996), pp. 591–600.
- [5] C. BERROU AND A. GLAVIEUX, *Near optimum error correcting coding and decoding: Turbo codes*, IEEE Trans. Commun., 44 (1996), pp. 1261–1271.
- [6] C. BERROU, A. GLAVIEUX, AND P. THITIMAJSHIMA, *Near Shannon limit error-correction coding and decoding: Turbo codes*, in Proceedings of the IEEE International Conference on Communications, 1993, pp. 1064–1070.
- [7] C. BERROU, M. JÉZÉQUEL, C. DOUILLARD, AND S. KEROUÉDAN, *The advantages of non-binary turbo codes*, in Proceedings of the IEEE Information Theory Workshop, 2001, pp. 61–63.
- [8] G. COMO, F. FAGNANI, AND F. GARIN, *ML performances of serial turbo codes do not concentrate*, in Proceedings of the International Symposium on Turbo Codes, 2006.
- [9] D. DIVSALAR, *A simple tight bound on error probability of block codes with application to turbo codes*, JPL TDA Progress Report, 42-139 (1999), pp. 1–35.
- [10] D. DIVSALAR, S. DOLINAR, AND F. POLLARA, *Iterative turbo decoder analysis based on density evolution*, IEEE J. Select. Areas Commun., 19 (2001), pp. 891–907.
- [11] C. DOUILLARD AND C. BERROU, *Turbo codes with rate- $m/m+1$ constituent convolutional codes*, IEEE Trans. Commun., 53 (2005), pp. 1630–1638.
- [12] T. M. DUMAN AND M. SALEHI, *New performance bounds of turbo codes*, IEEE Trans. Commun., 46 (1998), pp. 717–723.
- [13] S. EILENBERG, *Automata, Machines, and Languages, Vol. A*, Academic Press, New York, 1974.
- [14] H. EL-GAMAL AND A. R. HAMMONS, *Analyzing the turbo decoder using the Gaussian approximation*, IEEE Trans. Inform. Theory, 47 (2001), pp. 671–686.
- [15] F. FAGNANI, *Performance of parallel concatenated coding schemes*, IEEE Trans. Inform. Theory, 54 (2008), pp. 1521–1535.
- [16] F. FAGNANI, B. SCANAVINO, S. ZAMPIERI, AND R. GARELLO, *Some results on combined parallel concatenated schemes with trellis-coded modulation*, in Proceedings of the 2002 International Symposium on Information Theory, 2002, p. 444.
- [17] F. FAGNANI AND S. ZAMPIERI, *Convolutional codes over finitely generated Abelian groups: Some basic results*, in Codes, Systems and Graphical Models, B. Marcus and J. Rosenthal, eds., IMA Vol. Math. Appl. 123, Springer, New York, 2001, pp. 327–346.
- [18] F. FAGNANI AND S. ZAMPIERI, *System-theoretic properties of convolutional codes over rings*, IEEE Trans. Inform. Theory, 47 (2001), pp. 2256–2274.
- [19] F. FAGNANI AND S. ZAMPIERI, *Minimal and systematic convolutional codes over finite Abelian groups*, Linear Algebra Appl., 378 (2004), pp. 31–59.
- [20] G. D. FORNEY, *Geometrically uniform codes*, IEEE Trans. Inform. Theory, 37 (1991), pp. 1241–1260.
- [21] C. FRAGOULI AND R. D. WESEL, *Turbo-encoder design for symbol-interleaved parallel concatenated trellis-coded modulation*, IEEE Trans. Commun., 49 (2001), pp. 425–435.
- [22] R. G. GALLAGER, *Low Density Parity Check Codes*, MIT Press, Cambridge, MA, 1963.
- [23] R. GARELLO, G. MONTORSI, S. BENEDETTO, D. DIVSALAR, AND F. POLLARA, *Labelings and encoders with the uniform bit error property with applications to serially concatenated trellis codes*, IEEE Trans. Inform. Theory, 48 (2002), pp. 123–136.
- [24] F. GARIN, G. COMO, AND F. FAGNANI, *Staircase and other structured linear-time encodable LDPC codes: Analysis and design*, in Proceedings of the International Symposium on Information Theory, 2007, pp. 1226–1230.
- [25] A. GRAELL I AMAT, G. MONTORSI, AND F. VATTA, *Analysis and design of rate compatible serial concatenated convolutional codes*, in Proceedings of the International Symposium on Information Theory, 2005, pp. 607–611.

- [26] T. W. HUNGERFORD, *Algebra*, Holt, Rinehart, and Winston, New York, 1974.
- [27] H. JIN, A. KHANDEKAR, AND R. J. MCELIECE, *Irregular repeat-accumulate codes*, in Proceedings of the International Symposium on Turbo Codes, 2000, pp. 1–8.
- [28] H. JIN AND R. J. MCELIECE, *Coding theorems for turbo code ensembles*, IEEE Trans. Inform. Theory, 48 (2002), pp. 1451–1461.
- [29] R. JOHANNESSON, Z.-X. WAN, AND E. WITTENMARK, *Some structural properties of convolutional codes over rings*, IEEE Trans. Inform. Theory, 44 (1998), pp. 839–845.
- [30] R. E. KALMAN, P. L. FALB, AND M. A. ARBIB, *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969.
- [31] S. LE GOFF, A. GLAVIEUX, AND C. BERROU, *Turbo-codes and high spectral efficiency modulation*, in Proceedings of the IEEE International Conference on Communications, 1994, pp. 645–649.
- [32] H.-A. LOELIGER, *Signal sets matched to groups*, IEEE Trans. Inform. Theory, 37 (1991), pp. 1675–1682.
- [33] D. J. C. MACKAY, *Good error-correcting codes based on very sparse matrices*, IEEE Trans. Inform. Theory, 45 (1999), pp. 399–431.
- [34] H. OGIWARA, A. MIZUTOME, AND K. KOIKE, *Performance evaluation of parallel concatenated trellis-coded modulation*, IEICE Trans. Fundamentals, E84-A (2001), pp. 2410–2417.
- [35] T. RICHARDSON, *The geometry of turbo-decoding dynamics*, IEEE Trans. Inform. Theory, 46 (2000), pp. 9–23.
- [36] P. ROBERTSON AND T. WÖRZ, *A novel bandwidth efficient coding scheme employing turbo-codes*, in Proceedings of the IEEE International Conference on Communications, 1996, pp. 962–967.
- [37] I. SASON AND S. SHAMAI (SHITZ), *Improved upper bounds on the ML decoding error probability of parallel and serially concatenated turbo codes via their ensemble distance spectrum*, IEEE Trans. Inform. Theory, 46 (2000), pp. 24–47.
- [38] I. SASON, E. TELATAR, AND R. URBANKE, *The asymptotic input-output weight distributions and thresholds of convolutional and turbo-like encoders*, IEEE Trans. Inform. Theory, 48 (2002), pp. 3052–3061.
- [39] S. SHAMAI (SHITZ) AND I. SASON, *Variations on the Gallager bounds, connections, and applications*, IEEE Trans. Inform. Theory, 48 (2002), pp. 3029–3051.
- [40] E. TELATAR AND R. URBANKE, *On the ensemble performance of turbo codes*, in Proceedings of the IEEE International Symposium on Information Theory, 1997, p. 105.
- [41] D. V. TRUHACHEV, M. LENTMAIER, AND K. S. ZIGANGIROV, *Some results concerning design and decoding of turbo-codes*, Probl. Inf. Transm., 37 (2001), pp. 190–205.

TROPICAL LINEAR SPACES*

DAVID E. SPEYER†

Abstract. We define the tropical analogues of the notions of linear spaces and Plücker coordinates and study their combinatorics. We introduce tropical analogues of intersection and dualization and define a tropical linear space built by repeated dualization and transverse intersection to be constructible. Our main result is that all constructible tropical linear spaces have the same f -vector and are “series-parallel”. We conjecture that this f -vector is maximal for all tropical linear spaces, with equality precisely for the series-parallel tropical linear spaces. We present many partial results towards this conjecture.

Key words. tropical linear space, tropical geometry, matroid, valuated matroid, f -vector conjecture

AMS subject classification. 52B40

DOI. 10.1137/080716219

1. Introduction and summary. The last five years have shown an explosion of interest in the field of tropical geometry. In tropical geometry, multiplication is replaced by addition and addition by minimum. See [19] or [16] for an introduction. A polynomial is said to vanish tropically if, among the values of its terms, the minimum is not unique. For example, $xy + x + y$ vanishes tropically if $x = y \leq x + y$ or $x = x + y \leq y$ or $y = x + y \leq x$. One motivation for this definition is that this is the union of the three sets $x + y = \min(x, y)$, $x = \min(x + y, y)$, and $y = \min(x, x + y)$ which can be obtained by replacing multiplication by addition and addition by minimum in the various equations $-xy = x + y$, $-x = xy + y$, and $-y = xy + x$.

We will study the tropical analogues of the notions of linear spaces and Plücker coordinates. One of the pleasures of this theory is that it can be set up in a purely combinatorial manner without reference to general tropical theory. However, doing so requires defining slightly more inclusive notions of tropical linear space than if the theory were exactly set up to match the classical geometry. This is very similar to the situation of matroid theory, where the definition of a matroid is simple and almost every operation that can be performed with a vector arrangement has a corresponding operation for a general matroid. On the other hand, characterizing those matroids that come from vector arrangements—the realizable matroids—is very hard. Our approach in this paper will be to focus on pure combinatorics. In section 4.5 of my dissertation (see [18]) I give many examples and counter-examples regarding the subtle behaviors which can occur when one tries to study the tropical analogue of realizability.

Tropical linear spaces will be polyhedral complexes. Thus, we can ask about all of the invariants associated to polyhedral complexes. Ardila and Klivans [2] have studied the link of a vertex in a tropical linear space and have shown it to be homeomorphic to the chain complex of the lattice of flats of a certain matroid and thus homotopic to a wedge of spheres. The polytopes occurring in tropical linear spaces are Minkowski summands of permutahedra (Proposition 2.6). The topology of tropical linear spaces

*Received by the editors June 15, 2005; accepted for publication (in revised form) May 9, 2008; published electronically October 17, 2008.

<http://www.siam.org/journals/sidma/22-4/71621.html>

†Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139 (speyer@math.mit.edu).

is quite simple: they are contractible (Theorem 2.9). So we understand the local combinatorics and the local and global topology of tropical linear spaces.

The global combinatorics of tropical linear spaces, on the other hand, is quite intriguing. There is a great deal of theoretical and some experimental evidence for the following.

The f -vector conjecture. *The number of i -dimensional faces of a tropical d -plane in n -space which become bounded after being mapped to $\mathbb{R}^n/(1, \dots, 1)$ is at most*

$$\frac{(n-i-1)!}{(d-i)!(n-d-i)!(i-1)!}.$$

This conjecture is proven, in [17], in the case where the tropical linear space arises from a $\mathbb{C}((t))$ -valued point of the Grassmannian $G(d, n)$. That argument should probably be considered to be the best evidence for the f -vector conjecture at this point.

Remark 1.1. It would follow from the f -vector conjecture that the total number of i -dimensional faces of a tropical d -plane in n -space, without any boundedness condition, is at most $\binom{n-i-1}{d-i} \binom{2n-d-1}{i-1}$. See Proposition 2.11.

In section 2, we will define the basic concepts mentioned in this introduction and prove their essential properties. We will also introduce an operation called dualization, which is analogous to the classical operation of taking the orthogonal complement of a linear space.

In section 3, we will prove the f -vector conjecture in two special cases: where $i = 1$ and where $d = \lfloor n/2 \rfloor$. This section is largely independent of the rest of the paper and shows what we can accomplish with fairly direct arguments. We do warn the reader, however, that the notation $\Gamma(M, M')$ introduced in section 3 will recur throughout this paper. After this section, we move to the main tasks of this paper: defining constructible linear spaces and proving their main properties.

In section 4, we will introduce an operation of transverse intersection that produces new tropical linear spaces from old. We define a tropical linear space to be *constructible* if it can be built from tropical hyperplanes (equivalently, from points) by successive dualization and transverse intersection. Much of the rest of the paper will be devoted toward proving the following.

THEOREM 1.2. *Every constructible space achieves the f -vector of the f -vector conjecture.*

As we will explain in section 2, every tropical linear space is built out of a collection of matroids. A particularly nice class of matroids is the series-parallel matroids; the linear spaces that are built out of series-parallel matroids are called series-parallel linear spaces. In section 5 we will review standard results on series-parallel matroids and prove some new ones. In section 6, we will prove Theorem 1.3.

THEOREM 1.3. *Every constructible space is series-parallel.*

One unfortunate consequence of Theorem 1.3 is that it is quite hard to find a general method for writing down tropical linear spaces that are not series-parallel! More precisely, it is not so hard to produce degenerate limits of series-parallel linear spaces which are not themselves series-parallel. Writing down a large number of tropical linear spaces which are not limits of series-parallel tropical linear spaces, however, is a challenge—which is one of the reasons why experimentally testing the f -vector conjecture is tricky.

We pause in section 7 to prove a polyhedral lemma. In section 8 we undertake a very detailed study of how the combinatorics of the intersection of a pair of linear spaces changes when we perturb those linear spaces. We are then ready, in section 9, to prove Theorem 1.2. Our proof involves results from sections 5, 7, and 8, as well as several binomial coefficient identities.

Most of our vocabulary regarding matroids and polytopes is quite standard. We address a few issues here. The word *circuit* refers to a minimal dependent set in a matroid. This terminology is motivated by applications to graph theory which we will not discuss. There will be several graphs appearing in this paper, which we will not consider as matroids. When we want to refer to a cyclic induced subgraph of a given graph, we will use the term *cycle*. A *loop* of a matroid is an element which is not in any basis; a *coloop* is an element which is in every basis. The matroid consisting of an individual loop, or of an individual coloop, is connected. The *uniform matroid of rank d on a set S* , which we will denote by $\text{Uniform}(d, S)$, is the matroid on S whose bases are all d -element subsets of S . The next two definitions are not standard: a *filtration* of a matroid is an ascending chain of subsets $F_0 \subsetneq F_1 \subsetneq F_2 \subsetneq \cdots \subsetneq F_r$ of the ground set of M , with $F_0 = \emptyset$ and F_r the entire ground set of M . The *associated graded matroid of M with respect to $F_\bullet$* , denoted $\text{gr}(M, F_\bullet)$ is the matroid $\bigoplus (M|_{F_{i+1}})/F_i$. Note that each summand of $\text{gr}(M, F_\bullet)$ is a minor of M .

If P is a polytope, then $\mathbb{R}P$ is the vector space spanned by vectors of the form $p_1 - p_2$, with p_1 and p_2 in P . The dimension of P is the dimension of $\mathbb{R}P$. For Q a face of P , the *tangent cone* to P at Q denoted $T(P, Q)$ is the positive linear span of all vectors of the form $p - q$, with $p \in P$ and $q \in Q$. The *reduced tangent cone* to P at Q denoted $T_0(P, Q)$ is $T(P, Q)/\mathbb{R}Q$. The relative interior of P denoted $\text{Int}(P)$ is the set of points p in P which are not contained in any proper face of P .

2. Basic definitions and results. In this section, we introduce the definitions of a *tropical Plücker vector* and a *tropical linear space* by taking classical relations between Plücker coordinates and the linear spaces they describe and “tropicalizing” those relations. These are closely related to the valuated matroids and tight spans studied by Dress (see [9], [7], [10]). Tropical Plücker vectors give rise to subdivisions of a polytope known as the hypersimplex into smaller polytopes. Each of these smaller polytopes is associated to a combinatorial object known as a matroid, and for this reason we will frequently use tools from matroid theory. Just as ordinary linear spaces are determined by their Plücker coordinates, tropical linear spaces are determined by tropical Plücker coordinates. Specifically, if $\mathcal{D}$ is the subdivision of the hypersimplex associated to a tropical Plücker vector p , then the corresponding tropical linear space is a subcomplex of a dual subdivision $\mathcal{D}^\vee$ of $\mathbb{R}^n$.

For $f = \sum_{a \in A} f_a X^a$ a nonzero polynomial, where the coefficients f_a are nonzero and A is a finite subset of $\mathbb{Z}_{\geq 0}^n$, we denote by $\text{Trop}(f)$ the set of all points $(x_1, \dots, x_n)$ in $\mathbb{R}^n$ such that, if we form the collection of numbers $\sum_{i=1}^n a_i x_i$ for a ranging over A , then the minimum of this collection is not unique. Throughout the paper, we will use $\mathbf{X}$, $\mathbf{P}$, and $\mathbf{Q}$ to refer to variables in classical algebro-geometric coordinates and x , p , and q for the corresponding tropical variables. We will use classical algebraic geometry solely as an intuitive analogy to motivate our combinatorial constructions, but, of course, this analogy can be made far more precise.

Consider a collection of $\binom{n}{d}$ real numbers $p_{i_1 \dots i_d}$ indexed by the d element subsets of $[n] := \{1, \dots, n\}$. For any set S , we will write $\binom{S}{k}$ for the set of k -element subsets of S . We will say that p is a *tropical Plücker vector* if, for every $S \in \binom{[n]}{d-2}$ and every i, j, k , and l distinct members of $[n] \setminus S$, we have

$$(1) \quad (p_I) \in \text{Trop}(\mathbf{P}_{Sij}\mathbf{P}_{Sk l} - \mathbf{P}_{Sik}\mathbf{P}_{Sjl} + \mathbf{P}_{Sil}\mathbf{P}_{Sjk}).$$

(Here and elsewhere in this paper, we write Si as shorthand for $S \cup \{i\}$ and $\binom{S}{k}$ for the set of k -elements subsets of S .)

Remark 2.1. This definition is equivalent to saying that p_I is a valuated matroid with values in the semiring $(\mathbb{R}, +, \min)$. See [9] as well as the more general [7] and [10].

Let e_n be the standard basis of $\mathbb{R}^n$. Let $\Delta(d, n)$ denote the (d, n) -hypersimplex, defined as the convex hull of the points $e_{i_1} + \cdots + e_{i_d} \in \mathbb{R}^n$, where $\{i_1, \dots, i_d\}$ runs over $\binom{[n]}{d}$. (Here $e_1, \dots, e_n$ is the standard basis of $\mathbb{R}^n$.) We abbreviate $e_{i_1} + \cdots + e_{i_d}$ by $e_{i_1 \dots i_d}$. Consider a real-valued function $\{i_1, \dots, i_d\} \mapsto p_{i_1 \dots i_d}$ on the vertices of $\Delta(d, n)$. We define a polyhedral subdivision $\mathcal{D}_p$ of $\Delta(d, n)$ as follows: consider the points $(e_{i_1} + \cdots + e_{i_d}, p_{i_1 \dots i_d}) \in \Delta(d, n) \times \mathbb{R}$ and take their convex hull. Take the lower facets (those whose outward normal vector have last component negative) and project them back down to $\Delta(d, n)$; this gives us the subdivision $\mathcal{D}_p$. We will often omit the subscript p when it is clear from context. A subdivision which is obtained in this manner is called *regular*; see, for example, [24, Definition 5.3]. Another way of defining $\mathcal{D}$ is to say that a subset S of $\binom{[n]}{d}$ corresponds to a face of $\mathcal{D}$ if and only if there is an affine linear function $\lambda : \Delta(d, n) \rightarrow \mathbb{R}$, with $p_I - \lambda(e_I)$ minimized precisely at those I which lie in S .

Let P be a subpolytope of $\Delta(d, n)$. We say that P is *matroidal* if the vertices of P , considered as elements of $\binom{[n]}{d}$, are the bases of a matroid M . In this case, we write $P = \text{Poly}_M$. All faces of a matroidal polytope are also matroidal. Our references for matroid terminology and theory are White's anthologies [22] and [23].

PROPOSITION 2.2. *Let $p \mapsto p_I$ be a real-valued function on $\binom{[n]}{d}$. The following are equivalent:*

1. *The vector (p_I) is a tropical Plücker vector.*
2. *The one skeleta of $\mathcal{D}_p$ and $\Delta(d, n)$ are the same.*
3. *Every face of $\mathcal{D}_p$ is matroidal.*

Proof. $1 \implies 2$: Every edge of $\mathcal{D}$ joins two vertices of $\Delta(d, n)$. If e is an edge of $\mathcal{D}$ connecting e_I and e_J , we define the *length* of e , $\ell(e)$, to be $|I \setminus J| = |J \setminus I|$. Our claim is that every edge of e has length 1. We prove by induction on $\ell \geq 2$ that $\mathcal{D}$ has no edge of length ℓ .

We first consider the base case $\ell = 2$. If e is an edge with $\ell(e) = 2$, then $e = (e_{Sij}, e_{Skil})$ for some $S \in \binom{[n]}{d-2}$. The six vertices $e_{Sij}, e_{Sik}, e_{Sil}, e_{Sjk}, e_{Sjl}, e_{Skil}$ form the vertices of an octahedron O , with e_{Sij} and e_{Skil} opposite vertices. One can check that equation (1) implies that O is either a face of $\mathcal{D}$ or subdivided in $\mathcal{D}$ into two square pyramids (in one of three possible ways). In any of these cases, e is not an edge of $\mathcal{D}$.

Now consider the case where $\ell > 2$. Suppose (for contradiction) that e is an edge of $\mathcal{D}$. Let $e = (e_{ST}, e_{ST'})$, with $T \cap T' = \emptyset$, $|T| = |T'| = \ell$. Let F be the face of $\Delta(d, n)$ consisting of all vertices e_I , with $S \subset I \subset S \cup T \cup T'$. Then e must belong to some two-dimensional face of $\mathcal{D}$ contained in F ; call this two-dimensional face G .

Let γ be the path from e_{ST} to $e_{ST'}$ that goes around G the other way from e . No two vertices of F are more than distance ℓ apart, so the edges of γ have lengths less than or equal to ℓ . If γ contained an edge $(e_{SU}, e_{SU'})$ of length ℓ , then its midpoint $(e_{SU} + e_{SU'})/2$ would also be the midpoint of e , contradicting the convexity of G . Thus, every edge of γ has length less than ℓ and by induction must have length 1. So all of the edges of γ are in the direction $e_i - e_j$ for some i and $j \in [n]$. These vectors must span a two-dimensional space. This means either that there are $\{i_1, i_2, j_1, j_2\} \subset [n]$ such that all edges of γ are parallel to some $e_{i_r} - e_{j_s}$ or there are $\{i_1, i_2, i_3\} \subset [n]$ such that every edge of γ is parallel to some $e_{i_r} - e_{i_s}$. In either case, e has length at most 2, but that returns us to our base case.

2 $\implies$ 3: Let P be any polytope in $\mathcal{D}$. By assumption, all of the edges of P are edges of $\Delta(d, n)$. It is a theorem of Gelfand et al. ([13], Theorem 4.1) that this implies that P is matroidal.

3 $\implies$ 1: If 1 is false, $\mathcal{D}$ has a one-dimensional face of the form $\text{Hull}(e_{Sij}, e_{Skl})$, with i, j, k , and l distinct. This is not matroidal. $\square$

Now, suppose that $(p_I) \in \mathbb{R}^{\binom{[n]}{d}}$ is a tropical Plücker vector. Define $L(p) \subset \mathbb{R}^n$ by

$$(2) \quad L(p) := \bigcap_{1 \leq j_1 < \dots < j_{d+1} \leq n} \text{Trop} \left(\sum_{r=1}^{d+1} (-1)^r \mathbf{p}_{j_1 \dots \hat{j}_r \dots j_{d+1}} \mathbf{x}_{j_r} \right).$$

(The factor of $(-1)^r$ does not effect the right-hand side; we include it to emphasize the analogy with classical linear algebra.) We call any subset L of $\mathbb{R}^n$ which arises in this manner a *d-dimensional tropical linear space in n-space*. Rescaling all of the Plücker coordinates of a classical linear space by the same scalar does not alter the linear space. Similarly, it is clear from the definition that adding a constant to all of the tropical Plücker coordinates of a tropical linear space. See Proposition 2.8 for a converse result: A tropical linear space determines its Plücker coordinates up to this translation.

We can now define the term *series-parallel tropical Plücker vector*, used repeatedly in the introduction. A tropical Plücker vector p is called series-parallel if every facet of $\mathcal{D}_p$ corresponds to a series-parallel matroid; a tropical linear space $L(p)$ is called series-parallel if p is series-parallel. The definition and basic properties of series-parallel matroids are reviewed in section 5 and will be used continually in sections 6 and 8.

Let us compute two simple examples. First, let $d = 1$ and $n = 3$, and consider the tropical Plücker vector $(p_1, p_2, p_3) = (2, 5, 7)$. This is a tropical Plücker vector as equation (1) imposes no condition when $d = 1$ or $n - d = 1$. Then $L(p)$ is the set of points $(x_1, x_2, x_3) \in \mathbb{R}^3$ such that in each of the pairs $(x_1 + 5, x_2 + 2)$, $(x_1 + 7, x_3 + 2)$, and $(x_2 + 7, x_3 + 5)$ the minimum occurs more than once. The reader may check that this is precisely the set of triples $(x_1, x_2, x_3) = (2 + c, 5 + c, 7 + c)$, where c can be any real number. Alternatively, let $d = 2$, $n = 3$, and let $(p_{12}, p_{13}, p_{23}) = (3, 9, 13)$. Then $L(p)$ is the set of points $(x_1, x_2, x_3) \in \mathbb{R}^3$ such, among the triple $(x_1 + 13, x_2 + 9, x_3 + 3)$, the minimum occurs more than once. The reader may check that this is given by the union of three half-planes: $x_1 + 13 = x_2 + 9 \leq x_3 + 3$, $x_1 + 13 = x_3 + 3 \leq x_2 + 9$, and $x_2 + 9 = x_3 + 3 \leq x_1 + 13$.

The set $L(p)$ is essentially the same set as the tight span of p defined by Dress in [8]. There are two differences: The first is that Dress's sign conventions are opposite to ours. The second is that $L(p)$ is invariant under translation by $(1, \dots, 1)$, while Dress chooses a particular representative within each orbit for this translation. The choice makes many of the constructions in this paper more difficult to see.

While the defining equation (2) makes the connection to ordinary linear spaces most clear, for practically every purpose it is better to work with the alternate characterization which we now give. For any $w \in \mathbb{R}^n$, consider the subset of the vertices of $\Delta(d, n)$ at which $p_{i_1 \dots i_d} - \sum_{j=1}^d w_{i_j}$ is minimal. This is a face of $\mathcal{D}$ and thus is Poly_{M_w} for some matroid M_w .

PROPOSITION 2.3. *The point w is in $L(p)$ if and only if M_w is loop-free.*

Recall that a matroid is loop-free if every element of the matroid appears in at least one basis. There is a geometric way of recognizing when M is loop-free from the polytope Poly_M : M is loop-free if and only if Poly_M is not contained in any of the n facets of $\Delta(d, n)$ of the form $x_i = 0$. In particular, note that if Poly_M meets the interior of $\Delta(d, n)$, M is necessarily loop-free.

Proof. By replacing $p_{i_1 \dots i_d}$ by $p_{i_1 \dots i_d} - \sum_{j=1}^d w_{i_j} - (\text{constant})$, we may assume without loss of generality that $w = 0$ and $\min p_I = 0$. Then $M_w = M_0$ is simply the matroid whose bases correspond to those I for which $p_I = 0$. First, we assume that M_0 has a loop j and prove that $0 \notin L(p)$. Let $(i_1, \dots, i_d)$ be a basis of M_0 , clearly $j \notin (i_1, \dots, i_d)$. Then $p_{i_1 \dots i_d} = 0$, but $p_{j i_1 \dots \hat{i}_r \dots i_d} > 0$ for $1 \leq r \leq d$. Taking $(j_1, \dots, j_{d+1}) = (j, i_1, \dots, i_d)$ we see that $0 \notin \text{Trop}(\sum_{r=1}^{d+1} (-1)^r \mathbf{P}_{j_1 \dots \hat{j}_r \dots j_{d+1}} \mathbf{X}_{j_r})$.

The converse is more interesting. Fix $J = \{j_1, \dots, j_{d+1}\}$; our aim will be to prove $0 \in \text{Trop}(\sum_{r=1}^{d+1} (-1)^r \mathbf{P}_{j_1 \dots \hat{j}_r \dots j_{d+1}} \mathbf{X}_{j_r})$. Let $e_J = \sum_{r=1}^{d+1} e_{j_r}$ and, for any $s \in \mathbb{R}$, set $M_s = M_{-se_J}$. For s large enough, the bases of M_s will all be subsets of J ; more specifically, among the d -element subsets of J , they will be the subsets I at which p_I is minimized. In order to show that 0 is in $\text{Trop}(\sum_{r=1}^{d+1} (-1)^r \mathbf{P}_{j_1 \dots \hat{j}_r \dots j_{d+1}} \mathbf{X}_{j_r})$, it is equivalent to show that M_s has at least two bases for s sufficiently large. In other words, for s sufficiently large, we wish to show that j is not a loop of M_s , for j any element of J . By hypothesis, j is not a loop of M_0 , so it is enough to show that, if $j \in J$ is not a loop of M_0 , then it is not a loop of M_s for $s > 0$. We show the contrapositive—if j is a loop of M_s , then it is a loop of $M_{s'}$ for all $0 \leq s' < s$.

As s varies, M_s changes at a finite number of values of s ; call them $s_1 < s_2 < \dots < s_k$. Suppose that $s_i < s < s_{i+1}$, then Poly_{M_s} is a face of both $\text{Poly}_{M_{s_i}}$ and of $\text{Poly}_{M_{s_{i+1}}}$. The bases of M_s are precisely the bases of M_{s_i} which have the largest possible number of elements in common with J . Similarly, the bases of M_s are precisely the bases of $M_{s_{i+1}}$ which have the smallest possible number of elements in common with J . In other words,

$$(3) \quad M_s = M_{s_i}|_J \oplus M_{s_i}/J = M_{s_{i+1}}|_{[n] \setminus J} \oplus M_{s_{i+1}}/([n] \setminus J).$$

From equation (3), it follows that if $j \in J$ is a loop in $M_{s_{i+1}}$, then it is a loop in M_s . Similarly, if j is a loop in M_s , then it is a loop in M_{s_i} . Concatenating deductions of this sort, we see that, as promised, if j is a loop of M_s , then it is a loop of $M_{s'}$ for all $s' < s$ as desired. $\square$

It is clear from the defining equation (2) that $L(p)$ is invariant under translation by $(1, 1, \dots, 1)$. We will abuse notation by saying a face of $L(p)$ is bounded if its image in $\mathbb{R}^n/(1, \dots, 1)$ is bounded. However, when we refer to the dimensions of the faces of L we will always be speaking of L itself, without taking the quotient by $(1, \dots, 1)$.

Let $\mathcal{D}_p^\vee$ be the polyhedral subdivision of $\mathbb{R}^n$, where w and w' lie in the same face if $M_w = M_{w'}$. Then we see that L is a subcomplex of $\mathcal{D}_p^\vee$. We will often write $\mathcal{D}_p^\vee$ when p is clear from context. We warn the reader that, while the combinatorics of $\mathcal{D}_p^\vee$ is determined by the subdivision $\mathcal{D}_p$, the actual positions of the faces of $\mathcal{D}_p^\vee$ depend on p as well as on $\mathcal{D}_p$. In Figure 1, we show $\Delta(2, 4)$, which is an octahedron, subdivided into two square pyramids and draw the dual $L/(1, 1, 1, 1)$ in bold. We write $M^\vee$ for the face of $\mathcal{D}^\vee$ dual to Poly_M . If $v = M^\vee/(1, \dots, 1)$ is a vertex of $L/(1, \dots, 1)$, then we see that the link of $L/(1, \dots, 1)$ at v is the subcomplex of the normal fan of Poly_M . Specifically, the link of $L/(1, \dots, 1)$ at v consists of the cones dual to those faces of Poly_M which correspond to loop-free matroids. This fan was termed the Bergmann complex of M in [21, Chapter 9] and is studied in [2]. In Proposition 2.4, we summarize the main results of [2].

PROPOSITION 2.4. *Every face of Poly_M is of the form $\text{Poly}_{\text{gr}(M, F_\bullet)}$ for some filtration of M . The cones of the link $\text{link}_{M^\vee}(L)$ are dual to those faces of the form $\text{Poly}_{\text{gr}(M, F_\bullet)}$, where $F_\bullet$ is an ascending chain of flats. In this manner, the link $\text{link}_{M^\vee}(L)$ is homeomorphic to the chain complex of the lattice of flats of M and has a subdivision which is isomorphic (as a simplicial complex) to this chain complex.*

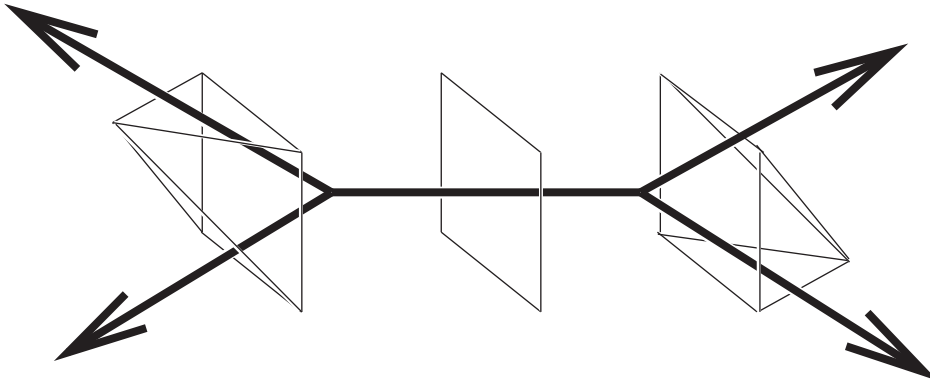


FIG. 1. A subdivision of $\Delta(2, 4)$ and the corresponding $L/(1, 1, 1, 1)$.

When $M^\vee/(1, \dots, 1)$ is positive-dimensional, Poly_M turns out to be a product of Poly_N 's.

PROPOSITION 2.5. *Let M be a matroid which decomposes as the direct sum of the connected matroids M_i . Then $\text{Poly}_M = \prod \text{Poly}_{M_i}$.*

Proof. By the definition of direct sum, the bases of M are all disjoint unions $\bigsqcup B_i$, where B_i is a basis of M_i . This is also precisely the definition of a product of polytopes. $\square$

We can now give a very restrictive condition on the faces of L . Recall that the braid arrangement $\mathcal{B}_r$ is the fan in $\mathbb{R}^r$ whose facets are the cones of the type $x_{\sigma(1)} \leq \dots \leq x_{\sigma(r)}$, where σ ranges over the symmetric group S_r .

PROPOSITION 2.6. *If $M^\vee$ is an r -dimensional face of L , then the normal fan of $M^\vee$ is a subfan of a coarsening of $\mathcal{B}_r$. When $M^\vee/(1, 1, \dots, 1)$ is bounded, $M^\vee/(1, 1, \dots, 1)$ is a Minkowski summand of some multiple of the r th permutahedron.*

For example, the two-dimensional faces of $L/(1, 1, \dots, 1)$ all have no more than 6 sides, and their sides point in at most 3 directions. The three-dimensional faces of L are the products with $\mathbb{R}$ of the two-dimensional faces of $L/(1, 1, \dots, 1)$.

Proof. The normal fan to $M^\vee$ is the same as the link of Poly_M in $\mathcal{D}$. Let Poly_N be a face of $\mathcal{D}$, with Poly_M a face of Poly_N . We must show that $T_0(\text{Poly}_N, \text{Poly}_M)$ is a union of cones of $\mathcal{B}_r$. But $T_0(\text{Poly}_N, \text{Poly}_M)$ is precisely the negative of the cone dual to Poly_M in the normal fan to Poly_N . From the description in [2], we see that the normal fan to Poly_N is a coarsening of $\mathcal{B}_n$, and that the cones dual to codimension r faces are the unions of cones of $\mathcal{B}_r$.

If $M^\vee$ is bounded, then its normal fan is complete and thus, by the last paragraph, a coarsening of $\mathcal{B}_r$. The fan $\mathcal{B}_r$ is the normal fan to the r th permutahedron. If P and Q are two polytopes such that the normal fan to Q refines that of P , then P is a Minkowski summand of tQ for t sufficiently large. $\square$

PROBLEM 2.7. *Let P be a polytope which is a Minkowski summand of some multiple of the r th permutahedron. Can P always occur as a face of a tropical linear space?*

For ordinary linear spaces, a linear space determines its Plücker coordinates up to scaling. For tropical linear spaces, a similar result holds.

PROPOSITION 2.8. *Let L be a tropical linear space with tropical Plücker coordinates p . Then L determines the p_I up to the addition of a common scalar. In other*

words, if p and p' are tropical Plücker vectors with $L(p) = L(p')$, then there is a constant c such that $p_I = p'_I + c$ for all $I \in \binom{[n]}{d}$.

Proof. Let S be a $d-1$ element subset of $[n]$, and let F denote the face of $\Delta(d, n)$ whose vertices are those e_I , where I contains S . Since F is an $(n-d)$ -simplex, it can not be subdivided in $\mathcal{D}$. Thus, F is a face of $\mathcal{D}$. The matroid Φ corresponding to F is loop-free, so it is dual to a face $\Phi^\vee$ of L . One can easily see that $\Phi^\vee$ is of the form $(\mathbb{R}_{\geq 0}^S + \mathbb{R}(1, \dots, 1)) + q$ for q a vector satisfying $q_i = p_{Si}$ for $i \notin S$. No other face of L is of the form $(\mathbb{R}_{\geq 0}^S \oplus \mathbb{R}(1, \dots, 1)) + q$, and q is unique up to translation by $(1, \dots, 1)$. Thus, L determines $p_{Si} - p_{Sj}$. As this is true for all S, i , and j , L determines the p_I up to adding the same constant to all of them. $\square$

THEOREM 2.9. *The polyhedral complex L is pure d -dimensional and contractible.*

Proof. To show that L is pure d -dimensional, it is enough to show that the link of any vertex v of $L/(1, \dots, 1)$ is a pure $(d-2)$ -dimensional complex. Let $v + (1, \dots, 1)$ be dual to Poly_M for some matroid M . By the results of [2] (Proposition 2.4), the link of v has a subdivision isomorphic to the order complex of the lattice of flats of M . This lattice is graded of length d , so its order complex is pure of dimension $d-2$.

We now show contractibility. It is easy to see that each tropical hyperplane is tropically convex in the sense of [6]. Thus, their intersection is tropically convex and hence contractible. (Theorem 2 of [6] contains all results on tropically convex spaces used in this argument.) $\square$

A slight extension of the argument in this theorem will show that a (nonempty) face $M^\vee$ is i -dimensional, where i is the number of connected components of M . We now discuss the bounded faces of L . Recall our abuse of language, that we call a face $M^\vee$ bounded when $M^\vee/(1, \dots, 1)$ is a bounded polytope.

PROPOSITION 2.10. *The following are equivalent:*

1. *The point w lies in a bounded face of L .*
2. *The polytope Poly_{M_w} is an interior face of $\mathcal{D}$.*
3. *The matroid M_w is loop-free and co-loop-free.*

Proof. $2 \iff 3$: To say that Poly_{M_w} is interior is to say that it is not contained in any facet of $\Delta(d, n)$. The facets of $\Delta(d, n)$ are of two types: $\{x \in \mathbb{R}^n : x_i = 1\}$ and $\{x \in \mathbb{R}^n : x_i = 0\}$ (where i is between 1 and n). Essentially by definition, Poly_{M_w} lies in the former type if and only if i is a coloop of M_w ; Poly_{M_w} lies in the latter type if and only if i is a loop of M_w .

$1 \iff 2$: If P is not an interior face of $\mathcal{D}$, the corresponding dual face is not bounded. Conversely, if P is internal, the corresponding dual face of $\mathcal{D}^\vee$ is bounded, so we just must check that this dual face is in $L(p)$. For this, we must check that M_w is loop-free. But we saw in the previous paragraph that this follows from Poly_{M_w} being internal. $\square$

At this point, we can prove our earlier remark.

PROPOSITION 2.11. *If the f -vector conjecture is true, then a tropical d -plane in n -space has no more than $\binom{n-i-1}{d-i} \binom{2n-d-1}{i-1}$ faces of dimension i .*

Proof. Let $\mathcal{D}$ be a matroidal decomposition of $\Delta(d, n)$; we must bound the number of faces Poly_M of $\mathcal{D}$ such that M has no loops and Poly_M has codimension $i-1$. For such a face Poly_M , let $C \subset [n]$ be the coloops of M , and let Δ be the face of $\Delta(d, n)$ whose vertices are the vertices e_I , where $I \in \binom{[n]}{d}$ contains C . The claim that M has no loops is equivalent to saying that Poly_M is an interior face of the decomposition of Δ . So we may count the total number of such faces by summing over Δ .

Let us look at those faces of $\Delta(d, n)$ which can play the role of Δ with $|C| = j$; there are $\binom{n}{j}$ such faces and each is isomorphic to $\Delta(d-j, n-j)$. A subpolytope

of such a face Δ has codimension $i - 1$ in $\Delta(d, n)$ if and only if it has codimension $i - j - 1$ in Δ . So, assuming the f -vector conjecture and writing the bound in the f -vector conjecture as $\binom{n-i-1}{d-i} \binom{n-d-1}{i-1}$, we have the following bound for the number of codimension $i - 1$ loop-free faces of $\mathcal{D}$:

$$\sum_j \binom{n}{j} \binom{(n-j)-(i-j)-1}{(d-j)-(i-j)} \binom{(n-j)-(d-j)-1}{(i-j)-1} = \binom{n-i-1}{d-i} \sum_j \binom{n}{j} \binom{n-d-1}{i-j-1} = \binom{n-i-1}{d-i} \binom{2n-d-1}{i-1}.$$

The last sum is evaluated by Vandermonde's convolution: $\sum_j \binom{p}{j} \binom{q}{r-j} = \binom{p+q}{r}$; see [14, (5.22)].

(Note that here and throughout this paper, we use the convention that $\binom{n}{k} = 0$ whenever $k < 0$ or $0 \leq n < k$. This permits us to often use, as we have above, the convention that $\sum_j f(j)$ denotes $\sum_{j=-\infty}^{\infty} f(j)$. See Chapter 5 of [14] for numerous examples of how these notations can simplify manipulations with binomial coefficients.) $\square$

Suppose that $p : \binom{[n]}{d} \rightarrow \mathbb{R}$ is a tropical Plücker vector. Define $p^\perp : \binom{[n]}{n-d} \rightarrow \mathbb{R}$ by $p_I^\perp = p_{[n] \setminus I}$. Then the subdivision $\mathcal{D}^\perp$ of $\Delta(n-d, n)$ induced by $p^\perp$ is simply the image of $\mathcal{D}$ under $(x_1, \dots, x_n) \mapsto (1 - x_1, \dots, 1 - x_n)$. This map replaces each matroidal polytope with the polytope of the dual matroid; in particular, $p^\perp$ also obeys the tropical Plücker relations.

We denote $L(p^\perp)$ by $L^\perp$. One can easily check that $L \mapsto L^\perp$ is an inclusion reversing bijection from the set of tropical linear spaces in $\mathbb{R}^n$ to itself. This fact is implicit in Example 6.2 and Proposition 6.3 of [20].

PROPOSITION 2.12. *The bounded part of $L^\perp$ is (-1) times the bounded part of L . If L is a tropical linear space of dimension d in n -space, then the bounded part of L is at most $\min(d, n-d)$ -dimensional.*

Proof. We saw above that $\mathcal{D}^\perp$ is the image of $\mathcal{D}$ under the map $(x_1, \dots, x_n) \mapsto (1 - x_1, \dots, 1 - x_n)$ from $\Delta(d, n)$ to $\Delta(n-d, n)$. So the dual subdivision $(\mathcal{D})^\vee$ is the negative of $(\mathcal{D}^\perp)^\vee$. The bounded faces of $(\mathcal{D}^\perp)^\vee$ are the negatives of the bounded faces of $(\mathcal{D})^\vee$ because being loop- and coloop-free is self dual. We know that L is d -dimensional and $L^\perp$ is $(n-d)$ -dimensional; as the bounded part of L is isomorphic to a subcomplex of both it is at most $\min(d, n-d)$ -dimensional. $\square$

3. Special cases of the f -vector conjecture. In this section we will prove the following results, which are special cases of the f -vector conjecture.

THEOREM 3.1. *Let L be a d -dimensional tropical linear space in n -space. Then L has at most $\binom{n-2}{d-1}$ vertices, with equality if and only if L is series-parallel.*

THEOREM 3.2. *Let L be a d -dimensional tropical linear space in n -space, with $n = 2d$ or $n = 2d + 1$. Then L has at most 1 bounded facet if $n = 2d$ and at most d if $n = 2d + 1$.*

We first recall the definition of the Tutte polynomial of a matroid. For M a rank d matroid on the ground set $[n]$ and Y a subset of $[n]$, let $\rho_M(Y)$ denote the rank of the flat spanned by Y . The polynomial

$$r_M(x, y) = \sum_{Y \subseteq [n]} x^{|Y| - \rho_M(Y)} y^{d - \rho_M(Y)}$$

is known as the rank-generating function of M . The polynomial

$$t_M(z, w) = r_M(z - 1, w - 1)$$

is known as the Tutte polynomial of M . Almost all matroid invariants can be computed in terms of the Tutte polynomial; see Chapter 6 of [23] for a survey of its importance. Write $t_M(z, w) = \sum t_{ij} z^i w^j$. Although not obvious from this definition, all of the t_{ij} are nonnegative. For $n \geq 2$, we have $t_{10} = t_{01}$; this number is known as the beta (or Crapo) invariant of M and denoted $\beta(M)$.

We will need the following result.

PROPOSITION 3.3. *Let M be a matroid on at least 2 elements. Then $\beta(M) = 0$ if and only if M is disconnected and $\beta(M) = 1$ if and only if M is series-parallel.*

Proof. See [4, Theorem II] for the first statement and [3, Theorem 7.6] for the second. $\square$

The key to proving Theorem 3.1 will be proving the following formula.

LEMMA 3.4. *Let M be a matroid, and let $\mathcal{D}$ be a matroidal subdivision of Poly_M . Let $\mathring{\mathcal{D}}$ denote the set of internal faces of $\mathcal{D}$. Then*

$$t_M(z, w) = \sum_{\text{Poly}_\gamma \in \mathring{\mathcal{D}}} (-1)^{\dim(\text{Poly}_M) - \dim(\text{Poly}_\gamma)} t_\gamma(z, w).$$

Before proving Lemma 3.4, let us see why it implies Theorem 3.1. Considering the case where $M = \text{Uniform}(d, n)$ and comparing the coefficients of z on each side, we see that

$$\beta(\text{Uniform}(d, n)) = \binom{n-2}{d-1} = \sum_{\text{Poly}_\gamma \in \mathring{\mathcal{D}}} \beta(\gamma) = \sum_{\text{Poly}_\gamma \text{ a facet of } \mathcal{D}} \beta(\gamma).$$

To go from the first sum to the second, note that all nonfacets of $\mathring{\mathcal{D}}$ correspond to disconnected matroids and hence have beta invariant 0, and that every facet of $\mathcal{D}$ is in $\mathring{\mathcal{D}}$.

Every term in the right-hand sum is a positive integer. Thus, $\mathcal{D}$ has at most $\binom{n-2}{d-1}$ facets, with equality if and only if all of the facets have beta invariant 1, i.e., if and only if $\mathcal{D}$ is a decomposition into series-parallel matroids. $\square$

We now prove Lemma 3.4. Ardila, Fink, and Rincón have, since the initial circulation of this paper, found an alternate proof; see [1].

Proof. Since $t_M(z, w) = r_M(z - 1, w - 1)$, it is enough to prove

$$r_M(x, y) = \sum_{\text{Poly}_\gamma \in \mathring{\mathcal{D}}} (-1)^{\dim(\text{Poly}_M) - \dim(\text{Poly}_\gamma)} r_\gamma(x, y).$$

Plugging in the definition of r_M and interchanging summation, it is enough to show that, for every $Y \subseteq [n]$,

$$x^{|Y| - \rho_M(Y)} y^{d - \rho_M(Y)} = \sum_{\text{Poly}_\gamma \in \mathring{\mathcal{D}}} (-1)^{\dim(\text{Poly}_M) - \dim(\text{Poly}_\gamma)} x^{|Y| - \rho_\gamma(Y)} y^{d - \rho_\gamma(Y)}.$$

Comparing coefficients of $x^{|Y| - r} y^{d - r}$, we are thus being asked to show that

$$\sum_{\substack{\text{Poly}_\gamma \in \mathring{\mathcal{D}} \\ \rho_\gamma(Y) = r}} (-1)^{\dim(\text{Poly}_M) - \dim(\text{Poly}_\gamma)} = \begin{cases} 1 & \text{if } r = \rho_M(Y) \\ 0 & \text{if } r \neq \rho_M(Y) \end{cases}.$$

The sum is empty if $r > \rho_M(Y)$, so it is enough to prove the claim when $\rho_M(Y) \leq r$. Equivalently, we will show

$$\sum_{\substack{\text{Poly}_\gamma \in \hat{\mathcal{D}} \\ \rho_\gamma(Y) \geq r}} (-1)^{\dim(\text{Poly}_M) - \dim(\text{Poly}_\gamma)} = 1$$

for all $r \leq \rho_M(Y)$.

Let ℓ_Y be the linear function $\Delta(d, n) \rightarrow \mathbb{R}$ mapping $(x_i) \mapsto \sum_{i \in Y} x_i$. Then $\rho_\gamma(Y) = \max_{x \in \text{Poly}_\gamma} \ell_Y(x)$. Thus, we see that $\rho_\gamma(Y) \geq r$ if and only if γ has a nonempty intersection with the half-space $\ell_Y > r - 1/2$. The promised equality now follows by the following lemma applied to the polytope $\text{Poly}_M \cap \{x \in \mathbb{R}^n : \ell_Y(x) > r - 1/2\}$. $\square$

LEMMA 3.5. *Let P be any bounded polytope, and let Γ be the set of internal faces of a decomposition of P . Then $\sum_{\gamma \in \Gamma} (-1)^{\dim(P) - \dim(\gamma)} = 1$.*

Proof. This sum is $(-1)^{\dim P}(\chi(P) - \chi(\partial P))$, where χ is the Euler characteristic. As P is contractible and ∂P is a sphere of dimension $\dim(P) - 1$, the result follows. $\square$

We now begin to prove Theorem 3.2.

Suppose that L is a d -dimensional tropical linear space in n -space. Let $M^\vee$ be a bounded facet of L , dual to a polytope Poly_M . Then the matroid M has rank d and has d connected components none of which are loops. So $M = \bigoplus_{i=1}^d M_i$, where each M_i has rank 1. Let S_i be the ground set of M_i ; because $M^\vee$ is bounded, we know that $|S_i|$ is at least 2. The vertices of Poly_M are those subsets of $[n]$ that meet each S_i precisely once.

Now, consider two such facets $M^\vee$ and $(M')^\vee$ occurring in the same tropical linear space, and define S_i and S'_i as above. The polytope $\text{Poly}_M \cap \text{Poly}_{M'}$ must be either matroidal or empty. We now introduce a graph which will be useful in reformulating the condition that $\text{Poly}_M \cap \text{Poly}_{M'}$ be either matroidal or empty, and also at many other points throughout this paper.

Let M and M' be two matroids on $[n]$, with connected decomposition $M = \bigoplus_{c \in C} M_c$, $M' = \bigoplus_{c \in C'} M'_c$. Let S_c and S'_c be the ground sets of M_c and M'_c . Let $\Gamma(M, M')$ be the bipartite graph (possibly with multiple edges) whose vertex set is $C \sqcup C'$ and which has an edge between c and c' for each element of $S_c \cap S'_{c'}$. A *perfect matching* of a graph Γ is a collection of edges such that, for every vertex v of Γ , exactly one edge in the collection contains v .

LEMMA 3.6. *Let M and M' be two rank d matroids on $[n]$, and let $\Gamma = \Gamma(M, M')$. The graph Γ may have multiple edges; let $\bar{\Gamma}$ be the simple graph obtained by replacing each multiple edge of Γ by a single edge. Then $\text{Poly}_M \cap \text{Poly}_{M'}$ is empty if and only if $\bar{\Gamma}$ has no perfect matchings and $\text{Poly}_M \cap \text{Poly}_{M'}$ is matroidal if and only if $\bar{\Gamma}$ has precisely one perfect matching.*

Proof. The vertices of $\text{Poly}_M \cap \text{Poly}_{M'}$ are in bijection with the perfect matchings of $\Gamma(M, M')$. There is an obvious surjection from perfect matchings of $\Gamma(M, M')$ to perfect matchings of $\bar{\Gamma}(M, M')$. This proves the first claim.

Suppose that $\text{Poly}_M \cap \text{Poly}_{M'}$ is nonempty and matroidal. Let (B_1, B_2) be an edge of $\text{Poly}_M \cap \text{Poly}_{M'}$, so B_1 and B_2 correspond to perfect matchings of $\Gamma(M, M')$ that differ only by a single edge—say, $B_1 \setminus \{e_1\} = B_2 \setminus \{e_2\}$. Then e_1 and e_2 must have the same image in $\bar{\Gamma}(M, M')$ in order for both B_1 and B_2 to be perfect matchings. Since every pair of vertices of $\text{Poly}_M \cap \text{Poly}_{M'}$ is connected by a chain of edges, we see that every perfect matching of $\Gamma(M, M')$ gives rise to the same perfect matching of

$\bar{\Gamma}(M, M')$, or, in other words, $\bar{\Gamma}(M, M')$ has exactly one perfect matching. Conversely, if $\bar{\Gamma}(M, M')$ has exactly one perfect matching, it is clear that $\text{Poly}_M \cap \text{Poly}_{M'}$ is a product of simplices and hence matroidal. $\square$

We now turn to the cases we are interested in, when $n = 2d$ or $2d + 1$. Let $M^\vee$ and $(M')^\vee$ be bounded facets of L , and let Γ and $\bar{\Gamma}$ be as defined in the lemma. First, suppose that $n = 2d$. Then every S_i must have order 2 so the graph Γ must consist of disjoint cycles. The only way that $\bar{\Gamma}$ can have at most one perfect matching is if all of those cycles have length 2 so that $\bar{\Gamma}$ is a disjoint union of edges. But then $\{S_i\} = \{T_i\}$ and $M = M'$. So we see that a linear space of dimension d in $2n$ -space can have at most one bounded facet.

Now suppose that $n = 2d + 1$. Let $M^\vee$ be a bounded facet of L . Then all of the connected components have order 2 except for one, which has order 3. We denote the connected components of M by $(R, S_2, \dots, S_d)$, with R the component of order 3. Let $(M')^\vee$ be another bounded facet of L , with connected components $(R', S'_2, \dots, S'_d)$ as above. The graph $\Gamma(M, M')$ consists of several disjoint cycles, each of even length, and one other component C . This other component consists of a pair of vertices v and w of degree 3 and three disjoint paths $\gamma_1, \gamma_2, \gamma_3$. Either these paths all connect v to w or there is one path each from v to v , from v to w , and from w to w . Moreover, after switching v and w if necessary, we may assume that v corresponds to R and w to R' , so any path from v to w is of odd length.

In each of these cases, Γ has at least one perfect matching. One can check that the only case where $\bar{\Gamma}$ has only one perfect matching is when all of the cycles of $\Gamma \setminus C$ have length 2 and C consists of a path from v to w and two cycles of length 2, one containing v and the other containing w . The condition on C is equivalent to requiring that there are some indices i and i' such that $S'_{i'} \subset R$ and $S_i \subset R'$.

Let f be the number of bounded facets of L , and let $(R^r, S^r_2, \dots, S^r_d)$ be the partition of $[n]$ corresponding to the r th facet of L . Note that, if there is any pair of matroids M and M' with $R = R'$, then $S_i = S'_i$ for every i (after reordering), as otherwise Γ would have a cycle of length more than 2. So all of the triplets $R^1, R^2, \dots, R^f$ are different. We may thus show that $f \leq d$ by showing that the number of distinct R^r 's is at most d .

We now build a graph G whose vertices are the various R^r 's and the labels $\{1, \dots, n\}$. Let there be an edge between R^r and k if $k \in R^r$. This graph has $f + n = f + 2d + 1$ vertices and $3f$ edges. Thus, if $f > d$, then G has at least as many edges as vertices and must contain a cycle; we will use this cycle to obtain a contradiction. We renumber the facets so that our cycle is $(R^1, i_2, R^2, i_3, \dots, R^{g-1}, i_g, R^g, i_1)$. Let $R^r = \{i_r, i_{r+1}, j_r\}$. (Our indices are cyclic modulo g .)

Now, for any r and r' , the compatibility condition on $\bar{\Gamma}(M^r, M^{r'})$ states that there must be some s , with $2 \leq s \leq d$, such that $S^r_s \subset R^{r'}$. Consider first the case where $(r, r') = (1, 2)$. So there is some $s_2 \geq 2$ so that $S^1_{s_2} \subset R^2$. Since we know that i_2 is in R^1 , the set $S^1_{s_2}$ must be $\{j_2, i_3\}$. Now, consider when $(r, r') = (1, 3)$. Again, there must be some $s_3 \geq 2$ so that $S^1_{s_3} \subset R^3$. Since j_2 is in $S^1_{s_2}$ but not in R^3 , the indices s_2 and s_3 must be different, and thus $S^1_{s_2}$ and $S^1_{s_3}$ must be disjoint. Since i_3 is already contained in $S^1_{s_2}$, we must have $S^1_{s_3} = \{j_3, i_4\}$. Continuing in this manner, we get a sequence of indices s_k , for $2 \leq k \leq g$, such that $S^1_{s_k} = \{j_k, i_{k+1}\}$. But, when $k = g$, then $S^1_{s_g} = \{j_g, i_1\}$ and $R^1 = \{i_1, i_2, j_1\}$ intersect, a contradiction. This concludes the proof of Theorem 3.2. $\square$

The f -vector conjecture predicts that, if $n = 2d + e$, L should have at most $\binom{d+e-1}{e}$ bounded facets. In the cases $e = 0$ and $e = 1$, the graph Γ always has at least one perfect matching, and the requirement that $\bar{\Gamma}$ have exactly one is very restrictive.

Once $e = 2$, it is possible for $\bar{\Gamma}$ to have no perfect matchings, and the problem grows much harder.

Nonetheless, it seems that this kind of reasoning, looking only at the facets of L and not at the smaller spaces that contain them, is adequate to resolve the f -vector conjecture for facets.

CONJECTURE 3.7. *Let $n = 2d + e$. It is not possible to find more than $\binom{d+e-1}{e}$ distinct partitions of $[n]$ into d disjoint subsets, $[n] = S_1^r \sqcup \cdots \sqcup S_d^r$, where each S_s^r has order at least 2, and, for every r and r' , the graph $\bar{\Gamma}$ formed from $\{S_1^r, \dots, S_d^r\}$ and $\{S_1^{r'}, \dots, S_d^{r'}\}$ has at most one perfect matching.*

Remark. When $d = 2$, the condition that (S_1, S_2) and (S_1', S_2') be compatible precisely says that one of the four intersections $S_i \cap S_j'$, $1 \leq i, j \leq 2$ is empty. It is well known that, given a collection of partitions $[n] = S_1 \sqcup S_2$ which are compatible in this manner, there is a tree T with no vertices of degree 2 and the leaves labeled by $[n]$ such that each partition consists of the leaves on one side or the other of an edge of T . Thus, the conjecture in this case is simply the (straightforward) statement that a tree with n leaves and no vertices of degree 2 has at most $n - 3$ edges which do not end at a leaf. When L is a two-dimensional tropical linear space, the tree in question is simply $L/(1, 1, \dots, 1)$; this picture is discussed in great detail in [20].

Remark. Consider the case where $d = 3$. In this case, there is an elegant way to describe the compatibility condition between (S_1, S_2, S_3) and (S_1', S_2', S_3') . Two partitions of $[n]$ into triples (S_1, S_2, S_3) and (S_1', S_2', S_3') are compatible if and only if, after reordering the S_i and S_i' , we have $S_1 \subseteq S_1'$ and $S_1 \cup S_2 \subseteq S_1' \cup S_2'$.

4. Stable intersections. In ordinary geometry, if a d plane and a d' plane meet transversely in n -space, then their intersection is a $d + d' - n$ plane. For tropical linear spaces, the same holds true, as we will show in Proposition 4.8. However, the situation for tropical linear spaces is better than that for ordinary linear spaces. In this section, for any tropical d -plane L and tropical d' -plane L' , we define a $d + d' - n$ plane called the stable intersection of L and L' . The stable intersection is always contained in $L \cap L'$ and is equal to $L \cap L'$ when L and L' meet transversely. The combined operations of orthogonal complement (changing L to $L^\perp$) and stable intersection allow us to generate a wide variety of tropical linear spaces and to mimic most standard constructions from linear algebra. Throughout this section, we take d and d' to be positive integers, with $0 < d, d' < n$ and $d + d' > n$.

If L and L' are two ordinary linear spaces of dimensions d and d' in n -space which meet transversely, then $L \cap L'$ is a $(d + d' - n)$ -dimensional space whose Plücker coordinates are given by

$$\mathbf{P}_J(L \cap L') = \sum_{\substack{I \cap I' = J \\ |I|=d \\ |I'|=d'}} \pm \mathbf{P}_I(L) \mathbf{P}_{I'}(L').$$

We omit the rule for determining the sign of each summand, as we will only use this formula for motivation. Note that the summation conditions on I and I' guarantee that $I \cup I' = [n]$.

We will discuss a tropical version of this formula. If L and L' are tropical linear spaces of dimension d and d' , define a real-valued function $p(L \cap_{\text{st}} L')$ on $\binom{[n]}{d+d'-n}$ by

$$p_J(L \cap_{\text{st}} L') = \min_{\substack{I \cap I' = J \\ |I|=d \\ |I'|=d'}} (p_I(L) + p_{I'}(L')).$$

At the moment, $p_J(L \cap_{\text{st}} L')$ is just defined as a formal symbol, there is not yet any tropical linear space called $L \cap_{\text{st}} L'$. We will need several combinatorial lemmas.

LEMMA 4.1. *Let M and M' be two matroids of rank r on $[n]$. If the polytopes Poly_M and $\text{Poly}_{M'}$ intersect, then their intersection contains a vertex of $\Delta(r, n)$.*

Proof. Let $R \subset \mathbb{R}^n$ be the convex hull of the characteristic vectors of those subsets of $[n]$ which are independent in M . (The characteristic vector of a subset S of $[n]$ is $\sum_{s \in S} e_s$.) Similarly, define R' using the independent subsets of M' . Let λ be the sum of the coordinate functions. Then Poly_M and $\text{Poly}_{M'}$ are the faces of R and R' (respectively) on which λ achieves its maximum value, namely, r . Since the intersection of Poly_M and $\text{Poly}_{M'}$ is nonempty, the maximum value of λ on the polytope $R \cap R'$ is also r . Thus, there must be a vertex v of $R \cap R'$ such that $\lambda(v) = r$. By [11, Theorem 45], all of the vertices of $R \cap R'$ are 0–1 vectors. The only 0–1 vectors in $\Delta(r, n)$ are the vertices of $\Delta(r, n)$. $\square$

We introduce several polyhedral objects which we will use throughout this section. We define a convex piecewise linear function $\mathbf{p}$ on $\Delta(d, n)$ by defining $\mathbf{p}(x)$ to be the minimum of $\sum_{I \in \binom{[n]}{d}} \alpha_I p_I$, where the minimum ranges over all $\binom{[n]}{d}$ -tuples α_I of nonnegative real numbers such that $\sum_{I \in \binom{[n]}{d}} \alpha_I = 1$ and $\sum_{I \in \binom{[n]}{d}} \alpha_I e_I = x$. In other words, as x ranges over $\Delta(d, n)$, the points $(x, \mathbf{p}(x))$ sweep out the bottom faces of the convex hull of $\{(e_I, p_I)\}_{I \in \binom{[n]}{d}}$. So $\mathbf{p}(e_I) = p_I$, and the faces of $\mathcal{D}$ are the regions on which $\mathbf{p}$ is linear. Similarly, we use $\mathbf{p}'$ to define a piecewise linear convex function $\mathbf{p}'$ on $\Delta(d', n)$. We then define a piecewise linear convex function $\mathbf{q}$ on the Minkowski sum $\Delta(d, n) + \Delta(d', n)$ by

$$\mathbf{q}(y) = \min_{\substack{x+x'=y \\ x \in \Delta(d, n) \\ x' \in \Delta(d', n)}} (\mathbf{p}(x) + \mathbf{p}'(x')).$$

We include $\Delta(d + d' - n, n)$ into the Minkowski sum $\Delta(d, n) + \Delta(d', n)$ by $\iota : e \mapsto e + (1, \dots, 1)$.

LEMMA 4.2. *For $J \in \binom{[n]}{d+d'-n}$, we have $p(L \cap_{\text{st}} L')_J = \mathbf{q}(\iota(e_J))$.*

Proof. By definition, $\mathbf{q}(\iota(e_J))$ is the minimum, over all $x \in \Delta(d, n)$ such that $\iota(e_J) - x$ lies in $\Delta(d', n)$, of $\mathbf{p}(x) + \mathbf{p}'(\iota(e_J) - x)$. Let F be the subset of points x of $\Delta(d, n)$ such that $\iota(e_J) - x$ is in $\Delta(d', n)$. It is easy to check that F is a face of $\Delta(d, n)$, isomorphic to $\Delta(2n - d - d', n)$. So $\mathbf{q}(\iota(e_J))$ is the maximum of the function $r(x) := \mathbf{p}(x) + \mathbf{p}'(\iota(e_J) - x)$ on F . Similarly, $p(L \cap_{\text{st}} L')_J$ is the minimum value of $r(x)$, when x ranges over the vertices of F . As a sum of convex piecewise linear functions, r is a convex piecewise linear function. Let $\mathcal{A}$, $\mathcal{A}'$, and $\mathcal{B}$ be the regular subdivisions of F induced by the functions $\mathbf{p}(x)$, $\mathbf{p}'(\iota(e_J) - x)$ and $r(x)$. Then (by the convexity of r) the function r achieves its minimum on a face of $\mathcal{B}$. Call this face Y . The subdivision $\mathcal{B}$ is the simultaneous refinement of $\mathcal{A}$ and $\mathcal{A}'$, so Y is the intersection of a face X of $\mathcal{A}$ and a face X' of $\mathcal{A}'$.

Now, X is also a face of $\mathcal{D}_p$ and X' is the image of a face of $\mathcal{D}_{p'}$ under the linear map $x \mapsto \iota(e_J) - x$. So X and X' are both matroidal polytopes in $\Delta(2n - d - d', n)$. Since $Y = X \cap X'$ is nonempty, we know by Lemma 4.1 that Y contains a vertex of F . So the minimum value of r on F is achieved at this vertex, as desired. $\square$

There is an operation $M \wedge M'$, called matroid intersection ([22, section 7.6]), which takes two matroids of ranks d and d' and produces a third matroid. By definition, the spanning sets of $M \wedge M'$ are precisely the sets of the form $U \cap U'$, where U and U' span M and M' respectively. The matroid $M \wedge M'$ always has rank at least $d + d' - n$. It has rank precisely $d + d' - n$ if and only if there exists a pair (I, I') of bases of M

and M' (respectively) such that $I \cup I' = [n]$. Let us restate this in the language of polyhedral geometry.

LEMMA 4.3. *Let M and M' be matroids of ranks d and d' on $[n]$. If $(\text{Poly}_M + \text{Poly}_{M'}) \cap \iota(\Delta(d + d' - n, n))$ is nonempty, then $M \wedge M'$ has rank $d + d' - n$ and $(\text{Poly}_M + \text{Poly}_{M'}) \cap \iota(\Delta(d + d' - n, n)) = \iota(\text{Poly}_{M \wedge M'})$.*

We can now prove the following result.

PROPOSITION 4.4. *The vector $p(L \cap_{\text{st}} L')$ is a tropical Plücker vector. Every face of $\mathcal{D}_{p(L \cap_{\text{st}} L')}$ is of the form $\text{Poly}_{M \wedge M'}$ where Poly_M and $\text{Poly}_{M'}$ are faces of $\mathcal{D}_p$ and $\mathcal{D}_{p'}$.*

Proof. First, consider the regular subdivision of $\Delta(d, n) + \Delta(d', n)$ induced by the convex piecewise linear function $\mathbf{q}$. Every face of this subdivision is of the form $\text{Poly}_M + \text{Poly}_{M'}$, where Poly_M and $\text{Poly}_{M'}$ are faces of $\mathcal{D}_p$ and $\mathcal{D}_{p'}$. Now, consider the regular subdivision of $\Delta(d + d', n)$ induced by the convex piecewise linear function $\iota \circ \mathbf{q}$. We will call this subdivision $\mathcal{E}$. By Lemma 4.3, every face of $\mathcal{E}$ is matroidal and of the form $\text{Poly}_{M \wedge M'}$ where Poly_M and $\text{Poly}_{M'}$ are faces of $\mathcal{D}_p$ and $\mathcal{D}_{p'}$. So, there is a convex piecewise linear function on $\Delta(d + d' - n, n)$, inducing a regular matroidal subdivision of the required form, and, by Lemma 4.2, this function evaluates to $p(L \cap_{\text{st}} L')_J$ at the vertices of $\Delta(d + d' - n, n)$. Let $\mathbf{r}$ be the convex piecewise linear function on $\Delta(d + d' - n, n)$ induced by $p(L \cap_{\text{st}} L')$, so $\mathcal{D}_{p(L \cap_{\text{st}} L')}$ is the regular subdivision induced by $\mathbf{r}$. Then $\mathbf{r}$ is the unique convex piecewise linear function such that $\mathbf{r}(e_J) = p(L \cap_{\text{st}} L')_J$, and every domain of linearity of $\mathbf{r}$ is a convex hull of a subset of the vertices of $\Delta(d + d' - n, n)$. Since $\iota \circ \mathbf{q}$ also has these properties, we know that $\mathbf{r} = \iota \circ \mathbf{q}$. So $\mathcal{D}_{p(L \cap_{\text{st}} L')} = \mathcal{E}$ and is a matroidal subdivision of the required form. $\square$

We thus see that there is a well-defined tropical linear space $L \cap_{\text{st}} L'$. Our aim now is to justify the notation $L \cap_{\text{st}} L'$ by showing that $L \cap_{\text{st}} L'$ deserves to be called the stable intersection of L and L' . We first need to introduce some additional matroid constructions.

Let N and N' be matroids on the underlying ground sets S and S' , with $S \cap S' = \{e\}$. Assume that e is not a loop of either M or M' . Then the *parallel connection* of N and N' is the matroid on $S \cup S'$ whose bases are those sets J such that either $J \cap S$ is a basis of M and $(J \cap S') \cup \{e\}$ is a basis of M' , or else $(J \cap S) \cup \{e\}$ is a basis of M and $J \cap S'$ is a basis of M' . (It is possible that both cases hold; this can only happen if e is in J .) For the fact that this is a matroid and various basic facts about this matroid; see [22, section 7.6]. When N and N' are graphical matroids (see [22, section 6.1]) associated to graphs G and G' , the parallel connection of N to N' is the graph formed by gluing G to G' at edge e . For later use, we also define the *series connection* of M and M' : its bases are those sets J such that either $J \cap S$ is a basis of M and $(J \cap S') \setminus \{e\}$ is a basis of M' , or else $(J \cap S) \setminus \{e\}$ is a basis of M and $J \cap S'$ is a basis of M' .

It is possible to compose several parallel connections along a tree structure; this is alluded to in Figure 7.10 of [22] and the corresponding discussion. We will need a precise definition. Let M^α be a collection of matroids indexed by α running through a finite set A , with S^α the ground set of M^α . Assume that, for α, β , and γ distinct elements of A , the intersection $S^\alpha \cap S^\beta$ contains at most one element and the intersection $S^\alpha \cap S^\beta \cap S^\gamma$ is empty. Form a graph G whose vertices are A and where there is an edge between α and β if and only if $S^\alpha \cap S^\beta$ is nonempty; we impose the condition that G is a forest. Finally, we impose that, if $S^\alpha \cap S^\beta = \{e\}$, then e is not a loop of M^α or M^β . Then we can define a matroid, the *parallel connection* of the set of matroids $\{M^\alpha\}_{\alpha \in A}$, as follows: a subset J of $\bigcup_{\alpha \in A} S^\alpha$ is a basis of the parallel connection of $\{M^\alpha\}_{\alpha \in A}$ if and only if there is an orientation of the graph G such that,

for each $\alpha \in A$, the set $(J \cap S^\alpha) \cup \bigcup_{\alpha \rightarrow \beta} (S^\alpha \cap S^\beta)$ is a basis of M^α . Here the union runs over edges of G oriented from α to β . So, when A has two elements and G is a single edge, this is the notion of parallel connection in the previous paragraph.

PROPOSITION 4.5. *The collection of subsets of $\bigcup_{\alpha \in A} S^\alpha$ described above forms the set of bases of a matroid.*

Proof. Our proof is by induction on the number of edges of G . If G is a collection of disjoint vertices, so the sets S^α are all disjoint, then the parallel collection of $\{M_\alpha\}_{\alpha \in A}$ is simply the direct sum $\bigoplus_{\alpha \in A} M^\alpha$. If G has some edge, say connecting β to γ , then modify the collection $\{M_\alpha\}_{\alpha \in A}$ by deleting M^β and M^γ and adding, in their place, the parallel connection of M^β and M^γ . This modifies the graph G by contracting the edge between α and β , so the new graph is still a forest, and, by induction, the parallel connection of this modified set of matroids is a matroid. It is straightforward to check that this parallel connection of this modified set is the same as the parallel connection of the original set $\{M_\alpha\}_{\alpha \in A}$. $\square$

We see from the proof that the parallel connection of a set of matroids is built by repeatedly replacing pairs of matroids by their parallel connection.

