H-colouring bipartite graphs

John Engbers   David Galvin

June 14, 2012[*]

Abstract

For graphs G and H, an H-*colouring* of G (or *homomorphism* from G to H) is a function from the vertices of G to the vertices of H that preserves adjacency. H-colourings generalize such graph theory notions as proper colourings and independent sets.

For a given H, $k \in V(H)$ and G we consider the proportion of vertices of G that get mapped to k in a uniformly chosen H-colouring of G. Our main result concerns this quantity when G is regular and bipartite. We find numbers $0 \leq a^-(k) \leq a^+(k) \leq 1$ with the property that for all such G, with high probability the proportion is between $a^-(k)$ and $a^+(k)$, and we give examples where these extremes are achieved. For many H we have $a^-(k) = a^+(k)$ for all k and so in these cases we obtain a quite precise description of the almost sure appearance of a randomly chosen H-colouring.

As a corollary, we show that in a uniform proper q-colouring of a regular bipartite graph, if q is even then with high probability every colour appears on a proportion close to $1/q$ of the vertices, while if q is odd then with high probability every colour appears on at least a proportion close to $1/(q + 1)$ of the vertices and at most a proportion close to $1/(q - 1)$ of the vertices.

Our results generalize to natural models of weighted H-colourings, and also to bipartite graphs which are sufficiently close to regular. As an application of this latter extension we describe the typical structure of H-colourings of graphs which are obtained from n-regular bipartite

graphs by percolation, and we show that $p = 1/n$ is a threshold function across which the typical structure changes.

The approach is through entropy, and extends work of J. Kahn, who considered the size of a randomly chosen independent set of a regular bipartite graph.

1 Introduction and statement of results

Let $G = (V(G), E(G))$ be a simple, loopless, finite graph, and let $H = (V(H), E(H))$ be a finite graph without multiple edges but perhaps with loops. An H-*colouring* of G, or *homomorphism* from G to H, is a function from $V(G)$ to $V(H)$ that preserves adjacency. The set of H-colourings of G is thus

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

H-colourings generalize a number of important graph theory notions. For example, when H is the complete graph on q vertices, $\mathrm{Hom}(G, H)$ coincides with the set of proper q-colourings of G, and when H consists of two vertices joined by an edge, with a loop at one of the vertices, then $\mathrm{Hom}(G, H)$ may be identified with the set of independent sets of G, via the preimage of the unlooped vertex.

H-colourings have a natural statistical physics interpretation as configurations in *hard-constraint spin models*. Here, the vertices of G are thought of as sites that are occupied by particles, with edges of G representing pairs of bonded sites. The vertices of H are the different types of particles (or spins), and the occupation rule is that bonded sites must be occupied by pairs of particles that are adjacent in H. A legal configuration in such a spin model is exactly an H-colouring of G.

From the statistical physics standpoint, there is a very natural family of probability distributions that can be put on $\mathrm{Hom}(G, H)$. Fix a set of positive weights $\Lambda = \{\lambda_i : i \in V(H)\}$ indexed by the vertices of H. We think of the magnitude of λ_k as measuring how likely particle k is to appear at each site. This can be formalized by giving each $f \in \mathrm{Hom}(G, H)$ weight $w_\Lambda(f) = \prod_{v \in V(G)} \lambda_{f(v)}$ and probability

$$p_\Lambda(f) = \frac{w_\Lambda(f)}{Z_\Lambda(G, H)}$$

where $Z_\Lambda(G, H) = \sum_{f \in \mathrm{Hom}(G,H)} w_\Lambda(f)$ is the appropriate normalizing constant or *partition function* of the model. For an introduction to statistical physics spin models from a combinatorial perspective, see for example [3].

The question to be addressed in this paper is the following. What can be said about an f that is drawn from $\mathrm{Hom}(G, H)$ according to the distribution p_Λ? Specifically, for each $f \in \mathrm{Hom}(G, H)$ and $k \in V(H)$ set

$$s(k, f) = \frac{|f^{-1}(k)|}{|V(G)|}$$

and

$$\bar{p}_\Lambda(k) = \frac{1}{|V(G)|} \sum_{v \in V(G)} p_\Lambda(f(v) = k) \ \ (= E_\Lambda\left(s(k, f)\right)).$$

The aim of this paper is to give fairly precise estimates for $\bar{p}_\Lambda(k)$ and the distribution of $s(k, f)$ for f chosen according to p_Λ, when G is bipartite and either regular or sufficiently close to regular.

The point of departure for this work is a result of Kahn on the hard-core model. When $H = H_{\mathrm{ind}}$ with $V(H_{\mathrm{ind}}) = \{0, 1\}$ and $E(H_{\mathrm{ind}}) = \{00, 01\}$, the set of vertices of G mapped to 1 forms an independent set in G, and $\mathrm{Hom}(G, H_{\mathrm{ind}})$ can be identified with $\mathcal{I}(G)$, the set of independent sets in G. For each $\lambda > 0$, the *hard-core* model on G is the probability distribution $\mathrm{hc}(\lambda)$ on $\mathcal{I}(G)$ that assigns to each $I \in \mathcal{I}(G)$ a probability proportional to $\lambda^{|I|}$. One of the oldest and most studied spin models in statistical physics, this is a simple mathematical model of the occupation of space (represented by G) by particles of non-negligible size. The model can easily be realized as a spin model of the kind described above by assigning weights $\lambda_0 = 1$ and $\lambda_1 = \lambda$ to the vertices of H_{ind}.

Kahn [7] studied this model on a regular bipartite graph G. He proved that for all fixed $\lambda > 0$, the model exhibits a *phase coexistence* in the sense that if G has equipartition $\mathcal{E} \cup \mathcal{O}$ then most $\mathrm{hc}(\lambda)$ independent sets tend to come either mostly from $\mathcal{E}$ or mostly from $\mathcal{O}$, in the sense that the size of an independent set chosen according to $\mathrm{hc}(\lambda)$ is concentrated close to $\lambda/(2(1 + \lambda))$, which is exactly the expected size of an independent set chosen according to the distribution that half the time picks a $\mathrm{hc}(\lambda)$ independent set from $\mathcal{E}$ and half the time picks from $\mathcal{O}$. The following theorem ([7, Theorem 1.4 & Corollary 1.5]) formalizes this.

Theorem 1.1 *Let $\lambda > 0$ be fixed. There are positive constants c_1, c_2, c_3 and c_4 (depending on λ) such that for every d-regular bipartite graph G on N vertices, the following two statements hold. Firstly, for every $\varepsilon \geq c_1/\sqrt{d}$, if I is chosen from $\mathcal{I}(G)$ according to the distribution $\mathrm{hc}(\lambda)$ then*

$$\Pr\left(\left||I| - \frac{\lambda N}{2(1+\lambda)}\right| \geq \varepsilon N\right) \leq c_2 \varepsilon^{-1} 2^{-c_3 \varepsilon^2 N}.$$

Secondly,

$$\left|\frac{E(|I|)}{N} - \frac{\lambda}{2(1+\lambda)}\right| \leq c_4 \zeta$$

where

$$\zeta = \max\left\{\frac{1}{\sqrt{d}}, \sqrt{\frac{\log N}{N}}\right\}. \tag{1}$$

In particular, a uniformly chosen independent set ($\lambda = 1$) from a regular bipartite graph consists, with high probability, of close to one quarter of the vertices. While this corollary may seem more natural than the formulation of Theorem 1.1, it is worth noting that in order to prove the theorem in the special case of $\lambda = 1$ it is necessary (at least using the entropy methods of [7]) to pass to the more general weighted model first. Similarly, it might seem more natural in the present paper to focus on the structure of uniform H-colourings, but we are unable to obtain any results without introducing weights.

From (1) we see that Theorem 1.1 only gives a concentration result when we consider families of graphs with d going to infinity. This is not just an artifact of the proof. For families of graphs with d fixed (and only N going to infinity), the behavior of $E(|I|)/N$ depends very much on the particular choice of family. As an example, consider the case $d = 2$. If G_N is the disjoint union of $N/4$ copies of the cycle C_4, and I is chosen uniformly from $\mathcal{I}(G)$, then $E(|I|)/N$ is easily seen to be concentrated close to $2/7$. If, however, G_N is the disjoint union of $N/6$ copies of the cycle C_6, then $E(|I|)/N$ is concentrated close to $5/18$. For this reason we implicitly assume throughout that d is going to infinity.

