A REVERSE SIDORENKO INEQUALITY

ASHWIN SAH, MEHTAAB SAWHNEY, DAVID STONER, AND YUFEI ZHAO

Abstract. Let H be a graph allowing loops as well as vertex and edge weights. We prove that, for every triangle-free graph G without isolated vertices, the weighted number of graph homomorphisms $\hom(G,H)$ satisfies the inequality

$$\hom(G,H) \le \prod_{uv\in E(G)} \hom(K_{d_u,d_v},H)^{1/(d_u d_v)},$$

where d_u denotes the degree of vertex u in G. In particular, one has

$$\hom(G,H)^{1/|E(G)|} \le \hom(K_{d,d},H)^{1/d^2}$$

for every d-regular triangle-free G. The triangle-free hypothesis on G is best possible. More generally, we prove a graphical Brascamp–Lieb type inequality, where every edge of G is assigned some two-variable function. These inequalities imply tight upper bounds on the partition function of various statistical models such as the Ising and Potts models, which includes independent sets and graph colorings.

For graph colorings, corresponding to $H = K_q$, we show that the triangle-free hypothesis on G may be dropped; this is also valid if some of the vertices of K_q are looped. A corollary is that among d-regular graphs, $G = K_{d,d}$ maximizes the quantity $c_q(G)^{1/|V(G)|}$ for every q and d, where $c_q(G)$ counts proper q-colorings of G.

Finally, we show that if the edge-weight matrix of H is positive semidefinite, then

$$\hom(G,H) \le \prod_{v\in V(G)} \hom(K_{d_v+1},H)^{1/(d_v+1)}.$$

This implies that among d-regular graphs, $G = K_{d+1}$ maximizes $\hom(G,H)^{1/|V(G)|}$. For 2-spin Ising models, our results give a complete characterization of extremal graphs: complete bipartite graphs maximize the partition function of 2-spin antiferromagnetic models and cliques maximize the partition function of ferromagnetic models.

These results settle a number of conjectures by Galvin–Tetali, Galvin, and Cohen–Csikvári–Perkins–Tetali, and provide an alternate proof to a conjecture by Kahn.

1. Introduction

1.1. **Independent sets, colorings, and graph homomorphisms.** Consider the following extremal questions. Given a graph G, let $i(G)$ denote the number of its independent sets, $c_q(G)$ the number of its proper q-colorings[1], and $\hom(G,H)$ the number of its graph homomorphisms to H (we allow H to have loops, and later, weights on its vertices and edges).[2]

Question 1.1. Fix d. Among d-regular graphs, which G maximizes $i(G)^{1/|V(G)|}$?

Question 1.2. Fix d and q. Among d-regular graphs, which G maximizes $c_q(G)^{1/|V(G)|}$?

Question 1.3. Fix d and H. Among d-regular graphs, which G maximizes $\hom(G,H)^{1/|V(G)|}$?

Date: September 2018 (initial); January 2020 (revised).

YZ was supported by NSF Awards DMS-1362326 and DMS-1764176, and the MIT Solomon Buchsbaum Fund.

[1] A *proper q-coloring* of G is an assignment of each vertex of G to $[q] := \{1,\dots,q\}$ so that no two adjacent vertices are assigned the same color (in particular, the colors are labeled).

[2] A *graph homomorphism* from G to H is a map of vertices $\phi\colon V(G)\to V(H)$ such that $\phi(u)\phi(v)$ is an edge of H whenever uv is an edge of G.

The third question encompasses the first two, as $i(G) = \hom(G, \bullet\!\!-\!\!\bullet)$ and $c_q(G) = \hom(G, K_q)$.

The exponential normalization is a natural choice. Indeed, replacing G by a disjoint union of copies of itself does not change the quantity $\hom(G, H)^{1/|V(G)|}$, as $\hom(G_1 \sqcup G_2, H) = \hom(G_1, H)\hom(G_2, H)$, where $\sqcup$ denotes a disjoint union.

Question 1.1 was initially raised by Granville in 1988 in connection with the Cameron–Erdős conjecture on the number of sum-free sets. Alon [1] and Kahn [27] conjectured that $G = K_{d,d}$ is the exact maximizer. Alon [1] proved an asymptotic version as $d \to \infty$. Kahn [27] proved the exact version under the additional hypothesis that G is bipartite using an elegant entropy argument that later became quite influential in the field. Zhao [38] later removed this bipartite assumption via a combinatorial argument. The results of Kahn [27] and Zhao [38] together answer Question 1.1: the maximizer is $K_{d,d}$ (unique up to taking disjoint unions of copies of itself).

Galvin and Tetali [22] initiated the study of Questions 1.2 and 1.3 and extended Kahn's entropy method [27] to prove that, under the additional hypothesis that G is bipartite, $G = K_{d,d}$ is also the maximizer for $\hom(G, H)^{1/|V(G)|}$. See Lubetzky and Zhao [32, Section 6] for a different proof using Hölder/Brascamp–Lieb type inequalities. Can the bipartite hypothesis on G also be dropped in this case? Not for all H: e.g., for $H = \bullet\ \bullet$, $G = K_{d+1}$ is the maximizer instead of $K_{d,d}$. Extending the technique for independent sets, Zhao [39] showed that the bipartite hypothesis can be dropped for certain classes of H, but the techniques failed for $H = K_q$, corresponding to colorings (Question 1.2). It remained a tantalizing conjecture to remove the bipartite hypothesis for colorings.

Recently, Davies, Jenssen, Perkins, and Roberts developed a novel technique called the "occupancy method" [15], which gave a new proof of the maximization problem for independent sets (Question 1.1). Their method reduces the problem to a (potentially large) linear program. Applying their method, they gave a computer-assisted proof of the coloring conjecture (answering Question 1.2) for $d = 3$ [16], later extended to $d = 4$ by Davies [14]. The occupancy method was later extended to other applications concerning independent sets [17, 33], as well as geometric applications concerning sphere packings [26] and spherical codes [25]. Despite its successes, the occupancy method has a number of drawbacks. Its progress on Question 1.2 requires extremely rapidly growing computational resources for larger values of d, and furthermore, the method appears to be ill-suited for irregular graphs.

Here, we answer Question 1.2 and show that $G = K_{d,d}$ is always the maximizer, thereby resolving the coloring conjecture.

Theorem 1.4. *Let G be a d-regular graph and q a positive integer. Then*

$$c_q(G)^{1/|V(G)|} \le c_q(K_{d,d})^{1/(2d)}.$$

We also prove a more general result for not necessarily regular graphs. It is analogous to our recent result [34] for independent sets, which resolved Kahn's conjecture [27]. Here is a way to phrase the question. Instead of ranging over d-regular graphs, what if we range over all graphs with a fixed *degree–degree distribution*, i.e., the distribution of the integer-pair $\{d_u, d_v\}$ over a uniform random edge $uv \in E(G)$, where d_u is the degree of $u \in V(G)$? Kahn conjectured that, for independent sets, the maximizing G, conditioned on a fixed degree-degree distribution, remains a disjoint union of complete bipartite graphs of possibly different sizes. We recently proved Kahn's conjecture, resulting in the following theorem.

Theorem 1.5 ([34])**.** *Let G be a graph without isolated vertices. Let d_v be the degree of vertex v in G. Then*

$$i(G) \le \prod_{uv \in E(G)} i(K_{d_u, d_v})^{1/(d_u d_v)}.$$

Galvin [20] conjectured (falsely) that Theorem 1.5 could be extended to $\hom(\cdot, H)$ for every H in place of $i(\cdot)$. Here we prove the extension for $H = K_q$, extending our Theorem 1.4 on the number of proper q-colorings to irregular graphs.

Theorem 1.6. *Let G be a graph without isolated vertices, and q a positive integer. Let d_v be the degree of vertex v in G. Then*

$$c_q(G) \le \prod_{uv \in E(G)} c_q(K_{d_u,d_v})^{1/(d_u d_v)}.$$

Let us state a more general version of Theorem 1.6 that interpolates between independent sets and proper colorings. Fix a finite set of *colors* Ω as well as a subset $\Omega_\circ \subseteq \Omega$, called the *looped colors*. A *semiproper coloring* of G is an assignment of each vertex of G to Ω so that for every non-looped color (i.e., a color in $\Omega \setminus \Omega_\circ$), the set of vertices of G of that color is an independent set. In other words, with $q = |\Omega|$ and $\ell = |\Omega_\circ|$, semiproper colorings correspond to homomorphisms from G to $K_q^{\ell\circ}$, where $K_q^{\ell\circ}$ is the complete graph on q vertices with exactly ℓ vertices looped. Proper colorings correspond to $\ell = 0$. Independent sets correspond to $(\ell, q) = (1, 2)$. The following theorem interpolates between Theorems 1.5 and 1.6.

Theorem 1.7. *Let G be a graph without isolated vertices, and $\ell \le q$ nonnegative integers. Let d_v be the degree of vertex v in G. Then*

$$\hom(G, K_q^{\ell\circ}) \le \prod_{uv \in E(G)} \hom(G, K_q^{\ell\circ})^{1/(d_u d_v)}.$$

Let us now move on to general graph homomorphisms. Here, Question 1.3 remains wide open. There has been a number of conjectures stated in the literature, though several of them have been falsified by counterexamples and then later revised [6, 20, 21, 22, 35]. For example, it was first conjectured [22] that the maximizer is always $G = K_{d,d}$, and then later revised [21] to $G \in \{K_{d+1}, K_{d,d}\}$, though this was later shown false too [35]. We do not even have a conjecture for what the set of possible maximizers G is. It is even unknown whether the set of potentially maximizing G is finite for each d. See the recent survey [40] for more discussion on this problem.

It is natural to restrict G in hope of a cleaner result. Cohen, Csikvári, Perkins, and Tetali [6] conjectured that among *triangle-free* graphs G, the maximizer is always $G = K_{d,d}$, extending the theorem of Galvin and Tetali [22] for bipartite G. We prove this conjecture.

Theorem 1.8. *Let G be a triangle-free d-regular graph, and H a graph allowing loops. Then*

$$\hom(G, H)^{1/|V(G)|} \le \hom(K_{d,d}, H)^{1/(2d)}.$$

We extend the result to irregular graphs and prove a corrected version of Galvin's conjecture [20].

Theorem 1.9. *Let G be a triangle-free graph without isolated vertices, and H a graph allowing loops. Then*

$$\hom(G, H) \le \prod_{uv \in E(G)} \hom(K_{d_u,d_v}, H)^{1/(d_u d_v)}.$$

Remark. Theorem 1.9 remains true even if H has vertex and edge weights, so that $\hom(G, H)$ is interpreted as the partition function for a certain "H-model" on G (e.g., the hard-core model generalizing independent sets, and the Potts model generalizing colorings). In fact, it follows by standard observations in graph limit theory [4, 31] (namely, approximating a graphon by a sequence of W-random graphs) that the weighted and unweighted versions of Theorem 1.9 are actually equivalent.

Furthermore, the triangle-free hypothesis is best possible in Theorems 1.8 and 1.9.

Proposition 1.10. *For every graph G with a triangle, there exists some graph H so that the inequality in Theorem 1.9 is false.*

The analogous minimization problem is also interesting and mysterious, though here we only mention a few known cases (see [12]). For both independent sets ($H =$ ) [13] and colorings ($H = K_q$) [2, Lemma A.1] (also see [40, Theorem 8.3]), the minimizer is K_{d+1}, whereas for the Widom–Rowlinson model ($H =$ ), the "minimizer" is the infinite d-regular tree [12].

1.2. **Graphons, norms, and reverse Sidorenko.** In the theory of graph limits [31], a *graphon* is a symmetric measurable function $W\colon \Omega \times \Omega \to [0,1]$ (symmetric means $W(x,y) = W(y,x)$), where Ω is some probability space. Define the *G-density in W* by

$$t(G,W) := \int_{\Omega^{V(G)}} \prod_{uv \in E(G)} W(x_u, x_v)\, d\boldsymbol{x}_{V(G)},$$

where $d\boldsymbol{x}_{V(G)} := \prod_{v \in V(G)} dx_v$ is the product probability measure on $\Omega^{V(G)}$.

Every graph H can be turned into a graphon $W_H\colon V(H) \times V(H) \to \{0,1\}$ by using the uniform probability measure on $V(H)$ and letting $W_H(x,y) = 1$ if $xy \in E(H)$, and $W_H(x,y) = 0$ if $xy \notin E(H)$. Then $t(G,H) := t(G,W_H) = \hom(G,H)/|V(H)|^{|V(G)|}$ is the homomorphism density of G to H. The graphon notation naturally allows us to consider edge and vertex weights on H.