We now introduce one more piece of terminology. Recall the definition of the graph $\Gamma(M, M')$ from section 3. We say that M and M' are *transverse* if $\Gamma(M, M')$ is a forest without multiple edges. When M and M' are transverse and loop-free, we may take the parallel connection of $\{M_c\}_{c \in C} \cup \{M'_{c'}\}_{c' \in C'}$, as the graph $\Gamma(M, M')$ will be the graph G which is required to be a tree in the definition of parallel connection. (We leave it to the reader to check the conditions on intersections.)

LEMMA 4.6. *Suppose that M and M' are transverse. Then $M \wedge M'$ is the parallel connection of the union of the set of connected components of M and the set of connected components of M' . In particular, $M \wedge M'$ is a loop-free matroid whose connected components are in bijection with the components of $\Gamma(M, M')$; each connected component of $M \wedge M'$ can be formed from connected components of M and M' by the repeated use of the operation of parallel connection and the rank of $M \wedge M'$ is $d + d' - n$.*

Proof. Let N be the parallel connection of the union of the set of connected components of M and the set of connected components of M' . Let J be a $(d + d' - n)$ -element subset of $[n]$. We will show that J is a basis of N if and only if there are bases I and I' of M , and I' such that $I \cap I' = J$. Once we have shown this, in particular, it will follow that there exists a pair (I, I') of bases of M and M' , with $|I \cap I'| = d + d' - n$. So the rank of $M \wedge M'$ is $d + d' - n$ and the bases of $M \wedge M'$ are precisely the sets J of this form. So this will show that $N = M \wedge M'$.

Assume that J is a basis of N . Fix a corresponding orientation of $G = \Gamma(M, M')$. Define a subset I of $[n]$ as follows: i is in I if either $i \in J$ or if the edge of $\Gamma(M, M')$ corresponding to i is oriented from its endpoint which is in C to its endpoint which is in C' . (It is possible that both are true, i in I in this case.) Similarly, i is in I' if either $i \in J'$ or if the edge of $\Gamma(M, M')$ corresponding to i is oriented from its endpoint which is in C' to its endpoint which is in C . Then $I \cup I' = [n]$ and $I \cap I' = J$. By construction, for each $c \in C$, the intersection of S_c with I is a basis of M_c . Thus, I is a basis of M . Similarly, I' is a basis of M' . This argument is easily reversed. $\square$

Remark. Without the hypothesis of transversality, it is possible that M and M' are loop-free but $M \wedge M'$ is not, even if $M \wedge M'$ has rank $d + d' - n$. For example, let M be the rank 2 matroid on $\{1, 2, 3\}$, where 1 and 2 are parallel and 3 is a coloop. Then $M \wedge M$ is the rank 1 matroid, where 1 and 2 are loops and 3 is still a coloop.

We will also occasionally need the following variant on Lemma 4.6.

LEMMA 4.7. *Let M and M' be two loop-free matroids on $[n]$, and let $\Gamma = \Gamma(M, M')$. Suppose that $M \wedge M'$ has rank $d + d' - n$ and is loop-free. Then the connected components of $M \wedge M'$ are the labels of the edges of the connected components of Γ .*

Proof. We first reduce to the case that $\Gamma(M, M')$ is connected. Let $G_1, G_2, \dots, G_r$ be the components of $\Gamma(M, M')$. For i between 1 and r , let M_i and M'_i be the summands of M and M' supported on the edge labels of G_i . So $M = \bigoplus M_i$ and $M' = \bigoplus M'_i$, and M_i has the same underlying ground set as M'_i , but M_i and M'_i are not necessarily connected. Then it is easy to check that $M \wedge M' = \bigoplus (M_i \wedge M'_i)$. So it is enough to check that each matroid $M_i \wedge M'_i$ is connected. Replacing M and M' by M_i and M'_i , we reduce to the case that $\Gamma(M, M')$ is connected.

Consider the polytope $\text{Poly}_M + \text{Poly}_{M'}$. We claim that this polytope has dimension $n - 1$. To see this, let λ be a linear function which is constant on $\text{Poly}_M + \text{Poly}_{M'}$, then λ must be constant on both Poly_M and $\text{Poly}_{M'}$. This means that we have $\lambda(e_i) = \lambda(e_j)$ whenever i and j label edges of $\Gamma(M, M')$ adjacent to the same vertex. Since the graph $\Gamma(M, M')$ is connected, this means that $\lambda(e_1) = \lambda(e_2) = \dots = \lambda(e_n)$. So the space of linear functions constant on $\text{Poly}_M + \text{Poly}_{M'}$ is one-dimensional, and we conclude that $\text{Poly}_M + \text{Poly}_{M'}$ has dimension $n - 1$. We want to conclude that $\text{Poly}_{M \wedge M'}$ also has dimension $n - 1$. Since $M \wedge M'$ has rank $d + d' - n$, the polytope $\text{Poly}_{M \wedge M'}$ is isomorphic (under the isometry ι) to $(\text{Poly}_M + \text{Poly}_{M'}) \cap \iota(\Delta(d + d' - n, n))$, so we want to show that $(\text{Poly}_M + \text{Poly}_{M'}) \cap \iota(\Delta(d + d' - n, n))$ has dimension $n - 1$.

If this intersection does not have dimension $n - 1$, then $\iota(\text{Poly}_{M \wedge M'}) = (\text{Poly}_M + \text{Poly}_{M'}) \cap \iota(\Delta(d + d' - n, n))$ must be contained in a facet of $\iota(\Delta(d + d' - n, n))$. Every facet of $\iota(\Delta(d + d' - n, n))$ is either of the form $\{x \in \mathbb{R}^n : x_i = 1\}$ or $\{x \in \mathbb{R}^n : x_i = 2\}$. If $\iota(\text{Poly}_{M \wedge M'})$ is contained in $\{x \in \mathbb{R}^n : x_i = 1\}$, then i is a loop of $M \wedge M'$, contrary to our hypothesis. If $\iota(\text{Poly}_{M \wedge M'})$ is contained in $\{x \in \mathbb{R}^n : x_i = 2\}$, then i is a coloop of $M \wedge M'$. We claim that this implies that i is a coloop of both M and M' . If i is not a coloop of M , then there is some basis B of M which does not contain i . Let B' be any basis of M' , then $B \cap B'$ is a spanning set of $M \wedge M'$ which does not contain i , contradicting our assumption that i is a coloop of $M \wedge M'$. So i is a coloop of both M and M' , and thus $\{i\}$ is a connected component of both M and M' . Then the edge of $\Gamma(M, M')$ labeled i is, in itself, a connected component of $\Gamma(M, M')$. Since we assumed that $\Gamma(M, M')$ is connected, this means that n is 1 and that M and M' both consist of a single coloop. In this case, $M \wedge M'$ is also a single coloop, and the claim is true in this case as well. $\square$

We now relate $L \cap_{\text{st}} L'$ to the intersection $L \cap L'$.

PROPOSITION 4.8. *Let $0 < d, d' < n$, with $d + d' > n$. Suppose that L and L' are tropical linear spaces in n -space of dimension d and d' . Then $L \cap_{\text{st}} L' \subseteq L \cap L'$. If L and L' meet transversely, then $L \cap_{\text{st}} L' = L \cap L'$.*

By “meet transversely” we mean that, if $w \in L \cap L'$ and $M^\vee$ and $(M')^\vee$ are the faces of L and L' whose relative interiors contain w , then the vector spaces $\mathbb{R}M^\vee$ and $\mathbb{R}(M')^\vee$ are transverse. We first must relate this notion of transversality to the notion of transversality of matroids defined above.

LEMMA 4.9. *Let L and L' be transverse tropical linear spaces in $\mathbb{R}^n$. Let w be a point of L and L' and $M := M_w$ and $M' := M'_w$ the corresponding loop-free matroids. Then the matroids M and M' are transverse.*

Proof. Let $M = \bigoplus_{c \in C} M_c$ and $M' = \bigoplus_{c' \in C'} M'_{c'}$ be the decompositions of M and M' into connected components. Let $|C| = r$ and $|C'| = r'$, so r and r' are the dimensions of $M^\vee$ and $(M')^\vee$. The linear space $\mathbb{R}M^\vee$ is cut out by the equations $x_i - x_j = 0$ when the edges of $\Gamma(M, M')$ labeled by i and j have the same endpoint in C ,

and similarly for $(M')^\vee$. Thus, $\mathbb{R}M^\vee \cap \mathbb{R}(M')^\vee$ is cut out by the equations $x_i - x_j = 0$ whenever the edges labeled by i and j are in the same component of $\Gamma(M, M')$. We assumed that L and L' meet transversely, so the dimension of $M^\vee \cap (M')^\vee$ must be $r + r' - n$, and Γ must have $r + r' - n$ connected components. A graph with $r + r'$ vertices, n edges, and $r + r' - n$ connected components must be a forest without multiple edges. $\square$

A very similar argument establishes the following more general result.

LEMMA 4.10. *Let L and L' be tropical linear spaces in $\mathbb{R}^n$, with w in $L \cap L'$. Suppose that $\mathbb{R}M^\vee + \mathbb{R}(M')^\vee$ has dimension $n - b$. Then the graph $\Gamma(M, M')$ has first betti number b .*

Proof of Proposition 4.8. Let $\mathcal{D}$, $\mathcal{D}'$, and $\mathcal{E}$ be the subdivisions of $\Delta(d, n)$, $\Delta(d', n)$, and $\Delta(d + d' - n, n)$ corresponding to L , L' , and $L \cap_{\text{st}} L'$. Let $\mathcal{D}^\vee$, $(\mathcal{D}')^\vee$, and $\mathcal{E}^\vee$ denote the dual subdivisions of $\mathbb{R}^n$. Recall the notations ι , $\mathbf{p}$, $\mathbf{p}'$, and $\mathbf{q}$ from the beginning of the section; so $\mathcal{D}$, $\mathcal{D}'$, and $\mathcal{E}$ are the regular matroidal subdivisions induced by the convex, piecewise linear functions $\mathbf{p}$, $\mathbf{p}'$, and $\iota \circ \mathbf{q}$.

Consider a particular $w \in \mathbb{R}^n$; we need to show that $w \in L \cap_{\text{st}} L'$ implies that $w \in L$ and $w \in L'$, and that the converse holds if L and L' meet transversely. Let $M^\vee$, $(M')^\vee$, and $N^\vee$ be the faces of $\mathcal{D}^\vee$, $(\mathcal{D}')^\vee$, and $\mathcal{E}^\vee$, respectively whose relative interiors contain w , and let Poly_M , $\text{Poly}_{M'}$, and Poly_N be the respective dual faces of $\mathcal{D}$, $\mathcal{D}'$, and $\mathcal{E}$. We must show that, if M or M' has a loop, then N has a loop, and that if L and L' meet transversely and N has a loop, then M or M' does as well.

First suppose that M has a loop; we will show that N has a loop as well. Without loss of generality, we may assume that $w = 0$. So the functions $\mathbf{p}$ and $\mathbf{p}'$ are minimized on Poly_M and $\text{Poly}_{M'}$, and hence $\mathbf{q}$ is minimized on $\text{Poly}_M + \text{Poly}_{M'}$; we will abbreviate $\text{Poly}_M + \text{Poly}_{M'}$ by A . Also, $\iota \circ \mathbf{q}$ is minimized on Poly_N . If A meets $\iota(\Delta(d + d' - n, n))$, then $M \wedge M'$ has rank $d + d' - n$. In that case, the minimum of $\iota \circ \mathbf{q}$ is achieved on $\text{Poly}_{M \wedge M'}$ and $N = M \wedge M'$. Since M has a loop, $N = M \wedge M'$ has a loop as well in that case. Now, let us consider the case where A does not meet $\iota(\Delta(d + d' - n, n))$. We adopt the abbreviation B for $\iota(\text{Poly}_N)$, so $\mathbf{q}$ is constant on both A and B , with $\mathbf{q}|_A < \mathbf{q}|_B$. Suppose (for the purpose of contradiction) that N does not have a loop, and hence B is not contained in $\{z \in \mathbb{R}^n : z_e = 1\}$ for any $e \in [n]$. Then there is some point, call it b , in B where none of the coordinate functions are 1. Let a be a point of A , and let $\overline{ab}$ be the line segment joining a to b . By assumption, $q(a) < q(b)$, and, as q is convex, we have $q(v) < q(b)$ for all $v \in \overline{ab} \setminus \{b\}$. This means that $\overline{ab}$ meets $\iota(\Delta(d + d' - n, n))$ only at b . So there is either some coordinate e for which $b_e = 1$ and $a_e < 1$, or there is some coordinate $e \in [n]$ for which $b_e = 2$ and $a_e > 2$. But the first possibility violates our assumption that none of the coordinates of b_e are 1, and the second violates the fact that a lies in $\Delta(d, n) + \Delta(d', n)$. So we have a contradiction, B is contained in $\{z \in \mathbb{R}^n : z_e = 1\}$ for some $e \in [n]$ and this e is a loop of N .

Now assume that M and M' have no loops. For every e in $[n]$, this implies that both Poly_M and $\text{Poly}_{M'}$ have points where $z_e = 1$, and hence there is a point of $\text{Poly}_M + \text{Poly}_{M'}$ where $z_e \geq 2$. Since the hyperplanes cutting out $\iota(\Delta(d + d' - n, n))$ inside of $\Delta(d, n) + \Delta(d', n)$ are all of the form $\{z \in \mathbb{R}^n : z_e \leq 1\}$, we know that $\text{Poly}_M + \text{Poly}_{M'}$ meets $\iota(\Delta(d + d' - n, n))$. Thus, $M \wedge M'$ has rank $d + d' - n$ and $N = M \wedge M'$. Now, assume further that L and L' are transverse. Then, by Lemma 4.9, the matroids M and M' are transverse. From Lemma 4.6, this implies that $M \wedge M'$ is loop-free. $\square$

The next theorem explains the motivation for the notation $L \cap_{\text{st}} L'$ and the geometric meaning of $L \cap_{\text{st}} L'$ when L and L' are not transverse.

THEOREM 4.11. *Let L and L' be tropical linear spaces. Then, for a generic $v \in \mathbb{R}^n$, L and $L' + v$ meet transversely. We have $L \cap_{\text{st}} L' = \lim_{v \rightarrow 0} (L \cap (L' + v))$ as long as v approaches 0 along a generic direction.*

By *generic* we mean “for an open dense subset of the appropriate parameter space.” That is to say, L meets $L' + v$ transversely for v in an open dense subset of $\mathbb{R}^n$, and $L \cap_{\text{st}} L' = \lim_{v \rightarrow 0} (L \cap (L' + v))$ as long as v approaches 0 along a ray whose direction lies in an open dense subset of S^{n-1} . This result will hold for virtually any reasonable topology on the set of polyhedral complexes in $\mathbb{R}^n$ —we use the Hausdorff metric D induced from the metric $d(x, y) := \max_{j \in [n]} |x_j - y_j|$. Explicitly, for X and Y closed subsets of $\mathbb{R}^n$, the distance $D(X, Y)$ is given by $\max(\sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y))$. (This is equivalent to the Hausdorff metric induced by the standard Euclidean metric on $\mathbb{R}^n$.)

Proof. Let F and F' be faces of L and L' . Suppose that $\mathbb{R}F + \mathbb{R}F'$ is not $\mathbb{R}^n$. Then, for v outside a proper subspace of $\mathbb{R}^n$, $F \cap (F' + v) = \emptyset$. Thus, for a generically chosen v , every such F and $F' + v$ fail to meet. The only faces of L and $L' + v$ that do meet, then, meet transversely.

We now need to check that the function $p \mapsto L(p)$ is continuous. More specifically, we show that $D(L(p), L(p')) \leq \delta(p, p') := \max_{I \in \binom{[n]}{d}} |p_I - p'_I|$. First, note that this is obvious when $\dim L(p) = 1$, as then $L(p) = (p_1, \dots, p_n) + \text{Span}_{\mathbb{R}}(1, \dots, 1)$. Now we handle the general case where $\dim L(p) = \dim L(p') = d$. Let $\delta = \delta(p, p')$, and let $x \in L(p)$. We must show that there is a point x' of $L(p')$, with $d(x, x') \leq \delta$. Let $L(p'')$ be any tropical linear space of dimension $n - d + 1$ passing through x and transverse to $L(p)$; we have $x + \text{Span}_{\mathbb{R}}(1, \dots, 1) = L(p) \cap_{\text{st}} L(p'')$. By adding a constant to all of the p'_I , we may assume that $x_j = \min_{I \cap I'' = \{j\}} (p_I + p''_{I''})$. (A priori, we only know this equality up to adding a common scalar to both sides.) Let x' be given by the formula $x'_j = \min_{I' \cap I'' = \{j\}} (p'_{I'} + p''_{I''})$. We have

$$x'_j \in x'_j + \text{Span}_{\mathbb{R}}(1, \dots, 1) = L(p') \cap_{\text{st}} L(p'') \subseteq L(p') \cap L(p'') \subseteq L(p').$$

We now check that $d(x, x')$ is sufficiently small. Choose an index j in $[n]$. Let (I, I'') be where the minimum in the equation $x_j = \min_{I \cap I'' = \{j\}} (p_I + p''_{I''})$ is achieved. Then we have

$$x_j - x'_j = (p_I + p''_{I''}) - \min_{K' \cap K'' = \{j\}} (p'_{K'} + p''_{K''}) \leq (p_I + p''_{I''}) - (p'_I + p''_{I''}) = p_I - p'_I \leq \delta,$$

and similarly $x'_j - x_j \leq \delta$. So we see that $d(x, x') = \max_{j \in [n]} |x_j - x'_j| \leq \delta$ as desired.

We have $p_I(L' + v) = p_I(L') + \sum_{i \in I} v_i$ so the tropical Plücker coordinates of $L' + v$ vary continuously with v . The tropical Plücker coordinates of $L \cap_{\text{st}} (L' + v)$ similarly vary continuously with v . As $L(p)$ is a continuous function of p , this means that $L \cap_{\text{st}} (L' + v)$ varies continuously with v . When v is chosen generically, L and $L' + v$ meet transversely so $L \cap (L' + v) = L \cap_{\text{st}} (L' + v)$. Take limits of both sides as $v \rightarrow 0$ and use the preceding continuity arguments to conclude that $\lim_{v \rightarrow 0} (L \cap (L' + v)) = L \cap_{\text{st}} L'$. $\square$

The notion of the stable intersection of two polyhedral complexes, defined by the formula $\lim_{v \rightarrow 0} (L \cap (L' + v))$, was introduced in [16] and shown to be well defined in the case of tropical hypersurfaces (defined in that paper). This operation for fans was described earlier, without this name, in [12] in the course of giving a combinatorial rule for multiplication in the cohomology ring of a toric variety.

5. Series-parallel matroids. Recall that we define a tropical linear space to be *constructible* if it can be built from hyperplanes (tropical linear spaces of dimension $n - 1$ in n -space) by successive dualization and transverse intersection. Let L be a constructible tropical linear space of dimension d in n -space, with $\mathcal{D}$ the corresponding division of $\Delta(d, n)$. Then every face of $\mathcal{D}$ is of the form Poly_M , where M is built from matroids of corank 1 by dualization and parallel connection. (We will check this in section 6.) It is therefore of interest to understand the class of matroids which can be built in this way. A connected matroid can be built in this way if and only if it is series-parallel. In general, a matroid can be built in this way if and only if its connected components are series-parallel matroids, loops, and coloops. In this section, we review the definition of series-parallel matroids and review several key lemmas about them. We will present a very streamlined version of this theory, ignoring the connections to graphs, electrical networks, and trees. The reader wanting more background on series-parallel matroids should consult [3]. We note that several of the references we cite in this section are written in the language of graph theory rather than that of matroid theory; for the connections between the two subjects, see [22, Chapter 6].

Let M be a matroid on the ground set E , and let e be an element of E . The *parallel extension* of M at e denoted $P(M, e)$ is a matroid on the ground set $\hat{E} := E \setminus \{e\} \cup \{e_1, e_2\}$. The bases of $P(M, e)$ are the subsets I of $\hat{E}$ such that either I does not contain either e_1 or e_2 and I is a basis of M , or I contains exactly one of e_1 and e_2 and $I \setminus \{e_1, e_2\} \cup \{e\}$ is a basis of M . Similarly, the *serial extension* of M at e denoted $S(M, e)$ is also a matroid on the ground set $\hat{E}$. The bases of $S(M, e)$ are the subsets I of $\hat{E}$ such that either I contains both e_1 and e_2 and $I \setminus \{e_1, e_2\} \cup \{e\}$ is a basis of M , or I contains precisely one of e_1 and e_2 , and $I \setminus \{e_1, e_2\}$ is a basis of M . We will want to know how series and parallel extensions effect the properties of being connected and loop-free. We leave the proof to the reader.

LEMMA 5.1. *Let M be a matroid with e an element of the ground set of M . If M has at least two elements, then M is connected if and only if $P(M, e)$ is connected if and only if $S(M, e)$ is connected. If any of these equivalent conditions hold, then M , $P(M, e)$, and $S(M, e)$ are all loop-free. We now describe the cases where M has one element. If M is a loop, then $S(M, e)$ is connected and loop-free, while $P(M, e)$ is neither. If M is a coloop, then $P(M, e)$ is connected and loop-free, while $S(M, e)$ is loop-free but disconnected.*

By definition, a *series-parallel* matroid is a matroid which can be produced from $\text{Uniform}(1, 2)$ by the repeated use of the operations of series and parallel extension. In this section, we prove several results on series-parallel matroids for future use.

PROPOSITION 5.2. *The dual of a series-parallel matroid is series-parallel.*

Proof. See [22, Proposition 6.4.5] for a proof in the language of graph theory. $\square$

PROPOSITION 5.3. *The parallel connection of two series-parallel matroids is series-parallel.*

Proof. In section 7 of [3], Brylawski gives a definition of “series-parallel network” which is tautologically closed under taking parallel connections. Brylawski’s definition clearly includes ours, as serial and parallel extensions are special cases of series and parallel connection. After Theorem 7.6 of [3], Brylawski notes that any matroid which is series-parallel by his definition can be built from $\text{Uniform}(1, 2)$ by repeated parallel extension and dualization. Since the dual of a parallel extension is a series extension and $\text{Uniform}(1, 2)$ is self-dual, this is equivalent to saying that every series-parallel network in Brylawski’s sense can be built from $\text{Uniform}(1, 2)$ by series and parallel extensions. This is our definition. $\square$

LEMMA 5.4. *Let M be a series-parallel matroid with at least three elements, and let f be an element of M . Then either M/f is connected and $M \setminus f$ is not, or vice versa.*

Proof. For the claim that at least one of M/f and $M \setminus f$ is connected, see “Lemma to Theorem II” in [4]. For the claim that at most one is, see Condition (6) in Theorem 7.6 of [3]. $\square$

PROPOSITION 5.5. *Let M be a series-parallel matroid, and let Poly_N be a face of Poly_M . Then the connected summands of N are all series-parallel, loops, or coloops.*

Proof. From Proposition 2.4, N is of the form $\text{gr}(M, F_\bullet)$ for some filtration F of M . By definition, $\text{gr}(M, F_\bullet)$ is a direct sum of the minors of M . By Lemma 7.7 of [3] and its dual, every minor of a series-parallel matroid is a direct sum of the matroids of the required forms. $\square$

Let M be a connected loop-free matroid. Following a suggestion of Sturmfels, we refer to proper, nonempty flats F of M such that $M|_F$ and M/F are both connected as *flacets*. With some redundancy, we can define a flacet as a proper, nonempty subset F of M so that both $M|_F$ and M/F are connected and loop-free. This is redundant (but correct) because F is a flat if and only if M/F is loop-free, while $M|_F$ never has loops because M doesn’t. The characteristic vectors of the flacets of M are the rays of the Bergman fan of M , as described in [2]. In the following straightforward, but tedious, lemma, we describe the effects of series and parallel extension on the set of flacets of a matroid.

LEMMA 5.6. *Let M be a connected loop-free matroid on the ground set E , with E of size at least two, and let e be an element of E . Let $\hat{E} = E \setminus \{e\} \cup \{e_1, e_2\}$. Let π be the function $\hat{E} \rightarrow E$ defined by $\pi(e_1) = \pi(e_2) = e$ and $\pi(e') = e'$ in all other cases. The flacets of $P(M, e)$ are precisely the sets $\pi^{-1}(F)$, where F is a flacet of M . Except for a few special cases which we will state shortly, a subset $\hat{F}$ of $\hat{E}$ is a flacet of $S(M, e)$ if and only if the set $\hat{F}$ is of the form $\pi^{-1}(F)$, where F is a flacet of M . The exceptions to this claim are the following: the singletons $\{e_1\}$ and $\{e_2\}$ are flacets of $\hat{M}$, the pair $\{e_1, e_2\}$ is not, and if $M \setminus e$ is connected, then $\hat{E} \setminus \{e_1, e_2\}$ is a flacet of $\hat{M}$.*

We use Lemma 5.1 repeatedly.

Proof. We treat the case of $P(M, e)$ first. Since e_1 is in the closure of e_2 and vice versa, every flat of $P(M, e)$ contains either both M_1 and M_2 or neither. Let $\hat{F}$ be a subset of $\hat{E}$ which contains either both e_1 and e_2 or neither; we will describe the circumstances under which $\hat{F}$ is a flacet of $P(M, e)$. First, suppose that $\hat{F}$ does not contain e_1 or e_2 . Then set $F = \hat{F}$ and observe that $\hat{M}|_{\hat{F}} = M|_F$ and $\hat{M}/\hat{F} = P(M/F, e)$, so $\hat{M}|_{\hat{F}}$ and $\hat{M}/\hat{F}$ are connected and loop-free if and only if $M|_F$ and M/F are. Alternatively, suppose that $\hat{F} = F \setminus \{e\} \cup \{e_1, e_2\}$ for a subset F of E containing e . Then $\hat{M}|_{\hat{F}} = P(M|_F, e)$ and $\hat{M}/\hat{F} = M/F$, so again $\hat{M}|_{\hat{F}}$ and $\hat{M}/\hat{F}$ are connected and loop-free if and only if $M|_F$ and M/F are.

Now, we consider the case of $S(M, e)$. If $\hat{F}$ is any subset of $\hat{E}$ which contains e_1 and not e_2 , then e_1 is a coloop of $\hat{M}|_{\hat{F}}$. So, in this case, the only way that $\hat{M}|_{\hat{F}}$ could be connected is if $\hat{F} = \{e_1\}$, and, on the other hand, it is easy to check in this case that $\{e_1\}$ is a flacet. Similarly, $\{e_2\}$ is the only flacet that contains e_2 and not e_1 . We now consider the case that $\hat{F}$ contains neither e_1 nor e_2 . Set $F = \hat{F}$. We have $\hat{M}|_{\hat{F}} = M|_F$ and $\hat{M}/\hat{F} = S(M/F, e)$. So $\hat{M}|_{\hat{F}}$ is connected and loop-free if and only if $M|_F$ is. The situation for $\hat{M}/\hat{F}$ is a little trickier: $\hat{M}/\hat{F}$ is connected and loop-free if and only if either $\hat{M}/\hat{F}$ is connected and loop-free or else if M/F is a loop. (It is impossible that M/F is a coloop, as then M would have a coloop, contradicting that

it is a connected matroid with at least two elements.) So, when $\hat{F}$ contains neither e_1 nor e_2 , we see that $\hat{F}$ is a facet if and only if either F is a facet or else $F = E \setminus \{e\}$ and $M|_F$ is connected. Finally, we consider the case where $\hat{F}$ contains both e_1 and e_2 . Set $F = \pi(\hat{F})$. Then $\hat{M}|_{\hat{F}} = S(M|_F, e)$ and $\hat{M}/\hat{F} = M/F$. This time, $\hat{M}/\hat{F}$ is connected and loop-free if and only if M/F and the difficulties come with $\hat{M}|_{\hat{F}}$. Specifically, $\hat{M}|_{\hat{F}}$ is loop-free and connected as long as $M|_F$ is loop-free and connected and $M|_F$ is not a coloop. The last criterion simply rules out the case $F = \{e\}$. (It is impossible that $M|_F$ is a loop, as that would imply that M had a loop, contradicting that M is connected and has multiple elements.) So, when $\hat{F}$ contains both e_1 and e_2 , then $\hat{F}$ is a facet if and only if F is a facet other than $\{e\}$. $\square$

LEMMA 5.7. *Let M be a series-parallel matroid, with f and g distinct elements of the ground set of M . Let $F_1, F_2, \dots, F_{d(f,g,M)}$ be the set of facets of M which contain f but not g . Then, after reordering, these facets form a nested chain $F_1 \subsetneq F_2 \subsetneq \dots \subsetneq F_{d(f,g,M)}$. For any $i < j$, the quotient F_j/F_i is connected. Furthermore, $d(f, g, M) = d(g, f, M)$.*

For any matroid M with elements f and g , we will define $d(f, g, M)$ to be the number of facets of M containing f and not g , even when M is not series-parallel.

Proof. Our proof is by induction on the cardinality of the ground set of M ; the base case where M has two elements is clear. So, consider a series-parallel matroid $\hat{M}$ on $\hat{E}$, with $\hat{M}$ equal to either $S(M, e)$ or $P(M, e)$ for some smaller series-parallel matroid M . By induction, the lemma has already been proven for M . We continue to use the notations e, e_1 , and e_2 as usual. We also write π for the function $\hat{E} \rightarrow E$ defined by $\pi(e_1) = \pi(e_2) = e$ and $\pi(e') = e'$ in all other cases. We break into several cases.

Case 1. The matroid $\hat{M}$ is a parallel extension of M . In this case, once again Lemma 5.6 shows that F is a facet of M containing f and not g if and only if $\pi^{-1}(F)$ is a facet of $\hat{M}$ doing the same. Taking inverse images under π preserves the property of being nested and preserves the length of nested chains. Furthermore, $\pi^{-1}(F_j)/\pi^{-1}(F_i)$ is either isomorphic to F_j/F_i or to a series or parallel extension thereof.

Case 2. The elements f, g, e_1 , and e_2 are all distinct, and $\hat{M} = S(M, e)$. Then Lemma 5.6 gives the same conclusion as in Case 1: F is a facet of M containing f and not g if and only if $\pi^{-1}(F)$ is a facet of $\hat{M}$ doing the same. We complete the argument as before.

Case 3. The matroid $\hat{M}$ is $S(M, e)$, and $(f, g) = (e_1, e_2)$. Then the only facet containing f and not g is $\{e_1\}$, and the only facet containing g and not f is $\{e_2\}$. Since there is only one facet separating f from g , the condition on quotients is empty.

Case 4. The matroid $\hat{M}$ is $S(M, e)$, with $e_1 = f$ and $g \neq e_2$. This is the only interesting case. Let $F_1 \subsetneq F_2 \subsetneq \dots \subsetneq F_{d(e,g,M)}$ be the chain of facets in M containing e and not g , while $G_1 \subsetneq G_2 \subsetneq \dots \subsetneq G_{d(g,e,M)}$ is the chain of facets in M containing g but not e . Now, by Lemma 5.4, either M/e is connected and $M \setminus e$ is not, or vice versa. We split into subcases.

Case 4a. The matroid M/e is connected and $M \setminus e$ is not. In this case, $\{e\}$ is a facet of M and is necessarily equal to F_1 . Consulting Lemma 5.6, the facets of $\hat{M}$ containing f and not g are

$$\{f\} = \{e_1\} \subsetneq \pi^{-1}(F_2) \subsetneq \dots \subsetneq \pi^{-1}(F_{d(e,g,M)}).$$

On the other hand, we also know in this case that $\hat{E} \setminus \{e_1, e_2\}$ is not a facet of $\hat{M}$,

so, consulting Lemma 5.6 again, the facets of $\hat{M}$ containing g and not f are

$$\pi^{-1}(G_1) \subsetneq \pi^{-1}(G_2) \subsetneq \cdots \subsetneq \pi^{-1}(G_{d(g,e,M)}).$$

We see that we have two nested chains of length $d(e, g, M) = d(g, e, M)$. We must check the condition on quotients: For $2 \leq i < j$, we have $\pi^{-1}(F_j)/\pi^{-1}(F_i) \cong F_j/F_i$, which is connected by the inductive hypothesis, and similarly $\pi^{-1}(G_j)/\pi^{-1}(G_i) \cong G_j/G_i$ is connected. Also, $\pi^{-1}(F_j)/e_1 \cong F_j$, which is connected by the definition of a facet.

Case 4b. The matroid M/e is not connected and $M \setminus e$ is. In this case, $\{e\}$ is not a facet of M , so the facets of $\hat{M}$ containing f and not g are

$$\{f\} = \{e_1\} \subsetneq \pi^{-1}(F_1) \subsetneq \pi^{-1}(F_2) \subsetneq \cdots \subsetneq \pi^{-1}(F_{d(e,g,M)}).$$

On the other hand, we also know in this case that $\hat{E} \setminus \{e_1, e_2\}$ is a facet of $\hat{M}$, while $E \setminus \{e\}$ is certainly not a facet of M . So the facets of $\hat{M}$ containing g and not f are

$$\pi^{-1}(G_1) \subsetneq \pi^{-1}(G_2) \subsetneq \cdots \subsetneq \pi^{-1}(G_{d(g,e,M)}) \subsetneq \hat{M} \setminus \{e_1, e_2\}.$$

We see that we have two nested chains of length $d(e, g, M) + 1 = d(g, e, M) + 1$. Once again, we must check the condition on quotients. The quotients $\pi^{-1}(F_j)/\pi^{-1}(F_i)$, $\pi^{-1}(F_j)/e_1$, and $\pi^{-1}(G_j)/\pi^{-1}(G_i)$ are all connected by the arguments in Case 4a. Finally, we must consider $(\hat{M} \setminus \{e_1, e_2\})/\pi^{-1}(G_i)$. First, note that $(\hat{M} \setminus \{e_1, e_2\})/\pi^{-1}(G_i)$ is isomorphic to $M \setminus e/G_i = (M/G_i) \setminus e$. So we must show that $(M/G_i) \setminus e$ is connected.

Assume for the sake of contradiction that $(M/e)/G_i$ is connected. We are currently assuming that M/e is disconnected so if $(M/e)/G_i$ is connected, then G_i must contain a connected component K of M/e . The closure of K , taken in M , contains e . (Otherwise, M would also be disconnected.) So e is in the closure of G_i (but not in G_i), contradicting that G_i is a flat of M . So we have a contradiction, and $(M/e)/G_i$ is, in fact, disconnected. Now, M/G_i is connected (because G_i is a facet), so, by Proposition 5.5, the quotient M/G_i is series-parallel. We have just shown that $(M/G_i)/e$ is disconnected, so, by Lemma 5.4, $(M/G_i) \setminus e$ is connected. This is what we wanted to show. $\square$

6. Proof of Theorem 1.3. We now have enough background to prove Theorem 1.3, which states that every constructible tropical linear space is series-parallel. We recall the definition of these concepts—a tropical linear space is called series-parallel if every facet of the corresponding subdivision of the hypersimplex is Poly_M for some series-parallel M . A tropical linear space is called constructible if it is built from tropical hyperplanes by the repeated use of the operations of orthogonal complement and transverse intersection. As the definition of constructibility suggests, we will prove Theorem 1.3 by showing that the orthogonal complement of a series-parallel linear space, or the transverse intersection of two series-parallel linear spaces, is series-parallel. Clearly, $L^\perp$ is series-parallel if L is, since the dual of a series-parallel matroid is series-parallel. So it is enough to show that if L and L' are two series-parallel tropical linear spaces that meet transversely, then $L \cap L'$ is series-parallel as well.

LEMMA 6.1. *If L is a series-parallel tropical linear space corresponding to the matroidal subdivision $\Delta(d, n)$, then, for every face Poly_M of $\mathcal{D}$, each connected component of M is series-parallel, a loop, or a coloop.*

Proof. Every face Poly_M of $\mathcal{D}$ is the face of some facet Poly_N of $\mathcal{D}$, where N is series-parallel by assumption. Now apply Proposition 5.5. $\square$

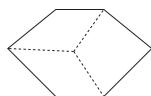


FIG. 2. A 3-simplicial division of a hexagon.

We now prove Theorem 1.3.

Proof of Theorem 1.3. Let L and L' be two series-parallel tropical linear spaces of dimensions d and d' in n -space which meet transversely at a point w . Let $\mathcal{D}$ and $\mathcal{D}'$ be the appropriate subdivisions of $\Delta(d, n)$ and $\Delta(d', n)$, respectively. Let $M^\vee$ and $(M')^\vee$ be the faces of $\mathcal{D}^\vee$ and $(\mathcal{D}')^\vee$ whose relative interiors contain w , then (by Corollary 4.3) $(M \wedge M')^\vee$ is the face of $L \cap L'$ containing w . Let $M = \bigoplus_{i \in I} M_i$ and $M' = \bigoplus_{i \in I'} M'_i$ be the connected components of M and M' ; note that none of the M_i or M'_i is a loop, so, by Lemma 6.1, the M_i and M'_i are series-parallel or coloops. By Lemma 4.6, the connected components of $M \wedge M'$ are built out of the M_i by a sequence of parallel connections. The parallel connection of N to a coloop is N again, and the parallel connection of two series-parallel matroids is series-parallel (Proposition 5.3), so we see that every component of N is series-parallel as desired. $\square$

7. A polyhedral lemma. The goal of this section is to prove the following lemma, which we will need in section 8.

LEMMA 7.1. *Let P be an N -dimensional polyhedron. Let $P_1, P_2, \dots, P_D$ be the facets of a polyhedral subdivision $\mathcal{E}$ of P . Suppose that, for every subset K of $[D]$, the intersection $\bigcap_{k \in K} P_k$ is a nonempty polytope of dimension at least $N - |K| + 1$. Then the number of $(N - g + 1)$ -dimensional faces of $\mathcal{E}$ whose relative interiors lie in $\text{Int}(P)$ is $\binom{D}{g}$.*

We will call a subdivision $\mathcal{E}$ as in Lemma 7.1 a D -simplicial subdivision. (The motivation for this terminology is that the neighborhood of the unique face of dimension $N - D + 1$ is the normal fan to a simplex of dimension $D - 1$.) Figure 2 shows a 3-simplicial division of a hexagon.

Some geometric experimentation should convince the reader that this lemma is extremely plausible. We will use a spherical variant of Helly's Theorem, due to Molnár. Recall that a convex cone is a subset C of $\mathbb{R}^N$ which is closed under addition, under multiplication by positive scalars, and contains the origin. A pointed convex cone is a convex cone C such that 0 can not be written as a positive combination of nonzero elements of C .

THEOREM 7.2. *Let $C_1, C_2, \dots, C_r$ be a collection of pointed convex cones of $\mathbb{R}^N$, with $r \geq N + 1$. Suppose that the intersection of any N of these cones contains a point other than the origin, and that the union of any $N + 1$ of these cones is not all of $\mathbb{R}^N$. Then the intersection of all of these cones contains a point other than the origin.*

Proof. For a proof in the case $N = 3$, see [15]. There is no difficulty in adapting Molnár's proof to arbitrary dimensions. The reader may also want to consult the survey article [5], which states this result in section 9 and sketches a proof. $\square$

We now prove Lemma 7.1. Recall from section 1 the notations $T(P, Q)$ and $T_0(P, Q)$.

Proof of Lemma 7.1. Without loss of generality, we may assume that the ambient space in which P lives is $\mathbb{R}^N$.

We first establish the following claim: for every nonempty subset K of $[D]$, the intersection $\bigcap_{k \in K} P_k$ has dimension $N - |K| + 1$ and meets $\text{Int}(P)$. We prove this

claim by induction on the cardinality of a subset K . When K is a singleton $\{k\}$, this is obvious as P_k is assumed to be a facet in a subdivision of P . Now, consider some K of cardinality greater than 1, and set $Q = \bigcap_{k \in K} P_k$. Let K' be a subset of K which contains one element less than K does. Since $Q \subseteq \bigcap_{k \in K'} P_k$ and since $P_1, \dots, P_D$ are faces of a subdivision, we know that Q is a face of $\bigcap_{k \in K'} P_k$. Since, by induction, $\bigcap_{k \in K'} P_k$ has dimension $N - |K'| + 2$ and, by hypothesis, Q has dimension at least $N - |K| + 1$, we know that Q has dimension either $N - |K| + 2$ or $N - |K| + 1$. We consider each case in turn.

First, suppose that Q has dimension $N - |K| + 2$. Consider the set of pointed cones $T_0(P_k, Q)$. Note that these cones live in a vector space of dimension $N - (N - |K| + 2) = |K| - 2$. We apply Theorem 7.2 to the cones $T_0(P_k, Q)$. For any $(|K| - 2)$ -element subset L of K , we know that $\bigcap_{\ell \in L} P_\ell$ has dimension $N - |L| + 2$, which is greater than that of Q , so $\bigcap_{\ell \in L} T_0(P_\ell, Q) = T_0(\bigcap_{\ell \in L} P_\ell, Q)$ has positive dimension and is not just the origin. Also, for any $(|K| - 1)$ -element subset $K \setminus \{x\}$ of K , we know that $\bigcup_{k \in K \setminus \{x\}} T_0(P_k, Q)$ can not be all of $(|K| - 2)$ -space, as it does not contain any points in $\text{Int}(T_0(P_x, Q))$. Thus, the hypotheses of Theorem 7.2 apply, and we conclude that $\bigcap_{k \in K} T_0(P_k, Q)$ contains a point other than the origin. But $\bigcap_{k \in K} T_0(P_k, Q) = T_0(\bigcap_{k \in K} P_k, Q) = T_0(Q, Q)$, which is just the origin. This is a contradiction.

So, we deduce that Q has dimension $N - |K| + 1$. If we can show that Q meets $\text{Int}(P)$ we are done, so assume for the sake of contradiction that Q lies in the boundary of P . Then there is some half-space H containing Q whose boundary contains P . Again, we consider the cones $T(Q, P_k)$, which we now know to lie in a space of dimension $|K| - 1$. Let L be any $(|K| - 1)$ -element subset of K . Then, by induction, $\bigcap_{\ell \in L} P_\ell$ has dimension $N - |L| + 2$ and hence is not Q . So $\bigcap_{\ell \in L} T_0(P_\ell, Q) = T_0(\bigcap_{\ell \in L} P_\ell, Q)$ is not just the origin. Also, it is impossible that $\bigcup_{k \in K} T_0(P_k, Q)$ is all of $(|K| - 1)$ -space, as every cone $T_0(P_k, Q)$ lies in $H/\mathbb{R}Q$. So Theorem 7.2 applies again, and we deduce that $\bigcap_{k \in K} T_0(P_k, Q)$ contains a point other than the origin. As before, this yields a contradiction. Thus, we conclude that Q has dimension $N - |K| + 1$ and meets $\text{Int}(P)$, as desired. This completes the inductive proof of the claim.

We now complete the proof. Every face of $\mathcal{E}$ whose relative interior lies in $\text{Int}(P)$ is of the form $\bigcap_{k \in K} P_k$ for some subset K of $[D]$. Our claim shows that every such intersection does meet the relative interior of P . Suppose two of these intersections are equal, let's say, $\bigcap_{k \in K} P_k = \bigcap_{k \in K'} P_k$. Then both these intersections are equal to $\bigcap_{k \in K \cup K'} P_k$. So, by our claim, the dimension of this intersection is equal to $N - |K| + 1$, $N - |K'| + 1$ and $N - |K \cup K'| + 1$. So $|K| = |K'| = |K \cup K'|$, and we deduce that $K = K'$. We have thus shown that all of the intersections $\bigcap_{k \in K} P_k$ give distinct faces whose relative interiors lie in $\text{Int}(P)$. So the number of such faces of dimension $N - g + 1$ is the number of g -element subsets of $[D]$. $\square$

8. Perturbing nearly transverse intersections. Our remaining goal is to prove Theorem 1.2—every constructible space achieves the f -vector of the f -vector conjecture. Our strategy is as follows: A constructible space is made by successive dualization and transverse intersection. From Proposition 2.12, the bounded parts of L and $L^\perp$ are negatives of each other, so they have the same f -vector. Thus, if L obeys the f -vector conjecture, so does $L^\perp$. So our main goal is to understand the effect of transverse intersection.

In this section, we will prove (Theorem 8.1) that, if L and L' are series-parallel tropical linear spaces and $v \in \mathbb{R}^n$, then the f -vector of $L \cap (L' + v)$ is independent of v as long as L and $L' + v$ meet transversely. The most technical part of this proof is the analysis of perturbing L' slightly when L and L' are almost transverse; this is

carried out in Lemma 8.4. Our result ultimately relies on a previously proven lemma (Lemma 5.7) about chains of flats in series-parallel matroids. In the next section, we will use Theorem 8.1 to prove Theorem 1.2. First, let us set out our goal for this section.

THEOREM 8.1. *Let L and L' be two series-parallel tropical linear spaces in n -space, of dimensions d and d' . Let a and $b \in \mathbb{R}^n$ be such that L meets $L' + a$ and $L' + b$ transversely. Then $L \cap (L' + a)$ and $L \cap (L' + b)$ have the same bounded f -vector.*

We first must prove some technical lemmas describing what happens when we perturb an intersection of two tropical linear spaces. We need to introduce a fair amount of notation. Let L and L' be tropical linear spaces of dimensions d and d' in n -space, with $0 < d, d' < n$ and $d + d' > n$. Let $\mathcal{D}$, $\mathcal{D}'$, and $\mathcal{E}$ be the regular subdivisions of $\Delta(d, n)$, $\Delta(d', n)$, and $\Delta(d + d' - n, n)$ corresponding to L , L' , and $L \cap_{\text{st}} L'$. Let v be a generic vector in $\mathbb{R}^n$, and consider $L \cap (L' + \epsilon v)$ for a ϵ sufficiently small positive number. Let $\mathcal{E}_v$ be the corresponding subdivision of $\Delta(d + d' - n, n)$, so $\mathcal{E}_v$ refines $\mathcal{E}$. For every face A of $\mathcal{E}_v$, there is a unique face B of $\mathcal{E}$ such that $\text{Int}(B) \supseteq \text{Int}(A)$. We will say that B is the face which *internally contains* A . We will focus on a particular face $\text{Poly}_{M \wedge M'}$ of $\mathcal{E}$ and try to determine how this face is subdivided in $\mathcal{E}_v$.

Let w be a point of $L \cap_{\text{st}} L'$, and let $M^\vee$ and $(M')^\vee$ be the faces of L and L' in whose relative interior w lies. So w lies in the face $\text{Int}((M \wedge M')^\vee)$ of $L \cap_{\text{st}} L'$. If $\mathbb{R}M^\vee + \mathbb{R}(M')^\vee = \mathbb{R}^n$, i.e., if $M^\vee$ and $(M')^\vee$ are transverse, then $M^\vee$ and $(M')^\vee + \epsilon v$ also meet for sufficiently small ϵ , so $\text{Poly}_{M \wedge M'}$ is a face of $\mathcal{E}_v$ as well as $\mathcal{E}$.

We will spend most of the rest of this section in a detailed analysis of the case where $M^\vee$ and $(M')^\vee$ are not transverse, but are nearly so. More specifically, suppose that $\mathbb{R}M^\vee + \mathbb{R}(M')^\vee$ has dimension $n - 1$. By Lemma 4.10, this means that $\Gamma(M, M')$ has first betti number 1. Let $c_1, c_2, \dots, c_{2r}$ be the vertices of this cycle, with $c_1, c_3, c_5, \dots$ indexing connected components of M and $c_2, c_4, c_6, \dots$ indexing connected components of M' . Let e_s be the edge joining c_s to c_{s+1} .¹ For future use, we define $\sigma := \sum_{j=1}^{2r} (-1)^j v_{e_j}$. The sign of σ will control the combinatorics of $\mathcal{E}_v$. We will use these notations throughout the section.

Now that $M^\vee$ and $(M')^\vee$ are not transverse, the polytopes $M^\vee$ and $(M')^\vee + \epsilon v$ will not meet for generic v . The faces of $\mathcal{E}_v$ which $\text{Poly}_{M \wedge M'}$ internally contains will all arise as $\text{Poly}_{N \wedge N'}$ where $N^\vee$ and $(N')^\vee$ are faces of L and L' containing $M^\vee$ and $(M')^\vee$ such that $N^\vee$ and $(N')^\vee + \epsilon v$ do meet (for ϵ a sufficiently small positive number.) If $N^\vee$ and $(N')^\vee + \epsilon v$ meet,² then $\text{Poly}_{N \wedge N'}$ is a face of $\mathcal{E}_v$ contained in $\text{Poly}_{M \wedge M'}$, but whether or not $\text{Poly}_{N \wedge N'}$ is internally contained in $\text{Poly}_{M \wedge M'}$ can be subtle.

We pause to introduce more notation. Let $M = \bigoplus_{c \in C} M_c$ and $M' = \bigoplus_{c' \in C'} M'_{c'}$ be the decompositions of M and M' into connected components, so the vertex set of $\Gamma(M, M')$ is $C \sqcup C'$. We will write S_c and S'_c for the elements of M_c and M'_c . Occasionally, we will write $M_c^{(\cdot)}$ or $S_c^{(\cdot)}$ when we don't want to state whether c lies in C or in C' .

We first wish to consider those faces internally contained in $\text{Poly}_{M \wedge M'}$ which are of the same dimension as $\text{Poly}_{M \wedge M'}$. Such faces must be of the form $N \wedge N'$, where either $N = M$ and $(N')^\vee$ has dimension one larger than $(M')^\vee$ or else $N' = M'$ and $N^\vee$ has dimension one larger than $M^\vee$. In the first case, $N' = \bigoplus_{d \neq c} M'_d \oplus M'_c|_V \oplus M'_c/V$ for

¹The indices on the c_i and the e_i are cyclic modulo $2r$.

²We will now begin to omit the statement “for ϵ a sufficiently small positive number” for brevity’s sake when it should be clear.

some component $c \in C'$ and some facet V of M'_c . Similarly, in the second case, N is described by giving a component c in C and a facet V of M_c . So, in general, we can describe a potential pair (N, N') by giving a vertex c of $\Gamma(M, M')$ and a facet V of $M_c^{(\cdot)}$. With this terminology, we can describe when $N^\vee$ and $(N')^\vee + \epsilon v$ meet.

LEMMA 8.2. *Let M and M' obey the above hypotheses and continue the above notation. Let N and N' be as above, with corresponding c and V . If σ is positive, then $\text{Int}(N^\vee)$ and $\text{Int}((N')^\vee) + \epsilon v$ meet if and only if c is in $\{c_1, c_2, \dots, c_{2r}\}$, with V containing e_{s-1} but not e_s . If σ is negative, then $\text{Int}(N^\vee)$ and $\text{Int}((N')^\vee) + \epsilon v$ meet if and only if c is in $\{c_1, c_2, \dots, c_{2r}\}$, with V containing e_s but not e_{s-1} .*

We first need the following general description of perturbed intersections.

LEMMA 8.3. *Let P and P' be two polytopes in $\mathbb{R}^n$, with $w \in P \cap P'$. Let Q and Q' be the faces of P and P' in whose relative interiors w lies. Let v be a vector in $\mathbb{R}^n$, and let $\bar{v}$ be the image of v in $\mathbb{R}^n/(\mathbb{R}Q + \mathbb{R}Q')$. Then $\text{Int}(P)$ and $\text{Int}(P') + \epsilon v$ meet if and only if $\bar{v}$ is in $\text{Int}((T(P, Q) - T(P', Q'))/(\mathbb{R}Q + \mathbb{R}Q'))$.*

Proof. Note that $\text{Int}(P)$ and $\text{Int}(P')$ lie in $w + \text{Int}(T(P, Q))$ and $w + \text{Int}(T(P', Q'))$, respectively. So if $\text{Int}(P)$ and $\text{Int}(P') + \epsilon v$ meet, then $\text{Int}(T(P, Q))$ and $\text{Int}(T(P', Q')) + \epsilon v$ meet implying that ϵv lies in $\text{Int}(T(P, Q)) - \text{Int}(T(P', Q'))$. As $\text{Int}(T(P, Q))$ and $\text{Int}(T(P', Q'))$ are cones, this means that v lies in $\text{Int}(T(P, Q)) - \text{Int}(T(P', Q'))$. This is equivalent to stating that $\bar{v}$ is in $\text{Int}((T(P, Q) - T(P', Q'))/(\mathbb{R}Q + \mathbb{R}Q'))$.

Now for the converse direction. Suppose that $v = x - x'$, with x and x' in $\text{Int}(T(P, Q))$ and $\text{Int}(T(P', Q'))$. Then, for ϵ sufficiently small, the points $w + \epsilon x$ and $w + \epsilon x'$ are in $\text{Int}(P)$ and $\text{Int}(P')$, respectively. So $w + \epsilon x = (w + \epsilon x') + \epsilon v$ is in the relative interiors of both P and $P' + \epsilon v$. $\square$

We now apply this description to the situation of Lemma 8.2.

Proof of Lemma 8.2. By assumption, $\mathbb{R}M^\vee + \mathbb{R}(M')^\vee$ has dimension $n - 1$, so the vector space $\mathbb{R}^n/(\mathbb{R}M^\vee + \mathbb{R}(M')^\vee)$ is one-dimensional. Let λ be the linear function $(x_1, x_2, \dots, x_n) \mapsto \sum_{j=1}^{2r} (-1)^j x_{e_j}$. We claim that λ is zero on $(\mathbb{R}M^\vee + \mathbb{R}(M')^\vee)$. The vector space $\mathbb{R}M^\vee$ is spanned by the characteristic vectors of the connected components of M . For any connected component M_c of M other than M_{c_1} , M_{c_3} , $\dots$, $M_{c_{2r-1}}$, none of the e_j lie in M_c , so λ is zero on the characteristic vector of M_c . Similarly, $M_{c_{2k-1}}$ contains e_{2k-1} , e_{2k} and none of the other e_j lie in $M_{c_{2k-1}}$, so λ is $1 - 1 = 0$ on the characteristic vector of $M_{c_{2k-1}}$. Thus, λ is zero on $\mathbb{R}M^\vee$. Similarly, λ is zero on $\mathbb{R}(M')^\vee$ and thus on $(\mathbb{R}M^\vee + \mathbb{R}(M')^\vee)$. So we can identify $\mathbb{R}^n/(\mathbb{R}M^\vee + \mathbb{R}(M')^\vee)$, with $\mathbb{R}$ by mapping every point x in $\mathbb{R}^n$ to $\lambda(x)$. Under this identification, $\bar{v}$ is identified, with $\lambda(v) = \sigma$. By Lemma 8.3, we need to determine whether or not σ is in the relative interior of $\lambda(T(N^\vee, M^\vee) - T((N')^\vee, (M')^\vee))$. For brevity, we abbreviate $\lambda(T(N^\vee, M^\vee))$ and $\lambda(T((N')^\vee, (M')^\vee))$ to K and K' .

From now on, we will consider the case where σ is positive. The case where σ is negative is precisely analogous. So, we need to determine the images of $T(N^\vee, M^\vee)$ and $-T((N')^\vee, (M')^\vee)$ under λ . Let us first consider the case where $c = c_{2j+1}$. Then $N' = M'$, while $T_0(N^\vee, M^\vee)$ is a ray spanned by the image in $\mathbb{R}^n/\mathbb{R}M^\vee$ of $\sum_{f \in V} e_f$. So K' is $\{0\}$, while K is spanned by $\mu(V) := \lambda(\sum_{f \in V} e_f)$. So σ is in $K - K'$ if and only if $\mu(V)$, like σ , is positive. This happens if and only if e_{2j} is in V and e_{2j+1} is not. Next, let us consider the case where $c = c_{2j}$. Then K is $\{0\}$, and K' is spanned by $\mu(V)$. So we see that, this time, σ is in $K - K'$ if and only if $\mu(V)$ is negative. This happens only if e_{2j-1} is in V and e_{2j} is not. In both cases, when $c = c_s$, we have that σ is in $K - K'$ if and only if e_{s-1} is in V and e_s is not. Finally, we need to consider the case where c is not one of $c_1, c_2, \dots, c_{2r}$. In this case, $\mu(V) = 0$, and both K and K' are $\{0\}$. So, in this case, σ is not in $K - K'$. $\square$

Recall from section 5 the notation $d(f, g, M)$ for the number of the facets of M containing f and not g . So we see that, when σ is positive, there are $\sum_{j=1}^{2r} d(e_{j-1}, e_j, M_{c_j}^{(\prime)})$ faces of $\mathcal{E}_v$, which are internally contained in and of the same dimension as $\text{Poly}_{M \wedge M'}$. Similarly, when σ is negative, there are $\sum_{j=1}^{2r} d(e_j, e_{j-1}, M_{c_j}^{(\prime)})$ such faces. In particular, assume that L and L' are series-parallel, so that each $M_i^{(\prime)}$ is series-parallel (by Lemma 6.1). Then, by Proposition 5.7, $d(e_{j-1}, e_j, M_{c_j}^{(\prime)}) = d(e_j, e_{j-1}, M_{c_j}^{(\prime)})$ for each j from 1 to $2r$. So $\sum_{j=1}^{2r} d(e_{j-1}, e_j, M_{c_j}^{(\prime)}) = \sum_{j=1}^{2r} d(e_j, e_{j-1}, M_{c_j}^{(\prime)})$, we define D to be the common value of these sums.

PROPOSITION 8.4. *Let L and L' be series-parallel linear spaces and continue the preceding assumptions and notation. For generic v , the subdivision of $\text{Poly}_{M \wedge M'}$ in $\mathcal{E}_v$ is D -simplicial.*

Proof. Since v is generic, σ is not zero. Our goal is to apply Lemma 7.1. Let $P_1, P_2, \dots, P_D$ be the maximal faces of the subdivision of $\text{Poly}_{M \wedge M'}$, with $P_k = \text{Poly}_{N(k) \wedge N'(k)}$. Our goal is to show that $\bigcap_{k \in K} P_k$ has dimension at least $\dim \text{Poly}_{M \wedge M'} - |K| + 1$. We first introduce some notation. Look at a particular index k . Suppose that $N'(k) = M'$ and $N = \bigoplus_{i \neq c_{2j+1}} M_i \oplus M_{c_{2j+1}}|_V \oplus M_{c_{2j+1}}/V$. In this case, we define $c(k) = c_{2j+1}$ and $V(k) = V$. Similarly, if $N(k) = M$ and $N' = \bigoplus_{i \neq c_{2j}} M_i \oplus M_{c_{2j}}|_V \oplus M_{c_{2j}}/V$, then we set $c(k) = c_{2j}$ and $V(k) = V$. Finally, we introduce the notation p for the number of connected components of $\Gamma(M, M')$, so $\text{Poly}_{M \wedge M'}$ has dimension $n - p$ (by Lemma 4.7).

Fix a particular nonempty set K of indices. Look at a particular component $M_c^{(\prime)}$ of $M^{(\prime)}$. (Throughout this paragraph, $(\prime)$ will either always denote a prime or always denote the absence of a prime.) Let $K_c = \{k \in K : c(k) = k\}$. Then, by Proposition 5.7, the set of facets $\{V(k)\}_{k \in K_c}$ forms a nested chain in $M_c^{(\prime)}$. Add the $\emptyset$ and $S_c^{(\prime)}$ to the ends of this chain to form a filtration on M_c . Define $N_c^{(\prime)}$ to be the associated graded matroid of $M_c^{(\prime)}$ with respect to this filtration. Set $\mathcal{N}^{(\prime)}(K) = \bigoplus_{c \in C^{(\prime)}} N_c^{(\prime)}$.

Consider the graph $\Gamma(\mathcal{N}(K), \mathcal{N}'(K))$. This graph is formed by taking $\Gamma(M, M')$ and separating each vertex c_j in the cycle of Γ into $|K_c| + 1$ vertices. So $\Gamma(\mathcal{N}(K), \mathcal{N}'(K))$ has $\sum_{c \in C \cup C'} |K_c| = |K|$ more vertices than $\Gamma(M, M')$ and has the same number of edges. Also, $\Gamma(\mathcal{N}(K), \mathcal{N}'(K))$ is a forest. (At least one K_{c_j} is nonempty and, for this j , the edges e_{j-1} and e_j of $\Gamma(\mathcal{N}(K), \mathcal{N}'(K))$ are connected to different vertices, splitting open the cycle of $\Gamma(M, M')$.) So $\Gamma(\mathcal{N}(K), \mathcal{N}'(K))$ is a forest, with $p + |K| - 1$ connected components. By Lemma 4.6, we know that $\mathcal{N}(K) \wedge \mathcal{N}'(K)$ has rank $d + d' - n$ and that $\text{Poly}_{\mathcal{N}(K) \wedge \mathcal{N}'(K)}$ has dimension $n - (p + |K| - 1) = \dim \text{Poly}_{M \wedge M'} - |K| + 1$.

For every k in K , every basis of $\mathcal{N}^{(\prime)}(K)$ is a basis of $N'(k)$. So $\text{Poly}_{\mathcal{N}^{(\prime)}(K)} \subseteq \text{Poly}_{N'(k)}$ and $\text{Poly}_{\mathcal{N}(K) \wedge \mathcal{N}'(K)} \subseteq \text{Poly}_{N(k) \wedge N'(k)} = P_k$. (We are using that we know $\mathcal{N}(K) \wedge \mathcal{N}'(K)$ has rank $d + d' - n$.) We deduce that $\text{Poly}_{\mathcal{N}(K) \wedge \mathcal{N}'(K)} \subseteq \bigcap_{k \in K} P_k$. So $\bigcap_{k \in K} P_k$ contains a polytope of dimension $\dim \text{Poly}_{M \wedge M'} - |K| + 1$. We have now checked the conditions of Lemma 7.1, and it follows that $\mathcal{E}_v$ is D -simplicial. $\square$

We now prove Theorem 8.1.

Proof of Theorem 8.1. Let $U_i \subseteq \mathbb{R}^n$ be the set of v such that, for any two faces F and F' of L and L' , either the relative interiors of F and $F' + v$ are disjoint or they span an affine linear space of dimension at least $n - i$. Thus, L and $L' + v$ meet transversely if and only if v is in U_0 and $\mathbb{R}^n \setminus U_i$ is a polyhedral complex of dimension at most $n - i - 1$. In particular, we see that U_1 is path connected. We may therefore join a and b by a path through U_1 . It is thus enough to show that if $w \in U_1 \setminus U_0$

and we perturb w into U_0 , then the f -vector we get is independent of the choice of perturbation. With loss of generality, we may take $w = 0$. Let $\mathcal{E}$ be the regular subdivision of $\Delta(d + d' - n, n)$ corresponding to $L \cap_{\text{st}} L'$, and let $\mathcal{E}_{\epsilon v}$ be the regular subdivision corresponding to $L \cap_{\text{st}} (L' + \epsilon v)$. Also, write $\mathcal{D}$ and $\mathcal{D}'$ for the subdivisions of $\Delta(d, n)$ and $\Delta(d', n)$. For ϵ small enough, $\mathcal{E}_{\epsilon v}$ refines $\mathcal{E}$. Let's focus on a face R of $\mathcal{E}$ and ask what faces of $\mathcal{E}_{\epsilon v}$ are internally contained in R .

The face R is of the form $\text{Poly}_{N \wedge N'}$ for Poly_N and $\text{Poly}_{N'}$ faces of $\mathcal{D}$ and $\mathcal{D}'$. If $N^\vee$ and $(N')^\vee$ are transverse, then $N^\vee$ and $(N')^\vee + \epsilon v$ will also meet, so R is also a face of $\mathcal{E}_{\epsilon v}$. If $N^\vee$ and $(N')^\vee$ are not transverse, then Lemma 8.4 applies. So there is an integer D so that, in $\mathcal{E}_{\epsilon v}$, there are $\binom{D}{i}$ faces of $\mathcal{E}_{\epsilon v}$ which have dimension $\dim R - i + 1$ and are internally contained in R . In particular, this number does not depend on the choice of v . Since each internal face of $\mathcal{E}_{\epsilon v}$ is internally contained in exactly one internal face of $\mathcal{E}$, this shows that the number of internal faces of $\mathcal{E}_{\epsilon v}$ is independent of v . $\square$

9. Proof of Theorem 1.2. We now are ready to begin our final assault on Theorem 1.2: that all constructible tropical linear spaces have the f -vector predicted by the f -vector conjecture. Our basic strategy is to use Theorem 8.1 to reduce to a particularly nice case. Our proof involves several binomial coefficient identities.

We will first need to study how the various faces of $\Delta(d, n)$ are subdivided in a constructible decomposition. Let S and T be disjoint subsets of $[n]$. Write $\Delta(d, n) \setminus S/T$ for the face of $\Delta(d, n)$ given by the equations $x_i = 0$ for $i \in S$ and $x_i = 1$ for $i \in T$. If p is a tropical Plücker vector, then define $p \setminus S/T$ to be the function on the vertices of $\Delta(d - |T|, n - |S| - |T|)$ given by $(p \setminus S/T)_I = p_{I \cup T}$. It is clear that $p \setminus S/T$ obeys the tropical Plücker relations (equation (1)), as they are a subset of the tropical Plücker relations for p . So $p \setminus S/T$ corresponds to a tropical $d - |T|$ plane in $n - |S| - |T|$ space. If $L = L(p)$, then we write $L \setminus S/T$ for the tropical linear space $L(p \setminus S/T)$.

LEMMA 9.1. *If L is a constructible d -plane in n -space and S and T are disjoint subsets of $[n]$, then $L \setminus S/T$ is also constructible.*

Proof. If L is a hyperplane, then every $L \setminus S/T$ is either a hyperplane or the whole space in which it sits and hence is constructible. Now suppose that the lemma holds for L . Then it holds for $L^\perp$ as $L^\perp \setminus S/T = (L \setminus T/S)^\perp$. Suppose that the lemma holds for L and L' and that L and L' meet transversely. Then it holds for $L \cap L' = L \cap_{\text{st}} L'$ as $(L \cap_{\text{st}} L') \setminus S/T = (L \setminus \emptyset / (S \cup T)) \cap_{\text{st}} (L' \setminus S/T)$. $\square$

We now begin to prove Theorem 1.2. Our proof is by induction, first on n and then on the number of steps used to construct L . We set

$$f_{i,d,n} = \frac{(n-i-1)!}{(i-1)!(d-i)!(n-d-i)!}.$$

First, suppose that the theorem is true for L . The bounded f -vectors of $L^\perp$ and L are the same by Proposition 2.12, so the theorem for $L^\perp$ follows from the symmetry $f_{i,d,n} = f_{i,n-d,n}$. For the rest of the proof, suppose that L and L'' are a constructible d -plane and d' plane in n -space which meet transversely and for which we know the theorem to hold. In addition, we know that the theorem holds for every $L \setminus S/T$ and $L'' \setminus S/T$ by our inductive hypotheses. Our aim will be to prove the theorem for $L \cap L''$.

By Theorem 8.1, it is enough to prove the result for $L \cap (L'' + w)$, where w is chosen generically enough. We will take a w for which $w_1 \gg w_2 \gg \cdots \gg w_n$ and which is otherwise generic. We will call such a w “large enough.” Set $L' = L'' + w$. Every face of $L \cap L'$ can be uniquely written as $(M^\vee) \cap ((M')^\vee)$ for some faces $M^\vee$ and $(M')^\vee$ which meet transversely. Our strategy will be to use the assumption

that w is large enough to figure out which pairs of faces meet in a bounded face of $L \cap L'$. We will need to consider the possibility that a bounded face of $L \cap L'$ is the intersection of unbounded faces of L and of L' ; to handle this situation we use the $L \setminus S/T$ construction to reduce enumerating unbounded faces of L (subject to certain restrictions) to enumerating bounded faces of tropical linear spaces for smaller n . By Lemma 9.1, all of these smaller tropical linear spaces will also be constructible, so our inductive hypotheses will apply.

Consider any face $M^\vee$ of L . We know M is loop-free; let T be the set of coloops of M so that Poly_M lies in the relative interior of $\Delta(d, n) \setminus \emptyset/T$. Then $M^\vee/(1, \dots, 1)$ is the product of $\mathbb{R}_{\geq 0}^T$ and a bounded polytope. Let $(M'')^\vee$ be a face of L'' and define T' similarly. We write $(M')^\vee = (M'')^\vee + w$, so M' and M'' are the same matroid.

Define an ordered pair (T, T') of the subsets of $[n]$ to be *facing* if, for all $1 \leq i < j \leq n$, either $i \in T$ or $j \in T'$.

LEMMA 9.2. *For $w_1 \gg w_2 \gg \dots \gg w_n$, the faces $M^\vee$ and $(M')^\vee$ meet if and only if (T, T') is facing. The intersection $M^\vee \cap (M')^\vee$ is bounded if and only if $T \cap T' = \emptyset$.*

Proof. First, suppose that (T, T') is not facing, so there exists some $i < j$, with $i \notin T$ and $j \notin T'$. On $M^\vee$, the function $x_i - x_j$ is bounded above (as $i \notin T$). On $(M'')^\vee$, $x_i - x_j$ is similarly bounded below. The value of $x_i - x_j$ on $(M')^\vee$ is $w_i - w_j$ larger than on $(M'')^\vee$ so if $w_j - w_i$ is large enough, then the values of $x_i - x_j$ on $M^\vee$ and $(M')^\vee$ will be disjoint. Thus, $M^\vee$ and $(M')^\vee$ will not meet.

Now suppose that (T, T') is facing. Let $(x_i) \in M^\vee$ and $(x'_i) \in (M'')^\vee$, set $x'_i = x''_i + w_i$, and set $d_i = x'_i - x_i$. For w large enough, we have $d_1 > d_2 > \dots > d_n$. Let b be the smallest element of $[n] \setminus T$, or let $b = n$ if $[n] = T$. Then, since (T, T') is facing, $j \in T'$ for every $j > b$ and, by construction, i is in T for every $i < b$. After translating (x_i) by $(1, \dots, 1)$, which will not change the assumption that $(x_i) \in M^\vee$, we may assume that $x_b = x'_b$, so $d_b = 0$. Consider the point

$$\begin{aligned} z &:= (x_1 + d_1, \dots, x_{b-1} + d_{b-1}, x_b, x_{b+1}, \dots, x_n) \\ &= (x'_1, \dots, x'_{b-1}, x'_b, x'_{b+1} - d_{b+1}, \dots, x'_n - d_n). \end{aligned}$$

Clearly, $z \in (x_i) + \mathbb{R}_{\geq 0}^T$ and $z \in (x'_i) + \mathbb{R}_{\geq 0}^{T'}$, so $z \in M^\vee \cap (M')^\vee$. We have now proved the first claim.

Now assume that (T, T') is facing, so there is some point $z = (z_i) \in M^\vee \cap (M')^\vee$. Suppose that $i \in T \cap T'$. Then $z + ue_i \in M^\vee \cap (M')^\vee$ for every $u > 0$ and $M^\vee \cap (M')^\vee$ is unbounded.

Suppose, on the other hand, that $T \cap T' = \emptyset$. Consider any distinct elements i and j , with $i < j$ in $[n]$. If neither i nor j is in T , then $x_i - x_j$ is bounded on $M^\vee$; if neither i nor j is in T' , then $x_i - x_j$ is bounded on $(M')^\vee$; if $i \notin T'$ and $j \notin T$, then $x_i - x_j$ is bounded below on $M^\vee$ and above on $(M')^\vee$. Given that (T, T') is facing and $T \cap T' = \emptyset$, it must be true that one of these three cases holds. We see that, in every case, $x_i - x_j$ is bounded on $M^\vee \cap (M')^\vee$. So all of the $x_i - x_j$ are bounded on $M^\vee \cap (M')^\vee$, which implies that $(M^\vee \cap (M')^\vee)/(1, \dots, 1)$ is bounded. $\square$

Keeping the above notation, $M^\vee$ corresponds to a bounded face of $L \setminus \emptyset/T$. If this bounded face is j -dimensional, then $M^\vee$ is $(j + t)$ -dimensional, where $t = |T|$. Similarly, let $t' = |T'|$, and let j' be the dimension of the face of $L \setminus \emptyset/T'$ corresponding to $(M')^\vee$, then the dimension of $(M')^\vee$ is $j' + t'$. The dimension of $M^\vee \cap (M')^\vee$ is then $j + j' + t + t' - n$. We thus see that the number of i -dimensional bounded faces of $L \cap L'$ is

$$\sum_{\substack{(T, T') \text{ facing} \\ T \cap T' = \emptyset}} \sum_{j+j'+t+t'-n=i} f_{j,d-t,n-t} f_{j',d'-t',n-t'}.$$

We have used our inductive assumption to know that all of the constructible tropical linear spaces $L \setminus \emptyset / T$ and $L' \setminus \emptyset / T'$ have the correct f -vector as they live either in a smaller ambient space, or if T or T' is empty, are constructed using a smaller number of steps.

The only facing pairs (T, T') , with $T \cap T' = \emptyset$ are $(\{1, \dots, b-1\}, \{b+1, \dots, n\})$ and $(\{1, \dots, b\}, \{b+1, \dots, n\})$. So we can write our sum as

$$\sum_b \left(\sum_{j+j'=i+1} f_{j,d-b+1,n-b+1} f_{j',d'-n+b,b} + \sum_{j+j'=i} f_{j,d-b,n-b} f_{j',d'-n+b,b} \right).$$

We will now show that this sum is $f_{i,D,n}$, where D is $d + d' - n$.