We now set up some notation that allows us to state our main result, which is an extension of Theorem 1.1 to arbitrary weighted H-colourings. From now on, whenever H and Λ are mentioned, it will be assumed that H is a finite graph without multiple edges but perhaps with loops, and that Λ

is a set of positive weights indexed by the vertices of H. For $A, B \subseteq V(H)$ write $A \sim B$ if for all $u \in A$ and $v \in B$ we have $uv \in E(H)$, and set

$$\eta_\Lambda(H) = \max\{w_\Lambda(A)w_\Lambda(B) : A \sim B\}$$

where $w_\Lambda(\cdot) = \sum_{i \in \cdot} \lambda_i$. Then set

$$\mathcal{M}_\Lambda(H) = \{(A, B) \in V(H)^2 : A \sim B, \ w_\Lambda(A)w_\Lambda(B) = \eta_\Lambda(H)\}.$$

Next define

$$a_\Lambda^+(k) = \frac{\max\{w_\Lambda(A)\lambda_k \mathbf{1}_{\{k \in B\}} + w_\Lambda(B)\lambda_k \mathbf{1}_{\{k \in A\}} : (A, B) \in \mathcal{M}_\Lambda(H)\}}{2\eta_\Lambda(H)}$$

and define $a_\Lambda^-(k)$ similarly, with max replaced by min. (After the statement of Theorem 1.4, we will give some explicit examples to illuminate these definitions.) Note that if k does not appear in any $(A, B) \in \mathcal{M}_\Lambda(H)$ then $a_\Lambda^+(k) = 0$ and that if there is a pair $(A, B) \in \mathcal{M}_\Lambda(H)$ in which k does not appear then $a_\Lambda^-(k) = 0$. Note also that $a_\Lambda^-(k) \le a_\Lambda^+(k)$. Finally, note that $a_\Lambda^+(k)$ and $a_\Lambda^-(k)$ both take the form

$$\frac{\lambda_k \mathbf{1}_{\{k \in A\}}}{2w_\Lambda(A)} + \frac{\lambda_k \mathbf{1}_{\{k \in B\}}}{2w_\Lambda(B)}$$

for some $(A, B) \in \mathcal{M}_\Lambda(H)$. We may interpret this quantity as the expected proportion of vertices mapped to k in a p_Λ-chosen H-colouring subject to the condition that all vertices from one partition class of G get mapped to A and all from the other class get mapped to B; we will refer to such a colouring as a *pure-(A, B)* colouring. Finally, for every $\varepsilon > 0$ and $k \in V(H)$ define

$$I_k(\varepsilon) = [0, a_\Lambda^-(k) - \varepsilon) \cup (a_\Lambda^+(k) + \varepsilon, 1].$$

Before stating our main result, we motivate it by considering weighted H-colourings of $K_{d,d}$, the complete bipartite graph with d vertices in each partition class, for some fixed H and Λ. The adjacency structure of $K_{d,d}$ ensures that all H-colourings are pure-(A, B) for some (A, B) with $A \sim B$, and that moreover all but a vanishing proportion (in d) of $Z_\Lambda(K_{d,d}, H)$ comes from pure-(A, B) colourings for some $(A, B) \in \mathcal{M}_\Lambda(H)$. It follows that for each $k \in V(H)$, in an H-colouring chosen according to p_Λ we have that with probability $1 - o(1)$ the proportion of vertices of $K_{d,d}$ mapped to k will be between $a_\Lambda^-(k) - o(1)$ and $a_\Lambda^+(k) + o(1)$. Our main result, which we now state, asserts that this property of $K_{d,d}$ is essentially shared by all d-regular graphs.

Theorem 1.2 *Fix H and Λ. There are positive constants c_1, c_2, c_3 and c_4 (depending on H and Λ) such that for every d-regular bipartite graph G on N vertices, the following two statements hold. Firstly, for every $\varepsilon \geq c_1/\sqrt{d}$ and $k \in V(H)$ we have*

$$p_\Lambda\left(s(k,f) \in I_k(\varepsilon)\right) \leq c_2 \varepsilon^{-1} 2^{-c_3 \varepsilon^2 N}. \tag{2}$$

Secondly, for each $k \in V(H)$ we have

$$\bar{p}_\Lambda(k) \in \left[a_\Lambda^-(k) - c_4\zeta, a_\Lambda^+(k) + c_4\zeta\right] \tag{3}$$

where ζ is as defined in (1).

In other words, for regular bipartite G the distribution p_Λ is concentrated on H-colourings for which, for every $k \in V(H)$, the proportion of vertices mapped to k is roughly between $a_\Lambda^-(k)$ and $a_\Lambda^+(k)$.

The proof of Theorem 1.2 goes along the following lines. We upper bound the contribution to $Z_\Lambda(G,H)$ from those $f \in \mathrm{Hom}(G,H)$ with $|f^{-1}(k)|/N = \gamma \geq a^+(k) + \varepsilon$ by $Z_{\Lambda(k,\delta)}(G,H)/(1+\delta)^{\gamma N}$ for some suitably small $\delta > 0$ (where $\Lambda(k,\delta)$ is obtained from Λ by multiplying λ_k by $1 + \delta$ and leaving all other λ_i unchanged). We in turn upper bound $Z_{\Lambda(k,\delta)}(G,H)$ using a result of Galvin and Tetali [6] to the effect that for all H and Λ and all d-regular bipartite graphs G on N vertices we have

$$Z_\Lambda(G,H) \leq Z_\Lambda(K_{d,d}, H)^{\frac{N}{2d}} \tag{4}$$

(where recall $K_{d,d}$ is the complete bipartite graph with d vertices in each partition class). We upper bound $Z_{\Lambda(k,\delta)}(K_{d,d}, H)$ in terms of $\eta_{\Lambda(k,\delta)}(H)$, and in the end we get, using our choice of $a_\Lambda^+(k)$ and for some sufficiently small δ, an upper bound on the contribution that is significantly smaller than a trivial lower bound on $Z_\Lambda(G,H)$, showing that those $f \in \mathrm{Hom}(G,H)$ with $|f^{-1}(k)|/N \geq a^+(k) + \varepsilon$ do not contribute greatly to the partition function. The same strategy works for $|f^{-1}(k)|/N$ falling significantly below $a^-(k)$. The details (in the more general setting of Theorem 1.6) are given in Section 3.

When $a_\Lambda^-(k) = a_\Lambda^+(k)$ for all k, we obtain a single vector around which $(s(k,f) : k \in V(H))$ is concentrated for f chosen according to p_Λ.

Corollary 1.3 *Fix H and Λ. Suppose that for all $k \in V(H)$ there is an $a_\Lambda(k)$ such that $a_\Lambda^-(k) = a_\Lambda^+(k) = a_\Lambda(k)$. Then there are positive constants c_1, c_2, c_3*

6

and c_4 (depending on H and Λ) such that for every d-regular, bipartite graph G on N vertices the following two statements hold. Firstly, for $\varepsilon \geq c_1/\sqrt{d}$ we have

$$p_\Lambda \left(\left\| (s(k,f))_{k \in V(H)} - (a_\Lambda(k))_{k \in V(H)} \right\|_\infty \geq \varepsilon \right) \leq c_2 \varepsilon^{-1} 2^{-c_3 \varepsilon^2 N}.$$

Secondly, we have

$$\left\| (\bar{p}_\Lambda(k))_{k \in V(H)} - (a_\Lambda(k))_{k \in V(H)} \right\|_\infty \leq c_4 \zeta$$

with ζ as in (1).

A situation in which Corollary 1.3 applies is when either $\mathcal{M}_\Lambda(H) = \{(A,A)\}$ or $\mathcal{M}_\Lambda(H) = \{(A,B),(B,A)\}$ (for some $A \neq B$). This is in a sense the generic situation. Indeed, for every H, if the weights λ_i are chosen from any continuous distribution supported on $\{x \in \mathbb{R}^{|V(H)|} : x > 0\}$, then with probability 1 we will have $\mathcal{M}_\Lambda(H)$ of the form described. As we will see in Example C below, Corollary 1.3 also applies in some other natural situations.

The gap between $a_\Lambda^-(k)$ and $a_\Lambda^+(k)$ (if there is one) cannot be closed in general, as the first part of the following theorem shows.

Theorem 1.4 *Fix H and Λ. There is a family $\{G_d\}_{d=1}^\infty$ of d-regular bipartite graphs, a function $g(d) = o(1)$ and a positive constant c (depending on H and Λ) such that for each $k \in V(H)$,*

$$\left. \begin{array}{c} p_\Lambda \left(\left| s(k,f) - a_\Lambda^+(k) \right| \leq g(d) \right) \\ p_\Lambda \left(\left| s(k,f) - a_\Lambda^-(k) \right| \leq g(d) \right) \end{array} \right\} \geq c - g(d).$$

There is also a family $\{G_d'\}_{d=1}^\infty$ of d-regular bipartite graphs, a function $g(d) = o(1)$ and (for each $k \in V(H)$) an $a_\Lambda(k)$ satisfying $a_\Lambda^-(k) \leq a_\Lambda(k) \leq a_\Lambda^+(k)$ such that for each k,

$$p_\Lambda \left(|s(k,f) - a_\Lambda(k)| \leq g(d) \right) \geq 1 - g(d)$$

and

$$|\bar{p}_\Lambda(k) - a_\Lambda(k)| \leq g(d).$$

We prove Theorem 1.4 in Section 5. The graphs G_d we exhibit will be suitably chosen random regular graphs, and we will use the expansion of these graphs to show that all but $o(1)$ of p_Λ is concentrated on pure-(A,B) colourings

7

for $(A, B) \in \mathcal{M}_\Lambda(H)$. The graphs G'_d will be disjoint unions of complete bipartite graphs on $2d$ vertices. Basic concentration estimates together with the independence of the components will give the claimed result.

