COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA 20 (2) (2000) 227–240

GRAPH PAIRS AND THEIR ENTROPIES: MODULARITY PROBLEMS

JÁNOS KÖRNER, GÁBOR SIMONYI*

Received June 24, 1997

Graph entropy is an information theoretic functional on a graph and a probability distribution on its vertex set. It is sub-additive with respect to graph union but not submodular in general. Here we give necessary and sufficient conditions for submodularity and supermodularity of graph entropy with respect to every probability distribution in case of those couples of graphs whose union is complete. Equality in this kind of inequalities can characterize important classes of graphs as shown by our earlier results with I. Csiszár, L. Lovász, K. Marton and Zs. Tuza for couples of edge-disjoint graphs.

1. Introduction

Graph entropy $H(G,P)$ is an information theoretic functional of a graph G with a probability distribution P on its vertex set, introduced in [6]. Before giving a formal definition, we would like to mention briefly those properties of this functional which motivate our present paper.

A crucial property of $H(G,P)$ is its sub-additivity with respect to graph union [8]; if F and G are two graphs on the same vertex set V and $F \cup G$ denotes the graph on V with edge set $E(F\cup G)=E(F)\cup E(G)$, then for every P we have

$$H(F\cup G,P) \le H(F,P)+H(G,P). \tag{1}$$

* Research of this author was partially supported by the Hungarian National Foundation for Scientific Research (OTKA) Grant Nos. F023442, T016386, T029255 and T032323.

Mathematics Subject Classification (1991): 05C15, 05C75, 94A17

If F and G are complementary graphs and thus $F \cup G$ is complete, then equality for every P above is equivalent to F and G being perfect, (cf. [2].) This makes one believe that equality in this kind of inequalities does express relevant structural properties of graph pairs. This conviction obtains further support from what happens in the case of two arbitrary edge-disjoint graphs F and G sharing the same vertex set V. In fact, in this latter case we have equality in (1) for every distribution P on V if and only if for all adjacent pairs of edges $\{x,y\} \in E(F)$ and $\{y,z\} \in E(G)$ their "private" endpoints x and z are adjacent in $F \cup G$, and further F (and thus G) induce perfect graphs on every complete subgraph of their union, ([11].) It is therefore interesting to continue on this road and try to generalize the above results to the case of graphs with some common edges. This is what we are going to do in this paper for a new case.

Our present situation is, in a certain sense, opposite to the one treated in [11]; two graphs F and G whose union is complete. Hence the complements of these graphs are edge-disjoint and thus if both F and G were perfect, the aforementioned results (1) and [2] would immediately yield

$$H(F,P) + H(G,P) \leq H(F \cap G, P) + H(P).$$

(Here $(F \cap G)$ stands for the graph with edge set $E(F \cap G) = E(F) \cap E(G)$ and vertex set V.) However, this complementation does not work in the remaining cases; in fact, the very inequality it would give fails to be true in general. Therefore, to treat our present problem some new idea will be needed.

Before giving the formal definition of graph entropy we formulate our main results. Let F and G be two graphs on the same vertex set V with possibly intersecting edge sets. We call F and G a submodular pair if for every probability distribution P on V we have

$$H(F \cup G, P) + H(F \cap G, P) \leq H(F,P) + H(G,P). \tag{2}$$

Likewise, we call F and G a supermodular pair if the sense of the last inequality is reversed.

Our main concern is to find conditions for F and G to make the graph couple $\{F,G\}$ submodular (supermodular) under the additional condition that $F \cup G = K_{|V|}$, i.e., the complete graph on $|V|$ vertices. It follows from the already mentioned result in [2] that if $F \cap G$ is empty and $F \cup G$ is the complete graph on V then (strict) submodularity is equivalent to F (and thus G) being imperfect. (A graph is called perfect if for each of its induced subgraphs G' its chromatic number is equal to its clique number. A graph is imperfect if it is not perfect.) This implies that F and G are certainly not a supermodular pair if on some clique of their union they are edge-disjoint and induce imperfect graphs. (From the definition we will give this can be seen by concentrating a probability distribution on this clique.)

On the other hand it takes an easy calculation to show that two different paths consisting of two edges each on the same three-element vertex set does not form a submodular pair. Thus we have two simple conditions that ensure the existence of probability distributions with strict inequalities in one or the other direction. Our main result states that actually there are no more conditions. Formally it is given by the following two statements.