Theorems 1.8 and 1.9 are equivalent to the following graphon formulation:

$$t(G,W) \le \prod_{uv \in E(G)} t(K_{d_u,d_v}, W)^{1/(d_u d_v)},$$

and in particular, for an n-vertex d-regular graph,

$$t(G,W) \le t(K_{d,d}, W)^{n/(2d)}.$$

Let us write

$$\|W\|_G := |t(G,W)|^{1/|E(G)|}.$$

Despite the suggestive notation, $\|\cdot\|_G$ is not always a norm. These quantities were first considered by Hatami [24] in connection to Sidorenko's conjecture. See the recent work of Conlon and Lee [11] addressing the question of which graphs G induce norms.

Our results above can now be written as

$$t(G,W) \le \prod_{uv \in E(G)} \|W\|_{K_{d_u,d_v}},$$

and, in particular, for d-regular graphs G,

$$\|W\|_G \le \|W\|_{K_{d,d}}.$$

In contrast, Sidorenko's conjecture says that for all bipartite graphs G, $t(G,W) \ge t(K_2,W)^{|E(G)|}$, or equivalently $\|W\|_G \ge \|W\|_{K_2}$. Sidorenko's conjecture [18, 36] has been proved for several families of graphs [3, 8, 9, 10, 24, 28, 29, 36, 37], though it remains open in general. The first open case of the conjecture is $G = K_{5,5} \setminus C_{10}$ (also known as the "Möbius strip" graph, for it is the incidence graph for a simplicial complex model of the Möbius strip viewed as gluing together five triangles). Sidorenko's conjecture proposes that $\|\cdot\|_{K_2}$ is a lower bound to $\|\cdot\|_G$, whereas our result proves an upper bound $\|\cdot\|_{K_{d,d}}$ for triangle-free d-regular graphs G. It is for this reason that we give the name *reverse Sidorenko inequality*.

1.3. **Graphical Brascamp–Lieb inequalities.** We prove a generalization of Theorem 1.9, allowing possibly different two-variable functions on every edge of G. This generalization corresponds to graph homomorphisms with list colorings, where every vertex of G is assigned an "allowable" subset of vertices of H, and we only consider homomorphisms assigning each vertex of G to one of its allowable vertices of H. This generality is actually needed as a strong induction hypothesis for our proof.

From now on, H will be a *weighted graph*, which we define to be a symmetric measurable function $H\colon \Omega \times \Omega \to \mathbb{R}_{\ge 0}$, where Ω is a measure space. We set[3]

$$\hom(G,H) := \int_{\Omega^{V(G)}} \prod_{uv\in E(G)} H(x_u,x_v)\, d\boldsymbol{x}_{V(G)}.$$

Here $d\boldsymbol{x}_{V(G)} = \prod_{v\in V(G)} dx_v$ and each dx_v is the measure on Ω, which is encoding vertex weights on H. Then Theorems 1.8 and 1.9 hold for weighted graphs H as well (see remark following Theorem 1.9).

In Section 1.2, in discussing graphons, it was important in the statement of Sidorenko's conjecture that Ω is a probability space, or else an extra normalizing factor is needed. In contrast, the inequalities that we prove in this paper are all scale-free in the sense that the measure of Ω does not have to be normalized.

The reader is welcome to think of H as an edge-weighted graph (allowing loops) on a finite set of vertices Ω (the "colors") equipped with the counting measure. By a standard graph limit argument, this case is equivalent to the general result.

Definition 1.11. For a two-variable function $f\colon \Omega_1 \times \Omega_2 \to \mathbb{R}$, define

$$\|f\|_{K_{a,b}} := \left| \int_{\Omega_1^a \times \Omega_2^b} \prod_{\substack{1\le i\le a\\ 1\le j\le b}} f(x_i,y_j)\, dx_1\cdots dx_a dy_1\cdots dy_b \right|^{1/(ab)}.$$

This quantity (again, not always a norm) can be viewed as a bipartite analog of the graph "norm" earlier, though here we do not require f to be symmetric. We have monotonicity in that if $a\le c$ and $b\le d$, and f is nonnegative-valued, then $\|f\|_{K_{a,b}} \le \|f\|_{K_{c,d}}$, which can be proved by two applications Hölder's inequality, as can be be seen after rewriting the above integral as $\int_{\Omega_1^a}\Big(\int_{\Omega_2} f(x_1,y)\cdots f(x_a,y)\,dy\Big)^b dx_1\cdots dx_a$.

The following theorem generalizes Theorem 1.9 upon taking the same $f_{uv} = H$ for all edges uv.

Theorem 1.12. *Let $G=(V,E)$ be a triangle-free graph. Let Ω_v be a measure space for each vertex $v\in V$. For each edge $uv\in E$, let $f_{uv}\colon \Omega_u\times\Omega_v\to\mathbb{R}_{\ge 0}$ be a measurable function, labeled so that $f_{uv}(x_u,x_v) = f_{vu}(x_v,x_u)$. We have*

$$\int_{\Omega_V} \prod_{uv\in E} f_{uv}(x_u,x_v)\, d\boldsymbol{x}_V \le \prod_{uv\in E} \|f_{uv}\|_{K_{d_v,d_u}},$$

where $\Omega_V := \prod_{v\in G}\Omega_v$ and $d\boldsymbol{x}_V := \prod_{v\in V} dx_v$, and d_u is the degree of u in G.

Remark. We have equality if (1) G is a disjoint union of complete bipartite graphs, or (2) if there are functions $g_v\colon \Omega_v\to\mathbb{R}_{\ge 0}$ such that $f_{uv}(x,y) = g_u(x)g_v(y)$ for every $uv\in E$.

By Proposition 1.10, the triangle-free hypothesis cannot be weakened.

For semiproper list colorings, Theorem 1.12 holds without the triangle-free hypothesis, generalizing Theorem 1.7. See Section 4 for the statement and proof.

Theorem 1.12 can be viewed as a graphical analog of the Brascamp–Lieb inequalities [5, 30], which have the form $\int f_1(B_1\boldsymbol{x})\cdots f_k(B_k\boldsymbol{x})d\boldsymbol{x} \lesssim \|f_1\|_{L^{p_1}}\cdots\|f_k\|_{L^{p_k}}$, where the B_i's are linear maps. The Brascamp–Lieb inequalities generalize classical inequalities such as Hölder's inequality and the Loomis–Whitney inequality, and have far reaching applications. Our inequality bounds a certain graphical integral in terms of graphical norm-like quantities that are in general weaker than L^p norms. It may be possible that these graphical Brascamp–Lieb inequalities have a rich theory yet to

[3]Such quantities are more commonly denoted $Z_H(G)$ for the partition function of a spin model with weights and interactions given by H. Here we prefer to extend $\hom(G,H)$ notation so as to be consistent with the case for simple graphs.

be uncovered, e.g., extensions to more general setups such as hypergraphs and simplicial complexes, allowing greater flexibility in the combinatorial form of the integral on the left-hand side of the inequality.

1.4. **Positive semidefinite models are clique-maximizing.** We have stated various results affirming that $G = K_{d,d}$ maximizes $\hom(G,H)^{1/|V(G)|}$ under various circumstances, e.g., among triangle-free G, or if H is a (partially looped) complete graph corresponding to (semi)proper colorings. However, as remarked following Theorem 1.7, $K_{d,d}$ is not always the correct answer to Question 1.3, and the general question remains very much open.

Given a weighted graph $H\colon \Omega \times \Omega \to \mathbb{R}_{\geq 0}$, we say that H is *biclique-maximizing* if it satisfies, for all graphs G without isolated vertices,

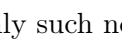
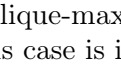
$$\hom(G,H) \leq \prod_{uv \in E(G)} \hom(K_{d_u,d_v},H)^{1/(d_u d_v)}, \tag{1.1}$$

where, as usual, d_v denotes the degree of v in G. We say that H is *clique-maximizing* if it satisfies, for all graphs G,

$$\hom(G,H) \leq \prod_{v \in V(G)} \hom(K_{d_v+1},H)^{1/(d_v+1)}. \tag{1.2}$$

Theorem 1.7 says that the partially looped complete graphs $K_q^{\ell\circ}$ are biclique-maximizing. On the other hand, it is not hard to check that a disjoint union of loops is clique-maximizing. It is known that there are graphs H that are neither biclique-maximizing nor clique-maximizing, even among d-regular graphs G (it is unknown which G achieves the maximum for such H) [35].

It was shown [39] that certain graphs H satisfy a "bipartite swapping trick" inequality: $\hom(G,H)^2 \leq \hom(G \times K_2, H)$ for all graphs G (extended to a larger class of H in [35]; also see Section 3.4 for a weighted variant). Here $G \times K_2$ is the graph with vertex set $V(G) \times \{0,1\}$ and an edge between (v,i) and $(u,1-i)$ for every $uv \in E(G)$ and $i \in \{0,1\}$. Every H satisfying this inequality is biclique-maximizing, since we can apply Theorem 1.12 to upper bound $\hom(G \times K_2, H)$ as $G \times K_2$ is bipartite and hence triangle-free.

In [6, 7, 35], it was shown that the Widom–Rowlinson model ($H =$ ) satisfies (1.2) for d-regular graphs G (this was the first and essentially only such non-trivial case that was known). However, it turns out that $H =$  is actually not clique-maximizing among irregular graphs (a counterexample is $G = K_{1,4}$, as $113 > 7^{4/2}63^{1/5}$). This case is interesting as there is a different maximization behavior between regular and irregular G.

Open problem 1.13. Determine all biclique-maximizing graphs H and all clique-maximizing graphs H, in each case, for d-regular G as well as for all G.

We say that a weighted graph (also called a *model*) $H\colon \Omega \times \Omega \to \mathbb{R}_{\geq 0}$ is *positive semidefinite* or *ferromagnetic* if the corresponding function is positive semidefinite (equivalently, the matrix $(H(x_i,x_j))_{i,j\in[n]}$ is positive semidefinite for every $x_1,\dots,x_n \in \Omega$). We say that H is *antiferromagnetic* if all eigenvalues (counting multiplicities) other than the top one are nonpositive. These definitions were taken from [19].

For example, a disjoint union of loops is ferromagnetic, whereas $K_q^{\ell\circ}$ is antiferromagnetic. For 2-spin models, i.e., $\Omega = \{0,1\}$ allowing vertex weights, H is ferromagnetic if $H(0,0)H(1,1) \geq H(0,1)^2$, and antiferromagnetic if $H(0,0)H(1,1) \leq H(0,1)^2$.

We prove the following result. See Theorem 5.1 for a list coloring type generalization.

Theorem 1.14. *Every ferromagnetic (i.e., positive semidefinite) model is clique-maximizing.*

We conjecture that the converse holds as well.

Every 2-spin model is either ferromagnetic or antiferromagnetic depending on the sign of the determinant of its 2×2 edge-weight matrix, though this is false for k-spin models for $k > 2$. As a

corollary, we completely characterize all 2-spin models, generalizing independent sets. See Section 3.4 for the antiferromagnetic part of the proof, which follows from the bipartite swapping trick [38, 39] and Theorem 1.12.

Corollary 1.15. *A 2-spin model is biclique-maximizing if it is antiferromagnetic and clique-maximizing if it is ferromagnetic.*

We close with a conjecture generalizing Theorem 1.7.

Conjecture 1.16. *Every antiferromagnetic model is biclique-maximizing.*

The converse of Conjecture 1.16 is false. The looped graph H given by the adjacency matrix $\left(\begin{smallmatrix}1&1&1\\1&1&0\\1&0&0\end{smallmatrix}\right)$ satisfies the bipartite swapping trick from [39] and hence is biclique-maximizing (see comment two paragraphs before Open problem 1.13) but H is not antiferromagnetic. It remains wide open to classify all biclique-maximizing H.

Theorem 1.7 establishes Conjecture 1.16 for $K_q^{\ell\circ}$. Though, even the following extension remains just out of reach of our current methods: $H\colon \Omega \times \Omega \to [0,1]$ with $H(x,y) = 1$ if $x \neq y$ and $0 \leq H(x,x) \leq 1$ arbitrary. This is a generalization of the antiferromagnetic Potts model. The usual Potts model has additionally the same diagonal values $H(x,x) = \beta \in [0,1]$ for all $x \in \Omega$, and for these H, the conjecture has been verified for 3-regular [16] and 4-regular [14] graphs G via the occupancy method with computer assistance.

1.5. **Relation to previous work.** This work builds on our earlier work [34] proving Kahn's conjecture on independent sets, Theorem 1.5, but requires several significantly new ideas. Our proof of Theorem 1.12 in Section 3 actually gives a new and more streamlined proof of Theorem 1.5. The new proof is significantly shorter, and it replaces a number of fairly technical inequality verifications in [34] (often involving checking repeated derivatives) by more conceptual inequalities primarily relying on Hölder's inequality and log-convexity considerations. In [34], as in the earlier [23], we relied on the recurrence $i(G) = i(G-v) + i(G-v-N(v))$ for the number independent sets, but such a relation is unavailable for colorings. Assigning a color to a vertex restricts the colors available to the neighborhoods, so it is natural to study the problem in the greater generality of list colorings and state a stronger induction hypothesis. By considering the effect of fixing a color on a vertex and carefully bounding contributions from far away vertices, we reduce the problem to more "local" inequalities. Section 2 of the paper discusses the general reduction to local inequalities in greater detail.

