Bounding the partition function of spin-systems

David J. Galvin*

Department of Mathematics
University of Pennsylvania
209 South 33th Street
Philadelphia PA 19104 USA
dgalvin@math.upenn.edu

Submitted: Aug 23, 2005; Accepted: Jul 28, 2006; Published: Aug 22, 2006
Mathematics Subject Classifications: 05C15, 82B20

Abstract

With a graph $G = (V, E)$ we associate a collection of non-negative real weights $\bigcup_{v \in V}\{\lambda_{i,v} : 1 \leq i \leq m\} \cup \bigcup_{uv \in E}\{\lambda_{ij,uv} : 1 \leq i \leq j \leq m\}$. We consider the probability distribution on $\{f : V \to \{1, \ldots, m\}\}$ in which each f occurs with probability proportional to $\prod_{v \in V} \lambda_{f(v),v} \prod_{uv \in E} \lambda_{f(u)f(v),uv}$. Many well-known statistical physics models, including the Ising model with an external field and the hard-core model with non-uniform activities, can be framed as such a distribution. We obtain an upper bound, independent of G, for the partition function (the normalizing constant which turns the assignment of weights on $\{f : V \to \{1, \ldots, m\}\}$ into a probability distribution) in the case when G is a regular bipartite graph. This generalizes a bound obtained by Galvin and Tetali who considered the simpler weight collection $\{\lambda_i : 1 \leq i \leq m\} \cup \{\lambda_{ij} : 1 \leq i \leq j \leq m\}$ with each λ_{ij} either 0 or 1 and with each f chosen with probability proportional to $\prod_{v \in V} \lambda_{f(v)} \prod_{uv \in E} \lambda_{f(u)f(v)}$. Our main tools are a generalization to list homomorphisms of a result of Galvin and Tetali on graph homomorphisms and a straightforward second-moment computation.

1 Introduction and statement of results

Let $G = (V(G), E(G))$ and $H = (V(H), E(H))$ be non-empty graphs. Set

$$Hom(G, H) = \{f : V(G) \to V(H) : uv \in E(G) \Rightarrow f(u)f(v) \in E(H)\}$$

(that is, $Hom(G, H)$ is the set of graph homomorphisms from G to H). In [4], the following result is obtained using entropy considerations (and in particular, Shearer's Lemma).

*This work was begun while the author was a member of the Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540 and was supported in part by NSF grant DMS-0111298.

Theorem 1.1 *For any graph H and any d-regular N-vertex bipartite graph G,*

$$|Hom(G,H)| \leq |Hom(K_{d,d},H)|^{\frac{N}{2d}}$$

where $K_{d,d}$ is the complete bipartite graph with d vertices in each partition class.

In [4] Theorem 1.1 is extended to a result on weighted graph homomorphisms. To each $i \in V(H)$ assign a positive pair of weights (λ_i, μ_i). Write (Λ, M) for the set of weights. For a bipartite graph G with bipartition classes $\mathcal{E}_G$ and $\mathcal{O}_G$ give each $f \in Hom(G,H)$ weight

$$w^{(\Lambda,\mathrm{M})}(f) := \prod_{v \in \mathcal{E}_G} \lambda_{f(v)} \prod_{v \in \mathcal{O}_G} \mu_{f(v)}.$$

The constant that turns this assignment of weights on $Hom(G,H)$ into a probability distribution is

$$Z^{(\Lambda,\mathrm{M})}(G,H) := \sum_{f \in Hom(G,H)} w^{(\Lambda,\mathrm{M})}(f).$$

The following is proved in [4].

Theorem 1.2 *For any graph H, any set (Λ, M) of positive weights on $V(H)$ and any d-regular N-vertex bipartite graph G,*

$$Z^{(\Lambda,\mathrm{M})}(G,H) \leq \left(Z^{(\Lambda,\mathrm{M})}(K_{d,d},H)\right)^{\frac{N}{2d}}.$$

Taking all weights to be 1, Theorem 1.2 reduces to Theorem 1.1.

In this note we consider a more general weighted model. Fix $m \in \mathbb{N}$ and a graph $G = (V,E)$. With each $1 \leq i \leq m$ and $v \in V$ associate a non-negative real weight $\lambda_{i,v}$ and with each $1 \leq i \leq j \leq m$ and $uv \in E$ associate a non-negative real weight $\lambda_{ij,uv}$. Set $\lambda_{ij,uv} := \lambda_{ji,uv}$ for $i > j$. Write W for the collection of weights and for each $f : V \rightarrow \{1,\ldots,m\}$ set

$$w^W(f) = \prod_{v \in V} \lambda_{f(v),v} \prod_{uv \in E} \lambda_{f(u)f(v),uv}$$

and

$$Z^W(G) = \sum_{f:V \rightarrow \{1,\ldots,m\}} w^W(f).$$

We may put all this in the framework of a well-known mathematical model of physical spin systems. We think of the vertices of G as particles and the edges as bonds between pairs of particles (typically a bond represents spatial proximity), and we think of $\{1,\ldots,m\}$ as the set of possible spins that a particle may take. For each $v \in V$ we think of the weights $\lambda_{\cdot,v}$ as a measure of the likelihood of seeing the different spins at v; furthermore, for each $uv \in E$ we think of the weights $\lambda_{\cdot,uv}$ as a measure of the likelihood of seeing the different spin-pairs across the edge uv. The probability of a particular spin configuration is thus proportional to the product over the vertices of G of the weights of the spins times the product over the edges of G of the weights of the spin-pairs. In this language $Z^W(G)$ is the partition function of the model – the normalizing constant that turns the above-described system of weights on the set of spin configurations into a probability measure.

