COMBINATORICA
Akadémiai Kiadó – Springer-Verlag

COMBINATORICA **16** (1) (1996) 59–85

GRAPHS OF SMALL DIMENSIONS

NANCY EATON* and VOJTĚCH RÖDL†

Received February 14, 1994
Revised March 16, 1995

Let $G=(V,E)$ be a graph with n vertices. The direct product dimension $\operatorname{pdim}(G)$ (c.f. [10], [12]) is the minimum number t such that G can be embedded into a product of t copies of complete graphs K_n.

In [10], Lovász, Nešetřil and Pultr determined the direct product dimension of matchings and paths and gave sharp bounds for the product dimension of cycles, all logarithmic in the number of vertices.

Here we prove that $\operatorname{pdim}(G) \le cd\log n$ for any graph with maximum degree d and n vertices and show that up to a factor of $1/\left(\log d+\log\log\frac{n}{2d}\right)$ this bound is the best possible.

We also study set representations of graphs. Let $G=(V,E)$ be a graph and $p\ge 1$ an integer. A family $\mathcal{F}=\{A_x, x\in V\}$ of (not necessarily distinct) sets is called a p-intersection representation of G if $|A_x\cap A_y|\ge p \Leftrightarrow \{x,y\}\in E$ for every pair x, y of distinct vertices of G. Let $\theta_p(G)$ be the minimum size of $|U\mathcal{F}|$ taken over all intersection representations of G. We also study the parameter $\theta(G)=\min\limits_p(\theta_p(G))$.

It turns out that these parameters can be bounded in terms of maximum degree and linear density of a graph G or its complement $\overline{G}$. While for example, $\theta_1(G)=|E(G)|$ holds if G contains no triangle, N. Alon proved that $\theta_1(G)\le c\Delta(\overline{G})\log n$, where $\Delta(\overline{G})$ denotes the maximum degree of $\overline{G}$. We extend this by showing that $\Delta(\overline{G})$ can be replaced by $\varrho(\overline{G})$, the linear density of $\overline{G}$. We also show that this bound is close to best possible as there are graphs with $\theta_1(G)\ge c_2\frac{\Delta^2(\overline{G})}{\log\Delta(\overline{G})}\log n$.

For the parameter θ we conjecture that

$$\theta(G)\le c\Delta(\overline{G})^{1+\varepsilon}\log n$$

for some constant c not dependent on $\Delta(\overline{G})$ or n and show that $\theta(G)\le c\Delta(\overline{G})\log^2\Delta(\overline{G})\log n$ if G is bipartite. This is, up to the factor $1/\log^2\Delta(\overline{G})$ best possible.

Finally, we give an upper bound on $\theta(G)$ in terms of $\Delta(G)$ and prove $\theta(G)\le c\Delta^2(G)\log n$.

Mathematics Subject Classification (1991): 05 C
* Supported by NSF Grant No. DMS-9310064
† Supported by NSF Grant No. DMS-9401559

1. Introduction

1.1 The direct product dimension

Let $G_i=(V_i,E_i)$, $i\in[t]$ be a family of graphs. The *direct product* denoted by $\prod_{i=1}^{t} G_i$ is the graph (V,E) where $V=\prod_{i=1}^{t} V_i$ and

$$\{(x_1,x_2,\dots,x_t)(y_1,y_2,\dots,y_t)\}\in E \Leftrightarrow \{x_i,y_i\}\in E_i, \quad \forall\, i\in[t].$$

The following definitions were introduced in [12].

Definition 1.1. The *direct product dimension* of a graph G denoted $\operatorname{pdim}(G)$ is the minimum integer t such that there exists a family K_{a_1}, K_{a_2}, …, K_{a_t} of complete graphs such that G can be embedded into $\prod_{i=1}^{t} K_{a_i}$.

A related concept is that of equivalence dimension.

Definition 1.2. Let $G=(V,E)$ be a graph. A family $\mathscr{E}=\{\mathscr{E}_1,\mathscr{E}_2,\dots,\mathscr{E}_t\}$ of partitions of V is called an *equivalence representation* of G if each equivalence class in each partition is a clique and

$$\{u,v\}\in E \Leftrightarrow \exists i\in[t] \text{ s.t. } u\sim v \text{ in } \mathscr{E}_i.$$

Then we define the equivalence dimension of G as,

$$\operatorname{edim}(G)=\min_{\mathscr{E}}(|\mathscr{E}|),$$

where $\mathscr{E}$ is taken over all equivalence representations of G.

The following fact (c.f. [12]) gives the relationship between equivalence dimension and product dimension. For any graph G,

$$\operatorname{edim}(\overline{G})\le \operatorname{pdim}(G)\le \operatorname{edim}(\overline{G})+1. \tag{1}$$

We list here some of the known facts about the dimension of particular graphs. It is shown by Lovász, Nešetřil and Pultr in [10]:

1. For $n\ge 2$, $\operatorname{pdim}(K_n+K_1)=n$ where the graph K_n+K_1 is the complete graph on n vertices with an extra vertex added which is non-adjacent to all of the others,
2. $\operatorname{pdim}(nK_2)=\lceil\log_2 n\rceil+1$, where nK_m is n copies of the complete graph on m vertices,
3. for $n\ge 3$, $\operatorname{pdim}(P_n)=\lceil\log_2 n\rceil$ and
4. for $n\ge 2$, $\lceil\log_2 n\rceil+1\le\operatorname{pdim}(C_{2n+1})\le\lceil\log_2 n\rceil+2$.

In [13], Poljak and Rödl showed that

1. For q, a power of a prime, $\operatorname{pdim}(qK_q)=q$ and
2. if $q\ge 3$, $\operatorname{pdim}(nK_q)\le\frac{q\log n}{\log q}(1+o(1))$.

Both $\operatorname{edim}(G)$ and $\operatorname{pdim}(G)$ depend on the maximum degree of G. The following is an easy consequence of Vizing's Theorem.

Proposition 1.1. *If* $\Delta(G) = d$ *then*

$$\operatorname{edim}(G) \le d+1$$

and moreover, if G *is triangle-free then*

$$\operatorname{edim}(G) \ge d.$$

To prove the corresponding relation between $\operatorname{pdim}(G)$ and maximum degree is not as straightforward. N. Alon [1] proved that for any graph G on n vertices with maximal degree $\le d$,

$$\operatorname{pdim}(G) \le c(d+1)^2 \log_2 n + 1 \tag{2}$$

holds where $c = 2e^2/\log_2 e$. On the other hand, he also showed that if $\delta(G) \ge 1$,

$$\log_2 n - \log_2 d \le \operatorname{pdim}(G). \tag{3}$$

Note that the bound (2) is based on a probabilistic approach while (3) uses exterior algebra. Also, (2) is obtained by the relationship $\operatorname{pdim}(G) \le \operatorname{edim}(\overline{G}) + 1 \le \theta_1(\overline{G}) + 1$ and (4), see below.

In here we study the dependence of product dimension on maximum degree and linear edge density. Recall that linear edge density, $\varrho(G)$, is a maximum average degree of a subgraph of G, that is,

$$\varrho(G) = \max\left\{\frac{2|E(G')|}{|V(G')|} : G' \le G\right\}.$$

We replace the upper bound in (2) by a bound which is linear in d. On the other hand, we prove the existence of graphs G, $|V(G)| = n$, $\Delta(G) = d$ which have $\operatorname{pdim}(G)$ essentially larger than (3) when $d \gg \log n$. Set

$$\mathscr{B}(n,d) = \{G : V(G) = \{1,2,\ldots,n\},\ \Delta(G) \le d\}$$

and

$$\mathscr{L}(n,d) = \{G : V(G) = \{1,2,\ldots,n\},\ \varrho(G) \le d\}.$$

Theorem 1. *There exists a positive constant* c *such that*

1. *For any graph* $G \in \mathscr{L}(n,d)$, $\operatorname{pdim}(G) \le 32 d \log n$ *while*
2. *There exists* $G \in \mathscr{B}(n,d)$ *such that*

$$\operatorname{pdim}(G) \ge cd \frac{\log n}{\log d + \log\log n}.$$

1.2. Intersection number

Definition 1.3. Let $G=(V,E)$ be a graph. A family $\mathcal{F}=\{A_x : x\in V\}$ of (not necessarily distinct) sets is called an *intersection representation* of G if

$$A_x \cap A_y \neq \emptyset \Leftrightarrow \{x,y\} \in E$$

for every pair x, y of distinct vertices of G; conversely G is called an *intersection graph* of $\mathcal{F}$.