We now explore the consequences of Theorem 1.2 for some specific choices of H and Λ.

Example A (Hard-core model) Let $H = H_{\text{ind}}$ be as described earlier, with $\lambda_0 = 1$ and $\lambda_1 = \lambda$. We have seen that an element of $\text{Hom}(G, H_{\text{ind}})$ chosen according to p_Λ is a configuration in the hard-core model on G with activity λ. With these choices we have $\mathcal{M}_\Lambda(H_{\text{ind}}) = \{(\{0\}, \{0, 1\}), (\{0, 1\}, \{0\})\}$ and

$$a_\Lambda^-(1) = a_\Lambda^+(1) = \frac{\lambda}{2(1 + \lambda)}$$

and so Theorem 1.2 indeed generalizes Theorem 1.1, as claimed.

Example B (Multistate hard-core model) Let $H = H_k$ be the graph on vertex set $\{0, \ldots, k\}$ with $ij \in E(H)$ if and only if $i + j \leq k$, and $\lambda_i = \lambda^i$ for some fixed $\lambda > 0$. An element of $\text{Hom}(G, H_k)$ chosen according to p_Λ is exactly a configuration of the *multistate hard-core* (or *multicast communications*) model on G with activity λ. This model allows multiple particles (up to and including k) at each site, with the restriction that there are no more than k particles in total across each edge. A generalization of the hard-core model (the case $k = 1$), it has been studied in a variety of contexts: in communications [12], statistical physics [10] and combinatorics [5]. For k even the unique pair $(A, B) \in \mathcal{M}_\Lambda(H_k)$ has $A = B = \{1, \ldots, k/2\}$, while for k odd, say $k = 2\ell + 1$, we have $\mathcal{M}_\Lambda(H_k) = \{(A, B), (B, A)\}$ with $A = \{1, \ldots, \ell\}$ and $B = \{1, \ldots, \ell + 1\}$. In either case Corollary 1.3 shows that for this model $(s(k, f) : k \in V(H))$ is concentrated close to a single value for f chosen according to p_Λ.

Example C (Uniform proper q-colourings) Let $H = K_q$, the complete graph on q vertices, and $\Lambda = (1, \ldots, 1)$. An element of $\text{Hom}(G, K_q)$ chosen according to p_Λ corresponds to a uniform proper q-colouring of G. In this case elements of $\mathcal{M}_\Lambda(K_q)$ consist of all partitions of $V(K_q)$ into two classes as near equal in size as possible, and an easy calculation gives that for all colours k

$$a_\Lambda^-(k) = \frac{1}{2\lceil q/2 \rceil} \quad \text{and} \quad a_\Lambda^+(k) = \frac{1}{2\lfloor q/2 \rfloor}$$

so that in particular $a_\Lambda^-(k) = a_\Lambda^+(k) = 1/q$ for q even, and we get the following corollary of Theorem 1.2.

8

Corollary 1.5 *Fix $q \in \mathbb{N}$. There are positive constants c_1, c_2 and c_3 (depending on q) such that for every d-regular, bipartite graph G on N vertices, the following statements hold. If χ is a uniformly chosen q-colouring of G and $\varepsilon \geq c_1/\sqrt{d}$ then for q even*

$$\Pr\left(\exists k \in V(H) : \left|\frac{|\chi^{-1}(k)|}{N} - \frac{1}{q}\right| \geq \varepsilon\right) \leq c_2 \varepsilon^{-1} 2^{-c_3 \varepsilon^2 N},$$

and for q odd

$$\left.\begin{array}{l} \Pr\left(\exists k \in V(H) : \frac{|\chi^{-1}(k)|}{N} \leq \frac{1}{q+1} - \varepsilon\right) \\ \Pr\left(\exists k \in V(H) : \frac{|\chi^{-1}(k)|}{N} \geq \frac{1}{q-1} + \varepsilon\right) \end{array}\right\} < c_2 \varepsilon^{-1} 2^{-c_3 \varepsilon^2 N}.$$

So for even q, almost all proper q-colourings of a regular bipartite graph are "almost equitable". Of course, by the symmetry of K_q we have $E(|\chi^{-1}(k)|) = N/q$ for all k in this case.

The condition that G be regular can be relaxed quite a bit; we simply require that G has not too many low degree vertices, that the sum of the degrees of high degree vertices is not too large, and that the difference between the sizes of the partition classes is not too great.

Theorem 1.6 *Fix H and Λ. There are positive constants c_1, c_2, c_3 and c_4 (depending on H and Λ) such that the following statements hold. Let G be a bipartite graph on N vertices with bipartition classes $\mathcal{E}$ and $\mathcal{O}$ (with $|\mathcal{O}| \geq |\mathcal{E}|$). Let d be an arbitrary positive parameter. Let ε satisfy $\varepsilon \geq c_1\sqrt{h(G,d)}$ where*

$$h(G,d) = \frac{1}{d} + \frac{|\{v \in \mathcal{E} : d(v) < d\}|}{N} + \frac{|\mathcal{O}| - |\mathcal{E}|}{N} + \frac{1}{dN}\sum_{v \in \mathcal{O}}(d(v) - d)\mathbf{1}_{\{d(v) \geq d\}}.$$

Then for each $k \in V(H)$ we have (2), as well as (3) with now

$$\zeta = \max\left\{\sqrt{h(G,d)}, \sqrt{\frac{\log N}{N}}\right\}.$$

If G is d-regular then $h(G,d) = 1/d$ and so Theorem 1.6 is a generalization of Theorem 1.2. The proof of Theorem 1.6 follows the same lines as already described for Theorem 1.2, except that we now require a new upper bound on $Z_\Lambda(G,H)$. In Section 2 we modify the entropy-based proof of (4) to obtain the following, which is just what we need for Theorem 1.6, the proof of which is then given in Section 3. Here $d(v) = |\{u \in V(G) : uv \in E(G)\}|$ is the degree of vertex v, and we write $w_\Lambda(H)$ for $w_\Lambda(V(H))$.

9

Theorem 1.7 *Fix H and Λ, and suppose that $\lambda_i > 1$ for all $i \in V(H)$. Let G be any bipartite graph on bipartition classes $\mathcal{E}$ and $\mathcal{O}$, with $|\mathcal{O}| \geq |\mathcal{E}|$, and let d be an arbitrary positive parameter. Then*

$$Z_\Lambda(G, H) \leq w_\Lambda(H)^{|\{w \in \mathcal{E}: d(w) < d\}|} \prod_{v \in \mathcal{O}} Z_\Lambda(K_{d(v),d}, H)^{\frac{1}{d}}.$$

Note that if G is d-regular then Theorem 1.7 reduces to (4). Note also that the condition imposed on the λ_i by Theorem 1.7 is not restrictive: if Λ' is obtained from Λ by multiplying all $\lambda_i \in \Lambda$ by the same positive constant then $p_\Lambda(N_1(f) = \cdot) = p_{\Lambda'}(N_1(f) = \cdot)$ and so we may assume without loss of generality that $\min\{\lambda_i : i \in V(H)\} > 1$.

Theorem 1.6 is only of interest in situations where $h(G, d)$ can be shown to be small (as, for example, when G is d-regular). A natural situation where we can say something about $h(G, d)$ is in percolation. Given a graph G and a parameter $0 \leq p \leq 1$, let G_p be a random subgraph of G obtained by deleting each edge independently with probability $1 - p$ (so the probability that $G_p = H$ is $p^{|E(H)|}(1 - p)^{|E(G)| - |E(H)|}$). A corollary of Theorem 1.6 (which we will prove in Section 4) is the following "phase transition" phenomenon for percolation on a regular bipartite graph. If G is an n-regular bipartite graph and p is much greater than $1/n$, then the typical appearance of a p_Λ-chosen H-colouring of G_p is similar to that of a p_Λ-chosen H-colouring of G, whereas if p is much smaller than $1/n$, then as long as there is some $k \in V(H)$ with $\lambda_k / w_\Lambda(H) \notin [a_\Lambda^-(k), a_\Lambda^+(k)]$, these two objects have different appearances.