Theorem 1. *Two graphs F and G on the same vertex set V with $F \cup G = K_{|V|}$ form a submodular pair if and only if there is no three-element subset of $|V|$ on which each of $F - G$, $G - F$ and $F \cap G$ has exactly one edge.*

Theorem 2. *The graphs F and G on the same vertex set $|V|$ with $F \cup G = K_{|V|}$ form a supermodular pair if and only if there is no subset $U \subseteq V$ on which F and G are imperfect and edge-disjoint.*

These two results immediately yield a full characterization of modularity. The graphs F and G on the same vertex set form a *modular pair* if they are both submodular and supermodular. We have

Corollary 1. *The graphs F and G on the same vertex set V with $F \cup G = K_{|V|}$ form a modular pair if and only if the following two conditions simultaneously hold:*

(i) *there is no three-element subset of $|V|$ on which each of $F - G$, $G - F$ and $F \cap G$ have exactly one edge,*

(ii) *there is no subset $U \subseteq V$ on which F and G are imperfect and edge-disjoint.* ∎

On the other hand, it must be clear that there are graph pairs that are neither submodular nor supermodular.

In this paper we limit ourselves to prove these results representing a new step in the analysis of the behavior of the sum of the entropies of two graphs, first raised in the paper [7] joint with G. Longo, and solved for graphs without common edges in a series of papers written in collaboration with K. Marton [9] (if one of the graphs is bipartite), with I. Csiszár, L. Lovász and K. Marton [2] (in the aforementioned case of two edge-disjoint graphs whose union is the complete graph on their common vertex set) and with Zs. Tuza [11] (for arbitrary edge-disjoint graph pairs).

These problems seem to be interesting even for yet another reason, for the sub-additivity of graph entropy is at the core of a bounding technique in combinatorics and computer science that has been succesfully applied by various authors, cf., e. g., Kahn and Kim [4] and Radhakrishnan [12]. For more applications and details we refer to the survey article [13]. When not stated otherwise, we adopt the terminology of [1]. In particular, exp's and log's are to the base 2.

2. Graph entropy

Graph entropy is formally defined as

$$H(G,P) = \min_{X \in Y \in S(G),\, P_X = P} I(X \wedge Y),$$

where $S(G)$ denotes the family of the stable sets of vertices in G. (A subset of the vertex set is called stable if it does not contain any edge.) For the basics in information theory the reader is referred to the book [1]. We recall that the mutual information $I(X \wedge Y)$ of the random variables X and Y equals $H(X)+H(Y)-H(X,Y)$, where e. g. $H(X,Y)$ is the entropy of the random variable (X,Y). (Notice that the entropy of a random variable is the entropy of its distribution.) It is immediate from this definition that if K is a complete graph and P an arbitrary distribution on its vertex set, then $H(K,P)=H(P)$. Likewise, the entropy of a graph without edges is 0.

This definition of graph entropy has the merit of being short but it is not particularly intuitive (for those who are not familiar with information theory). In a later part of this paper, based on a more complicated but purely combinatorial definition, we shall develop a combinatorial intuition for this quantity. For the time being we just anticipate that entropy is a kind of fractional chromatic number.

If $E(F) \cap E(G) \neq \emptyset$, we cannot have equality in (1) for a P concentrated on the two endpoints of some common edge. This raises the question of whether, introducing as a correcting term the entropy of the intersection graph $F \cap G$, submodularity of graph entropy is true for two graphs with non-disjoint edge sets. In general the answer is negative as shown by the already stated theorems. Their main content is that more can be said.

3. Submodularity

In this section we shall prove Theorem 1. We begin by recalling two earlier results needed in the proof. The first of these is a theorem of Gallai following from his Decomposition Theorem in [3] (cf. also [5] and [11].)