Let $G=(V,E)$ be a graph, we define

$$\theta_1(G) = \min_{\mathcal{F}}(|\cup \mathcal{F}|)$$

where $\mathcal{F}$ is taken over all intersection representations of G. Then $\theta_1(G)$ is called the *intersection number* of G. For an integer n, let

$$\theta_1(n) = \max\{\theta_1(G) : \ |V(G)| = n\}.$$

Erdős, Goodman and Pósa [4] showed that all graphs on n vertices have intersection number at most $n^2/4$ and the graph $K_{\frac{n}{2},\frac{n}{2}}$, has intersection number equal to $n^2/4$.

N. Alon [1] gave a relationship between the intersection number of a graph and its degree. For a given graph $G\in B(n,d)$

$$\theta_1(\overline{G}) \leq c(d+1)^2 \log_2 n \tag{4}$$

where $c=2e^2/\log_2 e$.

Extending this to the graphs whose complement have bounded linear density, we will show:

Theorem 2. *There exist positive absolute constants c_1, c_2 such that*

1. *For any graph G, such that $\overline{G}\in\mathcal{L}(n,d)$, $\theta_1(G)\leq c_1 d^2 \log n$ while*
2. *There exists G with $\overline{G}\in\mathcal{B}(n,d)$ such that*

$$\theta_1(G) \geq c_2 \frac{d^2}{\log d} \log\left(\frac{n}{d}\right).$$

Note, we can prove a version of Theorem 2 where we restrict the graphs to trees. Let T_n be a tree on n vertices. (All trees have edge density bounded by 2.) The proof of this special case gives

$$\theta_1(\overline{T_n}) \leq 31 \log n.$$

We consider a variant on intersection representations which is investigated in [2], [3], [6] and [9].

Definition 1.4. Let $G=(V,E)$ be a graph. A family $\mathcal{F}=\{A_x:\ x\in V\}$ of (not necessarily distinct) sets is called a *p-intersection representation* of G if

$$|A_x\cap A_y|\geq p\Leftrightarrow \{x,y\}\in E$$

for every pair x, y of distinct vertices of G. Define $\theta_p(G)$ to be

$$\theta_p(G)=\min_{\mathcal{F}}(|\cup\mathcal{F}|)$$

where $\mathcal{F}$ is taken over all p-intersection representations of G. Then $\theta_p(G)$ is called the *p-intersection number* of G.

We will set

$$\theta(G)=\min_p(\theta_p(G)).$$

Clearly, for any graph G and integer $p\geq 2$, a p-intersection representation of G exists since we know an intersection representation of G exists and given a $(p-1)$-intersection representation we may add 1 element to the universe and include it in each set. This gives $\theta_p(G)\leq\theta_{p-1}(G)+1\leq\theta_1(G)+p-1$.

There is a characterization of p-intersection representations.

Definition 1.5. We define a *p-edge cover* (also known as a p-generator or p-edge clique cover) of a graph $G=(V,E)$ to be a family $\mathcal{P}$ of V such that

$$e\in E\Leftrightarrow \exists\{P_1,P_2,\dots,P_p\}\subset\mathcal{P}\text{ s.t. }\forall i\in[p],\ e\subset P_i.$$

One can see that

$$\theta_p(G)=\min_{\mathcal{P}}(|\mathcal{P}|)$$

where the minimum is taken over all p-edge covers $\mathcal{P}$ of G.

For fixed p, the parameter θ_p was investigated in [2], [3], [6] and [9]. For example, it has been conjectured in [2] that $\theta_p(G)\leq(1+o(1))\theta_p\left(K_{\frac{n}{2},\frac{n}{2}}\right)$ for any graph G with n vertices. This is known to be true for $p=1$ by the result of Erdős, Goodman and Pósa [4]. Also, it was established in [6], [3], (see also [2]) that $\theta_p\left(K_{\frac{n}{2},\frac{n}{2}}\right)=\frac{n^2}{4p}(1+o(1))$.

For parameter θ we infer due to (4) that

$$\theta(G)\leq\theta_1(G)\leq c\Delta^2(\overline{G})\log n.$$

We believe that this can be improved and make the following conjecture:

Conjecture 1. *There exists an absolute constant c such that $\overline{G}\in\mathcal{B}(n,d)$ implies $\theta(G)\leq cd^{1+\varepsilon}\log n$, provided $n\geq n_0(\varepsilon)$.*

We were able to show it under the assumption that $\overline{G}$ is bipartite.

Theorem 3. *There exist absolute constants c_1, $c_2>0$ such that*

1. *If $\overline{G}$ is bipartite, $\overline{G}\in\mathcal{B}(n,d)$, then $\theta(G)\le c_1 d(\log d)^2\log n$.*
2. *There exists G with $\overline{G}$ bipartite, $\overline{G}\in\mathcal{B}(n,d)$, with $\frac{n}{2}$ black and $\frac{n}{2}$ white vertices, such that*

$$\theta(G)\ge c_2 d\log\left(\frac{n}{2d}\right).$$

Finally, we give an upper bound on $\theta(G)$ in terms of the maximum degree of G itself.

Theorem 4. *There exist absolute constants c_1, $c_2>0$ such that*

1. *If $G\in\mathcal{B}(n,d)$, then*

$$\theta(G)\le c_1 d^2\log n.$$

2. *There exists G, $G\in\mathcal{B}(n,d)$ such that*

$$\theta(G)\ge c_2 d\log\left(\frac{n}{2d}\right).$$

Note. As $\theta_1(G)=|E(G)|$ for triangle-free graphs, we cannot expect the same bound for $\theta_1(G)$.

2. Proof of Theorem 1

2.1. The lower bound

We start with the following proposition.

Proposition 2.1. *Let $\mathcal{B}$ be the set of all labeled bipartite graphs with maximum degree bounded by d and with $\frac{n}{2}$ white vertices $W=\{w_1,w_2,\dots,w_{\frac{n}{2}}\}$ and $\frac{n}{2}$ black vertices $B=\{b_1,b_2,\dots,b_{\frac{n}{2}}\}$. Then*

$$|\mathcal{B}|\ge\left(\frac{n}{2d}\right)^{\frac{dn}{4}}.$$

Proof. We consider graphs with the property that $\deg(w_j)=\frac{d}{2}$, while $\deg(b_i)\le d$ for each $i=1,2,\dots,\frac{n}{2}$ and $j=1,2,\dots,\frac{n}{2}$. Let

$$s=\binom{n/4}{d/2}.$$

We will be selecting successively neighborhoods of w_1, w_2, $\dots$, $w_{\frac{n}{2}}$ showing that each choice can be done in at least s ways. Indeed, suppose that we have decided on neighbors of $w_1, w_2, \ldots, w_j$, $j < \frac{n}{2}$ in such a way that $\deg(b_i) \le d$ for all $i=1,2,\ldots,\frac{n}{2}$. As at this point

$$\sum_{i=1}^{n/2} \deg(b_i) \le \frac{dj}{2},$$

there can be at most $\frac{j}{2} < \frac{n}{4}$ vertices b_i with degree equal to d. Out of the remaining (at least $\frac{n}{4}$) vertices we can choose neighbors of w_{j+1} which we can do in at least s ways. Since for each $j \le \frac{n}{2}$ there are more than s choices for neighbors of w_j, there are more than

$$\binom{n/4}{d/2}^{\frac{n}{2}} > \left(\frac{n}{2d}\right)^{\frac{dn}{4}}$$

labeled graphs of maximum degree at most d. ∎

We will show that we can restrict ourselves to graph with maximum degree d such that $d \le \sqrt{n}$. To that matter, assuming that $d \ne n-1$, consider a graph H, which is a union of a clique of size $d+1$ and one isolated vertex. By (1) and Proposition 1.1,

$$\operatorname{pdim} H \ge \operatorname{edim} \overline{H} \ge d+1.$$

Let G be a graph consisting of H and $n-2d-2$ additional isolated vertices. Then $G \in \mathscr{B}(n,d)$ and

$$\operatorname{pdim} G \ge \operatorname{pdim} H \ge d+1.$$

For $d > \sqrt{n}$ and $c \le 1/2$,

$$d+1 \ge cd\frac{\log n}{\log d + \log\log n}.$$

Thus, the lower bound follows whenever $d > \sqrt{n}$, and so, from now on we will assume that $d \le \sqrt{n}$.

Lemma 2.1. (The Grouping Lemma) *Let $1 < t < l$ and let $\{a_{i,j} : 1 \le i < j \le l\}$ be a set of nonnegative numbers, then there exists a partition*

$$L_1 \cup L_2 \cup \ldots \cup L_t = [l]$$

such that

$$\mu(L_1, L_2, \ldots, L_t) \equiv \frac{\sum_{u=1}^{t} \sum_{\{i,j\} \subset L_u} a_{i,j}}{\sum_{1 \le i < j \le l} a_{i,j}} < \frac{1}{t}.$$

Proof. We proceed by induction of l. The base is for $l=3$ and $t=2$.

$$a_{i,j} = \min\{a_{1,2}, a_{1,3}, a_{2,3}\}.$$

Letting $L_1=\{i,j\}$ and $L_2=\{1,2,3\}\setminus\{i,j\}$, we have that $\mu(L_1,L_2)\leq 1/3<1/2$.