An example of such a model is the *hard-core* (or independent set) model. Here $m = 2$ and the system of weights W_{hc} is given by

$$\lambda_{i,v} = \begin{cases} \lambda_v & \text{if } i = 1 \\ 1 & \text{if } i = 2 \end{cases} \qquad \text{and} \qquad \lambda_{ij,uv} = \begin{cases} 0 & \text{if } i = j = 1 \\ 1 & \text{otherwise,} \end{cases}$$

and so

$$Z^{W_{hc}}(G) = \sum_{f:V\to\{1,2\}} \Big(\prod_{v:f(v)=1} \lambda_v\Big) \left(\mathbf{1}_{\{\nexists uv\in E: f(u)=f(v)=1\}}\right) = \sum_{I\in\mathcal{I}(G)} \prod_{v\in I} \lambda_v$$

is a weighted sum of independent sets in G. (Recall that $I \subseteq V$ is *independent* in G if for all $u, v \in I$, $uv \notin E$. We write $\mathcal{I}(G)$ for the collection of independent sets in G.)

The hard-core model is a *hard-constraint* model in which all of the edge-weights are either 0 or 1, and the rôle of these weights is to exclude certain configurations from contributing to the partition function. We now consider the best known example of a *soft-constraint* model (one in which all configurations are potentially allowable), the *Ising* model. Here $m = 2$ and there are two parameters, $\beta, h \in \mathbb{R}$. It is convenient to take the set of spins to be $\{+1, -1\}$. The system $W_{Ising,\beta,h}$ of weights on $\{+1, -1\}$ is as follows.

$$\begin{aligned} \lambda_{+1,v} &= e^{h} \text{ for all } v \in V \\ \lambda_{-1,v} &= e^{-h} \text{ for all } v \in V \\ \lambda_{ii,uv} &= e^{-\beta} \text{ for } i \in \{+1,-1\} \text{ and all } uv \in E \text{ and} \\ \lambda_{+1-1,uv} &= e^{\beta} \text{ for all } uv \in E. \end{aligned}$$

For each $\sigma : V \to \{+1, -1\}$ we have

$$w^{W_{Ising,\beta,h}}(\sigma) = \exp\Big\{ -\beta \sum_{uv\in E} \sigma(u)\sigma(v) + h \sum_{v\in V(G} \sigma(v)\Big\}.$$

Then $Z^{W_{Ising,\beta,h}}(G) = \sum_\sigma w^{W_{Ising,\beta,h}}(\sigma)$ is the partition function of the Ising model on G with inverse temperature $|\beta|$ and external field h. (If $\beta > 0$, we are in the anti-ferromagnetic case, where configurations with many $+1$--1 edges are favoured; if $\beta < 0$, we are in the ferromagnetic case, where configurations with few $+1$--1 edges are favoured.)

Let us now set up the notation for our main result. For completeness, we choose to make the straightforward generalization from regular bipartite graphs to (a, b)-biregular graphs, that is, bipartite graphs in which one partition class, which we shall label $\mathcal{E}_G$, consists of vertices of degree a and the other class, $\mathcal{O}_G$, consists of vertices of degree b. For $v \in \mathcal{O}_G$ write $\{n_1(v), \ldots, n_b(v)\}$ for the set of neighbours of v.

Let G be such a graph and let W be a collection of weights on G. Give labels $w_1, \ldots, w_b$ to the degree a vertices of $K_{a,b}$ (the complete bipartite graph with a vertices in one partition class and b in the other) and labels $z_1,\ldots,z_a$ to the degree b vertices. For $v\in\mathcal{O}_G$ write W^v for the following system of weights on $K_{a,b}$:

$$\lambda^v_{i,u}=\left\{\begin{array}{ll}\lambda_{i,v} & \text{if } u=z_k \text{ for some } 1\leq k\leq a\\ \lambda_{i,n_k(v)} & \text{if } u=w_k \text{ for some } 1\leq k\leq b\end{array}\right.$$

and, for $1\leq k\leq b$ and $1\leq \ell\leq a$, $\lambda_{ij,w_kz_\ell}=\lambda_{ij,n_k(v)v}$. Our main result is a generalization of Theorem 1.2 to the following.

Theorem 1.3 *For any (a,b)-biregular graph G and any system of weights W,*

$$Z^W(G)\leq \prod_{v\in\mathcal{O}_G}\left(Z^{W^v}(K_{a,b})\right)^{\frac{1}{a}}.$$

Taking $a=b=d$,

$$\lambda_{i,v}=\left\{\begin{array}{ll}\lambda_i & \text{if } v\in\mathcal{E}_G\\ \mu_i & \text{if } v\in\mathcal{O}_G\end{array}\right. \quad\text{and}\quad \lambda_{ij,uv}=\left\{\begin{array}{ll}1 & \text{if } ij\in E(H)\\ 0 & \text{otherwise,}\end{array}\right.$$

Theorem 1.3 reduces to Theorem 1.2.