Corollary 1.8 *Fix H and Λ. Let $f(n) = \omega(1)$. There is a function $g(n) = o(1)$ (depending on $f(n)$) such that if $\{G^n\}_{n=1}^\infty$ is a sequence of n-regular bipartite graphs and p satisfies $p \geq f(n)/n$, then with probability at least $1 - g(n)$ the graph G_p^n satisfies that for each $k \in V(H)$ we have*

$$p_\Lambda(s(k, f) \in I_k(g(n))) \leq g(n)$$

and

$$\bar{p}_\Lambda(k) \in \left[a_\Lambda^-(k) - g(n), a_\Lambda^+(k) + g(n)\right].$$

If on the other hand $p \leq 1/(f(n)n)$ then with probability at least $1 - g(n)$ we have that for each $k \in V(H)$,

$$p_\Lambda\left(\left|s(k, f) - \frac{\lambda_k}{w_\Lambda(H)}\right| \leq g(n)\right) \geq 1 - g(n)$$

and

$$\left| \bar{p}_\Lambda(k) - \frac{\lambda_k}{w_\Lambda(H)} \right| \le g(n).$$

For the multicast model (Example B), for example, we have

$$a_\Lambda^-(0) = a_\Lambda^+(0) = \frac{1}{2\left(\sum_{i \le \lfloor k/2 \rfloor} \lambda^i\right)} + \frac{1}{2\left(\sum_{i \le \lceil k/2 \rceil} \lambda^i\right)} > \frac{1}{\sum_{i \le k} \lambda^i}$$

and so Corollary 1.8 shows a phase transition for this model. For the uniform q-colouring model (Example C), on the other hand, Corollary 1.8 gives no information about what happens as p crosses $1/n$.

2 Proof of Theorem 1.7

We will initially assume that for all $i \in V(H)$, we have $\lambda_i \in \mathbb{Q}$. Under this assumption, we can relate $Z_\Lambda(G, H)$ to a uniform model. We repeat an idea used in [6] and first introduced in [2]. Let C be any positive integer with the property that $C\lambda_i \in \mathbb{Z}$ for each $i \in V(H)$. Let H_Λ^C be the graph obtained from H by the following process: replace each vertex i with a set S_i of size $C\lambda_i$, replace each edge ij ($i \ne j$) with a complete bipartite graph between S_i and S_j, and replace each loop ii with a complete looped graph on S_i. It is easy to check that for any N vertex graph G we have

$$Z_\Lambda(G, H) = \frac{|\text{Hom}(G, H_\Lambda^C)|}{C^N}. \tag{5}$$

We now bound $|\text{Hom}(G, H_\Lambda^C)|$ using an entropy approach that was used in [7] to upper bound the number of independent sets in a regular bipartite graph, and was generalized in [6] to bound $|\text{Hom}(G, H)|$ for arbitrary H and regular bipartite G. We very briefly review the necessary entropy background here; see for example [9] for a more detailed treatment.

For a discrete random variable X, let $R(X)$ be the support of the mass function of X. Define the *entropy* of X to be

$$H(X) = \sum_{x \in R(X)} -P(X = x) \log P(X = x),$$

where here, and throughout the rest of this paper, logarithms have base 2. We may think of $H(X)$ as a measure of the randomness of X or as the amount

11

of information it contains. The *conditional entropy* of X given the discrete random variable Y is given by

$$H(X|Y) = \sum_{y \in R(Y)} P(Y = y) \sum_{x \in R(X)} -P(X = x|Y = y) \log P(X = x|Y = y).$$

Here are the basic facts about the entropy function that we will need. The inequality that makes entropy useful as a tool for enumeration is

$$H(X) \le \log |R(X)| \tag{6}$$

with equality if and only if X is uniform. For a vector $(X_1, \dots, X_n)$ of random variables (itself a discrete random variable) we have a chain rule

$$H(X_1, \dots, X_n) = H(X_1) + H(X_2|X_1) + \dots + H(X_n|X_1, \dots, X_{n-1}). \tag{7}$$

For random variables X, Y and Z we have

$$H(X|Y) \le H(X) \quad \text{and} \quad H(X|Y, Z) \le H(X|Y) \tag{8}$$

(so dropping conditioning does not decrease entropy). Finally, we have conditional subadditivity:

$$H(X_1, \dots, X_n|Y) \le H(X_1|Y) + H(X_2|Y) + \dots + H(X_n|Y). \tag{9}$$

Now let f be a uniformly chosen element of $\mathrm{Hom}(G, H_\Lambda^C)$. By (7) the entropy of f satisfies

$$H(f) = H(f(\mathcal{E})) + H(f(\mathcal{O})|f(\mathcal{E})). \tag{10}$$

We upper bound $H(f(\mathcal{O})|f(\mathcal{E}))$ using (8) and (9):

$$H(f(\mathcal{O})|f(\mathcal{E})) \le \sum_{v \in \mathcal{O}} H(f(v)|f(N(v))) \tag{11}$$

where $N(v) = \{u \in V(G) : uv \in E(G)\}$ is the neighbourhood of v. We upper bound $H(f(\mathcal{E}))$ using a form of Shearer's Lemma [4] derived from Radhakrishnan's proof of same (see for example [8]). Put a total order $<$ on the vertices of G. For each $v \in \mathcal{O}$ with $N(v) = \{n_1, \dots, n_{d(v)}\}$ where $n_1 < \dots < n_{d(v)}$ we have, by (7) and (8),

$$\begin{aligned} H(f(N(v))) &= \sum_{i=1}^{d(v)} H(f(n_i)|f(n_{i-1}), \dots, f(n_1)) \\ &\ge \sum_{i=1}^{d(v)} H(f(n_i)|\{f(u) : u < n_i\}) \end{aligned}$$

12

and so

$$\sum_{v \in \mathcal{O}} H(f(N(v))) \geq \sum_{w \in \mathcal{E}} d(w) H(f(w) | \{f(u) : u < w\})$$

$$= \sum_{w \in \mathcal{E}} (d + (d(w) - d)) H(f(w) | \{f(u) : u < w\}) \quad (12)$$

where d, as in the statement of Theorem 1.7, is any positive parameter. Since by (7) again we have

$$\sum_{w \in \mathcal{E}} H(f(w) | \{f(u) : u < w\}) = H(f(\mathcal{E}))$$

we rearrange the terms of (12) to get

$$H(f(\mathcal{E})) \leq \frac{1}{d} \sum_{v \in \mathcal{O}} H(f(N(v))) + \sum_{w \in \mathcal{E}} \left(1 - \frac{d(w)}{d}\right) H(f(w) | \{f(u) : u < w\}).$$
$$(13)$$

We combine (10), (11) and (13) to upper bound $H(f)$ as the sum of

$$\frac{1}{d} \sum_{v \in \mathcal{O}} \left(H(f(N(v))) + dH(f(v) | f(N(v))) \right) \quad (14)$$

and

$$\sum_{w \in \mathcal{E}} \left(1 - \frac{d(w)}{d}\right) H(f(w) | \{f(u) : u < w\}). \quad (15)$$

We deal first with (14). Fix $v \in \mathcal{O}$. For each $A \in V(H)^{N(v)}$ that occurs as a value of $f(N(v))$, let $p(A)$ be the probability that A occurs and let $e(A)$ be the number of possible ways of assigning an image to v given that $f(N(v))$ takes value A. Expanding out the entropy terms we have

$$H(f(N(v))) + dH(f(v) | f(N(v))) \leq \sum_{A} p(A) \log \frac{e(A)^d}{p(A)} \quad (16)$$

$$\leq \log \sum_{A} e(A)^d \quad (17)$$

$$\leq \log |\mathrm{Hom}(K_{d(v),d}, H_\Lambda^C)| \quad (18)$$

$$= \log \left(C^{d(v)+d} Z_\Lambda(K_{d(v),d}, H) \right). \quad (19)$$

13

We use (6) to obtain (16) and Jensen's inequality for (17), and the equality in (19) follows from (5). To see (18) note that we specify an element of $\mathrm{Hom}(K_{d(v),d}, H_\Lambda^C)$ by first choosing the restriction A of the homomorphism to the partition class of size $d(v)$ and then for each of the remaining d vertices choosing the value independently from $e(A)$. Summing over $v \in \mathcal{O}$ we see that (14) is bounded above by

$$\log \left(C^{\frac{|E(G)|}{d} + |\mathcal{O}|} \right) + \sum_{v \in \mathcal{O}} \log \left(Z_\Lambda(K_{d(v),d}, H)^{\frac{1}{d}} \right). \tag{20}$$

For (15), if $d(w) < d$, we upper bound

$$H(f(w)|\{f(u) : u < w\}) \leq \log |V(H_\Lambda^C)| = \log (Cw_\Lambda(H))$$

using (6) and (8). If $d(w) \geq d$ then we need a lower bound on $H(f(w)|\{f(u) : u < w\})$. Since f is a homomorphism, there is at least one i such that f can take values in S_i. Fix one such. If we add the condition that $f(w) \in S_i$ then, because the vertices of S_i are indistinguishable, $f(w)$ becomes uniform and its entropy is the logarithm of $|S_i|$. That is,

$$\begin{aligned} H(f(w)|\{f(u) : u < w\}) &\geq H(f(w)|\{f(u) : u < w\}, \{f(w) \in S_i\}) \\ &= \log C\lambda_i \\ &\geq \log C, \end{aligned}$$

the last inequality following from $\lambda_i > 1$ for all $i \in V(H)$. It follows that (15) is bounded above by

$$\log \left(C^{|\mathcal{E}| - \frac{|E(G)|}{d}} w_\Lambda(H)^{|\{w \in \mathcal{E}: d(w) < d\}|} \right). \tag{21}$$

Putting (20) and (21) into (10), using $H(f) = \log |\mathrm{Hom}(G, H_\Lambda^C)|$ (since f is uniform) and combining with (5), we obtain Theorem 1.7 for rational λ_i's. By continuity, this bound remains valid when the λ_i's are not necessarily rational.