Now suppose $l>3$. If $t<l-1$, by induction, we may assume that there exists L_1', L_2', ..., L_t' (some possibly empty) such that

$$L_1' \cup L_2' \cup \ldots \cup L_t' = [l-1]$$

and

$$\mu(L_1', L_2', \ldots, L_t') = \frac{\sum\limits_{u=1}^{t} \sum\limits_{\{i,j\}\subset L_u'} a_{i,j}}{\sum\limits_{1\leq i<j\leq l} a_{i,j}} < \frac{1}{t}. \tag{5}$$

If $t=l-1$ then (5) is easily accomplished by letting each set $L_i'=\{i\}$.

Set

$$b_u = \sum_{i\in L_u'} a_{i,l}$$

and without loss of generality assume $b_t \leq b_u$ for $1\leq u\leq t-1$, so that

$$\frac{b_t}{\sum\limits_{u=1}^{t} b_u} \leq \frac{1}{t}. \tag{6}$$

For $1\leq u\leq t-1$, set $L_u=L_u'$. Also set $L_t=L_t'\cup\{l\}$. Then

$$L_1 \cup L_2 \cup \ldots \cup L_t = [l]$$

and

$$\mu(L_1, L_2, \ldots, L_t) = \frac{\sum\limits_{u=1}^{t} \sum\limits_{\{i,j\}\subset L_u'} a_{i,j} + b_t}{\sum\limits_{1\leq i<j\leq l} a_{i,j} + \sum\limits_{u=1}^{t} b_u}.$$

And so by (5) and (6)

$$\mu(L_1, L_2, \ldots, L_t) < \frac{1}{t}.$$ ∎

We now proceed with the proof of the lower bound in Theorem 1. We will work with the equivalence dimension (see relationship (1)).

Fix k such that for every $G\in\mathcal{B}(n,d)$, $\operatorname{edim}(\overline{G})\leq k$. Set $t=10k$.

Definition 2.1. Let an equivalence relation be called *fine* if the number of equivalence classes is greater than t and *coarse* if the number of equivalence classes is at most t. If each class of a partition $\mathcal{F}$ is contained as a subset of some class of a partition $\mathcal{C}$, we set $\mathcal{F}<\mathcal{C}$.

Fix $G\in\mathscr{B}(n,d)$. Suppose $\{\mathscr{F}_1,\mathscr{F}_2,\dots,\mathscr{F}_p,\mathscr{C}_{p+1},\dots,\mathscr{C}_k\}$ is a family of equivalence relations representing $\overline{G}$. That is, each class of each equivalence relation is a clique in $\overline{G}$ and every edge in $\overline{G}$ is in some class. Suppose further that $\mathscr{F}_1$, $\mathscr{F}_2$, …, F_p are fine and $\mathscr{C}_{p+1}$, …, $\mathscr{C}_k$ are coarse.

Let $s\in\{1,2,\dots,p\}$ and consider $\mathscr{F}_s$. Suppose the number of classes in $\mathscr{F}_s$ is l and $a_{i,j}$ is the number of edges of G between the i^{th} and j^{th} class of $\mathscr{F}_s$. By the Grouping Lemma, there exists an equivalence relation $L_1\cup L_2\cup\ldots\cup L_t=[l]$ such that

$$\sum_u \sum_{\{i,j\}\subset L_u} a_{i,j} \le \frac{\sum\limits_{1\le i<j\le l} a_{i,j}}{t} \le \frac{dn}{2t}.$$

Let $\mathscr{C}_s$ be the equivalence relation obtained from $\mathscr{F}_s$ by grouping classes together according to the relation $\{L_1,L_2,\dots,L_t\}$.

Note, that $\mathscr{F}_s<\mathscr{C}_s$ and the classes in $\mathscr{C}_s$ are not necessarily independent sets in G, but altogether, at most $dn/2t$ edges are contained within the classes of $\mathscr{C}_s$.

Consider the following observations.

1. In the set of equivalence $\mathscr{C}_1$, …$\mathscr{C}_k$, there are at most $k\frac{dn}{2t}\le\frac{dn}{20}$ edges of G which are contained in some equivalence class.
2. All other edges of G are in no equivalence class.
3. All non-edges of G are covered by some equivalence class.

This list of observations implies the following fact.

Fact 2.1. *There exists a subgraph $H\subset G$ with $V(H)=V(G)$ and*

$$|E(G)-E(H)|=x\le\frac{dn}{20}$$

with edim$(\overline{H})\le k$ *which can be realized by k coarse equivalences.* ∎

For each $x\le dn/20$, let $\mathscr{G}_x\subset\mathscr{B}(n,d)$ be such that $\forall\ G\in\mathscr{G}_x$, Fact 2.1 is satisfied. By averaging there exists an $x\le dn/20$ such that

$$|\mathscr{G}_x|\ge\frac{1}{\frac{dn}{20}}|\mathscr{B}(n,d)|.$$

We fix a value x which satisfies the above and define a class $\mathscr{H}$.

Definition 2.2. Let $\mathscr{H}$ be the set of all graphs H on n vertices such that

1. there exists a graph $G\in\mathscr{G}_x$ such that $H\le G$ and $|E(G)-E(H)|=x$ and
2. There exists a set of k coarse equivalences $\mathscr{C}_1$, $\mathscr{C}_2$, …, $\mathscr{C}_k$ which represent H.

Since there are at most $\binom{\binom{n}{2}}{x}$ ways to extend $H \in \mathcal{H}$ to a graph G with x more edges than H and each $G \in \mathcal{G}_x$ contains such an H, we underestimate the size of $\mathcal{H}$ as follows:

$$|\mathcal{H}|\binom{\binom{n}{2}}{x} \geq |\mathcal{G}_x| \geq \frac{1}{\frac{dn}{20}}|\mathcal{B}(n,d)|,$$

$$|\mathcal{H}| \geq \frac{1}{\frac{dn}{20}}|\mathcal{B}(n,d)|\frac{1}{\binom{\binom{n}{2}}{x}}. \tag{7}$$

To underestimate the size of $\mathcal{B}(n,d)$ we reason that it is at least as big as the class $\mathcal{B}$ of all labeled bipartite graphs with maximum degree bounded by d and with $n/2$ white and $n/2$ black vertices. Thus, from Proposition 2.1,

$$|\mathcal{B}(n,d)| \geq \left(\frac{n}{2d}\right)^{\frac{dn}{4}}. \tag{8}$$

From (7) and (8) we now infer

$$|\mathcal{H}| \geq \frac{20}{dn}\left(\frac{n}{2d}\right)^{\frac{dn}{4}} \frac{1}{\binom{\binom{n}{2}}{\frac{dn}{20}}} \geq \frac{20}{dn}\left(\frac{n}{2d}\right)^{\frac{dn}{4}} \frac{1}{\left(\frac{10en}{d}\right)^{\frac{dn}{20}}} \geq \frac{20}{dn}\left(\frac{n}{6d}\right)^{\frac{4d}{5}}.$$

The number of sets $\{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_k\}$ where each $\mathcal{C}_i$ is a coarse equivalence is at most $\binom{(10k)^n}{k} < (10k)^{nk}$ and thus

$$(10k)^{nk} \geq |\mathcal{H}| \geq \frac{20}{dn}\left(\frac{n}{6d}\right)^{\frac{dn}{5}}$$

and hence there exists a constant c_1 such that

$$k\log(10k) \geq c_1 d \log\left(\frac{n}{6d}\right)$$

which implies there exists a constant c_2 such that

$$k \geq \frac{d\log\left(\frac{n}{6d}\right)}{\log d + \log\log\left(\frac{n}{6d}\right)}.$$

However, $d \leq \sqrt{n}$, and thus for c sufficiently small,

$$c_2\frac{d\log\left(\frac{n}{6d}\right)}{\log d + \log\log\left(\frac{n}{6d}\right)} \geq c\frac{d\log n}{\log d + \log\log n}. \qquad \blacksquare$$

2.2. The upper bound

The proof of the upper bound depends on a lemma which appeared in [14]. We present it here for the sake of thoroughness. Note that in an oriented graph, the notation $d_+(x)$ stands for the out-degree of a vertex x.

lemma 2.2. *Every graph $G=(V,E)$ admits an orientation $\vec{E}$ of the edges with*

$$d_+(G) \le \varrho(G).$$

Proof. This is proved by induction on $n=|V|$. If $n-1$, the result is trivial. When $n>1$, we select a vertex $v \in V$ such that $d(v) \le \varrho$. This can be done since the average degree in G itself must be at most ϱ. Orient the edges of $G\setminus\{v\}$ using the induction assumption and orient the edges $\{v,u\}$ so that $(u,v\} \in \vec{E}$. ∎

Proof. (Upper bound) We will use the fact that $\mathrm{pdim}(G) \le \mathrm{edim}(\overline{G})+1$ (see (1)), and prove that there are $k = cd\log n$ equivalences the union of which is the edge set of $\overline{G}$. This will be achieved by selecting k random equivalences $\mathcal{P}_1, \ldots, \mathcal{P}_k$, chosen from the space of equivalences with at most d classes, then, changing them according to the rule described below, to obtain equivalences $\mathcal{P}'_1, \mathcal{P}'_2, \ldots, \mathcal{P}'_k$ such that $E(\overline{G}) = \bigcup_{i=1}^{k} E(\mathcal{P}'_i)$ holds with positive probability.