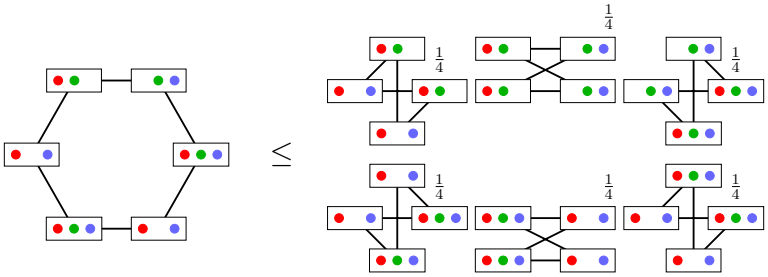
Organization. In Section 2 we give a toy calculation illustrating some proof ideas. In Section 3, we prove Theorem 1.12, the graphical Brascamp–Lieb inequality, and hence Theorems 1.8 and 1.9. In Section 4, we prove Theorem 1.7 concerning semiproper colorings, and hence Theorems 1.4 and 1.6. In Section 5, we prove Theorem 1.14 showing that ferromagnetic (i.e., positive semidefinite) models are clique-maximizing.

2. A toy calculation

In this section we sketch a toy calculation demonstrating our general inductive strategy on $G = C_6$ with target graph K_3, i.e., 3-list-coloring a 6-cycle. This is the dessert before the dinner, as the actual proof involves more difficult steps not shown here.

The inequality that we would like to prove is illustrated by the following diagram. This is a special case of Theorem 1.12 for list coloring, i.e, $f(x,y) = 1_{x\neq y}$ in Theorem 1.12 or $H = K_q$ in the graph homomorphism setup. See Theorem 4.1 for a statement of the list coloring inequality. This is an example of the strong induction hypothesis for upper bounding the number of list colorings, and we

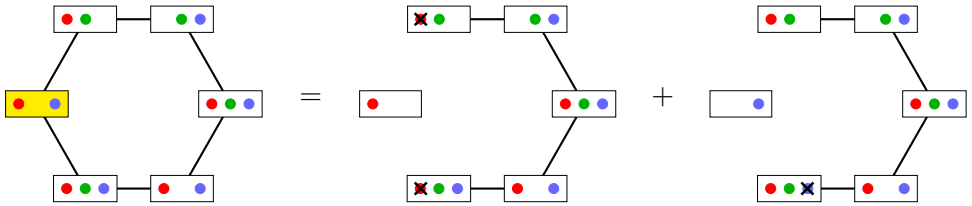
will apply induction on the number of vertices of G.



(2.1)

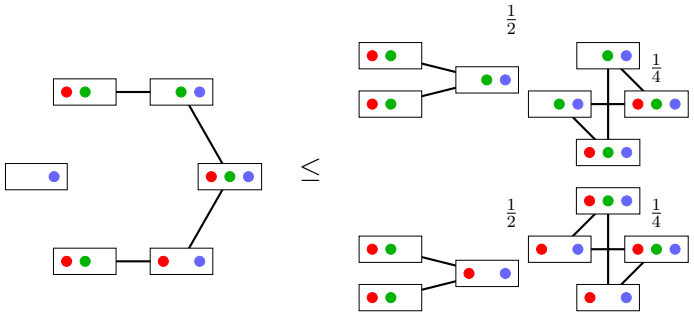
Let us explain the meaning of the above diagram. On the left-hand side, the figure should be interpreted as the number of valid list colorings of the 6-cycle where each vertex of the 6-cycle is assigned one of its listed colors, such that no adjacent vertices receive the same color. On the right hand side, we have a product of six quantities, each being the number of list colorings of a 4-cycle (with different color lists for each 4-cycle) raised to the power 1/4.

To prove (2.1), we begin by selecting the color of the left-most vertex of the 6-cycle, which gives the following:

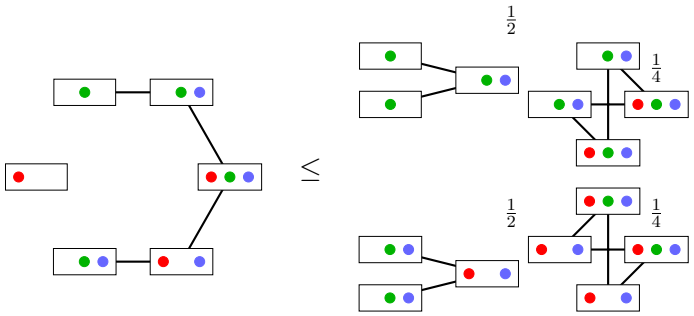


(2.2)

To upper bound the two terms on the right-hand side of (2.2), we apply the induction hypothesis to the yield the following inequalities:

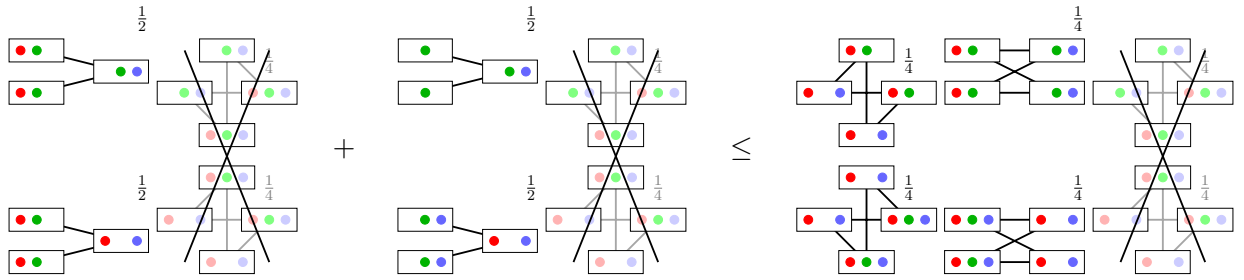


(2.3)



(2.4)

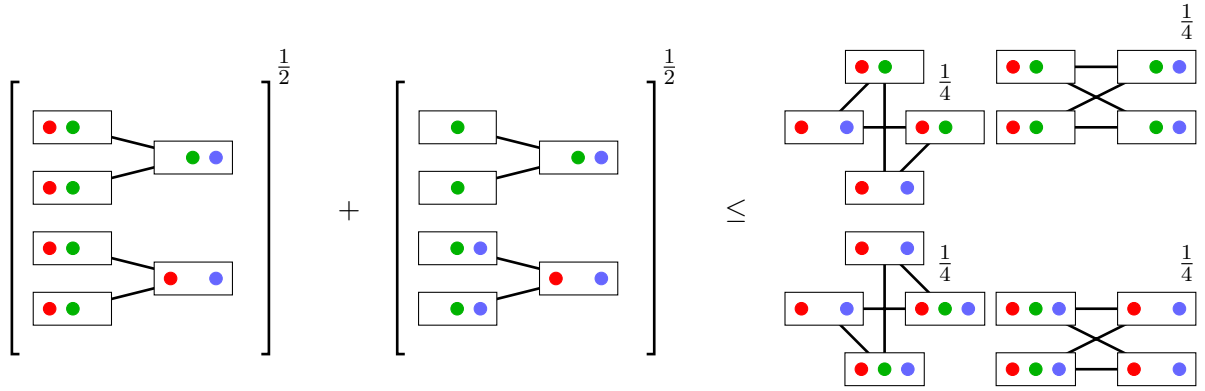
First localization. Here is the inequality that we are now left to prove:



(2.5)

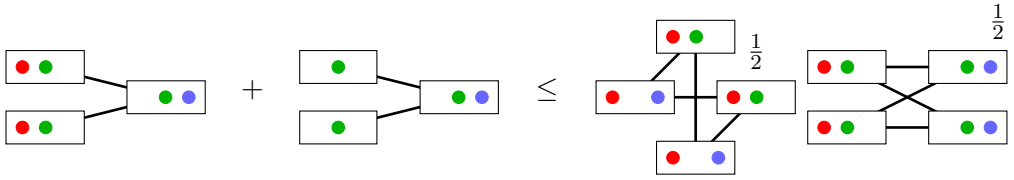
Note that all the factors associated to edges that are more than 2 edges away from the deleted vertex in $G = C_6$ are identical on both sides of the inequality, and thus they can be discarded (hence they are grayed and crossed out above).

We are left with showing the following inequality. Note that we have reduced the original inequality to a more local one involving only the edges of $G = C_6$ that are within two steps of the the deleted vertex (i.e., the edges of G incident to a neighbor of the left-most vertex)

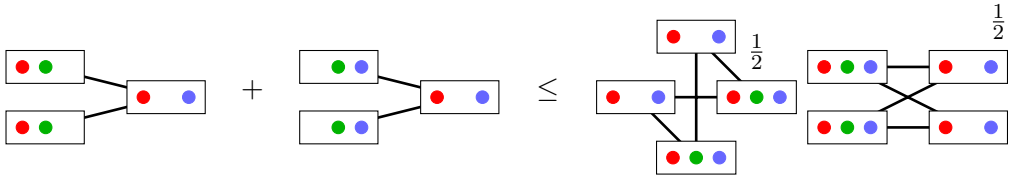


(2.6)

Second localization. Now let us apply the Cauchy–Schwarz inequality in the form of $\sqrt{a_1 b_1} + \sqrt{a_2 b_2} \le \sqrt{a_1 + a_2}\sqrt{b_1 + b_2}$ to the left-hand side above. The remaining inequality to show follows by taking the product of the following two inequalities (corresponding to the top and bottom halves of the above inequality after Cauchy–Schwarz):



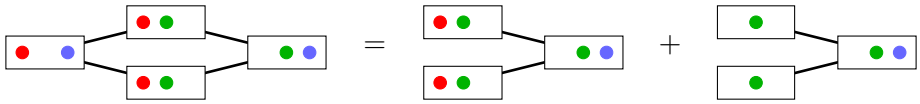
(2.7)



(2.8)

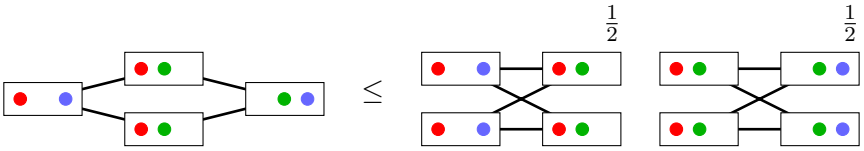
Note that each inequality now involves only a two-edge path in $G = C_6$ starting from the deleted vertex. Thus we have further localized the inequality that we wish to prove.

Let us explain how to prove (2.7), as the proof of (2.8) is analogous (in this specific example (2.8) is actually an equality). As in (2.2), the left-hand side of (2.7) can be rewritten as:



(2.9)

Now we are left with proving the following inequality. Note that it has the same form as the strong induction hypothesis (e.g., (2.1)). While it may be tempting to quote the induction hypothesis, it is instead more faithful to view this as a version of the local inequality which we need to prove.



(2.10)

This inequality follows from the Cauchy-Schwarz inequality applied as follows:

$$\begin{aligned}
&\sum_{a,b_1,b_2,c} f(a,b_1)f(a,b_2)g(b_1,c)g(b_2,c) \\
&= \sum_{b_1,b_2} \left(\sum_a f(a,b_1)f(a,b_2)\right)\left(\sum_c g(b_1,c)g(b_2,c)\right) \\
&\le \left(\sum_{b_1,b_2}\left(\sum_a f(a,b_1)f(a,b_2)\right)^2\right)^{\frac{1}{2}} \left(\sum_{b_1,b_2}\left(\sum_c g(b_1,c)g(b_2,c)\right)^2\right)^{\frac{1}{2}} \\
&= \left(\sum_{a_1,a_2,b_1,b_2} f(a_1,b_1)f(a_1,b_2)f(a_2,b_1)f(a_2,b_2)\right)^{\frac{1}{2}} \left(\sum_{b_1,b_2,c_1,c_2} g(b_1,c_1)g(b_2,c_1)g(b_1,c_2)g(b_2,c_2)\right)^{\frac{1}{2}}.
\end{aligned}$$

Here the variables $a, b_1, \dots$ range over $\{\text{red}, \text{green}, \text{blue}\}$ in the sums, $f(a,b)$ is the indicator function associated to coloring the top-left edge on the left-hand side of (2.10), i.e.,

$$f(a,b) = \begin{cases} 1 & \text{if } a \in \{\text{red}, \text{blue}\}, b \in \{\text{red}, \text{green}\}, \text{ and } a \neq b, \\ 0 & \text{otherwise,} \end{cases}$$

and $g(b,c)$ is analogously defined for the top-right edge on the left-hand side of (2.10). This completes the proof of (2.7).