Let us consider an application of Theorem 1.3 to the antiferromagnetic ($\beta>0$) Ising model without external field ($h=0$) on a d-regular, N-vertex bipartite graph G. A trivial lower bound on $Z^{W_{Ising,\beta,h}}(G)$,

$$\exp\left\{\frac{\beta dN}{2}\right\}\leq Z^{W_{Ising,\beta,h}}(G), \tag{1}$$

is obtained by considering the configuration in which one partition class of G is mapped entirely to $+1$ and the other entirely to -1. Applying Theorem 1.3 we obtain as an upper bound

$$\begin{aligned} Z^{W_{Ising,\beta,h}}(G) &\leq Z^{W_{Ising,\beta,h}}(K_{d,d})^{\frac{N}{2d}}\\ &\leq \left(2^{2d}e^{\beta d^2}\right)^{\frac{N}{2d}} &(2)\\ &= 2^N\exp\left\{\frac{\beta dN}{2}\right\}. &(3)\end{aligned}$$

In (2) we are using that there are 2^{2d} possible configurations on $K_{d,d}$ and that each has weight at most $e^{\beta d^2}$. Combining (1) and (3) we obtain the following bounds on the *free-energy* of the Ising model, the quantity $F^{W_{Ising,\beta,h}}(G):=\log(Z^{W_{Ising,\beta,h}}(G))/N$:

$$\frac{\beta d}{2}\leq F^{W_{Ising,\beta,h}}(G)\leq\frac{\beta d}{2}+\ln 2.$$

Note that these bounds are absolute (independent of G and N), and asymptotically tight in the case of a family of graphs satisfying $\beta d=\omega(1)$.

We give the proof of Theorem 1.3 in Section 3. An important tool in the proof is an extension of Theorem 1.1 to the case of list homomorphisms, which we now discuss. Let H and G be non-empty graphs. To each $v\in V(G)$ associate a set $L(v)\subseteq V(H)$ and write $\mathcal{L}(G,H)$ for $\{L(v):$

$v \in V(G)\}$. A *list homomorphism from G to H with list set $\mathcal{L}(G,H)$* is a homomorphism $f \in Hom(G,H)$ satisfying $f(v) \in L(v)$ for all $v \in V(G)$. Write $Hom^{\mathcal{L}(G,H)}(G,H)$ for the set of all list homomorphism from G to H with list set $\mathcal{L}(G,H)$.

The notion of a list homomorphism is a generalization of that of a homomorphism. Indeed, if $L(v) = V(H)$ for all $v \in V(G)$ then $Hom^{\mathcal{L}(G,H)}(G,H)$ is the same as $Hom(G,H)$. List homomorphisms also generalize the well-studied notion of list colourings of a graph (see *e.g.* [3, Chapter 5] for an introduction). Recall that if a list $L(v)$ of potential colours is assigned to each vertex v of a graph G, then a list colouring of G (with list set $\mathcal{L}(G) = \{L(v) : v \in V(G)\}$) is a function $\chi : V(G) \to \cup_{v \in V(G)} L(v)$ satisfying the property that χ is a proper colouring (*i.e.*, $\chi(u) \neq \chi(v)$ for all $uv \in E(G)$) that respects the lists (*i.e.*, $\chi(v) \in L(v)$ for all $v \in V(G)$). The set of list colourings of G with list set $\mathcal{L}(G)$ is exactly the set $Hom^{\mathcal{L}(G)}(G, H_{\mathcal{L}(G)})$ where $H_{\mathcal{L}(G)}$ is the complete loopless graph on vertex set $\cup_{v \in V(G)} L(v)$.

In the discussion that follows we fix an (a,b)-biregular graph G. We also fix H and $\mathcal{L}(G,H)$ and for convenience of notation we often suppress dependence on G and H. For $v \in \mathcal{O}_G$ write $\mathcal{L}^v$ for the list set on $K_{a,b}$ in which each vertex of degree b gets list $L(v)$ and the vertices of degree a get the lists $L(n_1(v)), \ldots, L(n_b(v))$ (each one occurring exactly once) where $\{n_1(v), \ldots, n_b(v)\}$ is the set of neighbours of v. (Recall that $K_{a,b}$ is the complete bipartite graph with a vertices in one partition class and b in the other.) We generalize Theorem 1.1 to the following result, whose proof is given in Section 2.

Theorem 1.4 *For any graph H, any (a,b)-biregular graph G and any list set $\mathcal{L}$,*

$$\left|Hom^{\mathcal{L}}(G,H)\right| \leq \prod_{v \in \mathcal{O}_G} \left(\left|Hom^{\mathcal{L}^v}(K_{a,b},H)\right|\right)^{\frac{1}{a}}.$$

Taking $a = b = d$ and $L(v) = V(H)$ for all $v \in V(G)$, Theorem 1.4 reduces to Theorem 1.1.