We work with the first sum first. Plugging in $f_{i,d,n} = \binom{n-d-1}{i-1} \binom{n-i-1}{n-d-1}$ and exchanging the order of summation, we have

$$\sum_{j+j'=i+1} \binom{n-d-1}{j-1} \binom{n-d'-1}{j'-1} \left(\sum_b \binom{n-b-j}{n-d-1} \binom{b-j'-1}{n-d'-1} \right).$$

Using the identity $\sum_{k+k'=p} \binom{k}{q} \binom{k'}{r} = \binom{p+1}{q+r+1}$, (see [14, equation (5.26)]) the inner sum evaluates to $\binom{n-i-1}{2n-d-d'-1} = \binom{n-i-1}{n-D-1}$. We are reduced to evaluating the sum

$$\binom{n-i-1}{n-D-1} \sum_{j+j'=i+1} \binom{n-d-1}{j-1} \binom{n-d'-1}{j'-1}.$$

The remaining sum can be evaluated by Vandermonde's convolution (see [14, equation (5.22)]) $\sum_{j+j'=r} \binom{p}{j} \binom{q}{j'} = \binom{p+q}{r}$. We get $\binom{n-i-1}{D-i} \binom{n-D-2}{i-1}$.

A similar argument shows that the second sum is $\binom{n-i-1}{D-i} \binom{n-D-2}{i-2}$. Adding these together, we get

$$\binom{n-i-1}{D-i} \left[\binom{n-D-2}{i-1} + \binom{n-D-2}{i-2} \right] = \binom{n-i-1}{D-i} \binom{n-D-1}{i-1} = f_{i,D,n}.$$

This concludes the proof of Theorem 1.2.

Acknowledgments. Much of this material originally appeared in my dissertation, [18], and I am grateful to my advisor Bernd Sturmfels for his guidance on this project. In addition, I have benefitted from conversations with many mathematicians; foremost among them Federico Ardila, Mike Develin, Michael Joswig, Frank Sottile, and Josephine Yu. I was supported by a research fellowship from the Clay Mathematics Institute for much of the writing of this paper.

REFERENCES

- [1] F. ARDILA, A. FINK, AND F. RINCÓN, *Valuations for matroid polytope subdivisions*, Canadian Mathematical Bulletin, to appear.
- [2] F. ARDILA AND C. KLIVANS, *The Bergman complex of a matroid and phylogenetic trees*, J. Combin. Theory Ser. B, 96 (2006), pp. 38–49.
- [3] T. BRYLAWSKI, *A combinatorial model for series-parallel networks*, Trans. AMS, 154 (1971), pp. 1–22.
- [4] H. CRAPO, *A higher invariant for matroids*, J. Combin. Theory, 2 (1967), pp. 406–417.
- [5] L. DANZER, B. GRÜNBAUM, AND V. KLEE, *Helly's theorem and its relatives*, in Proceedings of the Symposium on Pure Math., Vol. VII, Amer. Math. Soc., Providence, RI, 1963, pp. 101–180.

- [6] M. DEVELIN AND B. STURMFELS, *Tropical convexity*, Doc. Math., 9 (2004), pp. 1–27.
- [7] A. DRESS, *Duality theory for finite and infinite matroids with coefficients*, Adv. Math., 59 (1986), pp. 97–123.
- [8] A. DRESS AND W. TERHALE, *The tree of life and other affine buildings*, Doc. Math., 3 (1998), pp. 565–574.
- [9] A. DRESS AND W. WENZEL, *Valuated matroids*, Adv. Math., 93 (1992), pp. 214–250.
- [10] A. DRESS AND W. WENZEL, *Grassmann-Plücker relations and matroids with coefficients*, Adv. Math., 86 (1991), pp. 68–110.
- [11] J. EDMONDS, *Submodular functions, matroids, and certain polyhedra*, in Combinatorial Structures and their Applications, R. Guy et al., eds., Gordon and Breach, New York, 1970, pp. 69–87.
- [12] W. FULTON AND B. STURMFELS, *Intersection theory on toric varieties*, Topology, 36 (1997), pp. 335–353.
- [13] I. GELFAND, I. GORESKY, R. MACPHERSON, AND V. SERGANOVA, *Combinatorial geometries, convex polyhedra, and Schubert cells*, Adv. Math., 63 (1987), pp. 301–316.
- [14] R. GRAHAM, D. KNUTH, AND O. PATASHNIK, *Concrete Mathematics*, 2nd ed., Addison-Wesley Publishing Company, Reading, MA, 1994.
- [15] J. MOLNÁR, *Über eine verallgemeinerung auf die Kugelfläche eines topologischen Satzes von Helly* Acta Math. Acad. Sci. Hungar., 7 (1956), pp. 107–108.
- [16] J. RICHTER-GEBERT, B. STURMFELS, AND T. THEOBALD, *First steps in tropical geometry*, in Idempotent Mathematics and Mathematical Physics, Contemp. Math. 377, Amer. Math. Soc., Providence, RI, 2005, pp. 289–317.
- [17] D. SPEYER, *A matroid invariant via the K-theory of the Grassmannian*, preprint, MIT, Cambridge, MA, 2006 arXiv:math.AG/0603551.
- [18] D. SPEYER, *Tropical Geometry*, Ph.D. Dissertation, University of California at Berkeley, Berkeley, CA, 2005.
- [19] D. SPEYER AND B. STURMFELS, *Tropical Mathematics*, based on Clay Mathematics Institute Senior Scholar Lecture given at Park City, Utah, 2004 by second author arXiv:math.CO/0408099.
- [20] D. SPEYER AND B. STURMFELS, *The Tropical Grassmannian*, Adv. Geom., 4 (2004), pp. 389–411.
- [21] B. STURMFELS, *Solving systems of polynomial equations*, CBMS Reg. Conf. Ser. Math. 97, Amer. Math. Soc., Providence, RI, 2002.
- [22] N. WHITE (ed.), *Theory of Matroids—Encyclopedia of Mathematics and its Applications Vol. 26*, Cambridge University Press, London, 1986.
- [23] N. WHITE (ed.), *Matroid Applications—Encyclopedia of Mathematics and its Applications Vol. 40*, Cambridge University Press, London, 1992.
- [24] G. ZIEGLER, *Lectures on polytopes*, Grad. Texts in Math. 152, Springer-Verlag, New York, NY, 1995.

BOUNDS FOR THE REAL NUMBER GRAPH LABELLINGS AND APPLICATION TO LABELLINGS OF THE TRIANGULAR LATTICE*

DANIEL KRÁL'[†] AND PETR ŠKODA[‡]

Abstract. We establish new lower and upper bounds for the real number graph labelling problem. As an application, we consider a problem of Griggs to determine the optimum spans of $L(p, q)$ -labellings of the infinite triangular plane lattice and find (using a computer) the optimum spans for all p and q .

Key words. real number graph labelling, channel assignment problem, graph labelling with distance conditions

AMS subject classification. 05C15

DOI. 10.1137/07068552X

1. Introduction. One of the important applications of graph colorings lies in the area of the channel assignment problem. The significance of results on the channel assignment problem has increased with development of mobile phone networks—large activity in this area yielding numerous results is documented in several recent surveys: an excellent survey [1] is focused on applications, a broad overview of the mathematical models can be found in [5], and some more specific results can be found in [19].

One of the well-established models for the channel assignment problem is the distance constrained graph labelling, introduced by Griggs and Yeh [12]. This notion can be treated in a more general setting of real number graph labellings and λ -graphs introduced in a series of papers of Griggs and Jin [8, 9, 10] and further developed in [2, 15]. The reader can check a recent survey [13] for more details.

Throughout this paper, we will use the concept of λ -graphs. A λ -graph G is a graph with k types of edges; two vertices of G could be joined by several edges of different types. Given the parameters $x_1, \dots, x_k$, a labelling f of vertices of G by nonnegative real numbers is called *proper* if the labels of every two vertices u and v joined by an edge of type i differ by at least x_i . The *span* of the labelling f is the supremum of the labels of all vertices (note at this point that it can always be assumed that the infimum of the used labels is equal to zero). The infimum of the spans of proper labellings of G is denoted by $\lambda(G; x_1, \dots, x_k)$. The values of $\lambda(G; x_1, \dots, x_k)$ interpreted as a k -parameter function form the λ -function of G .

Note that throughout the article we often assume implicitly that for every choice of parameters $x_1, \dots, x_k$ the value of $\lambda(G; x_1, \dots, x_k)$ is finite. This is equivalent to the statement that G viewed as an ordinary graph can be colored by a finite number of colors.

*Received by the editors March 17, 2007; accepted for publication (in revised form) June 9, 2008; published electronically October 17, 2008.

<http://www.siam.org/journals/sidma/22-4/68552.html>

[†]Faculty of Mathematics and Physics, Institute for Theoretical Computer Science (ITI), Charles University, Malostranské náměstí 25, 118 00 Prague, Czech Republic (kral@kam.mff.cuni.cz). The Institute for Theoretical Computer Science is supported by the Ministry of Education of the Czech Republic as project 1M0545.

[‡]Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Malostranské náměstí 25, 118 00 Prague, Czech Republic (peskoj@kam.mff.cuni.cz).

Let us now recall the notion of *distance constrained labellings*. For a graph G and nonnegative integers $p_1, \dots, p_k$, an $L(p_1, \dots, p_k)$ -labelling of G is a labelling of the vertices of G with nonnegative integers such that vertices at distance i , $1 \leq i \leq k$, receive labels that differ by at least p_i . This notion can be relaxed from integer labellings to real number graph labellings yielding a special class of λ -graphs: the corresponding λ -graph is the graph with the same vertex set where the vertices at distance i are connected by an edge of type x_i .

The most studied case both in the settings of λ -graphs and distance constrained labellings is $k = 2$. Since the so-called scaling property [8, 13] asserts that

$$\alpha\lambda(G; x_1, \dots, x_k) = \lambda(G; \alpha x_1, \dots, \alpha x_k)$$

for all $\alpha \geq 0$, the λ -function could be viewed as a one parameter function $\lambda(G; x, 1)$ if $k = 2$. In section 3, we prove several lemmas that allow us to derive (matching) upper and lower bounds on $\lambda(G; x, 1)$ from the knowledge of the values of $\lambda(G; x, 1)$ for specific values of x . To the best of our knowledge, these bounds are the first general bounds that can be used to deduce upper and lower bounds on $\lambda(G; x, 1)$ besides a direct application of the scaling property (see [8, 9] for an example of such an application). As an application of our bounds, we were able to find optimal $L(p, q)$ -labellings of the infinite triangular lattice, which is a known open problem for which we now give more details.

In practical applications, infinite graphs often provide a convenient model for the underlying topology of the network: instead of considering a finite graph, one can model a network as an infinite regular tiling of the plane. Hence, the infinite triangular lattice, the infinite square lattice, and the infinite hexagonal lattice naturally appear in such applications. The optimum spans of $L(p, q)$ -labellings of the infinite square and hexagonal lattices have been determined by Griggs and Jin [9]; cf. Calamoneri [4]. The case of the infinite triangular lattice seems to be more difficult.

The problem of determining the values of the λ -function for infinite triangular lattice Γ_Δ was posed by Griggs [11] in the 2000 International Math Contest in Modeling (MCM). Five teams proved new results for particular choices of parameters [3, 6, 7, 17, 18]. In particular, Goodwin, Johnston, and Marcus [7] determined the value of $\lambda(\Gamma_\Delta; x, 1)$ for $x \geq 4$. Jin and Yeh [14], Zhu and Shi [20], and Calamoneri [5] studied further the values of $\lambda(\Gamma_\Delta; x, 1)$ for $x \geq 1$. Finally, Griggs and Jin [9] managed to determine the values of $\lambda(\Gamma_\Delta; x, 1)$ for all $x \notin (1/3, 4/5)$.

As we have already said, we apply our bounds to complete the behavior of the function $\lambda(\Gamma_\Delta; x, 1)$ —the reader can find the whole function $\lambda(\Gamma_\Delta; x, 1)$ in Figure 1. Our argument is computer assisted, which is also the case of some earlier bounds on $\lambda(\Gamma_\Delta; x, 1)$; see, e.g., [9]. Of course, it would be interesting to determine the function $\lambda(\Gamma_\Delta; x, 1)$ avoiding the use of computers.

2. Preliminary results. In this section, we recall basic results on λ -functions mainly from [2, 8, 15].

The results of Griggs and Jin [8] and Král' [15] characterize the general behavior of the λ -function that could be summarized in the following theorem.

THEOREM 2.1. *Let G be a (possibly infinite) λ -graph with k types of edges. The λ -function of G is a continuous, nondecreasing, and piecewise linear function of $x_1, \dots, x_k$ on $[0, \infty)$ with finitely many linear parts.*

Additionally, the values of the λ -function are restricted to a subset $D(x_1, \dots, x_p)$

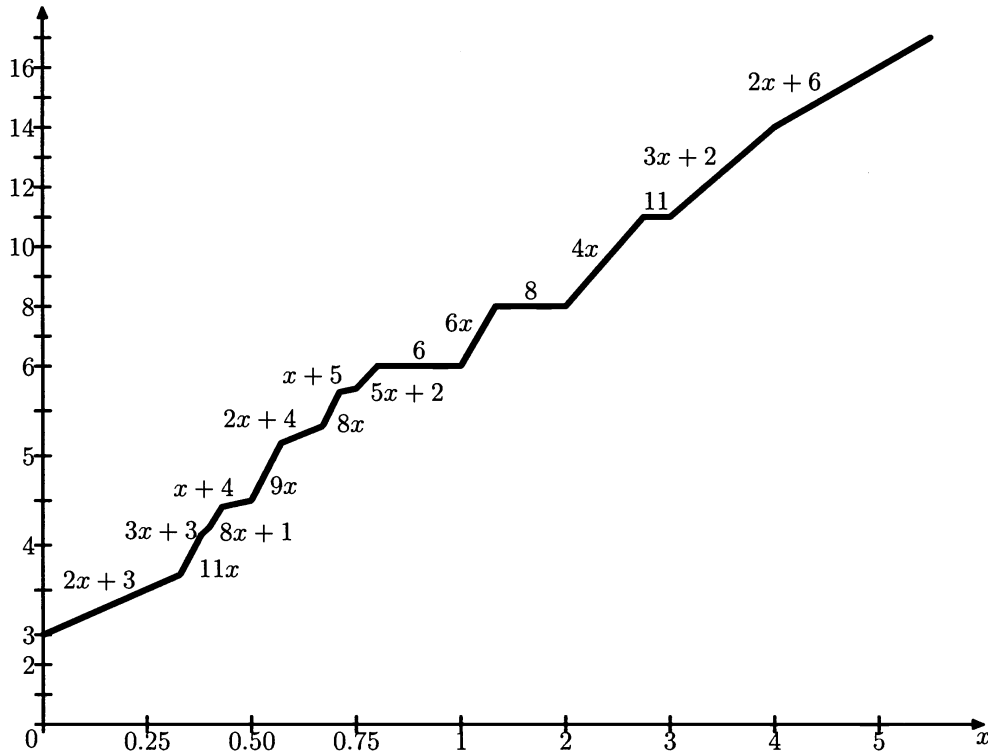


FIG. 1. The optimal value of $\lambda(\Gamma_\Delta; x, 1)$. Note that some parts of the graph are enlarged to reflect the behavior $\lambda(\Gamma_\Delta; x, 1)$ in more detail.

of real numbers defined as follows: $D(x_1, \dots, x_p)$ is a set of all linear combinations $\sum_{i=1}^p \alpha_i x_i$, where α_i are nonnegative integers. A theorem of Griggs and Jin [8] states that there exists a finite optimal labelling that uses only labels from set $D(x_1, \dots, x_p)$.

THEOREM 2.2 (Griggs and Jin [8]). *Let G be a λ -graph, possibly infinite, with k types of edges. For all nonnegative real numbers $x_1, \dots, x_k$, there is an optimal labelling f for the parameters $x_1, \dots, x_k$ such that $f(v) \in D(x_1, \dots, x_k)$ for each vertex v of G . In particular, $\lambda(G; x_1, \dots, x_k)$ belongs to the set $D(x_1, \dots, x_k)$.*

As we explained in section 1, the infinite coverings of the plane play a special role in the channel assignment problem. By the compactness principle, every λ -graph G contains a finite subgraph with the same λ -function.

THEOREM 2.3 (Král' [16]). *Every λ -graph G with k types of edges and a finite chromatic number contains a finite subgraph H such that $\lambda(G; x_1, \dots, x_k) = \lambda(H; x_1, \dots, x_k)$ for all $(x_1, \dots, x_k) \in [0, \infty)^k$.*

Note that it is easy to derive the existence of a subgraph H for a fixed k -tuple $(x_1, \dots, x_k)$ such that $\lambda(G; x_1, \dots, x_k) = \lambda(H; x_1, \dots, x_k)$, but this fact itself does not imply the existence of a universal subgraph H for all choices of $x_1, \dots, x_k$, which is claimed in Theorem 2.3.

Griggs and Jin [9] studied the function $\lambda(G; x, y)$ for G being a λ -graph obtained from the infinite triangular lattice Γ_Δ by connecting neighbors with edges of the first type and vertices at distance exactly two with edges of the second type. For the sake of simplicity, only the edges of the first type will be drawn in our figures. Griggs and Jin [9] determined $\lambda(\Gamma_\Delta; x, 1)$ for $x \in [0, 1/3] \cup [4/5, \infty)$.

THEOREM 2.4 (Griggs and Jin [9]). *For $k \geq 0$ the minimum span of any $L(x, 1)$ -labelling of triangular lattice is given by*

$$\lambda(\Gamma_{\Delta}; x, 1) \begin{cases} = 2x + 3 & \text{if } 0 \leq x \leq 1/3, \\ \in [2x + 3, 11x] & \text{if } 1/3 \leq x \leq 9/22, \\ \in [2x + 3, \frac{9}{2}] & \text{if } 9/22 \leq x \leq 3/7, \\ \in [9x, \frac{9}{2}] & \text{if } 3/7 \leq x \leq 1/2, \\ \in [\frac{9}{2}, \frac{16}{3}] & \text{if } 1/2 \leq x \leq 2/3, \\ \in [\frac{16}{3}, \frac{23}{4}] & \text{if } 2/3 \leq x \leq 3/4, \\ \in [\frac{23}{4}, 6] & \text{if } 3/4 \leq x \leq 4/5, \\ = 6 & \text{if } 4/5 \leq x \leq 1, \\ = 6x & \text{if } 1 \leq x \leq 4/3, \\ = 8 & \text{if } 4/3 \leq x \leq 2, \\ = 4x & \text{if } 2 \leq x \leq 11/4, \\ = 11 & \text{if } 11/4 \leq x \leq 3, \\ = 3x + 2 & \text{if } 3 \leq x \leq 4, \\ = 2x + 6 & \text{if } x \geq 4. \end{cases}$$

We complete the behavior of the function $\lambda(\Gamma_{\Delta}; x, 1)$ for $x \in [1/3, 4/5]$.

3. General bounds. In this section, we prove several bounds on $\lambda(G; x, 1)$ of the following type: given the values of $\lambda(G; x, 1)$ at two points x_1 and x_2 , we estimate the values of $\lambda(G; x, 1)$ in the whole interval (x_1, x_2) .

Since the λ -function $\lambda(G; x, 1)$ is a piecewise linear function of x , it is easy to complete the function inside the interval if there is only one acceptable linear function connecting the two known points of the function. One such case is described in the next lemma ($\mathbb{Z}^+$ stands for the set of nonnegative integers).

LEMMA 3.1. *Let G be a λ -graph and assume that $\lambda(G; x_i, 1) = \alpha + \beta x_i$ for $x_1, x_2 \in [0, \infty)$, $x_1 < x_2$, and $i = 1, 2$. If $\beta' \geq \beta$ whenever $\alpha + \beta x = \alpha' + \beta' x$ for $\alpha', \beta' \in \mathbb{Z}^+$ and $x \in [x_1, x_2]$, then $\lambda(G; x, 1) = \alpha + \beta x$ for all $x \in [x_1, x_2]$.*

Proof. By Theorem 2.1, the function $\lambda(G; x, 1)$ is a continuous piecewise linear function of x . It follows from Theorem 2.2 that the value of $\lambda(G; x, 1)$ is equal to an element of $D(x, 1)$.

Suppose that $\lambda(G; x, 1)$ is strictly greater than $\alpha + \beta x_0$ for some $x_0 \in (x_1, x_2)$. Since the function $\lambda(G; x, 1)$ is equal to $\alpha + \beta x$ for x_2 , there exists a linear function $\alpha' + \beta' x$ and $y_1, y_2 \in [x_1, x_2]$, $y_1 < y_2$, such that

- $\lambda(G; x, 1) = \alpha' + \beta' x$ for $x \in [y_1, y_2]$,
- $\alpha + \beta y_1 < \alpha' + \beta' y_1$, and
- $\alpha + \beta y_2 = \alpha' + \beta' y_2$.

Hence, $\beta'(y_2 - y_1)$ is strictly smaller than $\beta(y_2 - y_1)$, and thus $\beta' < \beta$, which is excluded by the assumptions of the lemma. We conclude that there is no $x_0 \in (x_1, x_2)$ such that $\lambda(G; x_0, 1) > \alpha + \beta x_0$.

An analogous argument yields that $\lambda(G; x_0, 1) < \alpha + \beta x_0$ for no $x_0 \in (x_1, x_2)$. We infer that $\lambda(G; x, 1)$ is equal to $\alpha + \beta x$ for all $x \in [x_1, x_2]$. $\square$

There is a variant of Lemma 3.1 where all the linear functions have a smaller slope than the segment connecting the two known points. The proof is analogous to that of Lemma 3.1.

LEMMA 3.2. *Let G be a λ -graph and assume that $\lambda(G; x_i, 1) = \alpha + \beta x_i$ for $x_1, x_2 \in [0, \infty)$, $x_1 < x_2$, and $i = 1, 2$. If $\beta' \leq \beta$ whenever $\alpha + \beta x = \alpha' + \beta' x$ for $\alpha', \beta' \in \mathbb{Z}^+$ and $x \in [x_1, x_2]$, then $\lambda(G; x, 1) = \alpha + \beta x$ for all $x \in [x_1, x_2]$.*

For every linear function $f(x) = \alpha + \beta x$, $\alpha, \beta \in \mathbb{Z}^+$, there exist intervals I where no linear function with a nonnegative integer offset and a nonnegative integer slope smaller than β is equal to $f(x)$ for any $x \in I$. We utilize this fact in our further considerations. For an interval $I \subseteq [0, \infty)$, $\xi(I)$ denotes the smallest positive integer q such that $p/q \in I$ for $p \in \mathbb{Z}$; i.e., $\xi(I)$ is the smallest denominator of a fraction with integral numerator and denominator contained in I .

PROPOSITION 3.3. *Let $I \subseteq [0, \infty)$ be an interval and α, β, α' , and β' nonnegative integers. If $\beta' < \beta < \xi(I)$, then $\alpha + \beta x \neq \alpha' + \beta' x$ for every $x \in I$.*

Proof. Assume the contrary; i.e., there exist nonnegative integers α, β, α' , and β' such that $\alpha + \beta x_0 = \alpha' + \beta' x_0$ for $x_0 \in I$ and $\beta' < \beta < \xi(I)$. A simple computation yields that

$$x_0 = -\frac{\alpha - \alpha'}{\beta - \beta'} = \frac{|\alpha - \alpha'|}{|\beta - \beta'|} \in I.$$

Hence, x_0 is a rational number with the denominator $|\beta - \beta'| \leq \beta$, which is impossible since $\beta < \xi(I)$. $\square$

Proposition 3.3 immediately yields the next simple lemma.

LEMMA 3.4. *Let $\lambda(G; x, 1) = \alpha + \beta x$ for $x = x_1, x_2 \in [0, \infty)$, $x_1 < x_2$. If $\xi([x_1, x_2]) > \beta$, then $\lambda(G; x, 1) = \alpha + \beta x$ for all $x \in [x_1, x_2]$.*

Proof. By Proposition 3.3, there is no linear function with nonnegative integer coefficients and a slope smaller than β that is equal to $\alpha + \beta x$ for $x \in [x_1, x_2]$. Therefore, the statement follows from the Lemma 3.1. $\square$

Lemma 3.1 handles the case when the interval contains no function with a smaller slope. The counterpart, Lemma 3.2, can be applied, too. It is enough to assume that the left endpoint of the interval is sufficiently large, as stated in the next lemma.

LEMMA 3.5. *Let $\lambda(G; x, 1) = \alpha + \beta x$ for $x = x_1, x_2 \in [0, \infty)$, $x_1 < x_2$. If x_1 is greater than $\frac{\alpha}{\xi([x_1, x_2])}$, then $\lambda(G; x, 1) = \alpha + \beta x$ for all $x \in [x_1, x_2]$.*

Proof. We want to apply Lemma 3.2. Assume that there is a linear function $\alpha' + \beta' x$, $\alpha', \beta' \in \mathbb{Z}^+$, that is equal to $\alpha + \beta x$ at $x_0 \in [x_1, x_2]$ and such that the slope β' is greater than β . A simple computation yields that

$$x_0 = -\frac{\alpha' - \alpha}{\beta' - \beta} = \frac{|\alpha' - \alpha|}{|\beta' - \beta|}.$$

Since $|\beta' - \beta| \geq \xi([x_1, x_2])$, it holds that

$$x_0 \leq \frac{|\alpha - \alpha'|}{\xi([x_1, x_2])} \leq \frac{\alpha}{\xi([x_1, x_2])} < x_1.$$

This implies that $x_0 < x_1$, contrary to the choice of x_0 . Therefore, all linear functions $\alpha' + \beta' x$ equal to $\alpha + \beta x$ for some $x \in [x_1, x_2]$ have a slope smaller than or equal to β . The statement of the lemma now follows from Lemma 3.2. $\square$

Consider an optimal labelling f of a λ -graph G for $x_1 = x_0$ and $x_2 = 1$. As we have explained in section 2, we can assume that $f(v) \in D(x_0, 1)$ for every vertex v . It is often the case that the labelling defined by considering the same integer combinations that appears in f is optimal for some $x_1 < x_0$ (and $x_2 = 1$). In the next lemma, we find an estimate of the interval to which the labelling f can be extended.

LEMMA 3.6. Let x_0 be a point equal to p_0/q_0 for some positive integers p_0 and q_0 that are relatively prime. If $\lambda(G; x_0, 1) = \alpha + \beta x_0$ and $\beta < q_0$, then

$$\lambda(G; x, 1) \leq \alpha + \beta x \quad \text{for all } x \in \left[\frac{\tilde{p}}{\tilde{q}}, x_0 \right],$$

where $\tilde{p}$ and $\tilde{q}$ are integers such that $\frac{\tilde{p}}{\tilde{q}} < \frac{p_0}{q_0}$, $1 \leq \tilde{q} \leq q_0$, and $\frac{\tilde{p}}{\tilde{q}}$ is maximal among all fractions formed by such a pair of integers $\tilde{p}$ and $\tilde{q}$.

Proof. Let $\tilde{x} = \tilde{p}/\tilde{q}$. By the choice of $\tilde{p}$ and $\tilde{q}$, $\xi((\tilde{x}, x_0])$ is equal to q_0 . By Proposition 3.3, there is no linear function with nonnegative integer coefficients and slope β' smaller than β that is equal to $\alpha + \beta x$ for $x \in (\tilde{x}, x_0]$.

Let x' be the smallest number such that $\lambda(G; x, 1) \leq \alpha + \beta x$ for all $x \in [x', x_0]$. Note that x' could be equal to x_0 . If $x' \leq \tilde{x}$, the statement of the lemma is true. Assume that $x' > \tilde{x}$. By the choice of x' , $\lambda(G; x, 1) = \alpha' + \beta' x$ for some integers α' and $\beta' < \beta$ (otherwise, the function $\lambda(G; x, 1)$ would not cross the function $f(x) = \alpha + \beta x$ at $x = x'$, and x' could be chosen smaller). Consequently, the choice of x' and the continuity of the function $\lambda(G; x, 1)$ imply that $\alpha + \beta x' = \alpha' + \beta' x'$, which is excluded by Proposition 3.3, as explained in the first paragraph. $\square$

In the rest of the paper, we will apply the bounds obtained in this section to complete the behavior of the function $\lambda(\Gamma_\Delta; x, 1)$.

4. Upper bounds for triangular lattice. The simplest way of obtaining an upper bound is a construction of a labelling. Let us start with introducing some additional notation related to the description of the lattice Γ_Δ . First, choose an arbitrary vertex of the lattice Γ_Δ to be the origin. The lattice contains two naturally defined axes (see Figure 2). This gives every vertex unique coordinates, as shown in the figure.

Our construction of labellings of Γ_Δ is similar to the technique of matrix labelling used in [9]. However, in addition, the patterns used to label the lattice can be “shifted” in the way described further.

A matrix M of size $n \times m$ specifies the labelling of nm vertices of Γ_Δ . We say that M is *placed* at coordinates $[a, b]$ if a vertex with coordinates $[a + i, b + j]$, $i \in \{0, \dots, m - 1\}$, $j \in \{0, \dots, n - 1\}$, is labelled with $M_{n-j, i+1}$. The labelling is produced with “horizontal offset” o if the matrix M is placed at all coordinates $[c \cdot m, d \cdot n + c \cdot o]$ for all integers c and d . The labelling constructed could also be created by the original matrix labelling technique, but our method allows using smaller matrices to describe the labelling.

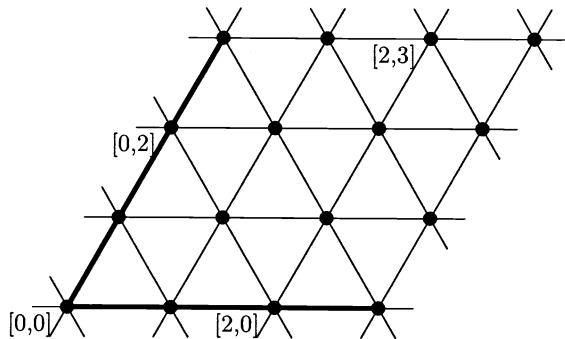


FIG. 2. The infinite triangular lattice Γ_Δ with an origin and two axes.

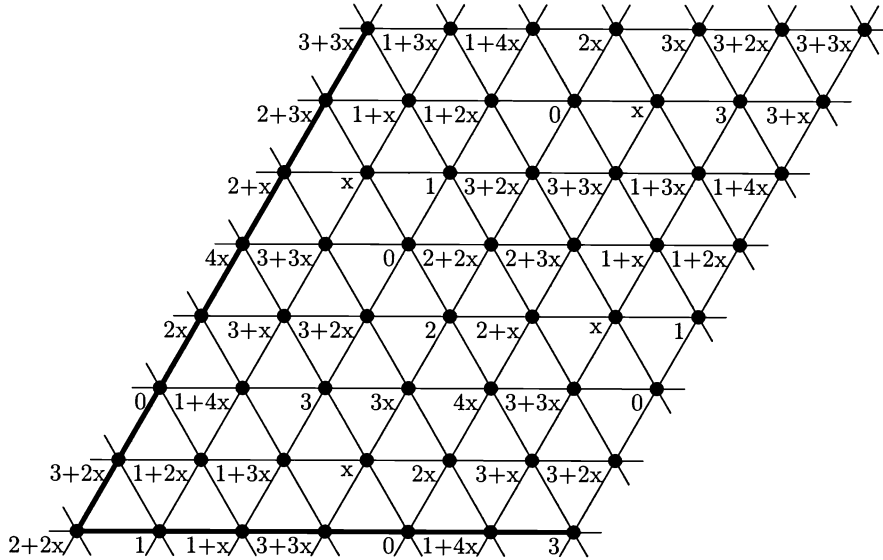


FIG. 3. A part of the labelling of the infinite triangular lattice.

The next lemma uses shifted matrix labelling to construct proper labellings of Γ_Δ for $x \in [3/8, 2/5]$.

LEMMA 4.1. *The value of $\lambda(\Gamma_\Delta; x, 1)$ for $x \in [3/8, 2/5]$ is at most $3 + 3x$.*

Proof. The labelling of the infinite triangular lattice Γ_Δ will be obtained through the matrix labelling technique with horizontal offsets.

The matrix M we use is the following:

$$\begin{bmatrix} 2 & 0 & x & 2+3x \\ 3x & 3+2x & 3+3x & 2+x \\ x & 3 & 3+x & 4x \\ 3+3x & 1+3x & 1+4x & 2x \\ 2+3x & 1+x & 1+2x & 0 \\ 2+x & x & 1 & 3+2x \\ 4x & 3+3x & 0 & 2+2x \\ 2x & 3+x & 3+2x & 2 \\ 0 & 1+4x & 3 & 3x \\ 3+2x & 1+2x & 1+3x & x \\ 2+2x & 1 & 1+x & 3+3x \end{bmatrix}.$$

The horizontal offset of the matrix labelling is 9 (see Figure 3 for a part of the labelling of Γ_Δ obtained). The reader is kindly asked to verify that the labelling produced assigns neighboring vertices labels that differ by at least x and those at distance two labels that differ by at least one. Since the span of the labelling of Γ_Δ constructed is equal to $3 + 3x$, the statement of the lemma follows. $\square$

5. Lower bounds. Our lower bound proofs are partly computer assisted. By the compactness principle, if Γ_Δ does not have a labelling with a certain span, then there exists a finite subgraph of Γ_Δ that does not have a labelling of that span either. We have independently prepared two programs that test all labellings of finite parts of the

triangular lattice Γ_Δ by brute force to reveal specific values of x ; one of the programs can be downloaded at <http://kam.mff.cuni.cz/~kral/lattice.c>. Both programs use the same coordinate system as we define in the proof of Lemma 4.1. Finally, for an integer n , $\Gamma_{\Delta,n}$ denotes the subgraph of Γ_Δ induced by the vertices with coordinates $[x, y]$ from set $\{0 \leq x \leq n, 0 \leq y \leq n\}$.

In order to decrease the running time, our program does not try to assign all values of the set $D(x, 1)$ to the points of the lattice, but it assigns only labels contained in a set $\tilde{D}_m(x) \subset D(x, 1)$. The set $\tilde{D}_m(x)$ is defined recursively. We start with $\tilde{D}_m^0(x)$, which is equal to the set of labels contained in $D(x, 1)$ smaller than or equal to m . If $\tilde{D}_m^i(x)$ contains two numbers a and b , $a < b$, such that for every label $c \in \tilde{D}_m^i(x)$ it holds that

- if $c \geq a + x$, then $c \geq b + x$, and
- if $c \geq a + 1$, then $c \geq b + 1$,

then we set $\tilde{D}_m^{i+1}(x) = \tilde{D}_m^i(x) \setminus \{a\}$. We proceed in this way while $\tilde{D}_m^i(x)$ contains a pair a and b of such labels. The final set $\tilde{D}_m^i(x)$ is denoted by $\tilde{D}_m(x)$.

The input of our programs is the size of the lattice n , the value of x , and the maximal label m . The programs return whether $\Gamma_{\Delta,n}$ can be labelled by $\tilde{D}_m(x)$. However, it remains to verify that restricting the set of possible labels to $\tilde{D}_m(x)$ does not affect the correctness of our algorithm. We do so in the next lemma.

LEMMA 5.1. *Let G be a λ -graph. If G has a proper labelling for $x_1 = x$ and $x_2 = 1$ of span at most m , then it has a proper labelling of span at most m that uses only the labels contained in the set $\tilde{D}_m(x)$.*

Proof. Assume that G has a proper labelling of span at most m . By Theorem 2.2, there also exists a proper labelling f that uses only the labels contained in the set $D(x, 1)$.

We now prove by induction on i that for every $i \geq 0$ there exists a proper labelling f_i that uses only labels from $\tilde{D}_m^i(x)$. Consider the pair of numbers a and b such that a was removed from $\tilde{D}_m^i(x)$ at the i th step. We claim that changing the labels of all the vertices labelled by f_i with a to b yields a proper labelling of G . Let f_{i+1} be the resulting labelling. If f_{i+1} is not proper, then there exist a vertex v such that $f_i(v) = a$ (and thus $f_{i+1}(v) = b$) and a vertex u adjacent to v such that the labels $f_{i+1}(v)$ and $f_{i+1}(u)$ are too close. Let $c = f_i(u) = f_{i+1}(u)$.

If $c < a$, then the difference between the labels of u and v increases after relabelling. Hence, we can assume $c > a$ since $c \neq a$ by the constraints on the labelling. If u and v are joined by an x_1 -edge, then $c \geq a + x$ (recall that $x_1 = x$). Consequently, by the choice of a and b , $c \geq b + x$. An analogous argument applies when u and v are joined by an x_2 -edge. Hence, we have inferred from the fact that f_i is a proper labelling of G that f_{i+1} is also a proper labelling of G .

We eventually conclude that there is a proper labelling of G that uses only the labels contained in $\tilde{D}_m(x)$. $\square$

The following four lemmas have been established using our computer-assisted technique.

LEMMA 5.2. *The optimal span $\lambda(\Gamma_\Delta; x, 1)$ of the infinite triangular lattice Γ_Δ is at least $4 + x$ for $x = 3/7$.*

Proof. We have verified by a computer that $\lambda(\Gamma_{\Delta,9}; x, 1) > 3 + 3x$ for $x = 3/7$; i.e., $\Gamma_{\Delta,9}$ has no labellings of span $30/7$ for $x = 3/7$. Therefore, $\lambda(\Gamma_\Delta; 3/7, 1) > 30/7$. By Theorem 2.2, the value of $\lambda(\Gamma_\Delta; 3/7, 1)$ is contained in the set $D(3/7, 1)$. The smallest element of $D(3/7, 1)$ greater than $30/7$ is $4 + 3/7$. This yields the statement of the lemma. $\square$

LEMMA 5.3. *The optimal span $\lambda(\Gamma_\Delta; x, 1)$ of the infinite triangular lattice Γ_Δ is at least $3 + 3x$ for $x = 3/8$ and $x = 2/5$.*

Proof. We proceed similarly as in Lemma 5.2. Again, we have verified using a computer that $\lambda(\Gamma_{\Delta,10}; x, 1) > 4$ for $x = 3/8$. Hence, $\lambda(\Gamma_\Delta; 3/8, 1) > 4$. Since the optimal span is a nondecreasing function of x , it holds that $\lambda(\Gamma_\Delta; 2/5, 1) > 4$. Again, by Theorem 2.2, $\lambda(\Gamma_\Delta; 3/8, 1) \in D(3/8, 1)$ and $\lambda(\Gamma_\Delta; 2/5, 1) \in D(2/5, 1)$. The smallest elements greater than 4 are $3 + 3 \cdot 3/8$ for $x = 3/8$ and $3 + 3 \cdot 2/5$ for $x = 2/5$. We conclude that $\lambda(\Gamma_\Delta; x, 1) \geq 3 + 3x$ for $x = 3/8$ and $x = 2/5$. $\square$

LEMMA 5.4. *The optimal span $\lambda(\Gamma_\Delta; x, 1)$ of the infinite triangular lattice Γ_Δ is at least $4 + 2x$ for $x = 4/7$ and $x = 3/5$.*

Proof. Using a computer, we have found out that $\lambda(\Gamma_{\Delta,13}; x, 1) > 5$ for $x = 4/7$. Therefore, $\lambda(\Gamma_\Delta; 4/7, 1) > 5$. It follows from the fact that the optimal span is a nondecreasing function that also satisfies $\lambda(\Gamma_\Delta; 3/5, 1) > 5$. By Theorem 2.2, the value of $\lambda(\Gamma_\Delta; 4/7, 1)$ is contained in the set $D(4/7, 1)$. The smallest element greater than 5 is $4 + 2x$ for both $x = 4/7$ and $x = 3/5$. The assertion of the lemma follows. $\square$

LEMMA 5.5. *The optimal span $\lambda(\Gamma_\Delta; x, 1)$ of the infinite triangular lattice Γ_Δ is at least $5 + x$ for $x = 5/7$.*

Proof. Again, we have verified that $\lambda(\Gamma_{\Delta,15}; x, 1) > 2 + 5x$ for $x = 5/7$. Therefore, the value of $\lambda(\Gamma_\Delta; 5/7, 1)$ is greater than $5 + 4/7$. By Theorem 2.2, the value of $\lambda(\Gamma_\Delta; 5/7, 1)$ is contained in $D(5/7, 1)$. The smallest element of $D(5/7, 1)$ greater than $5 + 4/7$ is $5 + 5/7 = 5 + x$, which yields the lemma. $\square$

6. Optimal labellings of triangular lattice. In this section we prove our main theorem, filling in the values of $\lambda(\Gamma_\Delta; x, 1)$ for $x \in [1/3, 4/5]$.

THEOREM 6.1. *The following values are the spans of optimal labelling of the infinite triangular lattice Γ_Δ :*

$$\lambda(\Gamma_\Delta; x, 1) = \begin{cases} 3 + 2x & \text{if } 0 \leq x \leq 1/3, \\ 11x & \text{if } 1/3 \leq x \leq 3/8, \\ 3 + 3x & \text{if } 3/8 \leq x \leq 2/5, \\ 1 + 8x & \text{if } 2/5 \leq x \leq 3/7, \\ 4 + x & \text{if } 3/7 \leq x \leq 1/2, \\ 9x & \text{if } 1/2 \leq x \leq 4/7, \\ 4 + 2x & \text{if } 4/7 \leq x \leq 2/3, \\ 8x & \text{if } 2/3 \leq x \leq 5/7, \\ 5 + x & \text{if } 5/7 \leq x \leq 3/4, \\ 2 + 5x & \text{if } 3/4 \leq x \leq 4/5, \\ 6 & \text{if } 4/5 \leq x \leq 1, \\ 6x & \text{if } 1 \leq x \leq 4/3, \\ 8 & \text{if } 4/3 \leq x \leq 2, \\ 4x & \text{if } 2 \leq x \leq 11/4, \\ 11 & \text{if } 11/4 \leq x \leq 3, \\ 3x + 2 & \text{if } 3 \leq x \leq 4, \\ 2x + 6 & \text{if } x \geq 4. \end{cases}$$

Proof. The values of $\lambda(\Gamma_\Delta; x, 1)$ for $x \in [0, 1/3] \cup [4/5, \infty)$ were determined by Griggs and Jin [9] (see Theorem 2.4). For $x \in [1/3, 4/5]$, we split the proof into

several parts dealing with subintervals of $[1/3, 4/5]$ separately. We do not consider the subintervals of $[1/3, 4/5]$ from left to right since some of the values we compute are used in our further considerations for neighboring subintervals.

- The interval $[3/8, 2/5]$.
It follows from Lemmas 5.3 and 4.1 that the value of $\lambda(\Gamma_\Delta; x, 1)$ is equal to $3 + 3x$ for $x = 3/8$ and $x = 2/5$. Since $\xi([3/8, 2/5]) = 5$, Lemma 3.4 implies that the value of $\lambda(\Gamma_\Delta; x, 1)$ is equal to $3 + 3x$ for all $x \in [3/8, 2/5]$.
- The interval $[1/3, 3/8]$.
By Theorem 2.4, the value of $\lambda(\Gamma_\Delta; 1/3, 1)$ is equal to $3 + 2/3$. Therefore, $\lambda(\Gamma_\Delta; x, 1)$ is equal to $11x$ for $x = 1/3$ and $x = 3/8$. Since $1/3$ is greater than 0, Lemma 3.5 asserts the value of $\lambda(\Gamma_\Delta; x, 1)$ being equal to $11x$ for all $x \in [1/3, 3/8]$.
- The interval $[3/7, 1/2]$.
By Theorem 2.4, $\lambda(\Gamma_\Delta; x, 1) = 4 + x$ for $x = 1/2$. It follows from Lemma 3.6 that the value of $\lambda(\Gamma_\Delta; x, 1)$ is less than or equal to $4 + x$ for $x \in [0, 1/2]$. On the other hand, by Lemma 5.2, the value of $\lambda(\Gamma_\Delta; x, 1)$ is greater than or equal to $4 + x$ for $x = 3/7$. Therefore, $\lambda(\Gamma_\Delta; 3/7, 1) = 4 + 3/7$. Since $\xi([3/7, 1/2]) = 2$, Lemma 3.4 yields that $\lambda(\Gamma_\Delta; x, 1) = 4 + x$ for all $x \in [3/7, 1/2]$.
- The interval $[2/5, 3/7]$.
We have shown above that $\lambda(\Gamma_\Delta; x, 1)$ is equal to $1 + 8x$ for $x = 2/5$ and $x = 3/7$. Since $\xi([2/5, 3/7]) = 5$ and $2/5$ is greater than $1/5$, Lemma 3.5 implies that the value of $\lambda(\Gamma_\Delta; x, 1)$ is equal to $1 + 8x$ for all $x \in [2/5, 3/7]$.
- The interval $[4/7, 2/3]$.
By Theorem 2.4, the value of $\lambda(\Gamma_\Delta; x, 1)$ is equal to $4 + 2x$ for $x = 2/3$. We infer from Lemma 3.6 that the value of $\lambda(\Gamma_\Delta; x, 1)$ is less than or equal to $4 + 2x$ for $x \in [1/2, 2/3]$. On the other hand, by Lemma 5.4, the value of $\lambda(\Gamma_\Delta; x, 1)$ is greater than or equal to $4 + 2x$ for $x = 4/7$. Therefore, $\lambda(\Gamma_\Delta; 4/7, 1)$ is equal to $4 + 2 \cdot 4/7$. Since $\xi([4/7, 2/3]) = 3$, we derive from Lemma 3.4 that $\lambda(\Gamma_\Delta; x, 1) = 4 + 2x$ for all $x \in [4/7, 2/3]$.
- The interval $[1/2, 4/7]$.
We have established that $\lambda(\Gamma_\Delta; x, 1) = 9x$ for $x = 1/2$ and $x = 4/7$. Since $1/2$ is greater than 0, Lemma 3.5 asserts that the value of $\lambda(\Gamma_\Delta; x, 1)$ is equal to $9x$ for all $x \in [1/2, 4/7]$.
- The interval $[5/7, 3/4]$.
By Theorem 2.4, $\lambda(\Gamma_\Delta; x, 1)$ is equal to $5 + x$ for $x = 3/4$. We infer from Lemma 3.6 that $\lambda(\Gamma_\Delta; x, 1)$ is less than or equal to $5 + x$ for $x \in [2/3, 3/4]$. On the other hand, by Lemma 5.5, the value of $\lambda(\Gamma_\Delta; x, 1)$ is greater than or equal to $5 + x$ for $x = 5/7$. Therefore, $\lambda(\Gamma_\Delta; 5/7, 1)$ equals $5 + 5/7$. Since $\xi([5/7, 3/4]) = 4$, Lemma 3.4 implies that $\lambda(\Gamma_\Delta; x, 1)$ is equal to $5 + x$ for all $x \in [5/7, 3/4]$.
- The interval $[2/3, 5/7]$.
We have shown that $\lambda(\Gamma_\Delta; x, 1)$ is equal to $8x$ for $x = 2/3$ and $x = 5/7$. Since $2/3$ is greater than 0, Lemma 3.5 yields that the value of $\lambda(\Gamma_\Delta; x, 1)$ is equal to $8x$ for all $x \in [2/3, 5/7]$.
- The interval $[3/4, 4/5]$.
By Theorem 2.4, $\lambda(\Gamma_\Delta; x, 1)$ is equal to $2 + 5x$ for $x = 3/4$ and $x = 4/5$. Since $\xi([3/4, 4/5]) = 4$ and $3/4$ is greater than $2/4$, we infer from Lemma 3.5 that the value of $\lambda(\Gamma_\Delta; x, 1)$ is equal to $2 + 5x$ for all $x \in [3/4, 4/5]$.

Since we have analyzed all subintervals of $[1/3, 4/5]$, the proof of the theorem is finished. $\square$

REFERENCES

- [1] K. I. AARDAL, S. P. M. VAN HOESEL, A. M. C. A. KOSTER, C. MANNINO, AND A. SASSANO, *Models and solution techniques for frequency assignment problems*, Ann. Oper. Res., 153 (2007), pp. 79–129.
- [2] R. BABILON, V. JELÍNEK, D. KRÁL', AND P. VALTR, *Labelings of graphs with fixed and variable edge-weights*, SIAM J. Discrete Math, 21 (2007), pp. 688–706.
- [3] R. E. BROADHURST, W. J. SHANAHAN, AND M. D. STEFFEN, *We're sorry you're outside the coverage area*, UMAP J., 21 (2000), pp. 327–342.
- [4] T. CALAMONERI, *Optimal $L(h, k)$ -labelling of regular grids*, Discrete Math. Theor. Comput. Sci., 8 (2006), pp. 141–158.
- [5] T. CALAMONERI, *The $L(h, k)$ -labelling problem: A survey and annotated bibliography*, Comput. J., 49 (2006), pp. 585–608.
- [6] D. J. DURAND, J. M. KLINE, AND K. M. WOODS, *Groovin' with the big band(width)*, UMAP J., 21 (2000), pp. 357–368.
- [7] J. GOODWIN, D. JOHNSTON, AND A. MARCUS, *Radio channel assignments*, UMAP J., 21 (2000), pp. 369–378.
- [8] J. R. GRIGGS AND X. T. JIN, *Real number graphs labellings with distance conditions*, SIAM J. Discrete Math., 20 (2006), pp. 302–327.
- [9] J. R. GRIGGS AND X. T. JIN, *Real number channel assignments for lattices*, SIAM J. Discrete Math., 22 (2008), pp. 996–1021.
- [10] J. R. GRIGGS AND X. T. JIN, *Real number graph labellings for paths and cycles*, Internet Math., to appear.
- [11] J. R. GRIGGS, *Author/Judge's commentary: The outstanding channel assignment problems papers*, UMAP J., 21 (2000), pp. 379–386.
- [12] J. R. GRIGGS AND R. K. YEH, *Labelling graphs with a condition at distance 2*, SIAM J. Discrete Math., 5 (1992), pp. 586–595.
- [13] J. R. GRIGGS AND D. KRÁL', *Graph labellings with variable weights, a survey*, Discrete Appl. Math., to appear.
- [14] X. T. JIN AND R. K. YEH, *Graph distance-dependent labelling related to code assignment in computer networks*, Naval Res. Logist., 52 (2005), pp. 159–164.
- [15] D. KRÁL', *The channel assignment problem with variable weights*, SIAM J. Discrete Math., 20 (2006), pp. 690–704.
- [16] D. KRÁL', *A Note on the Compactness Principle for the Channel Assignment Problem with Variable Weights*, in preparation.
- [17] J. MINTZ, A. NEWCOMER, AND J. C. PRICE, *Channel assignment model: The span without a face*, UMAP J., 21 (2000), pp. 311–326.
- [18] R. M. WHITAKER, S. HURLEY, AND S. M. ALLEN, *Optimizing channel assignments for private mobile radio networks in the UHF 2 band*, Wireless Networks, 8 (2002), pp. 587–595.
- [19] R. K. YEH, *A survey on labeling graphs with a condition at distance two*, Discrete Math., 306 (2006), pp. 1217–1231.
- [20] D. ZHU AND A. SHI, *Optimal Channel Assignments*, preprint, 2001.

ORTHOGONAL DRAWINGS OF SERIES-PARALLEL GRAPHS WITH MINIMUM BENDS*

XIAO ZHOU[†] AND TAKAO NISHIZEKI[†]

Abstract. In an orthogonal drawing of a planar graph G , each vertex is drawn as a point, each edge is drawn as a sequence of alternate horizontal and vertical line segments, and any two edges do not cross except at their common end. A bend is a point where an edge changes its direction. A drawing of G is called an optimal orthogonal drawing if the number of bends is minimum among all orthogonal drawings of G . In this paper we give an algorithm to find an optimal orthogonal drawing of any given series-parallel graph of the maximum degree at most three. Our algorithm takes linear time, while the previously known best algorithm takes cubic time. Furthermore, our algorithm is much simpler than the previous one. We also obtain a best possible upper bound on the number of bends in an optimal drawing.

Key words. orthogonal drawing, bend, series-parallel graph, planar graph

AMS subject classifications. 05C85, 05C10

DOI. 10.1137/060667621

1. Introduction. Automatic graph drawings have numerous applications in VLSI circuit layouts, networks, computer architecture, circuit schematics, etc. [3, 11]. Many graph drawing styles have been introduced [1, 3, 9, 11, 15, 17]. Among them, an “orthogonal drawing” has attracted much attention due to its various applications, especially in circuit schematics, entity relationship diagrams, data flow diagrams, etc. [14, 16, 19, 20]. An *orthogonal drawing* of a planar graph G is a drawing of G such that each vertex is mapped to a point, each edge is drawn as a sequence of alternate horizontal and vertical line segments, and any two edges do not cross except at their common end. A point where an edge changes its direction in a drawing is called a *bend* of the drawing. Figure 1.1(a) depicts an orthogonal drawing of the planar graph in Figure 1.1(b); the drawing has exactly one bend on the edge joining vertices g and t . If a planar graph G has a vertex of degree five or more, then G has no orthogonal drawing. On the other hand, if G has no vertex of degree five or more, that is, the maximum degree Δ of G is at most four, then G has an orthogonal drawing, but may need bends. If a planar graph represents a VLSI routing, then one may be interested in an orthogonal drawing such that the number of bends is as small as possible, because bends increase the manufacturing cost of a VLSI chip. An orthogonal drawing of a planar graph G is called an *optimal* orthogonal drawing if it has the minimum number of bends among all possible orthogonal drawings of G .

The problem of finding an optimal orthogonal drawing is one of the most famous problems in the graph drawing literature [3, 11] and has been studied both in the fixed embedding setting [7, 14, 16, 18, 20] and in the variable embedding setting [5, 6, 8, 12, 13]. A planar graph with a fixed embedding is called a *plane graph*. As a result in the fixed embedding, Tamassia [20] presented an algorithm to find an optimal orthogonal drawing of a plane graph G in time $O(n^2 \log n)$, where n is the

*Received by the editors August 16, 2006; accepted for publication (in revised form) May 15, 2008; published electronically October 17, 2008. This work is supported by JSPS grants. An early version of this paper has been presented in [22].

<http://www.siam.org/journals/sidma/22-4/66762.html>

[†]Graduate School of Information Sciences, Tohoku University, Aoba-yama 6-6-05, Sendai, 980-8579, Japan (zhou@ecei.tohoku.ac.jp, nishi@ecei.tohoku.ac.jp).

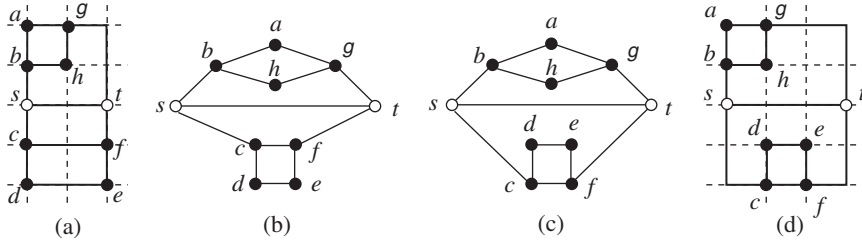


FIG. 1.1. (a) An optimal orthogonal drawing with one bend, (b), (c) two embeddings of the same planar graph, and (d) an orthogonal drawing with three bends.

number of vertices in G ; he reduced the optimal drawing problem to a min-cost flow problem. Then Garg and Tamassia improved the complexity to $O(n^{7/4}\sqrt{\log n})$ [7]. As a result in the variable embedding setting, Garg and Tamassia showed that the problem is NP-complete for planar graphs of $\Delta \leq 4$ [8]. However, Di Battista, Liotta, and Vargiu [5] showed that the problem can be solved in polynomial time for a planar graph G of $\Delta \leq 3$. Their algorithm finds an optimal orthogonal drawing among all possible plane embeddings of G . They use the properties of “spirality,” min-cost flow techniques, and a data structure, call a SPQ*R tree that implicitly represents all of the plane embeddings of G . The algorithm is complicated and takes time $O(n^5 \log n)$ for a planar graph of $\Delta \leq 3$. Using the algorithm, one can find more efficiently an optimal orthogonal drawing for a biconnected series-parallel simple graph; it takes time $O(n^4)$ if $\Delta \leq 4$ [5] and takes time $O(n^3)$ if $\Delta \leq 3$ [4, 5]. Note that every series-parallel graph is planar. Series-parallel graphs arise in a variety of problems such as scheduling, electrical networks, data-flow analysis, database logic programs, and circuit layout [21]. On the other hand, Garg and Liotta give an algorithm to find a nearly optimal orthogonal drawing for a biconnected planar graph of $\Delta \leq 3$ [6]. Their algorithm finds a drawing having at most three more bends than the optimal one in time $O(n^2)$. The complexities $O(n^5 \log n)$, $O(n^4)$, and $O(n^3)$ above for an exact algorithm in the variable embedding setting are very high, and it is expected to obtain an efficient algorithm for a particular class of planar graphs of $\Delta \leq 3$ [2].

In this paper we deal with the class of series-parallel (multi)graphs of $\Delta \leq 3$ and give a simple linear algorithm to find an optimal orthogonal drawing in the variable embedding setting. The graph G in Figure 1.1 is series-parallel and has various plane embeddings; two of them are illustrated in Figures 1.1(b) and (c); there is no plane embedding having an orthogonal drawing with no bend; however, the embedding in Figure 1.1(b) has an orthogonal drawing with one bend, as illustrated in Figure 1.1(a), and hence the drawing is optimal; the embedding in Figure 1.1(c) needs three bends, as illustrated in Figure 1.1(d); given G , our algorithm finds an optimal drawing in Figure 1.1(a). Our algorithm works well even if G has multiple edges or is not biconnected and is much simpler and faster than the algorithms for biconnected series-parallel simple graphs in [4, 5]; we use neither the min-cost flow technique nor the SPQ*R tree, but uses some structural features of series-parallel graphs, which have not been exploited in [21]. We furthermore obtain a best possible upper bound on the minimum number of bends. An early version of the paper has been presented at [22].

The rest of the paper is organized as follows. In section 2 we present some definitions and our main idea. In section 3 we present an algorithm and an upper

bound for biconnected series-parallel graphs. In section 4 we present an algorithm and an upper bound for nonbiconnected series-parallel graphs. Finally section 5 is a conclusion.

2. Preliminaries. In this section we present some definitions and our main idea.

Let $G = (V, E)$ be an undirected graph, with vertex set V and edge set E . We denote the number of vertices in G by $n(G)$ or simply by n . For a vertex $v \in V$, we denote by $G - v$ the graph obtained from G by deleting v . An edge joining vertices u and v is denoted by uv . We denote by $G - uv$ the graph obtained from G by deleting uv . We denote the degree of a vertex v in G by $d(v, G)$ or simply by $d(v)$. We denote the maximum degree of G by $\Delta(G)$ or simply by Δ . A connected graph is *biconnected* if there is no vertex whose removal results in a disconnected graph or a single-vertex graph K_1 . A *plane graph* is a fixed embedding of a planar graph.

Let G be a planar graph with $\Delta \leq 3$. We denote by $bend(G)$ the number of bends of an optimal orthogonal drawing of G in the variable embedding setting. (Thus $bend(G) = 1$ for the graph G in Figure 1.1.) Let D be an orthogonal drawing of G . The number of bends in D is denoted by $bend(D)$. Of course, $bend(G) \leq bend(D)$. Let $G(D)$ be a plane graph obtained from a drawing D by replacing each bend in D with a new vertex. Figures 2.1(a) and (b) depict $G(D)$ for the drawings D in Figures 1.1(a) and (d), respectively. An angle formed by two edges e and e' incident to a vertex v in $G(D)$ is called an *angle of vertex v* if e and e' appear consecutively around v . An angle of a vertex in $G(D)$ is called an *angle of the plane graph $G(D)$* . In an orthogonal drawing, every angle is $\pi/2$, π , $3\pi/2$, or 2π . Consider a labeling l , which assigns a label 1, 0, -1 , or -2 to every angle of $G(D)$. Labels 1, 0, -1 , and -2 correspond to angles $\pi/2$, π , $3\pi/2$, and 2π , respectively. We call l a *regular labeling* of $G(D)$ if l satisfies the following three conditions (a)–(c) [11, 20]:

- (a) for each vertex v of $G(D)$,
 - (a-1) if $d(v) = 1$, then the label of the angle of v is -2 ;
 - (a-2) if $d(v) = 2$, then the labels of the two angles of v total to 0; and
 - (a-2) if $d(v) = 3$, then the labels of the three angles of v total to 2;
- (b) the sum of the labels of each inner face is 4; and
- (c) the sum of the labels of the outer face is -4 .

Figures 2.1(a) and (b) illustrate regular labelings for the orthogonal drawings in Figures 1.1(a) and (d), respectively. If D is an orthogonal drawing of G , then clearly $G(D)$ has a regular labeling. Conversely, every regular labeling of $G(D)$ corresponds to an orthogonal drawing of G [20]. An orthogonal (geometric) drawing of G can be obtained from a regular labeling of $G(D)$ in linear time, that is, in time $O(n(G) + bend(D))$ [11, 20]. Therefore, from now on, we call a regular labeling of $G(D)$ an *orthogonal*

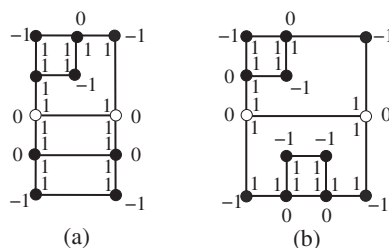


FIG. 2.1. Regular labelings of $G(D)$ corresponding to the drawings D in Figures 1.1(a) and (d), respectively.

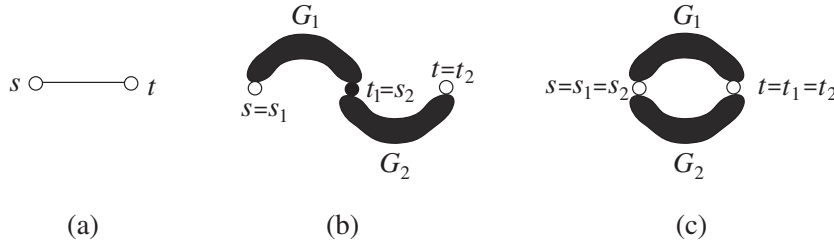


FIG. 2.2. (a) K_2 , (b) series, and (c) parallel connections.

drawing of a planar graph G or simply, a drawing of G , and obtain a regular labeling of G in place of an orthogonal (geometric) drawing of G .

For a drawing D of a planar graph G and for a subgraph G' of G , we denote by $D|G'$ the drawing of G' in D . Clearly

$$(2.1) \quad \text{bend}(G') \leq \text{bend}(D|G') \leq \text{bend}(D).$$

Let $\overline{G'}$ be the complement of G' , that is, the subgraph of G induced by all of the edges that are not contained in G' . Then

$$(2.2) \quad \text{bend}(D) = \text{bend}(D|G') + \text{bend}(D|\overline{G'}).$$

A series-parallel graph (with terminals s and t) is recursively defined as follows:

- (a) A graph G of a single edge is a series-parallel graph. The ends s and t of the edge are called the *terminals* of G . (See Figure 2.2(a).)
- (b) Let G_1 be a series-parallel graph with terminals s_1 and t_1 , and let G_2 be a series-parallel graph with terminals s_2 and t_2 .
 - (i) A graph G obtained from G_1 and G_2 by indentifying vertex t_1 with vertex s_2 is a series-parallel graph, whose terminals are $s = s_1$ and $t = t_2$. Such a connection is called a *series connection*. (See Figure 2.2(b).)
 - (ii) A graph G obtained from G_1 and G_2 by indentifying s_1 with s_2 and t_1 with t_2 is a series-parallel graph, whose terminals are $s = s_1 = s_2$ and $t = t_1 = t_2$. Such a connection is called a *parallel connection*. (See Figure 2.2(c).)

For example, the graph in Figure 1.1 is series-parallel.

Throughout the paper we assume that the maximum degree of a given series-parallel graph G is at most three, that is, $\Delta \leq 3$. We may assume without loss of generality that G is a simple graph, that is, G has no multiple edges, as follows. If a series-parallel multigraph G consists of exactly three multiple edges, then G has an optimal drawing of four bends; otherwise, insert a dummy vertex of degree two into an edge of each pair of multiple edges in G , and let G' be the resulting series-parallel simple graph, then an optimal drawing of the multigraph G can be immediately obtained from an optimal drawing of the simple graph G' by replacing each dummy vertex with a bend.

A drawing D of a series-parallel graph G is *outer* if the two terminals s and t of G are drawn on the outer face of D . A drawing D is called an *optimal outer drawing* of G if D is outer and $\text{bend}(D) = \text{bend}(G)$. The graph in Figure 1.1 has an optimal outer drawing, as illustrated in Figure 1.1(a). On the other hand, the graph in Figure 2.3(a) has no optimal outer drawing for the specified terminals s and t ; the no-bend drawing

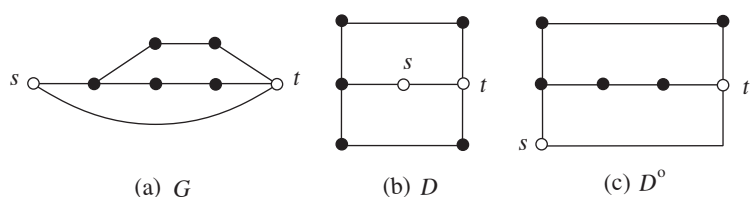


FIG. 2.3. (a) A biconnected series-parallel graph G , (b) an optimal drawing D , and (c) an outer drawing D^0 .

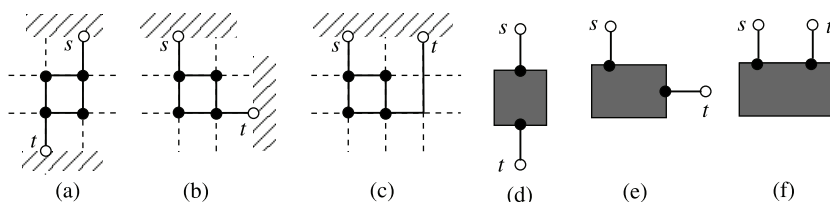


FIG. 2.4. (a)–(c) I-, L-, and U-shaped drawings, and (d)–(f) their schematic representations.

D in Figure 2.3(b) is optimal but is not outer, because s is not on the outer face; and the drawing D^0 with one bend in Figure 2.3(c) is outer but is not optimal.

Our main idea is to notice that a series-parallel graph G has an optimal outer drawing if G is “2-legged.” We say that G is 2-legged if $n(G) \geq 3$ and $d(s) = d(t) = 1$ for the terminals s and t of G . The edge incident to s or t is called a *leg* of G , and the neighbor of s or t is called a *leg-vertex*. For example, the series-parallel graphs in Figures 2.4(a)–(c) are 2-legged.

We will show in section 3 that every 2-legged series-parallel graph G has an optimal outer drawing, and the drawing has one of the three shapes: “I-shape,” “L-shape,” and “U-shape,” defined as follows. An outer drawing D of G is *I-shaped* if D intersects neither the north side of terminal s nor the south side of terminal t after rotating the drawing and renaming the terminals if necessary, as illustrated in Figure 2.4(a). D is *L-shaped* if D intersects neither the north side of s nor the east side of t after rotating the drawing and renaming the terminals if necessary, as illustrated in Figure 2.4(b). D is *U-shaped* if D does not intersect the north sides of s and t after rotating the drawing and renaming the terminals if necessary, as illustrated in Figure 2.4(c). In Figures 2.4(a)–(c) each side is shaded. The north side and the south side of a terminal contain the horizontal line passing through the terminal, while the east side of a terminal contains the vertical line passing through the terminal. The schematic representations of I-, L-, and U-shaped drawings are depicted in Figures 2.4(d), (e), and (f), respectively. D is called an *optimal X-shaped drawing*, $X = \text{I, L and U}$, if D is X-shaped and $\text{bend}(D) = \text{bend}(G)$.

More precisely, we will show in section 3 that every 2-legged series-parallel graph G , with $n(G) \geq 3$ has both an optimal I-shaped drawing and an optimal L-shaped drawing and that G has an optimal U-shaped drawing, too, unless G is a “diamond graph,” defined as follows. A *diamond graph* is either a graph in Figure 2.5(a) or obtained from two diamond graphs G' and G'' by connecting them in parallel and adding two legs, as illustrated in Figures 2.5(b) and (c).

For example, the 2-legged series-parallel graph in Figure 2.6(a) is a diamond graph and has both an optimal (no-bend) I-shaped drawing and an optimal (no-bend) L-

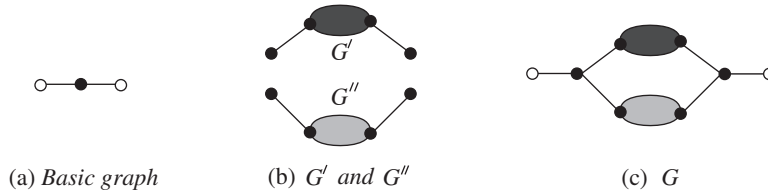


FIG. 2.5. Recursive definition of a diamond graph.

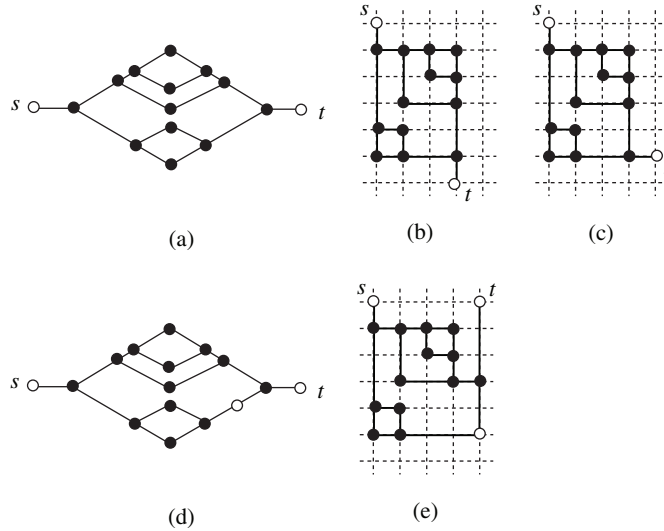


FIG. 2.6. (a) Diamond graph, (b) I-shaped drawing, (c) L-shaped drawing, (d) nondiamond graph, and (e) U-shaped drawing.

shaped drawing, as illustrated in Figures 2.6(b) and (c), but does not have an optimal (no-bend) U-shaped drawing. On the other hand, the 2-legged series-parallel graph in Figure 2.6(d) is obtained from the diamond graph in Figure 2.6(a) simply by inserting a new vertex of degree two in an edge and is not a diamond graph anymore. It has an optimal (no-bend) U-shaped drawing, too, as illustrated in Figure 2.6(e). Thus the diamond graph in Figure 2.6(a) has a U-shaped drawing with one bend.

3. Optimal drawing of biconnected series-parallel graph. In this section we give a linear algorithm to find an optimal drawing of a biconnected series-parallel graph G of $\Delta \leq 3$. We first give an algorithm for 2-legged series-parallel graphs in subsection 3.1. Using the algorithm, we then give an algorithm for biconnected series-parallel graphs in subsection 3.2.

3.1. 2-legged series-parallel graph. We first have the following lemma on a diamond graph.

LEMMA 3.1. *If G is a diamond graph, then*

- G has both a no-bend I-shaped drawing D_I and a no-bend L-shaped drawing D_L ;*
- D_I and D_L can be found in linear time; and*
- every no-bend drawing of G is either I-shaped or L-shaped, and hence G does not have a no-bend U-shaped drawing.*

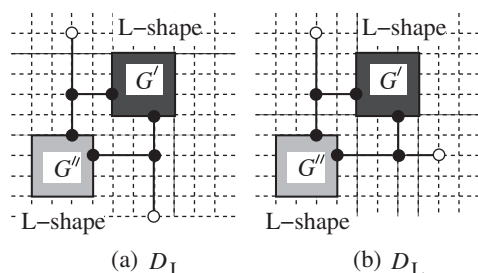


FIG. 3.1. Illustration of the proof of Lemma 3.1.

Proof. We prove the lemma by induction on the number $n(G)$ of vertices of G . Since G is a diamond graph, G has at least three vertices. If $n(G) = 3$, as illustrated in Figure 2.5(a), then (a), (b), and (c) trivially hold true. One may thus assume that $n(G) \geq 4$, and inductively assume that (a), (b), and (c) hold for every diamond graph of at most $n(G) - 1$ vertices.

We first prove that (a) holds for G . The definition of a diamond graph implies that one can obtain two diamond subgraphs G' and G'' from G by deleting the two terminals of G and splitting each of the two leg-vertices into two vertices, as illustrated in Figures 2.5(b) and (c). By the inductive hypothesis, (a) holds for G' , and hence G' has a no-bend L-shaped drawing D'_L . Similarly, G'' has a no-bend L-shaped drawing D''_L . Combining a flipped drawing of D'_L with the drawing D''_L and drawing the two legs appropriately, one can construct a no-bend I-shaped drawing D_I and a no-bend L-shaped drawing D_L of G , as illustrated in Figures 3.1(a) and (b), respectively. Thus (a) holds true.

One can obtain (regular labelings of) D_I and D_L from (those of) D'_L and D''_L in constant time by deciding the labels of the angles of the terminals and leg vertices of G . Thus (b) holds true for G .