As $\varrho(G) \le d$, there exists an orientation of edges of G, so that $\deg_+(v) \le d$ for any $v \in V(G)$. Let $\mathcal{P} = \{X_1, X_2, \ldots, X_d\}$ be an equivalence on $V(G)$.

For each $i = 1, 2, \ldots, d$ we select an independent set (of G), $X'_i \subset X_i$ by excluding each vertex which is the tail of some directed edge both endpoints of which are in X_i. Let the new equivalence relation $\mathcal{P}'$ consist of the equivalence classes $X'_1, X'_2, \ldots, X'_d$ together with the remaining singleton sets.

Consider the sample space of ordered partitions (equivalences) of V with at most d parts, where the partitions $\mathcal{P} = \{X_1, X_2, \ldots, X_d\}$ are chosen uniformly as follows. For each vertex $v \in V$ and each i, let $\mathrm{Prob}(v \in X_i) = 1/d$.

Now given $\{u,v\} \in E(\overline{G})$ and a partition $\mathcal{P} = \{X_1 \cup X_2 \cup \ldots \cup X_d\}$ define $B^{\mathcal{P}}_{u,v}$ to be the event

$$B^{\mathcal{P}}_{u,v} \equiv \exists X_i \ \{u,v\} \subset X_i \quad \text{and} \quad (N_+(u) \cup N_+(v)) \cap X_i = \emptyset \tag{9}$$

where $N_+(u)$ is the set of out neighbors of u. Observe that (9) is equivalent to the event

$$\exists i, \quad \{u,v\} \subset X'_i. \tag{10}$$

We select $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_k$ at random. Let E be the event defined by

$$E \equiv \forall \{u,v\} \in E(\overline{G}), \quad \exists \mathcal{P}_j, \quad \text{s.t.} \quad \mathcal{B}^{\mathcal{P}_j}_{u,v}.$$

We are going to show that Prob $(E) > 0$; this by (10) implies that $E(\overline{G}) = \bigcup_{i=1}^{k} E(\mathcal{P}'_i)$, which by (1) concludes the proof.

We will estimate the negation of E,

$$\overline{E} \equiv \exists\{u,v\} \in E(\overline{G}) \text{ s.t. } \forall \mathcal{P}_j \overline{\mathcal{B}^{\mathcal{P}_j}_{u,v}}.$$

First we consider Prob $(\mathcal{B}^{\mathcal{P}_i}_{u,v})$. The vertex u must be in one of the d classes and the probability that v is in that same class is $1/d$. The size of $N_+(u) \cup N_+(v)$ is at most $2d$ so that the probability that the set $N_+(u) \cup N_+(v)$ is not in that same class is at least $(1-1/d)^{2d}$. Thus,

$$\text{Prob } (\mathcal{B}^{\mathcal{P}_j}_{u,v}) \geq \frac{1}{d}\left(1-\frac{1}{d}\right)^{2d}.$$

Since $d \geq 2$

$$\left(\frac{1}{d}\left(1-\frac{1}{d}\right)^{2d}\right) \geq \frac{1}{16d}.$$

Thus,

$$\text{Prob } \overline{\mathcal{B}^{\mathcal{P}_j}_{u,v}} \leq 1 - \frac{1}{16d}$$

and

$$\text{Prob } (\overline{E}) \leq \sum_{u,v} \text{Prob } \left(\bigwedge_{j=1}^{k} \overline{\mathcal{B}^{\mathcal{P}_j}_{u,v}}\right) \leq \sum_{u,v} \prod_{j=1}^{k} \text{Prob } \left(\overline{\mathcal{B}^{\mathcal{P}_j}_{u,v}}\right)$$
$$\leq \binom{n}{2}\left(1-\frac{1}{16d}\right)^{k} \leq \binom{n}{2}\left(1-\frac{1}{16d}\right)^{32d\log n} < 1. \qquad \blacksquare$$

3. Proof of Theorem 2

3.1. Proof of the lower bound

We make the following assumptions:

$$60 \leq d = 2r \leq \frac{n}{2}, \tag{11}$$

r is even and $\frac{n}{d}$ is an integer. This is sufficient as for $d < 60$ our theorem ensures just the existence of a graph whose complement has minimum degree bounded by 60 and with θ_1 at least $c\log n$ where c is an absolute constant. This however follows easily as any graph G with distinct neighborhoods must be represented by distinct sets and hence $\theta_1(\overline{G}) \geq \log_2 n$ for any such graph G. Similarly, assuming that $d \leq n/2$ is no restriction since $\mathscr{B}(n,d) \subset \mathscr{B}(n,d')$ for $d < d'$ and thus having proved the bound under this restriction we may satisfy it for $d > n/2$ with the choice of small c_2.

Set $l = \frac{n}{d}$, and let $G_1 = 2lK_r$ be the graph which is the vertex disjoint union of $2l$ cliques $K_r^{(1)}, K_r^{(2)}, \ldots, K_r^{(2l)}$. Set $W = \bigcup_{i=1}^{l} V(K_r^{(i)})$ and $B = \bigcup_{i=l+1}^{2l} V(K_r^{(i)})$. For two graphs H and H' on the same vertex set V, let the graph $H \cup H'$ be the graph with vertex set, $V(H \cup H') = V$ and $E(H \cup H') = E(H) \cup E(H')$.

Consider the set $\mathscr{B}$ described in Proposition 2.1. We will show that there exists a graph $G_2 \in \mathscr{B}$ with

$$\theta_1(\overline{G_1 \cup G_2}) \geq c\frac{d^2}{\log d}\log n$$

for some absolute constant c.

Suppose $\theta_1(\overline{G_1 \cup G}) \leq t$ for each $G \in \mathscr{B}$. Given any $G \in \mathscr{B}$ consider a representation of $\overline{G_1 \cup G}$

$$\varphi = \varphi_G \,:\, W \cup B \to 2^{|T|}, \qquad |T| = t.$$

For $j = 1, 2, \ldots, 2l$, let $V(K_r^{(j)}) = \{u_1^j, u_2^j, \ldots, u_r^j\}$ and for $i = 1, 2, \ldots, r$, let $T_i^j = \varphi(u_i^j)$.

The sets $T_1^j, T_2^j, \ldots, T_r^j$ are pairwised disjoint subsets of the t-element set T and hence there are at least $\frac{19}{20}r$ of them with cardinality at most $\frac{20t}{r}$. Consider the set $W' \subset W$ $(B' \subset B)$ of all vertices $w \in W$ $(b \in B)$ such that the cardinality of $\varphi(w)$ $(\varphi(b)$ respectively) is at most $\frac{20t}{r}$. We have $|W'| \geq \frac{19}{40}n \equiv m$, $|B'| \geq m$ and if strict inequality holds, we choose a proper subset satisfying $|W'| = |B'| = m$.

For each $G \in \mathscr{B}$ (G_1 is fixed throughout the proof) we obtain vertex sets $W' \subset W$ and $B' \subset B$ with $|W'| = |B'| = m = \frac{19}{40}n$ such that each vertex in $W' \cup B'$ is represented by a set of cardinality at most $\frac{20t}{r}$. Hence there exist m-elements sets $W_0' \subset W$ and $B_0' \subset B$ which were obtained for at least

$$\frac{|\mathscr{B}|}{\binom{\frac{n}{2}}{m}^2}$$

different subgraphs from $\mathscr{B}$. Let $\mathscr{B}'$ be the set of all such graphs. Denote by ψ the mapping which for each graph of $G \in \mathscr{B}'$ assigns the subgraph G_0 induced on $W_0' \cup B_0'$. We will estimate the number z of distinct graphs in $\psi[\mathscr{B}']$.