Theorem G. ([3]) *Let the graphs $H_1, H_2, \dots, H_k$ be edge-disjoint with their union being the complete graph on their common vertex set. If for no 3 vertices does the resulting triangle have its three edges in three different H_i's then at most two of the H_i's are at the same time connected and spanning the whole vertex set.*

The second result we need establishes a kind of additivity of graph entropy ([11]), called the *Substitution Lemma*, cf. also [13]. To state the lemma, we need the concept of graph substitution. Let A and B be two graphs and consider any $v \in V(A)$. Substituting B for v means to delete v from A and join every vertex of a copy of B to exactly those vertices of A which were adjacent to v in A. The resulting graph is denoted by $A_{v\leftarrow B}$. Substitution is also extended to distributions as follows. Let P be a distribution on $V(A)$ and Q a distribution on $V(B)$. Then a distribution $P_{v\leftarrow Q}$ is obtained on $V(A_{v\leftarrow B})$ by putting $P_{v\leftarrow Q}(x) = P(x)$ if $x \in V(A) - \{v\}$ and $P_{v\leftarrow Q}(x) = P(v)Q(x)$ if $x \in V(B)$. The statement we need is the following:

Substitution Lemma. ([11]) *Let A and B be two vertex-disjoint graphs with probability distributions P and Q on their respective vertex sets and v be any vertex of A. Then*

$$H(A_{v\leftarrow B}, P_{v\leftarrow Q}) = H(A,P) + P(v)H(B,Q).$$

Now we are ready to prove our first result.

Proof of Theorem 1. Let us call a multicolored triangle (MCT) a configuration of three vertices of V on which each of $F-G$, $G-F$ and $F\cap G$ have exactly one edge. If F and G induce such a configuration, then defining P to be the uniform distribution on its 3 vertices an easy calculation shows that

$$H(F\cup G, P) + H(F\cap G, P) > H(F,P) + H(G,P).$$

We prove the other direction by induction. The statement is trivially true if the vertex set $V = V(F) = V(G)$ has less than 3 elements and it is also easy to check for $|V| = 3$. In fact, in the latter case at least one of the three graphs $F\cap G$, $F-G$ and $G-F$ has no edges. If it is the first one, then F and G are edge-disjoint and the statement follows by the sub-additivity of graph entropy. In any other case one of $E(F)$ and $E(G)$ contains the other and we have equality in (2).

In order to make the induction work, we have to deal separately with the case when one of the graphs $F\cap G$, $F-G$, $G-F$ has no edges (i.e., it is an empty graph). This case can be settled analogously to that of $|V| = 3$.

In fact, if $F \cap G$ is the empty graph, this means that F and G are edge-disjoint and we are done by the sub-additivity of graph entropy (1). In the other cases one of F and G contains the other and we have equality in our inequality needed for submodularity.

Assume the statement is true for $|V| < n$. Let F and G be two graphs on V with $|V| = n$ such that $F \cup G = K_n$ and with no MCT (i. e. no triangle having an edge in each of $F - G$, $G - F$, and $F \cap G$.) Then by Gallai's Theorem G above at least one of $T_1 = F - G$, $T_2 = G - F$ and $T_3 = F \cap G$ fails to be connected on V. Let T_i be one of the disconnected graphs among these three and let $U \subseteq V$ be the vertex set of some fixed non-empty connected component of T_i. (In case there is none, T_i is the empty graph and we are already done.) Note that $|U| < n$, and by the foregoing we can also assume $|U| > 1$.

Let v be an arbitrary vertex in $V - U$. Observe that either all edges between v and the vertices of U belong to T_{i+1} or they all belong to T_{i+2} (with addition in the indices intended modulo 3), otherwise an MCT would occur. This means that in all of the graphs F, G, $F \cap G$, and (trivially) $F \cup G$ the set U forms a so-called *autonomous set* (which means that every vertex outside U is either connected to all or none of the vertices in U). Let X denote any of the graphs F, G, $F \cap G$, $F \cup G$. Let X_U be the graph uniquely defined by the property of inducing the same graph as X on U but with all the points in $V - U$ being isolated. Furthermore, let $X_{\bar{U}}$ denote the graph we obtain from X if we replace (by an "inverse substitution") the set U by a single vertex u, i.e., $V(X)$ is $(V - U) \cup \{u\}$ and u is connected to exactly those points of $V - U$ which were adjacent to the vertices of U in X. Observe that $X_{\bar{U}}$ cannot contain any MCT. We set $P(u) = P(U) = \Sigma_{x \in U} P(x)$ and with a slight abuse of notation refer to the distribution obtained on $V(X_{\bar{U}})$ in this manner also as P. Now by the Substitution Lemma we have