3 Proof of Theorem 1.6

We begin by using Theorem 1.7 to put an upper bound on $Z_\Lambda(G, H)$. We first consider those $v \in \mathcal{O}$ with $d(v) \geq d$. For each of the $4^{|V(H)|}$ ordered pairs

$A \sim B$ of subsets of H, the contribution to $Z_\Lambda(K_{d(v),d}, H)$ from those f with the partition class of $K_{d(v),d}$ of size $d(v)$ mapped to A and the class of size d mapped to B is at most

$$w_\Lambda(A)^{d(v)} w_\Lambda(B)^d \leq (w_\Lambda(A) w_\Lambda(B))^d w_\Lambda(H)^{d(v)-d}$$

and so

$$Z_\Lambda(K_{d(v),d}, H) \leq 4^{|V(H)|} w_\Lambda(H)^{d(v)-d} \eta_\Lambda(H)^d.$$

Similarly, for those $v \in \mathcal{O}$ with $d(v) < d$ we have

$$Z_\Lambda(K_{d(v),d}, H) \leq Z_\Lambda(K_{d,d}, H) \leq 4^{|V(H)|} \eta_\Lambda(H)^d.$$

It follows from Theorem 1.7 that $Z_\Lambda(G, H)$ is upper bounded by

$$\eta_\Lambda(H)^{|\mathcal{O}|} 4^{\frac{|V(H)||\mathcal{O}|}{d}} w_\Lambda(H)^{|\{v \in \mathcal{E}: d(v) < d\}| + \frac{1}{d} \sum_{v \in \mathcal{O}} (d(v)-d) \mathbf{1}_{\{d(v) \geq d\}}}$$

and so, using $|\mathcal{O}| = N/2 + (|\mathcal{O}| - |\mathcal{E}|)/2$,

$$Z_\Lambda(G, H) \leq \eta_\Lambda(H)^{\frac{N}{2}} C^{Nh(G,d)} \tag{22}$$

where C is a positive constant depending only on H and Λ. On the other hand, we get a lower bound (with any $w_\Lambda(A) w_\Lambda(B) = \eta_\Lambda(H)$, and using $\lambda_i > 1$ for all $i \in V(H)$) by

$$\begin{aligned} Z_\Lambda(G, H) &\geq w_\Lambda(A)^{|\mathcal{E}|} w_\Lambda(B)^{|\mathcal{O}|} \\ &\geq (w_\Lambda(A) w_\Lambda(B))^{|\mathcal{E}|} \tag{23} \\ &= \eta_\Lambda(H)^{\frac{N}{2}} \eta_\Lambda(H)^{-\frac{|\mathcal{O}|-|\mathcal{E}|}{2}}. \tag{24} \end{aligned}$$

In (23) we are using $|\mathcal{O}| \geq |\mathcal{E}|$.

We now use (22) and (24) to prove (2). Fix $k \in V(H)$ and an integer N_k satisfying $0 \leq N_k \leq N$ and

$$\frac{N_k}{N} \in [0, a_\Lambda^-(k) - \varepsilon) \cup (a_\Lambda^+(k) + \varepsilon, 1] \ (= I_k(\varepsilon)).$$

Write $c_k(N_k)$ for the contribution to $Z_\Lambda(G, H)$ from those $f \in \mathrm{Hom}(G, H)$ with $|f^{-1}(k)| = N_k$. We aim to obtain an upper bound on $c_k(N_k)$ (via (22)) which is substantially lower than the lower bound (24), indicating that this term does not contribute greatly to $Z_\Lambda(G, H)$.

15

We begin by considering N_k for which

$$\gamma := \frac{N_k}{N} = a_\Lambda^+(k) + \varepsilon'$$

for some ε' satisfying $\varepsilon \leq \varepsilon' \leq 1 - a_\Lambda^+(k)$. For any $\delta > 0$ let $\Lambda(k, \delta)$ be obtained from Λ by replacing λ_k with $(1 + \delta)\lambda_k$ and leaving all other λ_i's unchanged. By (22) we have

$$
\begin{aligned}
(1 + \delta)^{N_k} c_k(N_k) &\leq Z_{\Lambda(k,\delta)}(G, H) \\
&\leq \eta_{\Lambda(k,\delta)}(H)^{\frac{N}{2}} C^{Nh(G,d)}
\end{aligned}
\tag{25}
$$

where now the constant C depends on δ as well as on H and Λ.

Before proceeding, we need to understand $\eta_{\Lambda(k,\delta)}(H)$. Viewed as a function of δ, the quantity $w_{\Lambda(k,\delta)}(A)w_{\Lambda(k,\delta)}(B)$ (for $(A, B) \in \mathcal{M}_\Lambda(H)$) is of the form $a + b\delta + c\delta^2$ where $a = \eta_\Lambda(H)$, $b = w_\Lambda(A)\lambda_k \mathbf{1}_{\{k \in B\}} + w_\Lambda(B)\lambda_k \mathbf{1}_{\{k \in A\}}$ and $c = \lambda_k^2 \mathbf{1}_{\{k \in A \cap B\}}$. From this formulation we can easily identify that set $\emptyset \neq \mathcal{S}_\Lambda^+(k, H) \subseteq \mathcal{M}_\Lambda(H)$ with the property that for all $\delta > 0$, all $(A, B) \in \mathcal{M}_\Lambda(H)$ and all $(A', B') \in \mathcal{S}_\Lambda^+(k, H)$ we have $w_{\Lambda(k,\delta)}(A')w_{\Lambda(k,\delta)}(B') \geq w_{\Lambda(k,\delta)}(A)w_{\Lambda(k,\delta)}(B)$: $\mathcal{S}_\Lambda^+(k, H)$ consists of all those $(A', B') \in \mathcal{M}_\Lambda(H)$ for which b is maximum and (subject to this condition) c is maximum. This latter condition simply means that if some of the pairs that maximize b have $c > 0$ we only take those pairs, and if they all have $c = 0$ we take all pairs.

It is easily seen that there is a sufficiently small $\delta_k^+ > 0$ (depending on H and Λ) with the property that for all $0 < \delta < \delta_k^+$ and $(A', B') \in \mathcal{S}_\Lambda^+(k, H)$ we have $(A', B') \in \mathcal{M}_{\Lambda(k,\delta)}(H)$. Choose one such, (A^+, B^+), arbitrarily. Note that by construction

$$a_\Lambda^+(k) = \frac{w_\Lambda(A^+)\lambda_k \mathbf{1}_{\{k \in B^+\}} + w_\Lambda(B^+)\lambda_k \mathbf{1}_{\{k \in A^+\}}}{2\eta_\Lambda(H)} = \frac{\lambda_k \mathbf{1}_{\{k \in A^+\}}}{2w_\Lambda(A^+)} + \frac{\lambda_k \mathbf{1}_{\{k \in B^+\}}}{2w_\Lambda(B^+)}.$$

Now combining (24) and (25) and choosing $\delta < \delta_k^+$ we have

$$
\begin{aligned}
p_\Lambda(|f^{-1}(k)| = N_k) &= \frac{c_k(N_k)}{Z_\Lambda(G, H)} \\
&\leq C^{h(G,d)N} \left(\frac{w_{\Lambda(k,\delta)}(A^+)w_{\Lambda(k,\delta)}(B^+)}{w_\Lambda(A^+)w_\Lambda(B^+)(1 + \delta)^{2(a_\Lambda^+(k)+\varepsilon')}} \right)^{\frac{N}{2}}
\end{aligned}
\tag{26}
$$

where, by our restriction on δ, C may be taken to depend only on H and Λ. Our aim is to show that there is a positive constant c (depending on H and Λ) such that for all $0 < \varepsilon' \leq 1 - a_\Lambda^+(k)$ we can find a $0 < \delta < \delta_k^+$ for which

$$\frac{w_{\Lambda(k,\delta)}(A^+)w_{\Lambda(k,\delta)}(B^+)}{w_\Lambda(A^+)w_\Lambda(B^+)(1+\delta)^{2(a_\Lambda^+(k)+\varepsilon')}} \leq 2^{-c\varepsilon'^2}. \tag{27}$$