Before turning to proofs, we pause to make a conjecture. The point of departure for this note and for [4] is a result of Kahn [5] bounding the number of independent sets in a d-regular, N-vertex bipartite graph G by

$$|\mathcal{I}(G)| \leq |\mathcal{I}(K_{d,d})|^{\frac{N}{2d}}. \tag{4}$$

Kahn conjectured in [5] that for an arbitrary graph G it should hold that

$$|\mathcal{I}(G)| \leq \prod_{uv \in E(G)} |\mathcal{I}(K_{d(u),d(v)})|^{\frac{1}{d(u)d(v)}}. \tag{5}$$

where $d(u)$ denotes the number of neighbours of u in G. Note that (4) is a special case of (5), and that (5), if true, would be tight for any G which is the union of complete bipartite graphs.

At the moment we see no reason not to conjecture the following, which stands in relation to Theorem 1.3 as (5) does to (4).

Conjecture 1.5 *Let G be any graph and W any collection of weights on G. For each $u \in V(G)$ let $\{n_1(u), \ldots, n_{d(u)}(u)\}$ be the set of neighbours of u. For each edge $uv \in E(G)$, label the degree $d(u)$ vertices of $K_{d(u),d(v)}$ by $w_1(u,v), \ldots, w_{d(v)}(u,v)$ and the degree $d(v)$ vertices by $z_1(u,v), \ldots, z_{d(u)}(u,v)$. Let W^{uv} be the collection of weights on $K_{d(u),d(v)}$ given by*

$$\lambda^{u,v}_{i,w_j(u,v)} = \lambda_{i,n_j(v)}, \quad \lambda^{u,v}_{i,z_j(u,v)} = \lambda_{i,n_j(u)} \quad \text{and} \quad \lambda^{u,v}_{ij,w_j(u,v)z_k(u,v)} = \lambda_{ij,n_j(v)n_k(u)}.$$

Then

$$Z^W(G) \leq \prod_{uv \in E(G)} Z^{W^{uv}}(K_{d(u),d(v)})^{\frac{1}{d(u)d(v)}}.$$

Exactly as Theorem 1.3 follows from Theorem 1.4 (as will be described in Section 3), Conjecture 1.5 would follow from the following conjecture concerning list homomorphisms.

Conjecture 1.6 *Let G and H be any graphs and $\mathcal{L}$ any list set. Let $\mathcal{L}^{uv}$ be the list set on $K_{d(u),d(v)}$ given by*

$$L^{u,v}(w_j(u,v)) = L(n_j(v)) \quad \textit{and} \quad L^{u,v}(z_j(u,v)) = L(n_j(u))$$

(with the notation as in Conjecture 1.5). Then

$$|Hom^{\mathcal{L}}(G,H)| \leq \prod_{uv \in E(G)} |Hom^{\mathcal{L}^{uv}}(K_{d(u),d(v)})|^{\frac{1}{d(u)d(v)}}.$$

2 Proof of Theorem 1.4

We derive Theorem 1.4 from the following more general statement.

Theorem 2.1 *Let G be a bipartite graph with partition classes $\mathcal{E}_G$ and $\mathcal{O}_G$, H an arbitrary graph and $\mathcal{L} = \mathcal{L}(G,H)$ a list set. Suppose that there is m, t_1 and t_2 and families $\mathcal{A} = \{A_i : 1 \leq i \leq m\}$ and $\mathcal{B} = \{B_i : 1 \leq i \leq m\}$ with each $A_i \subseteq \mathcal{E}_G$ and each $B_i \subseteq \mathcal{O}_G$ such that each $v \in \mathcal{E}_G$ is contained in at least t_1 members of $\mathcal{A}$ and each $u \in \mathcal{O}_G$ is contained in at least t_2 members of $\mathcal{B}$. Then*

$$|Hom^{\mathcal{L}}(G,H)| \leq \prod_{i=1}^{m} \left(\sum_{\underline{x} \in \prod_{v \in A_i} L(v)} |C^{\underline{x}}(A_i,B_i)|^{\frac{t_1}{t_2}} \right)^{\frac{1}{t_1}}$$

where, for each $1 \leq i \leq m$ and each $\underline{x} \in \prod_{v \in A_i} L(v)$,

$$C^{\underline{x}}(A_i,B_i) = \left\{ f : B_i \to V(H) : \begin{array}{l} \forall\, v \in B_i,\ f(v) \in L(v) \textit{ and} \\ \forall\, v \in B_i,\ u \in A_i \textit{ with } uv \in E(G),\ (\underline{x})_u f(v) \in E(H) \end{array} \right\}$$

is the set of extensions of the partial list homomorphism $\underline{x}$ on A_i to a partial list homomorphism on $A_i \cup B_i$.

To obtain Theorem 1.4 from Theorem 2.1 we take $\mathcal{A} = \{N(v) : v \in \mathcal{O}_G\}$ and $\mathcal{B} = \{\{v\} : v \in \mathcal{O}_G\}$ where $N(v) = \{u \in V(G) : uv \in E(G)\}$ so that $t_1 = a$ and $t_2 = 1$, and note that in this case $\sum_{\underline{x} \in \prod_{u \in A_v} L(u)} |C^{\underline{x}}(N(v),\{v\})|^a$ is precisely $|Hom^{\mathcal{L}^v}(K_{a,b},H)|$.