Further complications. In the general setting, the first and second localization steps are analogous to the toy calculation above. In particular, the induction proceeds by first selecting the color of a maximum degree vertex w (in the above calculation, the left-most vertex), and then updating the lists of colors in the neighborhood of w for each color selection. We then apply the inductive hypothesis followed by the first localization as in (2.5), reducing the problem to just considering edges in a radius 2 neighborhood of w. The second localization is in general an application of Hölder's inequality, which reduces the problem to inequalities on two-edge paths and triangles. The analysis is somewhat easier in the triangle-free case, which is done in Section 3, as we only need to prove one type of local inequality. When G contains triangles, which is done in Section 4 for colorings, the presence of triangles require the analysis of additional local inequalities that are more difficult to

handle. The additional difficulty is expected since the local inequalities involving triangles cannot be true for all targets/models H, so the proofs need to use more specific knowledge of the model.

Even in the triangle-free case, the local inequalities for two-edge paths are in general more involved than shown above in (2.7) and (2.8). This is because the equality (2.9) turns out to be coincidental to the graph $G = C_6$; in general, one side of these inequalities is a summation of terms with fractional exponents. We handle this difficulty by defining an interpolation between the terms of the local inequality, and proving log-convexity with respect to the underlying parameter of this interpolation.

In Section 4, we handle local inequalities for triangles in the case where H is the complete graph, possibly with loops on some of its vertices. In this particular case, the result follows from additional log-convexity results along with an intricate analysis of correlation inequalities on symmetric polynomials.

For the clique-maximization result Theorem 1.14, proved in Section 5, the spirit of the solution is similar, although the execution differs. In this case, since the upper bound is a product over vertices, the analogous localization steps result in statements for each *vertex* in the neighborhood of w. Here again we have fractional exponents in general, which are handled with a series of several interpolations, each of which is shown to be log-convex.

3. Inequality on triangle-free graphs

The goal of this section is to prove Theorem 1.12.

3.1. **Some preliminary inequalities.** In the lemmas, we omit stating the obvious integrability hypotheses.

Lemma 3.1. *For nonnegative functions $g(s,u)$ and $h(s,v)$, and real $q \geq 1$, one has*

$$\left(\int \left(\int g(s,u)h(s,v)\,duds\right)^q dv\right)^{2/q} \leq \left(\int g(s,u)g(s,u')\,dudu'ds\right)\left(\int \left(\int h(s,v)h(s,v')\,ds\right)^q dvdv'\right)^{1/q}.$$

Proof. Let $1/q + 1/q' = 1$. We have

$$\begin{aligned}
\text{LHS} &= \sup_{\|f\|_{L^{q'}}\leq 1}\left(\int g(s,u)h(s,v)f(v)\,dudvds\right)^2 && [L^p \text{ duality}]\\
&\leq \sup_{\|f\|_{L^{q'}}\leq 1}\left(\int\left(\int g(s,u)\,du\right)^2 ds\right)\left(\int\left(\int h(s,v)f(v)dv\right)^2 ds\right) && [\text{Cauchy–Schwarz}]\\
&= \left(\int g(s,u)g(s,u')\,dudu'ds\right)\sup_{\|f\|_{L^{q'}}\leq 1}\int h(s,v)h(s,v')f(v)f(v')\,dvdv'\,ds,
\end{aligned}$$

which is at most the RHS, by using L^p duality once again, noting that the function $(v,v') \mapsto f(v)f(v')$ has $L^{q'}$ norm $\|f\|_{L^{q'}}^2 \leq 1$. □

Remark. Define the mixed $L^{p,q}$ matrix norm of $A = (a_{ij})$ by

$$\|A\|_{L_{p,q}} := \left(\sum_i\left(\sum_j |a_{ij}|^p\right)^{q/p}\right)^{1/q}.$$

Lemma 3.1 is equivalent to the following inequality: If $q \geq 1$, $A \in \mathbb{R}_{\geq 0}^{m\times n}$ and $B \in \mathbb{R}_{\geq 0}^{m\times k}$, then

$$\|A^\intercal B\|_{L_{1,q}}^2 \leq \|A^\intercal A\|_{L_{q,q}}\|B^\intercal B\|_{L_{1,1}}. \tag{3.1}$$

Here is the above proof written out in the language of matrices and vectors, with $\mathbf{1}_m$ denoting the all ones vector:

$$\begin{aligned}\|A^\intercal B\|_{L_{1,q}}^2 = \|A^\intercal B\mathbf{1}_m\|_q^2 &= \sup_{\|\boldsymbol{u}\|_{q'}\le 1} \langle \boldsymbol{u}, A^\intercal B\mathbf{1}_m\rangle^2 = \sup_{\|\boldsymbol{u}\|_{q'}\le 1} \langle A\boldsymbol{u}, B\mathbf{1}_m\rangle^2 \\ &\le \sup_{\|\boldsymbol{u}\|_{q'}\le 1} \langle A\boldsymbol{u}, A\boldsymbol{u}\rangle\langle B\mathbf{1}_m, B\mathbf{1}_m\rangle = \sup_{\|\boldsymbol{u}\|_{q'}\le 1} \langle \boldsymbol{u}\boldsymbol{u}^\intercal, A^\intercal A\rangle\langle \mathbf{1}_m\mathbf{1}_m^\intercal, B^\intercal B\rangle \le \|A^\intercal A\|_{L_{q,q}} \|B^\intercal B\|_{L_{1,1}}.\end{aligned}$$

We do not know if the inequality can be extended to $\|A^\intercal B\|_{L_{p,q}}^2 \le \|A^\intercal A\|_{L_{q,q}} \|B^\intercal B\|_{L_{p,p}}$ for all reals $1 \le p \le q$ (this is true for positive integer p by a tensor-power argument). Note that the above proof holds for all real matrices A and B provided that $A^\intercal A, B^\intercal B, A^\intercal B$ have nonnegative entries.

Lemma 3.2. *For nonnegative functions $f(s,t)$, $g(s,t,u)$, $h(s,t,v)$, and real $q \ge 1$,*

$$\begin{aligned}&\left(\int\left(\int f(s,t)g(s,t,u)h(s,t,v)\,duds\right)^q dtdv\right)^2 \\ &\le \left(\int\left(\int f(s,t)g(s,t,u)g(s,t,u')\,dudu'ds\right)^q dt\right)\left(\int\left(\int f(s,t)h(s,t,v)h(s,t,v')\,ds\right)^q dtdvdv'\right)\end{aligned}$$

Proof. By replacing $g(s,t,u)$ by $f(s,t)^{1/2}g(s,t,u)$ and $h(s,t,v)$ by $f(s,t)^{1/2}h(s,t,v)$, we may assume that $f = 1$. By the Cauchy–Schwarz inequality with respect to dt, the right-hand side is at least

$$\left(\int\left(\int g(s,t,u)g(s,t,u')\,dudu'ds\right)^{q/2}\left(\int\left(\int h(s,t,v)h(s,t,v')\,ds\right)^q dvdv'\right)^{1/2} dt\right)^2.$$

It suffices to show that, for every fixed t, one has

$$\begin{aligned}&\int\left(\int g(s,t,u)h(s,t,v)\,duds\right)^q dv \\ &\qquad\le \left(\int g(s,t,u)g(s,t,u')\,dudu'ds\right)^{q/2}\left(\int\left(\int h(s,t,v)h(s,t,v')\,ds\right)^q dvdv'\right)^{1/2},\end{aligned}$$

which is Lemma 3.1 applied to the functions $g(s,u) = g(s,t,u)$ and $h(s,v) = h(s,t,v)$. □

The following lemma is a "local" inequality that the proof of Theorem 1.12 will reduce to.

Lemma 3.3 (Local inequality)**.** *Let $f_{12}\colon \Omega_1 \times \Omega_2 \to \mathbb{R}_{\ge 0}$ and $f_{23}\colon \Omega_2 \times \Omega_3 \to \mathbb{R}_{\ge 0}$ be measurable functions. Let $1 \le \beta \le \Delta$ and $\gamma \ge 2$ be integers. For $x \in \Omega_1$, define $f_{23}^x\colon \Omega_2 \times \Omega_3 \to \mathbb{R}_{\ge 0}$ by $f_{23}^x(y,z) := f_{12}(x,y)^{1/(\gamma-1)} f_{23}(y,z)$ and $\|\cdot\|_{K_{a,b}}$ as in Definition 1.11. Then*

$$\int_{\Omega_1} \|f_{23}^x\|_{K_{\beta,\gamma-1}}^{\Delta(\gamma-1)}\,dx \le \|f_{12}\|_{K_{\gamma,\Delta}}^{\Delta} \|f_{23}\|_{K_{\beta,\gamma}}^{\Delta(\gamma-1)}.$$

Proof. Define, for nonnegative integers a, b, c,

$$M_{a,b,c} = \int_{\Omega_1^a}\left(\int_{\Omega_2^b\times\Omega_3^c} \prod_{i\in[a],j\in[b]} f_{12}(x^{(i)},y^{(j)}) \prod_{j\in[b],k\in[c]} f_{23}(y^{(j)},z^{(k)})\,d\boldsymbol{y}^{[b]}d\boldsymbol{z}^{[c]}\right)^{\Delta/\beta} d\boldsymbol{x}^{[a]},$$

where $d\boldsymbol{x}^{[a]} = dx^{(1)}\cdots dx^{(a)}$, $d\boldsymbol{y}^{[b]} = dy^{(1)}\cdots dy^{(b)}$, and $d\boldsymbol{z}^{[c]} = dz^{(1)}\cdots dz^{(c)}$. By expanding f_{23}^x, we have

$$\int_{\Omega_1} \|f_{23}^x\|_{K_{\beta,\gamma-1}}^{\Delta(\gamma-1)}\,dx = M_{1,\beta,\gamma-1}.$$

Also

$$\|f_{12}\|_{K_{\gamma,\Delta}}^{\gamma\Delta} = M_{\gamma,\beta,0} \quad\text{and}\quad \|f_{23}\|_{K_{\beta,\gamma}}^{\beta\gamma} = M_{0,\beta,\gamma}^{\beta/\Delta}.$$

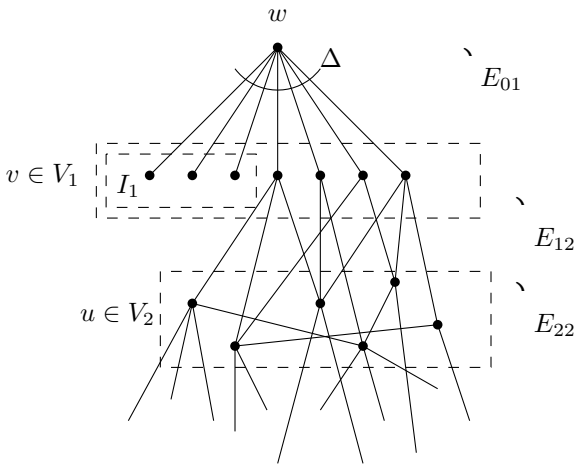


FIGURE 1. Labels of vertices and edges in the proof of Theorem 1.12.

Thus the claimed inequality can be written as

$$M_{1,\beta,\gamma-1} \le M_{\gamma,\beta,0}^{1/\gamma} M_{0,\beta,\gamma}^{1-1/\gamma},$$

which would follow from $(M_{i,\beta,\gamma-i})_{0\le i\le\gamma}$ being log-convex. Thus it suffices to prove that

$$M_{a+1,b,c+1}^2 \le M_{a+2,b,c} M_{a,b,c+2},$$

for all nonnegative integers a, b, c. This inequality follows from Lemma 3.2, after setting $q = \Delta/\beta \ge 1$, with

$$\begin{aligned}
\boldsymbol{s} &= (y^{(1)}, \dots, y^{(b)}, z^{(1)}, \dots, z^{(c)}),\\
\boldsymbol{t} &= (x^{(1)}, \dots, x^{(a)}),\\
\boldsymbol{u} &= (z^{(c+1)}), \quad \boldsymbol{u'} = (z^{(c+2)}),\\
\boldsymbol{v} &= (x^{(a+1)}), \quad \boldsymbol{v'} = (x^{(a+2)}),\\
f(\boldsymbol{s},\boldsymbol{t}) &= \prod_{i\in[a],j\in[b]} f_{12}(x^{(i)}, y^{(j)}) \prod_{j\in[b],k\in[c]} f_{23}(y^{(j)}, z^{(k)}),\\
g(\boldsymbol{s},\boldsymbol{t},\boldsymbol{u}) &= \prod_{j\in[b]} f_{23}(y^{(j)}, z^{(c+1)}), \qquad g(\boldsymbol{s},\boldsymbol{t},\boldsymbol{u'}) = \prod_{j\in[b]} f_{23}(y^{(j)}, z^{(c+2)})\\
h(\boldsymbol{s},\boldsymbol{t},\boldsymbol{v}) &= \prod_{j\in[b]} f_{12}(x^{(a+1)}, y^{(j)}), \qquad h(\boldsymbol{s},\boldsymbol{t},\boldsymbol{v'}) = \prod_{j\in[b]} f_{12}(x^{(a+2)}, y^{(j)}).
\end{aligned}$$

□

3.2. **Proof of Theorem 1.12.** We apply induction on the number of vertices in $G = (V, E)$. The base case $|E| = 0$ is trivial. Let Δ be the maximum degree of G, and let w be a vertex of degree Δ in G. The idea of the following calculation is to consider what happens when we condition on a certain color (i.e., element of Ω_w) assigned to w.