Combining this with (26) we see that if $\varepsilon > c\sqrt{h(G,d)}$ for some suitably large positive constant c (depending on Λ and H) then for all $\varepsilon < \varepsilon' \leq 1 - a_\Lambda^+(k)$ for which $a^+(k)N + \varepsilon'N$ is an integer we have

$$p_\Lambda\left(|f^{-1}(k)| = a^+(k)N + \varepsilon'N\right) \leq 2^{-c'\varepsilon'^2 N}$$

for a suitable positive c', and so

$$\begin{aligned} p_\Lambda\left(|f^{-1}(k)| \geq a^+(k)N + \varepsilon N\right) &\leq \sum_{\ell \geq \varepsilon N} 2^{-\frac{c'\ell^2}{N}} \\ &\leq 2^{-c'\varepsilon^2 N} \sum_{\ell \geq 0} 2^{-2\ell c'\varepsilon} \\ &\leq c''\varepsilon^{-1} 2^{-c'\varepsilon^2 N} \end{aligned} \tag{28}$$

for suitably large c'' (depending on c'). An almost identical argument (the details of which we leave to the reader) yields

$$p_\Lambda\left(|f^{-1}(k)| \leq a^-(k)N - \varepsilon N\right) \leq c''\varepsilon^{-1} 2^{-c'\varepsilon^2 N} \tag{29}$$

for $\varepsilon > c\sqrt{h(G,d)}$. Combining (28) and (29) gives (2).

We now turn to (27). Observe that it is enough to prove (27) for all $0 < \varepsilon' \leq \varepsilon_0$, where $\varepsilon_0 \leq 1 - a_\Lambda^+(k)$ may be any constant (perhaps depending on H and Λ). Indeed, for any $\varepsilon' \geq \varepsilon_0$ we know that there is a choice of $\delta < \delta_k^+$ for which

$$\begin{aligned} \frac{w_{\Lambda(k,\delta)}(A^+)w_{\Lambda(k,\delta)}(B^+)}{w_\Lambda(A^+)w_\Lambda(B^+)(1+\delta)^{2(a_\Lambda^+(k)+\varepsilon')}} &\leq \frac{w_{\Lambda(k,\delta)}(A^+)w_{\Lambda(k,\delta)}(B^+)}{w_\Lambda(A^+)w_\Lambda(B^+)(1+\delta)^{2(a_\Lambda^+(k)+\varepsilon_0)}} \\ &\leq 2^{-c\varepsilon_0^2}. \end{aligned}$$

Setting $c' = c\varepsilon_0^2$ we have $2^{-c\varepsilon_0^2} \leq 2^{-c'\varepsilon'^2}$ for $\varepsilon' \geq \varepsilon_0$ and $2^{-c\varepsilon'^2} \leq 2^{-c'\varepsilon'^2}$ for $\varepsilon' < \varepsilon_0$, so we may replace c with c' to obtain the result for the full range of

ε'. From now on we will assume that $\varepsilon' < \varepsilon_0$, for a certain ε_0 that will be specified later.

Setting

$$\gamma_A = \frac{\lambda_k \mathbf{1}_{\{k \in A^+\}}}{2w_\Lambda(A^+)}, \quad \gamma_B = \frac{\lambda_k \mathbf{1}_{\{k \in B^+\}}}{2w_\Lambda(B^+)}$$

(so $a_\Lambda^+(k) = \gamma_A + \gamma_B$) the left-hand side of (27) becomes

$$\frac{w_\Lambda(A^+) + \delta\lambda_k \mathbf{1}_{\{k \in A^+\}}}{(1+\delta)^{2\gamma_A + \varepsilon'} w_\Lambda(A^+)} \times \frac{w_\Lambda(B^+) + \delta\lambda_k \mathbf{1}_{\{k \in B^+\}}}{(1+\delta)^{2\gamma_B + \varepsilon'} w_\Lambda(B^+)}. \tag{30}$$

If either $A^+ = \{k\}$ or $k \notin A^+$ then the first term of (30) is $(1+\delta)^{-\varepsilon'}$ so that in this case we have that for any $\delta > 0$ depending only on H and Λ, and any $0 < \varepsilon' \le 1$,

$$\frac{w_\Lambda(A^+) + \delta\lambda_k \mathbf{1}_{\{k \in A^+\}}}{(1+\delta)^{2\gamma_A + \varepsilon'} w_\Lambda(A^+)} \le 2^{-c\varepsilon'} \le 2^{-c\varepsilon'^2},$$

where c is a positive constant depending on H and Λ (the last inequality using $\varepsilon' \le 1$). If $k \in A^+$ and $|A^+| > 1$ then the first term of (30) takes the form

$$\begin{aligned}
\frac{w_\Lambda(A^+) + \delta\lambda_k}{(1+\delta)^{2\gamma_A + \varepsilon'} w_\Lambda(A^+)} &\le& \frac{1 + \delta(\lambda_k/w_\Lambda(A^+))}{1 + \delta(2\gamma_A + \varepsilon')} \\
&=& 1 - \frac{\delta\varepsilon'}{1 + \delta((\lambda_k/w_\Lambda(A^+)) + \varepsilon')} \\
&\le& 1 - \frac{\delta\varepsilon'}{3}, \tag{31}
\end{aligned}$$

with (31) valid for sufficiently small ε'. Now taking $\delta = \varepsilon'$ (having chosen ε_0 small enough that this choice is allowed, and that (31) holds), we get a bound of $2^{-c\varepsilon'^2}$ on the first term of (30), where c is a positive constant depending on H and Λ only.

Repeating this analysis for the second term of (30), we obtain (27) and thus (2).

Applying (2) with $\varepsilon = c\sqrt{(\log N)/N}$ (if $(\log N)/N > h(G,d)$) and $\varepsilon = c\sqrt{h(G,d)}$ (otherwise), where $c \ge c_1$ satisfies $c^2 c_3 \ge 1$, we easily obtain (3), based on the observation that in both cases

$$E_\Lambda(s(k,f)) \le (a_\Lambda^+(k) + \varepsilon)\left(1 - c_2\varepsilon^{-1}2^{-c_3\varepsilon^2 N}\right) + c_2\varepsilon^{-1}2^{-c_3\varepsilon^2 N}$$

with a similar lower bound involving $a_\Lambda^-(k)$.

4 Proof of Corollary 1.8

We assume throughout that $|V(G^n)| = N$ (a function of n) and that G^n has fixed bipartition $\mathcal{E} \cup \mathcal{O}$.

We begin with the $p = \omega(1/n)$ regime. We take

$$d = np - \sqrt{2xnp}$$

with $x = \sqrt{f(n)}$. The choice of x is driven by the aim of making all of the terms of $h(G_p^n, d)$ be $o(1)$, with probability $1 - o(1)$; this is enough for both statements of the corollary in this regime. Note that since $|\mathcal{E}| = |\mathcal{O}|$ we immediately have $(|\mathcal{O}| - |\mathcal{E}|)/N = o(1)$.

By our choice of x we have

$$\sqrt{2xnp} \le \frac{np}{2}$$

(for large enough n) and so $d \ge np/2$ and $1/d = o(1)$.

For a given vertex $v \in \mathcal{E}$, let $d(v)$ be its degree in G_p^n. This is a binomial random variable with parameters n and p, and so by standard Chernoff-type bounds (see for example [1, Appendix A]) we have

$$P(d(v) < d) \le e^{-x}.$$

(The specific bound we are using here is

$$P(\mathrm{Bin}(n, p) - np < -a) < e^{-a^2/2pn}$$

for $a > 0$.) The distribution of vertices from $\mathcal{E}$ which have degree smaller than d is therefore binomial with parameters $N/2$ and $p' \le e^{-x}$. The expected number of such vertices is at most $Ne^{-x}/2$, and by Markov's inequality the probability that there are more than $Ne^{-x+\sqrt{x}}/2$ such is at most $e^{-\sqrt{x}}$. Since $x = \omega(1)$, this is $o(1)$, and so with probability $1 - o(1)$ we have

$$\frac{|\{v \in \mathcal{E} : d(v) < d\}|}{N} = o(1).$$

It remains to consider $S := \sum \{d(v) - d : v \in \mathcal{O}, d(v) \geq d\}$. We have

$$
\begin{aligned}
E(S) &= \sum_{v \in \mathcal{O}} E\left(d(v)\mathbf{1}_{\{d(v) \geq d\}}\right) - dE\left(\mathbf{1}_{\{d(v) \geq d\}}\right) \\
&\leq \sum_{v \in \mathcal{O}} \left(\sum_{j \geq d} j\binom{n}{j} p^j(1-p)^{n-j} - d(1 - e^{-x})\right) \\
&\leq N\left(np - d + de^{-x}\right) \\
&\leq N\left(\sqrt{2xnp} + npe^{-x}\right),
\end{aligned}
$$

and so

$$
E\left(\frac{S}{dN}\right) \leq 2\sqrt{\frac{2x}{f(n)}} + \frac{2}{e^x} \quad (= o(1))
$$

for large enough n (again using $d \geq np/2$). Again by Markov's inequality, with probability $1 - o(1)$ we have $S/dN = o(1)$ and so with probability $1 - o(1)$ we have $h(G_p^n, d) = o(1)$, as required.