Given a graph G_0 with vertex set $W_0' \cup B_0'$ we will overestimate the number of distinct $G \in \mathscr{B}'$ such that $G_0 = \psi(G)$. Since $|W \setminus W_0'| = |B \setminus B_0'| = \frac{n}{40}$ and we are free to pick any neighborhood of size r from the set W for those vertices in $B \setminus B_0'$ and from the set B for those vertices in $W \setminus W_0'$, we get that the number of graphs of the form $\psi^{-1}(G)$ does not exceed $\left[\binom{\frac{n}{2}}{r}^{\frac{n}{40}}\right]^2$.

Thus by Proposition 2.1,

$$z\binom{\frac{n}{2}}{r}^{\frac{n}{20}} \geq \frac{|\mathscr{B}|}{\binom{\frac{n}{2}}{m}^2} \geq \frac{\left(\frac{n}{2r}\right)^{\frac{rn}{4}}}{\binom{\frac{n}{2}}{m}^2}.$$

Also,

$$\binom{\frac{n}{2}}{r}^{\frac{n}{20}} \leq \left(\frac{en}{2r}\right)^{\frac{rn}{20}}$$

and

$$\binom{\frac{n}{2}}{m}^2 \leq 2^n.$$

So,

$$z \geq \frac{\left(\frac{n}{2r}\right)^{\frac{rn}{4}}}{2^n \left(\frac{n}{2r}\right)^{\frac{rn}{20}} e^{\frac{rn}{20}}} = \frac{\left(\frac{n}{2r}\right)^{\frac{rn}{5}}}{2^n e^{\frac{rn}{20}}}.$$

Consider representations of the z distinct graphs in $\psi[\mathscr{B}']$. Each vertex of each graph in $\psi[\mathscr{B}']$ is represented by a set of size between 1 and $\frac{20t}{r}$. Therefore, there are at most $\left[\sum_{j=1}^{\frac{20t}{r}} \binom{t}{j}\right]^{2m}$ different representations that can be achieved for such graphs. Some of these representations may give the same graph, but different graphs must have different representations.

Hence,

$$\left[\sum_{j=1}^{\frac{20t}{r}} \binom{t}{j}\right]^{2m} \geq z \geq \frac{\left(\frac{n}{2r}\right)^{\frac{rn}{5}}}{2^n e^{\frac{rn}{20}}}.$$

Using that $m=(19n)/40$, we have

$$\left[2\left(\frac{er}{20}\right)^{\frac{20t}{r}}\right]^{\frac{19n}{20}} \geq \frac{\left(\frac{n}{2r}\right)^{\frac{rn}{5}}}{2^n e^{\frac{rn}{20}}}.$$

which gives

$$\left(\frac{er}{20}\right)^{\frac{10tn}{r}} \geq \left(\frac{n}{2r}\right)^{\frac{rn}{5}} / \left(2^r e^{\frac{rn}{20}}\right),$$

$$\frac{19t}{r}\log\left(\frac{er}{20}\right) \geq \frac{r}{5}\log\left(\frac{n}{2r}\right) - \frac{r}{20} - \log 4,$$

$$t \geq \frac{r^2}{95\log\left(\frac{er}{20}\right)}\log\left(\frac{n}{2r}\right) - \frac{r^2}{380\log\left(\frac{er}{20}\right)} - \frac{r\log 4}{19\log\left(\frac{er}{20}\right)}.$$

As for any choice of n and r satisfying (11) one can show that

$$\frac{r^2}{380\log\left(\frac{er}{20}\right)} + \frac{r\log 4}{19\log\left(\frac{er}{20}\right)} \leq \frac{3}{4}\left[\frac{r^2}{95\log\left(\frac{er}{20}\right)}\log\left(\frac{n}{2r}\right)\right]$$

holds, we infer that

$$t \geq \frac{1}{4}\left[\frac{r^2}{95\log\left(\frac{er}{20}\right)}\log\left(\frac{n}{2r}\right)\right] \geq \frac{1}{380}\frac{r^2}{\log r}\log\left(\frac{n}{2r}\right).$$

Hence, (in view of (11)) there exists a graph $G_2 \in \mathcal{B}$ such that

$$\theta_1(\overline{G_1 \cup G_2}) \geq \frac{1}{380}\frac{r^2}{\log r}\log\left(\frac{n}{2r}\right) \geq \frac{1}{1520}\frac{d^2}{\log d}\log\left(\frac{n}{d}\right). \qquad \blacksquare$$

3.2. Proof of the upper bound

Set

$$t = 2e^2(1+d^2)\log n.$$

Apply Lemma 2.2 to orient the edges of G thus obtaining $\vec{E}$ with $d_+(v) \leq d$, $\forall v \in V(G)$.

Given an arbitrary family of sets, one for each vertex of G, $\mathcal{F} = \{F_x : x \in V(G)\}$, define

$$\mathcal{S}(\mathcal{F}) = \left\{ S_x = F_x \bigcap_{(x,u)\in\vec{E}} \overline{F}_u \, : \, x \in V(G) \right\}.$$

We will argue that there exists a family $\mathcal{F}$ of subsets of $[t]$ such that $\mathcal{S}(\mathcal{F})$ is a 1-intersection representation of $\overline{G}$.

If $\{x,y\} \in E$ then either $(x,y) \in \vec{E}$ or $(y,x) \in \vec{E}$. In either case, no matter what $\mathcal{F}$ is, $S_x \cap S_y = \emptyset$. On the other hand, we will show that for an appropriate choice of $\mathcal{F}$, the family $S(\mathcal{F}) = \{S_x : x \in V(G)\}$ satisfies $S_x \cap S_y \neq \emptyset$ whenever $\{x,y\} \notin E$.

Set

$$p = \frac{1}{1+d}$$

and consider a random family

$$\mathcal{F} = \{F_x : x \in V(G)\}$$

of subsets of $[t]$ where for all $a \in [t]$, and for all $x \in V(G)$, $\text{Prob}(a \in F_x) = p$.

Let A be the event that for all non-edge pairs, $\{x,y\}$, $|S_x \cap S_y| \geq 1$. We must show that the probability of event A is larger than zero. Actually we will show that

$$\text{Prob}\ (\overline{A}) < 1.$$

To do so, we wish to find an upper bound for $\text{Prob}(|S_x \cap S_y| = 0)$ for any arbitrary non-edge pair $\{x,y\}$.

We have

$$S_x \cap S_y = F_x \cap F_y \cap \bigcap_{x \in N_+(x) \cup N_+(y)} \overline{F}_z$$

where $N_+(v)$ is the set of out-neighbors of the vertex v and thus if $\{x,y\}$ is not an edge we have

$$\text{Prob}\ (a \in S_x \cap S_y) \geq p^2(1-p)^{2d}$$

(with this lower bound attained when $d_+(x) = d_+(y) = d$ and $N_+(x) \cap N_+(y) = \emptyset$.).

Hence,

$$\begin{aligned}\text{Prob}\ (|S_x \cap S_y| = 0) &= \prod_{a \in [t]} \text{Prob}\ (a \notin S_x \cap S_y) \\ &\leq \left(1 - p^2(1-p)^{2d}\right)^t \leq \left(1 - \left(\frac{p}{e}\right)^2\right)^t.\end{aligned}$$

Given that there are no more than $\binom{n}{2}$ non-edge pairs in G, we have that

$$\text{Prob}\ (\overline{A}) \leq \binom{n}{2}\left(1 - \left(\frac{p}{e}\right)^2\right)^t < 1. \qquad \blacksquare$$

4 Proof of Theorem 3

4.1. Lower bound

By Proposition 2.1, the number of labeled bipartite graphs with maximum degree d and $\frac{n}{2}$ black and $\frac{n}{2}$ white vertices is at least

$$\left(\frac{n}{2d}\right)^{\frac{nd}{4}}.$$

On the other hand, we determine the number of different p-intersection representations that can be achieved from a universe set T of size t. There are 2^{tn} sequences $(T_1,T_2,\dots,T_n)$, $T_i \subset T$ and given p, $1 \le p \le t$ each such sequence determines a labeled graph, G with vertex set $(v_1,v_2,\dots,v_n)$.

Suppose now that t is large enough to represent all bipartite graphs considered above. Then there is a mapping from the set of $n{+}1$-tuples $(T_1,T_2,\dots,T_n,p)$, where $T_i \subset T$, $i=1, 2, \dots, n$, $1 \le p \le t$ onto the set of all labeled bipartite graphs on n vertices with maximum degree d.

Hence,

$$t2^{tn} \ge \left(\frac{n}{2d}\right)^{\frac{nd}{4}} \tag{12}$$

which means

$$t + o(1) \ge \frac{d}{4} \log_2 \left(\frac{n}{2d}\right). \tag{13}$$

∎

4.2. Upper bound

4.2.1. Preliminaries

The following lemma is due to Hoeffding [8], *c.f.* [11].