The proof of Theorem 2.1 uses entropy considerations, which for completeness we very briefly review here. Our treatment is mostly copied from [5]. For a more thorough discussion, see *e.g.* [6]. In what follows $\mathbf{X}$, $\mathbf{Y}$ etc. are discrete random variables, which in our usage are allowed to take values in any finite set.

The *entropy* of $\mathbf{X}$ is

$$H(\mathbf{X}) = \sum_x p(x) \log \frac{1}{p(x)},$$

where we write $p(x)$ for $\mathbb{P}(\mathbf{X} = x)$ (and extend this convention in natural ways below). The *conditional entropy* of $\mathbf{X}$ given $\mathbf{Y}$ is

$$H(\mathbf{X}|\mathbf{Y}) = \mathbb{E}H(\mathbf{X}|\{\mathbf{Y} = y\}) = \sum_y p(y) \sum_x p(x|y) \log \frac{1}{p(x|y)}.$$

Notice that we are also writing $H(\mathbf{X}|Q)$ with Q an event (in this case $Q = \{\mathbf{Y} = y\}$):

$$H(\mathbf{X}|Q) = \sum p(x|Q) \log \frac{1}{p(x|Q)}.$$

We have the inequalities

$$H(\mathbf{X}) \leq \log |\mathrm{range}(\mathbf{X})| \quad \text{(with equality if } \mathbf{X} \text{ is uniform)},$$

$$H(\mathbf{X}|\mathbf{Y}) \leq H(\mathbf{X}),$$

and more generally,

$$\text{if } \mathbf{Y} \text{ determines } \mathbf{Z} \text{ then } H(\mathbf{X}|\mathbf{Y}) \leq H(\mathbf{X}|\mathbf{Z}). \tag{6}$$

For a random vector $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$ there is a chain rule

$$H(\mathbf{X}) = H(\mathbf{X}_1) + H(\mathbf{X}_2|\mathbf{X}_1) + \cdots + H(\mathbf{X}_n|\mathbf{X}_1, \ldots, \mathbf{X}_{n-1}). \tag{7}$$

Note that (6) and (7) imply

$$H(\mathbf{X}_1, \ldots, \mathbf{X}_n) \leq \sum H(\mathbf{X}_i) \tag{8}$$

We also have a conditional version of (8):

$$H(\mathbf{X}_1, \ldots, \mathbf{X}_n|\mathbf{Y}) \leq \sum H(\mathbf{X}_i|\mathbf{Y}).$$

Finally we use a lemma of Shearer (see [2, p. 33]). For a random vector $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_m)$ and $A \subseteq \{1, \ldots, m\}$, set $\mathbf{X}_A = (\mathbf{X}_i : i \in A)$.

Lemma 2.2 *Let $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_m)$ be a random vector, $\mathbf{Y}$ a random variable and $\mathcal{A}$ a collection of subsets (possibly with repeats) of $\{1, \ldots, m\}$, with each element of $\{1, \ldots, m\}$ contained in at least t members of $\mathcal{A}$. Then*

$$H(\mathbf{X}) \leq \frac{1}{t} \sum_{A \in \mathcal{A}} H(\mathbf{X}_A) \quad \textit{and} \quad H(\mathbf{X}|\mathbf{Y}) \leq \frac{1}{t} \sum_{A \in \mathcal{A}} H(\mathbf{X}_A|\mathbf{Y}).$$

Proof of Theorem 2.1: We follow closely the proof of [4, Lemma 3.1]. Let $\mathbf{f}$ be a uniformly chosen member of $Hom^{\mathcal{L}}(G,H)$. For each $1 \leq i \leq m$ and each $\underline{x} \in \prod_{v \in A_i} L(v)$ let $p_i(\underline{x})$ be the probability that $\mathbf{f}$ restricted to A_i is $\underline{x}$. With the key inequalities justified below (the remaining steps follow in a straightforward way from the properties of entropy just established) we have

$$\begin{aligned}
H(\mathbf{f}) &= H(\mathbf{f}|_{\mathcal{E}_G}) + H(\mathbf{f}|_{\mathcal{O}_G} \mid \mathbf{f}|_{\mathcal{E}_G}) \\
&\leq \frac{1}{t_1}\sum_{i=1}^m H(\mathbf{f}|_{A_i}) + \frac{1}{t_2}\sum_{i=1}^m H(\mathbf{f}|_{B_i} \mid \mathbf{f}|_{\mathcal{E}_G}) && (9)\\
&\leq \frac{1}{t_1}\sum_{i=1}^m \left(H(\mathbf{f}|_{A_i}) + \frac{t_1}{t_2}H(\mathbf{f}|_{B_i} \mid \mathbf{f}|_{A_i})\right) && (10)\\
&= \frac{1}{t_1}\sum_{i=1}^m \sum_{\underline{x}\in\prod_{v\in A_i} L(v)} \left(p_i(\underline{x})\log\frac{1}{p_i(\underline{x})} + \frac{t_1}{t_2}p_i(\underline{x})H(\mathbf{f}|_{B_i} \mid \{\mathbf{f}|_{A_i} = \underline{x}\})\right) \\
&\leq \frac{1}{t_1}\sum_{i=1}^m \sum_{\underline{x}\in\prod_{v\in A_i} L(v)} p_i(\underline{x})\log\frac{|C^{\underline{x}}(A_i,B_i)|^{\frac{t_1}{t_2}}}{p_i(\underline{x})} \\
&\leq \frac{1}{t_1}\sum_{i=1}^m \log\left(\sum_{\underline{x}\in\prod_{v\in A_i} L(v)} |C^{\underline{x}}(A_i,B_i)|^{\frac{t_1}{t_2}}\right). && (11)
\end{aligned}$$