We now prove that (c) holds for G . Let D^* be an arbitrary no-bend drawing of G . Then both the drawing $D^*[G']$ of G' in D^* and the drawing $D^*[G'']$ of G'' in D^* are no-bend drawings. By the inductive hypothesis, each of $D^*[G']$ and $D^*[G'']$ is either I-shaped or L-shaped. Since D^* is a no-bend drawing, both $D^*[G']$ and $D^*[G'']$ must be L-shaped and D^* must be either I-shaped or L-shaped. Thus (c) holds true. $\square$

The proof of Lemma 3.1 immediately yields a linear algorithm **Diamond**(G, D_I, D_L), which recursively finds both a no-bend I-shaped drawing D_I and a no-bend L-shaped drawing D_L of a given diamond graph G .

The following lemma holds for a 2-legged series-parallel graph G which is not a diamond graph.

LEMMA 3.2. *The following (a), (b), and (c) hold for a 2-legged series-parallel graph G , with $n(G) \geq 3$ unless G is a diamond graph:*

- (a) G has three optimal I-, L-, and U-shaped drawings D_I, D_L , and D_U ;
- (b) D_I, D_L , and D_U can be found in linear time; and
- (c) $\text{bend}(G) \leq (n(G) - 2)/3$.

Proof. We prove the lemma by induction on the number $n(G)$ of vertices of G . Assume that G is a 2-legged series-parallel graph, with $n(G) \geq 3$, and G is not a diamond graph. Then $n(G) \geq 4$. If $n(G) = 4$, that is, G is a path of four vertices, then the lemma trivially holds true, as illustrated in Figure 3.2. One may thus assume $n(G) \geq 5$, and inductively assume that the lemma holds for every 2-legged

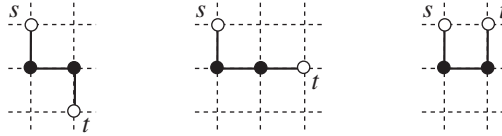


FIG. 3.2. Optimal I-, L-, and U-shaped drawings of a path of four vertices.

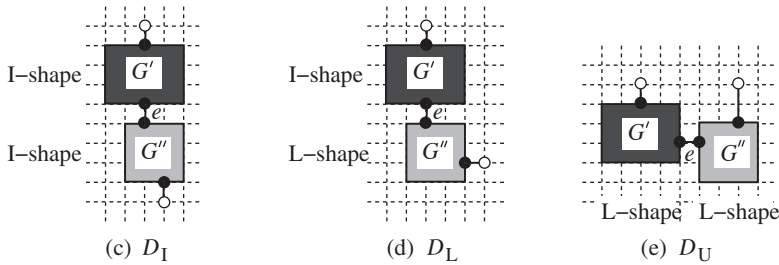
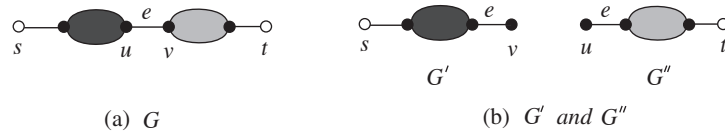


FIG. 3.3. Illustration of the proof of Lemma 3.2.

series-parallel graph of at most $n(G) - 1$ vertices. We now prove that the lemma holds for G .

(a) We first prove that (a) holds for G . Let s and t be the terminals of G , then $d(s) = d(t) = 1$. Since G is a 2-legged series-parallel graph, the graph $G - s - t$ obtained from G by deleting vertices s and t is a series-parallel graph having the leg-vertices as the terminals, and hence $G - s - t$ is either a series connection or a parallel connection of two subgraphs.

Consider first the case where $G - s - t$ is a series connection of two subgraphs. In this case, since $\Delta(G) \leq 3$, G has a bridge $e = uv$ other than the two legs, that is, $G - e$ is disconnected, as illustrated in Figure 3.3(a). Then G contains two subgraphs G' and G'' ; G' is a 2-legged series-parallel graph with terminals s and v , and G'' is a 2-legged series-parallel graph with terminals u and t , as illustrated in Figure 3.3(b). If G' is not a diamond graph, then $4 \leq n(G') < n(G)$, and hence by the inductive hypothesis, G' has three optimal drawings of I-, L-, and U-shapes. If G' is a diamond graph, then by Lemma 3.1(a), G' has two optimal (no-bend) drawings of I- and L-shapes. Hence, in either case, G' has two optimal drawings of I- and L-shapes. Similarly, G'' has two optimal drawings of I- and L-shapes. Combining these drawings of G' and G'' , one can easily construct three optimal drawings D_I , D_L , and D_U of G , as illustrated in Figures 3.3(c), (d), and (e). Thus (a) holds true.

Consider next the case where $G - s - t$ is a parallel connection of two subgraphs. Then G has exactly one biconnected component, and the degrees of the leg-vertices are three, as illustrated in Figure 3.4(a). Deleting the terminals of G and splitting the two leg-vertices of G , one can obtain two edge-disjoint 2-legged series-parallel

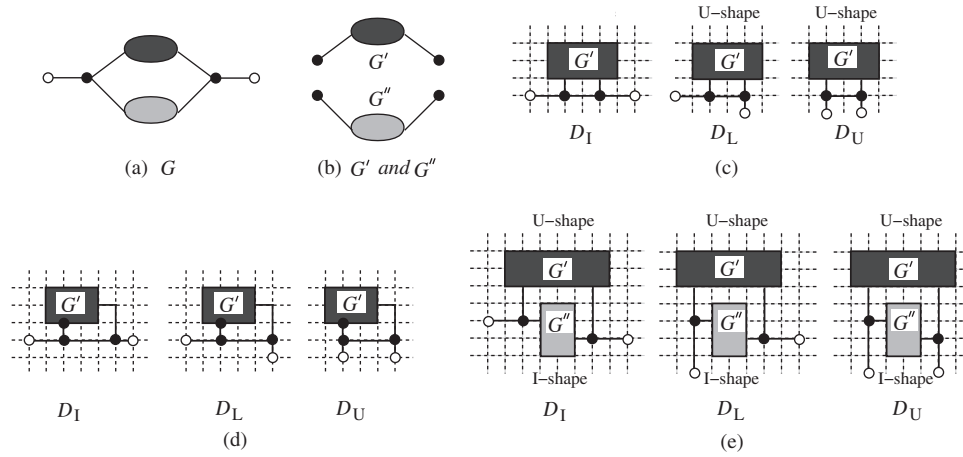


FIG. 3.4. Illustration of the proof of Lemma 3.2.

subgraphs G' and G'' , as illustrated in Figure 3.4(b). Since $n(G'), n(G'') \geq 2$, there are the following two cases to consider.

Case 1: Either $n(G') = 2$ or $n(G'') = 2$.

One may assume that $n(G'') = 2$, and hence G'' consists of a single edge. Since G is a simple graph, we have $n(G') \geq 3$.

Consider first the case where G' is not a diamond graph. Then $4 \leq n(G') < n(G)$, and hence by the inductive hypothesis, G' has an optimal U-shaped drawing D'_U . From a flipped drawing of D'_U , one can easily construct three optimal drawings D_I , D_L , and D_U of G , as illustrated in Figure 3.4(c).

Consider next the case where G' is a diamond graph. Then by Lemma 3.1(a), G' has a no-bend L-shaped drawing D'_L . From a flipped drawing of D'_L , one can easily construct three drawings D_I , D_L , and D_U of G , with one bend, as illustrated in Figure 3.4(d). Since G' is a diamond graph, by Lemma 3.1(c), every no-bend drawing of G' is either I-shaped or L-shaped. Therefore, we have $bend(G) \geq 1$; if G had a no-bend drawing D , then $D|G'$ would be I- or L-shaped, and hence D would have one or more bends on the edge in G'' , a contradiction. Thus the constructed drawings D_I , D_L , and D_U having exactly one bend are optimal.

Case 2: $n(G'), n(G'') \geq 3$.

Since G is not a diamond graph, either G' or G'' is not a diamond graph. One may assume without loss of generality that G' is not a diamond graph. Then $4 \leq n(G') < n(G)$, and hence by the inductive hypothesis, G' has an optimal U-shaped drawing D'_U . On the other hand, G'' has an optimal I-shaped drawing D'_I ; if G'' is a diamond graph, then by Lemma 3.1(a), G'' has an optimal (no-bend) I-shaped drawing D'_I ; otherwise, by the inductive hypothesis, G'' has an optimal I-shaped drawing D'_I . From the U-shaped drawing D'_U and the I-shaped drawing D'_I , one can easily construct I-, L-, and U-shaped drawings D_I , D_L , and D_U of G without introducing any new bends, as illustrated in Figure 3.4(e). Since D_I has no bend on the legs, by (2.1) and (2.2) we have

$$bend(D_I) = bend(D'_U) + bend(D'_I).$$

Since D'_U and D'_I are optimal drawings of G' and G'' , respectively,

$$bend(D'_U) = bend(G')$$



FIG. 3.5. A graph attaining the bound in Lemma 3.2(c).

and

$$\text{bend}(D_1'') = \text{bend}(G'').$$

Let D^* be an arbitrary optimal drawing of G , then clearly

$$\begin{aligned} \text{bend}(G') &\leq \text{bend}(D^*|G'), \\ \text{bend}(G'') &\leq \text{bend}(D^*|G''), \end{aligned}$$

and

$$\text{bend}(D^*|G') + \text{bend}(D^*|G'') \leq \text{bend}(D^*) = \text{bend}(G).$$

From these six equations we have $\text{bend}(D_I) \leq \text{bend}(G)$, and hence D_I is an optimal drawing of G . Similarly, D_L and D_U are optimal drawings of G . Thus (a) holds true.

(b) We now prove (b). One can construct the (regular labeling of) drawings D_I , D_L , and D_U above from the (regular labeling of) drawings of G' and G'' simply by deciding the labels of the terminals and leg-vertices in G ; this can be done in constant time. Thus (b) holds true for G .

(c) We finally prove (c). If a new bend is produced in a construction of G , as illustrated in Figures 3.3 and 3.4, then the construction is one in Figure 3.4(d), $n(G) \geq 5$, and $\text{bend}(G) = \text{bend}(G') + 1 = 1 \leq (n(G) - 2)/3$. In any other construction, no new bend is produced. Noting this fact, one can inductively prove that $n(G) \leq (n(G) - 2)/3$. $\square$

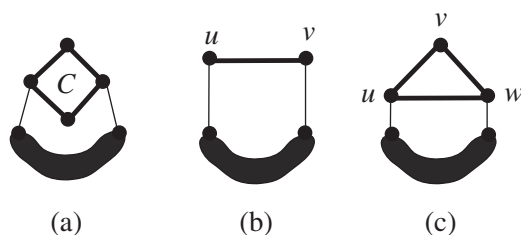
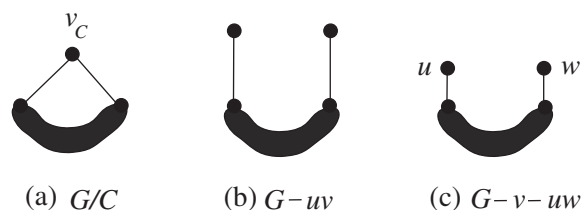
We denote by K_n a complete graph of $n(\geq 1)$ vertices. Let G be a 2-legged series-parallel graph obtained from copies of K_2 and K_3 by connecting them alternately in series, as illustrated in Figure 3.5. Then $\text{bend}(G) = (n(G) - 2)/3$. Thus the bound in Lemma 3.2(c) is best possible.

The proof of Lemma 3.2 immediately yields a linear algorithm **NonDiamond**(G, D_I, D_L, D_U), which recursively finds three optimal I-, L-, and U-shaped drawings D_I , D_L , and D_U of a given 2-legged series-parallel graph G unless G is a diamond graph. By algorithms **Diamond** and **NonDiamond**, one can find an optimal drawing of a 2-legged series-parallel graph G . Note that **NonDiamond** may call **Diamond**.

For a series-parallel graph G with $d(s), d(t) \leq 2$, one can easily find an optimal outer drawing D of G , as follows. Add to G a dummy edge $s's$ for a new vertex s' if $d(s) = 2$, and add to G a dummy edge tt' for a new vertex t' if $d(t) = 2$. Then the resulting graph G' is a 2-legged series-parallel graph, and $\Delta(G') \leq 3$. Find an optimal drawing D' of G' by **Diamond** or **NonDiamond** and delete the dummy edges from D' . Then the resulting drawing D of G is clearly optimal and outer.

3.2. Biconnected series-parallel graphs. A biconnected series-parallel graph G can be defined (without specifying terminals) as a biconnected graph which has no K_4 as a minor. For every edge uv in G , G is a series-parallel graph with terminals u and v .

A cycle C of four vertices in a graph G is called a *diamond* if two nonconsecutive vertices of C have degree two in G and the other two vertices of C have degree three

FIG. 3.6. Substructures contained in a biconnected series-parallel graph of $\Delta \leq 3$.FIG. 3.7. Smaller graphs G' .

and are not adjacent in G , as illustrated in Figure 3.6(a). We denote by G/C the graph obtained from G by contracting C to a new single vertex v_C , as illustrated in Figure 3.7(a). (Note that $G_C = G/C$ is series-parallel if G is series-parallel. One can observe that, from every diamond graph, one can obtain a graph in Figure 2.5(a) by repeatedly contracting a diamond.)

Noting that every biconnected series-parallel graph G has a vertex of degree two, one can easily observe that the following Lemma 3.3 holds. (Lemma 3.3 is also an immediate consequence of Lemma 2.1 in [10] on a general series-parallel graph.)

LEMMA 3.3. *Every biconnected series-parallel graph G of $\Delta \leq 3$ has, as a subgraph, one of the following three substructures (a)–(c) illustrated in Figure 3.6:*

- (a) a diamond C ;
- (b) two adjacent vertices u and v such that $d(u) = d(v) = 2$; and
- (c) a complete graph K_3 of three vertices u , v , and w such that $d(v) = 2$.

Our idea is to reduce the optimal drawing problem for a biconnected series-parallel graph G to that for a smaller graph G' illustrated in Figure 3.7, as in the following Lemmas 3.4 and 3.5.

LEMMA 3.4. *If a biconnected series-parallel graph G , with $n(G) \geq 6$ has a diamond C , then $\text{bend}(G) = \text{bend}(G')$ for $G' = G/C$.*

Proof. Assume that a biconnected series-parallel graph G has a diamond C and that s' and t' are the two vertices of C having degree three in G , as illustrated in Figure 3.8(a). Let $G' = G/C$ be the graph obtained from G by contracting C to a new single vertex v_C , as illustrated in Figure 3.8(b). Since $n(G) \geq 6$ and G is biconnected, the neighbor of s' outside C is different from that of t' . Therefore, G' is a biconnected series-parallel (simple) graph. Let D'^* be an optimal drawing of G' , then $\text{bend}(D'^*) = \text{bend}(G')$. (See Figure 3.8(e).) Replace the vertex v_C in D'^* with a rectangular drawing of C , and let D be the resulting drawing of G , as illustrated in Figure 3.8(d). We claim that D is an optimal drawing of G , and hence $\text{bend}(G) = \text{bend}(G')$.

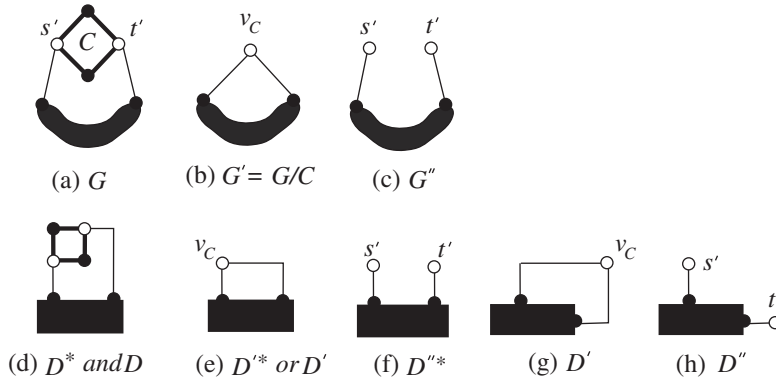


FIG. 3.8. Illustration for the proof of Lemma 3.4.

The (regular labeling of) drawing D is constructed from the (regular labeling of) optimal drawing D'^* of G' without introducing any new bend, and hence we have $\text{bend}(D) = \text{bend}(D'^*) = \text{bend}(G')$. Let D^* be an arbitrary optimal drawing of G , then $\text{bend}(D^*) = \text{bend}(G)$. If

$$(3.1) \quad \text{bend}(G') \leq \text{bend}(D^*),$$

then $\text{bend}(D) = \text{bend}(G') \leq \text{bend}(D^*) = \text{bend}(G)$, and hence D is an optimal drawing of G . It thus suffices to verify (3.1). There are the following two cases to consider.

Case 1: $\text{bend}(D^*|C) = 0$.

In this case, $D^*|C$ is a rectangle, as illustrated in Figure 3.8(d). Contracting C in D^* to a single vertex v_C and appropriately deciding the two labels of v_C , one can obtain (a regular labeling of) a drawing D' of $G' = G/C$ from D^* without introducing any new bend, as illustrated in Figure 3.8(e). We thus have $\text{bend}(G') \leq \text{bend}(D') = \text{bend}(D^*)$.

Case 2: $\text{bend}(D^*|C) \geq 1$.

In this case, $D^*|C$ is a rectilinear polygon having five or more (geometric) vertices including at least one bend. Delete from G the two vertices of C having degree two, and let G'' be the resulting graph, as illustrated in Figure 3.8(c). Since G is biconnected, G'' is a 2-legged series-parallel graph with terminals s' and t' . (Note that every biconnected series-parallel graph is a series-parallel graph having the ends of an arbitrary edge as terminals, and that a graph obtained from a series-parallel graph by contracting an edge is series-parallel.) Since G'' is the complement of C in G , by (2.2) we have

$$(3.2) \quad \text{bend}(D^*|G'') + \text{bend}(D^*|C) = \text{bend}(D^*).$$

There are the following two subcases to consider.

Case 2.1: G'' is not a diamond graph.

Since $n(G) \geq 6$, $n(G'') = n(G) - 2 \geq 4$. Since G'' is not a diamond graph, by Lemma 3.2(a), G'' has an optimal U-shaped drawing D''^* , as illustrated in Figure 3.8(f). D''^* can be modified to a drawing D' of G' by identifying s' with t' as a single vertex v_C and introducing a new bend, as illustrated in Figure 3.8(g), and

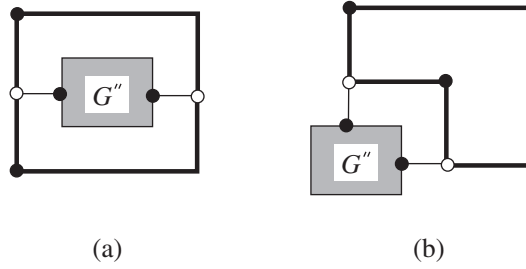


FIG. 3.9. An optimal drawing D^* of G such that $D^*[G'']$ is I-shaped in (a) or L-shaped in (b).

hence by (3.2),

$$\begin{aligned}
 \text{bend}(G') &\leq \text{bend}(D') \\
 &= \text{bend}(D''^*) + 1 \\
 &= \text{bend}(G'') + 1 \\
 &\leq \text{bend}(D^*[G'']) + \text{bend}(D^*[C]) \\
 &= \text{bend}(D^*).
 \end{aligned}$$

Case 2.2: G'' is a diamond graph.

Since $\text{bend}(D^*[C]) \geq 1$, by (2.1) we have $\text{bend}(D^*) \geq \text{bend}(D^*[C]) \geq 1$. We now claim that

$$(3.3) \quad \text{bend}(D^*) \geq 2.$$

Suppose for a contradiction that $\text{bend}(D^*) = 1$. Since $\text{bend}(D^*[C]) \geq 1$, by (3.2) we have $\text{bend}(D^*[C]) = 1$ and $\text{bend}(D^*[G'']) = 0$, and hence $D^*[G'']$ is a no-bend drawing of G'' . Then by Lemma 3.1(c), $D^*[G'']$ is either I-shaped or L-shaped, as illustrated in Figures 3.9(a) and (b), respectively. In either case, the drawing of C needs two or more bends in D^* , that is, $\text{bend}(D^*[C]) \geq 2$, as illustrated in Figure 3.9, where C is drawn by thick lines. This is contrary to $\text{bend}(D^*[C]) = 1$.

Since G'' is a diamond graph, by Lemma 3.1(a) G'' has a no-bend L-shaped drawing D'' , as illustrated in Figure 3.8(h). The drawing D'' can be modified to a drawing D' of $G' = G/C$ by identifying the two vertices s' and t' as a single vertex v_C and introducing two new bends, as illustrated in Figure 3.8(g), and hence we have

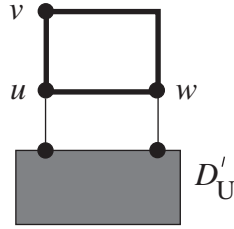
$$(3.4) \quad \text{bend}(G') \leq \text{bend}(D') = \text{bend}(D'') + 2 = 2.$$

By (3.3) and (3.4) we have $\text{bend}(G') \leq 2 \leq \text{bend}(D^*)$. $\square$

LEMMA 3.5. Assume that a biconnected series-parallel graph G , with $n(G) \geq 6$ has no diamond. If G has a substructure of type (b) in Figure 3.6(b), then $\text{bend}(G) = \text{bend}(G')$ for $G' = G - uv$. If G has a substructure of type (c) in Figure 3.6(c), then $\text{bend}(G) = \text{bend}(G') + 1$ for $G' = G - v - uw$. (See Figures 3.7(b) and (c).)

Proof. Since G has no diamond, by Lemma 3.3, G has either a substructure (b) or (c).

Suppose first that G has a substructure (b). Let $G' = G - uv$, as illustrated in Figure 3.7(b). Then G' is a 2-legged series-parallel graph with terminals u and v . Clearly $n(G') = n(G) \geq 6$. Since G has no diamond, the subgraph G' of G has no diamond. Therefore G' is not a diamond graph, and hence by Lemma 3.2(a), G' has


 FIG. 3.10. Construction of D from D'_U .

an optimal U-shaped drawing D'_U . D'_U can be extended to a drawing D of G simply by drawing uw as a straight line segment. Since $\text{bend}(D) = \text{bend}(D'_U) = \text{bend}(G') \leq \text{bend}(G)$, D is an optimal drawing of G . Thus $\text{bend}(G) = \text{bend}(D) = \text{bend}(G')$.

Suppose next that G has a substructure (c). Let $G' = G - v - uw$, as illustrated in Figure 3.7(c). Then G' is a 2-legged series-parallel graph with terminals u and w and is not a diamond graph. Therefore, by Lemma 3.2(a), G' has an optimal U-shaped drawing D'_U . D'_U can be extended to a drawing D of G by drawing the complete graph K_3 of three vertices u, v, w as a rectangle with one bend, as illustrated in Figure 3.10. Any orthogonal drawing of K_3 needs a bend, and hence $\text{bend}(D) = \text{bend}(D'_U) + 1 = \text{bend}(G') + 1 \leq \text{bend}(G)$. Thus D is an optimal drawing of G , and hence $\text{bend}(G) = \text{bend}(G') + 1$. $\square$

From the proofs of Lemmas 3.4 and 3.5 we have the following algorithm **Biconnected**(G, D) to find an optimal drawing D of a biconnected series-parallel graph G .

Biconnected(G, D);

begin

One may assume that $n(G) \geq 6$ (otherwise, one can easily find an optimal drawing D of G in linear time);

{By Lemma 3.3, G has one of the three substructures (a)–(c) in Figure 3.6.}

Case 1: G has a diamond C ;

Let $G' = G/C$; { G' is a biconnected series-parallel graph.}

Biconnected(G', D');

Extend an optimal drawing D' of G' to an optimal drawing D of G simply by replacing v_C by a rectangular drawing of C , as illustrated in

Figures 3.8(d) and (e); {cf. Lemma 3.4}

Case 2: G has no diamond but has a substructure (b);

Let $G' = G - uv$;

{ G' is a 2-legged series-parallel graph with terminals u and v and is not a diamond graph.}

Find an optimal U-shaped drawing D'_U of G' by **NonDiamond**;

{cf. Lemma 3.2}

Extend D'_U to an optimal drawing D of G by drawing uv as a straight line segment;

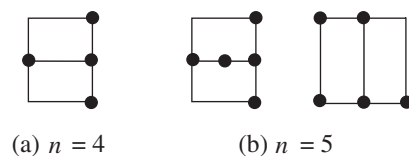
{cf. Lemma 3.5}

Case 3: G has neither a diamond nor a substructure (b) but has a substructure (c);

Let $G' = G - v - uw$;

{ G' is a 2-legged series-parallel graph with terminals u and w and is not a diamond graph.}

Find an optimal U-shaped drawing D'_U of G' by **NonDiamond**;

FIG. 3.11. Biconnected series-parallel graphs with $n = 4$ or $n = 5$.

Extend D'_U to an optimal drawing D of G by drawing $K_3 = uvw$ as a rectangle with one bend, as illustrated in Figure 3.10; {cf. Lemma 3.5}

end.

All substructures (a)–(c) can be found total in time $O(n)$ by a standard bookkeeping method to maintain all degrees of vertices together with all paths of length two with an intermediate vertex of degree two. One can thus observe that **Biconnected** can be executed in linear time.

We thus have the following theorem.

THEOREM 3.6. *An optimal orthogonal drawing of a series-parallel biconnected graph G of $\Delta \leq 3$ can be found in linear time.*

It should be noted that the (original) terminals s and t of G are not always on the outer face of a drawing D obtained by **Biconnected**, and hence D is not always an outer drawing, as known from the example in Figure 2.3.

Di Battista, Liotta, and Vargiu gave an $O(n^3)$ algorithm for biconnected series-parallel simple graphs using a min-cost flow technique and a data structure called an SPQ*R tree [4, 5]. Our linear algorithm is much simpler and faster than their algorithm.

We have the following lemma on the minimum number of bends.

LEMMA 3.7. *If a series-parallel graph G is biconnected and $\Delta \leq 3$, then*

$$(3.5) \quad \text{bend}(G) \leq \lceil n(G)/3 \rceil.$$

Proof. We prove (3.5) by induction on $n(G)$. Since G is biconnected, $n(G) \geq 3$.

1°. We first show that (3.5) holds if $3 \leq n(G) \leq 5$. If $n(G) = 3$, then $G = K_3$ and $\text{bend}(G) = 1 = \lceil n(G)/3 \rceil$. If $n(G) = 4$ and $\text{bend}(G) \geq 1$, then $G = K_4 - e$ in Figure 3.11(a), and hence $\text{bend}(G) = 2 \leq \lceil n(G)/3 \rceil$, where $K_4 - e$ is a graph obtained from K_4 by deleting an edge. If $n(G) = 5$ and $\text{bend}(G) \geq 1$, then G is one of the two graphs in Figure 3.11(b), and hence $\text{bend}(G) \leq 2 \leq \lceil n(G)/3 \rceil$.

2°. Assume that $n(G) \geq 6$, and assume inductively that (3.5) holds for every series-parallel biconnected graph of at most $n(G) - 1$ vertices.

3°. We then show that (3.5) holds for G . There are the following three cases to consider.

Case 1: G has a diamond C .

By Lemma 3.4 we have $\text{bend}(G) = \text{bend}(G')$ for $G' = G/C$. Since $n(G') = n(G) - 3$ and G' is a series-parallel biconnected graph, by the inductive hypothesis we have $\text{bend}(G') \leq \lceil n(G')/3 \rceil$. We thus have $\text{bend}(G) = \text{bend}(G') < \lceil n(G)/3 \rceil$.

Case 2: G has no diamond but has a substructure (b) in Figure 3.6(b).

By Lemma 3.5 we have $\text{bend}(G) = \text{bend}(G')$ and $n(G) = n(G')$ for $G' = G - uv$. Since G' is a 2-legged series-parallel graph and is not a diamond graph, by Lemma 3.2(c) we have $\text{bend}(G') \leq (n(G') - 2)/3$. We thus have $\text{bend}(G) = \text{bend}(G') \leq \lceil n(G)/3 \rceil$.

Case 3: G has neither a diamond nor a substructure (b) but has a substructure (c) in Figure 3.6(c).

By Lemma 3.5 we have $\text{bend}(G) = \text{bend}(G') + 1$ and $n(G) = n(G') + 1$ for $G' = G - v - uw$. Since G' is a 2-legged series-parallel and is not a diamond graph, by Lemma 3.2(c) we have $\text{bend}(G') \leq (n(G') - 2)/3$. We thus have

$$\begin{aligned} \text{bend}(G) &\leq (n(G') - 2)/3 + 1 \\ &= n(G)/3 \\ &\leq \lceil n(G)/3 \rceil. \quad \square \end{aligned}$$

4. Optimal drawing of nonbiconnected series-parallel graph. It has been left as open by Di Battista, Liotta, and Vargiu [4, 5] to obtain an algorithm for nonbiconnected series-parallel graphs G . In section 4.1 we give a linear algorithm for G , and in section 4.2 we give a best possible upper bound on $\text{bend}(G)$.

4.1. Algorithm. By Lemma 3.1, Lemma 3.2, and Theorem 3.6, one may assume that G is neither 2-legged nor biconnected. Therefore either $d(s) \geq 2$ or $d(t) \geq 2$, and hence one may assume without loss of generality that $d(s) \geq 2$.

We now claim that one may further assume that $d(t) \geq 2$, and hence G consists of biconnected components $B_1, B_2, \dots, B_p$, $p \geq 2$, and $p - 1$ copies of K_2 , alternately connected in series, as illustrated in Figure 4.1(a), after replacing each induced path contained in none of the biconnected components by a single edge. Suppose that $d(t) = 1$, as illustrated in Figure 4.1(b). Then $G' = G - t$ is a series-parallel graph with terminals s and t_p , and takes the form in Figure 4.1(a). Since $d(t_p, G') = d(t_p, G) - 1 = 2$, there is an angle of π or $3\pi/2$ around vertex t_p in any optimal drawing D' of G' . An optimal drawing D of G can be obtained from D' simply by inserting the edge $t_p t$ in the angle. One may thus assume that $d(t) \geq 2$.

For a series-parallel graph G in Figure 4.1(a), each biconnected component B_i , $1 \leq i \leq p$, is a series-parallel graph with terminals s_i and t_i , where $s_1 = s$ and $t_p = t$. Clearly

$$(4.1) \quad \sum_{i=1}^p \text{bend}(B_i) \leq \text{bend}(G).$$

However, (4.1) does not always hold with equality, as follows. Consider the series-parallel graph G in Figure 4.2(a); G has two biconnected components B_1 and B_2 , and hence $p = 2$. B_1 and B_2 have no-bend (optimal) drawings D_1 and D_2 , as illustrated in Figures 4.2(b) and (c), respectively. Thus $\text{bend}(B_1) = \text{bend}(B_2) = 0$, and hence $\text{bend}(B_1) + \text{bend}(B_2) = 0$. However, terminal t_1 of B_1 is not on the outer face of any no-bend drawing of B_1 . Similarly, s_2 is not on the outer face of any no-bend drawing of B_2 . Therefore, one cannot connect t_1 in D_1 and s_2 in D_2 by an edge without edge-crossing. Thus G does not have a no-bend drawing. On the other hand, B_2 has a drawing D_2' with one bend in which s_2 is on the outer face, as illustrated in Figure 4.2(d). Combining D_1 and D_2' , one can construct a drawing D of G with one bend, as illustrated in Figure 4.2(e). Thus $\text{bend}(G) = 1$, and hence

$$\text{bend}(B_1) + \text{bend}(B_2) < \text{bend}(G).$$

In section 4.3, we will prove

$$(4.2) \quad \text{bend}(G) = \sum_{i=1}^p \text{bend}(B_i) \quad \text{or} \quad \sum_{i=1}^p \text{bend}(B_i) + 1.$$

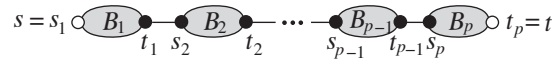
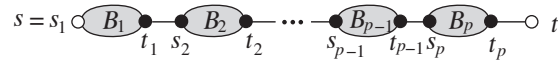
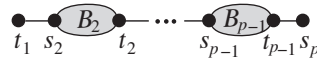
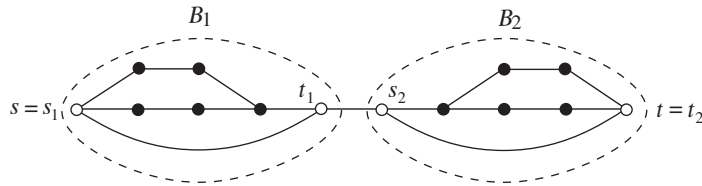
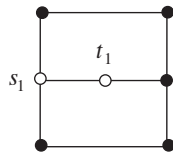
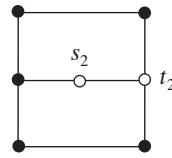
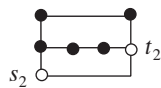
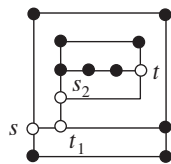
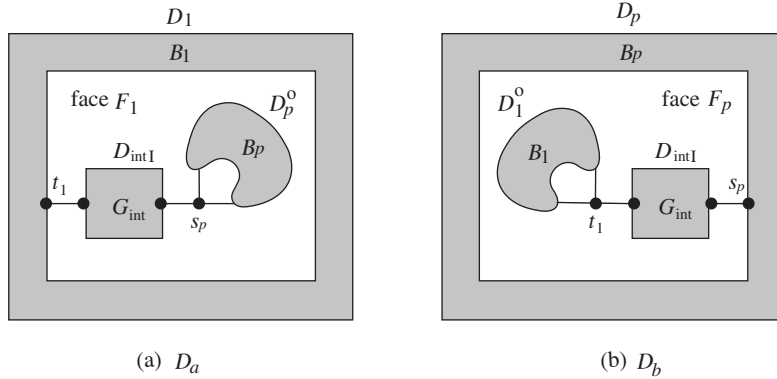
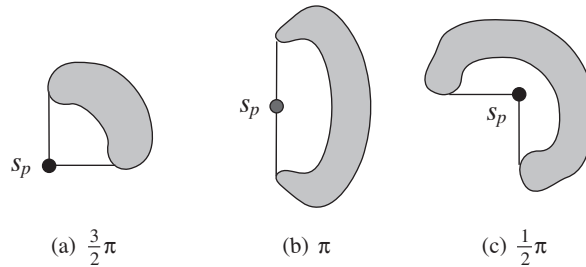
(a) G with $d(s), d(t) \geq 2$ (b) G with $d(t) = 1$ (c) G_{int}

FIG. 4.1. Nonbiconnected series-parallel graphs.

(a) G (b) D_1 (c) D_2 (d) D_2^o (e) D FIG. 4.2. Construction of an optimal drawing D of G for a case $p = 2$.

Delete from a series-parallel graph G in Figure 4.1(a) all vertices in B_1 and B_p except t_1 and s_p , and let G_{int} be the resulting series-parallel graph with terminals t_1 and s_p , as illustrated in Figure 4.1(c). Then we claim that G_{int} has an optimal I-shaped drawing. If $p = 2$, then $G_{\text{int}} = K_2$, and hence G_{int} has an optimal (no-bend) I-shaped drawing D_{intI} . If $p \geq 3$, then G is 2-legged, and hence by Lemmas 3.1 and 3.2, G_{int} has an optimal I-shaped drawing D_{intI} , in which both terminals t_1 and s_p are on the outer face.


 FIG. 4.3. Two alternative drawings D_a and D_b of G .

 FIG. 4.4. Three kinds of the outer angle of s_p in D_p^o .

As known from the example in Figure 4.2, we have the following lemma.

LEMMA 4.1. Let G be a series-parallel graph taking the form in Figure 4.1(a). Then an arbitrary drawing D of G has one of the following two alternatives forms:

- (a) the terminal s_p of B_p is on the outer face of the drawing $D_p^o = D|B_p$ of B_p (see Figure 4.3(a), for example); and
- (b) the terminal t_1 of B_1 is on the outer face of the drawing $D_1^o = D|B_1$ of B_1 (see Figure 4.3(b), for example).

In (a) above, the *outer angle* of s_p , that is, the angle around s_p in the outer face of D_p^o , must be either $3\pi/2$ or π , as illustrated in Figures 4.4(a) and (b); if it were $\pi/2$, as illustrated in Figure 4.4(c), then the edge $t_{p-1}s_p$ could not be drawn in the outer angle. Since $d(t_1, B_1) = 2$, one of the two angles of t_1 in any drawing of B_1 has an angle of $3\pi/2$ or π . Of course, in an arbitrary drawing D of G , both G_{int} and B_p are drawn in the same face F_1 of $D_1 = D|B_1$ having such an angle of t_1 , as illustrated in Figure 4.3(a). Note that face F_1 may be outer although it is drawn as an inner face in Figure 4.3(a). Similarly, in (b) above, both G_{int} and B_1 are drawn in the same face F_p of $D_p = D|B_p$, as illustrated in Figure 4.3(b).

Let B be a biconnected series-parallel graph, and let v be a vertex of degree two in B . A drawing D of B is (α, v) -*outer* if v is on the outer face of D and the outer angle of v is α , where $\alpha = \pi/2, \pi$, or $3\pi/2$. An (α, v) -outer drawing D is *optimal* if D has the minimum number of bends among all possible (α, v) -outer drawings of B . The minimum number of bends is denoted by $\text{bend}(B, \alpha, v)$. Clearly $\text{bend}(B) \leq \text{bend}(B, \alpha, v)$. However, the equation does not always hold with equality. For example, $\text{bend}(B_2) = 0 < 1 = \text{bend}(B_2, 3\pi/2, s_2)$ for B_2 in Figure 4.2(a).

Lemma 4.1 and the arguments above imply that we shall find optimal (α, t_1) -outer drawings of B_1 and optimal (α, s_p) -outer drawings of B_p for both $\alpha = 3\pi/2$ or $\alpha = \pi$. However, it suffices to find them only for $\alpha = 3\pi/2$, because $\text{bend}(B_1, 3\pi/2, t_1) \leq \text{bend}(B_1, \pi, t_1)$ and $\text{bend}(B_p, 3\pi/2, s_p) \leq \text{bend}(B_p, \pi, s_p)$ as known from the following, Lemma 4.2(a), whose proof will be given in section 4.3.

LEMMA 4.2. *If a biconnected series-parallel graph B has a vertex v of degree two, then*

(a)

$$(4.3) \quad \text{bend}(B, 3\pi/2, v) \leq \min\{\text{bend}(B, \pi/2, v), \text{bend}(B, \pi, v)\};$$

(b)

$$\text{bend}(B, 3\pi/2, v) = \text{bend}(B) \quad \text{or} \quad \text{bend}(B) + 1;$$

and

(c) *an optimal $(3\pi/2, v)$ -outer drawing of B can be found in linear time.*

The proof of Lemma 4.2 yields a linear algorithm **OuterDrawing** (B, v, D) to find an optimal $(3\pi/2, v)$ -outer drawing of B . Thus one can easily observe that the following algorithm **NonBiconnected** (G, D) finds an optimal drawing of G .

NonBiconnected (G, D) ;

begin

Find an optimal drawing D_1 of B_1 and an optimal drawing D_p of B_p
by **Biconnected**;

Find an optimal $(3\pi/2, t_1)$ -outer drawing D_1° of B_1 and an optimal $(3\pi/2, s_p)$ -outer drawing D_p° of B_p by **OuterDrawing**;

Find an optimal I-shaped drawing D_{intI} of G_{int} ; {cf. Lemmas 3.1 and 3.2}

Let F_1 be a face of D_1 such that the angle of t_1 in F_1 is π or $3\pi/2$, insert D_{intI} and D_p° in F_1 , and let D_a be the resulting drawing of G , as illustrated in Figure 4.3(a);

Similarly construct a drawing D_b of G , as illustrated in Figure 4.3(b);

if $\text{bend}(D_a) \leq \text{bend}(D_b)$, **then return** D_a as D

else return D_b as D ;

end.

One can easily observe that **NonBiconnected** finds an optimal drawing of G in linear time. We thus have the following theorem.

THEOREM 4.3. *An optimal orthogonal drawing of a series-parallel graph G of $\Delta \leq 3$ can be found in linear time.*

4.2. Upper bounds on bends. In this subsection we prove the following theorem.

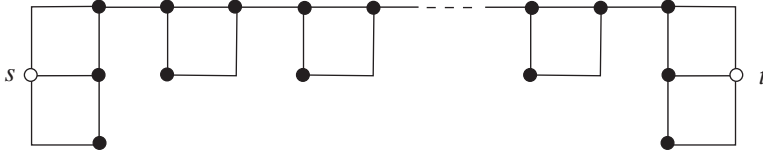
THEOREM 4.4. *If G is a series-parallel graph of $\Delta \leq 3$, then $\text{bend}(G) \leq (n + 4)/3$.*

The bound in Theorem 4.4 is best possible, because $\text{bend}(G) = (n + 4)/3$ for a series-parallel graph G consisting of two copies of $K_4 - e$ and several copies of K_2 and K_3 connected in series, as illustrated in Figure 4.5, where $K_4 - e$ is a graph obtained from K_4 by deleting an edge e .

We first have the following lemma.

LEMMA 4.5. *If a biconnected series-parallel graph B has a vertex v of degree two, then*

$$\text{bend}(B, 3\pi/2, v) \leq (n(B) + 2)/3.$$


 FIG. 4.5. A series-parallel graph G with $(n+4)/3$ bends.

Proof. Since B is biconnected, we have $n(B) \geq 3$. Split the vertex v into two vertices s' and t' , and let B_v be the resulting 2-legged series-parallel graph, as illustrated in Figure 4.6.

Consider first the case where B_v is a diamond graph. Then clearly $n(B) \geq 6$. By Lemma 3.1(a), B_v has a no-bend L-shaped drawing D_{vL} . From D_{vL} , one can construct a $(3\pi/2, v)$ -outer drawing $D_{3\pi/2}$ of B with two bends. (See Figures 3.8(g) and (h).) Therefore $\text{bend}(B, 3\pi/2, v) \leq \text{bend}(D_{3\pi/2}) = 2 \leq (n(B) + 2)/3$.

Consider next the case where B_v is not a diamond graph. By Lemma 3.2(a), B_v has an optimal U-shaped drawing D_{vU} , and $\text{bend}(D_{vU}) = \text{bend}(B_v)$. From D_{vU} , one can construct a $(3\pi/2, v)$ -outer drawing $D_{3\pi/2}$ of B with a new bend, and hence $\text{bend}(D_{3\pi/2}) = \text{bend}(D_{vU}) + 1$. (See Figures 3.8(e) and (f).) By Lemma 3.2(c), $\text{bend}(B_v) \leq (n(B_v) - 2)/3$. Clearly $n(B) = n(B_v) - 1$. We thus have

$$\begin{aligned} \text{bend}(B, 3\pi/2, v) &\leq \text{bend}(D_{3\pi/2}) \\ &= \text{bend}(D_{vU}) + 1 \\ &= \text{bend}(B_v) + 1 \\ &\leq (n(B_v) - 2)/3 + 1 \\ &= (n(B) + 2)/3. \quad \square \end{aligned}$$

We are now ready to present a proof of Theorem 4.4.

Proof of Theorem 4.4. By Lemma 3.7, $\text{bend}(G) \leq \lceil n/3 \rceil < (n+4)/3$ for every series-parallel biconnected graph G . We may thus assume that G is not biconnected. Then one may assume that G consists of biconnected components $B_1, B_2, \dots, B_p$, $p \geq 2$, and several copies of K_2 , alternately connected in series, as illustrated in Figure 4.1(a). **NonBiconnected** finds an optimal drawing D of G such that $\text{bend}(D) \leq \text{bend}(D_a) \leq \text{bend}(B_1) + \text{bend}(G_{\text{int}}) + \text{bend}(B_p, 3\pi/2, s_p)$.

By Lemma 3.7 we have $\text{bend}(B_1) \leq \lceil n(B_1)/3 \rceil \leq (n(B_1) + 2)/3$. We have $\text{bend}(G_{\text{int}}) \leq (n(G_{\text{int}}) - 2)/3$; if $n(G_{\text{int}}) = 2$, then $\text{bend}(G_{\text{int}}) = 0 = (n(G_{\text{int}}) - 2)/3$; if $n(G_{\text{int}}) \geq 3$ and G_{int} is a diamond graph, then by Lemma 3.1(a) $\text{bend}(G_{\text{int}}) = 0 < (n(G_{\text{int}}) - 2)/3$; if $n(G_{\text{int}}) \geq 3$ and G_{int} is not a diamond graph, then by Lemma 3.2(c) $\text{bend}(G_{\text{int}}) \leq (n(G_{\text{int}}) - 2)/3$. By Lemma 4.5 we have $\text{bend}(B_p, 3\pi/2, s_p) \leq (n(B_p) + 2)/3$. Clearly $n(G) = n(B_1) + n(G_{\text{int}}) + n(B_p) - 2$. We thus have

$$\begin{aligned} \text{bend}(G) &\leq (n(B_1) + 2)/3 + (n(G_{\text{int}}) - 2)/3 + (n(B_p) + 2)/3 \\ &\leq (n(G) + 4)/3. \quad \square \end{aligned}$$

One can easily observe that each edge has at most one bend in a drawing found by our algorithm.

4.3. Proof of Lemma 4.2. In this subsection, we first give a proof of Lemma 4.2, then present algorithm **OuterDrawing** to find an optimal $(3\pi/2, v)$ -outer drawing, and finally prove (4.2).

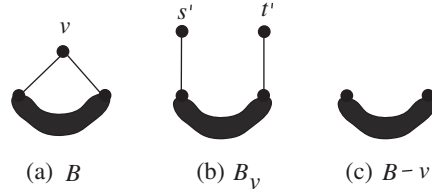


FIG. 4.6. (a) Biconnected graph B , (b) 2-legged graph B_v , and (c) graph $B - v$.

Let B be a biconnected series-parallel graph with $\Delta(B) \leq 3$, and let v be a vertex of degree two in B , as illustrated in Figure 4.6(a). Split the vertex v into two vertices s' and t' , and let B_v be the resulting 2-legged series-parallel graph with terminals s' and t' , as illustrated in Figure 4.6(b).

Consider first the case where B_v is a diamond graph. Then we have the following lemma.

LEMMA 4.6. *If B_v is a diamond graph, then $\text{bend}(B, 3\pi/2, v) = \text{bend}(B) = 2$.*

Proof. Since B_v is a diamond graph, by Lemma 3.1 B_v has a no-bend L-shaped drawing D_L , from which one can construct a $(3\pi/2, v)$ -outer drawing $D_{3\pi/2}$ of B with two bends (see Figures 3.8(g) and (h)), and hence

$$\text{bend}(B) \leq \text{bend}(B, 3\pi/2, v) \leq \text{bend}(D_{3\pi/2}) = 2.$$

Thus it suffices to verify $2 \leq \text{bend}(B)$.

Let D^* be an arbitrary optimal drawing of B . Let D_v be a drawing of B_v obtained from D^* without introducing any new bend as follows: erase, from D^* , point v together with very short segments of two edges incident to vertex v , and regard the ends of the erased segments other than v as vertices s' and t' of B_v . We then have

$$(4.4) \quad \text{bend}(D_v) = \text{bend}(D^*) = \text{bend}(B).$$

The construction of D_v implies that D_v has none of I-, L-, and U-shapes. We now claim $2 \leq \text{bend}(D_v)$; if $\text{bend}(D_v) = 0$, then by Lemma 3.1(c), D_v must be I-shaped or L-shaped, a contradiction; if $\text{bend}(D_v) = 1$, then one can easily observe from the example in Figure 2.6 that D_v is I-, L-, or U-shaped, a contradiction. Thus we have $2 \leq \text{bend}(D_v) = \text{bend}(B)$. $\square$

The proof of Lemma 4.6 immediately yields the following linear algorithm **OuterDiamond**(B, v, D) to find an optimal $(3\pi/2, v)$ -outer drawing D of B , with $\text{bend}(D) = \text{bend}(B) = 2$ if B_v is a diamond graph.

OuterDiamond(B, v, D);

begin

Find a no-bend L-shaped drawing D_{vL} of B_v by **Diamond**;

Construct from D_{vL} a $(3\pi/2, v)$ -outer drawing D of B with two bends;

(See Figures 3.8(g) and (h).)

end.

Consider next the case where $B - v$ is a series connection of subgraphs. We then have the following lemma.

LEMMA 4.7. *If $B - v$ is a series connection of subgraphs, then $\text{bend}(B, 3\pi/2, v) = \text{bend}(B)$.*

Proof. Since $\Delta(B) \leq 3$ and $B - v$ is a series connection of subgraphs, B_v has a bridge e other than the two legs, as illustrated in Figure 4.7(a). B_v has two subgraphs

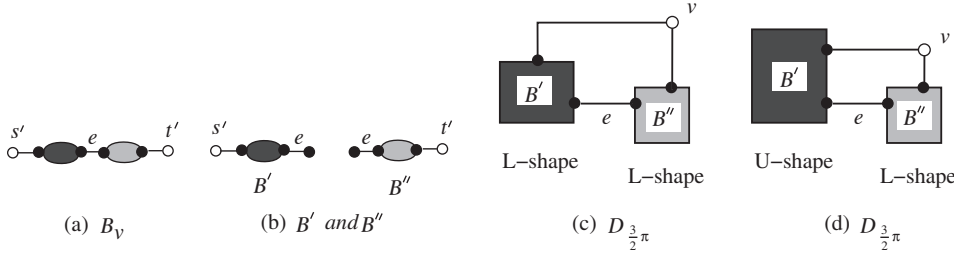


FIG. 4.7. Illustration of the proof of Lemma 4.7.

B' and B'' , each of which is a 2-legged series-parallel graph and has e as a leg, as illustrated in Figure 4.7(b). Clearly $n(B'), n(B'') \geq 3$.

Consider first the case where both B' and B'' are diamond graphs. Then by Lemma 3.1(a), B' and B'' have no-bend L-shaped drawings D'_L and D''_L , respectively. Merging D'_L and D''_L , one can construct a $(3\pi/2, v)$ -outer drawing $D_{3\pi/2}$ of B with one bend, as illustrated in Figure 4.7(c), and hence $\text{bend}(B, 3\pi/2, v) \leq 1$. By Lemma 3.1(c), any no-bend drawings of B' and B'' are not U-shaped but are I-shaped or L-shaped. Therefore, B does not have a no-bend drawing. Thus $D_{3\pi/2}$ is an optimal drawing of B , and hence $\text{bend}(B) = \text{bend}(B, 3\pi/2, v) = 1$.

Consider next the case of either B' or B'' is not a diamond graph. Assume without loss of generality that B' is not a diamond graph. Then by Lemma 3.2(a), B' has an optimal U-shaped drawing D'_U . On the other hand, B'' has an optimal L-shaped drawing D''_L ; if B'' is a diamond graph, then by Lemma 3.1(a), B'' has an optimal (no-bend) L-shaped drawing; otherwise, by Lemma 3.2(a), B'' has an optimal L-shaped drawing. Merging D'_U and D''_L , one can construct an optimal $(3\pi/2, v)$ -outer drawing $D_{3\pi/2}$ of B without introducing any new bend, as illustrated in Figure 4.7(d). Clearly $D_{3\pi/2}$ is an optimal drawing of B , and hence $\text{bend}(B, 3\pi/2, v) = \text{bend}(B)$. $\square$

The proof of Lemma 4.7 immediately yields the following linear algorithm **OuterSeries**(B, v, D) to find an optimal $(3\pi/2, v)$ -outer drawing D of B , with $\text{bend}(D) = \text{bend}(B)$ if $B - v$ is a series connection of subgraphs.

OuterSeries(B, v, D);

begin

 Define B' and B'' as in Figures 4.7(a) and (b);

if both B' and B'' are diamond graphs,

then

 Find no-bend L-shaped drawings D'_L of B' and D''_L of B'' by **Diamond**;

 Construct a drawing D with one bend from D'_L and D''_L as in Figure 4.7(c);

else

 Assume that B' is not a diamond graph;

 Find an optimal U-shaped drawing D'_U of B' by **NonDiamond**;

 Find an optimal L-shaped drawing D''_L of B'' by **Diamond**

or NonDiamond;

 Construct D from D'_U and D''_L as in Figure 4.7(d);

end.

We shall thus consider the remaining case where $B - v$ is a parallel connection of subgraph. Roughly speaking, for the case, we reduce the problem of finding an optimal $(3\pi/2, v)$ -outer drawing of B to those for smaller graphs B'_{id} and B''_{id} illustrated in Figure 4.8. The detail is given in the following proof of Lemma 4.2(a).

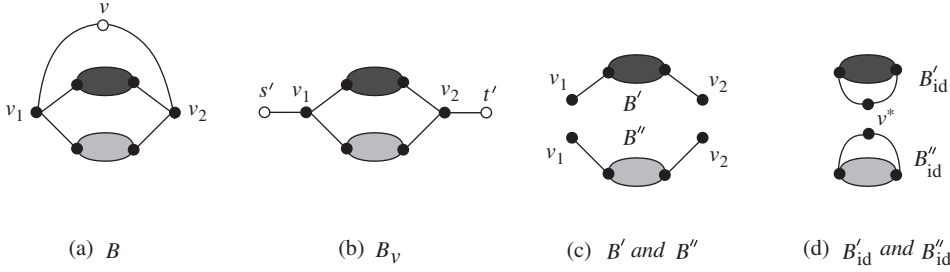


FIG. 4.8. Illustration of the proof of Lemma 4.2.

Proof of Lemma 4.2(a). We prove (4.3) by induction on the number $n(B)$ of vertices of B . If $n(B) = 3$, that is, $B = K_3$, then $\text{bend}(B, 3\pi/2, v) = 1$, $\text{bend}(B, \pi, v) = 2$, and $\text{bend}(B, \pi/2, v) = 3$, and hence (4.3) holds. One may thus assume that $n(B) \geq 4$, and inductively assume that (4.3) holds for every biconnected series-parallel graph of at most $n(B) - 1$ vertices. We now prove (4.3) for B . By Lemmas 4.6 and 4.7, one may assume that B_v is not a diamond graph and $B - v$ is a parallel connection of subgraphs. Then $B - v$ is biconnected, and the two vertices v_1 and v_2 that are adjacent to v in B have degree three in B , as illustrated in Figure 4.8(a). Splitting vertices v_1 and v_2 of $B - v$, one can obtain two edge-disjoint series-parallel subgraphs B' and B'' of B , as illustrated in Figure 4.8(c). Clearly $n(B'), n(B'') \geq 2$.

We prove only $\text{bend}(B, 3\pi/2, v) \leq \text{bend}(B, \pi, v)$, because one can similarly prove $\text{bend}(B, 3\pi/2, v) \leq \text{bend}(B, \pi/2, v)$. Let D_π^* be an optimal (π, v) -outer drawing of B , then $\text{bend}(D_\pi^*) = \text{bend}(B, \pi, v)$. It suffices to prove that $\text{bend}(B, 3\pi/2, v) \leq \text{bend}(D_\pi^*)$. There are the following three cases to consider.

Case 1: D_π^* has a bend on edge vv_1 or vv_2 .

Erasing from D_π^* a line segment connecting v and a bend in the edge, one can obtain a drawing D_v of B_v . Since a bend in D_π^* is regarded as a vertex in D_v , we have $\text{bend}(B_v) \leq \text{bend}(D_v) = \text{bend}(D_\pi^*) - 1$.

Since B_v is not a diamond graph, by Lemma 3.2(a) B_v has an optimal U-shape drawing D_{vU} . From D_{vU} , one can construct a $(3\pi/2, v)$ -outer drawing $D_{3\pi/2}$ of B with a new bend. (See Figures 3.8(e) and (f).) Thus we have

$$\begin{aligned} \text{bend}(B, 3\pi/2, v) &\leq \text{bend}(D_{3\pi/2}) \\ &= \text{bend}(D_{vU}) + 1 \\ &= \text{bend}(B_v) + 1 \\ &\leq \text{bend}(D_\pi^*). \end{aligned}$$

Case 2: D_π^* has no bend on edges vv_1 and vv_2 , and either $n(B') = 2$ or $n(B'') = 2$.

One may assume without loss of generality that $n(B') = 2$, and hence B' consists of a single edge v_1v_2 . Since D_π^* has no bend on the edges v_1v and vv_2 , they are drawn on the same straight line segment in D_π^* . Therefore the edge v_1v_2 in B' must have two or more bends in D_π^* , and hence $\text{bend}(D_\pi^*[B']) \geq 2$. Since B has no multiple edges, $n(B'') \geq 3$ and B'' is a 2-legged series-parallel graph.

If B'' is a diamond graph, then $\text{bend}(B'') = 0$ and B'' has a U-shaped drawing D_U'' with one bend as known from the example in Figure 2.6. If B'' is not a diamond graph, then by Lemma 3.2(a), B'' has an optimal U-shaped drawing D_U'' . In either case $\text{bend}(D_U'') \leq \text{bend}(B'') + 1$. One can easily extend D_U'' to a $(3\pi/2, v)$ -outer drawing

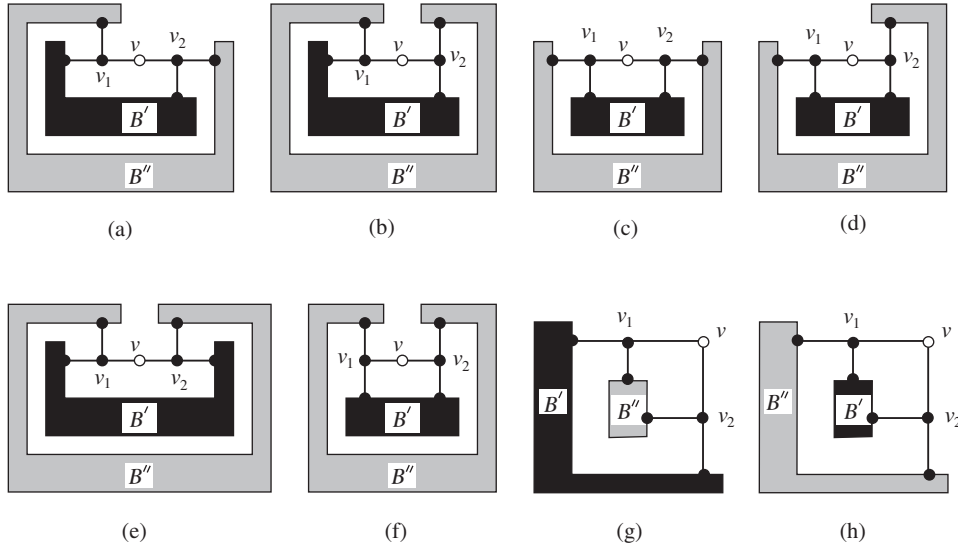


FIG. 4.9. (a)–(f) All of the possible (π, v) -outer drawings D_π^* , and (g), (h) constructions of $D_{3\pi/2}$.

$D_{3\pi/2}$ with a new bend. We thus have

$$\begin{aligned}
 \text{bend}(B, 3\pi/2, v) &\leq \text{bend}(D_{3\pi/2}) \\
 &= 1 + \text{bend}(D_U'') \\
 &\leq 1 + \text{bend}(B'') + 1 \\
 &\leq 2 + \text{bend}(D_\pi^*|B'') \\
 &\leq \text{bend}(D_\pi^*|B') + \text{bend}(D_\pi^*|B'') \\
 &= \text{bend}(D_\pi^*).
 \end{aligned}$$

Case 3: Otherwise.

In this case, $n(B'), n(B'') \geq 3$, and hence both B' and B'' are 2-legged series-parallel graphs. Identify v_1 with v_2 as a new vertex v^* in B' , and let B'_{id} be the resulting graph, as illustrated in Figure 4.8(d). Similarly define B''_{id} . Then both B'_{id} and B''_{id} are biconnected series-parallel graphs and have a vertex v^* of degree two. Since D_π^* has no bend on the two edges v_1v and vv_2 , all of the possible (π, v) -outer drawings of B are those illustrated in Figures 4.9(a)–(f) after interchanging the roles of B' and B'' and the roles of v_1 and v_2 , depending on the angles of v_1 and v_2 in D_π^* . There are the following two subcases to consider.

Case 3.1: D_π^* is a drawing illustrated in Figure 4.9(a) or (b).

In this case, one can construct from $D_\pi^*|B'$ a $(3\pi/2, v^*)$ -outer drawing of B'_{id} with no new bends, and hence $\text{bend}(B'_{\text{id}}, 3\pi/2, v^*) \leq \text{bend}(D_\pi^*|B')$. Let $D'_{\text{id}3\pi/2}$ be an optimal $(3\pi/2, v^*)$ -outer drawing of B'_{id} . Since B'' is 2-legged, B'' has an optimal L-shaped drawing D''_L . Merging $D'_{\text{id}3\pi/2}$ and D''_L , one can easily construct a $(3\pi/2, v)$ -

outer drawing $D_{3\pi/2}$ of B , as illustrated in Figure 4.9(g). We thus have

$$\begin{aligned} \text{bend}(B, 3\pi/2, v) &\leq \text{bend}(D_{3\pi/2}) \\ &= \text{bend}(D'_{\text{id}3\pi/2}) + \text{bend}(D'_L) \\ &= \text{bend}(B'_{\text{id}}, 3\pi/2, v^*) + \text{bend}(B'') \\ &\leq \text{bend}(D_\pi^*|B') + \text{bend}(D_\pi^*|B'') \\ &= \text{bend}(D_\pi^*). \end{aligned}$$

Case 3.2: D_π^* is a drawing illustrated in Figures 4.9(c)–(f).

We first prove that if D_π^* is a drawing in Figure 4.9(c) or (d), then

$$(4.5) \quad \text{bend}(B''_{\text{id}}, 3\pi/2, v^*) \leq \text{bend}(D_\pi^*|B'').$$

One can construct from $D_\pi^*|B''$ an (α, v^*) -outer drawing of B''_{id} with no new bend, where $\alpha = \pi$ for Figure 4.9(c) and $\alpha = \pi/2$ for Figure 4.9(d). We thus have $\text{bend}(B''_{\text{id}}, \alpha, v^*) \leq \text{bend}(D_\pi^*|B'')$. Since B''_{id} is a biconnected series-parallel graph with $d(v^*, B''_{\text{id}}) = 2$ and $n(B''_{\text{id}}) < n(B)$, by the inductive hypothesis we have $\text{bend}(B''_{\text{id}}, 3\pi/2, v^*) \leq \text{bend}(B''_{\text{id}}, \alpha, v^*)$. Thus we have (4.5).

We also have the following fact, whose proof will be given in section 4.4.

FACT 1. *If D_π^* is a drawing illustrated in Figures 4.9(e) and (f), then $\text{bend}(B''_{\text{id}}, 3\pi/2, v^*) \leq \text{bend}(D_\pi^*|B'')$.*

Let $D'_{\text{id}3\pi/2}$ be an optimal $(3\pi/2, v^*)$ -outer drawing of B''_{id} . Let D'_L be an optimal L-shaped drawing of B' . Merging D'_L and $D'_{\text{id}3\pi/2}$, one can easily construct a $(3\pi/2, v)$ -outer drawing $D_{3\pi/2}$ of B , as illustrated in Figure 4.9(h). Thus, using (4.5) and Fact 1 we have

$$\begin{aligned} \text{bend}(B, 3\pi/2, v) &\leq \text{bend}(D_{3\pi/2}) \\ &= \text{bend}(D'_L) + \text{bend}(D'_{\text{id}3\pi/2}) \\ &= \text{bend}(B') + \text{bend}(B''_{\text{id}}, 3\pi/2, v^*) \\ &\leq \text{bend}(D_\pi^*|B') + \text{bend}(D_\pi^*|B'') \\ &= \text{bend}(D_\pi^*). \quad \square \end{aligned}$$

We then have the following lemma.

LEMMA 4.8. *Suppose that B_v is not a diamond graph and that $B - v$ is a parallel connection of subgraphs. (See Figure 4.8.) Then the following (a) and (b) hold:*

- (a) *if $n(B') = 2$ or $n(B'') = 2$, then $\text{bend}(B, 3\pi/2, v) = \text{bend}(B_v) + 1$; and*
- (b) *if $n(B'), n(B'') \geq 3$, then*

$$(4.6) \quad \begin{aligned} \text{bend}(B, 3\pi/2, v) &= \min\{\text{bend}(B_v) + 1, \\ &\quad \text{bend}(B'_{\text{id}}, 3\pi/2, v^*) + \text{bend}(B''), \\ &\quad \text{bend}(B') + \text{bend}(B''_{\text{id}}, 3\pi/2, v^*)\}. \end{aligned}$$

Proof. We first prove

$$(4.7) \quad \text{bend}(B, 3\pi/2, v) \leq \text{bend}(B_v) + 1.$$

Since $n(B) \geq 3$, $n(B_v) \geq 4$. Since B_v is not a diamond graph, by Lemma 3.2(a), B_v has an optimal U-shaped drawing D_{vU} . From D_{vU} , one can easily construct a

$(3\pi/2, v)$ -outer drawing $D_{3\pi/2}$ of B with a new bend. (See Figures 3.8(e) and (f).) Therefore,

$$\begin{aligned} \text{bend}(B, 3\pi/2, v) &\leq \text{bend}(D_{3\pi/2}) \\ &= \text{bend}(D_{vU}) + 1 \\ &= \text{bend}(B_v) + 1. \end{aligned}$$

(a) We then prove (a). By (4.7), it suffices to prove

$$(4.8) \quad \text{bend}(B, 3\pi/2, v) \geq \text{bend}(B_v) + 1.$$

Since B has no multiple edges and either $n(B') = 2$ or $n(B'') = 2$, one may assume without loss of generality that $n(B') = 2$, and hence $n(B'') \geq 3$. Let $D_{3\pi/2}^*$ be an optimal $(3\pi/2, v)$ -outer drawing of B , then $\text{bend}(B, 3\pi/2, v) = \text{bend}(D_{3\pi/2}^*)$. If $D_{3\pi/2}^*$ has a bend on an edge incident to v , then $\text{bend}(D_{3\pi/2}^*) - 1 \geq \text{bend}(B_v)$, and hence $\text{bend}(B, 3\pi/2, v) = \text{bend}(D_{3\pi/2}^*) \geq \text{bend}(B_v) + 1$. One may thus assume that $D_{3\pi/2}^*$ has no bend on the two edges of B incident to v . Then $D_{3\pi/2}^*$ has one or more bends on the edge in B' , and hence $\text{bend}(D_{3\pi/2}^*) \geq 1$. Let $\overline{B''}$ be the complement of B'' in B . Then $\overline{B''}$ is $K_3 = v_1vv_2$, and hence

$$(4.9) \quad \text{bend}(B_{3\pi/2}^* | \overline{B''}) \geq \text{bend}(\overline{B''}) \geq 1.$$

By (2.2) we have

$$(4.10) \quad \text{bend}(D_{3\pi/2}^*) = \text{bend}(D_{3\pi/2}^* | B'') + \text{bend}(D_{3\pi/2}^* | \overline{B''}).$$

There are the following two cases to consider.

Case 1: B'' is a diamond graph.

We first claim that $\text{bend}(D_{3\pi/2}^*) \geq 2$. Suppose for a contradiction that $\text{bend}(D_{3\pi/2}^*) = 1$. Then by (4.9) and (4.10) we have

$$\text{bend}(D_{3\pi/2}^* | \overline{B''}) = 1$$

and

$$\text{bend}(D_{3\pi/2}^* | B'') = 0.$$

Therefore, by Lemma 3.1(c), the no-bend drawing $D_{3\pi/2}^* | B''$ of B'' must be I- or L-shaped. Since $D_{3\pi/2}^*$ has no bend on edges vv_1 and vv_2 , $D_{3\pi/2}^* | B''$ must be L-shaped. Then $D_{3\pi/2}^*$ must have three or more bends on the edge v_1v_2 in B' , and hence $\text{bend}(D_{3\pi/2}^* | \overline{B''}) \geq 3$, contrary to $\text{bend}(D_{3\pi/2}^* | \overline{B''}) = 1$.

Since B'' is a diamond graph, by Lemma 3.1(a), B'' has a no-bend L-shaped drawing D_L'' . From D_L'' , one can easily construct a drawing D_v of B_v with one bend, as illustrated in Figure 4.10(a), and hence $\text{bend}(B_v) \leq \text{bend}(D_v) = 1$. Thus we have $\text{bend}(B, 3\pi/2, v) = \text{bend}(D_{3\pi/2}^*) \geq 2 \geq \text{bend}(B_v) + 1$.

Case 2: B'' is not a diamond graph.

Since $\text{bend}(D_{3\pi/2}^* | B'') \geq \text{bend}(B'')$, by (4.9) and (4.10) we have

$$(4.11) \quad \text{bend}(D_{3\pi/2}^*) \geq \text{bend}(B'') + 1.$$

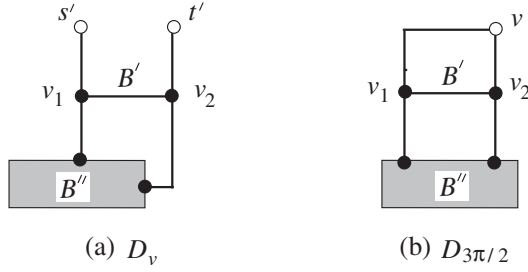


FIG. 4.10. Illustration for the proof of Lemma 4.8(a).

Since B'' is not a diamond graph, by Lemma 3.2(a), B'' has an optimal U-shaped drawing D''_U . From D''_U , one can easily construct a $(3\pi/2, v)$ -outer drawing $D_{3\pi/2}$ of B with one new bend, as illustrated in Figure 4.10(b). We then have

$$\begin{aligned}
 \text{bend}(D_{3\pi/2}) &= \text{bend}(D''_U) + 1 \\
 (4.12) \qquad \qquad &= \text{bend}(B'') + 1.
 \end{aligned}$$

Erasing from $D_{3\pi/2}$ the line segment connecting v and a bend, one can obtain a drawing D_v of B_v , and hence

$$\begin{aligned}
 \text{bend}(B_v) &\leq \text{bend}(D_v) \\
 (4.13) \qquad \qquad &= \text{bend}(D_{3\pi/2}) - 1.
 \end{aligned}$$

By (4.11)–(4.13) we have

$$\begin{aligned}
 \text{bend}(B, 3\pi/2, v) &= \text{bend}(D_{3\pi/2}^*) \\
 &\geq \text{bend}(B'') + 1 \\
 &= \text{bend}(D_{3\pi/2}) \\
 &\geq \text{bend}(B_v) + 1.
 \end{aligned}$$

(b) We then prove (b). We first prove that

$$\begin{aligned}
 \text{bend}(B, 3\pi/2, v) &\leq \min\{\text{bend}(B_v) + 1, \\
 (4.14) \qquad \qquad &\text{bend}(B'_{\text{id}}, 3\pi/2, v^*) + \text{bend}(B''), \\
 &\text{bend}(B') + \text{bend}(B''_{\text{id}}, 3\pi/2, v^*)\}.
 \end{aligned}$$

Since (4.7) holds, it suffices to prove that

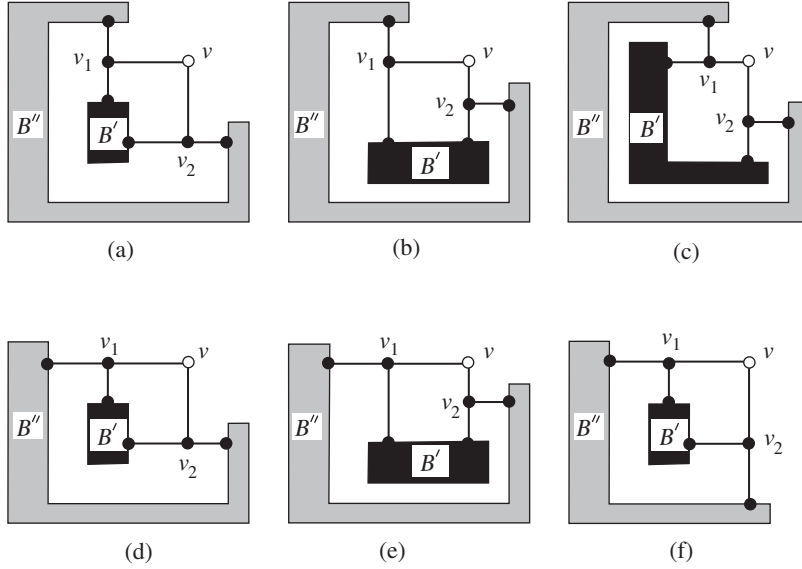
$$(4.15) \qquad \text{bend}(B, 3\pi/2, v) \leq \text{bend}(B'_{\text{id}}, 3\pi/2, v^*) + \text{bend}(B'')$$

and

$$(4.16) \qquad \text{bend}(B, 3\pi/2, v) \leq \text{bend}(B') + \text{bend}(B''_{\text{id}}, 3\pi/2, v^*).$$

We verify only (4.15), because one can similarly verify (4.16).