$$H(X, P) = H(X_{\bar{U}}, P) + H(X_U, P).$$

Notice that deleting the edges of X_U from X the entropy of the resulting graph is equal to the entropy $H(X_{\bar{U}}, P)$. (This also follows from the Substitution Lemma and the fact that the entropy of a graph without edges is zero. Besides, we have used the fact that $H(X_U, P)$ remains the same if we suppress in it the isolated points of $V - X$.) But since both U and $V(X_{\bar{U}})$ are strictly smaller than n, (in the case of the latter this follows from $|U| > 1$), we will be able to use the induction hypothesis on these sets. By adding the corresponding inequalities we are basically done, still let us formalize what we just said. Writing F, G, $F \cap G$, and $F \cup G$ in place of X and using the induction hypothesis we have

$$H(F,P)+H(G,P)=(H(F_{\bar{U}},P)+H(F_U,P))+(H(G_{\bar{U}},P)+H(G_U,P))=$$
$$=(H(F_{\bar{U}},P)+H(G_{\bar{U}},P))+(H(F_U,P)+H(G_U,P))\leq$$
$$\leq(H((F\cup G)_{\bar{U}},P)+H((F\cap G)_{\bar{U}},P))+$$
$$+(H((F\cup G)_U,P)+H((F\cap G)_U,P))=$$
$$=H(F\cup G,P)+H(F\cap G,P)$$

and this proves the theorem. ∎

4. Supermodularity

In this section we prove Theorem 2. The proof we present here is due to Kati Marton who suggested the following short proof to replace our original considerably longer argument. Later in the paper we also give our original proof for it uses the information theoretic concepts less extensively. However, the brevity of Kati Marton's proof convinced us that this proof should be given first.

Proof of Theorem 2. The "only if" part of the statement is an immediate consequence of the already mentioned information theoretic characterization of perfect graphs in [2]; if there is a subset $U\subseteq V$ on which F and G induce two edge-disjoint imperfect graphs the union of which is complete, then by Theorem 2 in [2] there is a distribution P concentrated on U for which

$$H(F,P)+H(G,P)>H(F\cup G,P).$$

Since $F\cap G$ has no edges in U, the right-hand side is further equal to $H(F\cup G,P)+H(F\cap G,P)$ and thus F and G fail to be a supermodular pair.

To see the "if" part, let F and G be two graphs satisfying the conditions in the statement and let again $S(A)$ denote the family of all stable sets of graph A. Let X and Z be random variables taking their values in V and in $S(F\cap G)$, respectively, such that $X\in Z$ with probability 1 and

$$H(F\cap G,P)=I(X\wedge Z).$$

Fix a value U of Z. Then U is a stable set of $F\cap G$. Let F_U and G_U denote the subgraphs induced on U in F and G, respectively, and P_U the conditional distribution of X given $Z=U$. Since F_U and G_U are edge disjoint, they are perfect. Therefore,

$$H(X|Z=U) = H(P_U) = H(G_U, P_U) + H(F_U, P_U).$$

(Here $H(X|Z=U)$ is the entropy of the conditional distribution of X given that $Z=U$. $H(X|Z)$ of the next formula is the avarage of these values with respect to the distribution of Z, called the conditional entropy of X (with respect to Z.) For further details on these notions the reader is referred again to [1].)

Averaging with respect to the distribution of Z gives

$$H(X|Z) = \sum_U Pr\{Z=U\}H(G_U, P_U) + \sum_U Pr\{Z=U\}H(F_U, P_U).$$

Adding $2I(X \wedge Z)$ to both sides we obtain

$$\sum_U Pr\{Z=U\}H(G_U, P_U) + I(X \wedge Z) +$$
$$+ \sum_U Pr\{Z=U\}H(F_U, P_U) + I(X \wedge Z) =$$
$$H(X|Z) + 2I(X \wedge Z) = H(P) + I(X \wedge Z) =$$
$$= H(P) + H(F \cap G, P) = H(F \cup G, P) + H(F \cap G, P).$$

Thus it is enough to show that

$$\sum_U Pr\{Z=U\}H(G_U, P_U) + I(X \wedge Z) \geq H(G, P).$$

This will imply the similar inequality for F and thus the statement.

Let $Y \in S(G)$ be a random variable such that the joint distribution of (X,Y,Z) satisfies the conditions that $X \in Y \subset Z$ with probability 1 and for every value U of Z

$$H(G_U, P_U) = I(X \wedge Y|Z=U).$$

(These conditions define the conditional distribution of Y given X and Z, and thus the joint distribution of (X,Z,Y), too.) We have then

$$\sum_U Pr\{Z=U\}H(G_U, P_U) + I(X \wedge Z) = I(X \wedge Y|Z) + I(X \wedge Z) =$$
$$= I(X \wedge YZ) \geq I(X \wedge Y) \geq H(G, P).$$

Thus the proof is completed. ∎

5. Graph entropy made simple

Much of the remaining part of our paper contains a second proof of Theorem 2, the proof we originally had. To present it we have to make some preparations.