In (9) we use Shearer's Lemma twice, once with $\mathcal{A}$ as the covering family and once with $\mathcal{B}$, and in (11) we use Jensen's inequality. In (10) we would have equality if it happened that for each i, A_i included all the neighbours of B_i, since $\mathbf{f}|_{B_i}$ depends only on the values of $\mathbf{f}$ on B_i's neighbours. It is easy, however, to construct examples where $H(\mathbf{f}|_{B_i} \mid \mathbf{f}|_{\mathcal{E}_G}) < H(\mathbf{f}|_{B_i} \mid \mathbf{f}|_{A_i})$ when A_i does not include all the neighbours of B_i.

The theorem now follows from the equality $H(\mathbf{f}) = \log|Hom^{\mathcal{L}}(G,H)|$. $\square$

3 Proof of Theorem 1.3

By continuity we may assume that all weights are rational and non-zero. By scaling appropriately we may also assume that $0 < \lambda_{ij,uv} \leq 1$ for all i, j and $uv \in E(G)$ (we will later think of the $\lambda_{ij,uv}$'s as probabilities).

Set $N = |V(G)|$ and

$$\lambda_{vmin} = \min_{i,w}\lambda_{i,w}, \quad \lambda_{vmax} = \max_{i,w}\lambda_{i,w} \quad \text{and} \quad \lambda_{emin} = \min_{ij,vw}\lambda_{ij,vw}.$$

Also, set $w_{min} = \lambda_{vmin}^N \lambda_{emin}^{abN/(a+b)}$; this is a lower bound on $w^W(f)$ for all $f : V(G) \to \{1,\ldots,m\}$ (observe that an (a,b)-biregular graph G on N vertices has $|\mathcal{E}_G| = bN/(a+b)$, $|\mathcal{O}_G| = aN/(a+b)$ and $|E(G)| = abN/(a+b)$) as well as a lower bound on $w^{W^v}(f)$ for all $v \in \mathcal{O}_G$ and all $f : V(K_{a,b}) \to \{1,\ldots,m\}$.

Choose $C \geq 1$ large enough that $C\lambda_{i,v} \in \mathbb{N}$ for all $1 \leq i \leq m$ and $v \in V(G)$. For each i and v let $S_{i,v}$ be a set of size $C\lambda_{i,v}$, with all the $S_{i,v}$'s disjoint. Let H be the graph on vertex set $\cup_{i,v} S_{i,v}$ with $xy \in E(H)$ iff $x \in S_{i,v}$ and $y \in S_{j,w}$ for some i, j, v, w with $vw \in E(G)$. For each $v \in V(G)$ let $L(v) = \cup_i S_{i,v}$ and set $\mathcal{L} = \{L(v) : v \in V(G)\}$. For each $g : V(G) \to \{1, \ldots, m\}$ and each subgraph $\widetilde{H}$ of H (on the same vertex set as H) set

$$\mathcal{H}_g(G, \widetilde{H}) = \{f \in Hom^{\mathcal{L}}(G, \widetilde{H}) : f(v) \in S_{g(v),v} \text{ for all } v \in V(G)\}.$$

Note that $\mathcal{H}_g(G, H)$ is exactly $\{f : V(G) \to V(H) : f(v) \in S_{g(v),v}$ for all $v \in V(G)\}$ and so $|\mathcal{H}_g(G, H)| = C^N \prod_{v \in V(G)} \lambda_{g(v),v}$. Note also that for $g \neq g'$ we have $\mathcal{H}_g(G, \widetilde{H}) \cap \mathcal{H}_{g'}(G, \widetilde{H}) = \emptyset$ and that $Hom^{\mathcal{L}}(G, \widetilde{H}) = \cup_g \mathcal{H}_g(G, \widetilde{H})$.

For each $v \in \mathcal{O}_G$, each $g : V(K_{a,b}) \to \{1, \ldots, m\}$ and each $\widetilde{H}$, set

$$\mathcal{H}^v_g(K_{a,b}, \widetilde{H}) = \left\{ f \in Hom^{\mathcal{L}^v}(K_{a,b}, \widetilde{H}) : \begin{array}{ll} f(w_k) \in S_{g(w_k),n_k(v)}, & 1 \leq k \leq b \\ f(z_k) \in S_{g(z_k),v}, & 1 \leq k \leq a \end{array} \right\},$$

where the notation is as established before the statements of Theorems 1.3 and 1.4. Note that for $g \neq g'$ we have $\mathcal{H}^v_g(K_{a,b}, \widetilde{H}) \cap \mathcal{H}^v_{g'}(K_{a,b}, \widetilde{H}) = \emptyset$ and that $Hom^{\mathcal{L}^v}(K_{a,b}, \widetilde{H}) = \cup_g \mathcal{H}^v_g(K_{a,b}, \widetilde{H})$.