Notation. For $k \in \{0, 1, 2, \dots\} \cup \{\infty\}$, let V_k be the set of vertices at distance exactly k from w. For $0 \le i \le j \le i+1$, let E_{ij} be the edges with one endpoint in V_i and the other in V_j. Let $V_{\ge k} = \bigcup_{i\ge k} V_i$, $E_{\ge k} = \bigcup_{k\le i\le j} E_{ij}$, and $E_{>k} = E_{k,k+1} \cup E_{\ge k+1}$. Note that $V = V_{\ge 0}$ and $E = E_{>0}$.

Let I_1 be the set of vertices in V_1 whose neighborhood is exactly $\{w\}$.

Although we treat edges as unordered pairs, when we write $vu \in E_{ij}$, we always mean $v \in V_i$ and $u \in V_j$. On the other hand, when we range over $uv \in E_{ii}$, we do not count uv and vu separately.

For any $S \subseteq V$, write $\Omega_S := \prod_{v\in S} \Omega_v$ and $d\boldsymbol{x}_S := \prod_{v\in S} dx_v$.

For $vu \in E_{12}$ with $v \in V_1$ and $u \in V_2$, and $x_w \in \Omega_w$, define $f_{vu}^{x_w} \colon \Omega_v \times \Omega_u \to \mathbb{R}$ by

$$f_{vu}^{x_w}(x_v, x_u) := f_{wv}(x_w, x_v)^{1/(d_v-1)} f_{vu}(x_v, x_u).$$

By distributing the E_{01} factors to E_{12}, we have

$$\begin{aligned}
&\int_{\Omega_V} \prod_{vu \in E} f_{vu}(x_v, x_u)\, d\boldsymbol{x}_V \\
&= \int_{\Omega_{V\setminus I_1}} \prod_{v \in I_1} \left(\int_{\Omega_v} f_{wv}(x_w, x_v)\, dx_v \right) \prod_{vu \in E_{12}} \left(f_{wv}(x_w, x_v)^{1/(d_v-1)} f_{vu}(x_v, x_u) \right) \prod_{vu \in E_{\geq 2}} f_{vu}(x_v, x_u)\, d\boldsymbol{x}_{V\setminus I_1} \\
&\leq \int_{\Omega_w} \prod_{v \in I_1} \|f_{wv}(x_w, \cdot)\|_1 \prod_{vu \in E_{12}} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}} \prod_{vu \in E_{\geq 2}} \|f_{vu}\|_{K_{d_u,d_v}}\, dx_w,
\end{aligned}$$

where in the last step we applied the induction hypothesis to $G - w$ (the graph G with the vertex w removed along with all its incident edges).

Now fix $x_w \in \Omega_w$. It remains to prove the bound

$$\int_{\Omega_w} \prod_{v \in I_1} \|f_{wv}(x_w, \cdot)\|_1 \prod_{vu \in E_{12}} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}} \prod_{vu \in E_{\geq 2}} \|f_{vu}\|_{K_{d_u,d_v}}\, dx_w \leq \prod_{vu \in E} \|f_{vu}\|_{K_{d_u,d_v}}.$$

First localization. Observing that the factor $\|f_{vu}\|_{K_{d_u,d_v}}$ appears on both sides whenever $vu \in E_{\geq 2}$, we see that it suffices to prove

$$\int_{\Omega_w} \prod_{v \in I_1} \|f_{wv}(x_w, \cdot)\|_1 \prod_{vu \in E_{12}} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}}\, dx_w \leq \prod_{vu \in E_{01} \cup E_{12}} \|f_{vu}\|_{K_{d_u,d_v}}.$$

By distributing the E_{01} factors on the RHS to E_{12}, we can rewrite the above inequality as

$$\int_{\Omega_w} \prod_{v \in I_1} \|f_{wv}(x_w, \cdot)\|_1 \prod_{vu \in E_{12}} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}}\, dx_w \leq \prod_{v \in I_1} \|f_{wv}\|_{K_{1,\Delta}} \prod_{vu \in E_{12}} \left(\|f_{wv}\|_{K_{d_v,\Delta}}^{1/(d_v-1)} \|f_{vu}\|_{K_{d_u,d_v}} \right). \tag{3.2}$$

From now on until the end of the proof, by convention, we use the letter v to denote a vertex in V_1 and u for a vertex in V_2.

Second localization. Applying Hölder's inequality with exponents given by the summands of

$$\sum_{v \in I_1} \frac{1}{\Delta} + \sum_{vu \in E_{12}} \frac{1}{\Delta(d_v - 1)} = 1,$$

we upper bound the left-hand side of (3.2) by

$$\begin{aligned}
&\int_{\Omega_w} \prod_{v \in I_1} \|f_{wv}(x_w, \cdot)\|_1 \prod_{uv \in E_{12}} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}}\, dx_w \\
&\qquad \leq \prod_{v \in I_1} \left(\int_{\Omega_w} \|f_{wv}(x_w, \cdot)\|_1^{\Delta}\, dx_w \right)^{\frac{1}{\Delta}} \prod_{vu \in E_{12}} \left(\int_{\Omega_w} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}}^{\Delta(d_v-1)}\, dx_w \right)^{\frac{1}{\Delta(d_v-1)}}.
\end{aligned}$$

Comparing with the desired right-hand side of (3.2), we have

$$\int_{\Omega_w} \|f_{wv}(x_w, \cdot)\|_1^{\Delta}\, dx_w = \|f_{wv}\|_{K_{1,\Delta}}^{\Delta}$$

and, by Lemma 3.3, the local inequality, for every $vu \in E_{12}$,

$$\int_{\Omega_w} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}}^{\Delta(d_v-1)}\, dx_w \leq \|f_{wv}\|_{K_{d_v,\Delta}}^{\Delta} \|f_{vu}\|_{K_{d_u,d_v}}^{\Delta(d_v-1)},$$

which proves (3.2). This concludes the proof of Theorem 1.12. □

3.3. **Necessity of the triangle-free hypothesis.** Now we prove Proposition 1.10, showing that triangle-free hypotheses on G in Theorem 1.9 (and hence also Theorem 1.12) cannot be removed.

Proof of Proposition 1.10. Let H_ϵ be a weighted graph on two vertices each with vertex weight $1/2$, and edge-weight "adjacency" matrix

$$\begin{pmatrix} 1+2\epsilon & 1 \\ 1 & 1+2\epsilon \end{pmatrix},$$

i.e., a loop with weight $1+2\epsilon$ on each vertex, and an edge of unit weight between the two vertices. For every graph G, one has, for small ϵ,

$$\begin{aligned} \hom(G, H_\epsilon) &= \mathbb{E}_{\boldsymbol{x}\in\{0,1\}^{V(G)}} (1+2\epsilon)^{|\{uv\in E(G): x_u = x_v\}|} \\ &= 1 + |E(G)|\,\epsilon + \binom{|E(G)|}{2}\epsilon^2 + \left(\binom{|E(G)|}{3} + |T(G)|\right)\epsilon^3 + O(\epsilon^4) \\ &= (1+\epsilon)^{|E(G)|} + |T(G)|\,\epsilon^3 + O(\epsilon^4), \end{aligned}$$

where $T(G)$ is the set of triangles in G. Indeed, the coefficient of ϵ^k comes from examining each k-edge subset of $E(G)$ and determining the probability that each connected component of this k-edge subset receives the same color in $\boldsymbol{x}$. Thus,

$$\hom(G, H_\epsilon)^{1/|E(G)|} = 1 + \epsilon + \frac{|T(G)|}{|E(G)|}\epsilon^3 + O(\epsilon^4).$$

On the other hand, since $K_{a,b}$ is always triangle-free,

$$\prod_{uv\in E(G)} \hom(K_{d_u,d_v}, H)^{1/(d_u d_v)} = (1 + \epsilon + O(\epsilon^4))^{|E(G)|}.$$

Thus, for every graph G with at least one triangle, for sufficiently small $\epsilon > 0$, the weighted graph $H = H_\epsilon$ satisfies

$$\hom(G, H) > \prod_{uv\in E(G)} \hom(K_{d_u,d_v}, H)^{1/(d_u d_v)}. \tag{3.3}$$

Finally, one can obtain a simple graph H satisfying (3.3) by a standard graph limit argument (e.g., [4] [31, Ch. 10]). Indeed, note that (3.3) is also true for $H = H_\epsilon/2$ (obtained by scaling all edge weights of H_ϵ by $1/2$), since scaling the edge weights of H by a factor of 2 changes both sides of (3.3) by the same factor $2^{|E(G)|}$. Now take a sequence of simple graphs H_n converging to $H_\epsilon/2$ in the sense that $\hom(F, H_n)/|V(H_n)|^{|V(F)|} \to \hom(F, H_\epsilon/2)$ as $n \to \infty$ for every graph F. Then for sufficiently large n, the simple graph $H = H_n$ satisfies (3.3), as both sides are scaled by $|V(H)|$ raised to the same exponent $|V(G)| = \sum_{uv\in E(G)} (d_u + d_v)/(d_u d_v)$. □

3.4. **Antiferromagnetic 2-spin models are biclique-maximizing.** Here we prove the part of the claim in Corollary 1.15 that every 2-spin antiferromagnetic model is biclique-maximizing. The proof relies on the bipartite swapping trick introduced in [38] for Question 1.1 on upper bounding the number of independent sets. It was proved that $i(G)^2 \le i(G \times K_2)$, thereby reducing Question 1.1 to the bipartite setting, which had been resolved earlier [27]. The following inequality is a weighted generalization of the same inequality for independent sets. The proof is adapted from [38], and we include it here for completeness. It can also be extended further as in [39] to a larger class of weighted H, though we omit the details. See [12, Theorem 1.13] for another proof based on similar ideas.

Lemma 3.4. *Let G be a graph and H be a 2-spin antiferromagnetic model. Then*

$$\hom(G, H)^2 \le \hom(G \times K_2, H). \tag{3.4}$$

Since $G \times K_2$ is bipartite, Lemma 3.4 followed by Theorem 1.12 (or Theorem 1.9 for weighted H) gives

$$\hom(G,H) \le \hom(G \times K_2, H)^{1/2} \le \prod_{uv \in E(G)} \hom(K_{d_u,d_v}, H)^{1/(d_u d_v)},$$

as the degree-degree distribution does not change when G is lifted to $G \times K_2$.

It remains open whether (3.4) holds for $H = K_q$, corresponding to proper colorings.

Proof of Lemma 3.4. Let $G = (V,E)$. Let $\Omega = \{0,1\}$ be a two-point measure space and let $H\colon \Omega \times \Omega \to \mathbb{R}_{\ge 0}$ be antiferromagnetic, or equivalently, $H(0,0)H(1,1) \le H(0,1)^2$. Let

$$S_= = \{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) : z_{uv} \le H(x_u,x_v)H(y_u,y_v)\ \forall uv \in E\} \subseteq \Omega^V \times \Omega^V \times \mathbb{R}^E_{\ge 0}$$

and

$$S_\times = \{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) : z_{uv} \le H(x_u,y_v)H(y_u,x_v)\ \forall uv \in E\} \subseteq \Omega^V \times \Omega^V \times \mathbb{R}^E_{\ge 0},$$

where $\boldsymbol{x} = (x_v)_{v\in V(G)} \in \Omega^V$, $\boldsymbol{y} = (y_v)_{v\in V} \in \Omega^V$ and $\boldsymbol{z} = (z_e)_{e\in E} \in \mathbb{R}^E_{\ge 0}$. Note that $\hom(G,H)^2$ is equal to the measure of $S_=$, and $\hom(G \times K_2, H)$ equals to the measure of $S_\times$. Thus the lemma reduces to constructing a measure-preserving injection $\phi\colon S_= \to S_\times$.