For any natural n the n'th (co-normal or OR) power of G is the graph with vertex set V^n in which two vertices, $\mathbf{x}=x_1x_2\dots x_n$ and $\mathbf{y}=y_1y_2\dots y_n$ are adjacent if for at least one $i\in\{1,2,\dots,n\}$ the relation $\{x_i,y_i\}\in E(G)$ holds. Given an arbitrary probability distribution P on the vertex set V of G, let $\mathcal{T}_P^n$ denote the (possibly empty) set of all sequences $\mathbf{x}\in V^n$ for which

$$|\{i;\quad x_i = a\ ,\quad \mathbf{x} = x_1x_2\dots x_n\}| = nP(a) \text{ for every } a\in V.$$

Further, let G_P^n denote the graph the set $\mathcal{T}_P^n$ induces in G^n. As usual, we denote by $\alpha(G)$ the maximum cardinality of a stable set and by $\chi(G)$ the chromatic number of the graph G. Graph entropy can be redefined in terms of the above concepts, following the approach of [6].

We shall call a distribution *rational* if all its probabilities are rational numbers. Let $k=k(P)$ be the smallest natural number for which all the numbers $kP(x)$, $x\in V$ are integers. It is well-known (cf. Lemma 1.2.3 in [1]) that

$$\lim_{n\to\infty}\frac{1}{kn}\log|\mathcal{T}_P^{kn}| = H(P). \tag{3}$$

Likewise, it is not hard to see (and it is implicit in [6]) that for any simple graph with vertex set $V(G)=V$ one has

$$\lim_{n\to\infty}\frac{1}{kn}\log\chi(G_P^{kn}) = H(G,P), \tag{4}$$

and, defining $S(G,P)=H(P)-H(G,P)$ also

$$\lim_{n\to\infty}\frac{1}{kn}\log\alpha(G_P^{kn}) = S(G,P). \tag{5}$$

More precisely, there exists a sequence of constants $r_n=r_n(|V|)$ with $r_n\to 0$ for which

$$\left|\frac{1}{kn}\log\alpha(G_P^{kn}) - S(G,P)\right| \le r_n,$$

and similarly in the preceding inequalities. The relation (4) could be used to give an alternative definition of graph entropy. It defines the entropy of a graph for rational distributions whence the general case would follow by continuity. (The original definition in [6] is in this spirit.) It easily follows from the above that

$$\max H(G,P) = \log\chi^*(G),$$

where the maximum is for all the probability distributions P on the vertex set of G and where $\chi^*(G)$ stands for the fractional chromatic number of the graph G. This fact is not needed in the sequel, and we quote it only to give more intuitive sense to the concept of graph entropy.

6. A more combinatorial proof

Second proof of Theorem 2.

We reprove only the more difficult "if" part of the statement. To this end, recall that $S(G,P)=H(P)-H(G,P)$. Further, one has to remember that since the union graph $F\cup G$ is complete, we have $S(F\cup G,P)=0$. Our statement is therefore equivalent to saying that under the conditions of the theorem

$$S(F\cap G,P)\le S(F,P)+S(G,P). \tag{6}$$

Clearly, it is sufficient to prove this statement for rational probability distributions since the rest follows by continuity. As an easy consequence of (5) one proves (by a simple "time sharing" argument) that for fixed G the quantity $S(G,P)$ is a cap-convex function of the distribution P. (The name "time sharing" is standard in information theory for the kind of argument we need here. What we do to get the proof of cap-convexity is to consider at first only convex combinations of distributions with rational coefficients, since the rest will follow by continuity. If we want to prove that

$$\lambda S(G,P_1)+(1-\lambda)S(G,P_2)\le S(G,\lambda P_1+(1-\lambda)P_2)$$

for rational distributions P_1 and P_2 on $V(G)$ and an arbitrary rational number $\lambda\in[0,1]$, it is sufficient to notice that the Cartesian product of any stable set of maximum cardinality of the graph induced by $G^{\lambda n}$ on $\mathcal{T}_{P_1}^{\lambda n}$ with any stable set of maximum cardinality of the graph induced by $G^{(1-\lambda)n}$ on $\mathcal{T}_{P_2}^{(1-\lambda)n}$ defines a stable set in the graph induced by G^n on $\mathcal{T}_P^n$, where $P=\lambda P_1+(1-\lambda)P_2$. We omit the details needed to guarantee the integrality of various exponents and similar technicalities.)