We will exhibit a subgraph $\widetilde{H}$ of H which satisfies

$$\left| C^N w^W(g) - |\mathcal{H}_g(G, \widetilde{H})| \right| \leq \delta(C) |\mathcal{H}_g(G, \widetilde{H})| \tag{12}$$

for all $g : V(G) \to \{1, \ldots, m\}$ and

$$\left| C^{a+b} w^{W^v}(g) - |\mathcal{H}^v_g(K_{a,b}, \widetilde{H})| \right| \leq \delta(C) |\mathcal{H}^v_g(K_{a,b}, \widetilde{H})| \tag{13}$$

for all $v \in \mathcal{O}_G$ and $g : V(K_{a,b}) \to \{1, \ldots, m\}$, where $\delta(C)$ depends also on N, a, b and W and tends to 0 as C tends to infinity (with N, a, b and W fixed). This suffices to prove the theorem, for we have

$$\begin{aligned} \left| C^N Z^W(G) - |Hom^{\mathcal{L}}(G, \widetilde{H})| \right| &\leq \sum_g \left| C^N w^W(g) - |\mathcal{H}_g(G, \widetilde{H})| \right| \\ &\leq \delta(C) \sum_g |\mathcal{H}_g(G, \widetilde{H})| \\ &= \delta(C) |Hom^{\mathcal{L}}(G, \widetilde{H})| \end{aligned}$$

and similarly, for each $v \in \mathcal{O}_G$,

$$\left| C^{a+b} Z^{W^v}(K_{a,b}) - |Hom^{\mathcal{L}^v}(K_{a,b}, \widetilde{H})| \right| \leq \delta(C) |Hom^{\mathcal{L}^v}(K_{a,b}, \widetilde{H})| \tag{14}$$

and so

$$\begin{aligned} C^N Z^W(G) &\leq (1 + \delta(C)) |Hom^{\mathcal{L}}(G, \widetilde{H})| \\ &\leq (1 + \delta(C)) \prod_{v \in \mathcal{O}_G} \left(|Hom^{\mathcal{L}^v}(K_{a,b}, \widetilde{H})| \right)^{\frac{1}{a}} \qquad (15) \\ &\leq C^N \frac{1 + \delta(C)}{(1 - \delta(C))^{\frac{N}{a+b}}} \prod_{v \in \mathcal{O}_G} \left(Z^{W^v}(K_{a,b}) \right)^{\frac{1}{a}}. \qquad (16) \end{aligned}$$

In (15) we use Theorem 1.4 while in (16) we use (14). Theorem 1.3 follows since the constant in front of the product in (16) can be made arbitrarily close to 1 (with N, a, b and W fixed) by choosing C sufficiently large.

The graph $\widetilde{H}$ will be a random graph defined as follows. For each $xy \in E(H)$ with $x \in S_{i,v}$ and $y \in S_{j,w}$ we put $xy \in E(\widetilde{H})$ with probability $\lambda_{ij,uv}$, all choices independent. The proofs of (12) and (13) involve a second moment calculation. For each $f \in \mathcal{H}_g(G,H)$, set $\mathbf{X}_f = \mathbf{1}_{\{f\in\mathcal{H}_g(G,\widetilde{H})\}}$ and $\mathbf{X} = \sum_{f\in\mathcal{H}_g(G,H)} \mathbf{X}_f$. Note that $\mathbf{X} = |\mathcal{H}_g(G,\widetilde{H})|$. For each $f \in \mathcal{H}_g(G,H)$ we have

$$\begin{aligned}
\mathbb{E}(\mathbf{X}_f) &= \mathbb{P}(f \in \mathcal{H}_g(G,\widetilde{H})) \\
&= \mathbb{P}(\{f(u)f(v) \in E(\widetilde{H})\ \forall uv \in E(G)\}) \\
&= \prod_{uv\in E(G)} \lambda_{g(u)g(v),uv}, \qquad (17)
\end{aligned}$$

with (17) following from the fact that $\{f(u)f(v) : uv \in E(G)\}$ is a collection of disjoint edges and so $\{\{f(u)f(v) \in E(\widetilde{H})\} : uv \in E(G)\}$ is a collection of independent events. By linearity of expectation we therefore have

$$\mathbb{E}(\mathbf{X}) = |\mathcal{H}_g(G,H)| \prod_{uv\in E(G)} \lambda_{g(u)g(v),uv} = C^N w^W(g) := \mu. \qquad (18)$$