Lemma 4.1. *Let* $\mathbf{X}$ *be a random variable having the binomial distribution,* $\mathscr{B}(n,p)$, *then for* $0<p\le 1/2$ *and* $\varepsilon \le 2/3$,

$$\text{Prob}\ (|\mathbf{X} - np| \ge \varepsilon np) \le 2\exp\left(-\varepsilon^2(1-\varepsilon)\frac{np}{2}\right). \tag{14}$$

Thus, for $p=1/2$ *and* $\varepsilon \le 1/2$,

$$\text{Prob}\ \left[|\mathbf{X} - np| \ge \frac{\varepsilon n}{2}\right] \le 2\exp\left(-\frac{1}{2}\left(\frac{\varepsilon n}{2}\right)^2\right). \tag{15}$$

Next we state the Lovász Local Lemma [5].

Theorem 5. *Let* $G=(V,E)$ *be a graph with maximum degree* d *and vertices* v_1, v_2, $\dots$, v_n. *For each* v_i *let us associate an event* A_i *and suppose that* A_i *is independent of the set*

$$\{A_j : \{v_i,v_j\} \notin E\}.$$

Also suppose

$$\text{Prob}\ (A_i) \le \frac{1}{4d}.$$

Then

$$\text{Prob}\ (\overline{A}_1 \wedge \overline{A}_2 \wedge \dots \wedge \overline{A}_n) > 0.$$

Next, two lemmas are given. The first lemma is a slight modification of a lemma of Füredi and Kahn, [7].

Lemma 4.2. *Let $\mathcal{H} = (X,\mathcal{E})$ be a hypergraph, $\mathcal{E} \subset [X]^{\leq d}$ with $\Delta(\mathcal{H}) \leq d$, where $d \geq 3$. Set*

$$s = \left\lceil \frac{d}{\log d} \right\rceil \quad \text{and} \quad k = \lceil e^2 \log d \rceil.$$

Then there is a partition of X,

$$X = X_1 \cup X_2 \cup \ldots \cup X_s$$

such that $\forall\ E \in \mathcal{E}$ and $i \in [s]$,

$$|E \cap X_i| \leq k.$$

Proof. Let $X = \{X_1, X_2, \ldots, X_s\}$ be a random ordered partition of the set X, such that $\text{Prob}\,(v \in X_i) = 1/s$, for all $i \in [s]$. Consider the probability space of all such partitions.

Let $E \in \mathcal{E}$ and let $\mathbf{Y}_i$ be the random variable which indicates the size of the intersection of X_i with E. That is, $\mathbf{Y}_i = |X_i \cap E|$. Then $\mathbf{Y}_i$ is a random variable with binomial distribution, $\mathcal{B}\left(|E|, \frac{1}{s}\right)$.

For $a = e^2 \log d$, we have

$$\text{Prob}\,(\mathbf{Y}_i) \leq \left(\frac{1}{s}\right)^a \binom{|E|}{a} < \left(\frac{|E|\frac{1}{s}e}{e^2 \log d}\right)^{e^2 \log d} < e^{-e^2 \log d}.$$

For each edge $E \in \mathcal{E}$, let A_E be the event that there exists an i such that $|X_i \cap E| \geq e^2 \log d$, then

$$\text{Prob}\,(A_E) \leq s e^{-e^2 \log d} = \left\lceil \frac{d}{\log d} \right\rceil e^{-e^2 \log d}. \tag{16}$$

We have

$$3 - \frac{\log \log d}{\log d} + \frac{3}{\log d} \leq e^2$$

which implies

$$3 \log d - \log \log d \leq e^2 \log d - 3$$

which is equivalent to

$$\frac{d^3}{\log d} \leq e^{e^2 \log d - 3}.$$

Hence, in view of (16)

$$\text{Prob}\,(A_E) \leq \left\lceil \frac{d}{\log d} \right\rceil e^{-e^2 \log d} \leq \frac{d}{\log d} e^{-e^2 \log d} \cdot e \leq \frac{1}{e^2 d^2} < \frac{1}{4d^2}.$$

Let G be the dependency graph described in the Lovász Local Lemma, associated with the events $\{A_E : E \in \mathscr{E}\}$, Then, for edges E and F such that $E \neq F$, the vertex corresponding to A_E is adjacent to the vertex corresponding to A_F if and only if $E \cap F \neq \emptyset$. The maximum degree in G is at most d^2 since $\Delta(\mathscr{H}) \leq d$, and $|E| \leq d$.

By the Lovász Local Lemma,

$$\text{Prob}\left(\underset{E \in \mathscr{E}}{\wedge} \overline{A}_e\right) > 0.$$

Thus, there exists a partition such that for all edges, E, and for all classes, X_i,

$$|X_i \cap E| < e^2 \log d. \quad \blacksquare$$

The next lemma provides a special set of partitions for a given set.

Lemma 4.3. *Let X be a set, $|X| = m$, $\varepsilon > 0$ and $M = \frac{2}{\varepsilon^2} \log m$. There exist M bipartitions of X with the following property: For each $\{x, y\} \in |X|^2$, if $t_{x,y}$ is the number of partitions with $\{x, y\}$ in different parts, then*

$$\frac{1}{2}M - \varepsilon M \leq t_{x,y} \leq \frac{1}{2}M + \varepsilon M.$$

Proof. Instead of proving this lemma directly we form an equivalent statement.

There exist m vectors, $\{\mathbf{a}_v : v \in X\}$, in $\{-1, 1\}^M$ such that for all $v \neq v'$,

$$|\langle \mathbf{a}_v, \mathbf{a}_{v'} \rangle| \leq 2\varepsilon M.$$

Consider the probability space of all vectors, $\{-1, 1\}^M$, such that for $\mathbf{u} \in \{-1, 1\}^m$, $\mathbf{u} = (u_1, u_2, \ldots, u_M)$ and $\forall i$,

$$\text{Prob}\,(u_i = -1) = \text{Prob}\,(u_i = 1) = 1/2.$$

The distribution of $|\langle \mathbf{u}, \mathbf{v} \rangle|$ is the same as the distribution of $\mathscr{B}(M, p) - pM$ with $p = 1/2$ so by Lemma 4.1, (15)

$$\text{Prob}\,(|\langle \mathbf{u}, \mathbf{v} \rangle| \geq 2\varepsilon M) \leq 2\exp(-2\varepsilon^2 M). \tag{17}$$

Randomly choose m vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ (with possible repetition) from the cube $\{-1, 1\}^M$. For each i, define the event A_i as

$$A_i \equiv \exists\, j \neq i,\ \left|\langle \mathbf{u}_i, \mathbf{u}_j \rangle\right| \geq 2\varepsilon M.$$

By (17), we infer that

$$\text{Prob}\,(A_i) < (m - 1) 2\exp(-2\varepsilon M)$$

and hence

$$\text{Prob } (\exists i, A_i) \leq \sum_{i=1}^{m} \text{Prob } (A_i) < 2m(m-1)\exp(-2\varepsilon^2 M) < 1. \qquad \blacksquare$$

4.2.2. Proof of the upper bound

Let G be a graph such that $\overline{G}=(V_1 \cup V_2, E)$ is a bipartite graph, $|V_1|+|V_2|=n$ and $\Delta(\overline{G}) \leq d$. Set

$$p = k\left(\frac{1}{2} - \varepsilon\right) M$$

where

$$\varepsilon = \frac{1}{4e^2 \log d} \quad \text{and} \quad M = \frac{2}{\varepsilon^2} \log n.$$

Our strategy will be to obtain a covering $\mathscr{C}$ of the vertices of $\overline{G}$ such that $|\mathscr{C}| \leq cd \log n$ and for p as above, all non-edges are covered at least p times and all edges are covered at most $p-1$ times. This will provide the upper bound, (c.f. Definition 1.5)

$$\theta(G) \leq \theta_p(G) \leq cd \log^2 d \log n.$$

Consider the hypergraph, $\mathscr{H} = (V_2, \{N(x) : x \in V_i\})$. The edges in $\mathscr{H}$ have maximum order d since $\forall x \in V_1$, $|N(x)| \leq d$, and $\forall x \in V_2$, $N_{\mathscr{H}}(x) \leq d$. Thus by Lemma 4.2, there exists a partition of V_2,

$$V_2 = X_1 \cup X_2 \cup \ldots \cup X_s$$

with

$$s = \left\lceil \frac{d}{\log d} \right\rceil$$

such that

$$\forall x \in V_1,\ \forall j \in [s], \quad |N(x) \cap X_j| \leq k$$

where

$$k = e^2 \log d.$$

We define s subgraphs of $\overline{G}$, $A_1, A_2, \ldots A_s$, to be the induced subgraphs,

$$A_i = \langle V_i \cup X_i \rangle.$$

Now, for each $i \in [s]$ we form the k subgraphs, $B_{i,1}, B_{i,2}, \ldots, B_{i,k}$ of the graph A_i so that $E(A_i) = E(B_{i,1}) \cup E(B_{i,2}) \cup \ldots \cup E(B_{i,k})$.