Fix some n and let $S=S(n)\subseteq V^n$ be a maximal stable set of $(F\cap G)^n$ for which $S\cap\mathcal{T}_P^n$ achieves $\alpha((F\cap G)_P^n)$. Clearly, $S=\prod_{i=1}^n S_i$ for some stable subsets $S_i\subseteq V$ of the graph $F\cap G$. Since the order of the sets in this Cartesian product is indifferent we can suppose that

$$S=\prod_{U\in\mathcal{Y}(F\cap G)} U^{nQ_n(U)} \tag{7}$$

where $\mathcal{Y}(F\cap G)$ is the family of all the maximal stable sets of the graph $F\cap G$ and Q_n is defined by

$$Q_n(U) = \frac{|\{i; S_i = U\}|}{n}. \tag{8}$$

Given an arbitrary sequence $\mathbf{x}\in S\cap\mathcal{T}_P^n$ define for every $a\in V$

$$P_{\mathbf{x}|U}(a) = \frac{|\{i; x_i = a, S_i = U\}|}{|\{i; S_i = U\}|} \tag{9}$$

and let the random variables (RV in the sequel) X and Z have the joint distribution

$$\Pr\{X = a, Z = U\} = Q_n(U)P_{\mathbf{x}|U}(a) \quad \forall a \in V, \quad \forall U \in \mathcal{Y}(F\cap G). \tag{10}$$

Now consider the set $\mathcal{Z}_n$ of all the joint distributions on $V\times\mathcal{Y}(F\cap G)$ emerging in this manner (i. e., with an arbitrary but fixed S as above and for some $\mathbf{x}\in S\cap\mathcal{T}_P^n$). An easy calculation shows that their number is only polynomial in n, and more precisely, (cf. e. g. Lemma 1.2.2 in [1]) we have

$$|\mathcal{Z}_n| \le (n+1)^{|V||\mathcal{Y}(F\cap G)|}, \tag{11}$$

and thus

$$|S(n)\cap\mathcal{T}_P^n| = \sum_{P_{XZ}\in\mathcal{Z}_n} |\{\mathbf{x}; \mathbf{x}\in S(n)\cap\mathcal{T}_P^n,\ P_{\mathbf{x}|U}(a)Q_n(U) = P_{XZ}(a,U)$$
$$\forall a\in V,\ \forall U\in\mathcal{Y}(F\cap G)\}| \le \sum_{P_{XZ}\in\mathcal{Z}_n} \exp[nH(X|Z)],$$

where the last inequality is immediate from Lemma 1.2.5 in [1]. We continue this chain of inequalities by using (11) to give

$$|S(n)\cap\mathcal{T}_P^n| \le (n+1)^{|V||\mathcal{Y}(F\cap G)|} \max_{P_{XZ}\in\mathcal{Z}_n} \exp[nH(X|Z)], \tag{12}$$

where the constraint $P_{XZ}\in\mathcal{Z}_n$ means that the joint distribution of the RV's X and Z is contained in the set of distributions $\mathcal{Z}_n$. Notice next that the conditional entropy $H(X|Z)$ can be written as

$$H(X|Z) = \sum_{U\in\mathcal{Y}(F\cap G)} Q_n(U)H(R_U) \tag{13}$$

for some distributions R_U on V satisfying the condition

$$\sum_{U \in \mathcal{Y}(F \cap G)} Q_n(U) R_U(a) = P(a) \qquad \forall a \in V. \tag{14}$$

Further recall that by our hypothesis F and G induce two edge-disjoint perfect graphs on U and thus by Theorem 2 of [2]

$$H(R_U) = S(F, R_U) + S(G, R_U). \tag{15}$$

Substituting (15) into (13) we see that

$$H(X|Z) = \sum_{U \in \mathcal{Y}(F \cap G)} Q_n(U)[S(F, R_U) + S(G, R_U)]. \tag{16}$$