We now consider the second moment. For $f, f' \in \mathcal{H}_g(G,H)$ write $f \sim f'$ if there is $uv \in E(G)$ with $f(u) = f'(u)$ and $f(v) = f'(v)$. Note that $\mathbf{X}_f$ and $\mathbf{X}_{f'}$ are not independent iff $f \sim f'$. By standard methods (see *e.g.* [1]) we have

$$\begin{aligned}
\mathrm{Var}(\mathbf{X}) &\le \mu + \sum_{(f,f')\in\mathcal{H}_g(G,H)^2\,:\, f\sim f'} \mathbb{P}(\{f \in \mathcal{H}_g(G,\widetilde{H})\} \wedge \{f' \in \mathcal{H}_g(G,\widetilde{H})\}) \\
&\le \mu + |\{(f,f') \in \mathcal{H}_g(G,H)^2 : f \sim f'\}|.
\end{aligned}$$

To estimate $|\{(f,f') \in \mathcal{H}_g(G,H)^2 : f \sim f'\}|$ note that there are $|\mathcal{H}_g(G,H)|$ choices for f, at most N^2 choices for a $uv \in E(G)$ on which f and f' agree, and finally at most

$$\frac{|\mathcal{H}_g(G,H)|}{C\lambda_{g(u),u}C\lambda_{g(v),v}} \le \frac{|\mathcal{H}_g(G,H)|}{C^2\lambda_{vmin}^2}$$

choices for the rest of f'. We therefore have

$$\begin{aligned}
\frac{\mathrm{Var}(\mathbf{X})}{\mu^2} &\le \frac{1}{\mu} + \frac{|\mathcal{H}_g(G,H)|^2N^2}{\mu^2C^2\lambda_{vmin}^2} \\
&= \frac{1}{C^2}\left(\frac{1}{C^{N-2}w^W(g)} + \frac{N^2}{\lambda_{vmin}^2\prod_{uv\in E(G)}\lambda_{g(u)g(v),uv}}\right) \\
&\le \frac{1}{C^2}\left(\frac{1}{w_{min}} + \frac{\lambda_{vmax}^N N^2}{\lambda_{vmin}^2 w_{min}}\right) \qquad (19) \\
&\le \frac{\alpha(N,a,b,W)}{C^2}
\end{aligned}$$

for some function α (independent of G and g). In (19) we use (18) and the fact that $N \geq 2$ (which holds since G is non-empty). By Tchebychev's inequality, we therefore have

$$\mathbb{P}\left(|\mu - \mathbf{X}| > \mu\sqrt{\frac{\alpha}{C}}\right) \leq \frac{1}{C}.$$

It follows that the probability that $\widetilde{H}$ fails to satisfy

$$\left|\prod_{uv\in E(G)} \lambda_{g(u)g(v),uv} - \frac{|\mathcal{H}_g(G,\widetilde{H})|}{|\mathcal{H}_g(G,H)|}\right| \leq \sqrt{\frac{\alpha}{C}} \prod_{uv\in E(G)} \lambda_{g(u)g(v),uv} \tag{20}$$

for a particular g is at most $1/C$, and so the probability that it fails to satisfy (20) for any g is at most m^N/C.

A similar argument gives that for a particular $v \in \mathcal{O}_G$ and $g : V(K_{a,b}) \to \{1,\ldots,m\}$ the probability that $\widetilde{H}$ fails to satisfy

$$\left|\prod_{w_jz_k\in E(K_{a,b})} \lambda_{g(w_j)g(z_k),n_k(v)v} - \frac{|\mathcal{H}^v_g(K_{a,b},\widetilde{H})|}{|\mathcal{H}^v_g(K_{a,b},H)|}\right| \leq \sqrt{\frac{\alpha}{C}} \prod_{w_jz_k\in E(K_{a,b})} \lambda_{g(w_j)g(z_k),n_k(v)v} \tag{21}$$

is at most $1/C$, and so the probability that it fails to satisfy (21) for any g is at most m^{a+b}/C.

As long as $C > m^N + aNm^{a+b}/(a+b)$ there is therefore an $\widetilde{H}$ for which (12) is satisfied for each $g : V(G) \to \{1,\ldots,m\}$ and (13) is satisfied for each $v \in \mathcal{O}_G$ and $g : V(K_{a,b}) \to \{1,\ldots,m\}$ with

$$\delta(C) = \frac{\sqrt{\alpha}}{\sqrt{C} - \sqrt{\alpha}}.$$

Since $\delta(C) \to 0$ as $C \to \infty$, we are done.

Acknowledgment We are grateful to Alex Scott [7] for suggesting the construction of $\widetilde{H}$.

References

[1] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, New York, 2000.

[2] F.R.K. Chung, P. Frankl, R. Graham and J.B. Shearer, Some intersection theorems for ordered sets and graphs, *J. Combin. Theory Ser. A.* **48** (1986), 23–37.

[3] R. Diestel, *Graph Theory*, Springer, New York, 1997.

[4] D. Galvin and P. Tetali, On weighted graph homomorphisms, DIMACS Series in Discrete Mathematics and Theoretical Computer Science **63** (2004) *Graphs, Morphisms and Statistical Physics*, 97–104.

[5] J. Kahn, An entropy approach to the hard-core model on bipartite graphs, *Combin. Prob. Comp.* **10** (2001), 219–237.

[6] R.J. McEliece, *The Theory of Information and Coding*, Addison-Wesley, London, 1977.

[7] A. Scott, personal communication.