To describe these graphs we start with an arbitrary order on the vertices in X_i (and hence for each $v \in V_1$, its neighborhood in X_i is ordered). Let

$$N(v) = \{v_1, v_2, \ldots, v_{t_v}\}$$

where $v_1 < v_2 < \ldots < v_{t_k}$ and $t_v \leq k$.

Consider k new dummy vertices, $u_1, u_2, \ldots, u_k$. For each $j \in [k]$, let the set

$$Y_{i,j} = X_i \cup \{u_j\}.$$

For $i = 1, 2, \ldots, s$, $j = 1, 2, \ldots, k$ define the graph $B_{i,j}$ as follows:

$$V(B_{i,j}) = V_1 \cup Y_{i,j},$$

and

$$E(B_{i,j}) = \{\{v, x_j\} \ : \ v \in V_1\}$$

where $x_j = v_j$ if $j \leq t_v$ and $x_j = u_j$, the dummy variable if $j > t_v$.

Note that $B_{i,j}$ is a bipartite graph such that every vertex in V_1 has exactly one neighbor in $Y_{i,j}$, namely, its j^{th} neighbor or the dummy vertex u_j. Thus, each graph $B_{i,j}$ is a star forest.

Fix $i \in [s]$ and $j \in [k]$. We proceed to use Lemma 4.3 to show the existence of a cover, $\mathscr{C}_{i,j}$, of the graph $B_{i,j}$ which will then be used in the construction of the p-edge cover $\mathscr{C}$ of G.

We note that $|Y_{i,j}| \leq n$ and $M = \frac{2}{\varepsilon^2} logn$ and so by Lemma 4.3 there exist M bi-partitions,

$$\mathscr{P}_1, \mathscr{P}_2, \ldots, \mathscr{P}_M$$

of $Y_{i,j}$ with the following property: For each $\{x, y\} \in [Y_{i,j}]^2$, if $t_{x,y}$ is the number of partitions with x and y in different parts, then

$$\frac{1}{2}M - \varepsilon M \leq t_{x,y} \leq \frac{1}{2}M + \varepsilon M.$$

For each $l \in \{1, 2, \ldots, M\}$, let $\mathscr{P}_l = \mathscr{P}_{l,1} \cup \mathscr{P}_{l,2}$. Now we are ready to define the cover $\mathscr{C}_{i,j}$ and eventually the desired cover $\mathscr{C}$.

For each $l = 1, 2, \ldots, M$ and partition $\mathscr{P}_l = \mathscr{P}_{l,1} \cup \mathscr{P}_{l,2}$ let

$$C_{l,1} = P_{l,1} \cup \{v \in V_1 \ : \ N_{B_{i,j}}(v) \in P_{l,2}\}$$

and

$$C_{l,2} = P_{l,2} \cup \{v \in V_1 \ : \ N_{B_{i,j}}(v) \in P_{l,1}\}.$$

We set

$$\mathscr{C}_{i,j} = \bigcup_{l=1}^{M} \{C_{l,1}, C_{l,2}\}.$$

The cover $\mathscr{C}$ for the entire graph G will consist of the coverings $\mathscr{C}_{i,j}$ together with p copies of V_1 and p copies of V_2. That is, if we define the sets $C_1, C_2, \ldots, C_{2p}$ by $C_1 = C_2 = \ldots = C_p = V_1$ and $C_{p+1} = C_{p+2} = \ldots = C_{2p} = V_2$ then,

$$\mathscr{C} = \bigcup_{i,j} \mathscr{C}_{i,j} \cup \{C_1, C_2, \ldots, C_{2p}\}.$$

The size of this cover is

$$2skM + 2p = 2\left(s + \frac{1}{2} - \varepsilon\right) kM < 3skM = 96e^6 d \log^2 d \log(n).$$

It remains only to verify that this cover works. That is, for any $\{x,y\} \in E(\overline{G})$ the number of sets in the cover, $\mathscr{C}$, containing both x and y is less than p and for any non-edge pair, $\{x,y\}$, the number of sets in $\mathscr{C}$ containing both x and y is at least p.

For any pair of vertices, $\{x,y\}$, let $c_{x,y}$ indicate the number of sets in the cover $\mathscr{C}$ which contain the pair $\{x,y\}$.

Suppose $\{x,y\}$ is an edge in G, say $x \in V_1$ and $y \in V_2$. The vertex y is in exactly one of the parts from the partition $V_2 = X_1 \cup X_2 \cup \ldots \cup X_s$, say $y \in X_i$, and so the pair $\{x,y\}$ of vertices is contained in the vertex set of exactly one of the induced subgraphs $\{A_i : i \in [s]\}$ of G and so the edge $\{x,y\}$ could only be contained in sets from the covers $\mathscr{C}_{i,1}$, $\mathscr{C}_{1,2}$, $\ldots$, $\mathscr{C}_{i,k}$. The pair of vertices $\{x,y\}$ is contained in the vertex set of each of the graphs, $B_{i,1}$, $B_{i,2}$, $\ldots$, $B_{i,k}$, but forms an edge in exactly one. Say, $\{x,y\} \in E(B_{ij,})$. By the definition of the cover $\mathscr{C}_{i,j}$ of $B_{i,j}$ we see that the number of sets in $\mathscr{C}_{i,j}$ containing the pair $\{x,y\}$ is zero. If $l \neq j$ then the vertex x of V_1 is "identified" with its (precisely one) $B_{i,l}$ neighbor $z \in Y_{i,l}$ $(z \neq y)$ and since $\frac{1}{2}M - \varepsilon M \leq t_{y,z} \leq \frac{1}{2}M + \varepsilon M$, we have that the number of sets in $\mathscr{C}_{i,l}$ containing $\{x,y\}$ is between $\frac{1}{2}M - \varepsilon M$ and $\frac{1}{2}M + \varepsilon M$.

Thus as $(2k-1)\varepsilon < \frac{1}{2}$ we infer

$$c_{x,y} \leq (k-1)\left(\frac{1}{2}M + \varepsilon M\right) < k\left(\frac{1}{2} - \varepsilon\right) M = p.$$

Now suppose the pair $\{x,y\}$ is a non-edge of G. There are 3 cases to consider.

Case 1. $\{x,y\} \subset V_1$.

We see that the pair is easily covered by the sets C_1, C_2, $\ldots$, C_p which are copies of the set V_1. Thus,

$$c_{x,y} \geq p.$$

Case 2. $\{x,y\} \subset V_2$.

The sets C_{p+1}, C_{p+2}, $\ldots$, C_{2p} which are copies of the set V_2 cover the pair $\{x,y\}$. Thus,

$$c_{x,y} \geq p.$$

Case 3. *$x \in V_1$ and $y \in X_i$ for some $i \in [k]$.*

The pair $\{x,y\}$ is contained in the vertex set of the graph A_i and in the vertex sets of each of the graphs $B_{i,1}$, $B_{i,2}$, $\ldots$, $B_{i,k}$. For a fixed $l \in [k]$, the vertex x is defined with one of the vertices z' of $Y_{i,j}$, but $z' \neq y$, since the pair $\{x,y\}$ is a non-edge. We know that $\frac{1}{2}M - \varepsilon M \leq t_{z',y} \leq \frac{1}{2}M + \varepsilon M$, and thus the number of sets in $\mathscr{C}_{i,l}$ containing $\{x,y\}$ is between $\frac{1}{2}M - \varepsilon M$ and $\frac{1}{2}M + \varepsilon M$. As this holds for all $l = 1, 2, \ldots, k$ we infer in this case,

$$c_{x,y} \geq k\left(\frac{1}{2}M - \varepsilon M\right) = p. \qquad \blacksquare$$

5. Proof of Theorem 4

To see the lower bound we follow the argument of the lower bound in Theorem 3 as follows: assuming $\theta(G) \leq t$ for any $G \in \mathscr{B}(n,d)$ we get in the same way $t \cdot w^{tn}$ as the upper bound for the number of distinct representations, while analogously to (12) and (13) we infer that $t + o(1) \geq \frac{d}{4}\log_2\left(\frac{n}{2d}\right)$ holds.

We start with the following lemma.