For any $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \in \Omega^V \times \Omega^V \times \mathbb{R}^E_{\ge 0}$, say that an edge $uv \in E$ is *unsafe* with respect to $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})$ if either $z_{uv} > H(x_u,x_v)H(y_u,y_v)$ or $z_{uv} > H(x_u,y_v)H(y_u,x_v)$. Fixing $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \in S_=$, if uv is unsafe, then $H(x_u,y_v)H(y_u,x_v) < z_{uv} \le H(x_u,x_v)H(y_u,y_v)$ (the former due to being unsafe, and the latter due to the definition of $S_=$). Recall $\Omega = \{0,1\}$. Since H is 2-spin antiferromagnetic, the only way to satisfy $H(x_u,y_v)H(y_u,x_v) < H(x_u,x_v)H(y_u,y_v)$ is that one of the endpoints of uv, say u, has $(x_u,y_u) = (0,1)$, and the other endpoint v satisfies $(x_v,y_v) = (1,0)$. This shows that the unsafe edges with respect to $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})$ form a bipartite subgraph of G.

Define $\phi\colon S_= \to S_\times$ as follows. Fix some arbitrary ordering of V. For any $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \in S_=$, let T be the lexicographically-first subset of V so that every unsafe edge with respect to $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})$ has exactly one endpoint in T. Such T exists since the unsafe edges form a bipartite subgraph. Define $\phi(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = (\boldsymbol{x}',\boldsymbol{y}',\boldsymbol{z})$ by setting

$$(x'_v, y'_v) = \begin{cases} (y_v, x_v) & \text{if } v \in T, \\ (x_v, y_v) & \text{if } v \notin T. \end{cases}$$

In other words, the map ϕ swaps (x_v,y_v) for each $v \in T$.

Let us check that the image of ϕ lies in $S_\times$, we need to check that $z_{uv} \le H(x'_u,y'_v)H(y'_u,x'_v)$ for all $uv \in E$. Only unsafe edges have a chance of violating the inequality. If uv is an unsafe edge, then exactly one of (x_u,y_u) and (x_v,y_v) is swapped by ϕ, and so $H(x'_u,y'_v)H(y'_u,x'_v) = H(x_u,x_v)H(y_u,y_v) \ge z_{uv}$. Thus the image of ϕ lies in $S_\times$.

To see that ϕ is injective, note that given $\phi(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})$, we can identify the unsafe edges, which are unaffected by swapping, and then recover the lexicographically-first subset T of vertices that contains exactly one vertex from every unsafe edge, and then swap the pair (x'_v,y'_v) for every $v \in T$ to recover $(\boldsymbol{x},\boldsymbol{y})$. It is also easy to see that ϕ is a measure-preserving map, as we can partition $S_=$ into regions indexed by the set T of swapped vertices. Thus $\phi\colon S_= \to S_\times$ is a measure-preserving injection. □

4. Colorings

4.1. **Semiproper colorings.** Let us state a generalization of Theorem 1.7 to semiproper list colorings, where every vertex in G has a possibly different set of allowable colors. Recall that "colors" are synonymous with vertices of $H = K_q^{\ell\circ}$.

To state the theorem, we will need to set up some notation. Throughout this entire section, we fix a finite set of colors Ω, as well as a subset $\Omega_\circ \subseteq \Omega$ of looped colors. Recall that a *semiproper coloring* of G is an assignment of vertices of G to colors so that no two adjacent vertices of G share

the same non-looped color. The sets of colors $\Omega_\circ \subseteq \Omega$ are fixed throughout, and we omit mentioning them explicitly in the statements below.

For $A, B \subseteq \Omega$ and nonnegative integers a, b, define

$$\left\langle \begin{matrix} A & B \\ a & b \end{matrix} \right\rangle$$

to be the number of semiproper colorings of $K_{a,b}$, where the a vertices in the first vertex part of $K_{a,b}$ have their colors chosen from A, and the b vertices in the second vertex part of $K_{a,b}$ have their colors chosen from B. Observe that $\left\langle \begin{matrix} A & B \\ a & b \end{matrix} \right\rangle = \|H|_{A\times B}\|^{ab}_{K_{a,b}}$ where $H|_{A\times B}$ is the restriction of the associated partially looped complete graph H (viewed as function $\Omega \times \Omega \to \{0,1\}$) to $A \times B$.

Here is the main theorem of this section. It implies Theorem 1.7 after taking $\Omega_v = \Omega$ for all $v \in V$.

Theorem 4.1. *Let $G = (V, E)$ be a graph without isolated vertices. Assign a subset of colors $\Omega_v \subseteq \Omega$ to each $v \in V$. Then the total number of semiproper colorings of G where each $v \in V$ is assigned some color from Ω_v is at most*

$$\prod_{uv\in E} \left\langle \begin{matrix} \Omega_u & \Omega_v \\ d_v & d_u \end{matrix} \right\rangle^{1/(d_u d_v)}.$$

Here are some conventions about notation that will be maintained throughout this section:

- $A \ominus B := A \setminus (B \setminus \Omega_\circ)$, i.e., remove from A all non-looped colors in B. This is a handy operation when we consider what happens to the list of colors at the vertex after we assign colors to its neighbors.
- In $A \setminus x$, $A \cup x$, $A \ominus x$, for $x \in \Omega$, we treat x as a singleton set $\{x\}$.
- $\boldsymbol{x}$ and $\boldsymbol{y}$ refer to a vector of colors (colors are elements of Ω), and x_i refers to the i-th coordinate of $\boldsymbol{x}$.
- Given a vector $\boldsymbol{x}$, we often treat $\boldsymbol{x}$ as a subset of Ω. So $|\boldsymbol{x}|$ is the number of distinct colors appearing in $\boldsymbol{x}$, $y \cup \boldsymbol{x}$ (where $y \in \Omega$) is the union of the elements in $\boldsymbol{x}$ along with y, and $A \ominus \boldsymbol{x}$ is the set of colors left in A after we remove all non-looped colors appearing in $\boldsymbol{x}$.

4.2. **Some correlation inequalities for symmetric polynomials.** The main result of this section is the following inequality of symmetric polynomials. We will need it later for our proof of Theorem 4.1.

Proposition 4.2. *Let $\alpha_1, \dots, \alpha_n \geq 0$ be reals, and k a nonnegative integer. Let $|\boldsymbol{x}|$ denote the number of distinct entries in $\boldsymbol{x}$. Set*

$$m_\ell := \mathop{\mathbb{E}}_{\substack{\boldsymbol{x}\in[n]^k \\ |\boldsymbol{x}|=\ell}} \prod_{i=1}^{k} \alpha_{x_i}.$$

Then $m_1 \geq \cdots \geq m_{\min\{n,k\}}$.

For example, with $n = 3$ and $k = 4$, we have

$$m_1 = \tfrac{1}{3}(\alpha_1^4 + \alpha_2^4 + \alpha_3^4),$$
$$m_2 = \tfrac{1}{42}(4\alpha_1^3\alpha_2 + 4\alpha_1^3\alpha_3 + 4\alpha_2^3\alpha_1 + 4\alpha_2^3\alpha_3 + 4\alpha_3^3\alpha_1 + 4\alpha_3^3\alpha_2 + 6\alpha_1^2\alpha_2^2 + 6\alpha_1^2\alpha_3^2 + 6\alpha_2^2\alpha_3^2),$$
$$m_3 = \tfrac{1}{3}(\alpha_1^2\alpha_2\alpha_3 + \alpha_1\alpha_2^2\alpha_3 + \alpha_1\alpha_2\alpha_3^2).$$

For $S \subseteq [n]$ and $|S| \leq k$, define

$$f_{k,S} = \sum_{\substack{\boldsymbol{x}\in S^k \\ |\boldsymbol{x}|=|S|}} \prod_{i=1}^{k} \alpha_{x_i}.$$

Here $|\boldsymbol{x}| = |S|$ in the index of the summation simply says that all elements of S appear in $\boldsymbol{x}$. In other words, $f_{k,S}$ is the sum of all monomials whose set of indices is exactly S. For example,

$$f_{5,\{1,2,3\}} = 20\alpha_1^3\alpha_2\alpha_3 + 20\alpha_1\alpha_2^3\alpha_3 + 20\alpha_1\alpha_2\alpha_3^3 + 30\alpha_1^2\alpha_2^2\alpha_3 + 30\alpha_1^2\alpha_2\alpha_3^2 + 30\alpha_1\alpha_2^2\alpha_3^2.$$

Observe that $f_{k,S}$ satisfies the recursion

$$f_{k,S} = \sum_{x\in S} \alpha_x (f_{k-1,S} + f_{k-1,S\setminus x}). \tag{4.1}$$

We introduce the following averaging notation. For any polynomial P in the variables $\alpha_1, \dots,$ write $\overline{P} := P/c$ where c is the normalizing constant chosen so that $\overline{P} = 1$ whenever $\alpha_1 = \alpha_2 = \cdots = 1$, i.e., c is the sum of all coefficients (if $c = 0$, we set $\overline{P} = 0$). For example, $\overline{\alpha_1 + \alpha_2} = (\alpha_1 + \alpha_2)/2$. For this notation to make sense, we view the $\alpha_1, \dots, \alpha_n$ as formal unassigned variables. When we say that an inequality is true, we mean that it is true for all nonnegative assignments of the α_i's. This averaging notation has the convenience that we do not have to keep track of the unimportant normalization factor.

The proof of Proposition 4.2 proceeds in several steps.

Lemma 4.3. *Let $S \subseteq [n]$ and $1 \le |S| < k$. Then*

$$\overline{\sum_{x\in S} \alpha_x f_{k-1,S}} \le \overline{f_{k,S}}.$$

Proof. We apply induction on $|S| + k$. The claimed inequality is an equality when $|S| = 1$ or $k = |S| + 1$. So assume that $|S| > 1$ and $k > |S| + 1$.

Note that $(\alpha_x : x \in S)$ and $(\alpha_x f_{k-1,S\setminus x} : x \in S)$ are oppositely sorted (meaning, whenever evaluated at nonnegative assignment of the α_x's). Indeed, note that $\alpha_x f_{k-1,S\setminus x} = \alpha_x\alpha_y Q$ where Q is some polynomial with nonnegative coefficients and it does not involve the variable α_x, so that if $\alpha_x \le \alpha_y$, then swapping the two variables α_x and α_y cannot increase Q. In particular, this sortedness implies, via the rearrangement inequality,

$$\overline{\sum_{x\in S} \alpha_x^2 f_{k-1,S\setminus x}} \le \overline{\sum_{x,y\in S} \alpha_x\alpha_y f_{k-1,S\setminus x}},$$

and thus, using that $\overline{A} \le \overline{A+B}$ implies $\overline{A+B} \le \overline{B}$, we have

$$\overline{\sum_{x,y\in S} \alpha_x\alpha_y f_{k-1,S\setminus x}} \le \overline{\sum_{x\in S}\sum_{y\in S\setminus x} \alpha_x\alpha_y f_{k-1,S\setminus x}}. \tag{4.2}$$

Applying the recursion (4.1), we have

$$\overline{f_{k-1,S}} = \overline{\sum_{x\in S} \alpha_x f_{k-2,S} + \sum_{x\in S} \alpha_x f_{k-2,S\setminus x}}.$$

On the other hand, by the induction hypothesis with $(k-1, S)$, we have

$$\overline{\sum_{x\in S} \alpha_x f_{k-2,S}} \le \overline{f_{k-1,S}}.$$

Using that $\overline{A} \le \overline{A+B}$ implies $\overline{A+B} \le \overline{B}$, we have

$$\overline{f_{k-1,S}} \le \overline{\sum_{x\in S} \alpha_x f_{k-2,S\setminus x}}. \tag{4.3}$$

Therefore, we have

$$\begin{aligned}\overline{\sum_{y\in S}\alpha_y f_{k-1,S}} &\le \overline{\sum_{x,y\in S}\alpha_x\alpha_y f_{k-2,S\setminus x}} && \text{[by (4.3)]}\\ &\le \overline{\sum_{x\in S}\sum_{y\in S\setminus x}\alpha_x\alpha_y f_{k-2,S\setminus x}} && \text{[by (4.2)]}\\ &\le \overline{\sum_{x\in S}\alpha_x f_{k-1,S\setminus x}}. && \text{[by induction with } (k-1,S\setminus x)]\end{aligned}$$

The lemma then follows by using the recursion (4.1) and that $\overline{A}\le\overline{B}$ implies $\overline{A}\le\overline{A+B}$. □

Lemma 4.4. *Let $S\subseteq[n]$ and $2\le|S|\le k$. Then*

$$\overline{f_{k,S}}\le\overline{\sum_{x\in S}f_{k,S\setminus x}}.$$

Proof. We apply induction on $|S|$. When $|S|=2$, the lemma follows by noting that $\alpha_1^i\alpha_2^{k-i}+\alpha_1^{k-i}\alpha_2^i\le\alpha_1^k+\alpha_2^k$ for all $0\le i\le k$. Now assume that $|S|>2$.