Let $D'_{\text{id}3\pi/2}$ be an optimal $(3\pi/2, v^*)$ -outer drawing of B'_{id} . Since $n(B'') \geq 3$, there is an optimal L-shaped drawing D''_L of B'' . Merging $D'_{\text{id}3\pi/2}$ and D''_L , one can


 FIG. 4.11. All of the possible $(3\pi/2, v)$ -outer drawings $D_{3\pi/2}^*$ of B .

easily construct a $(3\pi/2, v)$ -outer drawing $D_{3\pi/2}$ of B with no new bend, as illustrated in Figure 4.9(g). We thus have

$$\begin{aligned} \text{bend}(B, 3\pi/2, v) &\leq \text{bend}(D_{3\pi/2}) \\ &= \text{bend}(D'_{\text{id}, 3\pi/2}) + \text{bend}(D''_{\text{L}}) \\ &= \text{bend}(B'_{\text{id}}, 3\pi/2, v^*) + \text{bend}(B''). \end{aligned}$$

We next prove that

$$\begin{aligned} \text{bend}(B, 3\pi/2, v) &\geq \min\{\text{bend}(B_v) + 1, \\ &\quad \text{bend}(B'_{\text{id}}, 3\pi/2, v^*) + \text{bend}(B''), \\ &\quad \text{bend}(B') + \text{bend}(B'_{\text{id}}, 3\pi/2, v^*)\}. \end{aligned} \quad (4.17)$$

Let $D_{3\pi/2}^*$ be an optimal $(3\pi/2, v)$ -outer drawing of B , then $\text{bend}(B, 3\pi/2, v) = \text{bend}(D_{3\pi/2}^*)$. If $D_{3\pi/2}^*$ has a bend on an edge incident to v , then $\text{bend}(D_{3\pi/2}^*) - 1 \geq \text{bend}(B_v)$, and hence $\text{bend}(B, 3\pi/2, v) = \text{bend}(D_{3\pi/2}^*) \geq \text{bend}(B_v) + 1$ and (4.17) holds. One may thus assume that $D_{3\pi/2}^*$ has no bend on the two edges of B incident to v . Then all of the possible $(3\pi/2, v)$ -outer drawings of B are those illustrated in Figures 4.11(a)–(f) after interchanging the roles of B' and B'' and the roles of v_1 and v_2 .

We now claim

$$\text{bend}(B''_{\text{id}}, 3\pi/2, v^*) \leq \text{bend}(D_{3\pi/2}^*|B''). \quad (4.18)$$

If $D_{3\pi/2}^*$ is a drawing illustrated in Figure 4.11(f), then from $D_{3\pi/2}^*|B''$, one can easily construct a $(3\pi/2, v^*)$ -outer drawing of B''_{id} with no new bend, and hence (4.18) holds. One may thus assume that $D_{3\pi/2}^*$ is a drawing illustrated in Figures 4.11(a)–(e). Then one can construct from $D_{3\pi/2}^*|B''$ an (α, v^*) -outer drawing of B''_{id} with no new bend, where $\alpha = \pi/2$ for Figures 4.11(a)–(c) and $\alpha = \pi$ for Figures 4.11(d) and (e). We thus

have $\text{bend}(B''_{\text{id}}, \alpha, v^*) \leq \text{bend}(D^*_{3\pi/2}|B'')$. By Lemma 4.2(a), $\text{bend}(B''_{\text{id}}, 3\pi/2, v^*) \leq \text{bend}(B''_{\text{id}}, \alpha, v^*)$. These two equations imply (4.18).

By (4.18) we have

$$\begin{aligned} \text{bend}(B, 3\pi/2, v) &= \text{bend}(D^*_{3\pi/2}) \\ &= \text{bend}(D^*_{3\pi/2}|B') + \text{bend}(D^*_{3\pi/2}|B'') \\ &\geq \text{bend}(B') + \text{bend}(B''_{\text{id}}, 3\pi/2, v^*), \end{aligned}$$

and hence (4.17) holds true. $\square$

By Lemmas 4.6, 4.7, and 4.8 and the proofs of (4.7) and (4.14) we have the following recursive algorithm **OuterDrawing**($B, v, D, D_{vI}, D_{vL}, D_{vU}$) to find an optimal $(3\pi/2, v)$ -outer drawing D of a biconnected series-parallel graph B with a vertex v of degree two. It finds an optimal I-shaped drawing D_{vI} , an optimal L-shaped drawing D_{vL} , and an optimal U-shaped drawing D_{vU} of B_v , too.

OuterDrawing($B, v, D, D_{vI}, D_{vL}, D_{vU}$);

begin

Case 1: B_v is a diamond graph.

OuterDiamond(B, v, D); {cf. Lemma 4.6}

Find a no-bend I-shaped drawing D_{vI} and a no-bend L-shaped drawing D_{vL} of B_v by **Diamond**;

Insert a dummy vertex of degree two in an edge of B_v , as illustrated in Figure 2.6(e);

Obtain a U-shaped drawing D_{vU} of B_v with one bend by regarding the dummy vertex as a bend;

Case 2: B_v is not a diamond graph, and $B - v$ is a series connection of subgraphs.

OuterSeries(B, v, D); {cf. Lemma 4.7}

Find an optimal I-shaped drawing D_{vI} , an optimal L-shaped drawing D_{vL} , and an optimal U-shaped drawing D_{vU} of B_v by **NonDiamond**;

Case 3: B_v is not a diamond graph, and $B - v$ is a parallel connection of subgraphs.

{cf. the proofs of (4.7) and (4.14).}

Case 3.1: Either $n(B') = 2$ or $n(B'') = 2$.

Find an optimal I-shaped drawing D_{vI} , an optimal L-shaped drawing D_{vL} , and an optimal U-shaped drawing D_{vU} of B_v by **NonDiamond**;

Extend D_{vU} to a $(3\pi/2, v)$ -outer drawing D of B with a new bend;

{See Figures 3.8(e) and (f).}

Case 3.2: $n(B'), n(B'') \geq 3$.

{One may assume without loss of generality that B' is not a diamond graph.}

OuterDrawing($B'_{\text{id}}, v^*, D'_{\text{id}}, D'_{v^*I}, D'_{v^*L}, D'_{v^*U}$);

OuterDrawing($B''_{\text{id}}, v^*, D''_{\text{id}}, D''_{v^*I}, D''_{v^*L}, D''_{v^*U}$);

Combine a U-shaped drawing D'_{v^*U} of $B' (= B'_{\text{id}v^*})$ and an I-shaped drawing D''_{v^*I} of $B'' (= B''_{\text{id}v^*})$ to an optimal I-shaped drawing D_{vI} , an optimal L-shaped drawing D_{vL} , and an optimal U-shaped drawing D_{vU} of B_v , as illustrated in Figure 3.4(e);

{cf. Lemma 3.2}

Let $x = \min\{\text{bend}(D_{vU}) + 1, \text{bend}(D'_{\text{id}}) + \text{bend}(D''_{\text{L}}), \text{bend}(D'_{\text{L}}) + \text{bend}(D''_{\text{id}})\}$, where $D'_{\text{L}} = D'_{v^*L}$ and $D''_{\text{L}} = D''_{v^*L}$ are optimal L-shaped drawings of $B' = B'_{\text{id}v^*}$ and $B'' = B''_{\text{id}v^*}$, respectively;

Case 3.2.1: $x = \text{bend}(D_{vU}) + 1$.

Extend D_{vU} to a $(3\pi/2, v)$ -outer drawing D of B with a new bend;

{See Figure 3.8(e).}

Case 3.2.2: $x = \text{bend}(D'_{id}) + \text{bend}(D''_L)$.

Construct a $(3\pi/2, v)$ -outer drawing D of B by merging D'_{id} and D''_L , as illustrated in Figure 4.9(g);

Case 3.2.3: $x = \text{bend}(D'_L) + \text{bend}(D''_{id})$.

Construct a $(3\pi/2, v)$ -outer drawing D of B by merging D'_L and D''_{id} , as illustrated in Figure 4.9(h);

end.

Clearly **OuterDrawing** correctly finds an optimal $(3\pi/2, v)$ -outer drawing D of B . We now show that **OuterDrawing** runs in linear time. Let $T(B)$ be the execution time of **OuterDrawing** $(B, v, D, D_{vI}, D_{vL}, D_{vU})$. Let $T_1(B)$ be the execution time of **OuterDrawing** when **Case 1** occurs, that is, B_v is a diamond graph. We similarly define $T_2(B)$, $T_{3.1}(B)$, and $T_{3.2}(B)$ for **Cases 2, 3.1, and 3.2**, respectively. Then clearly

$$T(B) = \max\{T_1(B), T_2(B), T_{3.1}(B), T_{3.2}(B)\},$$

$$T_1(B), T_2(B), T_{3.1}(B) = O(n(B)),$$

$$T_{3.2}(B) = T(B'_{id}) + T(B''_{id}) + O(1),$$

and

$$n(T(B'_{id})) + n(T(B''_{id})) < n(B).$$

We hence have $T(B) = O(n(B))$. We have thus completed a proof of Lemma 4.2(c).

Since $\text{bend}(B_v) \leq \text{bend}(B)$, by Lemmas 4.6, 4.7, and 4.8, we immediately have $\text{bend}(B, 3\pi/2, v) \leq \text{bend}(B) + 1$, completing a proof of Lemma 4.2(b).

Let a series-parallel graph G take the form in Figure 4.1(a), and let D_a be a drawing of G , as illustrated in Figure 4.3(a). Then we have

$$\begin{aligned} \text{bend}(G) &\leq \text{bend}(D_a) \\ &= \text{bend}(B_1) + \text{bend}(G_{\text{int}}) + \text{bend}(B_p, 3\pi/2, s_p). \end{aligned}$$

Clearly

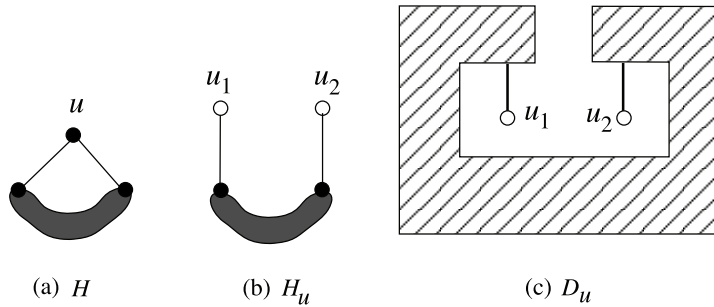
$$\begin{aligned} \text{bend}(G_{\text{int}}) &= \text{bend}(D_{\text{intI}}) \\ &= \sum_{i=2}^{p-1} \text{bend}(B_i) \end{aligned}$$

and

$$\text{bend}(B_p, 3\pi/2, s_p) \leq \text{bend}(B_p) + 1.$$

We thus have

$$(4.19) \quad \text{bend}(G) \leq \sum_{i=1}^p \text{bend}(B_i) + 1.$$

FIG. 4.12. Biconnected graph H , 2-legged graph H_u , and C-shaped drawing D_u .

Equations (4.1) and (4.19) immediately imply (4.2).

Omitting the operations in Cases 3.2.2 and 3.2.3 in **OuterDrawing**, one can obtain a nonrecursive algorithm **ApproOuterDrawing** to find a $(3\pi/2, v)$ -outer drawing D of B such that D is not always an optimal $(3\pi/2, v)$ -outer drawing, but $\text{bend}(D) \leq \text{bend}(B) + 1$. Replacing **OuterDrawing** in **NonBiconnected** by **ApproOuterDrawing** and finding only the drawing D_a in Figure 4.3, one can obtain a simple linear algorithm to find a drawing D of G such that

$$\text{bend}(D) \leq \sum_{i=1}^p \text{bend}(B_i) + 1.$$

4.4. Proof of Fact 1. In this subsection we give a proof of Fact 1 in Case 3.2 of the proof of Lemma 4.2(a).

A 2-legged series-parallel graph B'' and a biconnected series-parallel graph B''_{id} in Fact 1 are illustrated in Figures 4.8(c) and (d), respectively. B'' can be obtained from B''_{id} by splitting a vertex v^* of degree two to two vertices v_1 and v_2 . The drawing D_π^* in Fact 1 is illustrated in Figure 4.9(e) or (f). We call a drawing of a 2-legged series-parallel graph illustrated in Figure 4.12(c) a *C-shaped drawing*. Then $D_\pi^*[B''$ in Figure 4.9(e) or (f) is a C-shaped drawing of B'' . Thus it suffices to prove the following fact. (Regard B''_{id} as H , B'' as H_u , and v^* as u .)

FACT 2. Let H be a biconnected series-parallel graph with a vertex u of degree two, let H_u be a 2-legged series-parallel graph obtained from H by splitting u to two vertices u_1 and u_2 . If D_u is a C-shaped drawing of H_u , then

$$(4.20) \quad \text{bend}(H, 3\pi/2, u) \leq \text{bend}(D_u).$$

(See Figure 4.12.)

Proof. We prove (4.20) by induction on the number $n(H)$ of vertices of H . If $n(H) = 3$, that is, $H = K_3$, then $\text{bend}(H, 3\pi/2, v) = 1$, $\text{bend}(D_u) \geq 4$, and hence (4.20) holds true. One may thus assume that $n(H) \geq 4$, and inductively assume that (4.20) holds for every biconnected series-parallel graph of at most $n(H) - 1$ vertices. We now prove that (4.20) holds for H . There are the following two cases to consider.

Case 1: $H - u$ is a series connection of subgraphs.

In this case, H_u has a bridge e , and has two subgraphs H' and H'' , each of which is a 2-legged series-parallel graph and has e as a leg. (See Figures 4.7(a) and (b).) Then we have the following two subcases to consider.

Case 1.1: Both H' and H'' are diamond graphs.

In this case, H' and H'' have no-bend L-shaped drawings D'_L and D''_L , respectively. Merging D'_L and D''_L , one can easily construct a $(3\pi/2, u)$ -outer drawing $D_{3\pi/2}^*$ of H with one bend, and hence $\text{bend}(D_{3\pi/2}^*) = 1$. (See Figure 4.7(c).) Since both H' and H'' are diamond graphs, one can easily observe that $D_{3\pi/2}^*$ is an optimal $(3\pi/2, u)$ -outer drawing of H and that a C-shaped drawing D_u of H_u needs one or more bends. We thus have $\text{bend}(H, 3\pi/2, u) = \text{bend}(D_{3\pi/2}^*) = 1 \leq \text{bend}(D_u)$.

Case 1.2: Either H' or H'' is not a diamond graph.

In this case, one may assume without loss of generality that H' is not a diamond graph. Then H' has an optimal U-shaped drawing D'_U . H'' has an optimal L-shaped drawing D''_L . Merging D'_U and D''_L , one can easily construct a $(3\pi/2, u)$ -outer drawing $D_{3\pi/2}$ of H with no new bends. (See Figure 4.7(d).) Thus

$$\begin{aligned} \text{bend}(H, 3\pi/2, u) &\leq \text{bend}(D_{3\pi/2}) \\ &= \text{bend}(D'_U) + \text{bend}(D''_L) \\ &= \text{bend}(H') + \text{bend}(H'') \\ &\leq \text{bend}(D_u|H') + \text{bend}(D_u|H'') \\ &= \text{bend}(D_u). \end{aligned}$$

Case 2: $H - u$ is a parallel connection of subgraphs.

Consider first the case where H_u is a diamond graph. By Lemma 4.6, $\text{bend}(H, 3\pi/2, u) = \text{bend}(H)$. Since H_u has a U-shaped drawing with one bend, we have $\text{bend}(H) \leq 2$. Since D_u is a C-shaped drawing of a diamond graph H_u , one can easily observe that $\text{bend}(D_u) \geq 2$. We thus have

$$\text{bend}(H, 3\pi/2, u) = \text{bend}(H) \leq 2 \leq \text{bend}(D_u).$$

Consider next the case where H_u is not a diamond graph. Regarding H as B in Figure 4.8, we define H' , H'' , H'_{id} , and H''_{id} , as illustrated in Figure 4.8. Then we have the following three subcases.

Case 2.1: D_u has a bend on a leg of H_u .

In this case, erasing a line segment connecting a terminal and a bend from D_u , one can obtain a drawing $\widetilde{D}_u$ of H_u such that $\widetilde{D}_u$ is not always C-shaped and $\text{bend}(H_u) \leq \text{bend}(\widetilde{D}_u) = \text{bend}(D_u) - 1$. Since H_u is not a diamond graph, H_u has an optimal U-shape drawing D_{uU} . From D_{uU} , one can construct a $(3\pi/2, u)$ -outer drawing $D_{3\pi/2}$ of H with a new bend. (See Figures 3.8(e) and (f).) We thus have

$$\begin{aligned} \text{bend}(H, 3\pi/2, u) &\leq \text{bend}(D_{3\pi/2}) \\ &= \text{bend}(D_{uU}) + 1 \\ &= \text{bend}(H_u) + 1 \\ &\leq \text{bend}(\widetilde{D}_u) + 1 \\ &= \text{bend}(D_u). \end{aligned}$$

Case 2.2: D_u has no bend on the two legs of H_u , and either $n(H') = 2$ or $n(H'') = 2$.

One may assume without loss of generality that $n(H') = 2$, and hence H' consists of a single edge. Then H'' is a 2-legged series-parallel graph. Since the C-shaped drawing D_u has no bend on the two legs of H_u , we have $\text{bend}(D_u|H') \geq 4$.

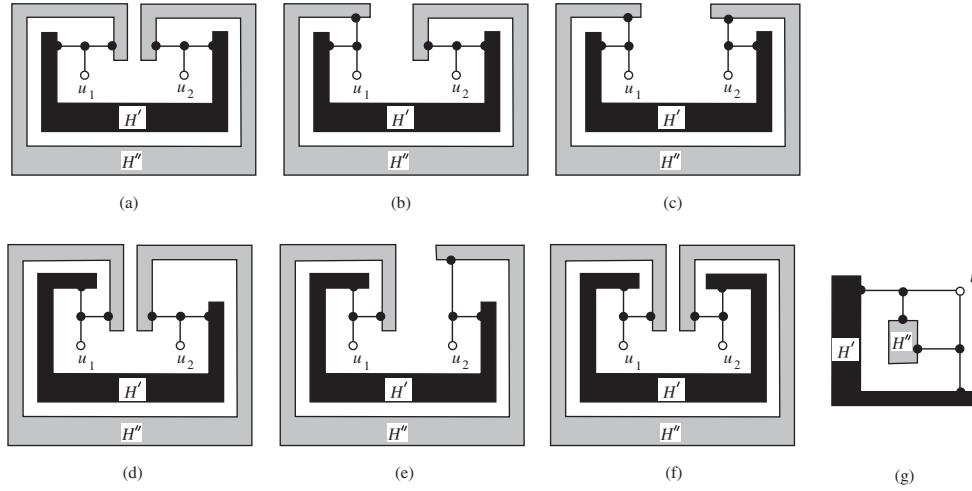


FIG. 4.13. (a)–(f) All of the possible C-shaped drawings D_u of H_u , and (g) construction of a $(3\pi/2, u)$ -outer drawing $D_{3\pi/2}$ of H .

Consider first the case where H'' is a diamond graph. Then H_u has a U-shaped drawing D_{uU} with one bend. From D_{uU} , one can construct a $(3\pi/2, u)$ -outer drawing $D_{3\pi/2}$ of H with two bends. Thus $\text{bend}(H, 3\pi/2, u) \leq \text{bend}(D_{3\pi/2}) = 2 < \text{bend}(D_u|H') \leq \text{bend}(D_u)$.

Consider next the case where H'' is not a diamond graph. Then by Lemma 3.2(a), H'' has an optimal U-shaped drawing D''_U . From D''_U , one can construct a $(3\pi/2, u)$ -outer drawing $D_{3\pi/2}$ of H with a new bend. We thus have

$$\begin{aligned}
 \text{bend}(H, 3\pi/2, u) &\leq \text{bend}(D_{3\pi/2}) \\
 &= 1 + \text{bend}(D''_U) \\
 &= 1 + \text{bend}(H'') \\
 &< \text{bend}(D_u|H') + \text{bend}(D_u|H'') \\
 &= \text{bend}(D_u).
 \end{aligned}$$

Case 2.3: D_u has no bend on the two legs of H_u , and $n(H'), n(H'') \geq 3$.

In this case H' and H'' are 2-legged series-parallel graphs. All of the possible C-shaped drawings D_u of H_u for this case are those illustrated in Figures 4.13(a)–(f).

Let $D'_{\text{id}3\pi/2}$ be an optimal $(3\pi/2, u^*)$ -outer drawing of H'_{id} . Let D''_L be an optimal L-shaped drawing of H'' . Merging $D'_{\text{id}3\pi/2}$ and D''_L , one can easily construct a $(3\pi/2, u)$ -outer drawing $D_{3\pi/2}$ of H , as illustrated in Figure 4.13(g). We thus have

$$\begin{aligned}
 \text{bend}(H, 3\pi/2, u) &\leq \text{bend}(D_{3\pi/2}) \\
 &= \text{bend}(D'_{\text{id}3\pi/2}) + \text{bend}(D''_L) \\
 (4.21) \quad &= \text{bend}(H'_{\text{id}}, 3\pi/2, u^*) + \text{bend}(H'').
 \end{aligned}$$

We now claim

$$(4.22) \quad \text{bend}(H'_{\text{id}}, 3\pi/2, u^*) \leq \text{bend}(D_u|H').$$

If D_u is a drawing illustrated in Figure 4.13(f), then $D_u|H'$ is a C-shaped drawing of H' , and hence the inductive hypothesis of the proof of Fact 2 implies (4.22). One

may thus assume that D_u is a drawing illustrated in Figures 4.13(a)–(e). Then one can construct from $D_u|H'$ an (α, u^*) -outer drawing with no new bend, where $\alpha = \pi$ for Figures 4.13(a)–(c) and $\alpha = \pi/2$ for Figures 4.13(d) and (e). Hence

$$(4.23) \quad \text{bend}(H'_{\text{id}}, \alpha, u^*) \leq \text{bend}(D_u|H').$$

Since $n(H'_{\text{id}}) < n(B)$ for the graph B in Fact 1, by the inductive hypothesis of the proof of Lemma 4.2(a) we have

$$(4.24) \quad \text{bend}(H'_{\text{id}}, 3\pi/2, u^*) \leq \text{bend}(H'_{\text{id}}, \alpha, u^*).$$

Equations (4.23) and (4.24) imply (4.22).

By (4.21) and (4.22) we have

$$\begin{aligned} \text{bend}(H, 3\pi/2, u) &\leq \text{bend}(D_u|H') + \text{bend}(D_u|H'') \\ &= \text{bend}(D_u). \quad \square \end{aligned}$$

5. Conclusions. In this paper, we gave a linear algorithm to find an optimal orthogonal drawing of a series-parallel graph G of $\Delta \leq 3$ in the variable embeddings setting. Our algorithm works well even if G has multiple edges or is not biconnected and is simpler and faster than the previously known one for biconnected series-parallel simple graphs [4, 5]. One can easily extend our algorithm so that it finds an optimal orthogonal drawing of a partial 2-tree of $\Delta \leq 3$. Note that the so-called block-cutvertex graph of a partial 2-tree is a tree although the block-cutvertex graph of a series-parallel graph is a path. We gave a best possible bound on $\text{bend}(G)$: $\text{bend}(G) \leq (n(G) + 4)/3$.

In an *orthogonal grid drawing*, every vertex has an integer coordinate. The *size* of an orthogonal grid drawing is the sum of the width and height of the minimum axis-parallel rectangle enclosing the drawing. Using an argument which is similar to that in sections 3 and 4 but is more lengthy, one can prove that every series-parallel graph G of $\Delta \leq 3$ has an optimal orthogonal grid drawing of size $\leq 3N/4$, and every biconnected series-parallel graph G of $\Delta \leq 3$ has an optimal orthogonal grid drawing of size $\leq 2N/3 + 1$, where $N = n(G) + \text{bend}(G)$. The proof is omitted in the paper.

It is left as a future work to obtain a linear algorithm for a larger class of planar graphs.

Acknowledgment. We thank Dr. Md. Saidur Rahman for fruitful discussions.

REFERENCES

- [1] P. BERTOLAZZI, R. F. COHEN, G. DI BATTISTA, R. TAMASSIA, AND I. G. TOLLIS, *How to draw a series-parallel digraph*, Internat. J. Comput. Geom. Appl., 4 (1994), pp. 235–282.
- [2] F. BRANDENBURG, D. EPPSTEIN, M. T. GOODRICH, S. KOBouROV, G. LIOTTA, AND P. MUTZEL, *Selected open problems in graph drawing*, in Proceedings of GD '03, Lect. Notes Comput. Sci. 2912, Springer-Verlag, New York, 2004, pp. 515–539.
- [3] G. DI BATTISTA, P. EADES, R. TAMASSIA, AND I. G. TOLLIS, *Graph Drawing: Algorithm for the Visualization of Graphs*, Prentice-Hall Inc., Upper Saddle River, NJ, 1999.
- [4] G. DI BATTISTA, G. LIOTTA, AND F. VARGIU, *Spirality and Optimal Orthogonal Drawings*, Technical report 07.94, Dipartimento di Informatica e Sistemistica, Universita' di Roma "La Sapienza," Rome, Italy, 1994.
- [5] G. DI BATTISTA, G. LIOTTA, AND F. VARGIU, *Spirality and optimal orthogonal drawings*, SIAM J. Comput., 27 (1998), pp. 1764–1811.
- [6] A. GARG AND G. LIOTTA, *Almost bend-optimal planar orthogonal drawings of biconnected degree-3 planar graphs in quadratic time*, in Proceedings of GD '99, Lect. Notes Comput. Sci. 1731, Springer-Verlag, New York, 1999, pp. 38–48.

- [7] A. GARG AND R. TAMASSIA, *A new minimum cost flow algorithm with applications to graph drawing*, in Proceedings of GD '96, Lect. Notes in Comput. Sci. 1190, Springer-Verlag, New York, 1997, pp. 201–226.
- [8] A. GARG AND R. TAMASSIA, *On the computational complexity of upward and rectilinear planarity testing*, SIAM J. Comput., 31 (2001), pp. 601–625.
- [9] S. H. HONG, P. EADES, AND S. H. LEE, *Drawig series-parallel digraphs symmetrically*, Comput. Geom., 17 (2000), pp. 165–188.
- [10] M. JUVAN, B. MOHAR, AND R. THOMAS, *List edge-colorings of series-parallel graphs*, Electron. J. Combin., 6 (1999), pp. 1–6.
- [11] T. NISHIZEKI AND M. S. RAHMAN, *Planar Graph Drawing*, World Scientific, Singapore, 2004.
- [12] K. NOMURA, S. TAYA, AND S. UENO, *On the orthogonal drawing of outerplanar graphs*, in Proceedings of COCOON '04, Lect. Notes Comput. Sci. 3106, Springer-Verlag, New York, 2004, pp. 300–308.
- [13] M. S. RAHMAN, N. EGI, AND T. NISHIZEKI, *No-bend orthogonal drawings of subdivisions of planar triconnected cubic graphs*, IEICE Trans. Inform. Syst., E88-D (2005), pp. 23–30.
- [14] M. S. RAHMAN AND T. NISHIZEKI, *Bend-minimum orthogonal drawings of plane 3-graphs*, in Proceedings of WG '02, Lect. Notes Comput. Sci. 2573, Springer-Verlag, New York, 2002, pp. 367–378.
- [15] M. S. RAHMAN, S. NAKANO, AND T. NISHIZEKI, *Rectangular grid drawings of plane graphs*, Comput. Geom. Theory Appl., 10 (1998), pp. 203–220.
- [16] M. S. RAHMAN, S. NAKANO, AND T. NISHIZEKI, *A linear algorithm for bend-optimal orthogonal drawings of triconnected cubic plane graphs*, J. Graph Algorithms Appl., 3 (1999), pp. 31–62.
- [17] M. S. RAHMAN, S. NAKANO, AND T. NISHIZEKI, *Box-rectangular drawings of plane graphs*, J. Algorithms, 37 (2000), pp. 363–398.
- [18] M. S. RAHMAN, S. NAKANO, AND T. NISHIZEKI, *Orthogonal drawings of plane graphs without bends*, J. Graph Algorithms Appl., 7 (2003), pp. 335–362.
- [19] J. STORER, *On minimum node-cost planar embeddings*, Networks, 14 (1984), pp. 181–212.
- [20] R. TAMASSIA, *On embedding a graph in the grid with the minimum number of bends*, SIAM J. Comput., 16 (1987), pp. 421–444.
- [21] K. TAKAMIZAWA, T. NISHIZEKI, AND N. SAITO, *Linear-time computability of combinatorial problems on series-parallel graphs*, J. Assoc. Comput. Mach., 29 (1982), pp. 623–641.
- [22] X. ZHOU AND T. NISHIZEKI, *Orthogonal drawings of series-parallel graphs with minimum bends*, in Proceedings of ISAAC '05, Lect. Notes Comput. Sci. 3827, Springer-Verlag, New York, 2005, pp. 166–175.

HOW MANY POINTS CAN BE RECONSTRUCTED FROM k PROJECTIONS?*

JIRÍ MATOUŠEK[†], ALEŠ PRÍVĚTIVÝ[‡], AND PETR ŠKOVROŇ[‡]

Abstract. Let A be an n -point set in the plane. A *discrete X-ray* of A in direction u gives the number of points of A on each line parallel to u . We define $F(k)$ as the maximum number n such that there exist k directions $u_1, \dots, u_k$ such that every set of at most n points in the plane can be uniquely reconstructed from its discrete X-rays in these directions. A simple “cube” construction shows $F(k) \leq 2^{k-1}$. We establish the lower bound $F(k) \geq 2^{\Omega(k/\log k)}$ by reducing the problem through linear algebra to a graph-theoretic question, for which we then obtain an almost tight bound. As a part of the proof we establish a result in extremal theory that allows one to conclude that, under certain conditions, a graph has only at most a logarithmic density, which may be of independent interest. We also improve the upper bound to $F(k) \leq O(1.81712^k)$ (or $O(1.79964^k)$ if we allow A to be a multiset).

Key words. geometric tomography, discrete X-ray, Cayley tree-matrix formula, probabilistic method, extremal graph theory

AMS subject classifications. 52C99, 51A99, 05D40

DOI. 10.1137/080715706

1. Introduction. Let A be a finite set of points in the plane. We consider A to be chosen by an adversary and inaccessible by itself, and the information available to us consists only of k discrete X-rays of A . The *discrete X-ray of A in direction u* specifies the number of points of A on every line parallel to u (this convenient terminology is borrowed from the field of geometric tomography). We say that A is *uniquely reconstructible for directions $u_1, u_2, \dots, u_k$* if there is no $B \neq A$ such that, for all $i = 1, 2, \dots, k$, the X-rays of A and B in direction u_i are identical.

It is not hard to see that every set A of $k - 1$ or fewer points is uniquely reconstructible for *any* k distinct directions. (To see this, we suppose that some $A \neq B$ have the same X-rays in directions $u_1, \dots, u_k$, we fix a point $a \in A \setminus B$, and we note that each line through a parallel to some u_i has to contain a point of B , forcing $|B| \geq k$.) This has been observed many times; the earliest reference seems to be Rényi [12]. If the directions are chosen by an adversary, then we cannot do any better in general: Figure 1 shows directions u_1, u_2 , and u_3 and two distinct point sets A and B (the black points and the white points, respectively) that cannot be distinguished. This can be generalized to any number k of equally spaced directions, where A and B are obtained by coloring the vertices of a regular $2k$ -gon alternately black and white.

Intuition suggests that the equally spaced directions in this example are “exceptionally bad” and that other sets of directions should allow for unique reconstruction of much larger sets. For given directions $u_1, u_2, \dots, u_k$ let us define

$$f(u_1, \dots, u_k) := \max\{n : \text{every } n\text{-point set is uniquely reconstructible for } u_1, \dots, u_k\}$$

*Received by the editors February 14, 2008; accepted for publication May 16, 2008; published electronically October 17, 2008.

<http://www.siam.org/journals/sidma/22-4/71570.html>

[†]Department of Applied Mathematics and Institute of Theoretical Computer Science (ITI), Charles University, Malostranské nám. 25, 118 00 Praha 1, Czech Republic (matousek@kam.mff.cuni.cz).

[‡]Department of Applied Mathematics, Charles University, Malostranské nám. 25, 118 00 Praha 1, Czech Republic (privetivy@kam.mff.cuni.cz, xofon@kam.mff.cuni.cz).

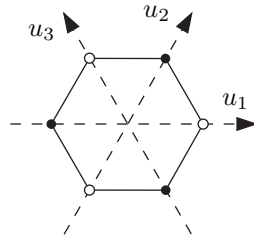


FIG. 1. Two 3-point sets with identical X-rays in three directions.

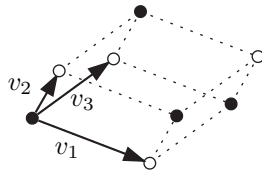


FIG. 2. A construction of two sets with identical X-rays in given directions.

and

$$F(k) := \max_{u_1, u_2, \dots, u_k} f(u_1, \dots, u_k).$$

We have just noted that $f(u_1, \dots, u_k) \geq k - 1$ for all k -tuples of distinct $u_1, \dots, u_k$.

The following simple construction (also rediscovered several times) shows that, perhaps counterintuitively, $F(k)$ is finite for every k ; see Figure 2. Let $u_1, u_2, \dots, u_k$ be given distinct directions. For suitable nonzero scalars $\alpha_1, \dots, \alpha_k$ we set $v_i := \alpha_i u_i$, and we define

$$(1) \quad A := \{v(I) : I \subseteq [k], |I| \text{ even}\}, \quad B := \{v(I) : I \subseteq [k], |I| \text{ odd}\},$$

where we use the notation $[k] := \{1, 2, \dots, k\}$ and $v(I) := \sum_{i \in I} v_i$. First, it is easy to see that almost all choices of the α_i guarantee that all of the 2^k points $v(I)$, $I \subseteq [k]$, are distinct, and thus $|A| = |B| = 2^{k-1}$. Second, for every $i \in [k]$, each point $v(I) \in A$ can be paired with the point $v(I \triangle \{i\}) = v(I) \pm v_i \in B$ (where $\triangle$ denotes symmetric difference), and this shows that A and B have identical X-rays in direction u_i . Thus $F(k) \leq 2^{k-1} - 1$.

The only published lower bound on $F(k)$ we could find is roughly $k + \Omega(\sqrt{k})$, due to Bianchi and Longinetti [3]. After we started working on the problem, we learned from Pór that Tóth [13] announced an $\Omega(k^{3/2})$ lower bound, which has remained unpublished.

We have the following lower bound.

THEOREM 1.1. *There are constants $c > 0$ and k_0 such that*

$$F(k) > 2^{ck/\log k}$$

for all $k \geq k_0$. Moreover, for every $k \geq k_0$ there exists a finite set $\mathcal{P}_k$ of nonzero polynomials in $2k$ variables and with integer coefficients such that if $u_1 = (x_1, y_1), \dots, u_k = (x_k, y_k)$ are directions with $f(u_1, \dots, u_k) \leq 2^{ck/\log k}$, then $(x_1, x_2, \dots, x_k, y_1, \dots, y_k)$ is in the zero set of some polynomial in $\mathcal{P}_k$. Consequently, almost all (in the sense of measure) k -tuples of directions $u_1, \dots, u_k$ satisfy the stated lower bound.

Thus, the true order of magnitude of $F(k)$ is between $2^{ck/\log k}$ and 2^{k-1} . We don't have any good guess of what it might be, but at least we can improve the upper bound from 2^{k-1} to C^k for a suitable $C < 2$.

To prove an upper bound we need to construct, for any k -tuple $u_1, \dots, u_k$, two sets A and B with identical X-rays in these directions. It seems technically much easier to construct examples if we allow for *multisets* A and B (formally, an n -point multiset A can be regarded as an arbitrary mapping $[n] \rightarrow \mathbb{R}^2$, and in an X-ray of A , the points are counted with the appropriate multiplicities). We thus define $f_{\text{mult}}(u_1, \dots, u_k)$ as the largest n such that every n -point multiset is uniquely reconstructible for $u_1, \dots, u_k$, and $F_{\text{mult}}(k) := \max_{u_1, \dots, u_k} f_{\text{mult}}(u_1, \dots, u_k)$. Let us remark that the proof of the lower bound in Theorem 1.1 applies equally well to F_{mult} .

THEOREM 1.2. *We have $F(k) \leq O(C^k)$ for $C = 6^{1/3} \approx 1.81712$ and $F_{\text{mult}}(k) \leq O(C_1^k)$ for $C_1 = 198^{1/9} \approx 1.79964$.*

Higher dimensions. The reconstruction problem can naturally be generalized to point sets in $\mathbb{R}^d$. There are actually several possible generalizations, since we can x-ray the point set with lines as we did in the plane, but also with k -flats for some k between 1 and $d-1$. Here we comment only on the case of line X-rays, for which some simple observations below show that it is not too different from the planar case.

We thus say that $A, B \subset \mathbb{R}^d$ have the same X-rays in direction u if $|A \cap \ell| = |B \cap \ell|$ for every line ℓ parallel to u , and we define the function $F_d(k)$ analogous to $F(k)$.

The simple construction above for the upper bound on $F(k)$ works in any dimension, so $F_d(k) \leq 2^{k-1} - 1$ for all $d \geq 2$ and all k . By the method explained in section 4 below, it can be shown that for every fixed d there exists $\delta_d > 0$ with $F_d(k) = O((2 - \delta_d)^k)$.

On the other hand, we have $F_d(k) \geq F(k)$ for all d and k with $d \geq 2$. Indeed, let $u_1, \dots, u_k$ be directions in $\mathbb{R}^2$ with $f(u_1, \dots, u_k) = F(k)$. We identify $\mathbb{R}^2$ with a (2-dimensional) plane ρ in $\mathbb{R}^d$, so that we can consider the u_i as directions in $\mathbb{R}^d$. Now if $A, B \subset \mathbb{R}^d$ are distinct sets with identical X-rays in directions $u_1, \dots, u_k$, then for every plane σ parallel to ρ the sets $A \cap \sigma$ and $B \cap \sigma$ have identical X-rays in directions $u_1, \dots, u_k$ as well, and they have to be distinct for at least one σ . For such a σ we get $|A \cap \sigma| = |B \cap \sigma| > F(k)$ by the choice of $u_1, \dots, u_k$, and this concludes the proof of $F_d(k) \geq F(k)$.

Related work. Problems similar to those investigated in the present paper have been studied in a lively area called *geometric tomography*; see, e.g., the book by Gardner [8]. The classical tomography problem deals with reconstructing a set, or more generally a density function, from X-rays in all directions. *Discrete tomography* investigates the reconstruction problem for a finite (discrete) set of X-ray directions. Since reconstructing an arbitrary set is generally impossible, most of the work deals with special sets, say, convex ones.

For reconstructing finite sets A , most of the results in the literature concern the case where A is a lattice set, $A \subseteq \mathbb{Z}^2$, and the directions of the X-rays are integer vectors. A seminal paper by Gardner and Gritzmann [6] thoroughly examines the case where A is guaranteed to be a convex lattice set (that is, the intersection of $\mathbb{Z}^2$ with a convex set). In this case, they show that any 7 lattice directions suffice for unique reconstruction of every convex lattice set, while 6 directions need not suffice. Dulio, Gardner, and Peri [5] studied a variant of the problem ("point X-rays"): Instead of k directions, we have k points $p_1, \dots, p_k$, and we obtain the number of points of A on every line passing through one of the p_i . They show, for example, that the analogue of

our function $F(k)$ is unbounded in their setting. A few other papers with somewhat related results are [9], [7], [2].

2. A reduction of the lower bound to a combinatorial problem. This section contains the first part of the proof of the lower bound for $F(k)$. Given two sets A and B with identical X-rays in given k directions, we construct a graph with edges colored by k colors, and we show that for a generic choice of the directions this colored graph satisfies a combinatorial condition; namely, it doesn't contain two color-disjoint spanning trees. The subsequent section then deals with the resulting graph-theoretic problem.

The interchange graph. Let $u_1, \dots, u_k$ be fixed direction vectors, and let $A \subset \mathbb{R}^2$ be a set with the minimum number of points n that is not uniquely reconstructible for $u_1, \dots, u_k$. Thus there exists another set $B \subset \mathbb{R}^2$ with the same X-rays in these directions. We note that $A \cap B = \emptyset$ (if not, $A \setminus B$ and $B \setminus A$ would be smaller indistinguishable sets).

For each $i = 1, 2, \dots, k$, we construct a perfect matching E_i between A and B in such a way that, for each edge $\{a, b\} \in E_i$, the points $a \in A$ and $b \in B$ lie on the same line parallel to u_i . Thus, if a line ℓ in direction u_i contains a single point of A , and hence also a single point of B , these two points necessarily form an edge of E_i . If ℓ contains several points of A , these are matched arbitrarily, in a one-to-one fashion, to the points of $\ell \cap B$ (hence E_i need not be determined uniquely, but if there are several choices, we fix one once and for all). We note that $E_i \cap E_j = \emptyset$ for $i \neq j$.

We call the bipartite graph H with vertex set $A \cup B$ and with edge set $E(H) = E_1 \cup \dots \cup E_k$ an *interchange graph for directions* $u_1, \dots, u_k$. It has $2n$ vertices and kn edges.

For our exposition it will be convenient to consider the edges of an interchange graph to have colors: We say that an edge $e \in E_i$ has *color* i , and we write $c(e) = i$.

2.1. An algebraic necessary condition for interchange graphs. We begin by recalling several notions from algebraic graph theory.

Let $\vec{G} = (V, \vec{E})$ be a directed graph. For an edge $e = (u, v) \in \vec{E}$, we write $u = \text{tail}(e)$ and $v = \text{head}(e)$. The *incidence matrix* $D = D_{\vec{G}}$ is a $V \times \vec{E}$ matrix, with rows indexed by vertices of $\vec{G}$ and columns indexed by edges, given by

$$d_{v,e} := \begin{cases} +1 & \text{if } v = \text{head}(e), \\ -1 & \text{if } v = \text{tail}(e), \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

A *circulation* in $\vec{G}$ is any vector in $\ker D$, that is, a vector $\xi \in \mathbb{R}^{\vec{E}}$ with $D\xi = 0$.

Further we recall that an *orientation* of an undirected graph $G = (V, E)$ is a directed graph $(V, \vec{E})$ obtained by choosing, for every undirected edge $\{u, v\} \in E$, one of its vertices as a head and one as a tail (more formally, for every $\{u, v\} \in E$ we have exactly one of the directed edges (u, v) and (v, u) in $\vec{E}$). For notational convenience we will use the same letter for an undirected edge $e = \{u, v\}$ and for a directed version (u, v) of it.

We begin with the following necessary condition for interchange graphs.

LEMMA 2.1. *Let $H = (V, E)$ be a subgraph of an interchange graph for directions $u_1, u_2, \dots, u_k$, and let $\vec{H}$ be an arbitrary orientation of H . Then there exist reals α_e ,*

$e \in E$, all of them nonzero, such that for every circulation ξ in $\vec{H}$ we have

$$\sum_{e \in E} \alpha_e \xi_e x(u_{c(e)}) = 0,$$

and for every circulation ψ in $\vec{H}$ we have

$$\sum_{e \in E} \alpha_e \psi_e y(u_{c(e)}) = 0.$$

Here $c(e)$ denotes the color of the edge e (inherited from the interchange graph), and $x(u)$ and $y(u)$ denote the x -coordinate and y -coordinate, respectively, of a vector $u \in \mathbb{R}^2$.

Proof. First we note that if $G = (V, \vec{E})$ is a directed graph, ξ is a circulation in it, and $z \in \mathbb{R}^V$ is a vector indexed by the vertices of G , we have

$$(2) \quad \sum_{e \in \vec{E}} (z_{\text{head}(e)} - z_{\text{tail}(e)}) \xi_e = 0.$$

This is immediate by using $D\xi = 0$, because the sum equals $z^T D\xi$. (Or, perhaps more intuitively, the equality is obvious if ξ is an elementary circulation along a cycle, and the general case follows since every circulation is a linear combination of elementary circulations.)

Next, let us consider $\vec{H}$ as in the lemma, and let ξ be a circulation in it. The vertices of H are points in the plane, and for every edge $e \in E$ we have $(\text{head}(e) - \text{tail}(e)) = \alpha_e u_{c(e)}$ for some $\alpha_e \neq 0$, since the segment with end points $\text{head}(e)$ and $\text{tail}(e)$ is parallel to $u_{c(e)}$. Then the first equality in the claim of the lemma follows by applying (2) with $z_v := x(v)$ for all vertices v , while the second equality is obtained similarly by using the y -coordinates. $\square$

Let D be the incidence matrix of a directed graph. By D^- we denote D with the last row deleted. We note that since every column of D has one $+1$ and one -1 , the sum of all rows of D is the zero vector, and thus each row is a linear combination of the others. Consequently, we have $\ker D^- = \ker D$, and thus it suffices to “test” circulations by using D^- .

Let $H = (V, E)$ be a graph with n vertices and m edges and with edges colored by k colors. We define a polynomial matrix $P_H = P_H(x_1, \dots, x_k, y_1, \dots, y_k)$ with $m + 2n - 2$ rows and $2m$ columns, where the x_i and the y_i are variables:

$$P_H := \begin{pmatrix} X & Y \\ D^- & 0 \\ 0 & D^- \end{pmatrix},$$

where X is an $E \times E$ diagonal matrix with the entry (e, e) equal to $x_{c(e)}$, Y is defined analogously with $y_{c(e)}$, and D is the incidence matrix of some orientation of H (thus, P_H is not defined uniquely, since it depends on the orientation).

Here is an algebraic condition for certain subgraphs of interchange graphs.

LEMMA 2.2. *Let $H = (V, E)$ be a subgraph of an interchange graph for directions $u_1, \dots, u_k$. Let us assume that $|E| = 2n - 2$, where $n = |V|$. Then P_H is a square matrix and*

$$\det(P_H(x(u_1), \dots, x(u_k), y(u_1), \dots, y(u_k))) = 0$$

(we substitute the coordinates of the directions $u_1, \dots, u_k$ for the variables in the matrix P_H).

Proof. Let us write $\tilde{P}_H = P_H(x(u_1), \dots, x(u_k), y(u_1), \dots, y(u_k))$. By supposing for contradiction that $\det(\tilde{P}_H) \neq 0$, we get that the linear system $\tilde{P}_H q = b$ has a (unique) solution q for every right-hand side $b \in \mathbb{R}^{2m}$. Let us choose, in particular, $b = \begin{pmatrix} w \\ 0 \end{pmatrix}$, where $w \in \mathbb{R}^E$ is arbitrary (and 0 stands for an m -component zero vector), and let us also write $q = \begin{pmatrix} \xi \\ \psi \end{pmatrix}$, with $\xi, \psi \in \mathbb{R}^E$. By using the definition of P_H , we get that the system $\tilde{P}_H q = b$ is then equivalent to

$$D^- \xi = 0, \quad D^- \psi = 0, \quad x(u_{c(e)})\xi_e + y(u_{c(e)})\psi_e = w_e \quad \text{for all } e \in E.$$

The first two equations say that ξ and ψ are circulations.

Lemma 2.1 tells us that there exist nonzero α_e , $e \in E$, such that for every two circulations ξ, ψ we have

$$\sum_{e \in E} \alpha_e \left(x(u_{c(e)})\xi_e + y(u_{c(e)})\psi_e \right) = 0.$$

However, in view of the previous paragraph, for every $w \in \mathbb{R}^E$ we can choose circulations ξ, ψ with $x(u_{c(e)})\xi_e + y(u_{c(e)})\psi_e = w_e$ for all $e \in E$, and thus we have $\sum_{e \in E} \alpha_e w_e = 0$ for all w . This forces $\alpha_e = 0$ for all e , though—a contradiction proving the lemma. $\square$

In order that the last lemma provide a nontrivial condition for the directions u_i , we need that the determinant of the matrix $P_H(x_1, \dots, x_k, y_1, \dots, y_k)$ is not the zero polynomial. The following section provides a sufficient combinatorial condition for this.

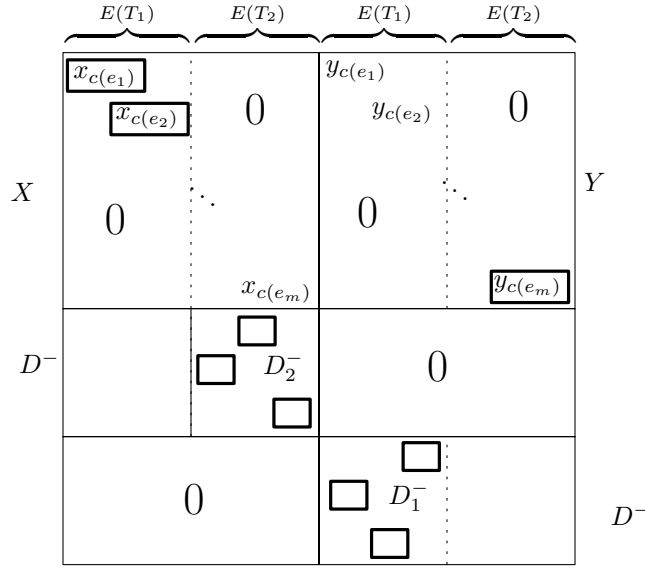
2.2. A combinatorial necessary condition for interchange graphs.

LEMMA 2.3. *Let $H = (V, E)$ be a subgraph of an interchange graph for directions $u_1, \dots, u_k$, and suppose that H has two color-disjoint spanning trees T_1 and T_2 (that is, $c(e_1) \neq c(e_2)$ for all $e_1 \in E(T_1)$, $e_2 \in E(T_2)$). Then $(x(u_1), \dots, x(u_k), y(u_1), \dots, y(u_k))$ is in the zero set of a nonzero polynomial with integer coefficients (depending on H).*

Proof. Without loss of generality we may assume that $E = E(T_1) \cup E(T_2)$. By writing $n = |V|$, we have $|E| = 2n - 2$, and so Lemma 2.2 tells us that the vector of the coordinates of the u_i is in the zero set of $\det(P_H)$. It remains to check that $\det(P_H)$ is not identically 0. To this end, we verify that the monomial

$$\mu := \left(\prod_{e \in E(T_1)} x_{c(e)} \right) \left(\prod_{e \in E(T_2)} y_{c(e)} \right)$$

appears in $\det(P_H)$ with coefficient ± 1 . Indeed, let us consider the expansion of $\det(P_H)$ according to the definition of the determinant as a sum over all permutations of products. Since T_1 and T_2 are color-disjoint, if such a product should be a nonzero multiple of the monomial μ , it must use all of the entries $x_{c(e)}$ of the diagonal matrix X for $e \in E(T_1)$ and also all of the entries $y_{c(e)}$ of the diagonal matrix Y for $e \in E(T_2)$; see Figure 3. This uniquely determines which entries are used from the top m rows of P_H . In the next $n - 1$ rows, the entries must be taken from the submatrix D_2^- of D^- corresponding to $E(T_2)$, and in the last $n - 1$ rows the entries must be taken from the submatrix D_1^- of D^- corresponding to $E(T_1)$. Hence the coefficient of μ in $\det(P_H)$ equals $\det(D_1^-) \det(D_2^-)$. By a standard fact (a special case of the matrix-tree theorem; see, e.g., Biggs [4]), the incidence matrix of any directed tree has determinant ± 1 , and the lemma follows. $\square$

FIG. 3. Terms in $\det(\tilde{P}_H)$ contributing to the monomial μ .

3. Existence of two color-disjoint spanning trees. The last building block in the proof of our lower bound for $F(k)$ is a result stating that every sufficiently dense graph has a subgraph with two color-disjoint spanning trees.

The graphs that we will deal with throughout the proof will be disjoint unions of perfect matchings, so it is convenient to introduce the following terminology. A *DUPM-graph* is a graph $G = (V, E)$ together with a fixed partition $E = E_1 \cup \dots \cup E_k$ of the edge set into perfect matchings on V (we also keep referring to *colors* of edges as in the previous sections). In particular, we note that each vertex is incident to edges of all colors in a DUPM-graph. For an index set $I \subseteq [k]$ and a subset $W \subseteq V$, we write $G[I, W]$ for the subgraph induced by W on edges with colors in I , that is, $G[I, W] = (W, \binom{W}{2} \cap \bigcup_{i \in I} E_i)$ (instead of $G[I, V]$ we write just $G[I]$). This $G[I, W]$ need not be a DUPM-graph in general, but in our considerations we will choose I and W so that we do get DUPM-graphs.

Before proving the result needed for the proof of Theorem 1.1, we first establish a quantitatively weaker result with a simpler proof (section 3.1), which can serve as a gradual introduction to the subsequent more involved proof of the stronger result.

3.1. A weaker bound. The weaker result with the simpler proof is the following.

PROPOSITION 3.1. *Let $G = (V, E_1 \cup \dots \cup E_k)$ be a DUPM-graph on $n \geq 2$ vertices. If $k \geq 5 \log_2^3 n$, then there exist disjoint index sets $I_1, I_2 \subset [k]$ and a subset $W \subseteq V$, $|W| \geq 2$, such that the graphs $G[I_1, W]$ and $G[I_2, W]$ are both connected.*

The construction of the desired two subgraphs proceeds in two stages. In the first stage, captured by the next lemma, we find a DUPM-subgraph of G in which every edge cut has sufficiently many colors.

LEMMA 3.2. *Let G and k be as in Proposition 3.1. Then there exist nonempty $J \subseteq [k]$ and $W \subseteq V$, $|W| \geq 2$, such that $H := G[J, W]$ is a DUPM-graph in which every edge cut has edges of at least $r := 5 \log_2^2 n$ colors. That is, for every partition of W into nonempty sets A and B , the edges in $E(H; A, B) := \{e \in E(H) : e \cap A \neq \emptyset \neq e \cap B\}$ have at least r distinct colors.*

The proof proceeds by a recursive partitioning: As long as we can find an edge cut F with fewer than r colors in the current graph, we discard the *larger* of the two pieces defined by F , together with all edges that do not have both end points in the new vertex set or whose color appears in F . We repeat this step as long as we can. Since the size of the vertex set is halved (at least) in each iteration, the number of iterations is at most $\log_2 n$. The key observation is that, since we started with $5 \log_2^3 n$ colors and we lose fewer than r colors in every iteration, we must finish with a nontrivial graph satisfying the required condition. A more formal proof follows.

Proof. We start with $J_1 = [k]$, $V_1 = V$, and $G_1 = G$, and for $j = 1, 2, \dots$ we repeat the following step:

(Step j) We have a current DUPM-graph G_j on vertex set V_j with edge colors in J_j . If there is a partition $V_j = A \cup B$ such that the edges in $E(G_j; A, B)$ have fewer than r colors, we construct J_{j+1} , V_{j+1} , G_{j+1} as follows:

- V_{j+1} is the smaller of the sets A and B (a possible tie broken arbitrarily);
- J_{j+1} is J_j minus all colors appearing on edges in $E(G_j; A, B)$;
- $G_{j+1} = G_j[V_{j+1}, J_{j+1}]$.

Otherwise, if $E(G_j; A, B)$ has at least r colors for every partition $V_j = A \cup B$, then we set $W = V_j$, $J = J_j$ and finish.

It is clear that the resulting $H = G[J, W]$ is a DUPM-graph, and actually that every vertex of H is incident to $|J|$ edges, and also that every edge cut has at least r colors. It remains to check that $J \neq \emptyset$. As we have observed, the algorithm must finish in at most $\log_2 n$ steps since $|V_{j+1}| \leq \lfloor \frac{1}{2} |V_j| \rfloor$ for all j . In each step the number of colors in J_j decreases by less than r , and hence $|J| > k - r \log_2 n \geq 0$. The lemma is proved. $\square$

In the second stage of the proof of Proposition 3.1, having a subgraph where every edge cut has at least r colors, we exhibit two color-disjoint connected subgraphs. To build the first connected subgraph, we start with an empty edge set and add the edge sets E_i one by one, always taking one that decreases the current number of components the most. Since every edge cut contains at least about a $1/\log n$ fraction of all colors, it can be shown that adding a random color decreases the current number of component by a factor of at least roughly $(1 - 1/\log n)$. Consequently, after adding $s = O(\log^2 n)$ colors we obtain a connected graph. By setting the constants so that $s < r$, the remaining colors form a graph where every edge cut still has at least one color, and hence the graph is connected. As was pointed to us by Alon, this stage of our proof resembles his proof in [1].

The next lemma formalizes this second stage.

LEMMA 3.3. *Let H be a DUPM-graph on $m \leq n$ vertices and with $k \leq 6 \log_2^3 n$ colors in which every edge cut has edges of at least $r := 5 \log_2^2 n$ colors. Then there exists a set I of at most $s := 4 \log_2^2 n$ colors such that $H[I]$ is connected.*

Proof. We set $I_0 = \emptyset$, and for $j = 0, 1, 2, \dots$, we do the following: If $H[I_j]$ is connected, we set $I := I_j$ and finish. Otherwise, we let i_j be a color i maximizing $\text{cc}(H[I_j]) - \text{cc}(H[I_j \cup \{i\}])$, where $\text{cc}(G)$ denotes the number of connected components of a graph G . Then we set $I_{j+1} := I_j \cup \{i_j\}$, and we continue with the next step.

We need to show that we obtain a connected graph before exhausting more than the claimed number of colors.

Let $m_j = \text{cc}(H[I_j])$. We have $m_1 = m \leq n$, and we claim that

$$m_{j+1} \leq m_j \left(1 - \frac{r}{2k}\right)$$

whenever $m_j \geq 2$. To see this, let $K_1, K_2, \dots, K_{m_j}$ be the connected components of $H[I_j]$. Let us say that K_ℓ gets connected by color i if there is an edge of color i connecting K_ℓ to the rest of H . Since the edge cut $E(H; V(K_\ell), W \setminus V(K_\ell))$ contains at least r colors (none of them in I_j , since K_ℓ is a connected component of $H[I_j]$), the probability that any particular K_ℓ gets connected by a randomly chosen color i is at least r/k . Hence in expectation at least $q := m_j r/k$ components get connected, and thus there is a particular color i by which at least q components get connected. It is not difficult to check that by adding such a color i to I_j the number of components decreases by at least $q/2$, and thus $m_{j+1} \leq m_j - q/2 = m_j(1 - r/2k)$ as claimed.

Applying this inequality inductively gives $m_j \leq n(1 - r/2k)^j \leq ne^{-jr/2k}$. This is at most 1 for $j = 2k(\ln n)/r \leq 2 \cdot 6 \log_2^3 n \cdot \log_2 e \cdot (\log_2 n)/(5 \log_2^2 n) \leq s = 4 \log_2^2 n$, and the lemma is proved. $\square$

Proof of Proposition 3.1. Given a DUPM-graph $G = (V, E_1 \cup \dots \cup E_k)$ as in the proposition, we first discard colors if needed so that the remaining number of colors is between $5 \log_2^3 n$ and $6 \log_2^3 n$ (we'd actually like to say that we may assume that there are exactly $5 \log_2^3 n$ colors, but that need not be an integer). Then we use Lemma 3.2 and obtain a DUPM-graph $H = G[J, W]$ where every edge cut contains at least r colors (and, of course, H has at most n vertices and at most $6 \log_2^3 n$ colors). We apply Lemma 3.3, and we obtain a set I_1 of at most $s = 4 \log_2^2 n < r$ colors with $H[I_1]$ connected. In the DUPM-graph induced in H by colors not belonging to I_1 , every edge cut still has at least one color (since we have removed fewer than r colors), and hence this DUPM-graph is connected as well. We can thus set $I_2 := J \setminus I_1$. This proves the proposition. $\square$

3.2. An improved bound. Here we improve the bound of $\log^3 n$ from Proposition 3.1 to $\log n \log \log n$. The improvement is based on two observations, which are more or less independent (applying each of them alone would save a factor of roughly $\log n$).

The first observation is that in the first stage, i.e., in the proof of Lemma 3.2, we can force unbalanced cuts (with A much smaller than B) to have considerably more colors than balanced ones (with A and B of roughly equal size). Indeed, if we encounter highly unbalanced cuts in most steps of the algorithm, then the algorithm has to finish considerably sooner than after $\log_2 n$ steps, and as a result fewer colors get discarded. This is expressed quantitatively in Lemma 3.6 below.

In the second stage we build a connected graph by iteratively merging components. Here we note that at the beginning, when the components are small, the property “cuts having sufficiently many colors” is applied with very much unbalanced cuts. Only as the typical size of the current components grows do we have to deal with more balanced cuts. As a result, if we know that unbalanced cuts are richer in colors than balanced ones, the merging procedure progresses faster at the beginning than towards the end. Technically, we are going to divide the merging procedure into roughly $\log \log n$ phases. In each phase, the number of components is typically reduced to approximately its square root (except for the last phase, where the number of components is reduced to 1).

This first improvement alone would bring the original $\log^3 n$ factor down to $\log^2 n \log \log n$, where the $\log \log n$ factor arises because the contribution of each of the $\log \log n$ phases turns out to be roughly the same.

We now turn to the second improvement. The procedure for selecting colors and merging components in the proof of Lemma 3.3 can conceptually be regarded as a randomized algorithm that selects colors one by one at random. (We phrased

it deterministically, by saying that we always pick a color that reduces the number of components as much as possible, but the reason why this worked was that the expected number of components merged by adding a *random* color is sufficiently large.) Instead of picking the colors one by one, we can think of this as selecting a random subset of s colors, with a suitable s . If we can show that such a random set of colors induces a connected subgraph with a probability strictly greater than $\frac{1}{2}$, we can use the following simple observation to assert the existence of *two disjoint* sets of s colors, each inducing a connected subgraph.

OBSERVATION 3.4. *Let s and n be natural numbers, $1 \leq s \leq n/2$, and let $\mathcal{P}$ be a property of s -element subsets of $[n]$. If a random s -element subset of $[n]$ (where all s -element subsets are chosen with equal probability) has property $\mathcal{P}$ with probability $p > \frac{1}{2}$, then there exist two disjoint s -element subsets of $[n]$ with property $\mathcal{P}$.*

This observation, as well as the following neat proof, may very well be folklore.

Proof. Let us choose a random permutation of $[n]$. Let S_1 consist of the first s elements in this permutation, and let S_2 consist of the next s elements. We have $S_1 \cap S_2 = \emptyset$, and both S_1 and S_2 can be regarded as random s -element subsets of $[n]$ (not independent, of course). The probability that either of S_1 and S_2 fails to have $\mathcal{P}$ is, by the union bound, at most $2(1 - p) < 1$. $\square$

Using this observation appropriately will allow us to save another $\log n$ factor in the number of colors.

Here is a formal statement of the improved version of Proposition 3.1.

PROPOSITION 3.5. *There exists a constant C such that the following holds. Let $G = (V, E_1 \cup \dots \cup E_k)$ be a DUPM-graph on $n \geq 4$ vertices.¹ If $k \geq \lfloor C \log_2 n \log_2 \log_2 n \rfloor$, then there exist disjoint index sets $I_1, I_2 \subset [k]$ and a subset $W \subseteq V$, $|W| \geq 2$, such that the graphs $G[I_1, W]$ and $G[I_2, W]$ are both connected.*

We begin with the first stage of the proof—a refined version of Lemma 3.2. For a partition of a vertex set into two nonempty subsets A and B , we define the *imbalance* of (A, B) as

$$\text{imb}(A, B) := \frac{|A| + |B|}{\min(|A|, |B|)},$$

and we set

$$\gamma(n, k, A, B) := \frac{k}{2 \log_2 n} \cdot \log_2 \text{imb}(A, B);$$

this is going to be the lower bound on the number of colors in every edge cut $E(H; A, B)$.

LEMMA 3.6. *Let $G = (V, E_1 \cup \dots \cup E_k)$ be a DUPM-graph on $n \geq 2$ vertices. Then there exist $J \subseteq [k]$, $|J| \geq k/2$, and $W \subseteq V$, $|W| \geq 2$, such that $H := G[J, W]$ is a DUPM-graph in which every edge cut $E(H; A, B)$ has edges of at least $\gamma(n, k, A, B)$ colors (note that n is the number of vertices of G , not of H).*

Proof. The proof is a simple modification of that for Lemma 3.2, so we proceed quickly. We assume $k > 0$. We start with $J_1 = [k]$, $V_1 = V$, and $G_1 = G$. If a current DUPM-graph G_j on vertex set V_j with edge colors in J_j has already been constructed and if there is a partition $V_j = A \cup B$ where $E(G_j; A, B)$ has fewer than $\gamma(n, k, A, B)$ colors, we let V_{j+1} be the smaller of the sets A and B , we let J_{j+1} be J_j minus all colors appearing in $E(G_j; A, B)$, and we set $G_{j+1} = G_j[V_{j+1}, J_{j+1}]$. If there is no such partition, we set $W = V_j$, $J = J_j$, $t = j$ and finish.

¹We assume that $n \geq 4$ so that $\log_2 \log_2 n \geq 1$ and the $C \log_2 n \log_2 \log_2 n$ term can absorb additive constants, for example, if C is chosen sufficiently large.

The only condition whose checking needs some work is $|J| \geq k/2$. In the j th step the number of colors in J_j decreases by less than

$$\gamma(n, k, V_{j+1}, V_j \setminus V_{j+1}) = \frac{k}{2\log_2 n} \log_2 \frac{|V_j|}{|V_{j+1}|},$$

and hence

$$\begin{aligned} |J| &\geq k - \sum_{j=1}^{t-1} \frac{k}{2\log_2 n} (\log_2 |V_j| - \log_2 |V_{j+1}|) \\ &= k - \frac{k}{2\log_2 n} (\log_2 |V_1| - \log_2 |V_t|) \\ &\geq k - \frac{k}{2\log_2 n} \log_2 n = \frac{k}{2}. \end{aligned}$$

The lemma is proved. $\square$

Next, we state a counterpart of Lemma 3.3.

LEMMA 3.7. *Let H be a DUPM-graph on $m \leq n$ vertices and with k colors, where $k_0/2 \leq k \leq k_0$ for $k_0 := \lfloor C \log_2 n \log_2 \log_2 n \rfloor$. Let us suppose that every edge cut $E(H; A, B)$ has at least $\gamma(n, k_0, A, B)$ colors. Then a set I of $s := \lfloor \frac{1}{4}k_0 \rfloor$ colors chosen uniformly at random from $[k]$ guarantees $H[I]$ to be connected with probability at least $\frac{3}{4}$.*

It is easy to derive Proposition 3.5 from Lemma 3.7.

Proof of Proposition 3.5. Given a DUPM-graph as in the proposition, we may assume (after deleting colors) that the number of colors is exactly k_0 . Lemma 3.6 yields a DUPM-subgraph H with at most n vertices, a number of colors between $k_0/2$ and k_0 , and every edge cut $E(H; A, B)$ having at least $\gamma(n, k_0, A, B)$ colors. Lemma 3.7 shows that a random subset of s colors induces a connected subgraph in H with a probability of at least $\frac{3}{4}$, and Observation 3.4 thus yields the existence of two disjoint sets of colors inducing connected subgraphs. $\square$

Now we start working towards the proof of Lemma 3.7. Instead of choosing the random set I of colors all at once, we think of picking the colors one by one. So we set $I_0 = \emptyset$, and, for $j = 1, 2, 3, \dots$, we choose a color $i_j \in [k] \setminus I_{j-1}$ uniformly at random and set $I_j := I_{j-1} \cup \{i_j\}$.

Moreover, for the purposes of the analysis, we conceptually divide the choice of I into *phases*. We set $b := \lceil 64 \log_2 n \rceil$, and, for $i = 0, 1, 2, \dots$, we let the sets $I_{ib}, I_{ib+1}, \dots, I_{(i+1)b-1}$ belong to the i th phase.

Let $X_j := \text{cc}(H[I_j])$ be the number of connected components of the graph $H[I_j]$; this is a random variable. The next lemma wants to claim that each phase is reasonably likely to reduce the current number of connected components to no more than its square root. However, if the number of components is already rather small, concretely, below 2^4 , we require the phase to finish with a connected graph. So let us call the i th phase *successful* if either $X_{ib} \geq 2^4$ and $X_{(i+1)b} \leq \sqrt{X_{ib}}$, or $2 \leq X_{ib} < 2^4$ and $X_{(i+1)b} = 1$.

LEMMA 3.8. *For every i we can estimate the conditional probability of the i th phase being successful, conditioned on $X_{ib} \geq 2$, as follows:*

$$\text{Prob}[\textit{ith phase is successful} \mid X_{ib} \geq 2] \geq \frac{1}{2}.$$

Proof. Let us assume that the set I_{ib} is fixed in such a way that $x := X_{ib} \geq 2$. Let $p^* = p^*(I_{ib})$ be the probability that the i th phase is *unsuccessful*, conditioned on I_{ib} fixed as above.² It suffices to prove that $p^* \leq \frac{1}{2}$ for every I_{ib} with $X_{ib} \geq 2$.

For contradiction, let us assume that $p^* > \frac{1}{2}$. We will show that this implies that

$$\mathbb{E}[\log_2 X_{(i+1)b} \mid I_{ib}] < 0,$$

which is a contradiction, since obviously $X_{(i+1)b} \geq 1$ always. To this end, we define $D_j := \log_2(X_j/X_{j+1})$, and we establish the following.

Claim. For all $j = ib, ib+1, \dots, (i+1)b-1$ we have

$$\mathbb{E}[D_j \mid I_{ib}] > \frac{1}{64} \frac{\log_2 x}{\log_2 n}.$$

First, assuming this claim, we can finish the proof of the lemma quickly. We calculate

$$\begin{aligned} \mathbb{E}[\log_2 X_{(i+1)b} \mid I_{ib}] &= \log_2 x - \mathbb{E}[\log_2(X_{ib}/X_{(i+1)b}) \mid I_{ib}] \\ &= \log_2 x - \sum_{j=ib}^{(i+1)b-1} \mathbb{E}[D_j \mid I_{ib}] \\ &< \log_2 x - \sum_{j=ib}^{(i+1)b-1} \frac{1}{64} \frac{\log_2 x}{\log_2 n} \\ &\leq (\log_2 x) \left(1 - \frac{b}{64 \log_2 n}\right) \leq 0 \end{aligned}$$

since $b = \lceil 64 \log_2 n \rceil$. We have shown $\mathbb{E}[\log_2 X_{(i+1)b} \mid I_{ib}] < 0$, which is the contradiction announced above.

Proof of the claim. Let us define y as the desired target value of $X_{(i+1)b}$ making the phase successful; that is, $y := \sqrt{x}$ for $x \geq 2^4$ and $y := 1$ for $2 \leq x < 2^4$. We have

$$\begin{aligned} \mathbb{E}[D_j \mid I_{ib}] &= \mathbb{E}[D_j \mid I_{ib} \text{ and } X_j > y] \cdot \text{Prob}[X_j > y \mid I_{ib}] \\ &\quad + \mathbb{E}[D_j \mid I_{ib} \text{ and } X_j \leq y] \cdot (1 - \text{Prob}[X_j > y \mid I_{ib}]) \\ (3) \quad &\geq \mathbb{E}[D_j \mid I_{ib} \text{ and } X_j > y] \cdot p^*, \end{aligned}$$

since $\mathbb{E}[D_j \mid I_{ib} \text{ and } X_j \leq y] \cdot (1 - \text{Prob}[X_j > y \mid I_{ib}]) \geq 0$ and $\text{Prob}[X_j > y \mid I_{ib}] \geq \text{Prob}[X_{(i+1)b} > y \mid I_{ib}] = p^*$.

We now want to bound below $\mathbb{E}[D_j \mid I_{ib} \text{ and } X_j > y]$. Let W denote the vertex set of H , and let us recall that $m = |W|$. Let us fix I_j such that $X_j > y$, and, for every color $\ell \in [k] \setminus I_j$, let $a_\ell := X_j - \text{cc}(H[I_j \cup \{\ell\}])$ be the decrease in the number of connected components caused by adding color ℓ .

For each component A of $H[I_j]$, the number of colors in the complement of I_j that connect A to the rest of the graph is at least

$$\gamma(n, k_0, A, W \setminus A) = \frac{k_0}{2 \log_2 n} \log_2 \text{imb}(A, W \setminus A),$$

²More precisely, we should say that we first fix an ib -element set $J \subset [k]$ with $\text{cc}(H[J]) \geq 2$, set $x := \text{cc}(H[J])$, and then condition on $I_{ib} = J$. However, we allow ourselves the luxury of the more concise notation.

and hence

$$(4) \quad \sum_{\ell \in [k] \setminus I_j} a_\ell \geq \frac{1}{2} \cdot \frac{k_0}{2 \log_2 n} \sum_A \log_2 \text{imb}(A, W \setminus A),$$

where the sum is over all components A of $H[I_j]$.

Now for the case $2 \leq x < 2^4$, we simply use $\text{imb}(A, W \setminus A) \geq 2$ for every A , and (4) yields

$$(5) \quad \sum_{\ell \in [k] \setminus I_j} a_\ell \geq \frac{1}{4} X_j \frac{k_0}{\log_2 n}.$$

To derive a similar inequality for $x \geq 2^4$ as well, we note that the number of components of $H[I_j]$ of size exceeding some threshold t is at most m/t , and, by setting $t = 2m/\sqrt{x}$, we find that if $X_j > \sqrt{x}$, then at least $\frac{1}{2}X_j$ of the components have size at most t . For each such component A , we have $\text{imb}(A, W \setminus A) \geq m/t = \frac{1}{2}\sqrt{x}$, and so

$$(6) \quad \sum_{\ell \in [k] \setminus I_j} a_\ell \geq \frac{k_0}{4 \log_2 n} \cdot \frac{1}{2} X_j \cdot \log_2(\sqrt{x}/2) \geq \frac{1}{32} X_j \frac{k_0 \log_2 x}{\log_2 n},$$

where we have used $\log_2(\sqrt{x}/2) \geq \frac{1}{4} \log_2 x$, which is valid for all $x \geq 2^4$.

In the case $2 \leq x < 2^4$, we have $\log_2 x \leq 4$, and so (5) shows that (6) holds in this case, too.