Lemma 5.1. *Let G be a graph with $\Delta(G) \leq d$ and $|V(G)| = n$, then there exists a d-regular graph H with $\tilde{n} = 2dn$ vertices that contains G as an induced subgraph.*

Proof. Let $G = (V,E)$ and let

$$(V_1, E_1),\ (V_2, E_2),\ \ldots,\ (V_{2d}, E_{2d})$$

be $2d$ vertex disjoint copies of G.

Let $i \in [d]$ then there exists an i-regular graph on $2d$ vertices, For example an i-regular bipartite graph on independent sets both of size d can easily be constructed.

For $v \in V$, set $d_v = \deg_G(v)$ and let E_v be the edge set of a $(d - d_v)$-regular graph imposed on the vertices $v_1, v_2, \ldots, v_{2d}$ which are the copies of v.

Set

$$V(H) = \bigcup_{i=1}^{2d} V_i$$

and

$$E(H) = \bigcup_{v \in V} E_v \cup \bigcup_{i=1}^{2d} E_i.$$

Then H satisfies the statement of the lemma. $\blacksquare$

We are now ready to prove Theorem 4. Let $G = (V,E)$ be a graph such that $\Delta(G) \leq d$ and $|V| = n$. Let H be a d-regular graph with $\tilde{n} = 2dn$ vertices that contains G as an induced subgraph the existence of which is guaranteed by Lemma 5.1. We will show that

$$\theta(H) \le c'd^2 \log(\tilde{n})$$

for some c' and as G is an induced subgraph we get the upper bound, $\theta(G) \le cd^2 \log(n)$ for some constant c as intended.

Set

$$m = ad^2 \log \tilde{n}$$

where a is a sufficiently large constant. (Say $a = 128e^2$.)

Set

$$\alpha = 1 - \frac{1}{2d},$$

$$\varepsilon = \frac{1}{2de}$$

and

$$p = (1-\varepsilon)\alpha^{2d-1}m. \tag{18}$$

For the fixed graph H we will find a family $\{Y_v : v \in V(H)\}$ of subsets of $[m]$, such that $\{u,v\} \in E$ if and only if $|Y_u \cap Y_v| \ge p$. To each edge, $\{u,v\}$ of H, we randomly assign a set $X_{u,v} \subset [m]$, with

$$\text{Prob}\ (x \in X_{u,v}) = \alpha.$$

Given a particular assignment $\{X_{u,v} : \{u,v\} \in E(H)\}$, for each $v \in V$, we set

$$Y_v = \bigcap_{\{u,v\} \in E(H)} X_{u,v}.$$

For each $\{u,v\} \in [V(H)]^2$, let $\mathbf{Y}_{u,v}$ be the random variable,

$$\mathbf{Y}_{u,v} = |Y_u \cap Y_v|$$

giving the size of the intersection of the sets assigned to the pair of vertices u and v.

We will prove that for $\{u,v\} \in E(H)$

$$\text{Prob}\ (\mathbf{Y}_{u,v} \ge p) > 1 - \frac{1}{\tilde{n}^2} \tag{19}$$

and for $\{u,v\} \notin E(H)$

$$\text{Prob}\ (\mathbf{Y}_{u,v} < p) > 1 - \frac{1}{\tilde{n}^2} \tag{20}$$

from which we may conclude that, as there are only $\binom{\tilde{n}}{2}$ pairs $\{u,v\}\in[V(G)]^2$, there is a system

$$\{X_{u,v} \ : \ \{u,v\} \in E(H)\}$$

such that the family

$$\left\{Y_u \ : \ v \in V(H) \text{ and } Y_v = \bigcup_{u,v} X_{u,v}\right\}$$

forms a representation of H.

Suppose first that u and v are two adjacent vertices. Let the neighbors of u be

$$v, u_1, u_2, \dots, u_{d-1}$$

and similarly let

$$u, v_1, v_2 \dots, v_{d-1}$$

be the neighbors of v. The random variable

$$\mathbf{Y}_{u,v} = |Y_u \cap Y_v| = \left| X_{u,v} \cap \left(\bigcap_{i=1}^{d-1} x_{u,u_i}\right) \cap \left(\bigcap_{i=1}^{d-1} x_{v,v_i}\right)\right|$$

has binomial distribution $\mathscr{B}(m,\beta_1)$ where

$$\beta_1 = \alpha^{2d-1}.$$

Hence, the expectation

$$E(\mathbf{Y}_{u,v}) = \alpha^{2d-1}m.$$

Similarly, if u, v is a pair of non-adjacent vertices we infer that $\mathbf{Y}_{u,v}$ has binomial distribution $\mathscr{B}(m,\beta_2)$ where

$$\beta_2 = \alpha^{2d}$$

and

$$E(\mathbf{Y}_{u,v}) = \alpha^{2d}m.$$

To conclude the proof of (19) and (20) we will apply Lemma 4.1, (14) to the random variable $\mathbf{Y}_{u,v}$.

Thus if u and v are adjacent then

$$\text{(21)} \qquad \text{Prob}\,(\mathbf{Y}_{u,v} < (1-\varepsilon)m\beta_1) < 2\exp\left(-\frac{\varepsilon^2}{2}(1-\varepsilon)m\beta_1\right)$$

and due to the choices of ε, m and α, see (18), (with the constant a sufficiently large) one can further bound the RHS of (21) from above by $(1.\tilde{n})^2$.

Thus

$$\text{Prob}\,(\mathbf{Y}_{u,v} < (1-\varepsilon)m\beta_1) < 2\left(\frac{1}{\tilde{n}}\right)^2 \tag{22}$$

whenever u and v are adjacent.

Similarly we infer that

$$\text{Prob}\,(\mathbf{Y}_{u,v} > (1+\varepsilon)m\beta_2) < 2\left(\frac{1}{\tilde{n}}\right)^2 \tag{23}$$

whenever u and v are nonadjacent. Due to the choice of p and ε we see that

$$m\beta_1 - \varepsilon m\beta_1 = p > m\beta_2 + \varepsilon m\beta_2$$

and this together with (22) and (23) concludes the proof of (19) and (20) and hence of the theorem. ∎

A proof using similar techniques is shown to give that there exists an absolute constant c such that

$$\theta_p(G) \leq cpd^{\frac{2}{p}} n^{\frac{1}{p}}$$

for graphs G with maximum degree bounded by d, see [3].

References

[1] N. ALON: Covering graphs by the minimum number of equivalence relations. *Combinatorica*, **6** (3) (1986), 201–206.

[2] M. S. CHUNG, and D. B. WEST: The p-intersection number of a complete bipartite graph and orthogonal double coverings of a clique, to appear.

[3] N. EATON, R. GOULD, and V. RÖDL: On p-intersection representations, to appear.

[4] P. ERDŐS, A. W. GOODMAN, and L. PÓSA: The representation of a graph by set intersections, *Canad. J. of Math.* **18** (1966), 106–112.

[5] P. ERDŐS, and L. LOVÁSZ: Problems and results on 4-chromatic hypergraphs and some related questions, In: *Colloquia Mathematica Societatis János Bolyai, 10. Infinite and Finite Sets, Vol. II.*, North-Holland Publishing, 1975.

[6] Z. FÜREDI: A random representation of the complete bipartite graph, to appear.

[7] Z. FÜREDI, and J. KAHN: On the dimensions of ordered sets of bounded degree, *Order*, **3** (1986), 15–20.

[8] W. HOEFFDING, Probability inequalities for sums of bounded random variables, *J. of the Amer. Math. Stat. Assoc.* **58** (1963), 13–30.

[9] M. S. JACOBSON, A. E. KÉZDY, and D. B. WEST: The 2-intersection number of paths and bounded-degree trees, to appear.

[10] L. LOVÁSZ, J. NEŠETŘIL, and A. PULTR: On a product dimension of graphs, *J. of Combin. Theory, Ser. B*, **29** (1980), 47–67.

[11] T. ŁUCZAK, A. RUCIŃSKI, and B. VOIGT: Ramsey properties of random graphs, *J. of Combin. Theory, Ser. B*, **56** (1) (1992), 66.

[12] J. NEŠETŘIL, and V. RÖDL: A simple proof of the Galvin–Ramsey Property of the class of all finite graphs and a dimension of a graph, *Discrete Mathematics*, **23** (1978), 49–55.

[13] S. POLJAK, and V. RÖDL: Orthogonal partitions and covering of graphs, *Czechoslovak Mathematical Journal*, **30** (105) (1980), 475–485.

[14] J. REITMAN, V. RÖDL, and E. ŠIŇAJOVÁ: On embedding of graphs into Euclidean Spaces of small dimension, *J. of Combin. Theory, Ser. B*, **54** (1992).

Nancy Eaton

University of Rhode Island
Kingston, RI 02881
`eaton@math.uri.edu`

Vojtěch Rödl

Emory University
Atlanta, GA 30322, USA
`rodl@mathcs.emory.edu`