Note that $(\alpha_x : x\in S)$ and $(f_{k-1,S\setminus x} : x\in S)$ are oppositely sorted. Indeed, comparing $f_{k-1,S\setminus x}$ with $f_{k-1,S\setminus y}$, we see that $f_{k-1,S\setminus x}$ does not involve α_x, and swapping all its α_y to α_x would yield $f_{k-1,S\setminus y}$. Thus, the rearrangement inequality gives

$$\overline{\sum_{x\in S}\alpha_x f_{k-1,S\setminus x}}\le\overline{\sum_{x\in S}\sum_{y\in S\setminus x}\alpha_x f_{k-1,S\setminus y}}.$$

Also, applying the induction hypotheses on $S\setminus x$ for each $x\in S$, we have

$$\overline{\sum_{x\in S}\alpha_x f_{k-1,S\setminus x}}\le\overline{\sum_{x\in S}\sum_{y\in S\setminus x}\alpha_x f_{k-1,S\setminus\{x,y\}}}.$$

Using that $\overline{A}\le\overline{B}$ and $\overline{A}\le\overline{C}$ imply $\overline{A}\le\overline{B+C}$, followed by the recursion (4.1), we obtain

$$\overline{\sum_{x\in S}\alpha_x f_{k-1,S\setminus x}}\le\overline{\sum_{x\in S}\sum_{y\in S\setminus x}\alpha_x(f_{k-1,S\setminus y}+f_{k-1,S\setminus\{x,y\}})}=\overline{\sum_{x\in S}f_{k,S\setminus x}}.$$

Lemma 4.3 and with the recursion (4.1) imply (using that $\overline{A}\le\overline{A+B}$ implies $\overline{A+B}\le\overline{B}$)

$$\overline{f_{k,S}}\le\overline{\sum_{x\in S}\alpha_x f_{k-1,S\setminus x}}.$$

The Lemma then follows from the above two inequalities. □

Proof of Proposition 4.2. Observe that

$$m_\ell=\mathop{\mathbb{E}}_{\substack{S\subseteq[n]\\|S|=\ell}}\overline{f_{k,S}}.$$

So Proposition 4.2 then follows from Lemma 4.4, as

$$m_{\ell-1}=\mathop{\mathbb{E}}_{\substack{T\subseteq[n]\\|T|=\ell-1}}\overline{f_{k,T}}=\mathop{\mathbb{E}}_{\substack{S\subseteq[n]\\|S|=\ell}}\overline{\sum_{x\in S}f_{k,S\setminus x}}\ge\mathop{\mathbb{E}}_{\substack{S\subseteq[n]\\|S|=\ell}}\overline{f_{k,S}}=m_\ell.$$ □

Proposition 4.2 has the following corollary that we will need next.

Corollary 4.5. *Let D be a finite set. Let $t \geq 1$ be real. Let $\alpha_x \geq 0$ for each $x \in D$. Let $|\boldsymbol{x}|$ denote the number of distinct elements in $\boldsymbol{x}$. Let $\tau \colon \mathbb{N}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be some non-increasing function. Then*

$$\mathop{\mathbb{E}}_{\boldsymbol{x} \in D^k} [\tau(|\boldsymbol{x}|)] \mathop{\mathbb{E}}_{\boldsymbol{x} \in D^k} \left[\prod_{i=1}^{k} \alpha_{x_i} \right] \leq \mathop{\mathbb{E}}_{\boldsymbol{x} \in D^k} \left[\tau(|\boldsymbol{x}|) \prod_{i=1}^{k} \alpha_{x_i} \right].$$

Proof. The inequality can be rewritten as

$$\mathop{\mathbb{E}}_{\boldsymbol{x} \in D^k} [\tau(|\boldsymbol{x}|)] \mathop{\mathbb{E}}_{\boldsymbol{x} \in D^k} \left[m_{|\boldsymbol{x}|} \right] \leq \mathop{\mathbb{E}}_{\boldsymbol{x} \in D^k} \left[\tau(|\boldsymbol{x}|) m_{|\boldsymbol{x}|} \right]$$

which follows from the rearrangement inequality, as $|\boldsymbol{x}|$ and $m_{|\boldsymbol{x}|}$ are oppositely sorted due to Proposition 4.2. □

4.3. **Inequalities for semiproper colorings of complete bipartite graphs.** Now we prove some "local" inequalities that will be needed in the next section.

Lemma 4.6. *Let $A, B \subseteq \Omega$. For any nonnegative integers k and $r \leq s \leq t$, we have*

$$\left\langle \begin{matrix} A & B \\ k & s \end{matrix} \right\rangle \leq \left\langle \begin{matrix} A & B \\ k & r \end{matrix} \right\rangle^{\frac{t-s}{t-r}} \left\langle \begin{matrix} A & B \\ k & t \end{matrix} \right\rangle^{\frac{s-r}{t-r}}.$$

Proof. The Lemma follows by Hölder's inequality, after expanding, for each $i \in \{r, s, t\}$,

$$\left\langle \begin{matrix} A & B \\ k & i \end{matrix} \right\rangle = \sum_{\boldsymbol{x} \in A^k} |B \ominus \boldsymbol{x}|^i.$$ □

Lemma 4.7. *Let $D \subseteq C \subseteq \Omega$ and $B \subseteq \Omega$, and suppose that $D \subseteq \Omega \setminus \Omega_\circ$. For positive integers b, c, k with $b \geq 2$, and real $t \geq 1$, we have*

$$\sum_{\boldsymbol{x} \in D^k} \prod_{i=1}^{k} \left\langle \begin{matrix} B \setminus x_i & C \setminus x_i \\ c-1 & b-1 \end{matrix} \right\rangle^{\frac{t}{b-1}} \leq \sum_{\boldsymbol{x} \in D^k} \frac{|C \setminus \boldsymbol{x}|^t}{|C|^{(1-\frac{k}{c})t}} \prod_{i=1}^{k} \left\langle \begin{matrix} B \setminus x_i & C \\ c & b-1 \end{matrix} \right\rangle^{\frac{t(c-1)}{(b-1)c}}.$$

Proof. We have

$$\left\langle \begin{matrix} B \setminus x_i & C \setminus x_i \\ c-1 & b-1 \end{matrix} \right\rangle = \sum_{\boldsymbol{y} \in (C \setminus x_i)^{b-1}} |B \ominus (x_i \cup \boldsymbol{y})|^{c-1} = (|C|-1)^{b-1} \mathop{\mathbb{E}}_{\boldsymbol{y} \in (C \setminus x_i)^{b-1}} \left[|B \ominus (x_i \cup \boldsymbol{y})|^{c-1} \right].$$

We used above that $(B \setminus x_i) \ominus \boldsymbol{y} = B \ominus (x_i \cup \boldsymbol{y})$ since $x_i \in D \subseteq \Omega \setminus \Omega_\circ$ is always non-looped. Similarly,

$$\left\langle \begin{matrix} B \setminus x_i & C \\ c-1 & b-1 \end{matrix} \right\rangle = \sum_{\boldsymbol{y} \in C^{b-1}} |B \ominus (x_i \cup \boldsymbol{y})|^{c-1} = |C|^{b-1} \mathop{\mathbb{E}}_{\boldsymbol{y} \in C^{b-1}} \left[|B \ominus (x_i \cup \boldsymbol{y})|^{c-1} \right].$$

By linearity of expectations,

$$\mathop{\mathbb{E}}_{\boldsymbol{x} \in C^k} \left[1 - \frac{|\boldsymbol{x}|}{|C|} \right] = \left(1 - \frac{1}{|C|} \right)^k.$$

It follows that the claimed inequality is equivalent to

$$\left(\mathop{\mathbb{E}}_{\boldsymbol{x} \in C^k} \left[1 - \frac{|\boldsymbol{x}|}{|C|} \right] \right)^t \mathop{\mathbb{E}}_{\boldsymbol{x} \in D^k} \left[\prod_{i=1}^{k} \mathop{\mathbb{E}}_{\boldsymbol{y} \in (C \setminus x_i)^{b-1}} \left[|B \ominus (x_i \cup \boldsymbol{y})|^{c-1} \right]^{\frac{t}{b-1}} \right]$$
$$\leq \mathop{\mathbb{E}}_{\boldsymbol{x} \in D^k} \left[\left(1 - \frac{|\boldsymbol{x}|}{|C|} \right)^t \prod_{i=1}^{k} \mathop{\mathbb{E}}_{\boldsymbol{y} \in C^{b-1}} \left[|B \ominus (x_i \cup \boldsymbol{y})|^c \right]^{\frac{t(c-1)}{c(b-1)}} \right]. \quad (4.4)$$

Applying Corollary 4.5 with $\alpha_i = \mathbb{E}_{\boldsymbol{y}\in C^{b-1}}\left[|B\ominus(x_i\cup\boldsymbol{y})|^{c-1}\right]^{\frac{t}{b-1}}$ for each $1\le i\le k$, we obtain

$$\mathop{\mathbb{E}}_{\boldsymbol{x}\in D^k}\left[\left(1-\frac{|\boldsymbol{x}|}{|C|}\right)^t\right]\mathop{\mathbb{E}}_{\boldsymbol{x}\in D^k}\left[\prod_{i=1}^k\mathop{\mathbb{E}}_{\boldsymbol{y}\in C^{b-1}}\left[|B\ominus(x_i\cup\boldsymbol{y})|^{c-1}\right]^{\frac{t}{b-1}}\right]$$
$$\le\mathop{\mathbb{E}}_{\boldsymbol{x}\in D^k}\left[\left(1-\frac{|\boldsymbol{x}|}{|C|}\right)^t\prod_{i=1}^k\mathop{\mathbb{E}}_{\boldsymbol{y}\in C^{b-1}}\left[|B\ominus(x_i\cup\boldsymbol{y})|^{c-1}\right]^{\frac{t}{b-1}}\right]. \quad (4.5)$$

For each $1\le i\le k$,

$$\mathop{\mathbb{E}}_{\boldsymbol{y}'\in(C\setminus x_i)^{b-1}}\left[\left|B\ominus(x_i\cup\boldsymbol{y}')\right|^{c-1}\right]\le\mathop{\mathbb{E}}_{\boldsymbol{y}\in C^{b-1}}\left[|B\ominus(x_i\cup\boldsymbol{y})|^{c-1}\right],$$

since we can couple $(\boldsymbol{y}',\boldsymbol{y})$ so that $x_i\cup\boldsymbol{y}'\supseteq x_i\cup\boldsymbol{y}$ for each $(\boldsymbol{y}',\boldsymbol{y})$: sample $\boldsymbol{y}$ uniformly from C^{b-1} and obtain $\boldsymbol{y}'$ by replacing every coordinate of $\boldsymbol{y}$ equal to x_i by an independent uniformly random element of $C\setminus x_i$.

Also, we have

$$\left(\mathop{\mathbb{E}}_{\boldsymbol{x}\in D^k}\left[1-\frac{|\boldsymbol{x}|}{|C|}\right]\right)^t\le\mathop{\mathbb{E}}_{\boldsymbol{x}\in D^k}\left[\left(1-\frac{|\boldsymbol{x}|}{|C|}\right)^t\right].$$

Combining the above two inequalities, we see that LHS of (4.5) upper bounds the LHS of (4.4). Also, the RHS of (4.5) lower bounds the RHS of (4.4) by the monotonicty of L^p norms. Thus (4.4) holds. □

Lemma 4.8. *Let $A,B,C\subseteq\Omega$. Let a,b,c be positive integers with $\max\{b,c\}\le a$. We have*

$$\sum_{x\in A}\left\langle\begin{matrix}B\ominus x & C\ominus x\\ c-1 & b-1\end{matrix}\right\rangle^{\frac{a}{b+c-2}}\le\left\langle\begin{matrix}A & C\\ c & a\end{matrix}\right\rangle^{\frac{b-1}{c(b+c-2)}}\left(\sum_{x\in A}\left\langle\begin{matrix}B\ominus x & C\\ c & b-1\end{matrix}\right\rangle^{\frac{a}{c}}\right)^{\frac{c-1}{b+c-2}}.$$

Proof. When $b=1$ this is in fact an equality as both sides are equal to $\sum_{x\in A}|B\ominus x|^a$. From now on assume $b\ge 2$. Raising both sides to exponent c and using $\left\langle\begin{matrix}A & C\\ c & a\end{matrix}\right\rangle=\sum_{\boldsymbol{x}\in A^c}|C\ominus\boldsymbol{x}|^a$, the inequality can be rewritten as

$$\sum_{\boldsymbol{x}\in A^c}\prod_{i=1}^c\left\langle\begin{matrix}B\ominus x_i & C\ominus x_i\\ c-1 & b-1\end{matrix}\right\rangle^{\frac{a}{b+c-2}}\le\left(\sum_{\boldsymbol{x}\in A^c}|C\ominus\boldsymbol{x}|^a\right)^{\frac{b-1}{b+c-2}}\left(\sum_{\boldsymbol{x}\in A^c}\prod_{i=1}^c\left\langle\begin{matrix}B\ominus x_i & C\\ c & b-1\end{matrix}\right\rangle^{\frac{a}{c}}\right)^{\frac{c-1}{b+c-2}}.$$