Now we can calculate

$$\begin{aligned} \mathbb{E}[D_j | I_j] &= \mathbb{E}[\log_2(X_j/X_{j+1}) | I_j] \\ &= \sum_{\ell \in [k] \setminus I_j} \frac{1}{k-j} \log_2 \frac{X_j}{X_j - a_\ell} \\ &\geq \frac{1}{k_0} \sum_{\ell \in [k] \setminus I_j} \log_2 \frac{1}{1 - a_\ell/X_j} \\ &\geq \frac{1}{k_0} \sum_{\ell \in [k] \setminus I_j} \frac{a_\ell}{X_j}, \end{aligned}$$

where we have used the inequality $\ln(1+z) \leq z$, with $z = -a_\ell/X_j$. By using (6), we arrive at $\mathbb{E}[D_j | I_j] \geq \frac{1}{32} \frac{\log_2 x}{\log_2 n}$, and averaging over all choices of $I_j \supseteq I_{ib}$ and plugging into (3) finally lead to the inequality in the claim. This also concludes the proof of the lemma. $\square$

Proof of Lemma 3.7. Let us put $q := \lceil \log_2 \log_2 n \rceil + 1$. First we note that there cannot be more than q successful phases in the algorithm. Indeed, each successful phase starting with at least 2^4 components reduces the current number of components to at most its square root, so there can't be more than $\log_2 \log_2 n$ such phases, and, after the number of components drops below 2^4 , there can be at most one additional successful phase.

Let Y_i be the indicator variable of the event that the i th phase is successful, and let $p_i := \text{Prob}[X_{ib} \geq 2]$. We have

$$\mathbb{E}[Y_i] = \text{Prob}[Y_i = 1 | X_{ib} \geq 2] p_i \geq \frac{1}{2} p_i$$

by Lemma 3.8.

Let us consider the first $8q$ phases, and let $p := p_{8q}$. For proving the lemma it suffices to show that $p \leq \frac{1}{4}$. Indeed, then after $8q$ phases we have $X_{8q} = 1$ with probability at least $\frac{3}{4}$.

For contradiction we thus assume that $p > \frac{1}{4}$. Since $p_i \geq p > \frac{1}{4}$ for all $i < 8q$, we have $\mathbb{E}[Y_0 + Y_1 + \dots + Y_{8q-1}] > 8q \cdot \frac{1}{8} = q$. But this contradicts the fact noted above, namely, that at most q of the Y_i can be 1. This finishes the proof. $\square$

3.3. Proof of Theorem 1.1. Theorem 1.1 is now a simple consequence of Lemma 2.3 and Proposition 3.5. Given directions $u_1, \dots, u_k$, we let $n := 2f(u_1, \dots, u_k)$. Assuming $k \geq \lfloor C \log_2 n \log_2 \log_2 n \rfloor$, where C is as in Proposition 3.5, we consider an interchange graph G on n vertices for $u_1, \dots, u_k$. By Proposition 3.5 it has a subgraph H possessing two color-disjoint connected subgraphs and hence two color-disjoint spanning trees. Then by Lemma 2.3 the vector of the coordinates of the u_i has to be in the zero set of a nonzero polynomial (depending on H). $\square$

3.4. A limitation of the proof method. Our strategy for finding the two color-disjoint spanning trees was to reduce the graph as long as possible by partitioning along “few-colored” cuts and then to choose two random disjoint sets of colors. Here we present an example showing that our version of the second part of this strategy, namely, Lemma 3.7, is not far from optimal.

PROPOSITION 3.9. *For infinitely many values of n , there exists a DUPM-graph H on n vertices and with k colors, where $k \geq \frac{1}{2} \log n \log \log n / \log^{(3)} n$ (with $\log^{(3)} n = \log \log \log n$), such that every edge cut $E(H; A, B)$ has at least $\gamma(n, k, A, B)$ colors and the probability that $H[I]$ is connected, where $I \subset [k]$ is a random subset of colors obtained by picking each color of $[k]$ independently with probability $\frac{1}{2}$, tends to 0 as $n \rightarrow \infty$. (There is nothing special about probability $\frac{1}{2}$; a similar construction with parameters slightly adjusted works for any constant probability $p < 1$.)*

Let us remark that the graph H that we construct in the proof of this proposition does have two color-disjoint spanning trees. Hence the proposition shows only that, in a setting as in Lemma 3.7 but with considerably fewer than $\log n \log \log n$ colors, we cannot get the desired trees by a purely random choice of colors—it might still be that they could be obtained by some more clever strategy.

One might also wonder whether the initial reduction, partitioning the graph along few-colored cuts, couldn't be improved. We believe that an improvement cannot be achieved in this way either, but we prefer not to formalize our heuristic reasoning here, since it doesn't seem to bring anything useful towards resolving the problem (while the example in Proposition 3.9 might give some inspiration for future research).

Sketch of proof of Proposition 3.9. We recall that the Cartesian product (or graph product) $G_1 \square G_2$ of two graphs has vertex set $V(G_1) \times V(G_2)$, and two vertices (v_1, v_2) and (v'_1, v'_2) in this product are connected by an edge if $v_1 = v'_1$ and $\{v_2, v'_2\} \in E(G_2)$ or if $v_2 = v'_2$ and $\{v_1, v'_1\} \in E(G_1)$. (So the d -dimensional cube is the d th Cartesian power of an edge.)

We assume n to be sufficiently large, and we set (ignoring integrality issues) $b := \frac{1}{2} \log \log n$, assuming b even, and $d := \log n / \log b$. We let H be the d -fold Cartesian product $H_1 \square H_2 \square \dots \square H_d$, where each H_i is a complete DUPM-graph on b vertices such that the sets of colors for H_i and H_j are disjoint for $i \neq j$. Then H is a DUPM-graph on $b^d = 2^{d \log b} = n$ vertices with $k = d(b - 1) \in [\frac{1}{2} \log n \log \log n / \log^{(3)} n, \log n \log \log n / \log^{(3)} n]$ colors (each edge in the product “comes from” a particular edge of a particular H_i , and it inherits the color of that edge).

If $I \subset [k]$ is a random set of colors as in the proposition, then with probability larger than 2^{-b} the set I contains no color among those used on H_i , and these events for $i = 1, 2, \dots, d$ are independent.³ Hence $H[I]$ is connected with probability at most $(1 - 2^{-b})^d \leq \exp(-2^{-b}d) = o(1)$ as $n \rightarrow \infty$.

It remains to verify that all cuts in H have sufficiently many colors. To this end, we show that any subgraph of H on a set $A \subseteq V(H)$ of m vertices has an average degree of at most $((b-1)/\log b) \log m$, generalizing the well-known isoperimetric inequality for the cube corresponding to the case $b = 2$ (then the average number of edges going from a vertex of A to the complement of A , $|A| \leq \frac{1}{2}|V(H)|$, is at least $d(b-1) - ((b-1)/\log b) \log |A| = ((b-1)/\log b) \log(n/|A|) \geq \gamma(n, k, A, V(H) \setminus A)$). This can be proved by induction; namely, by induction on i we prove that any subgraph on at most m vertices of $H_1 \square \dots \square H_i$ has an average degree of at most $((b-1)/\log b) \log m$. For $i = 1$ this holds because the average degree is at most $m-1$, and for $1 \leq m \leq b$ we have $(m-1)/\log m \leq (b-1)/\log b$. For $i > 1$ we consider the partition $A = A_1 \dot{\cup} \dots \dot{\cup} A_b$ of a set $A \subseteq V(H_1 \square \dots \square H_i)$ according to the projection on the last coordinate, we estimate the number of edges on each A_j by induction, and we add the edges among distinct A_j , whose number is at most $\sum_{j_1 < j_2} \min(|A_{j_1}|, |A_{j_2}|)$. Elementary calculations, which we omit, finish the inductive step. $\square$

4. The upper bound. Here we are going to prove Theorem 1.2. In order to prove an upper bound for $F(k)$, we need to construct, for any given k -tuple of distinct directions $u_1, u_2, \dots, u_k$, two point sets A and B with identical X-rays in all of these directions.

The basic approach. All of the constructions are modifications of the basic example mentioned in the introduction, which provides the upper bound $F(k) < 2^{k-1}$. Given $u_1, u_2, \dots, u_k$, we choose suitable nonzero scalars $\alpha_1, \alpha_2, \dots, \alpha_k$, and we define the multisets A', B' as follows:

$$A' := \left\{ v(I) : I \subseteq [k], |I| \text{ even} \right\}, \quad B' = \left\{ v(I) : I \subseteq [k], |I| \text{ odd} \right\},$$

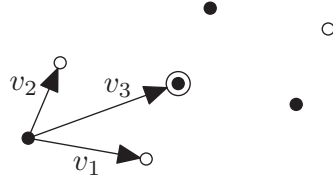
where $v(I) = \sum_{i \in I} v_i$ and $v_i = \alpha_i u_i$. For easier exposition we will call the points of A' *black* and the points of B' *white*. We have $|A'| = |B'| = \frac{1}{2}2^k$ (where each point is counted with multiplicity), and, as was discussed in the introduction, A' and B' have identical X-rays in each of the directions u_i .

In order to improve the upper bound of 2^{k-1} , we try to choose the α_i so that A' and B' have many points in common. After removing the common points the resulting multisets still have identical X-rays, but they are smaller. More precisely, when a black point coincides with a white point, they cancel each other out (so we have to respect multiplicities; if, say, five black points and three white points all occupy the same location, after cancellation we obtain two black points at that location). The resulting multisets after the cancellation will be denoted by A and B .

An example for $k = 3$. The simplest nontrivial example, with $k = 3$, is depicted in Figure 4, and it shows that $F(3) < 3$ (the circled point in the middle is in both A' and B' , and thus it cancels out).

To make this construction work, we need to choose nonzero $\alpha_1, \alpha_2, \alpha_3$ so that $v_1 + v_2 = v_3$, and we should check that this is possible for any triple of distinct directions

³Here the model of picking the colors independently is convenient; if we fixed the number of colors picked, we would not get independence, and the calculations would become more complicated.

FIG. 4. A construction for $k = 3$, using $v_1 + v_2 = v_3$.

u_1, u_2, u_3 . This is easy: Since the vectors u_1 and u_2 define distinct directions, they are linearly independent, and each vector in the plane is a linear combination of them. In particular, we can write $u_3 = \alpha_1 u_1 + \alpha_2 u_2$, and neither of α_1 and α_2 can be 0, for otherwise, u_3 would define the same direction as u_2 or u_1 .

We should also note that A' and B' are always sets (no multiple points) in this particular construction, and hence the upper bound holds for $F(3)$ and not only for $F_{\text{mult}}(3)$. Indeed, if two black points were equal or two white points were equal, then there would be two distinct sets $I, J \subseteq [3]$ with $|I| \equiv |J| \pmod{2}$ and with $v(I) = v(J)$. This yields a linear combination of the u_j equal to 0 and supported on the symmetric difference $I \Delta J$. But we necessarily have $|I \Delta J| = 2$, which would mean that one of the u_j is a multiple of another, which is excluded.

A product argument. Before considering more involved constructions of the type just described, we present a general “product” trick.

LEMMA 4.1. *Let us set $G(k) := 2(F(k) + 1)$. Then we have $G(k + \ell) \leq G(k)G(\ell)$ for any $k, \ell \geq 1$, and in particular, $F(mk) \leq \frac{1}{2}(2F(k) + 2)^m - 1$. An analogous statement holds for F_{mult} .*

Proof. Let $u_1, \dots, u_k$ and $w_1, \dots, w_\ell$ be directions, all of them distinct. Let A_u and B_u be sets of $F(k) + 1$ points each, $A_u \cap B_u = \emptyset$, that have identical X-rays in each of the directions u_i , and, similarly, A_w and B_w are disjoint $(F(\ell) + 1)$ -point sets with identical X-rays in each of the directions w_j .

We choose a sufficiently large real number t and set

$$A := (A_u + tA_w) \dot{\cup} (B_u + tB_w), \quad B := (A_u + tB_w) \dot{\cup} (B_u + tA_w),$$

where “+” represents the Minkowski addition, i.e., $X + Y = \{x + y : x \in X, y \in Y\}$ (and $X + Y$ is understood as a multiset). We have $|A|, |B| = 2(F(k) + 1)(F(\ell) + 1)$, and it is easy to check that A and B have identical X-rays in each of the directions u_i and w_j . Indeed, for the direction u_1 , for example, let $\varphi : A_u \rightarrow B_u$ be a bijection assigning to each point of A_u a point of B_u lying on the same line parallel to u_1 . Each point $a \in A$ can be written as $a_u + ta_w$, $a_u \in A_u$, $a_w \in A_w$, or as $b_u + tb_w$, $b_u \in B_u$, $b_w \in B_w$. Then we define a bijection $\psi : A \rightarrow B$ by $\psi(a_u + ta_w) := \varphi(a_u) + ta_w$, $\psi(b_u + tb_w) := \varphi^{-1}(b_u) + tb_w$, and it is immediate that a and $\psi(a)$ lie on the same line parallel to u_1 .

It remains to verify that, with an appropriate choice of t , A and B are sets and $A \cap B = \emptyset$. We choose t so large that the diameter of $A_u \cup B_u$ is smaller than the distance of any two points in $t(A_w \cup B_w)$. Then if two points $a, a' \in A \cup B$ coincide, we can write $a = x + ty$, $a' = x' + ty'$, $x, x' \in A_u \cup B_u$, $y, y' \in A_w \cup B_w$, and by the choice of t we obtain $y = y'$ and $x = x'$. $\square$

Proof of Theorem 1.2 for sets. By combining Lemma 4.1 with the example above showing $F(3) \leq 2$, we have $F(3m) \leq \frac{1}{2}6^m - 1$, and by monotonicity of F we then

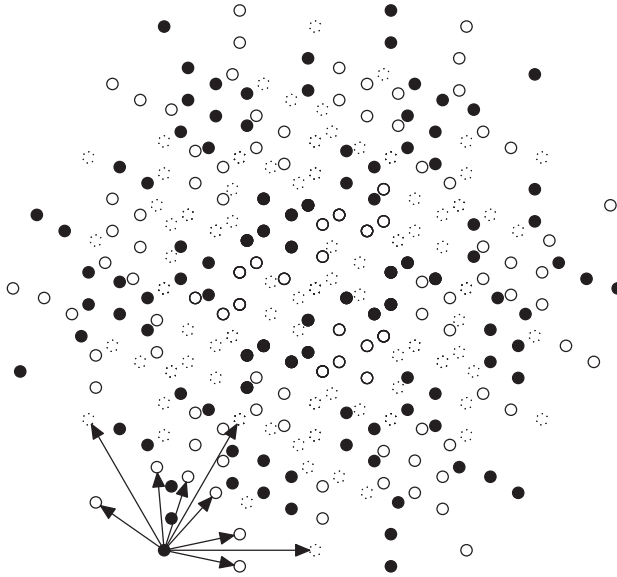


FIG. 5. Sets A and B with the same X -rays in 9 directions, with $|A| = |B| = 99$. Dotted circles denote points that have been cancelled out.

obtain $F(k) = O((6^{1/3})^k)$ for all k . This proves the claim in Theorem 1.2 concerning sets. $\square$

A better construction for multisets. Here we return to the basic construction, depending on a good choice of the α_i . We consider the case $k = 9$. We set $\alpha_9 = 1$, say, and we choose nonzero α_3 and α_6 so that $v_9 = v_3 + v_6$ (by the same argument as in the example for $k = 3$). Then we set the remaining α_i so that $v_1 + v_2 = v_3$, $v_4 + v_5 = v_6$, and $v_7 + v_8 = v_9$.

As we have verified by a computer, in this way 157 pairs of points always cancel out, and we are left with $|A| = |B| \leq 99$ (see Figure 5). Hence $F_{\text{mult}}(9) < 99$, and by Lemma 4.1 we then have $F_{\text{mult}}(k) = O((198^{1/9})^k)$, which verifies the part of Theorem 1.2 concerning F_{mult} .

An example where the construction gives multisets. It can easily be shown that, for a generic choice of $u_1, \dots, u_9$, the construction just given produces sets A' and B' (no two black points and no two white points coincide). However, it turns out that if we insist on the particular dependences among the v_i used in the construction, i.e., on the relations $v_9 = v_3 + v_6$, $v_3 = v_1 + v_2$, $v_4 + v_5 = v_6$, and $v_7 + v_8 = v_9$, then for some choices of the u_i multiple points in A or B cannot be avoided. Indeed, let us consider the following 9 distinct directions:

$$\begin{aligned} u_1 &= (-1, 4), & u_2 &= (2, 3), & u_3 &= (1, 7), & u_4 &= (-1, 5), & u_5 &= (2, 1), \\ u_6 &= (1, 6), & u_7 &= (1, 1), & u_8 &= (1, 12), & u_9 &= (2, 13). \end{aligned}$$

Up to scaling, the only choice of the α_i providing the desired relations among the v_i is $\alpha_1 = \alpha_2 = \dots = \alpha_9 = 1$. But then we also have $u_1 + u_3 = (0, 13) = u_4 + u_6$; i.e., two points in A' coincide. We have also checked that these two black points don't cancel out with any white points, and so A also has multiple points

It seems that, for the choices of u_i that lead to multiple points in this particular construction, we can use alternative constructions, obtained from the present construction by a suitable renumbering of the u_i . But so far we have no systematic approach to such problems.

Concluding remarks. It is clear that the upper bound for F_{mult} can be improved further, by using a larger number of directions in the basic construction. For example, we can take $k = 27$, divide the 27 vectors u_i into 3 groups by 9, apply the choice of the α_i as in the construction for $k = 9$ for each of these groups separately, and then force the additional relation $v_{27} = v_9 + v_{18}$, say. We can then continue and get better and better constructions for k 's of the form 3^m .

But analyzing these constructions seems demanding; already the case $k = 27$ looks challenging even with the help of a computer. Moreover, it appears that the improvements get very small as m grows. There might be a nice theory hiding behind such constructions, but, since this is only one particular class of examples, it is not clear whether understanding them better would be of much help for estimating the functions F and F_{mult} , which was our main objective in this paper.

5. Open problems. Our work was motivated by a question of Holub [11], which in our language can be reformulated as estimating $f(u_1, \dots, u_k)$ for $u_i = (1, i)$ (i.e., u_i is the direction with slope i). The problem emerged in a study of certain systems of equations over free semigroups; originally it deals with a set $F = \{x \mapsto a_i x + b_i : i = 1, 2, \dots, n\}$ of n linear functions, and it asks for what k we can be sure that every F is uniquely determined by the multiset of its values for $x = 1, 2, \dots, k$. An almost identical question appeared a few years later as a problem by Hillar and Levine in the American Mathematical Monthly [10]. In our geometric context, these special directions $(1, i)$ do not seem to have any special significance; yet, even disregarding the algebraic motivation, it would be interesting to have a simple explicit set of directions that allows unique reconstruction of large sets.⁴ Unfortunately, our current methods are unlikely to yield any nontrivial lower bound in this case.

Acknowledgments. We thank Štěpán Holub for raising a problem that brought us to the research reported here, Attila Pór for fruitful discussions concerning the upper bounds, Géza Tóth for the idea of considering the interchange graph, and Noga Alon for pointing out [1].

REFERENCES

- [1] N. ALON, *A note on network reliability*, in Discrete Probability and Algorithms, IMA Vol. Math. Appl. 72, D. Aldous, P. Diaconis, J. Spencer, and J. M. Steele, eds., Springer-Verlag, Berlin, 1995, pp. 11–14.
- [2] E. BARCUCCI, A. DEL LUNGO, M. NIVAT, AND R. PINZANI, *X-rays characterizing some classes of discrete sets*, Linear Algebra Appl., 339 (2001), pp. 3–21.
- [3] G. BIANCHI AND M. LONGINETTI, *Reconstructing plane sets from projections*, Discrete Comput. Geom., 5 (1990), pp. 223–242.
- [4] N. BIGGS, *Algebraic Graph Theory*, 2nd ed., Cambridge University Press, Cambridge, 1993.
- [5] P. DULIO, R. J. GARDNER, AND C. PERI, *Discrete point X-rays*, SIAM J. Discrete Math., 20 (2006), pp. 171–188.
- [6] R. J. GARDNER AND P. GRITZMANN, *Discrete tomography: Determination of finite sets by X-rays*, Trans. Amer. Math. Soc., 349 (1997), pp. 2271–2295.

⁴We can get an “explicit” set of directions from Theorem 1.1, namely, k directions whose coordinates are algebraically independent real numbers, but this is perhaps not what one might call “simple.”

- [7] R. J. GARDNER, *X-rays of polygons*, Discrete Comput. Geom., 7 (1992), pp. 281–293.
- [8] R. J. GARDNER, *Geometric Tomography*, 2nd ed., Western Washington University, Bellingham, 2006.
- [9] A. HEPPEs, *On the determination of probability distributions of more dimensions by thier projections*, Acta Math. Acad. Sci. Hungar., 7 (1956), pp. 403–410.
- [10] C. HILLAR AND L. LEVINE, *Problem 11223*, Amer. Math. Monthly, 5 (2006), p. 459.
- [11] Š. HOLUB, *private communication*, 2003.
- [12] A. RÉNYI, *On projections of probability distributions*, Acta Math. Acad. Sci. Hungar., 3 (1952), pp. 131–142.
- [13] G. TÓTH, *private communication* (through A. Pór), 2003.

MINIMUM COST HOMOMORPHISMS TO SEMICOMPLETE BIPARTITE DIGRAPHS*

GREGORY GUTIN[†], ARASH RAFIEY[‡], AND ANDERS YEO[§]

Abstract. For digraphs D and H , a mapping $f : V(D) \rightarrow V(H)$ is a homomorphism of D to H if $uv \in A(D)$ implies $f(u)f(v) \in A(H)$. If, moreover, each vertex $u \in V(D)$ is associated with costs $c_i(u)$, $i \in V(H)$, then the cost of the homomorphism f is $\sum_{u \in V(D)} c_{f(u)}(u)$. For each fixed digraph H , we have the *minimum cost homomorphism problem for H* . The problem is to decide, for an input graph D with costs $c_i(u)$, $u \in V(D)$, $i \in V(H)$, whether there exists a homomorphism of D to H and, if one exists, to find one of minimum cost. Minimum cost homomorphism problems encompass (or are related to) many well-studied optimization problems. We describe a dichotomy of the minimum cost homomorphism problem for semicomplete bipartite digraphs H . This solves an open problem from an earlier paper. To obtain the dichotomy of this paper, we introduce and study a new notion, a k -Min-Max ordering of digraphs.

Key words. homomorphisms, minimum cost homomorphisms, semicomplete bipartite digraphs

AMS subject classifications. 05C15, 05C20, 90C27

DOI. 10.1137/060668316

1. Introduction.

Motivation. We consider only directed (undirected) graphs that have neither loops nor multiple arcs (edges). In this paper we solve a problem raised in [6] to find a dichotomy for the computational complexity of the minimum cost homomorphism problem (MCH) for semicomplete bipartite digraphs (we define this problem below). In fact, our result leads to a complete dichotomy for the computational complexity of MCH for semicomplete k -partite digraphs ($k \geq 2$) as a (much simpler) dichotomy for the case $k \geq 3$ was obtained in [6] (see also section 5). Our result uses and significantly extends a dichotomy for the computational complexity of MCH for bipartite undirected graphs obtained in [4].

In our previous papers we used properties of an important notion of Min-Max ordering of digraphs. To obtain the dichotomy of this paper, we introduce and study a new notion, a k -Min-Max ordering of digraphs. We believe that properties of this notion and, in particular, Theorem 2.2 can be used to obtain further results on MCH and its special cases. Recent results obtained in [3, 7] and other papers led us to conjecture in [7] that, unless $P=NP$, MCH is polynomial time solvable only when H admits either a Min-Max ordering or a k -Min-Max ordering for some $k \geq 2$.

The MCH was introduced in [8], where it was motivated by a real-world problem in defense logistics. We believe it offers a practical and natural model for the optimization of weighted homomorphisms. MCH's special cases include the well-known

*Received by the editors August 25, 2006; accepted for publication (in revised form) May 17, 2008; published electronically October 17, 2008.

<http://www.siam.org/journals/sidma/22-4/66831.html>

[†]Corresponding author. Department of Computer Science, Royal Holloway University of London, Egham, Surrey TW20 0EX, UK (gutin@cs.rhul.ac.uk). Research of the first author was supported in part by the IST Program of the European Community, under the PASCAL Network of Excellence, IST-2002-506778.

[‡]School of Computing Science, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada (arashr@cs.sfu.ca).

[§]Department of Computer Science, Royal Holloway University of London, Egham, Surrey TW20 0EX, UK (anders@cs.rhul.ac.uk).

list homomorphism problem [10, 12] and the general optimum cost chromatic partition problem, which has been intensively studied [9, 13, 14] and has a number of applications, [16, 17]. A problem more general than MCH was considered in [2].

Minimum cost homomorphisms. For directed or undirected graphs G and H , a mapping $f : V(G) \rightarrow V(H)$ is a *homomorphism of G to H* if $uv \in E(G)$ implies $f(u)f(v) \in E(H)$. The recent treatments of homomorphisms in directed and undirected graphs can be found in [10, 12]. Let H be a fixed directed or undirected graph. The *homomorphism problem* for H asks whether a directed or undirected input graph G admits a homomorphism to H . The *list homomorphism problem* for H asks whether a directed or undirected input graph G with lists (sets) $L_u \subseteq V(H)$, $u \in V(G)$ admits a homomorphism f to H in which $f(u) \in L_u$ for each $u \in V(G)$.

Suppose G and H are directed (or undirected) graphs, and $c_i(u)$, $u \in V(G)$, $i \in V(H)$ are nonnegative *costs*. The *cost of a homomorphism f* of G to H is $\sum_{u \in V(G)} c_{f(u)}(u)$. If H is fixed, the *MCH for H* $\text{MinHOM}(H)$ is the following optimization problem. Given an input graph G , together with costs $c_i(u)$, $u \in V(G)$, $i \in V(H)$, we wish to find a minimum cost homomorphism of G to H or state that none exists.

A bipartite digraph is *semicomplete* if there is at least one arc between every two vertices belonging to different partite sets. In this paper, we study the *MCH for semicomplete bipartite digraphs*, i.e., $\text{MinHOM}(H)$ when H is a *semicomplete bipartite digraph*. Observe that MCH for semicomplete bipartite digraphs extends MCH for bipartite undirected graphs. Indeed, let B be a semicomplete bipartite digraph, with partite sets U, V and arc set $A(B) = A_1 \cup A_2$, where $A_1 = \{uv : u \in U, v \in V\}$ and $A_2 \subseteq \{vu : v \in V, u \in U\}$. Let B' be a bipartite graph, with partite sets U, V and edge set $E(B') = \{uv : vu \in A_2\}$. Notice that $\text{MinHOM}(B)$ is equivalent to $\text{MinHOM}(B')$.

Min-Max ordering. Let H be a digraph. We say that an ordering $v_1, v_2, \dots, v_p$ of $V(H)$ is a *Min-Max ordering* of H if $v_i v_r, v_j v_s \in A(H)$ implies $v_{\min\{i,j\}} v_{\min\{s,r\}} \in A(H)$ and $v_{\max\{i,j\}} v_{\max\{s,r\}} \in A(H)$. One can easily see that $v_1, v_2, \dots, v_p$ of $V(H)$ is a Min-Max ordering of H if $i < j$, $s < r$ and $v_i v_r, v_j v_s \in A(H)$, then $v_i v_s \in A(H)$ and $v_j v_r \in A(H)$. We can define a Min-Max ordering for a bipartite undirected graph G , with partite sets V and U as follows: We orient all edges from V to U and apply the above definition for digraphs. The importance of Min-Max ordering for $\text{MinHOM}(H)$ is indicated in the following two theorems.

THEOREM 1.1 (see [5]). *Let a digraph H have a Min-Max ordering. Then $\text{MinHOM}(H)$ is polynomial time solvable.*

A bipartite graph H with vertices $x_1, x_2, x_3, x_4, y_1, y_2, y_3$ is called

- (a) a *bipartite claw* if its edge set $E(H) = \{x_4 y_1, y_1 x_1, x_4 y_2, y_2 x_2, x_4 y_3, y_3 x_3\}$;
- (b) a *bipartite net* if its edge set $E(H) = \{x_1 y_1, y_1 x_3, y_1 x_4, x_3 y_2, x_4 y_2, y_2 x_2, y_3 x_4\}$;
- (c) a *bipartite tent* if its edge set $E(H) = \{x_1 y_1, x_1 y_2, x_1 y_3, x_2 y_1, x_2 y_3, x_3 y_1, x_4 y_1, x_4 y_2\}$.

See Figure 1.

THEOREM 1.2 (see [4]). *Let H be an undirected bipartite graph. If H contains a cycle C_{2k} , $k \geq 3$ or a bipartite claw or a bipartite net or a bipartite tent as an induced subgraph, then $\text{MinHOM}(H)$ is NP-hard.*

Assume that $P \neq \text{NP}$. Then the following three assertions are equivalent:

- (i) H has a Min-Max ordering;
- (ii) $\text{MinHOM}(H)$ is polynomial time solvable;

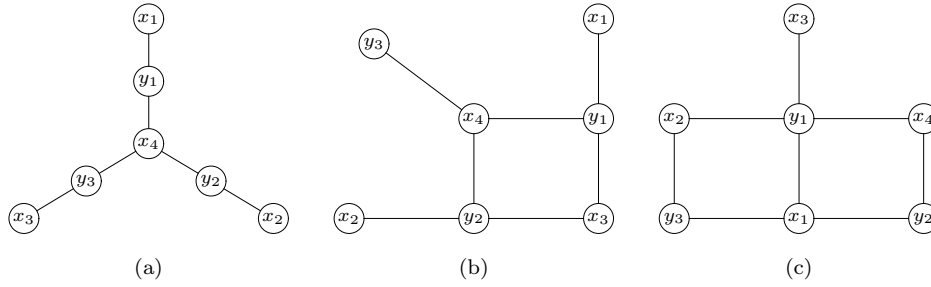


FIG. 1. A bipartite claw (a), a bipartite net (b), and a bipartite tent (c).

- (iii) H does not contain a cycle C_{2k} , $k \geq 3$, a bipartite claw, a bipartite net, or a bipartite tent as an induced subgraph.

Additional terminology and notation. For a graph H , $V(H)$ and $E(H)$ denote its vertex and edge sets, respectively. For a digraph H , $V(H)$ and $A(H)$ denote its vertex and arc sets, respectively. For a pair X, Y of vertex sets of a digraph, H , $(X, Y)_H$ denotes the set of all arcs of the form xy , where $x \in X, y \in Y$. We omit the subscript when it is clear from the context. Also, $X \times Y = \{xy : x \in X, y \in Y\}$. For a set $X \subseteq V(H)$, let $N^+(X) = \{y : \exists x \in X \text{ with } xy \in A(H)\}$ and $N^-(X) = \{y : \exists x \in X \text{ with } yx \in A(H)\}$.

If xy is an arc of a digraph H , we will say that y is an *out-neighbor* of x , and x is an *in-neighbor* of y . We also denote it by $x \rightarrow y$. For disjoint sets $X, Y \subseteq V(H)$, $X \rightarrow Y$ means that $x \rightarrow y$ for each $x \in X$ and $y \in Y$.

An *extension* of a digraph G is a digraph D obtained from G by replacing each vertex u of G by a set of independent vertices $u_1, u_2, \dots, u_{n(u)}$ such that, for a pair u, v of vertices in G , $u_i \rightarrow v_j$ in D if and only if $u \rightarrow v$ in G .

For a bipartite digraph $H = (V, U; A)$, where V and U are its partite sets, $H^{\rightarrow}$ is the subdigraph induced by all arcs directed from V to U , $H^{\leftarrow}$ is the subdigraph induced by all arcs directed from U to V , and $H^{\leftrightarrow}$ is the subdigraph induced by all 2-cycles of H , i.e., by the set $\{xy : xy \in A, yx \in A\}$. A digraph H is an *oriented graph* if it does not contain a 2-cycle, i.e., if $H^{\leftrightarrow}$ is empty. The *converse* of H is the digraph obtained from H by replacing every arc xy with the arc yx .

We denote a directed cycle with p vertices by $\vec{C}_p$. For a set X of vertices of a digraph H , $D[X]$ denotes the subdigraph of H induced by X . For a digraph H , $\text{UN}(H)$ denotes the *underlying graph* of H , i.e., an undirected graph obtained from H by disregarding all orientations and deleting multiple edges.

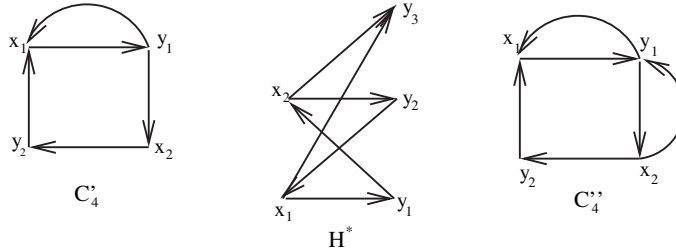
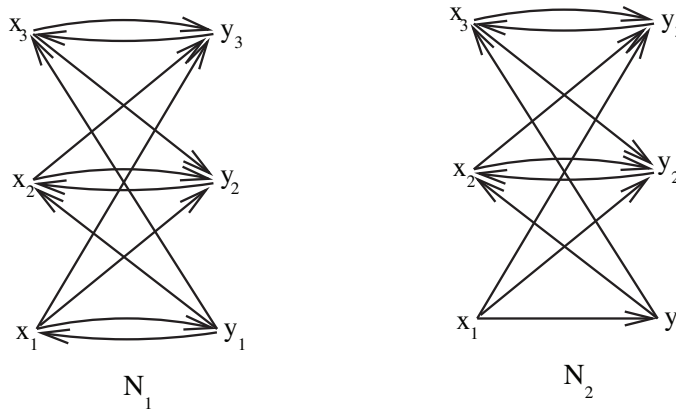
A digraph D is *strong* (or, *strongly connected*) if there is a directed path from x to y and a directed path from y to x for every pair x, y of vertices of D . A *strong component* of a digraph D is a strong maximal induced subdigraph of D .

Forbidden family. Let us introduce five special digraphs for which, as we will see later, the minimum homomorphism problem is NP-hard. The digraphs are depicted in Figures 2 and 3.

The digraph C'_4 has vertices $\{x_1, x_2, y_1, y_2\}$ and arcs $\{x_1y_1, y_1x_2, x_2y_2, y_2x_1, y_1x_1\}$. The digraph C''_4 has the same vertex set, but its arc set is $A(C'_4) \cup \{x_2y_1\}$. The digraph H^* has vertex set $\{x_1, x_2, y_1, y_2, y_3\}$ and arc set $\{x_1y_1, y_1x_2, x_2y_2, y_2x_1, x_1y_3, x_2y_3\}$.

Let N_1 be a digraph, with $V(N_1) = \{x_1, x_2, x_3, y_1, y_2, y_3\}$ and

$$A(N_1) = \{x_1y_1, y_1x_1, x_2y_2, y_2x_2, x_3y_3, y_3x_3, y_1x_2, y_1x_3, x_1y_2, x_1y_3, x_3y_2, x_2y_3\}.$$


 FIG. 2. The digraphs C'_4 , C''_4 , and H^* .

 FIG. 3. The digraphs N_1 and N_2 .

Let N_2 be a digraph, with $V(N_2) = \{x_1, x_2, x_3, y_1, y_2, y_3\}$ and

$$A(N_2) = \{x_1y_1, x_2y_2, y_2x_2, x_3y_3, y_3x_3, y_1x_2, y_1x_3, x_1y_2, x_1y_3, x_3y_2, x_2y_3\}.$$

A digraph H belongs to the family $\mathcal{HFORB}$ if H or its converse is isomorphic to one of the five digraphs above or $\text{UN}(H^s)$ is isomorphic to bipartite claw, bipartite net, bipartite tent, or even cycle with at least 6 vertices, where $s \in \{\rightarrow, \leftarrow, \leftrightarrow\}$.

k -Min-Max ordering. A collection $V_1, V_2, \dots, V_k$ of subsets of a set V is called a k -partition of V if $V = V_1 \cup V_2 \cup \dots \cup V_k$, $V_i \cap V_j = \emptyset$ provided $i \neq j$.

DEFINITION 1.3. Let $H = (V, A)$ be a digraph, and let $k \geq 2$ be an integer. We say that H has a k -Min-Max ordering if there is a k -partition of V into subsets $V_1, V_2, \dots, V_k$ and there is an ordering $v_1^i, v_2^i, \dots, v_{\ell(i)}^i$ of V_i for each i such that

- (i) Every arc of H is an (V_i, V_{i+1}) -arc for some $i \in \{1, 2, \dots, k\}$,
- (ii) $v_1^i, v_2^i, \dots, v_{\ell(i)}^i, v_1^{i+1}, v_2^{i+1}, \dots, v_{\ell(i+1)}^{i+1}$ is a Min-Max ordering of the subdigraph $H[V_i \cup V_{i+1}]$ for all $i \in \{1, 2, \dots, k\}$,

where all indices $i+1$ are taken modulo k . We call the ordering

$$v_1^1, v_2^1, \dots, v_{\ell(1)}^1, v_1^2, v_2^2, \dots, v_{\ell(2)}^2, \dots, v_1^k, v_2^k, \dots, v_{\ell(k)}^k$$

a k -Min-Max ordering of H .

Note that, if H is a strong digraph in which the greatest common divisor of all cycle lengths is k , then $V(H)$ has a k -partition, $k \geq 2$, satisfying (i) (see Theorem 10.5.1 in [1]). A simple example of a digraph having a k -Min-Max ordering is an extension of $\vec{C}_k$.

Dichotomy and paper organization. The main result of this paper is the following.

THEOREM 1.4. *Let H be an semicomplete bipartite digraph. If H contains a digraph from $\mathcal{HFORB}$ as an induced subdigraph, then $\text{MinHOM}(H)$ is NP-hard.*

Assume that $P \neq \text{NP}$. Then the following three assertions are equivalent:

- (i) *$\text{MinHOM}(H)$ is polynomial time solvable;*
- (ii) *H does not contain a digraph from $\mathcal{HFORB}$ as an induced subdigraph;*
- (iii) *Each component of H has a k -Min-Max ordering for $k = 2$ or 4 .*

Theorem 1.4 follows from Corollaries 3.5 and 4.6.

The rest of the paper is organized as follows. In section 2, we study properties of k -Min-Max orderings. In section 3, we prove polynomial cases of $\text{MinHOM}(H)$, when H is a semicomplete bipartite digraph. In section 4, we establish NP-hard cases of the problem. In section 5 we formulate a dichotomy for the computational complexity of $\text{MinHOM}(H)$, when H is a semicomplete multipartite digraph. Section 6 provides a short discussion of further research.

2. Properties of k -Min-Max orderings. Digraphs having k -Min-Max ordering have a very special structure as described in the following lemma.

LEMMA 2.1. *If a digraph $H = (V, A)$ has a k -Min-Max ordering as described in Definition 1.3 and $I = \{x \in V : d^-(x) = 0\}$, then we can define $0 \leq L(i, j) < R(i, j) \leq \ell(j) + 1$ for all i and j such that the following holds:*

- (a) *We have $N^+(v_i^j) = \{v_{L(i,j)+1}^{j+1}, v_{L(i,j)+2}^{j+1}, \dots, v_{R(i,j)-1}^{j+1}\} \setminus I$ for each $v_i^j \in V$ with positive out-degree, where all superscripts are taken modulo k ;*
- (b) *For all j and $i < i'$, we have $R(i, j) \leq R(i', j)$ and $L(i, j) \leq L(i', j)$.*

Proof. If $d_H^+(v_i^j) > 0$, then let $L(i, j)$ be the maximum number such that $N^+(v_i^j) \cap \{v_1^{j+1}, v_2^{j+1}, \dots, v_{L(i,j)}^{j+1}\} = \emptyset$, and let $R(i, j)$ be the minimum number such that $N^+(v_i^j) \cap \{v_{R(i,j)}^{j+1}, v_{R(i,j)+1}^{j+1}, \dots, v_{\ell(j)}^{j+1}\} = \emptyset$. If $d_H^+(v_i^j) = 0$, then if $i = 1$, we let $L(i, j) = 1 = R(i, j) - 1$ and if $i > 1$, we let $L(i, j) = L(i-1, j)$ and $R(i, j) = R(i-1, j)$.

Suppose that, for some m and $p \geq 2$, we have $v_m^j v_m^{j+1}, v_i^j v_{m+p}^{j+1} \in A$, but $v_i^j v_{m+1}^{j+1}, \dots, v_i^j v_{m+p-1}^{j+1} \notin A$ and $v_{m+s}^{j+1} \notin I$, where $1 \leq s \leq p-1$. Since $v_m^{j+s} \notin I$, there is an arc $v_i^j v_m^{j+s}$ in H . By the definition of a k -Min-Max ordering $v_i^j v_m^{j+1} \in A$, a contradiction that proves (a).

We now prove (b). If $d^+(v_i^j) > 0$ and $d^+(v_{i'}^j) > 0$, then this follows from the k -Min-Max ordering using an analogous argument as above. If $i = 1$ and $d^+(v_i^j) = 0$, then we also see that (b) holds. We now prove the remaining cases by induction on $i + i'$. If $d^+(v_i^j) = 0$, consider $i-1$ instead of i and if $d^+(v_{i'}^j) = 0$, then consider $i'-1$ instead of i' . If the new i and i' are equal, then (b) holds, and otherwise we are done by induction. $\square$

The construction used in the following theorem was inspired by somewhat similar constructions in [15] and [2].

THEOREM 2.2. *If a digraph H has a k -Min-Max ordering, then $\text{MinHOM}(H)$ is polynomial time solvable.*

Proof. Let H have a k -Min-Max ordering. Let $V_1, V_2, \dots, V_k$ be defined as in Definition 1.3. If there is a homomorphism h of D to H , then each $h^{-1}(V_i)$, $i = 1, 2, \dots, k$ is an independent set. Thus, we may say that the vertices of $h^{-1}(V_i)$ are of color i . We may assume that $\text{UN}(D)$ is connected, as otherwise we consider a minimum cost homomorphism of each component of D (corresponding to a component of $\text{UN}(D)$) separately.

Let $z \in V(D)$ be an arbitrary vertex. We will show how to decide whether there is a homomorphism f of D to H such that $f(z) \in V_i$ for each $i \in \{1, 2, 3, \dots, k\}$. If such a homomorphism exists for a given i , then we will furthermore show how to find a minimum cost homomorphism of D to H such that $f(z) \in V_i$. We may assume without loss of generality that $i = 1$.

Now assign color 1 to z . Assign every out-neighbor of z color 2 and each in-neighbor of z color k . For every vertex z' with color i , we assign every out-neighbor of z' color $i + 1$ modulo k and every in-neighbor of z' color $i - 1$ modulo k . If some vertex of D is assigned different colors, then clearly there is no homomorphism of D to H , so assume that this is not the case.

Consider the k -partition $G_1, G_2, \dots, G_k$ of $V(D)$ such that all vertices of G_j were assigned color j for each $j \in \{1, 2, \dots, k\}$. Clearly, all arcs in D are (G_j, G_{j+1}) -arcs, where all indices are taken modulo k and $j \in \{1, 2, 3, \dots, k\}$. Note that, if f is a homomorphism of D to H such that $f(z) \in V_1$, then for all $y_j \in G_j$, we must have $f(y_j) \in V_j$.

We will build a weighted digraph (network) $\mathcal{L}$, with vertex set $\cup_{j=1}^k (G_j \times V_j)$ together with two other vertices denoted by s and t . We will also denote t by $(x, v_{\ell(j)+1}^j)$ for every $j \in \{1, 2, \dots, k\}$. The weighted arcs of $\mathcal{L}$ are as follows, where M is any constant greater than the cost of a minimum cost homomorphism of D to H .

- An arc from s to (x, v_1^j) , of weight ∞ , for each $x \in G_j$.
- An arc, $e_{i,j}^x$, from (x, v_i^j) to (x, v_{i+1}^j) , for each $x \in G_j$ and $i \in \{1, 2, \dots, \ell(j)\}$. Recall that, when $i = \ell(j)$, the arc enters t . We define the weight of $e_{i,j}^x$ as follows. If $d_D^-(x) > 0$ and $d_H^-(v_i^j) = 0$, then the weight is ∞ . If $d_D^+(x) > 0$ and $d_H^+(v_i^j) = 0$, then the weight is ∞ . Otherwise, the weight is $c_{v_i}(x) + M$.
- An arc from (x, v_i^j) to $(y, v_{L(i,j)+1}^{j+1})$ and an arc from $(y, v_{R(i,j)}^{j+1})$ to (x, v_{i+1}^j) for every $xy \in A(D)$, with $x \in G_j$ and every $i = 1, 2, \dots, \ell(j)$. Furthermore the weight of these arcs are ∞ .

A *cut* in $\mathcal{L}$ is a partition of the vertices into two sets S and T such that $s \in S$ and $t \in T$ and the weight of a cut is the sum of the weights of all arcs going from a vertex of S to a vertex of T . We will show that a minimum weight cut in $\mathcal{L}$ has weight equal to the minimum cost homomorphism f of D to H such that $f(z) \in V_1$ plus $|V(D)|M$, if such a homomorphism exists. Otherwise, a minimum weight cut in $\mathcal{L}$ is of weight at least $|V(D)|M + M$.

Assume f is a minimum cost homomorphism of D to H such that $f(z) \in V_1$, and assume that $f(x) = v_{a(x)}^j$ for each $x \in G_j$ and for all $j = 1, 2, \dots, k$. Define a cut in $\mathcal{L}$ as follows: $S = \{(x, v_i^j) : i \leq a(x), j = 1, 2, \dots, k\} \cup \{s\}$ and $T = V(\mathcal{L}) - S$. Since f is a homomorphism of D to H , we note that, if $d_D^-(x) > 0$ then $d_H^-(v_{a(x)}^j) > 0$ and if $d_D^+(x) > 0$, then $d_H^+(v_{a(x)}^j) > 0$. Therefore the arcs from $(x, v_{a(x)}^j)$ to $(x, v_{a(x)+1}^j)$ belong to the cut and contribute $c_{f(x)}(x) + M$ to the weight of the cut (and not ∞). We will now show that there are no arcs of infinite weight in the cut, which would imply that the weight of S is exactly the cost of a minimum cost homomorphism from D to H plus $|V(D)|M$.

Clearly no arc out of s belongs to the cut S . Assume for the sake of contradiction that the arc (x, v_i^j) to $(y, v_{L(i,j)+1}^{j+1})$ belongs to the cut S for some $xy \in A(D)$, with $x \in G_j$. This implies that $a(x) \geq i$ (as $(x, v_i^j) \in S$) and $a(y) < L(i, j) + 1$ (as $(y, v_{L(i,j)+1}^{j+1}) \notin S$). By Lemma 2.1(b), this implies that $a(y) \leq L(i, j) \leq L(a(x), j)$. Thus, there is no arc from $v_{a(x)}^j$ to $v_{a(y)}^{j+1}$ in H , by the definition of $L(i, j)$. This is a contradiction to f being a homomorphism.

Now assume for the sake of contradiction that the arc $(y, v_{R(i,j)}^{j+1})$ to (x, v_{i+1}^j) belongs to the cut S for some $xy \in A(D)$, with $x \in G_j$. This implies that $a(x) < i+1$ (as $(x, v_{i+1}^j) \notin S$) and $a(y) \geq R(i, j)$ (as $(y, v_{R(i,j)}^{j+1}) \in S$). By Lemma 2.1(b), this implies that $a(y) \geq R(i, j) \geq R(a(x), j)$. Thus, there is no arc from $v_{a(x)}^j$ to $v_{a(y)}^{j+1}$ in H , by the definition of $R(i, j)$. This is a contradiction to f being a homomorphism. We have now proved that the cut S has the stated weight when there is a homomorphism of D to H , with $f(z) \in V_1$.

Assume that S' is a cut of weight less than $M + |V(D)|M$, and note that the cut S' contains exactly one arc of the form $(x, v_i^j)(x, v_{i+1}^j)$ for each $x \in V(D)$. Therefore we may define a mapping f' from $V(D)$ to $V(H)$ by letting $f'(x) = v_i^j$ if and only if $(x, v_i^j) \in S'$ and $(x, v_{i+1}^j) \notin S'$. We will now show that f' is a homomorphism of D to H of cost equal to the weight of the cut S' minus $|V(D)|M$.

Let xy be any arc in D , and assume without loss of generality that $x \in G_j$. Let i_x and i_y be defined such that $(x, v_{i_x}^j) \in S'$ and $(x, v_{i_x+1}^j) \notin S'$ (i.e., $f'(x) = v_{i_x}^j$) and $(y, v_{i_y}^{j+1}) \in S'$ and $(y, v_{i_y+1}^{j+1}) \notin S'$ (i.e., $f'(y) = v_{i_y}^{j+1}$). As the arc $(x, v_{i_x}^j)(y, v_{L(i_x,j)+1}^{j+1})$ is not in the cut, we must have $(y, v_{L(i_x,j)+1}^{j+1}) \in S$, which implies that $i_y \geq L(i_x, j) + 1$. Furthermore as the arc $(y, v_{R(i_x,j)}^{j+1})(x, v_{i_x+1}^j)$ is not in the cut, we must have $(y, v_{R(i_x,j)}^{j+1}) \notin S$, which implies that $i_y < R(i_x, j)$. We have now shown that $L(i_x, j) + 1 \leq i_y < R(i_x, j)$. Since $xy \in A(D)$ and the arc $(x, v_{i_x}^j)(x, v_{i_x+1}^j)$ has finite weight, we observe that $d_H^+(v_{i_x}^j) > 0$ and $v_{i_y}^{j+1} \notin I$, where I is defined in Lemma 2.1. Lemma 2.1 now implies that $v_{i_x}^j v_{i_y}^{j+1}$ is an arc in H . Therefore f' is a homomorphism. It is not difficult to see that the cost of f' is indeed equal to weight of the cut S' minus $|V(D)|M$.

It remains to recall that a minimum weight cut in a network can be found in polynomial time [1]. $\square$

3. Polynomial cases. We start from a special case which is of importance when H contains no induced $\vec{C}_4$.

LEMMA 3.1. *Let $H = (V, U; A)$ be a semicomplete bipartite digraph, which does not contain an induced subdigraph belonging to $\mathcal{HFORB}$. Suppose for every $v, v' \in V$, we have $N^+(v) \subseteq N^+(v')$ or $N^+(v') \subseteq N^+(v)$. Then H has a 2-Min-Max ordering.*

Proof. We say that vertices $v_i, v_j \in V$ are *similar* if $N^+(v_i) = N^+(v_j)$. Consider similarity classes $V_1, V_2, \dots, V_s$ of V . Moreover, assume that $N^+(V_i) \subset N^+(V_j)$ for $i < j$. Set $U_1 = N^+(V_1)$ and $U_i = N^+(V_i) - \bigcup_{j=1}^{i-1} U_j$ for each $i > 1$. If $s = 1$, then $\text{UN}(H^{\rightarrow})$ is a complete bipartite graph. By Theorem 1.2, $H^{\leftarrow}$ has a Min-Max ordering π . Observe that the orderings of vertices in V and U given by the relative order of vertices in π generate a 2-Min-Max ordering of H . Assume $s > 1$. We prove the following two claims:

- (1) Let $u_i \in U_i$, $v_j \in V_j$, $j > i$, and $u_i v_j \in A$. Then $U_r \rightarrow v_j$ for each $r > i$ and $u_i \rightarrow V_t$ for each $t < j$.

Proof of (1). By the definition of U_i and V_i , if $r > j$, then $U_r \rightarrow v_j$. Now suppose that $i < r \leq j$ and $u_r v_j \notin A$ for some $u_r \in U_r$. Let $v_i \in V_i$ be arbitrary, and note that $H[\{u_i, v_i, u_r, v_j\}]$ is isomorphic to C'_4 or C''_4 , a contradiction. So $u_r v_j \in A$.

By the definition of U_i and V_i , if $t < i$, then $u_i \rightarrow V_t$. Now suppose that $i \leq t < j$ and that $u_i v_t \notin A$ for some $v_t \in V_t$. Let $u_j \in U_j$ be arbitrary, and note that $H[\{u_i, v_t, u_j, v_j\}]$ is isomorphic to C'_4 or C''_4 , a contradiction.

- (2) If $u, u' \in U$, then $N^+(u) \subseteq N^+(u')$ or $N^+(u') \subseteq N^+(u)$.

Proof of (2). Assume that this is not the case, and there exist $v \in N^+(u) - N^+(u')$ and $v' \in N^+(u') - N^+(u)$. This implies that $u \neq u', v \neq v', uv, u'v' \in A$, and $uv', u'v \notin A$. If $u \in U_i$ and $u' \in U_j$ and $i < j$, then by (1) we note that $u'v \in A$, a contradiction. So, for some i , we must have $\{u, u'\} \subseteq U_i$. Analogously, if $v \in V_a$ and $v' \in V_b$ and $a < b$, then by (1) we note that $u'v \in A$, a contradiction. So, for some j , we must have $\{v, v'\} \subseteq V_j$. If $i > j$, then $U_i \rightarrow V_j$, which is a contradiction, so we must have $i \leq j$.

If $i < j$, then let $u'' \in U_j$ and $v'' \in V_i$ be arbitrary. Note that by (1) we must have the arcs $u''v, u''v', uv'', u'v''$ in H . By the construction of the sets U_i, U_j, V_i, V_j we now note that the underlying graph of $H[\{u, u', u'', v, v', v''\}]^{\leftrightarrow}$ is the 6-cycle $v''uvu''v'u'v''$, a contradiction.

Therefore we must have $i = j$. First assume that $i < s$. Now consider $v_s \in V_s$ and $u_s \in U_s$. By (1) there is no arc from $\{u, u'\}$ to v_s . However, $H[\{u, u', v, v', v_s, u_s\}]$ is either N_1 or N_2 , a contradiction.

Similarly, for the case $i = s$, we derive a contradiction.

Now consider an ordering $u_1, u_2, \dots, u_a$ of the vertices in U and an ordering $v_1, v_2, \dots, v_b$ of the vertices in V , defined as follows. If $i < j$, then $d^+(v_i) \leq d^+(v_j)$ and if $d^+(v_i) = d^+(v_j)$, then $d^-(v_i) \geq d^-(v_j)$. Furthermore when $i < j$, then $d^+(u_i) \leq d^+(u_j)$ and if $d^+(u_i) = d^+(u_j)$, then $d^-(u_i) \geq d^-(u_j)$. Note that the ordering $v_1, v_2, \dots, v_b$ first contains vertices from V_1 , then from V_2 , etc.

We will now show that, for every $i \in \{1, 2, \dots, b\}$, there exists an integer α_i such that $N^+(u_i) = \{v_1, v_2, \dots, v_{\alpha_i}\}$. Suppose this is not the case. Then there exists an arc $u_i v_j$ in H such that $u_i v_{j-1}$ is not an arc in H . Thus, both v_j and v_{j-1} belong to some V_k , as otherwise we have a contradiction to (1). This implies that $d^+(v_j) = d^+(v_{j-1})$ and $d^-(v_{j-1}) \geq d^-(v_j)$. Note that every vertex $u \in N^-(v_{j-1})$ has $N^+(u_i) \subseteq N^+(u)$, by (2). Therefore $u \rightarrow v_j$. However, this implies that $d^-(v_{j-1}) < d^-(v_j)$ (as $u_i \rightarrow v_j$ but $u_i v_{j-1} \notin A(H)$), a contradiction.

Using the fact that $N^+(u_i) = \{v_1, v_2, \dots, v_{\alpha_i}\}$ for each $i \in \{1, 2, \dots, b\}$ and that $\alpha_i \geq \alpha_{i-1}$ for each $i \in \{2, 3, \dots, b\}$ (as $d^+(u_i) \geq d^+(u_{i-1})$) and the similar relations for the vertices of V , we can readily conclude that H has a 2-Min-Max ordering. $\square$

The *distance* $\text{dist}(x, y)$ between a pair x, y of vertices in an undirected graph G is the length of the shortest path between x and y . The *diameter* of G is the maximal distance between a pair of vertices in G .

The following theorem shows when $\text{MinHOM}(H)$ is polynomial time solvable if H is strong and does not contain $\tilde{C}_4$ as an induced subdigraph.

THEOREM 3.2. *Let H be a strongly connected semicomplete bipartite digraph. Assume that H does not contain a digraph from $\mathcal{HFORB}$ or $\tilde{C}_4$ as an induced subdigraph. Then H has a 2-Min-Max ordering, and $\text{MinHOM}(H)$ is polynomial time solvable.*

Furthermore, either $\text{UN}(H^{\rightarrow})$ or $\text{UN}(H^{\leftarrow})$ are complete bipartite graphs, or the following holds. For every pair u, u' of distinct vertices of U , we have $N^+(u) \subseteq N^+(u')$ or $N^+(u') \subseteq N^+(u)$ and for every pair v, v' of distinct vertices of V , we have $N^+(v) \subseteq N^+(v')$ or $N^+(v') \subseteq N^+(v)$.

Proof. By Theorem 2.2, to prove the first part of this theorem (before “Furthermore”), it suffices to show that H has a 2-Min-Max ordering. Let V and U be partite sets of H . Denote $H_1 = H^{\rightarrow}$ and $H_2 = H^{\leftarrow}$. It follows from Theorem 1.2 that $\text{UN}(H_1)$ and $\text{UN}(H_2)$ have Min-Max orderings and so do H_1 and H_2 .

Let d_i be the diameter of $\text{UN}(H_i)$, $i = 1, 2$. Observe that if $\text{UN}(H_1)$ ($\text{UN}(H_2)$) is a complete bipartite graph, then, as in the proof of Lemma 3.1, a Min-Max ordering of

$H_2(H_1)$ generates a 2-Min-Max ordering of H . Therefore, $\text{MinHOM}(H)$ is polynomial time solvable by Theorem 2.2. Notice that $\text{UN}(H_i)$ is complete bipartite if and only if $d_i = 2$. Thus, we may assume that both $d_1 \geq 3$ and $d_2 \geq 3$.

We consider the following cases for the value of $d_1 \geq 3$.

Case 1: $d_1 > 4$.

We will show that $\text{UN}(H_2)$ is a complete bipartite graph, or equivalently, $d_2 = 2$. Assume that d_1 is odd, as the case of d_1 even can be considered similarly. Let $P = v_1 u_1 v_2 u_2 \dots v_{k-1} u_{k-1} v_k u_k$ be a shortest path of length d_1 between v_1 and u_k in $\text{UN}(H_1)$. Let $\hat{U} = \{u_1, u_2, \dots, u_k\}$, and $\hat{V} = \{v_1, v_2, \dots, v_k\}$. We will first prove that $\hat{U} \rightarrow \hat{V}$. Since P is a shortest path, we have $u_i \rightarrow v_j$ for each $u_i \in \hat{U}$ and $v_j \in \hat{V}$ provided $j \notin \{i, i+1\}$. Thus, it sufficient to prove

$$(1) \quad u_i \rightarrow v_i \ (i = 1, 2, \dots, k) \text{ and } u_i \rightarrow v_{i+1} \ (i = 1, 2, \dots, k-1).$$

Consider the subdigraph H' of H induced by four vertices $v_i, u_i, v_{i+2}, u_{i+1}$, where $i \in \{1, 2, \dots, k-2\}$. By the definition of P (including the fact that P is a shortest path), we have $v_i u_i, v_{i+2} u_{i+1}, u_{i+1} v_i, u_i v_{i+2} \in A(H)$, but $v_i u_{i+1}, v_{i+2} u_i \notin A(H)$. Since H' is not isomorphic to either $\vec{C}_4$ or C'_4 , we have $u_i v_i, u_{i+1} v_{i+2} \in A(H)$. This proves that $u_i \rightarrow v_i$ provided $i = 1, 2, \dots, k-2$ and $u_i \rightarrow v_{i+1}$ provided $i = 2, 3, \dots, k-1$. Thus, to prove (1) it remains to show that

$$(2) \quad u_{k-1} \rightarrow v_{k-1}, \ u_k \rightarrow v_k, \ u_1 \rightarrow v_2.$$

Consider the subdigraph H'' of H induced by four vertices $v_{k-1}, u_{k-1}, v_k, u_k$. By the definition of P , we have $v_{k-1} u_{k-1}, v_k u_k, v_k u_{k-1}, u_k v_{k-1} \in A(H)$, but $v_{k-1} u_k \notin A(H)$. We have proved that $u_{k-1} \rightarrow v_k$. Since H'' is not isomorphic to C'_4 or C''_4 , we have $u_{k-1} v_{k-1}, u_k v_k \in A(H)$.

Consider $H[\{v_2, v_3, u_1, u_2\}]$. By the definition of P , we have $v_2 u_1, v_2 u_2, v_3 u_2, u_1 v_3 \in A(H)$, but $v_3 u_1 \notin A(H)$. We have proved that $u_2 \rightarrow v_2$. Since $H[\{v_2, v_3, u_1, u_2\}]$ is not isomorphic to C'_4 or C''_4 , we have $u_1 v_2, u_2 v_3 \in A(H)$. This implies that (2), and thus (1), has been proved.

Consider vertex $u \in U - \hat{U}$; we will show that $u \rightarrow \hat{V}$. Suppose this is not true. Let j be the smallest index such that $uv_j \notin A(H)$. We have $v_j u \in A(H)$. Suppose $j > 1$. Since $H[\{u, v_1, u_1, v_j\}]$ is not isomorphic to C'_4 or C''_4 , we have $v_j u_1, v_1 u \in A(H)$. Since P is a shortest path, we have $j = 2$, as otherwise $v_1 u v_j u_{j+1} \dots v_j u_k$ is shorter than P . We have $v_3 u \notin A(H)$, as otherwise $v_1 u v_3 u_3 \dots v_k u_k$ is shorter than P . However, C'_4 is isomorphic to $H[\{v_2, u, v_3, u_3\}]$, a contradiction.

Now assume that $j = 1$. We have $v_3 u \notin A(H)$, as otherwise we have a shorter path. Since H is semicomplete bipartite, we have $uv_3 \in A(H)$. However, $H[\{v_1, u, v_3, u_2\}] \cong C'_4$, a contradiction.

Analogously we can prove that $\hat{U} \rightarrow v$ for every $v \in V - \hat{V}$. Consider $u \in U - \hat{U}$, $v \in V - \hat{V}$. We show that $uv \in A(H)$. Suppose this is not true. We have $vu \in A(H)$. Since $H[\{v, u, v_1, u_1\}]$ is not isomorphic to $\vec{C}_4$, C'_4 , or C''_4 , we have $v_1 u, v u_1 \in A(H)$. Since $H[\{v, u, v_k, u_k\}]$ is not isomorphic to C'_4 or C''_4 , we have $v_k u, v u_k \in A(H)$. But now $\text{dist}(v_1, u_k) = 3$ in $\text{UN}(H_1)$, a contradiction.

Case 2: $d_1 = 4$.

We will show that again $\text{UN}(H_2)$ is a complete bipartite graph. Assume that $v_1 u'_1 v'_2 u'_2 v'_3$ is a shortest path between a pair $v_1 \in V$ and $v'_3 \in U$ in $\text{UN}(H_1)$. Let $U_1 = N^+(v_1)$, $V_2 = N^-(U_1) - \{v_1\}$, $U_2 = N^+(V_2) - U_1$, and $V_3 = N^-(U_2) - V_2$. By the definitions, $V = \{v_1\} \cup V_2 \cup V_3$ and $U = U_1 \cup U_2$. Observe also that $(U_1, V_3) = U_1 \times V_3$,

$(V_3, U_1) = \emptyset$, $(U_2, \{v_1\}) = U_2 \times \{v_1\}$, and $(\{v_1\}, U_2) = \emptyset$. Let $u_1 \in U_1$, $u_2 \in U_2$, $v_2 \in V_2$, and $v_3 \in V_3$ be arbitrary. If $u_2v_3 \notin A(H)$, then $v_3u_2 \in A(H)$, and now $H[\{u_1, v_3, u_2, v_1\}]$ is either $\vec{C}_4$ or C'_4 , a contradiction. Therefore, we have $u_2v_3 \in A(H)$, and consequently $(U_2, V_3) = U_2 \times V_3$. Consider v_3, u_2 , where $v_3u_2 \in A(H)$. Since $H[\{u_1, v_3, u_2, v_1\}]$ is not C'_4 or $\vec{C}_4$, we conclude that $u_1v_1 \in A(H)$, and consequently $(U_1, \{v_1\}) = U_1 \times \{v_1\}$.

Note that we have already proved that the underlying graph of $H_2[U_1 \cup U_2 \cup V_3 \cup \{v_1\}]$ is a complete bipartite graph. Suppose $u_1v_2 \in A(H)$. Then $u_2v_2 \in A(H)$, as otherwise $H[\{u_1, v_1, u_2, v_2\}]$ is isomorphic to C'_4 or C''_4 , a contradiction. Therefore, if $u_1v_2 \in A(H)$, then $(U_2, \{v_2\}) = U_2 \times \{v_2\}$. Thus, to show that $(U_2, V_2) = U_2 \times V_2$, it suffices to prove that every vertex in V_2 has an in-neighbor in U_1 . Suppose this is not true, and let X_2 be the set of all vertices in V_2 that does not have an in-neighbor in U_1 . Let Y_2 be the set of all vertices in U_2 that have an arc into X_2 . As H is strong, some vertex in H must have an arc into X_2 , which implies that $Y_2 \neq \emptyset$.

If there is an arc v_3y_2 from V_3 to Y_2 , then let x_2 be an out-neighbor of y_2 in X_2 , and let $u_1 \in U_1$ be arbitrary. However, this is a contradiction to $H[\{v_3, y_2, x_2, u_1\}]$ not being isomorphic to C'_4 and C''_4 , which implies that there is no arc from V_3 to Y_2 .

As H is strong and there is no arc from V_3 into U_1 or Y_2 , there must be an arc, say v_3u_2 , from V_3 into $U_2 - Y_2$. As H is strong, there must also be an arc from $V(H) - X_2 - Y_2$ into $X_2 \cup Y_2$. By the above this arc, say v_2y_2 , must be from $V_2 - X_2$ to Y_2 . As y_2 belongs to Y_2 , there must be a vertex, say $x_2 \in X_2$, such that $y_2x_2 \in A(H)$. As $v_2 \notin X_2$, we note that there is a vertex, say $u_1 \in U_1$, such that $u_1v_2 \in A(H)$. As $u_1x_2 \notin A(H)$ and $H[\{v_2, x_2, u_1, y_2\}]$ is not isomorphic to C'_4 and C''_4 , we note that $x_2y_2, v_2u_1 \in A(H)$. If $u_2v_2 \notin A(H)$, then $H[\{v_1, v_2, u_1, u_2\}]$ is isomorphic to C'_4 and C''_4 (as $v_1u_2 \notin A(H)$), a contradiction. As $u_2v_2 \in A(H)$ and $H[\{v_2, v_3, u_2, y_2\}]$ is not isomorphic to C'_4 and C''_4 , we note that $y_2v_2 \in A(H)$ (as $v_3y_2 \notin A(H)$). As $H[\{v_2, x_2, u_2, y_2\}]$ is not isomorphic to C'_4 and C''_4 , we note that $v_2u_2 \in A(H)$ (as $u_2x_2 \notin A(H)$). However, the underlying graph of $H[v_1, u_1, v_2, v_3, x_2, y_2, u_2]^{\leftrightarrow}$ is now a bipartite claw, with edges $\{v_2u_1, v_2y_2, v_2u_2, u_1v_1, y_2x_2, u_2v_3\}$, a contradiction. Therefore $U_2 \rightarrow V_2$.

We show that $u_1v_2 \in A(H)$ for every $u_1 \in U_1$ and $v_2 \in V_2$. Suppose this is not true for some $v_2 \in V_2$ and $u_1 \in U_1$. Then we have $v_2u_1 \in A(H)$. Let $v_3 \in V_3$ be arbitrary, and $u_2 \in U_2 \cap N^+(v_3)$. Then $H[\{u_1, v_3, u_2, v_2\}]$ is isomorphic to C'_4 or C''_4 . This completes our proof that $\text{UN}(H_2)$ is a complete bipartite graph.

Case 3: $d_1 = 3$.

Consider a Min-Max ordering π for H_1 . Define vertices x, t, y , and z as follows. Let $\pi(x) = \min\{\pi(x') : x' \in V\}$, $\pi(t) = \max\{\pi(t') : t' \in U\}$, $\pi(y) = \max\{\pi(y') : y' \in U, x \rightarrow y'\}$, and $\pi(z) = \min\{\pi(z') : z' \in V, z' \rightarrow t\}$.

Let $T = N^-(t)$. Since H is strong, every vertex of V has an out-neighbor. Since π is a Min-Max ordering, $z' \rightarrow t$ for each $z' \in V$, with $\pi(z') \geq \pi(z)$. Thus, $T = \{z' \in V : \pi(z') \geq \pi(z)\}$. Let $X = N^+(x)$. Since H is strong, every vertex of U has an in-neighbor. Since π is a Min-Max ordering, $x \rightarrow y'$ for each $y' \in U$, with $\pi(y') \leq \pi(y)$. Thus, $X = \{y' \in U : \pi(y') \leq \pi(y)\}$.

Since $\text{dist}(x, z) = 2$ in $\text{UN}(H_1)$, we have $z \rightarrow y'$ for some $y' \in N^+(x)$. Since π is a Min-Max ordering, $z \rightarrow y$ (consider the arcs zy' and xy). Now, for every $z'' \in V$ with $\pi(z'') \leq \pi(z)$, we have $z'' \rightarrow y$ (as π is a Min-Max ordering). Similarly, for every $y'' \in U$ with $\pi(y'') \geq \pi(y)$, we have $z \rightarrow y''$.

Let $\pi(w) = \min\{\pi(w') : w' \in U\}$. Notice that $\text{dist}(w, t) = 2$ in $\text{UN}(H_1)$. Thus, for some $z' \in T$, we have $z' \rightarrow w$. Hence, $z \rightarrow w$ and $z \rightarrow X$ (as π is a Min-Max ordering). We conclude that $z \rightarrow U$. Similarly, we can obtain that $V \rightarrow y$.

Let $Y = V - T$, and $Z = U - X$. Let $x' \in X$, $y' \in Y$. We have $z \rightarrow x'$ and $y' \rightarrow y$. Hence (as π is a Min-Max ordering), $y' \rightarrow x'$. Thus, $Y \rightarrow X$. Analogously, we can prove that $T \rightarrow Z$.

Let $x' \in X$, $t' \in T$, $y' \in Y$, and $x' \rightarrow t'$. Then $t't$, ty' , $y'x' \in A(H)$. Since $H[\{x', t, t', y'\}]$ is not isomorphic to $\vec{C}_4$, C'_4 , or C''_4 , we have $x' \rightarrow y'$ and $t \rightarrow t'$. Thus, if $x' \rightarrow t'$, we have $x' \rightarrow Y$. Analogously, if $x' \rightarrow t'$, we have $Z \rightarrow t'$. Hence,

$$(3) \quad x' \rightarrow t' \text{ implies } x' \rightarrow Y \text{ and } Z \rightarrow t'.$$

We will now prove the following: for a pair u, u' of distinct vertices of U , we have $N^+(u) \subseteq N^+(u')$ or $N^+(u') \subseteq N^+(u)$. By Lemma 3.1, this implies that H has a 2-Min-Max ordering, and we are done. Suppose that we have neither $N^+(u) \subseteq N^+(u')$ nor $N^+(u') \subseteq N^+(u)$. Thus, there is a pair v, v' of vertices in V such that $u \rightarrow v$, $u' \rightarrow v'$, but uv' and $u'v$ are not arcs in H . Since $H[\{u, u', v, v'\}]$ is not isomorphic to $\vec{C}_4$, C'_4 , and C''_4 , we have $v \rightarrow u$ and $v' \rightarrow u'$. Now consider four cases.

Case 3.1: $v, v' \in Y$. Let $t' \in T$. By the definition of t , $t \notin \{u, u'\}$. If $u' \rightarrow t'$, then $H[\{t, t', v, u'\}]$ is isomorphic to $\vec{C}_4$, C'_4 , or C''_4 , which is impossible. Thus, $u't' \notin A(H)$, and $t' \rightarrow u'$. Analogously, $ut' \notin A(H)$, and $t' \rightarrow u$. By the fact that $t' \rightarrow t$ and the existence and nonexistence of previously considered arcs, we conclude that $H[\{v, v', u, u', t, t'\}]$ is isomorphic to N_1 or N_2 , which is impossible.

Case 3.2: $u, u' \in Z$. We can show that this case is impossible similarly to Case 3.1 but considering x, w instead of t, t' .

Case 3.3: $v \in Y$, $v' \in T$. By (3), $u' \in Z$. By Case 3.2, we may assume that $u \in X$. Then $H[\{v, v', u', t\}]$ is isomorphic to $\vec{C}_4$, C'_4 , or C''_4 , which is impossible.

Case 3.4: $v, v' \in T$. By Case 3.2, we may assume that $u \in X$. By (3), $Z \rightarrow v$, and thus $u' \in X$. By (3), we conclude that uxu , $u'xu'$, vtv , and $v'tv'$ are 2-cycles. Notice that $t \rightarrow x$, but $xt \notin A(H)$. Now it follows that $H[\{x, v, v', u, u', t\}]^{\leftrightarrow}$ is isomorphic to C_6 , a contradiction.

It follows from Cases 1, 2, and 3 that if neither $\text{UN}(H_1)$ nor $\text{UN}(H_2)$ are complete bipartite graphs, then we must have $d_1 = d_2 = 3$. In this case we have shown that, for every pair u, u' of distinct vertices of U , we have $N^+(u) \subseteq N^+(u')$ or $N^+(u') \subseteq N^+(u)$. However, by swapping the roles of U and V , we also get that, for every pair v, v' of distinct vertices of V , we have $N^+(v) \subseteq N^+(v')$ or $N^+(v') \subseteq N^+(v)$. $\square$

The following theorem shows when $\text{MinHOM}(H)$ is polynomial time solvable for the case when H is not strong and does not contain $\vec{C}_4$ as an induced subdigraph.