We now deal with the $p = o(1/n)$ regime. The probability that a particular vertex is isolated in G_p^n is $(1-p)^n \geq 1 - 2f(n)$ (for large enough n), so the number of non-isolated vertices in $\mathcal{E}$ is a binomial random variable with parameters $N/2$ and $p' \leq 2f(n)$. By the Chernoff bound, asymptotically almost surely (with probability tending to one as n tend to infinity) $\mathcal{E}$ has fewer than $2f(n)N$ non-isolated vertices and so also asymptotically almost surely G_p^n has fewer than $4f(n)N$ non-isolated vertices. For each $k \in V(H)$, the number of isolated vertices mapped to k is a binomial random variable with parameters $m \geq N(1 - 4f(n))$ and $p'' = \lambda_k/w_\Lambda(H)$ and so (again by Chernoff bounds) asymptotically almost surely there are at least $N(1 - 5f(n))\lambda_k/w_\Lambda(H)$ vertices of G_p^n mapped to k. Since $\sum_{k \in V(H)} \lambda_k/w_\Lambda(H) = 1$, we also have that asymptotically almost surely there are at most

$$
N\left(\frac{\lambda_k}{w_\Lambda(H)} + 5f(n)\left(1 - \frac{\lambda_k}{w_\Lambda(H)}\right)\right)
$$

vertices of G_p^n mapped to k. This completes the proof of the corollary.

5   Proof of Theorem 1.4

The graph G_d will be a random d-regular bipartite graph on $N = c^{d/\log d}$ vertices (where $c > 1$ will depend on the particular H and Λ under consideration).

20

A standard method of constructing such a graph is as follows. We begin with a set of size Nd consisting of $Nd/2$ type I vertices $\{u_{ij} : 1 \le i \le N/2, 1 \le j \le d\}$ and $Nd/2$ type II vertices $\{v_{ij} : 1 \le i \le N/2, 1 \le j \le d\}$. We then choose a uniformly random perfect matching from the type I vertices to the type II vertices, and turn this into a d-regular bipartite multigraph on N vertices with bipartition classes $\mathcal{E} = \{u_1, \ldots, u_{N/2}\}, \mathcal{O} = \{v_1, \ldots, v_{N/2}\}$ by, for each $i = 1, \ldots, N/2$, identifying $u_{i,1}, \ldots, u_{i,d}$ with u_i and $v_{i,1}, \ldots, v_{i,d}$ with v_i. Finally, we condition on the result being a simple graph. This process generates a d-regular bipartite graph on N vertices with bipartition classes $\mathcal{E}$, $\mathcal{O}$, uniformly (see for example [13]).

O'Neil [11] showed that the probability that the multigraph produced by this process is simple is (for large enough d) at least $e^{-d^2/3}$. It follows that if we establish that the multigraph produced (before conditioning on being simple) has a certain property with probability at least $1 - e^{-d^2}$ (say), then there is a simple d-regular graph with that property.

We want to establish that for large enough d the multigraph has a number of desirable expansion properties. First, we want to show that for each $C \log d \le j \le 3N(\log d)/d$ (for some constant $C > 0$, depending on c), every subset of $\mathcal{E}$ of size j and every subset of $\mathcal{O}$ of size j has at least αj distinct neighbours where $\alpha = d/(C \log d)$. For a particular such j, the probability that the graph fails to have this property is (by a union bound) at most

$$2 \binom{N/2}{j} \binom{N/2}{\alpha j} \frac{(\alpha j d)_{jd}}{(Nd/2)_{jd}} \le \left(\frac{eN}{2\alpha j} \right)^{2\alpha j} \left(\frac{2\alpha j}{N} \right)^{jd}$$

$$= e^{\frac{2jd}{C \log d}} \left(\frac{2jd}{CN \log d} \right)^{jd - \frac{2jd}{C \log d}}$$

$$\le e^{\frac{2jd}{C \log d}} \left(\frac{2jd}{CN \log d} \right)^{jd/2}$$

(for large enough d, depending on C) with the first inequality using $\binom{n}{r} \le (en/r)^r$. For $j \ge d \log d$ we bound $2jd/(CN \log d) \le 1/2$ (valid for $C \ge 12$) so that for large enough d (depending on C)

$$e^{\frac{2jd}{C \log d}} \left(\frac{2jd}{CN \log d} \right)^{jd/2} \le 1.4^{-jd} \le e^{-2d^2}.$$

For $j \le d \log d$ we instead bound $(2dj)/(CN \log d) \le d^2/N$ (valid for $C \ge 2$).

21

We now have

$$e^{\frac{2jd}{C\log d}} \left(\frac{2jd}{CN\log d}\right)^{jd/2} \leq \exp\left\{2jd\log d - \frac{jd^2\log c}{2\log d}\right\} \leq \exp\left\{\frac{-jd^2\log c}{3\log d}\right\}$$

(again for large d, recalling $N = c^{d/\log d}$), which is at most e^{-2d^2} for $j \geq C\log d$ for suitable C depending on c. Since there are at most $N = c^{d\log d}$ choices for j, the probability that the graph fails to have the desired property for some j is at most e^{-d^2}. If the process results in a simple graph, then we trivially get the same expansion for subsets of $\mathcal{E}$ or $\mathcal{O}$ of size at most $C\log d$, since for $1 \leq j \leq C\log d$ there is a trivial lower bound of d on the neighbourhood size of a set of size j, and we have $d \geq jd/(C\log d)$ for j in this range.

Next we establish that the graph has the property that for every subset A of $\mathcal{E}$ of size $3N(\log d)/d$ and every subset B of $\mathcal{O}$ of size $3N(\log d)/d$, there is an edge joining a vertex of A to a vertex of B. By a union bound, the probability that the multigraph fails to have the property is at most

$$\binom{N/2}{\beta N}^2 \frac{(Nd/2 - \beta Nd)_{\beta Nd}}{(Nd/2)_{\beta Nd}} \leq \exp\left\{2\beta N\log(e/(2\beta)) - 2\beta^2 dN\right\}$$

where $\beta = 3(\log d)/d$. With $N = c^{d/\log d}$, this is at most e^{-d^2} for large enough d (depending on c). We have shown the following.

Lemma 5.1 *Fix $c > 1$. There are $d_0 \geq 1$ and positive C, both depending on c, such that for all $d \geq d_0$ there is a d-regular, bipartite graph G_d on $N = c^{d/\log d}$ vertices with bipartition classes $\mathcal{E}$ and $\mathcal{O}$ satisfying the following:*

1. *Every subset of $\mathcal{E}$ or $\mathcal{O}$ of size j, with $1 \leq j \leq 3N(\log d)/d$, has at least $jd/(C\log d)$ neighbours.*

2. *Every pair of subsets each of size $3N(\log d)/d$, one from $\mathcal{E}$ and one from $\mathcal{O}$, have an edge between them.*

We now fix such a G_d and study $Z_\Lambda(G_d, H)$. Given $f \in \mathrm{Hom}(G_d, H)$ set

$$\mathcal{E}(f) = \{k \in V(H) : |f^{-1}(k) \cap \mathcal{E}| \geq 3N(\log d)/d\}$$

and

$$\mathcal{O}(f) = \{k \in V(H) : |f^{-1}(k) \cap \mathcal{O}| \geq 3N(\log d)/d\}.$$