At this point we apply the cap-convexity of $S(F,R)$ and $S(G,R)$ and the relation (14) to imply

$$H(X|Z) \le S(F, P) + S(G, P).$$

Substituting this into (12) we see that

$$|S(n) \cap \mathcal{T}_P^n| \le (n+1)^{|V||\mathcal{Y}(F \cap G)|} \exp\{n[S(F,P) + S(G,P)]\}. \tag{17}$$

However,

$$|S(n) \cap \mathcal{T}_P^n| = \alpha((F \cap G)_P^n)$$

by the definition of $S(n)$, whence recalling the asymptotics of $\alpha((F \cap G)_P^n)$ from (5) the inequality (17) implies (6). ∎

7. Outlook

We have just begun to study the sub-(super)modularity properties of more general graph pairs than treated above, and at this point all we can report are some rather intriguing preliminary observations. The entropies of bipartite graphs can be expressed in terms of the binary entropy function and thus in these cases the relevant inequalities for submodularity might be proved by elementary convexity arguments. We have proved e. g. that if $|E(F-G)| = |E(G-F)| = 1$, and if further $F \cup G$ is a simple path on k vertices whose middle edges (those with endpoints of degree 2) belong to $F \cap G$, then the graphs F and G form a submodular pair for $k = 3, 5$ and a supermodular pair if $k = 4$. We are not ready to make any conjecture in more generality, yet it is clear already that a complete understanding of this kind of properties may put graph pairs into an interesting new perspective.

Our discussion can be extended to hypergraphs. In terms of applications, e. g., in computer science the resulting entropy is as natural as graph entropy, since sub-additivity remains valid [10]. For conditions of additivity of hypergraph entropy in case of uniform hypergraphs the reader is referred to [14].

Acknowledgement. We thank Kati Marton for suggesting the simple proof of Theorem 2 presented in Section 4.

References

[1] I. Csiszár, J. Körner: *Information theory: Coding theorems for discrete memoryless systems*, Academic Press, New York, 1982 and Akadémiai Kiadó, Budapest, 1981.

[2] I. Csiszár, J. Körner, L. Lovász, K. Marton, G. Simonyi: Entropy splitting for antiblocking pairs and perfect graphs, *Combinatorica*, **10** (1) (1990), 27–40.

[3] T. Gallai: Transitiv orientierbare Graphen, *Acta Math. Acad. Sci. Hungar.* **18** (1967), 25–66.

[4] J. Kahn, J. H. Kim: Entropy and sorting, *J. Computer and System Sciences*, **51** (1995), 390–399.

[5] D. Kelly: Comparability graphs, in: *Graphs and Orders* (I. Rival ed.), D. Reidel Publ. Co. (1985), 3–40.

[6] J. Körner: Coding of an information source having ambiguous alphabet and the entropy of graphs, *Transactions of the 6th Prague conference on Information Theory, etc., 1971*, Academia, Prague, (1973), 411–425.

[7] J. Körner, G. Longo: Two-step encoding of finite sources, *IEEE Trans. Inform. Theory,* **19**, Nov. 1973, 778–782.

[8] J. Körner: Fredman–Komlós bounds and information theory, *SIAM J. Algebraic and Discrete Methods*, **7** (4), (1986), 560–570.

[9] J. Körner, K. Marton: Graphs that split entropies, *SIAM J. Discrete Math.*, **1** (1), (1988), 71–79.

[10] J. Körner, K. Marton: New bounds for perfect hashing via information theory, *European J. Combinatorics*, **9** (1988), 523–530.

[11] J. Körner, G. Simonyi, Zs. Tuza: Perfect couples of graphs, *Combinatorica*, **12** (2) (1992), 179–192.

[12] J. Radhakrishnan: $\Sigma\Pi\Sigma$ threshold formulas, *Combinatorica*, **14** (3) (1994), 345–374.

[13] G. Simonyi: Graph entropy: a survey, in: *Combinatorial Optimization*, DIMACS Series on Discrete Math. and Computer Science Vol. 20 (W. Cook, L. Lovász, P. D. Seymour eds.), AMS, 1995, 399–441.

[14] G. Simonyi: Entropy splitting hypergraphs, *J. Combin. Theory Ser. B* **66** (1996), 310–323.

János Körner

Department of Computer Science
Università "La Sapienza"
00198 Roma, Italy
korner@dsi.uniroma1.it

Gábor Simonyi

Alfréd Rényi Institute of Mathematics
Hungarian Academy of Sciences
H-1364 Budapest,
P.O.B. 127, Hungary
simonyi@renyi.hu