THEOREM 3.3. *Let $H = (V, U; A)$ be a semicomplete bipartite digraph, with strong components $C_1, C_2, \dots, C_p$ ($p \geq 2$) satisfying the following:*

- *There is no arc from C_i to C_j for $i > j$;*
- *H does not contain an induced subdigraph belonging to $\mathcal{HFORB}$ or an induced directed 4-cycle.*

Then H has a 2-Min-Max ordering, and $\text{MinHOM}(H)$ is polynomial time solvable.

Proof. If there are no $v, v' \in V$ and $u, u' \in U$ such that $A(H^{\rightarrow}[\{v, v', u, u'\}]) = \{vu, v'u'\}$, then $H^{\rightarrow}$ satisfies the condition of Lemma 3.1. Therefore H has a 2-Min-Max ordering.

Thus, we may assume that there exist $v, v' \in V$ and $u, u' \in U$ such that $A(H^{\rightarrow}[\{v, v', u, u'\}]) = \{vu, v'u'\}$. Note that $\{v, v', u, u'\}$ belong to a strong component of H as they are contained in a 4-cycle. Let $\{v, v', u, u'\} \subseteq V(C_t)$ for some t , and let $V_1 = \{x \in V \mid x \in C_i, i < t\}$, $U_1 = \{y \in U \mid y \in C_i, i < t\}$, $V_2 = V \cap C_t$, $U_2 = U \cap C_t$, $V_3 = V - V_1 - V_2$, and $U_3 = U - U_1 - U_2$.

If there is a $w' \in U_3$ and $w \in V_3$, such that $w' \rightarrow w$, then $H[\{u, v, u', v', w, w'\}]$ is the converse of either N_1 or N_2 . Therefore we must have $A(H[U_3 \cup V_3]) = V_3 \times U_3$. Analogously we must have $A(H[U_1 \cup V_1]) = V_1 \times U_1$. Consider $H' = H[U_2 \cup V_2]$, and note that H' is strong and does not contain a digraph from $\mathcal{HFORB}$ or $\vec{C}_4$ as an induced subdigraph. Therefore Theorem 3.2 implies that $U_2 \rightarrow V_2$ (as $V_2 \rightarrow U_2$ is not true), which furthermore implies that $(U_1 \cup U_2) \rightarrow (V_2 \cup V_3)$.

Let π be a Min-Max ordering of $H^\rightarrow$. Let $uv \in A(H)$ and $u'v' \in A(H)$ be two distinct arcs from U to V . As $d^+(u) > 0$, we note that $u \in U_1 \cup U_2$, by the above. Analogously we note that $u' \in U_1 \cup U_2$, $v \in V_2 \cup V_3$ (as $d^-(v) > 0$), and $v' \in V_2 \cup V_3$. By the above, we therefore have $uv', u'v \in A(H)$. As u, u', v, v' were chosen arbitrarily, this implies that π is a 2-Min-Max ordering. $\square$

The following theorem shows when $\text{MinHOM}(H)$ is polynomial time solvable for the case when H does contain $\vec{C}_4$ as an induced subdigraph.

THEOREM 3.4. *Let H be a semicomplete bipartite digraph. Assume that H does not contain a digraph from $\mathcal{HFORB}$ as an induced subdigraph but contains $\vec{C}_4$ as an induced subdigraph. Then H is an extension of a $\vec{C}_4$, and $\text{MinHOM}(H)$ is polynomial time solvable.*

Proof. Observe that an extension L of any cycle $\vec{C}_p$, $p \geq 2$, has a p -Min-Max ordering. Thus, $\text{MinHOM}(L)$ is polynomial time solvable by Theorem 2.2. Let $C = v_1u_1v_2u_2v_1$ be an induced 4-cycle of H . It suffices to prove that H is an extension of C .

For $i = 1, 2$, let $M^+(v_i) = \{u \in U : v_iu \in A(H), uv_i \notin A(H)\}$, $M^-(v_i) = \{u \in U : uv_i \in A(H), v_iu \notin A(H)\}$, and $M(v_i) = \{u \in U : uv_i, v_iu \in A(H)\}$. We have $M^+(v_1) \cap M^+(v_2) = \emptyset$, as otherwise $H[\{v_1, v_2, u_1, u_2, u_3\}] \cong H^*$, where $u_3 \in M^+(v_1) \cap M^+(v_2)$. We have $M^-(v_1) \cap M^-(v_2) = \emptyset$, as otherwise $H[\{v_1, v_2, u_1, u_2, u_3\}] \cong H^{**}$, where $u_3 \in M^-(v_1) \cap M^-(v_2)$ and H^{**} is the converse of H^* .

We have $M(v_1) \cap (M^+(v_2) \cup M^-(v_2)) = \emptyset$, as otherwise $H[\{v_1, u_1, v_2, u\}] \cong C'_4$, where $u \in M(v_1) \cap M^+(v_2)$ or $H[\{v_1, u_2, v_2, u\}] \cong C'_4$, where $u \in M(v_1) \cap M^-(v_2)$. Moreover, $M(v_1) \cap M(v_2) = \emptyset$, as otherwise $H[\{v_1, u_1, v_2, u\}] \cong C''_4$, where $u \in M(v_1) \cap M(v_2)$. The arguments above imply that $M(v_1) = M(v_2) = \emptyset$, $M^+(v_1) = M^-(v_2)$, and $M^-(v_1) = M^+(v_2)$.

Similarly, we can define $M^+(u_i)$, $M^-(u_i)$, and $M(u_i)$, $i = 1, 2$, and prove the relations analogous to those for $M^+(v_i)$, $M^-(v_i)$, and $M(v_i)$, $i = 1, 2$.

Let $v \in V - \{v_1, v_2\}$ and $u \in U - \{u_1, u_2\}$ be arbitrary. Without loss of generality, assume that $u \in M^+(v_1) = M^-(v_2)$ and $v \in M^+(u_2) = M^-(u_1)$ (all other cases can be treated similarly). To show that H is an extension of C , it suffices to prove that $v \rightarrow u$, but $uv \notin A(H)$. Suppose first that $u \rightarrow v$ and $v \rightarrow u$. Then $H[\{v, u, v_2, u_2\}] \cong C'_4$, a contradiction. Now suppose that $u \rightarrow v$, but $vu \notin A(H)$. Then $H[\{v, v_1, v_2, u, u_2\}] \cong H^*$, a contradiction. Thus, $v \rightarrow u$, but $uv \notin A(H)$, and we are done. $\square$

The three theorems of this section and the fact that $\vec{C}_4$ has a 4-Min-Max ordering imply the following.

COROLLARY 3.5. *Let H be a semicomplete bipartite digraph not containing a digraph from $\mathcal{HFORB}$ as an induced subdigraph. Then $\text{MinHOM}(H)$ is polynomial time solvable, and H has a k -Min-Max ordering for $k = 2$ or 4 .*

4. NP-hardness cases. It is well known that the problem of finding a maximum size independent set in an undirected graph G is NP-hard. We say that a set I in a digraph D is *independent* if no vertices in I are adjacent. Clearly, the problem of finding a maximum size independent set in an oriented graph D (MISO) is NP-hard.

LEMMA 4.1. *$\text{MinHOM}(C'_4)$ is NP-hard.*

Proof. Let H be isomorphic to C'_4 as follows:

$$V(H) = \{1, 2, 3, 4\}, \quad A(H) = \{12, 21, 23, 34, 41\}.$$

Let D be an arbitrary oriented graph. We replace every arc uv of D by the digraph G_{uv} , with $V(G_{uv}) = \{u, v, x_{uv}, y_{uv}, z_{uv}\}$ and $A(G_{uv}) = \{ux_{uv}, x_{uv}y_{uv}, y_{uv}z_{uv}, vz_{uv}\}$. Here $x_{uv} \neq x_{u'v'}$, $y_{uv} \neq y_{u'v'}$, and $z_{uv} \neq z_{u'v'}$ for each $uv \neq u'v'$. Let D' be the obtained oriented graph. Define the cost function as follows: $c_1(w) = 1$, $c_2(w) = 1$, $c_3(w) = 0$, and $c_4(w) = 1$ for each $w \in V(D)$, and $c_i(t) = 0$ for each $t \in V(D') \setminus V(D)$ and $i \in \{1, 2, 3, 4\}$.

Let g be a homomorphism of D' to H , and let $S = \{s \in V(D) : g(s) = 3\}$. Observe that the cost of g is $|V(D)| - |S|$, and S is an independent set in D . Indeed, suppose that $uv \in A(D)$ and $\{u, v\} \subseteq S$. Then $g(y_{uv}) = 1$ and $g(z_{uv}) = 4$, a contradiction to the fact that g is a homomorphism of D' to H .

Let I be an independent set in D of maximum size. Set $f(w) = 3$ for each $w \in I$, and $f(w) = 1$ for each $w \in V(D) \setminus I$. For each $uv \in A(D)$, set

- (a) $f(x_{uv}) = 2, f(y_{uv}) = 1$ and $f(z_{uv}) = 2$ if $\{u, v\} \cap I = \emptyset$;
- (b) $f(x_{uv}) = 4, f(y_{uv}) = 1$ and $f(z_{uv}) = 2$ if $u \in I$ and $v \notin I$;
- (c) $f(x_{uv}) = 2, f(y_{uv}) = 3$ and $f(z_{uv}) = 4$ if $u \notin I$ and $v \in I$.

Observe that f is a homomorphism of D' to H , and the cost of f is $|V(D)| - |I|$. By the arguments above, f is a homomorphism of D' to H of minimum cost. Thus, we have reduced MISO to MinHOM(H). $\square$

LEMMA 4.2. *MinHOM(C''_4) is NP-hard.*

Proof. Let H be isomorphic to C''_4 as follows:

$$V(H) = \{1, 2, 3, 4\}, \quad A(H) = \{12, 21, 23, 32, 34, 41\}.$$

Let D be an arbitrary oriented graph. We replace every arc uv of D by the digraph G_{uv} , with $V(G_{uv}) = \{u, v, x_{uv}\}$ and $A(G_{uv}) = \{ux_{uv}, x_{uv}v\}$. Here $x_{uv} \neq x_{u'v'}$ for each $uv \neq u'v'$. Let D' be the obtained oriented graph. Define the cost function as follows: $c_1(w) = 1$, $c_2(w) = 1$, $c_3(w) = 1$, and $c_4(w) = 0$ for each $w \in V(D)$, and $c_i(x) = 0$ for each $x \in V(D') \setminus V(D)$ and $i = 1, 2, 3, 4$.

Let g be a homomorphism of D' to H , and let $S = \{s \in V(D) : g(s) = 4\}$. Observe that the cost of g is $|V(D)| - |S|$, and S is an independent set in D .

Let I be an independent set in D of maximum size. Set $f(w) = 4$ for each $w \in I$, $f(w) = 2$ for each $w \in V(D) \setminus I$, $f(x_{uv}) = 3$ if $v \in I$ and $f(x_{uv}) = 1$, otherwise, for each $uv \in A(D)$. Observe that f is a homomorphism of D' to H , and the cost of f is $|V(D)| - |I|$. By the arguments above, f is a homomorphism of D' to H of minimum cost. Thus, we have reduced MISO to MinHOM(H). $\square$

The following lemma was stated in [6]. We give a proof here for the sake of completeness.

LEMMA 4.3. *MinHOM(H^*) is NP-hard.*

Proof. Let H be isomorphic to H^* as follows:

$$V(H) = \{1, 2, 3, 4, 5\}, \quad A(H) = \{12, 23, 34, 41, 15, 35\}.$$

We replace every arc uv of an oriented graph D by the digraph G_{uv} , with $V(G_{uv}) = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$, where $u = v_6$ and $v = v_7$, and

$$A(G_{uv}) = \{v_1v_2, v_2v_3, v_3v_4, v_4v_1, v_5v_6, v_5v_7, v_1v_6, v_3v_7\}.$$

Let D' be the obtained oriented graph. Define the cost function as follows: $c_4(w) = 0$ and $c_i(w) = 1$ for each $w \in V(D)$ and $i \in \{1, 2, 3, 5\}$, and $c_i(t) = 0$ for each $t \in V(D') \setminus V(D)$ and $i \in \{1, 2, 3, 4, 5\}$.

Let g be a homomorphism of D' to H , and let $S = \{s \in V(D) : g(s) = 4\}$. Observe that the cost of g is $|V(D)| - |S|$, and S is an independent set in D . Indeed, suppose that $uv \in A(D)$ and $\{u, v\} \subseteq S$. Then $g(v_1) = g(v_3) = 3$, and $g(v_2)$ cannot be given any value in $V(H)$ since g is a homomorphism.

Let I be an independent set in D of maximum size. Set $f(w) = 4$ for each $w \in I$ and $f(w) = 5$ for each $w \in V(D) \setminus I$. For each $uv \in A(D)$, set

- (a) $f(v_5) = 3$ and $f(v_i) = i$ for every $i = 1, 2, 3, 4$ if $u \notin I$,
- (b) $f(v_5) = 3$ and $f(v_i) = i + 2$ for every $i = 1, 2, 3, 4$, where $i + 2$ is taken modulo 4, if $u \in I$.

Observe that f is a homomorphism of D' to H , and the cost of f is $|V(D)| - |I|$. By the arguments above, f is a homomorphism of D' to H of minimum cost. Thus, we have reduced MISO to MinHOM(H). $\square$

LEMMA 4.4. *MinHOM(N_1) is NP-hard.*

Proof. Let H be the following digraph isomorphic to N_1 : $V(H) = \{1, 2, 3, 4, 5, 6\}$ and

$$A(H) = \{12, 21, 34, 43, 56, 65, 23, 25, 14, 16, 54, 36\}.$$

Let D be an arbitrary oriented graph. We replace every arc uv of D with the digraph G_{uv} , with $V(G_{uv}) = \{x, y, z, u, v\}$ and $A(G_{uv}) = \{ux, vz, xy, yz\}$. Consider the following cost function: $c_2(u) = c_2(v) = 1$, $c_6(u) = c_6(v) = 0$, $c_i(u) = c_i(v) = M + 1$ for $i \neq 2, 6$ and $c_6(y) = M + 1$, where $M = |V(D)|$. In all remaining cases the cost is zero.

Let D' be the obtained oriented graph, let f be a mapping from $V(D')$ to $V(H)$, and let uv be an arc in D . Assume that $f(u) = f(v) = 2$. Then with $f(x) = f(z) = 1$ and $f(y) = 2$, we obtain a homomorphism from G_{uv} to H of cost 2. This implies that there is a homomorphism of D' to H of cost $M < M + 1$, and thus every vertex of D in D' must be colored either 2 or 6 in any minimum cost homomorphism of D' to H . Let f be a homomorphism of D' to H , and let us consider the remaining options for coloring the vertices of D in D' .

Assume that $f(v) = 6$ and $f(u) = 2$. Then with $f(z) = 5$, $f(y) = 2$, and $f(x) = 1$, we obtain a homomorphism from G_{uv} to H of cost 1. Assume that $f(v) = 2$ and $f(u) = 6$. Then with $f(x) = 5$, $f(y) = 4$, and $f(z) = 3$, we obtain a homomorphism from G_{uv} to H of cost 1. Note that if $f(u) = f(v) = 6$, then $f(x) = f(z) = 5$ and $f(y) = 6$. Then the cost of f will be at least $M + 1$ implying we cannot color both vertices u and v in color 6 in any minimum cost homomorphism of D' to H .

Now let f be a minimum cost homomorphism, let S be the vertices of D in D' colored 6, and let $T = V(D) \setminus S$. Recall that the vertices of T are colored 2. Notice that S is an independent set, and the cost of f equals $|T|$. Thus, S is an independent set of maximum size, and we have reduced MISO to MinHOM(H). $\square$

LEMMA 4.5. *MinHOM(N_2) is NP-hard.*

Proof. Let H be the following digraph isomorphic to N_2 : $V(H) = \{1, 2, 3, 4, 5, 6\}$ and

$$A(H) = \{12, 34, 43, 56, 65, 23, 25, 14, 16, 54, 36\}.$$

Let D be an arbitrary oriented graph. We replace every arc uv of D by the digraph G_{uv} , with $V(G_{uv}) = \{x, y, z, u, v\}$ and $A(G_{uv}) = \{ux, vz, xy, zy\}$. We introduce the following cost function: $c_1(u) = c_1(v) = 1$, $c_5(u) = c_5(v) = 0$, $c_i(u) = c_i(v) = M + 1$ for $i \neq 1, 5$ and $c_4(x) = c_6(z) = M + 1$, where $M = |V(D)|$. In any other cases the

cost is zero. Let D' be the obtained oriented graph, let f be a mapping from $V(D')$ to $V(H)$, and let uv be an arc in D .

Assume that $f(u) = f(v) = 1$. With $f(x) = f(z) = 2$ and $f(y) = 3$, we obtain a homomorphism from H' to H with cost 2. Thus, there is a homomorphism of D' to H of cost at most M (assign all vertices of D in D' color 1), and no vertex of D in D' must not be assigned any color other than 1 and 5.

Let f be a homomorphism of D' to H , and let us consider the remaining options for coloring the vertices of D in D' . Assume that $f(v) = 1$ and $f(u) = 5$. With $f(z) = 2$, $f(x) = 6$, and $f(y) = 5$, we obtain a homomorphism of G_{uv} to H of cost 1. Assume that $f(v) = 5$ and $f(u) = 1$. With $f(x) = 2$, $f(z) = 4$, and $f(y) = 3$, we obtain a homomorphism of G_{uv} to H of cost 1. Note that if $f(u) = f(v) = 5$, then $f(x) \in \{4, 6\}$ and $f(z) \in \{4, 6\}$. Thus, f has cost at least $M + 1$ implying that a minimum cost homomorphism of D' to H does not assign adjacent vertices of D color 5 (in D').

Now let f be a minimum cost homomorphism, let S be the vertices of D in D' colored 5 and $T = V(D) \setminus S$. Recall that the vertices of T are colored 1. Notice that S is an independent set, and the cost of f equals $|T|$. Thus, S is an independent set of maximum size, and we have reduced MISO to MinHOM(H). $\square$

COROLLARY 4.6. *MinHOM(H) is NP-hard for every $H \in \mathcal{HFORB}$.*

Proof. If H is isomorphic to C'_4 , C''_4 , H^* , N_1 or N_2 , or the converse of one of the five digraphs, then MinHOM(H) is NP-hard due to the lemmas of this section and the simple fact that if MinHOM(H) is NP-hard and H' is the converse of H , then MinHOM(H') is NP-hard as well.

Let $\mathcal{B}$ be the set consisting of the following bipartite graphs: bipartite claw, bipartite net, bipartite tent, and every even cycle with at least 6 vertices. If $\text{UN}(H^s)$, where $s \in \{\rightarrow, \leftarrow\}$, is isomorphic to a graph in $\mathcal{B}$, then MinHOM(H) is NP-hard due to Theorem 1.2 and the transformation from a bipartite undirected graph to a semi-complete bipartite digraph described in the last paragraph of subsection “Minimum Cost Homomorphisms” of section 1. If $\text{UN}(H^{\leftrightarrow})$ is isomorphic to a graph in $\mathcal{B}$, then MinHOM(H) is NP-hard as, for each bipartite undirected graph L , MinHOM(L) is equivalent to MinHOM(L^+), where L^+ is the digraph obtained from L by replacing every edge xy with two arcs xy and yx . $\square$

5. Dichotomy for semicomplete multipartite digraphs. A digraph D is called *semicomplete k -partite* if D can be obtained from a complete k -partite (undirected) graph G by replacing every edge xy of G by either the arc xy or the arc yx or the pair xy, yx of arcs. Let TT_p denote the acyclic tournament on $p \geq 1$ vertices. Let $p \geq 3$, and let TT_p^- be a digraph obtained from TT_p by deleting the arc from the vertex of in-degree zero to the vertex of out-degree zero. Combining the main result of this paper with the main result of [6], we obtain the following.

THEOREM 5.1. *Let H be a semicomplete k -partite digraph. If $k = 2$ and H does not contain a digraph from $\mathcal{HFORB}$ as an induced subdigraph or if $k \geq 3$ and H is an extension of either TT_k or TT_{k+1}^- or $\vec{C}_3$, then MinHOM(H) is polynomial time solvable. Otherwise, MinHOM(H) is NP-hard.*

6. Further research. In the case of undirected graphs H , the well-known theorem of Hell and Nešetřil [11] on the homomorphism problem implies that MinHOM(H) is NP-hard for each nonbipartite graph H . The authors of [4] obtained a complete dichotomy of the computational complexity of MinHOM(H) when H is undirected. The dichotomy obtained in this paper significantly extends the dichotomy of [4]. This indicates that the problem of obtaining a dichotomy for the computational complex-

ity of $\text{MinHOM}(H)$ when H is a bipartite digraph is a very difficult problem. Note that $\text{MinHOM}(H)$ is polynomial time solvable for some nonbipartite digraphs, for example, for acyclic tournaments [6]. Thus, a dichotomy for the bipartite-directed case does not coincide with a dichotomy for the general-directed case. The problem of obtaining dichotomy for both cases is a very interesting open problem.

Acknowledgments. We are grateful to Pavol Hell and Eun Jung Kim for several useful discussions.

REFERENCES

- [1] J. BANG-JENSEN AND G. GUTIN, *Digraphs: Theory, Algorithms and Applications*, Springer, London, 2000.
- [2] D. COHEN, M. COOPER, P. JEAUVONS, AND A. KROKHIN, *A maximal tractable class of soft constraints*, J. Artificial Intelligence Res., 22 (2004), pp. 1–22.
- [3] A. GUPTA, P. HELL, M. KARIMI, AND A. RAFIEY, *Minimum cost homomorphisms to reflexive digraphs*, in Proc. LATIN 2008, Lecture Notes Comput. Sci. 4957, Springer, New York, 2008, pp. 182–193.
- [4] G. GUTIN, P. HELL, A. RAFIEY, AND A. YEO, *Dichotomy for minimum cost graph homomorphisms*, European J. Combin., 29 (2008), pp. 900–911.
- [5] G. GUTIN, A. RAFIEY, AND A. YEO, *Minimum cost and list homomorphisms to semicomplete digraphs*, Discrete Appl. Math., 154 (2006), pp. 890–897.
- [6] G. GUTIN, A. RAFIEY, AND A. YEO, *Minimum cost homomorphisms to semicomplete multipartite digraphs*, Discrete Appl. Math., 156 (2008), pp. 2429–2435.
- [7] G. GUTIN, A. RAFIEY, AND A. YEO, *Minimum cost homomorphism dichotomy for oriented cycles*, Proc. AAIM 2008, Lecture Notes Comput. Sci. 5034, Springer, New York, 2008, pp. 224–234.
- [8] G. GUTIN, A. RAFIEY, A. YEO, AND M. TSO, *Level of repair analysis and minimum cost homomorphisms of graphs*, Discrete Appl. Math., 154 (2006), pp. 881–889.
- [9] M. M. HALLDORSSON, G. KORTSARZ, AND H. SHACHNAI, *Minimizing average completion of dedicated tasks and interval graphs*, in Approximation, Randomization, and Combinatorial Optimization, Lecture Notes Comput. Sci. 2129, Springer, New York, 2001, pp. 114–126.
- [10] P. HELL, *Algorithmic aspects of graph homomorphisms*, in Survey in Combinatorics 2003, London Math. Soc. Lecture Note Ser. 307, Cambridge University Press, Cambridge, 2003, pp. 239–276.
- [11] P. HELL AND J. NEŠETŘIL, *On the complexity of H -coloring*, J. Combin. Theory Ser. B, 48 (1990), pp. 92–110.
- [12] P. HELL AND J. NEŠETŘIL, *Graphs and Homomorphisms*, Oxford University Press, Oxford, 2004.
- [13] K. JANSEN, *Approximation results for the optimum cost chromatic partition problem*, J. Algorithms, 34 (2000), pp. 54–89.
- [14] T. JIANG AND D.B. WEST, *Coloring of trees with minimum sum of colors*, J. Graph Theory, 32 (1999), pp. 354–358.
- [15] S. KHANNA, M. SUDAN, L. TREVISAN, AND D. WILLIAMSON, *The approximability of constraint satisfaction problems*, SIAM J. Comput., 30 (2001), pp. 1863–1920.
- [16] L.G. KROON, A. SEN, H. DENG, AND A. ROY, *The optimal cost chromatic partition problem for trees and interval graphs*, in WG 1996, Lecture Notes Comput. Sci. 1197, Springer, New York, 1997, pp. 279–292.
- [17] K. SUPOWIT, *Finding a maximum planar subset of a set of nets in a channel*, IEEE Trans. Comput. Aided Design, 6 (1987), pp. 93–94.

ON THE POLYTOPE OF THE (1,2)-SURVIVABLE NETWORK DESIGN PROBLEM*

MOHAMED DIDI BIHA[†], HERVÉ KERIVIN[‡], AND A. RIDHA MAHJOUB[§]

Abstract. This paper deals with the survivable network design problem where each node v has a connectivity type $r(v)$ equal to 1 or 2, and the survivability conditions require the existence of at least $\min\{r(s), r(t)\}$ edge-disjoint paths for all distinct nodes s and t . We consider the polytope given by the trivial and cut inequalities together with the partition inequalities. More precisely, we study some structural properties of this polytope which leads us to give some sufficient conditions for this polytope to be integer in the class of series-parallel graphs. With both separation problems for the cut and partition inequalities being polynomially solvable, we then obtain a polynomial time algorithm for the (1,2)-survivable network design problem in a subclass of series-parallel graphs including the outerplanar graph class. We also introduce a new class of facet-defining inequalities for the polytope associated to the (1,2)-survivable network design problem.

Key words. survivable network, cut and partition inequalities, polytope, facet, series-parallel graphs

AMS subject classifications. 68M10, 90C10, 90C57

DOI. 10.1137/050639600

1. Introduction. In order to protect telecommunication networks from equipment failures, one must maintain the survivability of networks when links are severed or nodes fail. As failures are not very common, robust networks are designed to withstand a single network equipment failure. Moreover, in practice a node usually fails completely because of major incidents (e.g., power outages), and thus it is more frequent to encounter link interface failures or severed links than node failures. Therefore, one of the main concerns when designing telecommunication networks is to devise network topologies that provide protection against single-link failures. The network topology problem is usually the first stage of the overall network design optimization process, and the second one involves traffic and routing issues.

Furthermore, some network nodes may be more important than others because of their specific functions. This fact thus leads to considering two kinds of nodes: the *specific nodes*, also called *terminals*, for which a “high” degree of survivability has to be guaranteed, and the *ordinary nodes*, which simply have to be connected to the network. The network topology problem then consists of selecting links such that the sum of their cost is minimized and the failure of any single link does not disconnect any two terminal nodes.

*Received by the editors September 6, 2005; accepted for publication (in revised form) June 2, 2008; published electronically October 17, 2008.

<http://www.siam.org/journals/sidma/22-4/63960.html>

[†]Laboratoire d'Analyse Non Linéaire et Géométrie, Université d'Avignon, 339 chemin des Meina-jaries, 84911 Avignon, France (mohamed.didi-biha@univ-avignon.fr).

[‡]Institute for Mathematics and its Applications, University of Minnesota, 357 Lind Hall, 207 Church Street S.E., Minneapolis, MN 55455. Current address: Laboratoire LIMOS, CNRS UMR 6158, Université Blaise Pascal, Clermont-Ferrand II, Complexe Scientifique des Cézeaux, 63177 Aubière Cedex, France (kerivin@math.univ-bpclermont.fr).

[§]Laboratoire LIMOS, CNRS UMR 6158, Université Blaise Pascal, Clermont-Ferrand II, Complexe Scientifique des Cézeaux, 63177 Aubière Cedex, France. Current address: Laboratoire LAMSADE, CNRS UMR 7024, Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France (mahjoub@lamsade.dauphine.fr).

More precisely, based on a model first introduced by Grötschel and Monma [8] (see also Stoer [19]), this problem can be stated as follows. Consider an undirected graph $G = (V, E)$ where V represents the node set and E represents the set of edges or potential links. The set V is partitioned into two subsets T and O corresponding, respectively, to the terminal and ordinary node sets. By associating to each node $u \in V$ a *connectivity type* $r(u)$ which is equal to 1 if u is an ordinary node and to 2 if u is a terminal, we have $O = \{u \in V : r(u) = 1\}$, $T = \{u \in V : r(u) = 2\}$, and $V = O \cup T$. A subgraph H of G fulfills the survivability conditions if there exist at least $\min\{r(s), r(t)\}$ edge-disjoint paths (i.e., paths sharing no edges) in H for any pair of nodes $s, t \in V$. Such a subgraph is then called *survivable*. Suppose that each edge $e \in E$ has a certain cost $c(e) \in \mathbb{R}_+$. Our network topology problem, called the *(1,2)-survivable network design problem* (SNDP), then consists of finding a survivable subgraph of G with minimum total cost. (The cost of a subgraph of G is equal to the sum of its edge costs.)

The optimization problem SNDP is NP-hard since it includes as a special case the 2-edge connected network problem (i.e., $r(u) = 2$ for all $u \in V$), which has been extensively studied in the past. Some heuristics have thus been devised such as the one of Monma and Shallcross [18] which was used to obtain near optimal solutions to both real-world and randomly generated problems. The SNDP has also been proved to be polynomially solvable in special cases. Particularly, if $T = \emptyset$ (i.e., $r(u) = 1$ for all $u \in V$, and r is then called a *unit connectivity type vector*), then the SNDP is nothing but the minimum-cost spanning tree problem which is well known to be polynomially solvable [15]. Furthermore, if the underlying graph G is series-parallel and $O = \emptyset$ (i.e., $r(u) = 2$ for all $u \in V$), then we have a linear time algorithm for the SNDP devised by Winter [20]. Many survivability problems related to the SNDP have received particular attention and complete surveys over the existing approaches can be found in Grötschel, Monma, and Stoer [11], Stoer [19], and Kerivin and Mahjoub [13].

Grötschel, Monma, and Stoer [9] studied the general model related to the SNDP from a polyhedral point of view. (They consider $r(u) \in \mathbb{Z}_+$ for all $u \in V$.) They introduced several families of valid inequalities for the polytope associated with this problem. They also derived some necessary and/or sufficient conditions under which these inequalities are facet-defining. Among all of the inequalities considered in [9], the so-called partition inequalities are of interest for solving the SNDP as pointed out in [10, 14]. Kerivin and Mahjoub [12] actually showed that the separation problem for the partition inequalities is polynomially solvable for the SNDP, even though this separation problem is NP-hard for general connectivity type vectors $r \in \mathbb{Z}_+^V$. Furthermore, Grötschel and Monma [8] showed that the partition inequalities together with the trivial lower-bound and upper-bound inequalities suffice to describe the polytope associated with the SNDP when $r(u) = 1$ for all $u \in V$. When the underlying graph G is series-parallel, Mahjoub [17] described the polytope associated with the 2-edge connected network problem by the trivial inequalities and the cut ones, the latter being a special case of the partition inequalities.

Let $\text{SNDP}(G, r)$ be the convex hull of incidence vectors of all survivable subgraphs. This polytope is called the *survivable network polytope*. In this paper, we are interested in the polytope $\text{CPP}(G, r)$ given by the trivial lower-bound and upper-bound inequalities and the so-called partition inequalities for connectivity type vectors $r \in \{1, 2\}^V$. This polytope is a strengthened linear relaxation of $\text{SNDP}(G, r)$. Here, we give sufficient conditions for the $\text{CPP}(G, r)$ to be an integer on series-parallel graphs. This

study leads us to give a polynomial-time algorithm for solving the $(1, 2)$ -survivable network design problem on a class of graphs including the outerplanar ones.

This paper is organized as follows. In the next section, we formulate the problem SNDP as an integer linear program and introduce its associated polytope $\text{SNDP}(G, r)$ as well as the polytope $\text{CPP}(G, r)$. Section 3 is devoted to the study of some structural properties of the polytope $\text{CPP}(G, r)$. In section 4, we consider the $\text{CPP}(G, r)$ when G can be decomposed by one-node cutsets. We show in section 5 that the polytope $\text{CPP}(G, r)$ may have noninteger extreme points even if the underlying graph G is series-parallel, and give sufficient conditions which make the $\text{CPP}(G, r)$ integer on series-parallel graphs. In section 6, we prove this last result. A new class of facet-defining inequalities for the $\text{SNDP}(G, r)$ is introduced in section 7. Finally, some concluding remarks are given in section 8.

The rest of this section is devoted to more definitions and notation. Throughout this paper, the graphs are undirected, finite, loopless, and connected. We consider a graph $G = (V, E)$ and denote by n the number of nodes of G , that is, $n = |V|$. For $W \subseteq V$, let $\bar{W} = V \setminus W$, and for $F \subseteq E$, let $\bar{F} = E \setminus F$. Given two distinct nodes u and v of V , an edge between u and v is denoted by uv . For a nonempty node subset $W \subsetneq V$, the set of edges having exactly one endnode in W is called a *cut* and is denoted by $\delta_G(W)$. If $W = \{u\}$, we then write $\delta_G(u)$ for $\delta_G(\{u\})$. A *partition* of V is a collection of disjoint subsets of V with union V . The elements of the partition are called its *classes*. Given a partition $\{V_1, \dots, V_p\}$ of the node set V , we denote by $\delta_G(V_1, \dots, V_p)$ the set of edges with endnodes in two different classes. Given a collection $(W_1, \dots, W_q)$ of node subsets, we write $[W_1, \dots, W_q]_G$ for the set of edges with endnodes in two different subsets. We notice that if $(W_1, \dots, W_q)$ is a partition of V , then we have $[W_1, \dots, W_q]_G = \delta_G(W_1, \dots, W_q)$. If u and v are two distinct nodes of V , we then write $[u, v]_G$ for $\{u\}, \{v\}_G$. For all of our notation, we don't use the subscript G whenever the graph G can be deduced from the context. For $F \subseteq E$, we denote by $V(F)$ the set of nodes which are spanned by the edges in F . For $W \subseteq V$, we denote by $E(W)$ the set of edges with both endnodes in W , and $G(W) = (W, E(W))$ is called the *subgraph induced by W* . A maximal connected non-empty subgraph of G is called a *connected component*. (Here, "maximal" is taken with respect to inclusion.) A graph G is called *2-node-connected* if for any node $u \in V$, the subgraph $G - u$ induced by $V \setminus \{u\}$ is connected. Given a ground set S , a set-function $f : 2^S \rightarrow \mathbb{R} \cup \{\infty\}$ is called *fully submodular* if

$$(1.1) \quad f(A) + f(B) \geq f(A \cap B) + f(A \cup B)$$

for all $A, B \subseteq S$. A pair of subsets A and B of S is said to be *crossing* if none of $A \setminus B$, $B \setminus A$, $A \cap B$, $S \setminus (A \cup B)$ is empty. A set-function f is called *submodular on crossing pairs* if the inequality (1.1) is required only for crossing pairs. Moreover, if f satisfies the inequality (1.1) with equality for crossing pairs, then f is called *modular on crossing pairs*. For a vector $x \in \mathbb{R}^S$ and a subset $A \subseteq S$, we denote $\sum_{a \in A} x(a)$ by $x(A)$. For $F \subseteq E$, its *incidence vector* $x^F \in \mathbb{R}^S$ is defined by $x^F(e) = 1$ if $e \in F$, and $x^F(e) = 0$ if $e \in S \setminus F$. An *integer vector* is a vector with all entries integer. A polytope $P \subseteq \mathbb{R}^S$ is *integer* if and only if each extreme point of P is integer. An inequality $a^T x \geq \alpha$, where $a \in \mathbb{R}^S$ and $\alpha \in \mathbb{R}$, is *tight* for a point x^* if $a^T x^* = \alpha$. If $a^T x \geq \alpha$ is induced by a cut (respectively, a partition), we equivalently say that the cut (respectively, partition) is tight for x^* .

2. The $(1, 2)$ -survivable network design problem and some related polytopes. Let $G = (V, E)$ be an undirected graph and $r \in \{1, 2\}^V$ be a connectiv-

ity type vector. Without loss of generality, we may assume throughout this paper that there exist at least two nodes having the largest connectivity type (i.e., $|T| \geq 2$). For a nonempty node subset $W \subsetneq V$, let $r(W) = \max\{r(u) : u \in W\}$ and $\text{con}(W) = \min\{r(W), r(V \setminus W)\}$. Given an edge subset $F \subseteq E$, if (V, F) is a survivable subgraph of G , then its incidence vector x^F satisfies

$$\begin{aligned} (2.1) \quad & x(e) \geq 0 && \text{for all } e \in E, \\ (2.2) \quad & x(e) \leq 1 && \text{for all } e \in E, \\ (2.3) \quad & x(\delta(W)) \geq \text{con}(W) && \text{for all } \emptyset \neq W \subsetneq V, \\ (2.4) \quad & x(e) \in \{0, 1\} && \text{for all } e \in E. \end{aligned}$$

The inequalities (2.1) and (2.2) are, respectively, called *lower-bound* and *upper-bound trivial inequalities* (or more generally *trivial inequalities*), and inequalities (2.3) are called *cut inequalities*.

For a class of inequalities, the separation problem is as follows: given a vector y , find a violated inequality in the class or prove that none exists. An algorithm for the separation problem associated with a class of inequalities is a key ingredient for being able to use those inequalities within a branch-and-cut algorithm. The separation problem for the cut inequalities (2.3) is polynomially solvable using a polynomial-time maximum flow algorithm (e.g., preflow-push algorithm of Goldberg and Tarjan [6] running in $O(n^3)$ time).

In [9], Grötschel, Monma, and Stoer introduced a class of valid inequalities for the polytope $\text{SNDP}(G, r)$ which can be stated as follows. Let $\{V_1, \dots, V_p\}$, $p \geq 2$, be a partition of V . Let $I_2 = \{i : \text{con}(V_i) = 2, i = 1, \dots, p\}$ be the set of subscripts whose corresponding classes of the partition contain at least one terminal. The *partition inequalities* induced by $\{V_1, \dots, V_p\}$ is

$$(2.5) \quad x(\delta(V_1, \dots, V_p)) \geq \begin{cases} p-1 & \text{if } I_2 = \emptyset, \\ p & \text{otherwise.} \end{cases}$$

The inequalities (2.5) are a generalization of the cut inequalities (2.3). (The latter correspond to the case where $p = 2$.) Therefore, if we do not specify $p \geq 3$, then a partition inequality (2.5) may be a cut inequality (2.3). In the remainder of this paper, a partition of V with $I_2 = \emptyset$ (respectively, $I_2 \neq \emptyset$) will be called a *partition of type 1* (respectively, *partition of type 2*), and the inequality (2.5) induced by it will be called a *partition inequality of type 1* (respectively, a *partition inequality of type 2*). Grötschel, Monma, and Stoer [9] gave sufficient conditions and necessary conditions for the inequalities (2.5) to define facets of $\text{SNDP}(G, r)$.

In [12], Kerivin and Mahjoub showed that the separation problem for the partition inequalities (2.5) reduces to minimizing a particular submodular function, and then is polynomially solvable. Later, Barahona and Kerivin [2] reduced the separation problem for those inequalities to a sequence of $O(n^4)$ minimum cut problem.

Let $\text{CPP}(G, r)$ be the polytope given by the inequalities (2.1), (2.2), (2.3), and (2.5). This polytope, called the *cut and partition inequalities polytope*, is the linear relaxation of $\text{SNDP}(G, r)$ strengthened by the partition inequalities (2.5). As we mentioned above, the separation problems for both cut and partition inequalities are polynomially solvable, and then this implies by the ellipsoid method [7] that the (1,2)-survivable network design problem can be solved in polynomial time on graphs for which $\text{SNDP}(G, r) = \text{CPP}(G, r)$.

The definition of the connectivity type of node subsets implies the following remarks.

Remark 2.1. The function $r : 2^V \longrightarrow \{1, 2\}$ is nondecreasing, that is, r satisfies $r(A) \leq r(B)$ for all $A \subseteq B \subseteq V$.

Remark 2.2. Let $A \subseteq V$ such that $r(A) = 1$. Then, we have

(a) $r(A \cup B) = r(B)$ for any $B \subseteq V$, and

(b) $\text{con}(A \cup B) = \text{con}(B)$ for any $B \subseteq V$ such that $A \cup B \neq V$.

PROPOSITION 2.3. *The function $\text{con} : 2^V \longrightarrow \{1, 2\}$ is submodular on crossing pairs.*

Proof. Let $A, B \subseteq V$ such that $A \cap B \neq \emptyset$, $A \setminus B \neq \emptyset$, $B \setminus A \neq \emptyset$, and $\overline{A \cup B} \neq \emptyset$. We must show that

$$(2.6) \quad \text{con}(A) + \text{con}(B) \geq \text{con}(A \cap B) + \text{con}(A \cup B).$$

If $\text{con}(A) + \text{con}(B) = 4$, then it is obvious that inequality (2.6) holds.

If $\text{con}(A) = \text{con}(B) = 1$, then $\text{con}(A \cap B) = \text{con}(A \cup B) = 1$ and hence, inequality (2.6) is satisfied.

Suppose now that $\text{con}(A) + \text{con}(B) = 3$. Without loss of generality, we may assume that $\text{con}(A) = 1$ and $\text{con}(B) = 2$. Consider first the case $r(A) = 1$. From Remark 2.1, we then have $r(A \cap B) = r(A)$. By the definition of the function con , we thus obtain $\text{con}(A \cap B) = \text{con}(A) = 1$. From Remark 2.2(b), we have $\text{con}(A \cup B) = \text{con}(B)$. Thus, inequality (2.6) holds. If $r(A) = 2$, then $r(A \cap B) = r(A \setminus B) = 2$ and $r(B \setminus A) = r(\overline{A \cup B}) = 1$ and hence, inequality (2.6) is satisfied. $\square$

From the proof of Proposition 2.3, we deduce the remark below.

Remark 2.4. The function $\text{con} : 2^V \longrightarrow \{1, 2\}$ is modular on crossing pairs $A, B \subseteq V$ if one of the two following properties hold:

(a) $\text{con}(A) + \text{con}(B) \leq 3$, or

(b) $\text{con}(A) = \text{con}(B) = 2$ and $r(A \cap B) = r(\overline{A \cup B}) = 2$.

To conclude this section, we define the connectivity type vectors associated with the subgraphs of G obtained by contractions or deletions of edges. For any subset F of E , deleting F gives rise to the graph $G - F = (V(\overline{F}), \overline{F})$. The connectivity type vector $r_F^d \in \{1, 2\}^{V(\overline{F})}$ is then obviously the restriction of $r \in \{1, 2\}^V$ on $V(\overline{F})$, that is,

$$(2.7) \quad r_F^d(u) = r(u) \quad \text{for all } u \in V(\overline{F}).$$

Given an edge $e = uv \in E$, contracting e means deleting e and identifying u and v . If $F \subseteq E$ induces a connected subgraph of G , then G/F denotes the graph obtained from G by contracting F , that is, by contracting all of the edges in F . Let w be the node that arises from the contraction of F . The connectivity type vector $r_F^c \in \{1, 2\}^{V'}$ associated with the node set $V' = (V \setminus V(F)) \cup \{w\}$ is defined as follows:

$$(2.8) \quad r_F^c(u) = \begin{cases} r(u) & \text{if } u \in V' \setminus \{w\}, \\ \text{con}_G(V(F)) & \text{if } u = w. \end{cases}$$

A couple (H, r_H) is called a *minor* of (G, r) if H arises from G by a series of deletions and contractions of edges. The vector r_H is the connectivity type vector associated with H , and it is obtained from r by applying at each deletion/contraction the corresponding (2.7)/(2.8). This notion of minor will be important in section 5. In the next section, we study some structural properties of the $\text{CPP}(G, r)$ which will be also useful later on.

3. Structural properties of the polytope $\text{CPP}(G, r)$. One way to solve the (1,2)-survivable network design problem is to use a branch-and-cut framework. In such an approach, one may first consider a linear program whose constraints are given by $\text{CPP}(G, r)$ which is a relaxation of $\text{SNDP}(G, r)$. This linear program provides a lower bound for the SNDP , and in some special cases an optimal solution. This linear program has an exponential number of constraints, and one needs to use a cutting-plane algorithm to solve it. The knowledge of structural properties of the points (especially the extreme points) of $\text{CPP}(G, r)$ may provide useful information which is important in order to determine violated inequalities more efficiently. The structural properties established in this section may also be used to characterize the polytope $\text{SNDP}(G, r)$ on certain classes of graphs (see section 5).

Let us consider a point $\bar{x}$ of $\text{CPP}(G, r)$.

PROPOSITION 3.1. *Let $F \subseteq E$ be an edge subset of E that induces a connected subgraph of G . Then $\bar{x}' \in \mathbb{R}^{E \setminus F}$, the restriction of $\bar{x}$ on $E \setminus F$, is a point of $\text{CPP}(G/F, r_F^c)$.*

Proof. The result comes directly from the fact that any inequality of $\text{CPP}(G/F, r_F^c)$ is also an inequality of $\text{CPP}(G, r)$. $\square$

Let us denote by π a partition $\{V_1, \dots, V_p\}$ of V , $p \geq 2$, which is tight for $\bar{x}$.

PROPOSITION 3.2. *Suppose $p \geq 3$. Consider $i, j \in \{1, \dots, p\}$ such that $i < j$. Let $\pi' = \{V'_1, \dots, V'_{p-1}\}$ be the partition defined below*

$$\begin{aligned} V'_t &= V_t, & t &= 1, \dots, i-1, i+1, \dots, j-1, \\ V'_i &= V_i \cup V_j, \\ V'_t &= V_{t+1}, & t &= j, \dots, p-1. \end{aligned}$$

- (a) *If π is of type 1, then $\bar{x}[V_i, V_j] \leq 1$.*
- (b) *If π is of type 2 and π' is of type 1, then $\bar{x}[V_i, V_j] \leq 2$.*
- (c) *If π and π' are of type 2, then $\bar{x}[V_i, V_j] \leq 1$.*

Moreover, if $\bar{x}[V_i, V_j] = 1$ (cases (a) and (c)) or $\bar{x}[V_i, V_j] = 2$ (case (b)), then the partition π' is also tight for $\bar{x}$.

Proof. We are going to prove (a). (The proofs of (b) and (c) are similar.) Since π is of type 1, then clearly π' is also of type 1. Hence, we have

$$\bar{x}(\delta(V'_1, \dots, V'_{p-1})) \geq p-2.$$

We thus obtain

$$\begin{aligned} \bar{x}[V_i, V_j] &= \bar{x}(\delta(V_1, \dots, V_p)) - \bar{x}(\delta(V'_1, \dots, V'_{p-1})) \\ &\leq p-1 - (p-2) \\ &= 1. \end{aligned}$$

Moreover, if $\bar{x}[V_i, V_j] = 1$, then the above inequalities are all satisfied as equalities, and we get $\bar{x}(\delta(V'_1, \dots, V'_{p-1})) = p-2$. $\square$

PROPOSITION 3.3. *Consider a partition $\{V_i^1, V_i^2\}$ of V_i for some $i \in \{1, \dots, p\}$. Let $\pi' = \{V'_1, \dots, V'_{p+1}\}$ be the partition of V given by*

$$\begin{aligned} V'_j &= V_j, & j &= 1, \dots, i-1, \\ V'_i &= V_i^1, \\ V'_{i+1} &= V_i^2, \\ V'_j &= V_{j-1}, & j &= i+2, \dots, p+1. \end{aligned}$$

- (a) If π and π' are of type 1, then $\bar{x}[V_i^1, V_i^2] \geq 1$.
- (b) If π is of type 1 and π' is of type 2, then $\bar{x}[V_i^1, V_i^2] \geq 2$.
- (c) If π is of type 2, then $\bar{x}[V_i^1, V_i^2] \geq 1$.

Moreover, if $\bar{x}[V_i^1, V_i^2] = 1$ (cases (a) and (c)) or $\bar{x}[V_i^1, V_i^2] = 2$ (case (b)), then the partition π' is also tight for $\bar{x}$.

Proof. The proof is omitted because of its similarity with the one of Proposition 3.2. $\square$

An immediate consequence of Proposition 3.3 is the following.

Remark 3.4. The subgraphs $G(V_i)$ for $i = 1, \dots, p$ are all connected.

We notice that the two previous propositions could be generalized as described in the remark below.

Remark 3.5. In Proposition 3.2, we may consider a subscript subset $I \subsetneq \{1, \dots, p\}$ with $|I| \geq 2$ (instead of only two distinct subscripts i and j), and in Proposition 3.3, we may partition the node subset V_i into at least (rather than exactly) two subsets.

PROPOSITION 3.6. *Let $\delta(W_1)$ and $\delta(W_2)$ be two cuts tight for $\bar{x}$ such that W_1 and W_2 are two crossing subsets of V . Then, $\delta(W_1 \cap W_2)$ and $\delta(\overline{W_1 \cup W_2})$ are tight for $\bar{x}$, and $\bar{x}[W_1 \setminus W_2, W_2 \setminus W_1] = 0$ if one of the two following properties holds:*

- (a) $\text{con}(W_1) + \text{con}(W_2) \leq 3$, or
- (b) $\text{con}(W_1) = \text{con}(W_2) = 2$ and $r(W_1 \cap W_2) = r(\overline{W_1 \cup W_2}) = 2$.

Proof. We have

$$\begin{aligned} \bar{x}(\delta(W_1)) + \bar{x}(\delta(W_2)) &= \text{con}(W_1) + \text{con}(W_2) \\ &= \bar{x}(\delta(W_1 \cap W_2)) + \bar{x}(\delta(\overline{W_1 \cup W_2})) + 2\bar{x}[W_1 \setminus W_2, W_2 \setminus W_1] \\ &\geq \text{con}(W_1 \cap W_2) + \text{con}(\overline{W_1 \cup W_2}), \end{aligned}$$

where the last inequality follows from $\bar{x} \in \text{CPP}(G, r)$. Since one of the properties (a) and (b) holds and $\text{con}(\overline{W_1 \cup W_2}) = \text{con}(W_1 \cup W_2)$, by Remark 2.4 we have $\text{con}(W_1) + \text{con}(W_2) = \text{con}(W_1 \cap W_2) + \text{con}(\overline{W_1 \cup W_2})$. We then obtain the result. $\square$

PROPOSITION 3.7. *Let $\pi_1 = \{V_1, \dots, V_p\}$ and $\pi_2 = \{W_1, \dots, W_q\}$, $p \geq 2$ and $q \geq 2$, be two partitions of V which are tight for $\bar{x}$. Consider two distinct subscripts $i, j \in \{1, \dots, p\}$.*

- (a) *If π_1 is of type 1 or π_1 is of type 2 with $p \geq 3$ and $r(\overline{V_i \cup V_j}) = 2$, then at most one class of π_2 only intersects both V_i and V_j .*
- (b) *If π_1 is of type 2 with either $p = 2$ or $r(\overline{V_i \cup V_j}) = 1$, then at most two classes of π_2 only intersect both V_i and V_j .*

Proof. Assume that π_1 is a partition of type 1. (The proofs for the other cases are similar.) Let $K \subseteq \{1, \dots, q\}$, $|K| > 1$, be the set of subscripts such that $W_k \cap V_i \neq \emptyset$, $W_k \cap V_j \neq \emptyset$, and $W_k = (W_k \cap V_i) \cup (W_k \cap V_j)$ for all $k \in K$. From Proposition 3.3, we have

$$\bar{x}[W_k \cap V_i, W_k \cap V_j] \geq 1 \quad \text{for all } k \in K.$$

These last inequalities together with Proposition 3.2(a) lead to

$$\begin{aligned} 1 &\geq \bar{x}[V_i, V_j] \geq \sum_{k \in K} \bar{x}[W_k \cap V_i, W_k \cap V_j] \\ &\geq |K| \\ &> 1, \end{aligned}$$

a contradiction. $\square$

We remark that if we have $p = 2$ in Proposition 3.7, then at most $\text{con}(V_1)$ classes of the partition π_2 intersect both V_1 and V_2 .

PROPOSITION 3.8. *Let $\pi_1 = \{V_1, \dots, V_p\}$ and $\pi_2 = \{W_1, \dots, W_q\}$, $p \geq 2$ and $q \geq 3$, be two partitions of V which are tight for $\bar{x}$. Given a subscript $i \in \{1, \dots, p\}$, define $J = \{j_1, \dots, j_k\} = \{j \in \{1, \dots, q\} \mid V_i \cap W_j \neq \emptyset\}$. Assume $2 \leq k < q$. We then have*

- (a) $\bar{x}[V_i \cap W_{j_1}, \dots, V_i \cap W_{j_k}] = \bar{x}[W_{j_1}, \dots, W_{j_k}] = k - 1$, if π_1 is of type 1, and one of the following conditions hold:
 - (a.1) π_2 is of type 1, or
 - (a.2) π_2 is of type 2, $r(V_1) = 1$ and $r(V \setminus \bigcup_{j \in J} W_j) = 2$.
- (b) $\bar{x}[V_i \cap W_{j_1}, \dots, V_i \cap W_{j_k}] = \bar{x}[W_{j_1}, \dots, W_{j_k}] = k$, if π_1 and π_2 are of type 2 and $r(V_i) = 2$.

Proof. We are going to prove the case (a.1). (The proofs for the other cases use similar arguments.) Without loss of generality, we suppose $J = \{1, \dots, k\}$. Since $[V_i \cap W_1, \dots, V_i \cap W_k] \subseteq [W_1, \dots, W_k]$ and $\bar{x} \geq 0$, we have

$$(3.1) \quad \bar{x}[V_i \cap W_1, \dots, V_i \cap W_k] \leq \bar{x}[W_1, \dots, W_k].$$

From the definition of J and the fact that π_2 is a partition of V , $\{V_i \cap W_1, \dots, V_i \cap W_k\}$ is a partition of V_i . By Proposition 3.3(a) and Remark 3.5, we have

$$(3.2) \quad \bar{x}[V_i \cap W_1, \dots, V_i \cap W_k] \geq k - 1.$$

Furthermore, as π_2 is of type 1, the partition $\{\bigcup_{j \in J} W_j, W_{k+1}, \dots, W_q\}$ is also of type 1. From Proposition 3.2(a) and Remark 3.5, we have

$$\bar{x}[W_1, \dots, W_k] \leq k - 1.$$

This last inequality combined with the inequalities (3.1) and (3.2) gives $\bar{x}[V_i \cap W_1, \dots, V_i \cap W_k] = \bar{x}[W_1, \dots, W_k] = k - 1$. $\square$

We remark that in Proposition 3.8, if we do not fulfill the conditions (a) and (b), we still have $k - 1 \leq \bar{x}[V_i \cap W_{j_1}, \dots, V_i \cap W_{j_k}] \leq \bar{x}[W_{j_1}, \dots, W_{j_k}] \leq k$.

PROPOSITION 3.9. *Let $\pi = \{V_1, \dots, V_p\}$, $p \geq 2$, be a partition of V which is tight for $\bar{x}$. Consider two distinct subscripts $i, j \in \{1, \dots, p\}$ such that $i < j$, $|V_i| \geq 2$, and $|V_j| \geq 2$. Given a nonempty node set $W \subsetneq V_i$ and a partition $\{V_j^1, \dots, V_j^q\}$, $q \geq 2$, of V_j , let $\pi' = \{V'_1, \dots, V'_{p'}\}$ be the following partition of V :*

$$\begin{aligned} V'_t &= V_t, & t &= 1, \dots, i-1, \\ V'_i &= V_i \setminus W, \\ V'_t &= V_t, & t &= i+1, \dots, j-1, \\ V'_j &= V_j^1 \cup W, \\ V'_{t+j} &= V_j^{t+1}, & t &= 1, \dots, q-1, \\ V'_{t+q-1} &= V_t, & t &= j+1, \dots, p, \end{aligned}$$

where $p' = p + q - 1$.

- (a) *If π and π' are of the same type, then $\bar{x}[V_j^1, \dots, V_j^q] \geq q - 1 + \bar{x}[W, V_j^1] - \bar{x}[W, V_i \setminus W]$.*
- (b) *If π is of type 1 and π' is of type 2, then $\bar{x}[V_j^1, \dots, V_j^q] \geq q + \bar{x}[W, V_j^1] - \bar{x}[W, V_i \setminus W]$.*

(c) If π is of type 2 and π' is of type 1, then $\bar{x}[V_j^1, \dots, V_j^q] \geq q - 2 + \bar{x}[W, V_j^1] - \bar{x}[W, V_i \setminus W]$.

Moreover, if $\bar{x}[V_j^1, \dots, V_j^q] = q - 1 + \bar{x}[W, V_j^1] - \bar{x}[W, V_i \setminus W]$ (case (a)), $\bar{x}[V_j^1, \dots, V_j^q] = q + \bar{x}[W, V_j^1] - \bar{x}[W, V_i \setminus W]$ (case (b)) or $\bar{x}[V_j^1, \dots, V_j^q] = q - 2 + \bar{x}[W, V_j^1] - \bar{x}[W, V_i \setminus W]$ (case (c)), then the partition π' is also tight for $\bar{x}$.

Proof. Using $\bar{x}(\delta(V'_1, \dots, V'_{p'})) = \bar{x}(\delta(V_1, \dots, V_p)) + \bar{x}[V_j^1, \dots, V_j^q] - \bar{x}[W, V_j^1] + \bar{x}[W, V_i \setminus W]$, the proof is similar to the one of Proposition 3.2, and thus it is omitted. $\square$

In the remainder of this section, let us assume that $\bar{x}$ is an extreme point of $\text{CPP}(G, r)$. We denote by $E_0(\bar{x})$, $E_1(\bar{x})$ and $E_f(\bar{x})$ the set of edges $e \in E$ such that $\bar{x}(e) = 0$, $\bar{x}(e) = 1$ and $0 < \bar{x}(e) < 1$, respectively. Let $P_1(\bar{x})$ and $P_2(\bar{x})$ be the sets of partitions of type 1 and 2, respectively, which are tight for $\bar{x}$.

Since $\bar{x}$ is an extreme point of $\text{CPP}(G, r)$, there exist $P_1^*(\bar{x}) \subseteq P_1(\bar{x})$ and $P_2^*(\bar{x}) \subseteq P_2(\bar{x})$ such that $\bar{x}$ is the unique solution of the system

$$S(\bar{x}) \begin{cases} x(e) = 0 & \text{for all } e \in E_0(\bar{x}), \\ x(e) = 1 & \text{for all } e \in E_1(\bar{x}), \\ x(\delta(V_1, \dots, V_p)) = p - 1 & \text{for all } \{V_1, \dots, V_p\} \in P_1^*(\bar{x}), \\ x(\delta(W_1, \dots, W_q)) = q & \text{for all } \{W_1, \dots, W_q\} \in P_2^*(\bar{x}), \end{cases}$$

where $|E_0(\bar{x})| + |E_1(\bar{x})| + |P_1^*(\bar{x})| + |P_2^*(\bar{x})| = |E|$. Since the system $S(\bar{x})$ is not unique, we give the following remarks which will be useful later.

Remark 3.10. From any system of equations, induced by inequalities of $\text{CPP}(G, r)$ and whose unique solution is $\bar{x}$, we may extract a nonsingular subsystem having exactly $|E|$ equations.

Remark 3.11. For any $e \in E$, there exists at least one equation of $S(\bar{x})$ which contains $x(e)$ with a nonzero coefficient.

PROPOSITION 3.12. *If $\bar{x}$ is fractional, then $|E_f(\bar{x})| \geq 2$.*

Proof. Every equation of the system $S(\bar{x})$ has coefficients equal to 0 or 1 and an integer right-hand side. Since $\bar{x}$ is fractional, Remark 3.11 then implies that $\bar{x}$ must contain at least two fractional components. $\square$

PROPOSITION 3.13. *For any edge pair of $E_f(\bar{x})$, there exists at least one equation of $S(\bar{x})$ which contains exactly one of the two edges.*

Proof. Suppose that there exist two edges $e_1, e_2 \in E_f(\bar{x})$ such that any equation of $S(\bar{x})$ contains either both or none of them. Let $x' \in \mathbb{R}^E$ be the point such that

$$x'(e) = \begin{cases} \bar{x}(e) + \epsilon & \text{if } e = e_1, \\ \bar{x}(e) - \epsilon & \text{if } e = e_2, \\ \bar{x}(e) & \text{if } e \in E \setminus \{e_1, e_2\}, \end{cases}$$

where $\epsilon \neq 0$. The point x' is also a solution of $S(\bar{x})$, which contradicts the extremality of $\bar{x}$. $\square$

A direct consequence of the previous proposition is the following.

Remark 3.14. If u, v are two nodes of V , then $[u, v]$ contains at most one edge in $E_f(\bar{x})$.

PROPOSITION 3.15. *Let $W \subsetneq V$ be a nonempty node subset such that $\bar{x}(\delta(W)) = 1$. We then have $\bar{x}(e) \in \{0, 1\}$ for all $e \in \delta(W)$.*

Proof. Suppose there exists $e \in \delta(W) \cap E_f(\bar{x})$. Since $\bar{x}(\delta(W)) = 1$, there must exist another edge $f \in \delta(W) \cap E_f(\bar{x})$. We are going to prove that the system $S(\bar{x})$ can

be chosen such that any of its equations contains either e and f or none of them. Let $\{V_1, \dots, V_p\}$ be a partition of V inducing an equation of the system $S(\bar{x})$. Assume that there exist two distinct subscripts $i, j \in \{1, \dots, p\}$ such that $V_i \cap W \neq \emptyset \neq V_j \cap W$. From Proposition 3.3, we obtain $\bar{x}[V_i \cap W, V_i \setminus W] \geq 1$ and $\bar{x}[V_j \cap W, V_j \setminus W] \geq 1$ which give $\bar{x}(\delta(W)) \geq \bar{x}[V_i \cap W, V_i \setminus W] + \bar{x}[V_j \cap W, V_j \setminus W] \geq 2$, a contradiction. Therefore, we have $W \subseteq V_k$ for some $k \in \{1, \dots, p\}$. If $W \neq V_k$, then from Proposition 3.3 and $\bar{x}(\delta(W)) = 1$ we must have $\delta(W) = [W, V_k \setminus W]$. We thus conclude that any equation of $S(\bar{x})$ contains either e and f or none of them. But this contradicts Proposition 3.13. $\square$

PROPOSITION 3.16. *Let $W \subsetneq V$ be a nonempty node subset such that $\text{con}(W) = 2$. If $\bar{x}(\delta(W)) = 2$, then the system $S(\bar{x})$ can be chosen such that $P_1^*(\bar{x}) = \emptyset$.*

Proof. Suppose there exists a partition $\{V_1, \dots, V_p\} \in P_1^*(\bar{x})$. Without loss of generality, we suppose that V_1 contains all the terminals. Since $\text{con}(W) = 2$, we obtain that $V_1 \cap W \neq \emptyset \neq V_1 \cap \bar{W}$. From Proposition 3.3(b), we have $\bar{x}[V_1 \cap W, V_1 \cap \bar{W}] \geq 2$ which implies $\delta(W) = [V_1 \cap W, V_1 \cap \bar{W}]$. Therefore by Proposition 3.2, we deduce that either W or $\bar{W}$ is a subset of V_1 . Without loss of generality, we assume that $W \subset V_1$. The partition $\{W, V_1 \cap \bar{W}, V_2, \dots, V_p\}$ is of type 2, then since $\bar{x}[W, V_1 \cap \bar{W}] = 2$ Proposition 3.3 gives $\{W, V_1 \cap \bar{W}, V_2, \dots, V_p\} \in P_2^*(\bar{x})$.

We can thus replace in the system $S(\bar{x})$ the partition $\{V_1, \dots, V_p\}$ by the partitions $\{W, \bar{W}\}$ and $\{W, V_1 \cap \bar{W}, V_2, \dots, V_p\}$. The obtained system may have more than $|E|$ equations, yet from Remark 3.10 we can choose $|E|$ equations of this system whose unique solution is $\bar{x}$. $\square$

PROPOSITION 3.17. *Let $W \subsetneq V$ be a nonempty node subset such that $\text{con}(W) = 2$. If $\bar{x}(\delta(W)) = 2$, then the system $S(\bar{x})$ can be chosen such that at most one class of any partition of $P_2^*(\bar{x})$ intersects both W and $\bar{W}$.*

Proof. From Proposition 3.7(b), any partition of $P_2^*(\bar{x})$ has at most two classes which intersect both W and $\bar{W}$. Suppose that there exists a partition $\{V_1, \dots, V_p\}$ of $P_2^*(\bar{x})$ such that, without loss of generality, V_1 and V_2 intersect both W and $\bar{W}$. Proposition 3.3(c) then states that the partition $\{V_1 \cap W, V_1 \cap \bar{W}, V_2 \cap W, V_2 \cap \bar{W}, V_3, \dots, V_p\}$ belongs to $P_2(\bar{x})$. We can thus replace in the system $S(\bar{x})$ the partition $\{V_1, \dots, V_p\}$ by the partitions $\{W, \bar{W}\}$ and $\{V_1 \cap W, V_1 \cap \bar{W}, V_2 \cap W, V_2 \cap \bar{W}, V_3, \dots, V_p\}$. The result thus follows from Remark 3.10. $\square$

PROPOSITION 3.18. *Let u, v be two distinct nodes such that $\bar{x}[u, v] \geq 1$. The system $S(\bar{x})$ can thus be chosen such that*

- (a) $\delta(V_1, \dots, V_p) \cap [u, v] = \emptyset$, for all $\{V_1, \dots, V_p\} \in P_1^*(\bar{x})$, and
- (b) if $\{V_1, \dots, V_p\} \in P_2^*(\bar{x})$ with $p \geq 3$ and $[u, v] \subseteq [V_i, V_j]$ for some distinct subscripts $i, j \in \{1, \dots, p\}$, then $r(V \setminus (V_i \cup V_j)) = 1$.

Proof. We will prove (a). (The proof of (b) uses similar arguments.) Let $\pi = \{V_1, \dots, V_p\}$, $p \geq 2$, be a partition of type 1, tight for $\bar{x}$ and such that without loss of generality, $[u, v] \subseteq [V_1, V_2]$. If $p = 2$, then from Proposition 3.15, we have $\bar{x}(e) \in \{0, 1\}$ for all $e \in [u, v]$. The equation induced by π can thus be obtained from the equations $x(e) = 0$ for all $e \in [u, v] \cap E_0(\bar{x})$, and $x(e) = 1$ for all $e \in [u, v] \cap E_1(\bar{x})$.

Suppose now that $p \geq 3$. From Proposition 3.2, we have $\bar{x}[V_1, V_2] \leq 1$. Since $[u, v] \subseteq [V_1, V_2]$ and $\bar{x}[u, v] \geq 1$, we then obtain $\bar{x}[u, v] = 1$. Therefore, the partition $\pi' = \{V_1 \cup V_2, V_3, \dots, V_p\}$ is also tight for $\bar{x}$. Moreover, by Remark 3.14, $\bar{x}(e) \in \{0, 1\}$ for all $e \in [u, v]$. The equation induced by π can thus be obtained from the one induced by π' together with the equations $x(e) = 0$ for all $e \in [u, v] \cap E_0(\bar{x})$ and $x(e) = 1$ for all $e \in [u, v] \cap E_1(\bar{x})$. $\square$

PROPOSITION 3.19. *Let u be a node having exactly two neighbors, namely u_1*

and u_2 . The system $S(\bar{x})$ can then be chosen such that any partition $\{V_1, \dots, V_p\} \in P_1^*(\bar{x}) \cup P_2^*(\bar{x})$ with $p \geq 3$ and $[u, u_1] \subset \delta(V_1, \dots, V_p)$ has the following property:

$$|V_i \cap \{u_1, u_2\}| \leq 1 \text{ for all } i \in \{1, \dots, p\}.$$

Moreover, if $\bar{x}[u, u_2] < 1$, then we also have

$$|V_i \cap \{u, u_2\}| \leq 1 \text{ for all } i \in \{1, \dots, p\}.$$

Proof. Let $\pi = \{V_1, \dots, V_p\} \in P_1^*(\bar{x}) \cup P_2^*(\bar{x})$ with $p \geq 3$ and $[u, u_1] \subset \delta(V_1, \dots, V_p)$. Without loss of generality, we assume $u \in V_1$ and $u_1 \in V_2$. Suppose $u_2 \in V_2$. By Remark 3.4, we have $V_1 = \{u\}$. If $\bar{x}(\delta(u)) > r(u)$, by considering the partition $\{V_1 \cup V_2, V_3, \dots, V_p\}$, we then contradict Proposition 3.2. Hence, we have $\bar{x}(\delta(u)) = r(u)$. Let π' be the partition $\{V_1 \cup V_2, V_3, \dots, V_p\}$. If $r(u) = 2$, then either $\bar{x}[u, u_1]$ or $\bar{x}[u, u_2]$ is greater than 1, and by Proposition 3.18, we have $\pi \in P_2^*(\bar{x})$, $r(V \setminus (V_1 \cup V_2)) = 1$, and then $\pi' \in P_1(\bar{x})$. If $r(u) = 1$, then we clearly have $\pi' \in P_1(\bar{x}) \cup P_2(\bar{x})$. We can thus replace in the system $S(\bar{x})$ the partition π by the partition π' and the cut $\delta(u)$. The obtained system may have more than $|E|$ equations, yet from Remark 3.10 we can choose $|E|$ equations of this system whose unique solution is $\bar{x}$. Therefore, the system $S(\bar{x})$ can be chosen such that no partition has one class containing u and a different one containing both u_1 and u_2 .

We now suppose that $\bar{x}[u, u_2] < 1$. If $u_2 \in V_1$, we then have $\bar{x}[u, V_1 \setminus \{u\}] = \bar{x}[u, u_2] < 1$, a contradiction with Proposition 3.3. $\square$

PROPOSITION 3.20. Let u, v be two distinct nodes such that $\bar{x}[u, v] \geq 2$. We have the following:

- (a) the system $S(\bar{x})$ can be chosen such that the variables $x(e)$, for all $e \in [u, v]$, only appear in equations induced by $E_0(\bar{x}) \cup E_1(\bar{x})$ with a nonzero coefficient, and
- (b) $\bar{x}(e) \in \{0, 1\}$ for all $e \in [u, v]$.

Proof. Using the same arguments as in the proof of Proposition 3.18, we get (a). Consequently, from Remark 3.11, we obtain $\bar{x}(e) \in \{0, 1\}$ for all $e \in [u, v]$. $\square$

4. Composition of G by one-node cutset and the $\text{CPP}(G, r)$. Given a graph $G = (V, E)$ and two subgraphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ of G , if $V = V_1 \cup V_2$ and $|V_1 \cap V_2| = 1$, then $G = (V, E_1 \cup E_2)$ is called the 1-sum of G_1 and G_2 . In that case, the singleton $V_1 \cap V_2$ is called a *one-node cutset* of G .

LEMMA 4.1. Let $G = (V, E)$ be a graph and $r \in \{1, 2\}^V$ be a connectivity type vector. Suppose that G is the 1-sum of the graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$. Let u be the only node of $V_1 \cap V_2$. For $i = 1, 2$, let $r_i \in \{1, 2\}^{V_i}$ be the connectivity type vector such that $r_i(v) = r(v)$ if $v \in V_i \setminus \{u\}$ and $r_i(u) = \max\{r(u), r(V \setminus V_i)\}$. If $\bar{x}$ is an extreme point of $\text{CPP}(G, r)$, then the restriction $\bar{x}_i$ of $\bar{x}$ on G_i is also an extreme point of $\text{CPP}(G_i, r_i)$ for $i = 1, 2$.

Proof. First, when we write $\bar{x}_i$, we refer to one of the two restrictions $\bar{x}_1$ and $\bar{x}_2$, and then the subscript i may be either 1 or 2 in the proof. From Proposition 3.1, $\bar{x}_i \in \text{CPP}(G_i, r_i)$. (We remark that the couples (G_1, r_1) and (G_2, r_2) are nothing but the couples $(G/E_2, r_{E_2}^c)$ and $(G/E_1, r_{E_1}^c)$, respectively.) To prove that $\bar{x}_i$ is an extreme point of $\text{CPP}(G_i, r_i)$, it suffices to show that we can choose the system $S(\bar{x})$ such that for any pair of edges $e_1 \in E_1$ and $e_2 \in E_2$, none of its equations involves both $x(e_1)$ and $x(e_2)$ with nonzero coefficients.

Let $\pi = \{W_1, \dots, W_p\}$ be a partition which is tight for $\bar{x}$. Suppose $u \in W_q$ for some $q \in \{1, \dots, p\}$. Since $\{u\}$ is a one-node cutset, by Remark 3.4 W_q is the only

class which may intersect both V_1 and V_2 . Without loss of generality, we can assume that $W_j \subseteq V_1$ for all $j \in \{1, \dots, q-1\}$ and $W_k \subseteq V_2$ for all $k \in \{q+1, \dots, p\}$.

Consider the two partitions $\pi_1 = \{W_1, \dots, W_{q-1}, W_q^1\}$ and $\pi_2 = \{W_q^2, W_{q+1}, \dots, W_p\}$ of V , where

$$W_q^1 = \bigcup_{j \geq q} W_j \quad \text{and} \quad W_q^2 = \bigcup_{k \leq q} W_k.$$

We clearly have

$$(4.1) \quad \bar{x}(\delta(W_1, \dots, W_p)) = \bar{x}(\delta(W_1, \dots, W_{q-1}, W_q^1)) + \bar{x}(\delta(W_q^2, W_{q+1}, \dots, W_p)).$$

We remark that we cannot have both π_1 and π_2 of type 2. The partition π would be of type 2 otherwise, and then (4.1) would give

$$\begin{aligned} p &= \bar{x}(\delta(W_1, \dots, W_p)) \\ &\geq q + (p - q + 1) \\ &= p + 1, \end{aligned}$$

which is impossible.