22

Clearly both $\mathcal{E}(f)$ and $\mathcal{O}(f)$ are non-empty, and by Lemma 5.1, we have $\mathcal{E}(f) \sim \mathcal{O}(f)$ (that is, everything in $\mathcal{E}(f)$ is adjacent to everything in $\mathcal{O}(f)$). So we can partition $\mathrm{Hom}(G_d, H)$ into classes indexed by pairs (A, B) with $A \sim B$. Write $\mathcal{C}(A, B)$ for the class corresponding to (A, B). We want to establish that for $(A, B) \in \mathcal{M}_\Lambda(H)$ we have

$$\sum_{f \in \mathcal{C}(A,B)} w_\Lambda(f) = (1 + o(1))\eta_\Lambda(H)^{N/2} \tag{32}$$

while for all other (A, B) we have

$$\sum_{f \in \mathcal{C}(A,B)} w_\Lambda(f) = o\left(\eta_\Lambda(H)^{N/2}\right), \tag{33}$$

where all asymptotic terms are (unless stated otherwise) as $d \to \infty$. From this we see that

$$Z_\Lambda(G_d, H) = |\mathcal{M}_\Lambda(H)|(1 + o(1))\eta_\Lambda(H)^{N/2},$$

and that all but a vanishing proportion of $Z_\Lambda(G_d, H)$ comes from pure-(A, B) colourings (with $(A, B) \in \mathcal{M}_\Lambda(H)$) in which $\mathcal{E}$ is mapped to A and $\mathcal{O}$ to B, with each such (A, B) contributing equally to $Z_\Lambda(G_d, H)$; this is enough to give the first part of Theorem 1.4. Indeed, fix $(A, B) \in \mathcal{M}_\Lambda(H)$. A proportion $(1 + o(1))/|\mathcal{M}_\Lambda|$ of $Z_\Lambda(G_d, H)$ is obtained by independently colouring $\mathcal{E}$ from A and $\mathcal{O}$ from B according to the given weights. Fix $k \in A$. We claim that with very high probability, a proportion very close to $\lambda_k/w_\Lambda(A)$ of $\mathcal{E}$ gets mapped to k. Set $p = \lambda_k/w_\Lambda(A)$ and $m = N/2$. The number U_k of vertices of $\mathcal{E}$ mapped to k is a binomial random variable with parameters m and p. So by Tchebychev's inequality,

$$\mathrm{Pr}\left(|U_k - pm| \geq \log m \sqrt{mp(1-p)}\right) \leq \frac{1}{\log^2 m}.$$

This shows that the proportion of vertices mapped to k in a pure-(A, B) colouring is very close to

$$\frac{\lambda_k \mathbf{1}_{\{k \in A\}}}{2w_\Lambda(A)} + \frac{\lambda_k \mathbf{1}_{\{k \in B\}}}{2w_\Lambda(B)}$$

with high probability. Applying this with $(A, B) = (A^+, B^+)$ and $(A, B) = (A^-, B^-)$, the first part of Theorem 1.4 follows.

23

The lower bound in (32) is obtained by considering pure-(A, B) colourings with $\mathcal{E}$ mapped to A and $\mathcal{O}$ to B. To establish (33) and the upper bound in (32), fix $0 \leq j \leq 3N(\log d)/d$, let $q = |V(H)|$, and assume that d is large. We consider the contribution to $\sum_{f \in \mathcal{C}(A,B)} w_\Lambda(f)$ from those $f \in \mathcal{C}(A, B)$ in which, for each $k \notin A \cup B$, we have at most j vertices mapped to k, and we have at least one $k' \notin A \cup B$ whose preimage has size j. To bound the contribution from these f, we first bound the number of ways of locating the vertices that are mapped to k for each $k \notin A \cup B$ by $\left(\sum_{i \leq j} \binom{N}{i} \right)^q$. The contribution to the sum of the weights from these exceptional vertices is at most $w_\Lambda(H)^{qj}$. For the contribution from the remaining vertices, we deal separately with the cases $(A, B) \in \mathcal{M}_\Lambda(H)$ and $(A, B) \notin \mathcal{M}_\Lambda(H)$. For $(A, B) \notin \mathcal{M}_\Lambda(H)$, we simply upper bound the contribution by $(w_\Lambda(A) w_\Lambda(B))^{N/2}$, leading to

$$
\sum_{f \in \mathcal{C}(A,B)} w_\Lambda(f) \quad \leq \quad (w_\Lambda(A) w_\Lambda(B))^{\frac{N}{2}} \left(\sum_{i \leq j} \binom{N}{i} \right)^q (w_\Lambda(H))^{qj}
$$

$$
= \quad o\left(\eta_\Lambda(H)^{N/2} \right),
$$

as required. For $(A, B) \in \mathcal{M}_\Lambda(H)$, consider a k' that has preimage size j. We claim that there are at least $jd/(2C \log d)$ vertices which, in the specification of f, need to be mapped to $A \cup B$ and which are adjacent to at least one of the j vertices mapped to k'. Indeed, by Lemma 5.1, the neighbourhood size of the j vertices mapped to k' is at least $jd/(C \log d)$, and at most qj vertices have been mapped to vertices from outside $A \cup B$, so there are at least $jd/(C \log d) - qj > jd/(2C \log d)$ vertices that are adjacent to a vertex mapped to k' and need to be mapped to vertices from $A \cup B$. Since k' cannot be adjacent to everything in A, nor can it be adjacent to everything in B (else we would not have $(A, B) \in \mathcal{M}_\Lambda(H)$), our choice on these at least $jd/(2C \log d)$ vertices is restricted to a proper subset of $A \cup B$; the contribution we get from the remaining vertices (those mapped to $A \cup B$) is therefore at most

$$
\frac{(w_\Lambda(A) w_\Lambda(B))^{\frac{N}{2}}}{(1 + \varepsilon)^{\frac{jd}{2C \log d}}}
$$

where $\varepsilon > 0$ (depending on H and Λ) can be chosen uniformly for all A, B. Combining these observations we get that

$$
\sum_{f \in \mathcal{C}(A,B)} w_\Lambda(f) \leq \eta_\Lambda(H)^{\frac{N}{2}} \frac{\left(\sum_{i \leq j} \binom{N}{i} \right)^q (w_\Lambda(H))^{qj}}{(1 + \varepsilon)^{\frac{jd}{2C \log d}}}.
$$

24

If $j = 0$, the right-hand side above is $(w_\Lambda(A)w_\Lambda(B))^{N/2}$. For $j > 0$ it can be bounded above by

$$\left(\frac{1}{(1+\varepsilon')^{\frac{d}{\log d}}}\right)^j$$

for some $\varepsilon' > 0$ (depending on H and Λ) for all j in the range $1 \leq j \leq 3N(\log d)/d$, as long as c is sufficiently small (recall $N = c^{d/\log d}$). Summing over j gives the upper bound in (32).

We now turn to the second part of Theorem 1.4. We take G'_d to be the disjoint union of m copies of $K_{d,d}$ where $m = m(d) = \omega(1)$. Fix $k \in V(H)$. Let X be the number of vertices mapped to k in a p_Λ-chosen H-colouring of G'_d, and X_i the number mapped to k in the ith copy of $K_{d,d}$. Define $a_\Lambda(k)$ by $E(X_i) = 2da_\Lambda(k)$, and note that $\mathrm{Var}(X_i) \leq d^2$. Since $X = \sum_{i=1}^m X_i$ we have $E(X) = 2dma_\Lambda(k)$ and $\mathrm{Var}(X) \leq md^2$. By Tchebychev's inequality,

$$P(|X - 2dma_\Lambda(k)| > 2dm\varepsilon) = P(|X/2dm - a_\Lambda(k)| > \varepsilon) \leq 1/4m\varepsilon^2.$$

So choosing $\varepsilon = o(1)$ with $m\varepsilon^2 = \omega(1)$ (for example, $\varepsilon = 1/m^{1/3}$), the probability that the proportion of vertices mapped to k in a p_Λ-chosen H-colouring of G'_d differs from $a_\Lambda(k)$ by more than $o(1)$ is at most $o(1)$. The claimed bound on $s(k, f)$ follows, as does the estimate of $\bar{p}_\Lambda(k)$.

References

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

[2] G. Brightwell and P. Winkler, Graph homomorphisms and phase transitions, *J. Combin. Theory Ser. B* **77** (1999), 221–262.

[3] G. Brightwell and P. Winkler, Hard constraints and the Bethe lattice: adventures at the interface of combinatorics and statistical physics, *Proc. Int'l. Congress of Mathematicians Vol. III* (Li Tatsien, ed.), Higher Education Press, Beijing (2002), 605–624.

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

[5] D. Galvin, F. Martinelli, K. Ramanan and P. Tetali, The multistate hard core model on a regular tree, *SIAM J. Discrete Math.* **25** (2011), 894–916.

[6] 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.

[7] J. Kahn, An Entropy Approach to the Hard-Core Model on Bipartite Graphs, *Combin. Probab. Comput.* **10** (2001), 219–237.

[8] J. Kahn, Range of the cube-indexed random walk, *Israel J. Math.* **124** (2001), 189–201.

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

[10] A. Mazel and Y. Suhov, Random surfaces with two-sided constraints: an application of the theory of dominant ground states, *J. Stat. Phys.* **64** (1991), 111–134.

[11] P. O'Neil, Asymptotics and random matrices with row-sum and column sum-restrictions, *Bulletin of the American Mathematical Society* **75** (1969), 1276-1282.

[12] K. Ramanan, A. Sengupta, I. Ziedins and P. Mitra, Markov random field models of multicasting in tree networks, *Adv. App. Prob.* **34** (2002), 58–84.

[13] N. Wormald, Models of random regular graphs, *Surveys in Combinatorics* (J.D. Lamb and D.A. Preece, eds., London Mathematical Society Lecture Note Series **276**) (1999), 239–298.