Without loss of generality, we now suppose that π_1 is of type 1. We notice that, in this case, π and π_2 have the same type. Assume that π_2 is of type 2. (The case where π_2 is of type 1 is similar.) The inequality (4.1) then gives

$$\begin{aligned} p &= \bar{x}(\delta(W_1, \dots, W_p)) \\ &\geq (q-1) + (p - q + 1) \\ &= p. \end{aligned}$$

We thus have that both partitions π_1 and π_2 are tight for $\bar{x}$. Hence, we can replace in the system $S(\bar{x})$ the partition π by the two partitions π_1 and π_2 , and we then get a new system $S'(\bar{x})$. By Remark 3.10, there is a nonsingular subsystem of $S'(\bar{x})$ with exactly $|E|$ equations and whose solution is $\bar{x}$. Since $\delta(W_1, \dots, W_{q-1}, W_q^1) \subseteq E_1$ and $\delta(W_q^2, W_{q+1}, \dots, W_p) \subseteq E_2$, any equation of this subsystem only contains variables induced by edges from either E_1 or E_2 . The proof is thus complete. $\square$

Using the same notations as those introduced in Lemma 4.1, the following is an immediate consequence of this lemma.

COROLLARY 4.2. *If $CPP(G_1, r_1)$ and $CPP(G_2, r_2)$ are integer, then so is $CPP(G, r)$.*

Suppose that G is decomposable by one-node cutsets into $G_1, G_2, \dots, G_t$ with $t \geq 2$. Let $r_i, i = 1, \dots, t$, be the connectivity type vector associated to G_i , defined as in Lemma 4.1. A direct consequence of Corollary 4.2 is that if $CPP(G_i, r_i)$ is integer for all $i \in \{1, \dots, t\}$, then so is $CPP(G, r)$.

5. The polytope $CPP(G, r)$ on series-parallel graphs. In this section, we are interested in the cut and partition inequalities polytope $CPP(G, r)$ on series-parallel graphs. A graph is called *series-parallel* if and only if it does not contain K_4 (i.e., the complete graph with 4 nodes) as a minor [4]. We are going to give sufficient conditions for this polytope to be an integer on this class of graphs, that is, sufficient conditions for the polytope $SNDP(G, r)$ to be completely described by the trivial inequalities (2.1) and (2.2) together with the partition inequalities (2.5). (We recall that the cut inequalities (2.3) are partition inequalities induced by partitions having exactly two classes.)

We can remark that if the graph G contains at most one terminal (i.e., $|T| \leq 1$), then the SNDP is nothing but the spanning tree problem. In this case and for general graphs G , Fulkerson [5] completely described the polytope $\text{SNDP}(G, r)$ by the trivial inequalities and the inequalities (2.5) induced by partitions of type 1. For general graphs G , we then have the following.

THEOREM 5.1 (see [5]). *If $|T| \leq 1$, then $\text{CPP}(G, r)$ is integer.*

On the other hand, Mahjoub [17] showed that the polytope $\text{SNDP}(G, r)$ is completely described by the trivial inequalities (2.1) and (2.2), and the cut inequalities (2.3) on series-parallel graphs when $r(u) = 2$ for all $u \in V$. We notice that for such connectivity type vectors, the partition inequalities (2.5) are dominated by the cut inequalities (2.3). Therefore we have the following theorem.

THEOREM 5.2 (see [17]). *If G is series-parallel and $T = V$, then $\text{CPP}(G, r)$ is integer.*

In the remainder of this paper, we are going to restrict our attention to graphs G having at least one ordinary node and two terminals. We remark that according to the definition of the survivability conditions, the case with exactly one terminal can be trivially reduced to the one with $T = \emptyset$.

In view of Theorems 5.1 and 5.2, one would have expected the integrality of the polytope $\text{CPP}(G, r)$ if G is series-parallel and $r \in \{1, 2\}^V$. Unfortunately, it turns out that this result does not hold. In fact, let us consider the two graphs $G_p^1 = (V_p^1, E_p^1)$ and $G_p^2 = (V_p^2, E_p^2)$ of Figure 5.1, where the terminals are represented by black circles. (G_p^1 has three terminals $\{u_1, u_2, u_3\}$ and two ordinary nodes $\{v_1, v_2\}$, while G_p^2 has four terminals $\{u_1, u_2, u_3, u_4\}$ and one ordinary node $\{v_1\}$.) These two graphs are clearly series-parallel. Moreover, the fractional solutions of $\frac{1}{2}$ for all of the dashed edges and 1 for all of the solid ones are extreme points of the corresponding polytopes $\text{CPP}(G_p^1, r_p^1)$ and $\text{CPP}(G_p^2, r_p^2)$, where $r_p^1 \in \{1, 2\}^{V_p^1}$ and $r_p^2 \in \{1, 2\}^{V_p^2}$ are the connectivity type vectors associated to G_p^1 and G_p^2 , respectively. This implies that the partition inequalities (2.5) together with the trivial ones (2.1) and (2.2) do not suffice to completely describe the survivable network design polytope $\text{SNDP}(G, r)$ on series-parallel graphs.

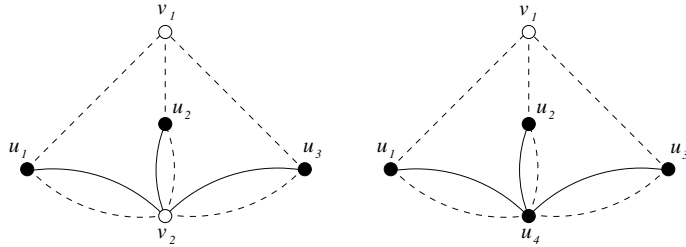


FIG. 5.1. Fractional extreme points of $\text{CPP}(G_p^1, r_p^1)$ and $\text{CPP}(G_p^2, r_p^2)$.

Let us now consider the following inequalities:

$$(5.1) \quad x(E_p^i) + x(\delta(v_1)) \geq 8 \quad \text{for } i = 1, 2.$$

It is not hard to see that the inequality (5.1) for $i = 1$ (respectively, $i = 2$) should be satisfied by any point of the polytope $\text{SNDP}(G_p^1, r_p^1)$ (respectively, $\text{SNDP}(G_p^2, r_p^2)$). Moreover, this inequality cuts off the fractional extreme point of $\text{CPP}(G_p^1, r_p^1)$ (respectively, $\text{CPP}(G_p^2, r_p^2)$) given above. We will see in section 7 that inequalities (5.1) are actually special cases of a more general class of facet-defining inequalities of $\text{SNDP}(G, r)$.

In the next theorem, we give sufficient conditions based on both graphs (G_p^1, r_p^1) and (G_p^2, r_p^2) for the cut and partition inequalities polytope $\text{CPP}(G, r)$ to be an integer on series-parallel graphs.

THEOREM 5.3. *Let $G = (V, E)$ be a series-parallel graph and $r \in \{1, 2\}^V$ be its associated connectivity type vector. If (G, r) does not have either (G_p^1, r_p^1) or (G_p^2, r_p^2) as a minor, then $\text{CPP}(G, r)$ is integer.*

Before giving the proof of the theorem, let us first note that its converse does not hold as shown by the following example. (The black circles still depict the terminal nodes.)

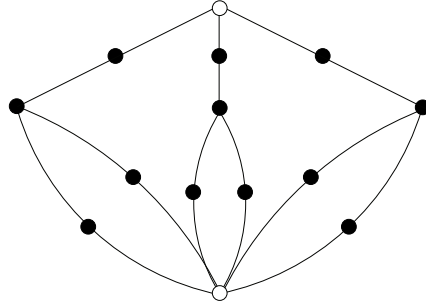


FIG. 5.2. Counterexample for the converse of Theorem 5.3.

Let $G_0 = (V_0, E_0)$ be the graph given in Figure 5.2 and $r_0 \in \{1, 2\}^{V_0}$ its associated connectivity type vector. The polytope $\text{CPP}(G_0, r_0)$ is integer. In fact it is reduced to the point with all components equal to 1. The couple (G_0, r_0) clearly has (G_p^1, r_p^1) as a minor, proving that the converse of Theorem 5.3 is not true.

Proof of Theorem 5.3. The proof is by induction on the number of edges. It is not hard to see that the statement holds for any graph with no more than two edges. Suppose that for any series-parallel graph G having no more than m edges and any connectivity type vector $r \in \{1, 2\}^V$ such that (G, r) has neither (G_p^1, r_p^1) nor (G_p^2, r_p^2) as a minor, we have that $\text{CPP}(G, r)$ is integer. Let us consider a series-parallel graph $G = (V, E)$ and a connectivity type vector $r \in \{1, 2\}^V$ such that G has $m+1$ edges, (G, r) has neither (G_p^1, r_p^1) nor (G_p^2, r_p^2) as a minor and $\text{CPP}(G, r)$ is not integer. There thus exists a fractional extreme point $\bar{x}$ of $\text{CPP}(G, r)$. We can choose $\bar{x}$ among all the fractional extreme points of $\text{CPP}(G, r)$ such that $|E_0(\bar{x}) \cup E_1(\bar{x})|$ is maximum. (Associated with $\bar{x}$, we consider the system $S(\bar{x})$ as well as the sets $E_0(\bar{x})$, $E_1(\bar{x})$ and $E_f(\bar{x})$ as defined in section 3.)

If $r(v) = 2$ for all $v \in V$, then from Theorem 5.2 the polytope $\text{CPP}(G, r)$ is integer. Therefore, without loss of generality, we can also suppose that the cardinality of the terminal set T is maximum. It means that for any series-parallel graph $G' = (V', E')$ and any connectivity type vector $r' \in \{1, 2\}^{V'}$ such that $|E'| = m + 1$, (G', r') has neither (G_p^1, r_p^1) nor (G_p^2, r_p^2) as a minor and $|\{v \in V' : r'(v) = 2\}| > |T|$, the polytope $\text{CPP}(G', r')$ is integer.

From the induction hypothesis, it follows that

$$(5.2) \quad \bar{x}(e) > 0 \quad \text{for all } e \in E,$$

that is, $E_0(\bar{x}) = \emptyset$.

CLAIM 1. *Any variable $x(e)$, $e \in E$, has a nonzero coefficient in at least two equations of $S(\bar{x})$.*

Proof. Let us consider an edge $f = uv$ of E , and denote $F = [u, v]$. From Remark 3.11, there exists at least one equation of $S(\bar{x})$ which contains $x(f)$ with a nonzero coefficient. Assume that there is exactly one such equation. Let $S'(\bar{x})$ be the system obtained from $S(\bar{x})$ by deleting all of the equations involving $x(e)$ for $e \in F$. We notice that the system $S'(\bar{x})$ contains exactly $|E| - |F|$ equations. Let x' be the restriction of $\bar{x}$ on $E \setminus F$. By Proposition 3.1, $x' \in \text{CPP}(G/F, r_F)$. Furthermore, x' is a solution of the system $S'(\bar{x})$. Since $S'(\bar{x})$ is nonsingular and its equations come from constraints of $\text{CPP}(G/F, r_F)$, this implies that x' is an extreme point of $\text{CPP}(G/F, r_F)$. By Proposition 3.12 and Remark 3.14, the point x' is fractional. As $(G/F, r_F)$ is a series-parallel graph having neither (G_p^1, r_p^1) nor (G_p^2, r_p^2) as a minor, this contradicts the induction hypothesis. $\square$

CLAIM 2. *The graph G is 2-node connected.*

Proof. Suppose the graph G is not 2-node connected. There then exists a node $u \in V$ defining a one-node cutset of G . Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be the two subgraphs of G such that $V_1 \cap V_2 = \{u\}$ and G is the 1-sum of G_1 and G_2 . For $i = 1, 2$ we associate with G_i the connectivity type vector $r_i \in \{1, 2\}^{V_i}$ such that $r_i(v) = r(v)$ if $v \in V_i \setminus \{u\}$, and $r_i(u) = \text{con}_G(V \setminus V_i)$. Let $\bar{x}_1$ and $\bar{x}_2$ be the two restrictions of $\bar{x}$ on E_1 and E_2 , respectively. From Lemma 4.1, both $\bar{x}_1$ and $\bar{x}_2$ are extreme points of $\text{CPP}(G_1, r_1)$ and $\text{CPP}(G_2, r_2)$, respectively. Since $E_f(\bar{x}) \neq \emptyset$, at least one among both restrictions is clearly fractional. Without loss of generality, we assume that $\bar{x}_1$ is fractional. We point out that (G_1, r_1) is nothing but the minor of (G, r) obtained by the contraction of E_2 . Therefore, since the couple (G, r) has neither (G_p^1, r_p^1) nor (G_p^2, r_p^2) as a minor, then neither does (G_1, r_1) . The latter combined with $|E_1| < |E|$ and $\bar{x}_1$ fractional contradicts the induction hypothesis. $\square$

The proof now proceeds by successively establishing a sequence of claims which build on each other. Therefore, for the sake of clarity, we only mention the highlights of our argument, and the detailed sequences and proofs are deferred to section 6.

Let $u \in V$ be a node having exactly two neighbors, say u_1 and u_2 . Since G is series-parallel, such a node u must exist [4]. Let us denote by F_1 and F_2 the set of edges between u and u_1 , and u and u_2 , respectively. Without loss of generality, we suppose

$$(5.3) \quad \bar{x}(F_1) \geq \bar{x}(F_2).$$

Suppose that $\bar{x}(F_1) < 1$. From the inequality (5.3), we also have $\bar{x}(F_2) < 1$, and then, by Remark 3.14, we obtain $F_1 = \{f_1 = uu_1\}$ and $F_2 = \{f_2 = uu_2\}$. From Proposition 3.13, there exists a partition $\{V_1, \dots, V_p\}$ of $P_1^*(\bar{x}) \cup P_2^*(\bar{x})$ such that $|\delta(V_1, \dots, V_p) \cap \{f_1, f_2\}| = 1$. Without loss of generality, assume that $f_1 \in \delta(V_1, \dots, V_p)$, and then $\{u, u_2\} \subseteq V_i$ for some $i \in \{1, \dots, p\}$. Hence, we obtain $\bar{x}[u, V_i \setminus \{u\}] = \bar{x}(f_2) < 1$, which contradicts Proposition 3.3. Therefore, $\bar{x}(F_1) \geq 1$.

CLAIM 3. $r(u) = 2$.

CLAIM 4. $r(u_2) = 1$.

CLAIM 5. *There does not exist a node of degree 2.*

CLAIM 6. $1 < \bar{x}(F_1) < 2$.

Let $F_1 = \{e_1, f_1\}$ with $\bar{x}(e_1) = 1$ and $0 < \bar{x}(f_1) < 1$. From Claims 3, 4, 5, 6, we can make the following remarks.

Remark 5.4. If v is a terminal adjacent to exactly two nodes, then between v and one of its neighbors there are exactly two edges.

Remark 5.5. Given two nodes v_1 and v_2 , then any path P between v_1 and v_2 , whose internal nodes have exactly two neighbors in G , satisfies the following.

- (a) All of the internal nodes of P are terminals.
- (b) P has at most two internal nodes.

Since $G = (V, E)$ is series-parallel and 2-connected, there is a 2-node cutset $\{v_1, v_2\}$ such that G decomposes with respect to $\{v_1, v_2\}$ into two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, where $\{v_1, v_2\} = V_1 \cap V_2$ and G_1 is a cycle, possibly with parallel edges (see Figure 5.3). Let L_1 and L_2 be the two walks of G_1 having only v_1 and v_2 in common. (In our context, a walk is a path between v_1 and v_2 , possibly with parallel edges.) We remark that if $|V| \geq 3$, then there exists such a decomposition with $|V_1| \geq 3$.

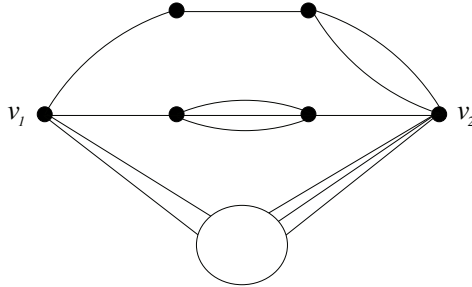


FIG. 5.3. Decomposition of G by the 2-node cutset $\{v_1, v_2\}$.

By Remark 5.5, all of the internal nodes of L_1 and L_2 are terminals and both L_1 and L_2 have each at most two internal nodes. Throughout the rest of the proof, we only consider the case where both L_1 and L_2 have each at most one internal node. The case of two internal nodes in L_1 or L_2 can be handled using the same arguments.

Without loss of generality, we can consider that L_1 has an internal node, say v . Under the induction assumptions, we remark that the claims 3, 4, and 6 are the results of the fact that u has exactly two neighbors. From now on, we consider $u = v$, $u_1 = v_1$, and $u_2 = v_2$.

CLAIM 7. $|F_2| = 1$

Let $F_2 = \{f_2 = uu_2\}$.

CLAIM 8. $\bar{x}(f_2) < 1$

CLAIM 9. L_2 contains an internal node.

Let $W = V \setminus \{u, u_1, u_2, u'\}$, where u' is an internal node in L_2 . We can suppose, without loss of generality, that $|\delta(u)| \leq |\delta(u')|$. From Remark 5.5(a), we have $r(u') = 2$. To finish the proof of our theorem, we give the following three claims.

CLAIM 10. $||u', u_1|| > ||u', u_2||$.

From Claims 6 and 7, we obtain $|\delta(u')| = |\delta(u)| = 3$. Using similar arguments for u' as those used for u , we have $F'_1 = \{e'_1, f'_1\}$ and $F'_2 = \{f'_2\}$ with $\bar{x}(e'_1) = 1$, $0 < \bar{x}(f'_1) < 1$, and $0 < \bar{x}(f'_2) < 1$.

CLAIM 11. $\bar{x}(f_1) = \bar{x}(f_2) = \bar{x}(f'_1) = \bar{x}(f'_2) = \frac{1}{2}$.

CLAIM 12. The couple (G, r) has the following properties:

- (a) $W \cap T \neq \emptyset$. (We recall that $T = \{u \in V : r(u) = 2\}$.)
- (b) $||S, u_1|| \geq 2$ and $||S, u_2|| \geq 1$ for every $S \subseteq W$ such that $G(S)$ is a connected component of $G(W)$ and $S \cap T \neq \emptyset$.

Among the series-parallel graphs, the class of outerplanar graphs has received particular attention (see, for instance, [3, 20]). A graph is called *outerplanar* if it can be embedded in the plane such that all nodes lie on the boundary of its exterior region. In order to give a complete linear description of the survivable network polytope on

outerplanar graphs, we give this second characterization devised from Kuratowski's theorem [16]. A graph is outerplanar if and only if it does not contain K_4 or $K_{2,3}$ as a minor. (We recall that $K_{2,3}$ is the complete bipartite graph having its node set decomposed into V_1 and V_2 with $|V_1| = 2$, $|V_2| = 3$, and $E(V_1) = E(V_2) = \emptyset$.)

First of all, we remark that the graph G_p of Figure 5.1 contains $K_{2,3}$ as a minor. In fact, $K_{2,3}$ can be obtained from G_p by deleting the three solid edges. Therefore, from the second definition of outerplanar graphs and Theorems 5.1, 5.2, and 5.3, we can deduce the following result.

THEOREM 5.6. *Let $G = (V, E)$ be an outerplanar graph and $r \in \{1, 2\}^V$ be an associated connectivity type vector. The survivable network polytope $SNDP(G, r)$ is then completely described by the trivial inequalities (2.1) and (2.2) together with the partition inequalities (2.5).*

We notice that there exist some series-parallel graphs which are not outerplanar and for which the polytope $CPP(G, r)$ is integer. For instance, from Theorem 5.3, we know that the cut and partition inequalities polytope is an integer for $K_{2,3}$ with exactly three terminals, and clearly, $K_{2,3}$ is not outerplanar. Therefore, the class of outerplanar graphs is strictly included in the subclass of series-parallel graphs implying an integer cut and partition inequalities polytope.

Since the separation problem for the partition inequalities (2.5) is polynomially solvable, a direct consequence of Theorem 5.6 is the following corollary.

COROLLARY 5.7. *The (1, 2)-survivable network design problem can be solved in polynomial time on outerplanar graphs.*

6. Proof of Theorem 5.3. In order to allow a better understanding of the proof of Theorem 5.3, we have just presented its main ideas. This section is thus devoted to give the details of the proof.

Proof of Claim 3. Suppose that $r(u) = 1$ and the system $S(\bar{x})$ satisfies Proposition 3.18. Consider a partition $\pi = \{V_1, \dots, V_p\}$ of $P_1^*(\bar{x}) \cup P_2^*(\bar{x})$ such that $F_1 \subseteq \delta(V_1, \dots, V_p)$. From Claim 1, we know that such a partition exists. Since $\bar{x}(F_1) \geq 1$, by Proposition 3.18(a), we have $\pi \in P_2^*(\bar{x})$. Without loss of generality, assume $u \in V_1$ and $u_1 \in V_2$. From Proposition 3.18(b) and the assumption $r(u) = 1$, we obtain $r(V_1 \setminus \{u\}) = 2$. We can then deduce from Remark 3.4 that $u_2 \in V_1$. With π , we associate the partition $\pi_a = \{V_1 \setminus \{u\}, V_2 \cup \{u\}, V_3, \dots, V_p\}$ which clearly is of type 2. Since $\bar{x}(\delta(V_1 \setminus \{u\}, V_2 \cup \{u\}, V_3, \dots, V_p)) \geq p$, we then obtain $\bar{x}(F_2) \geq \bar{x}(F_1)$. This last inequality combined with inequality (5.3) gives $\bar{x}(F_2) = \bar{x}(F_1)$, and thus, $\pi_a \in P_2(\bar{x})$.

Let $S_{F_1}(\bar{x})$ be the system arisen from $S(\bar{x})$ when we delete every equation induced by $e \in E_1(\bar{x}) \cap F_1$ and the ones induced by the partitions π containing F_1 , and we add the equations induced by the associated partitions π_a . (We remark that π_a might already belong to $P_2^*(\bar{x})$.) From Proposition 3.1, the restriction $\bar{x}_{F_1}$ of $\bar{x}$ on G/F_1 belongs to $CPP(G/F_1, r_{F_1})$. Moreover, the couple $(G/F_1, r_{F_1})$ clearly has neither (G_p^1, r_p^1) nor (G_p^2, r_p^2) as a minor. Since G/F_1 has less edges than G , by the induction hypothesis, the polytope $CPP(G/F_1, r_{F_1})$ is an integer. Therefore $\bar{x}_{F_1}$, which is clearly fractional, is not an extreme point of $CPP(G/F_1, r_{F_1})$. There must thus exist an extreme point $y \in CPP(G/F_1, r_{F_1})$ which is also a solution of $S_{F_1}(\bar{x})$. Let $\bar{y} \in \mathbb{R}^E$ be the unique point such that $\bar{y}(e) = y(e)$ if $e \in E \setminus F_1$, $\bar{y}(e) = 1$ if $e \in E_1(\bar{x}) \cap F_1$, and $\bar{y}(F_1) = y(F_2)$. We clearly have $\bar{y} \neq \bar{x}$. Moreover, it is obvious that $\bar{y}$ is also a solution of the system $S(\bar{x})$. This contradicts the fact that the system $S(\bar{x})$ is nonsingular. $\square$

Proof of Claim 4. Assume on the contrary that $r(u_2) = 2$. Let us consider a partition $\pi = \{V_1, V_2, \dots, V_p\}$ of $P_1^*(\bar{x}) \cup P_2^*(\bar{x})$ such that $u \in V_1$ and $u_1 \in V_2$.

From Claim 1, such a partition must exist and from Proposition 3.18, $\pi \in P_2^*(\bar{x})$. We claim that this partition exists providing $\bar{x}(F_1) = \bar{x}(F_2)$. In fact, suppose that $\bar{x}(F_1) > \bar{x}(F_2)$. Since $\bar{x}(\delta(u)) = \bar{x}(F_1) + \bar{x}(F_2) \geq 2$, we clearly get $\bar{x}(F_1) > 1$. If $u_2 \notin V_1 \cup V_2$, Remark 3.4 implies that the class V_1 is then reduced to the single node u . The partition $\pi_1 = \{V_1 \cup V_2, V_3, \dots, V_p\}$ is of type 2 because of $u \in V_1 \cup V_2$, $r(u) = 2$, $u_2 \notin V_1 \cup V_2$, and $r(u_2) = 2$. We then have

$$\begin{aligned}\bar{x}(\delta(V_1 \cup V_2, V_3, \dots, V_p)) &= \bar{x}(\delta(V_1, V_2, \dots, V_p)) - \bar{x}(F_1) \\ &= p - \bar{x}(F_1) \\ &< p - 1,\end{aligned}$$

which contradicts $\bar{x} \in \text{CPP}(G, r)$.

If $u_2 \in V_1$, the partition $\pi_2 = \{V_1 \setminus \{u\}, V_2 \cup \{u\}, V_3, \dots, V_p\}$ is obviously of type 2. We then get

$$\begin{aligned}\bar{x}(\delta(V_1 \setminus \{u\}, V_2 \cup \{u\}, V_3, \dots, V_p)) &= \bar{x}(\delta(V_1, V_2, \dots, V_p)) - \bar{x}(F_1) + \bar{x}(F_2) \\ &= p - \bar{x}(F_1) + \bar{x}(F_2) \\ &< p,\end{aligned}$$

which contradicts $\bar{x} \in \text{CPP}(G, r)$. We therefore deduce that the node u_2 belongs to V_2 , and thus $V_1 = \{u\}$. From Proposition 3.19, we could have chosen the system $S(\bar{x})$ such that $p = 2$, and thus partition π is nothing but the cut $\delta(u)$. Since $\bar{x}(F_1) > 1$, $\bar{x}(F_2) > 0$, and $\bar{x}(\delta(u)) = 2$, we clearly deduce that there exists $f_1 \in F_1 \cap E_f(\bar{x})$. This edge belongs to exactly one equation of $S(\bar{x})$, contradicting Claim 1. Therefore $\bar{x}(F_1) = \bar{x}(F_2)$.

We still consider the partition π introduced at the beginning of this proof. If $p = 2$, then we have $\bar{x}(F_1) = \bar{x}(F_2) = 1$ which implies that the equation induced by π is redundant with respect to $\bar{x}(e_1) = \bar{x}(e_2) = 1$, where $F_1 = \{e_1\}$ and $F_2 = \{e_2\}$. Hence, we can consider $p \geq 3$. From Proposition 3.19, we have either $u_2 \notin V_1 \cup V_2$ or $u_2 \in V_1$. If $u_2 \notin V_1 \cup V_2$ (respectively, $u_2 \in V_1$), then by considering the partition π_1 (respectively, π_2) previously defined, we obtain $\bar{x}(F_1) = 1$ and $\pi_1 \in P_2(\bar{x})$ (respectively, $\pi_2 \in P_2(\bar{x})$). Using the same arguments as the ones of the proof of Claim 3, we can exhibit a vector $\bar{y} \neq \bar{x}$ which is also a solution of $S(\bar{x})$. This contradiction completes our proof. $\square$

Proof of Claim 5. Suppose that the statement does not hold. Let v be a node of degree 2 adjacent to exactly two different nodes, say v_1 and v_2 . (Remark that the 2-node connectivity of G implies the existence of v_1 and v_2 .) Without loss of generality, we may suppose that $\bar{x}(vv_1) \geq \bar{x}(vv_2)$. By Claims 3 and 4, it then follows that $r(v) = 2$ and $r(v_2) = 1$. Let $r^* \in \mathbb{R}^V$ be the connectivity type vector such that $r^*(w) = r(w)$ if $w \neq v_2$, and $r^*(w) = 2$ if $w = v_2$. We claim that (G, r^*) contains neither (G_p^1, r_p^1) nor (G_p^2, r_p^2) as a minor. In fact, suppose, on the contrary, that (G, r^*) contains one of these minors, and let us denote that (H, r_H) contains one of these minors. Let $\pi = (S_1, \dots, S_5)$ be the partition of V that induces H , where S_1 corresponds to the ordinary node of H of degree 3, S_2, S_3, S_4 to the terminal nodes of H of degree 3, and S_5 to the node of H of degree 6. As $r^*(v_2) = 2$, v_2 belongs to a class S_i of connectivity type 2, that is, $i \neq 1$. If S_i contains a node $w \neq v_2$ with $r^*(w) = 2$, then (H, r_H) would also be a minor of (G, r) , a contradiction. Therefore $S_i \setminus \{v_2\}$ only contains nodes of connectivity type 1 with respect to r^* . In consequence, v belongs to S_j with $j \neq i$ and $j \neq 1$. Note that $v_1 \in S_j$. Otherwise as $G(S_j)$ is

connected, one would have $S_j = \{v\}$, and thus, H would contain a node of degree 2. Since neither (G_p^1, r_p^1) nor (G_p^2, r_p^2) has a node of degree 2, this is impossible.

If $i = 5$, then (H, r_H) is nothing but (G_p^2, r_p^2) . Since $r^*(S_i) = 2$ and $r^*(S_i \setminus \{v_2\}) = 1$, we have $r(S_i) = 1$, and (G_p^1, r_p^1) is a minor of (G, r) . This contradicts our hypothesis on (G, r) . Therefore $2 \leq i \leq 4$.

If $j = 5$, let us consider the partition $\pi' = (S'_1, \dots, S'_5)$ such that

$$\begin{aligned} S'_t &= S_t & t \in \{1, \dots, 4\} \setminus \{i\}, \\ S'_i &= S_i \cup \{v\}, \\ S'_5 &= S_5 \setminus \{v\}. \end{aligned}$$

We can easily see that the partition π' induces either (G_p^1, r_p^1) or (G_p^2, r_p^2) with respect to r . This yields to a contradiction, and then $2 \leq j \leq 4$.

Since $i \neq j$, $v \in S_j$, and $v_2 \in S_i$, we have $[S_i, S_j] \neq \emptyset$. Therefore, G is not series-parallel.

Consequently, (G, r^*) contains neither (G_p^1, r_p^1) nor (G_p^2, r_p^2) as a minor. Let us now prove that $\bar{x}$ is an extreme point of $\text{CPP}(G, r^*)$. We first show that $\bar{x}$ belongs to $\text{CPP}(G, r^*)$. Let $\pi = (V_1, V_2, \dots, V_p)$ be a partition of V . It is obvious that if v and v_2 are in the same class of π or if π is of type 2 with respect to r , then the type of π doesn't change by considering r^* instead of r . The inequality induced by π with respect to r^* is then satisfied by $\bar{x}$. Therefore, we only have to focus on the case where π is a partition of type 1 with respect to r and $e_2 \in \delta(V_1, V_2, \dots, V_p)$. Without loss of generality, suppose that $v \in V_1$ and $v_2 \in V_2$. Since π is of type 1 with respect to r , all of the nodes u with $r(u) = 2$ belong to V_1 . (We recall that (G, r) has at least two terminals which implies that $|V_1| \geq 2$.) The partition $(\{v\}, V_1 \setminus \{v\}, V_2, \dots, V_p)$ is of type 2 with respect to r , and hence, we have

$$\begin{aligned} \bar{x}(\delta(\{v\}, V_1 \setminus \{v\}, V_2, \dots, V_p)) &= \bar{x}(\delta(V_1, V_2, \dots, V_p)) + \bar{x}[\{v\}, V_1 \setminus \{v\}] \\ &\geq p + 1. \end{aligned}$$

Since $\bar{x}[\{v\}, V_1 \setminus \{v\}] \leq 1$, we obtain $\bar{x}(\delta(V_1, V_2, \dots, V_p)) \geq p$. Therefore, $\bar{x}$ belongs to $\text{CPP}(G, r^*)$. Using similar arguments, we can prove that $S(\bar{x})$ is a system of tight inequalities of $\text{CPP}(G, r^*)$. Thus, $\bar{x}$ is an extreme point of $\text{CPP}(G, r^*)$. As (G, r^*) has more terminals than (G, r) , this contradicts the maximality of T . $\square$

Proof of Claim 6. From Claim 5, we have $\bar{x}(F_1) > 1$. If $\bar{x}(F_1) \geq 2$, then by Proposition 3.20(a), the variable $x(e)$ belongs to exactly one equation of the system $S(\bar{x})$ with a nonzero coefficient, for all $e \in F_1$. Yet, this contradicts Claim 1. $\square$

Since by Claim 6, $1 < \bar{x}(F_1) < 2$, it follows from Remark 3.14 that there exist two edges $e_1, f_1 \in E$ such that $F_1 = \{e_1, f_1\}$, $e_1 \in E_1(\bar{x})$, and $f_1 \in E_f(\bar{x})$, that is, $\bar{x}(e_1) = 1$ and $0 < \bar{x}(f_1) < 1$.

Proof of Claim 7. Suppose that $|F_2| = 2$. We then have $F_2 = \{e_2, f_2\}$ with $\bar{x}(e_2) = 1$ and $0 < \bar{x}(f_2) < 1$. A consequence of Proposition 3.18(b) is that F_1 and F_2 cannot belong to a same partition inducing an equation of $S(\bar{x})$. Let $F_0 = [u_1, u_2]$. We are going to consider two cases. We first consider $F_0 \neq \emptyset$. Using Proposition 3.2(b) and the previous remark about F_1 and F_2 , we obtain $\bar{x}(F_0) < 1$. We then have $F_0 = \{f_0 = u_1 u_2\}$. Let $y \in \mathbb{R}^E$ be the point defined as follows:

$$y(e) = \begin{cases} \bar{x}(e) & \text{if } e \in E \setminus \{f_0, f_1, f_2\}, \\ \bar{x}(f_0) - \epsilon & \text{if } e = f_0, \\ \bar{x}(f_1) + \epsilon & \text{if } e = f_1, \\ \bar{x}(f_2) + \epsilon & \text{if } e = f_2, \end{cases}$$

where ϵ is any arbitrary scalar. The point y is also a solution of $S(\bar{x})$. In fact, any partition inducing an equation of $S(\bar{x})$ contains either none of the edges in $\{f_0, f_1, f_2\}$, or f_0 and exactly one among f_1 and f_2 . Since $y \neq \bar{x}$, this contradicts the extremality of $\bar{x}$.

We now suppose that $F_0 = \emptyset$. Let us denote by u^* the internal node in L_2 . Without loss of generality, we can suppose that $|\delta(u)| \leq |\delta(u^*)|$, which makes us deduce that $||[u^*, u_1]| = |[u^*, u_2]| = 2$. Proposition 3.2(b) then implies that there is no partition of $P_1(\bar{x}) \cup P_2(\bar{x})$ containing an edge in $\delta(u)$. This is a contradiction with Claim 1. $\square$

Proof of Claim 8. Let us suppose that $\bar{x}(f_2) = 1$. From Proposition 3.2(b) and $\bar{x}(F_1) > 1$, the system $S(\bar{x})$ can be chosen such that any of its equations containing $x(f_1)$ with a nonzero coefficient does not contain $x(f_2)$. In fact, suppose that there is a partition $\pi = \{V_1, \dots, V_p\}$ of $S(\bar{x})$ such that $\delta(u) \subset \delta(V_1, \dots, V_p)$. Thus, π is of type 2. Since $\bar{x}(\delta(u)) > 2$, we have $p \geq 3$. Without loss of generality, suppose that $V_1 = \{u\}$. If u_1 and u_2 are both in the same class of π , say V_2 , then $\bar{x}(\delta(V_1 \cup V_2, V_3, \dots, V_p)) = p - \bar{x}(\delta(u)) < p - 2$, which is a contradiction. Now suppose, without loss of generality, that $u_1 \in V_2$ and $u_2 \in V_3$. Since $\bar{x}[V_1, V_2] > 1$, then by Proposition 3.2, the partition $\{V_1 \cup V_2, V_3, \dots, V_p\}$ is of type 1. Thus, the partition $\pi' = \{V_1 \cup V_3, V_4, \dots, V_p\}$ is of type 1. Moreover, π' is tight for $\bar{x}$. We can thus replace in the system $S(\bar{x})$ the partition π by the partition π' and the equation $x(f_2) = 1$.

Let $\pi^1 = \{V_1, V_2, \dots, V_p\}$ be a partition of $S(\bar{x})$ such that $F_1 \subseteq \delta(V_1, \dots, V_p)$. We thus have $f_2 \notin \delta(V_1, V_2, \dots, V_p)$. By Proposition 3.18, we have $\pi^1 \in P_2^*(\bar{x})$. Without loss of generality, we suppose $\{u, u_2\} \subseteq V_1$ and $u_1 \in V_2$.

From Claim 1, there also exists a partition $\pi^2 = \{W_1, \dots, W_q\}$ inducing an equation of $S(\bar{x})$ such that $f_2 \in \delta(W_1, \dots, W_q)$. We are going to prove that the system $S(\bar{x})$ can be chosen such that $q \geq 3$. From above, we clearly have $F_1 \cap \delta(W_1, \dots, W_q) = \emptyset$. Suppose that $q = 2$. Without loss of generality, we assume that $u_2 \in W_1$. We then have $\{u, u_1\} \subseteq W_2$. From $\bar{x}(f_2) = 1$, it is obvious that $\text{con}(W_1) = 2$. Proposition 3.17 then implies that at most one class of the partition π^1 intersects both W_1 and W_2 . From the definition of π^1 , we have $V_1 \cap W_1 \neq \emptyset$ and $V_1 \cap W_2 \neq \emptyset$. We then obtain $V_2 \subseteq W_2$. Since $r(V \setminus (V_1 \cup V_2)) = 1$ and $\text{con}(W_1) = 2$, we get $r(V_1 \setminus \{u\}) = 2$. Thus, the partition $\{V_1 \setminus \{u\}, V_2 \cup \{u\}, V_3, \dots, V_p\}$ is clearly of type 2, and from Claim 7, we have

$$\begin{aligned} \bar{x}(\delta(V_1 \setminus \{u\}, V_2 \cup \{u\}, V_3, \dots, V_p)) &= \bar{x}(\delta(V_1, \dots, V_p)) - \bar{x}(F_1) + \bar{x}(f_2) \\ &= p - \bar{x}(F_1) + \bar{x}(f_2) \\ &< p, \end{aligned}$$

a contradiction.

Without loss of generality, we suppose $\{u, u_1\} \subseteq W_1$ and $u_2 \in W_2$. Since $\bar{x}(f_2) = 1$, Proposition 3.18 implies that $\pi^2 \in P_2^*(\bar{x})$ and $r(V \setminus (W_1 \cup W_2)) = 1$. Without loss of generality, let us assume that $W_i \cap V_2 \neq \emptyset$ for $i = 1, \dots, k$, and $W_i \cap V_2 = \emptyset$ for $i = k + 1, \dots, q$. In fact, we have $u_1 \in W_1$ and $u_1 \in V_2$. Moreover, since $\bar{x}(f_2) < \bar{x}(F_1)$ by Claim 7, the partition $\{V_1 \setminus \{u\}, V_2 \cup \{u\}, V_3, \dots, V_p\}$ must be of type 1, and then $r(V_1 \setminus \{u\}) = 1$. As $\pi^2 \in P_2^*(\bar{x})$ and $r(V \setminus (W_1 \cup W_2)) = 1$, we have $r(W_2) = 2$, which, combined with $r(V \setminus (V_1 \cup V_2)) = 1$ and $u \notin W_2$, implies that $W_2 \cap V_2 \neq \emptyset$. Let $\{U_1, \dots, U_k\}$ be the partition of V_2 such that $U_i = W_i \cap V_2$ for $i = 1, \dots, k$. It comes directly from the previous argument that $r(U_1) = r(U_2) = 2$. Therefore, by Proposition 3.9(a), we have $\bar{x}[U_1, \dots, U_k] \geq k - 1 + \bar{x}(F_1) - \bar{x}(f_2)$. Since

$\bar{x}(F_1) = \bar{x}(e_1) + \bar{x}(f_1) = 1 + \bar{x}(f_1)$, we obtain

$$(6.1) \quad \bar{x}[U_1, \dots, U_k] \geq k + \bar{x}(f_1) - \bar{x}(f_2).$$

Furthermore, we notice that $F_2 \cap [U_1, \dots, U_k] = \emptyset$. Hence, from (5.2), we have

$$(6.2) \quad \bar{x}[U_1, \dots, U_k] \leq \bar{x}(\delta(W_1, \dots, W_q)) - \bar{x}(f_2).$$

If $k = q$, by the inequalities (6.1) and (6.2), we then obtain $k + \bar{x}(f_1) - \bar{x}(f_2) \leq k - \bar{x}(f_2)$. This implies that $\bar{x}(f_1) \leq 0$ which contradicts (5.2). Suppose now that $k < q$. Since $r(V \setminus (W_1 \cup W_2)) = 1$ and $k \geq 2$, it is straightforward that the partition $\{Z_1, \dots, Z_{q-k+1}\}$ defined as

$$\begin{aligned} Z_1 &= \bigcup_{i=1}^k W_i, \\ Z_i &= W_{i+k-1}, \quad i = 2, \dots, q-k+1, \end{aligned}$$

is of type 1. We then have

$$\begin{aligned} \bar{x}(\delta(W_1, \dots, W_q)) - \bar{x}[U_1, \dots, U_k] - \bar{x}(f_2) &\geq \bar{x}(\delta(Z_1, \dots, Z_{q-k+1})) \\ &\geq q - k + 1 - 1 \\ &= q - k. \end{aligned}$$

Thus, we obtain $\bar{x}[U_1, \dots, U_k] + \bar{x}(f_2) \leq k$. From this last inequality combined with (6.1), we then get $\bar{x}(f_1) \leq 0$. This contradicts (5.2). Consequently, we obtain $\bar{x}(f_2) < 1$. $\square$

Proof of Claim 9. Suppose that $F_0 = [u_1, u_2] \neq \emptyset$. Suppose that $\bar{x}(F_0) \geq 1$. Let be $\pi = \{V_1, \dots, V_p\} \in S(\bar{x})$ such that $F_1 \in \delta(V_1, \dots, V_p)$. Since $\bar{x}(F_1) = \bar{x}[u, u_1] > 1$, by Proposition 3.18(a), π is of type 2. Since $\bar{x}(F_0) \geq 1$, by Proposition 3.18(b), u and u_2 are in the same class of π , say V_1 . Suppose, without loss of generality, that $u_1 \in V_2$. Since $F_0 \cup F_1 \in \delta(V_1, \dots, V_p)$ and $\bar{x}(F_0 \cup F_1) > 2$, we then have $p \geq 3$ and $\bar{x}(\delta(V_1 \cup V_2, V_3, \dots, V_p)) = p - \bar{x}(\delta(u)) < p - 2$, a contradiction. We therefore obtain $\bar{x}(F_0) < 1$. Let $F_0 = \{f_0 = u_1 u_2\}$. Let $y \in \mathbb{R}^E$ be the point defined as follows:

$$y(e) = \begin{cases} \bar{x}(e) & \text{if } e \in E \setminus \{f_0, f_2\}, \\ \bar{x}(f_0) - \epsilon & \text{if } e = f_0, \\ \bar{x}(f_2) + \epsilon & \text{if } e = f_2, \end{cases}$$

where $\epsilon = \min\{\bar{x}(f_0), 1 - \bar{x}(f_2)\}$. We first remark that any partition of $P_1(\bar{x}) \cup P_2(\bar{x})$, different from $\{\{u\}, V \setminus \{u\}\}$, contains either both f_0 and f_2 or none of them. If the partition $\{\{u\}, V \setminus \{u\}\}$ doesn't belong to $P_2(\bar{x})$, then the point y is a solution of $S(\bar{x})$ which is different to $\bar{x}$. This is a contradiction with the extremality of $\bar{x}$.

We suppose now that $\{\{u\}, V \setminus \{u\}\}$ belongs to $P_2(\bar{x})$. We first show that y is a point of $\text{CPP}(G, r)$. To obtain that, we only need to prove that for any partition $\pi = \{V_1, V_2, \dots, V_p\}$ with $f_0 \in \delta(V_1, V_2, \dots, V_p)$ and $f_2 \notin \delta(V_1, V_2, \dots, V_p)$, we have $y(\delta(V_1, V_2, \dots, V_p)) \geq t$, where $t = p - 1$ if π is of type 1 and $t = p$ otherwise. Without loss of generality, we suppose that $u \in V_1$. Let us consider the partition $\{\{u\}, V_1 \setminus \{u\}, V_2, \dots, V_p\}$. Note that this partition is of the same type as $\{V_1, V_2, \dots, V_p\}$.

Hence we have

$$\begin{aligned}\bar{x}(\delta(\{u\}, V_1 \setminus \{u\}, V_2, \dots, V_p)) &= y(\delta(\{u\}, V_1 \setminus \{u\}, V_2, \dots, V_p)) \\ &= y(\delta(V_1, V_2, \dots, V_p)) + y(f_2) \\ &= y(\delta(V_1, V_2, \dots, V_p)) + \bar{x}(f_2) + \epsilon \\ &\geq t + 1.\end{aligned}$$

This implies that

$$\begin{aligned}y(\delta(V_1, V_2, \dots, V_p)) &\geq t + 1 - (\bar{x}(f_2) + \epsilon) \\ &\geq t + 1 - 1 = t.\end{aligned}$$

From the definition of ϵ , we clearly have $0 \leq y(f_0) \leq 1$ and $0 \leq y(f_2) \leq 1$. Therefore, y belongs to $\text{CPP}(G, r)$.

We remark that $|E_f(y)| < |E_f(\bar{x})|$ since at least one variable among $y(f_0)$ and $y(f_2)$ is an integer. Moreover, as $0 < y(f_1) < 1$, y is fractional. By the induction hypothesis, y isn't an extreme point of $\text{CPP}(G, r)$. Hence, there exist $t \geq 2$ extreme points $y^1, \dots, y^t$ of $\text{CPP}(G, r)$ and t scalars $0 < \alpha_i < 1$, $i = 1, \dots, t$, such that

$$y = \sum_{i=1}^t \alpha_i y^i \quad \text{and} \quad \sum_{i=1}^t \alpha_i = 1.$$

It is clear that $|E_f(y^i)| < |E_f(\bar{x})|$ for $i = 1, \dots, t$. From the extremality of the y^i and the induction hypothesis on $\bar{x}$, we obtain that the points y^i are integer. All of the constraints that are tight for y are also tight for y^i . Moreover, since $y(\delta(u)) = 2 + \epsilon < 3$, there exists $i_0 \in \{1, \dots, t\}$ such that $y^{i_0}(\delta(u)) < 3$. The integrality of y^{i_0} then implies that $y^{i_0}(\delta(u)) = 2$. Therefore, the point y^{i_0} is also a solution of $S(\bar{x})$ which contradicts the extremality of $\bar{x}$. We then conclude that $F_0 = \emptyset$. $\square$

Proof of Claim 10. Denote $F'_1 = [u', u_1]$ and $F'_2 = [u', u_2]$. Suppose that $|F'_2| \geq 2$. From (5.2) and Remark 3.14, we have $\bar{x}(F'_2) > 1$. Let $\pi = \{V_1, \dots, V_p\}$ be a partition of $P_1^*(\bar{x}) \cup P_2^*(\bar{x})$ different from $\{\{u\}, V \setminus \{u\}\}$. By Claim 1, such a partition exists. Moreover, Claim 8 together with Proposition 3.3 implies that $p \geq 3$. From Proposition 3.19, u , u_1 , and u_2 belong to three different classes of π . Without loss of generality, suppose that $u \in V_1$, $u_1 \in V_2$, and $u_2 \in V_3$. We remark that $V_1 = \{u\}$. Using Proposition 3.18, we obtain that $u' \in V_3$. Therefore, π is a partition of type 2. The partition $\pi' = \{V_1 \cup V_2, V_3, \dots, V_p\}$ is of type 2, and is such that

$$\begin{aligned}\bar{x}(\delta(V_1 \cup V_2, V_3, \dots, V_p)) &= \bar{x}(\delta(V_1, \dots, V_p)) - \bar{x}(F_1) \\ &< p - 1.\end{aligned}$$

The last inequality comes from $\bar{x}(F_1) > 1$. We then get a contradiction. We conclude that $|F'_2| = 1$. By Claims 5 and 6, we have $|F'_1| = 2$. $\square$

To make the proofs of the next two claims clearer and shorter, we introduce additional notation. Given an edge subset $F \subseteq E$, we denote by $P(\bar{x}, F)$ the subset of partitions $\pi = \{V_1, \dots, V_p\} \in P_1^*(\bar{x}) \cup P_2^*(\bar{x})$ such that $p \geq 3$ and $F \subseteq \delta(V_1, \dots, V_p)$.

Proof of Claim 11. In order to prove the result, it is enough to prove that $\bar{x}(F_1) = \bar{x}(F'_1)$, $\bar{x}(\delta(u)) = \bar{x}(\delta(u')) = 2$, and $\bar{x}(f_2) + \bar{x}(f'_2) = 1$.

We first prove that $\bar{x}(F_1) = \bar{x}(F'_1)$. Without loss of generality, we suppose $\bar{x}(F_1) \geq \bar{x}(F'_1)$. Since $0 < \bar{x}(f_1) < 1$, by Claim 1, there must exist a partition $\pi = \{V_1, \dots, V_p\}$ such that $\pi \in P(\bar{x}, F_1)$. By Proposition 3.18(a), $\pi \in P_2^*(\bar{x})$. Without loss of generality,

assume $u \in V_1$ and $u_1 \in V_2$. From $\bar{x}(F_2) < 1$ and Proposition 3.3(c), we obtain $V_1 = \{u\}$. Since $\bar{x}[u, u_1] > 1$, by Proposition 3.18(b), we have $r(V \setminus (V_1 \cup V_2)) = 1$, and then $u' \in V_2$. Consider then the partition $\{\{u'\}, (V_2 \setminus \{u'\}) \cup \{u\}, V_3, \dots, V_p\}$. It clearly is a partition of type 2, and then we have

$$\begin{aligned} \bar{x}(\delta(\{u'\}, (V_2 \setminus \{u'\}) \cup \{u\}, V_3, \dots, V_p)) &= \bar{x}(\delta(V_1, \dots, V_p)) + \bar{x}(F'_1) - \bar{x}(F_1) \\ &= p + \bar{x}(F'_1) - \bar{x}(F_1) \\ &\geq p. \end{aligned}$$

We get $\bar{x}(F_1) \leq \bar{x}(F'_1)$, and consequently $\bar{x}(F_1) = \bar{x}(F'_1)$.

Now we are going to prove that $\bar{x}(\delta(u)) = \bar{x}(\delta(u')) = 2$. Since $0 < \bar{x}(f_1) < 1$, as we have shown in the proof of Claim 10, the system $S(\bar{x})$ can be chosen such that there exists a partition $\pi^0 = \{V_1^0, \dots, V_{p_0}^0\} \in P(\bar{x}, F_1)$ such that u, u_1 , and u_2 belong to three different classes of π_0 . Without loss of generality, suppose that $V_1^0 = \{u\}$, $u_1 \in V_2^0$, and $u_2 \in V_3^0$. Since $r(u') = 2$, Proposition 3.18(b) implies $u' \in V_2^0$. Let $\pi'_0 = \{V'_1, V'_2, V'_3, \dots, V'_{p_0}\}$, where $V'_1 = \{u'\}$ and $V'_2 = (V_2^0 \setminus \{u'\}) \cup \{u\}$. This partition clearly is of type 2, and since $\bar{x}(F_1) = \bar{x}(F'_1)$, we get

$$\begin{aligned} \bar{x}(\delta(V'_1, V'_2, V'_3, \dots, V'_{p_0})) &= \bar{x}(\delta(V_1^0, \dots, V_{p_0}^0)) + \bar{x}(F'_1) - \bar{x}(F_1) \\ &= p_0 + \bar{x}(F'_1) - \bar{x}(F_1) \\ &= p_0. \end{aligned}$$

Let π be any partition of $P(\bar{x}, F_1) \setminus \{\pi_0\}$. The partition π' obtained from π by switching u and u' , as π'_0 was obtained above from π_0 , belongs to $P_2(\bar{x})$. Consider the system $S'(\bar{x})$ obtained from $S(\bar{x})$ by adding the equation induced by π'_0 and replacing those induced by $\pi \in P(\bar{x}, F_1) \setminus \{\pi_0\}$ by the ones induced by π' . This system is nonsingular, and $\bar{x}$ is its unique solution. The system $S(\bar{x})$ can then be chosen such that $P(\bar{x}, F_1) = \{\pi_0\}$. Since $0 < \bar{x}(f_1) < 1$, Claim 1 implies $\bar{x}(\delta(u)) = 2$, and thus, $\bar{x}(\delta(u')) = 2$. Moreover, we have $\bar{x}(F_1) = \bar{x}(F'_1)$, and we then get $\bar{x}(f_2) = \bar{x}(f'_2)$.

Finally, we prove that $\bar{x}(f_2) + \bar{x}(f'_2) = 1$. Suppose that $\bar{x}(f_2) + \bar{x}(f'_2) > 1$. (The case $\bar{x}(f_2) + \bar{x}(f'_2) < 1$ is similar.) Let $S_{F_1}(\bar{x})$ be the system obtained from $S(\bar{x})$ by removing the equations $x(e_1) = 1$, $x(\delta(u)) = 2$, and the one induced by π_0 . Let $\bar{x}_1$ be the restriction of $\bar{x}$ on G/F_1 . By the induction hypothesis, the polytope $\text{CPP}(G/F_1, r_{F_1})$ is an integer. Since $\bar{x}_1$ is fractional, it is not an extreme point of $\text{CPP}(G/F_1, r_{F_1})$. Hence, there exist $k \geq 2$ extreme points $y^1, \dots, y^k$ of $\text{CPP}(G/F_1, r_{F_1})$ and k scalars $0 < \alpha_i < 1$, $i = 1, \dots, k$, such that

$$\bar{x}_1 = \sum_{i=1}^k \alpha_i y^i \quad \text{and} \quad \sum_{i=1}^k \alpha_i = 1.$$

By the integrality of $\text{CPP}(G/F_1, r_{F_1})$, the points y^i , $i = 1, \dots, k$, are integer. Since $\bar{x}$ is a solution of $S_{F_1}(\bar{x}_1)$ and is a convex combination of the points y^i , $i = 1, \dots, k$, it is clear that y^i is also a solution of $S_{F_1}(\bar{x}_1)$ for all $i \in \{1, \dots, k\}$. There must then exist some $j \in \{1, \dots, k\}$ such that $y^j(f_2) + y^j(f'_2) > 1$. Since y^j is integer, we have $y^j(f_2) + y^j(f'_2) \geq 2$, which gives $y^j(f_2) = y^j(f'_2) = 1$. Let $\bar{y} \in \mathbb{R}^E$ be the point defined below

$$\bar{y}(e) = \begin{cases} y^j(e) & \text{if } e \in E \setminus F_1, \\ 1 & \text{if } e = e_1, \\ 0 & \text{if } e = f_1. \end{cases}$$

This point is also a solution of $S(\bar{x})$. In fact, additionally, to be a solution of $S_{F_1}(\bar{x})$, y^j clearly satisfies the equations $x(e_1) = 1$, $x(\delta(u)) = 2$, and $x(\delta(V_1^0, \dots, V_{p_0}^0)) = p_0$, which are the only three equations removed from $S(\bar{x})$ to get $S_{F_1}(\bar{x})$. Since $\bar{y} \neq \bar{x}$, this contradicts the fact that $\bar{x}$ is an extreme point of $\text{CPP}(G, r)$. $\square$

Proof of Claim 12. We first prove (a), that is, $W \cap T \neq \emptyset$. From Claim 4 we know that $u_2 \notin T$. Suppose now that $T \subseteq \{u, u', u_1\}$. Let $G^* = (V^*, E^*)$ be the graph obtained from G by contracting F_1 and F'_1 , and deleting f_2 as well. Let $w \in V^*$ be the node arising from the contraction, that is, $V^* \setminus V = \{w\}$. We have $r(v) = 1$ for all $v \in V \setminus \{u, u', u_1\}$ and therefore, the connectivity type vector r^* associated with V^* can be defined such that $r^*(v) = 1$ for all $v \in V^*$. Let x^* be the restriction of $\bar{x}$ on E^* .

We are going to show that $x^* \in \text{CPP}(G^*, r^*)$. Consider a partition $\pi^* = \{V_1^*, \dots, V_p^*\}$, $p \geq 2$, of V^* . (We remark that any partition of V^* is obviously of type 1 with respect to r^* .) Without loss of generality, we suppose $w \in V_1^*$. Let $\{U_1, \dots, U_p\}$ be the partition of V such that $U_1 = (V_1^* \setminus \{w\}) \cup \{u, u', u_1\}$ and $U_i = V_i^*$ for all $i = 2, \dots, p$. This partition clearly is of type 1 with respect to r . If $f'_2 \notin \delta(V_1^*, \dots, V_p^*)$, then we get

$$\begin{aligned} x^*(\delta(V_1^*, \dots, V_p^*)) &= \bar{x}(\delta(U_1, \dots, U_p)) \\ &\geq p - 1. \end{aligned}$$

If $f'_2 \in \delta(V_1^*, \dots, V_p^*)$, we suppose, without loss of generality, that $u_2 \in V_2^*$. By considering the partition $\{\{u\}, U_1 \setminus \{u\}, U_2, \dots, U_p\}$ of V which is of type 2 with respect to r , we then have

$$\begin{aligned} x^*(\delta(V_1^*, \dots, V_p^*)) &= \bar{x}(\delta(\{u\}, U_1 \setminus \{u\}, U_2, \dots, U_p)) - \bar{x}(\delta(u)) \\ &\geq p + 1 - 2 \\ &= p - 1. \end{aligned}$$

Therefore, we conclude that $x^* \in \text{CPP}(G^*, r^*)$.

Proposition 3.16 and $\bar{x}(\delta(u)) = 2$ imply that the system $S(\bar{x})$ can be chosen such that $P_1^*(\bar{x}) = \emptyset$. Let $\pi_0^* = \{U_1^0, \dots, U_{p_0-1}^0\}$ be the partition of V^* such that $U_1^0 = (V_2^0 \setminus \{u', u_1\}) \cup \{w\}$ and $U_i^0 = V_{i+1}^0$ for all $i = 2, \dots, p_0 - 1$, where $\pi^0 = \{V_1^0, \dots, V_{p_0}^0\}$ is the unique partition in $P(\bar{x}, F_1)$. (See the proof of Claim 11 for the definition of π^0 .) We then get

$$\begin{aligned} x^*(\delta(U_1^0, \dots, U_{p_0-1}^0)) &= \bar{x}(\delta(V_1^0, \dots, V_{p_0}^0)) - \bar{x}(\delta(u)) \\ &= p_0 - 2. \end{aligned}$$

Let $\pi_1 = \{V_1, \dots, V_p\}$ be a partition in $P(\bar{x}, F'_1)$ and then, in $P_2^*(\bar{x})$ too. We suppose, without loss of generality, that $\{u, u_1\} \subseteq V_1$ and $V_2 = \{u'\}$. Let $\pi_1^* = \{V'_1, \dots, V'_{p-1}\}$ be the partition of V^* such that $V'_1 = (V_1 \setminus \{u, u_1\}) \cup \{w\}$ and $V'_i = V_{i+1}$ for all $i = 2, \dots, p - 1$. This partition clearly is of type 1 with respect to r^* , and then

$$\begin{aligned} x^*(\delta(V'_1, \dots, V'_{p-1})) &= \bar{x}(\delta(V_1, \dots, V_p)) - (\bar{x}(f_2) + \bar{x}(F'_1)) \\ &= p - \left(\frac{1}{2} + \frac{3}{2}\right) \\ &= p - 2. \end{aligned}$$

Let $Q^*(F'_1) = \{\pi_1^* : \pi_1 \in P(\bar{x}, F'_1)\}$ be the set of partitions of V^* obtained from the partitions in $P(\bar{x}, F'_1)$ as described above. The point x^* is a solution of the following

system:

$$S(x^*) \begin{cases} x(e) = 1 & \text{for all } e \in E_1(\bar{x}) \setminus \{e_1, e'_1\}, \\ x(\delta(U_1^0, \dots, U_{p_0-1}^0)) = p_0 - 2, \\ x(\delta(V_1^*, \dots, V_p^*)) = p - 2 & \text{for all } \{V_1^*, \dots, V_p^*\} \in Q^*(F'_1). \end{cases}$$

Since x^* is fractional, from Theorem 5.1, it cannot be an extreme point of $\text{CPP}(G^*, r^*)$. There must then exist an extreme point y^* of $\text{CPP}(G^*, r^*)$ such that y^* is a solution of $S(x^*)$ and $y^*(f_2^*) > 0$. By Theorem 5.1, y^* is an integer, and thus, $y^*(f_2^*) = 1$. Let $\bar{y} \in \mathbb{R}^E$ be the point defined as

$$\bar{y}(e) = \begin{cases} y^*(e) & \text{if } e \notin \delta(u) \cup F'_1, \\ 1 & \text{if } e \in \{e_1, e'_1, f_2\}, \\ 0 & \text{if } e \in \{f_1, f'_1\}. \end{cases}$$

The point $\bar{y}$ clearly satisfies the two equations $x(\delta(u)) = 2$ and $x(\delta(u')) = 2$. Therefore, since $P_2^*(\bar{x}) = P(\bar{x}, F_1) \cup P(\bar{x}, F'_1)$ and y^* is a solution of $S(x^*)$, the point $\bar{y}$ is also a solution of $S(\bar{x})$. This leads to a contradiction, and we conclude that $W \cap T \neq \emptyset$.

To prove (b), we consider a node subset $S \subseteq W$ such that $G(S)$ is a connected component of $G(W)$ and $S \cap T \neq \emptyset$. Suppose that $[S, u_1] = \emptyset$. (The proof for the case $[S, u_2] = \emptyset$ is similar.) The node u_2 then defines a one-node cutset of G which contradicts Claim 2. Therefore $|[S, u_1]| \geq 1$.

Suppose that $|[S, u_1]| = 1$. From Proposition 3.18(b), we have $r(V \setminus (V_1^0 \cup V_2^0)) = 1$, where $\pi_0 = \{V_1^0, \dots, V_{p_0}^0\}$ is the only partition in $P(\bar{x}, F_1)$ as previously defined. We then have $(W \cap T) \subset V_2^0$. Thus, $S \cap T \subset V_2^0$. Let $S^0 = S \cap V_2^0$. We obviously have $r(S^0) = 2$. Since $|[S, u_1]| = 1$, we get $\bar{x}[S^0, u_1] \leq 1$. By considering the partition $\{S^0, (V_2^0 \setminus S^0) \cup \{u\}, V_3^0, \dots, V_{p_0}^0\}$ which is of type 2, we have

$$\begin{aligned} \bar{x}(\delta(S^0, (V_2^0 \setminus S^0) \cup \{u\}, V_3^0, \dots, V_{p_0}^0)) &= \bar{x}(\delta(V_1^0, \dots, V_{p_0}^0)) - \bar{x}(F_1) + \bar{x}[S^0, u_1] \\ &\leq p_0 - \frac{3}{2} + 1 \\ &= p_0 - \frac{1}{2}. \end{aligned}$$

This leads to a contradiction with $\bar{x} \in \text{CPP}(G, r)$. Consequently, we have $|[S, u_1]| \geq 2$. $\square$

7. New facet-defining inequalities for the polytope $\text{SNDP}(G, r)$. In this section, we introduce a new family of facet-defining inequalities of the polytope $\text{SNDP}(G, r)$. Given a graph $G = (V, E)$, not necessarily series-parallel, and a connectivity type vector $r \in \{1, 2\}^V$, let $\{V_0, V_1, \dots, V_t, V_{t+1}\}$ be a partition of V such that $t \geq 2$ and

- (a) $r(V_0) = 1$,
- (b) $r(V_i) = 2$ for $i = 1, 2, \dots, t$,
- (c) $[V_i, V_j] = \emptyset$ for $i = 1, 2, \dots, t-1$ and $j = i+1, \dots, t$.

Consider the inequality

$$(7.1) \quad x(\delta(V_0, V_1, \dots, V_t, V_{t+1})) + x(\delta(V_0)) \geq 2t + 2$$

THEOREM 7.1. *Inequality (7.1) is valid for the polytope $\text{SNDP}(G, r)$.*

Proof. Let $H = (V, F)$ be a survivable subgraph of G . Let $F' = F \cap \delta(V_0, V_1, \dots, V_t, V_{t+1})$. It suffices to prove that $|F'| + |F' \cap \delta(V_0)| \geq 2t + 2$. Since $r(V_0) = 1$, we have $|F' \cap \delta(V_0)| \geq 1$. If $|F' \cap \delta(V_0)| = 1$, let $i_0 \in \{1, \dots, t\}$ such that $F' \cap \delta(V_0) = [V_0, V_{i_0}]$. Since H is survivable, we have $|F' \cap \delta(V_0 \cup V_{i_0})| \geq 2$ and $|F' \cap \delta(V_i)| \geq 2$ for all $i \in \{1, \dots, t\} \setminus \{i_0\}$. We then obtain that $|F' \setminus \delta(V_0)| \geq 2t$. From $|F' \setminus \delta(V_0)| + 2|F' \cap \delta(V_0)| = |F'| + |F' \cap \delta(V_0)|$, we get $|F'| + |F' \cap \delta(V_0)| \geq 2t + 2$. If $|F' \cap \delta(V_0)| \geq 2$, by summing up the $t + 1$ inequalities $|F' \cap \delta(V_i)| \geq 2$ for all $i \in \{0, 1, \dots, t\}$, we obtain $|F'| + |F' \cap \delta(V_0)| \geq 2t + 2$. $\square$

We call inequalities of type (7.1) *spinning-top inequalities*. In the next theorem, we give necessary and sufficient conditions for spinning-top inequalities to define facets of $\text{SNDP}(G, r)$ when $G(V_{t+1})$ is 2-edge connected. We denote by $\mathcal{F}$ the face of $\text{SNDP}(G, r)$ induced by a spinning-top inequality. From the proof of Theorem 7.1, we give the following remark which is useful to prove the theorem.

Remark 7.2. Let (V, F) be a survivable subgraph of G . If $|F \cap \delta(V_0)| \geq 3$ or $|F \cap \delta(V_0)| \geq 4$ for some $i \in \{1, \dots, t\}$, then $x^F \notin \mathcal{F}$.

THEOREM 7.3. Suppose that $G(V_{t+1})$ is 2-edge connected. Inequality (7.1) then defines a facet of the polytope $\text{SNDP}(G, r)$ if and only if the following holds:

- (a) $G(V_i)$ is 2-edge connected for $i = 0, 1, \dots, t$,
- (b) $|[V_0, V_i]| \geq 1$ for $i = 1, \dots, t$, and
- (c) $|[V_i, V_{t+1}]| \geq 2$ for $i = 1, \dots, t$.

Proof. The proof uses Remark 7.2 and standard polyhedral techniques. $\square$

A direct consequence of Theorem 7.3 is that for a general couple (G, r) , inequality (7.1) must be considered to obtain a complete linear description of the polytope $\text{SNDP}(G, r)$. The previous statement remains true even when G is a series-parallel graph, as we showed in section 5.

8. Final remarks. In this paper, we studied the polytope $\text{CPP}(G, r)$ given by the trivial inequalities and the partition ones. We first gave some structural properties of the extreme points of $\text{CPP}(G, r)$. Using these, we proved that the polytope $\text{CPP}(G, r)$ is an integer on a nontrivial subclass of series-parallel graphs, which includes the outerplanar graph class. This result leads to a polynomial-time algorithm, based on the ellipsoid method, for solving the (1, 2)-survivable network design problem in that class of graphs. To the best of our knowledge, the complete linear description of $\text{SNDP}(G, r)$ given in this paper is the first one which combines even and odd connectivity types for an important class of graphs. We also introduced a new family of facet-defining inequalities for the polytope $\text{SNDP}(G, r)$, called the *spinning-top inequalities*. This class of inequalities must be considered in linear descriptions of the survivable network polytope $\text{SNDP}(G, r)$.

From Theorem 5.3, we can deduce that $\text{CPP}(G, r)$ is an integer when G is a series-parallel graph and $|T| \leq 2$. (We remind that T is the set of terminal nodes, that is, nodes u such that $r(u) = 2$.) For a general graph G and $|T| \leq 1$, we know that $\text{SNDP}(G, r) = \text{CPP}(G, r)$ [5]. Moreover, for a general graph G , a positive cost function, and $|T| = 2$, the survivable network design problem can be solved in polynomial time, since Arkin and Hassin [1] have shown that this special case of the SNDP can be reduced to the matroid intersection problem. In view of the previous discussion, we give the following conjecture.

CONJECTURE 8.1. Let $G = (V, E)$ be a graph and $r \in \{1, 2\}^V$ its connectivity type vector such that $|\{u \in V : r(u) = 2\}| = 2$. The polytope $\text{SNDP}(G, r)$ is then completely described by both the trivial and the partition inequalities (i.e., $\text{CPP}(G, r)$ is an integer).

An important problem which deserves to be addressed is to study the separation problem of the spinning-top inequalities. We think that this problem is polynomial on series-parallel graphs and NP-hard in general.

Our work has started with the objective of giving a linear description of the polytope $\text{SNDP}(G, r)$ on series-parallel graphs. Our study then leads us to give the following conjecture.

CONJECTURE 8.2. *Let $G = (V, E)$ be a series-parallel graph and $r \in \{1, 2\}^V$ its connectivity type vector. The polytope $\text{SNDP}(G, r)$ is then completely described by the trivial, partition, and spinning-top inequalities.*

REFERENCES

- [1] E. ARKIN AND R. HASSIN, *The k -path tree matroid and its application to survivable network design*, J. Combin. Optim., 5 (2008), pp. 314–322.
- [2] F. BARAHONA AND H. KERIVIN, *Separation of partition inequalities with terminals*, Discrete Optim., 1 (2004), pp. 129–140.
- [3] S. CHOPRA, *The k -edge-connected spanning subgraph polyhedron*, SIAM J. Discrete Math., 7 (1994), pp. 245–259.
- [4] R. J. DUFFIN, *Topology of series-parallel networks*, J. Math. Anal. Appl., 10 (1965), pp. 303–318.
- [5] D. R. FULKERSON, *Blocking polyhedra*, in Graph Theory and its Applications (Proc. Adv. Sem., University of Wisconsin, Madison, WI, 1969; B. Harris, ed.), Academic Press, New York, 1970, pp. 93–112.
- [6] A. V. GOLDBERG AND R. E. TARJAN, *A new approach to the maximum flow problem*, J. Assoc. Comput. Mach., 35 (1988), pp. 921–940.
- [7] M. GRÖTSCHEL, L. LOVÁSZ, AND A. SCHRIJVER, *The ellipsoid method and its consequences in combinatorial optimization*, Combinatorica, 1 (1981), pp. 169–197.
- [8] M. GRÖTSCHEL AND C. L. MONMA, *Integer polyhedra arising from certain network design problems with connectivity constraints*, SIAM J. Discrete Math., 3 (1990), pp. 502–523.
- [9] M. GRÖTSCHEL, C. L. MONMA, AND M. STOER, *Facets for polyhedra arising in the design of communication networks with low-connectivity constraints*, SIAM J. Optim., 2 (1992), pp. 474–504.
- [10] M. GRÖTSCHEL, C. L. MONMA, AND M. STOER, *Computational results with a cutting plane algorithm for designing communication networks with low-connectivity constraints*, Oper. Res., 40 (1992), pp. 309–330.
- [11] M. GRÖTSCHEL, C. L. MONMA, AND M. STOER, *Design of survivable networks*, Handbook Oper. Res. Management Sci. 7, M. O. Ball et al., eds., North-Holland, Amsterdam, 1995, pp. 617–671.
- [12] H. KERIVIN AND A. R. MAHJOUB, *Separation of partition inequalities for the $(1, 2)$ -survivable network design problem*, Oper. Res. Lett., 30 (2002), pp. 265–268.
- [13] H. KERIVIN AND A. R. MAHJOUB, *Design of Survivable Networks: A Survey*, Networks, 46 (2005), pp. 1–21.
- [14] H. KERIVIN, C. NOCQ, AND A. R. MAHJOUB, *$(1, 2)$ -survivable networks: Facets and branch-and-cut*, in The Sharpest Cut, M. Grötschel, ed., MPS/SIAM Ser. Optim., SIAM, Philadelphia, 2004, pp. 121–152.
- [15] J. B. KRUSKAL, *On the shortest spanning subtree of a graph and the traveling salesman problem*, Proc. Amer. Math. Soc., 7 (1956), pp. 48–50.
- [16] C. KURATOWSKI, *Sur le problème des courbes gauches en topologie*, Fund. Math., 15 (1930), pp. 271–283.
- [17] A. R. MAHJOUB, *Two-edge connected spanning subgraphs and polyhedra*, Math. Programming, 64 (1994), pp. 199–208.
- [18] C. L. MONMA AND D. F. SHALLCROSS, *Methods for designing communication networks with certain two-connected survivability constraints*, Oper. Res., 37 (1989), pp. 531–541.
- [19] M. STOER, *Design of Survivable Networks*, Lecture Notes in Math. 1531, Springer-Verlag, Berlin, 1992.
- [20] P. WINTER, *Generalized Steiner problem in outerplanar networks*, BIT, 25 (1985), pp. 485–496.