Applying Hölder's inequality to the right-hand side, we see that it suffices to prove

$$\sum_{\boldsymbol{x}\in A^c}\prod_{i=1}^c\left\langle\begin{matrix}B\ominus x_i & C\ominus x_i\\ c-1 & b-1\end{matrix}\right\rangle^{\frac{a}{b+c-2}}\le\sum_{\boldsymbol{x}\in A^c}|C\ominus\boldsymbol{x}|^{\frac{a(b-1)}{b+c-2}}\prod_{i=1}^c\left\langle\begin{matrix}B\ominus x_i & C\\ c & b-1\end{matrix}\right\rangle^{\frac{a(c-1)}{(b+c-2)c}}.$$

Let $D:=(A\cap C)\setminus\Omega_\circ\subseteq\Omega\setminus\Omega_\circ$. It suffices to show that the above inequality holds with a partial summation where we hold fixed the coordinates of $\boldsymbol{x}$ lying outside D and let the other coordinates range over D. In other words, let $K\subseteq[c]$. Fix $x_i\in A\setminus D$ for each $i\notin K$. Writing $\boldsymbol{x}_K=(x_i)_{i\in K}$, it suffices to show that

$$\sum_{\boldsymbol{x}_K\in D^K}\prod_{i=1}^c\left\langle\begin{matrix}B\ominus x_i & C\ominus x_i\\ c-1 & b-1\end{matrix}\right\rangle^{\frac{a}{b+c-2}}\le\sum_{\boldsymbol{x}_K\in D^K}|C\ominus\boldsymbol{x}|^{\frac{a(b-1)}{b+c-2}}\prod_{i=1}^c\left\langle\begin{matrix}B\ominus x_i & C\\ c & b-1\end{matrix}\right\rangle^{\frac{a(c-1)}{(b+c-2)c}}. \quad (4.6)$$

Let $k = |K|$ Applying Lemma 4.7 with $t = \frac{a(b-1)}{b+c-2} \geq \frac{c(b-1)}{b+c-2} \geq 1$, we obtain

$$\sum_{\boldsymbol{x}_K \in D^K} \prod_{i \in K} \left\langle \begin{matrix} B \setminus x_i & C \setminus x_i \\ c-1 & b-1 \end{matrix} \right\rangle^{\frac{a}{b+c-2}} \leq \sum_{\boldsymbol{x}_K \in D^K} \left(\prod_{i \in K} \left\langle \begin{matrix} B \setminus x_i & C \\ c & b-1 \end{matrix} \right\rangle^{\frac{a(c-1)}{(b+c-2)c}} \right) \frac{|C \setminus \boldsymbol{x}_K|^{\frac{a(b-1)}{b+c-2}}}{|C|^{\frac{a(b-1)(c-k)}{c(b+c-2)}}}. \quad (4.7)$$

For each $i \notin K$, either $x_i \in \Omega_\circ$ or $x_i \notin C$, so $C \ominus x_i = C$. Applying Lemma 4.6,

$$\left\langle \begin{matrix} B \ominus x_i & C \ominus x_i \\ c-1 & b-1 \end{matrix} \right\rangle = \left\langle \begin{matrix} B \ominus x_i & C \\ c-1 & b-1 \end{matrix} \right\rangle$$
$$\leq \left\langle \begin{matrix} B \ominus x_i & C \\ c & b-1 \end{matrix} \right\rangle^{\frac{c-1}{c}} \left\langle \begin{matrix} B \ominus x_i & C \\ 0 & b-1 \end{matrix} \right\rangle^{\frac{1}{c}} = \left\langle \begin{matrix} B \ominus x_i & C \\ c & b-1 \end{matrix} \right\rangle^{\frac{c-1}{c}} |C|^{\frac{b-1}{c}}.$$

By multiplying onto (4.7) the above inequality raised to exponent $a/(b+c-2)$ and ranged over all $i \in [c] \setminus K$, we obtain (4.6). Note that $C \setminus \boldsymbol{x}_K = C \ominus \boldsymbol{x}$ since $x_i \in D \subseteq C \setminus \Omega_\circ$ for all $i \in K$. □

The following lemma is the "local" inequality that the proof of Theorem 4.1 will reduce to.

Lemma 4.9 (Local inequality for semiproper colorings)**.** *Let $A, B, C \subseteq \Omega$. Let a, b, c be positive integers with $\max\{b, c\} \leq a$. We have*

$$\sum_{x \in A} \left\langle \begin{matrix} B \ominus x & C \ominus x \\ c-1 & b-1 \end{matrix} \right\rangle^{\frac{a}{b+c-2}} \leq \left\langle \begin{matrix} A & B \\ b & a \end{matrix} \right\rangle^{\frac{c-1}{b(b+c-2)}} \left\langle \begin{matrix} A & C \\ c & a \end{matrix} \right\rangle^{\frac{b-1}{c(b+c-2)}} \left\langle \begin{matrix} B & C \\ c & b \end{matrix} \right\rangle^{\frac{a(b-1)(c-1)}{(b+c-2)bc}}.$$

Proof. When $b = 1$ we have an equality: $\sum_{x \in A} |B \ominus x|^a = \left\langle \begin{smallmatrix} A & B \\ 1 & a \end{smallmatrix} \right\rangle$. Now let $b \geq 2$. Applying Lemma 3.3 with $\Delta = a$, $\beta = c$, $\gamma = b$, and $f_{12} \colon A \times B \to \{0, 1\}$ and $f_{23} \colon B \times C \to \{0, 1\}$ the color compatibility functions (i.e., $f(x, y) = 1$ if $x = y \notin \Omega_\circ$ and $f(x, y) = 1$ otherwise; so $\|f_{12}\|^{st}_{K_{s,t}} = \left\langle \begin{smallmatrix} A & B \\ s & t \end{smallmatrix} \right\rangle$ and likewise with f_{23}), we obtain

$$\sum_{x \in A} \left\langle \begin{matrix} B \ominus x & C \\ c & b-1 \end{matrix} \right\rangle^{\frac{a}{c}} \leq \left\langle \begin{matrix} A & B \\ b & a \end{matrix} \right\rangle^{\frac{1}{b}} \left\langle \begin{matrix} B & C \\ c & b \end{matrix} \right\rangle^{\frac{a(b-1)}{bc}}.$$

The Lemma follows by bounding the right-hand side of the inequality in Lemma 4.8 using the above inequality. □

4.4. **Proof of Theorem 4.1.** Now we are ready to prove Theorem 4.1. We proceed similarly to Section 3.2, with an important new twist, namely that the neighborhood of a vertex is no longer necessarily an independent set. This explains the needs for the more involved inequalities for semiproper colorings seen earlier. We begin by recalling the notation used in Section 3.2 which will also be used in this section.

We apply induction on the number of vertices in $G = (V, E)$. The base case $|E| = 0$ is trivial. Let Δ be the maximum degree of G, and fix a vertex w of degree Δ in G.

For $k \in \{0, 1, 2, \dots\} \cup \{\infty\}$, let V_k be the set of vertices at distance exactly k from w. For $0 \leq i \leq j \leq i + 1$, let E_{ij} be the edges with one endpoint in V_i and the other in V_j. Let $V_{\geq k} = \bigcup_{i \geq k} V_i$, $E_{\geq k} = \bigcup_{k \leq i \leq j} E_{ij}$, and $E_{>k} = E_{k,k+1} \cup E_{\geq k+1}$. Note that $V = V_{\geq 0}$ and $E = E_{>0}$. Note that unlike in the triangle-free setting, now E_{11} may be nonempty.

Let I_1 be the set of vertices in V_1 whose neighborhood is exactly $\{w\}$.

Although we treat edges as unordered pairs, when we write $vu \in E_{ij}$, we always mean $v \in V_i$ and $u \in V_j$. On the other hand, when we range over $uv \in E_{ii}$, we do not count uv and vu separately.

We use the variable x_v to denote an element of Ω_v. For any $S \subseteq V$, write $\Omega_S := \prod_{v \in S} \Omega_v$ and $d\boldsymbol{x}_S := \prod_{v \in S} dx_v$.

We associate to each edge uv the function $f_{uv}\colon \Omega_u \times \Omega_v \to \{0,1\}$ encoding validity of color assignments:

$$f_{uv}(x_u,x_v) = \begin{cases} 0 & \text{if } x_u = x_v \notin \Omega_\circ, \\ 1 & \text{otherwise.} \end{cases}$$

Define

$$f_{vu}^{x_w}(x_v,x_u) := f_{wv}(x_w,x_v)^{1/(d_v-1)} f_{vu}(x_v,x_u), \quad \text{for } vu \in E_{12} \cup E_{12},$$

and

$$f_{vu}^{x_w,x_w}(x_v,x_u) := f_{wv}(x_w,x_v)^{1/(d_v-1)} f_{wv}(x_w,x_u)^{1/(d_u-1)} f_{vu}(x_v,x_u), \quad \text{for } vu \in E_{11}.$$

We have

$$\begin{aligned}
&\int_{\Omega_V} \prod_{vu\in E} f_{vu}(x_v,x_u)\, d\boldsymbol{x}_V \\
&= \int_{\Omega_{V\setminus I_1}} \prod_{v\in I_1} \left(\int_{\Omega_v} f_{wv}(x_w,x_v)\, dx_v \right) \prod_{vu\in E_{11}} f_{vu}^{x_w,x_w}(x_v,x_u) \prod_{vu\in E_{12}} f_{vu}^{x_w}(x_v,x_u) \prod_{vu\in E_{\geq 2}} f_{vu}(x_v,x_u)\, d\boldsymbol{x}_{V\setminus I_1} \\
&\leq \int_{\Omega_w} \prod_{v\in I_1} \|f_{wv}(x_w,\cdot)\|_1 \prod_{vu\in E_{11}} \|f_{vu}^{x_w,x_w}\|_{K_{d_u-1,d_v-1}} \prod_{vu\in E_{12}} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}} \prod_{vu\in E_{\geq 2}} \|f_{vu}\|_{K_{d_u,d_v}}\, dx_w,
\end{aligned}$$

where in the last step we applied the induction hypothesis to $G-w$. It remains to prove the bound

$$\begin{aligned}
\int_{\Omega_w} \prod_{v\in I_1} \|f_{wv}(x_w,\cdot)\|_1 \prod_{vu\in E_{11}} \|f_{vu}^{x_w,x_w}\|_{K_{d_u-1,d_v-1}} \prod_{vu\in E_{12}} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}} \prod_{vu\in E_{\geq 2}} \|f_{vu}\|_{K_{d_u,d_v}}\, dx_w \\
\leq \prod_{vu\in E} \|f_{vu}\|_{K_{d_u,d_v}}.
\end{aligned}$$

First localization. Observing that the factor $\|f_{vu}\|_{K_{d_u,d_v}}$ appears on both sides whenever $vu \in E_{\geq 2}$, we see that it suffices to prove

$$\begin{aligned}
\int_{\Omega_w} \prod_{v\in I_1} \|f_{wv}(x_w,\cdot)\|_1 \prod_{vu\in E_{11}} \|f_{vu}^{x_w,x_w}\|_{K_{d_u-1,d_v-1}} \prod_{vu\in E_{12}} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}}\, dx_w \\
\leq \prod_{vu\in E_{01}\cup E_{11}\cup E_{12}} \|f_{vu}\|_{K_{d_u,d_v}}.
\end{aligned}$$

By distributing the E_{01} factors on the RHS to $E_{11}\cup E_{12}$, we can rewrite the above inequality as

$$\begin{aligned}
&\int_{\Omega_w} \prod_{v\in I_1} \|f_{wv}(x_w,\cdot)\|_1 \prod_{vu\in E_{11}} \|f_{vu}^{x_w,x_w}\|_{K_{d_u-1,d_v-1}} \prod_{vu\in E_{12}} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}}\, dx_w \\
&\qquad \leq \prod_{v\in I_1} \|f_{wv}\|_{K_{1,\Delta}} \prod_{vu\in E_{11}} \left(\|f_{wv}\|_{K_{d_v,\Delta}}^{1/(d_v-1)} \|f_{wv}\|_{K_{d_u,\Delta}}^{1/(d_u-1)} \|f_{vu}\|_{K_{d_u,d_v}} \right) \\
&\qquad\qquad \cdot \prod_{vu\in E_{12}} \left(\|f_{wv}\|_{K_{d_v,\Delta}}^{1/(d_v-1)} \|f_{vu}\|_{K_{d_u,d_v}} \right). \qquad (4.8)
\end{aligned}$$

Second localization. Applying Hölder's inequality with exponents given by the summands of

$$\sum_{v\in I_1} \frac{1}{\Delta} + \sum_{vu\in E_{11}} \left(\frac{1}{\Delta(d_v-1)} + \frac{1}{\Delta(d_u-1)} \right) + \sum_{vu\in E_{12}} \frac{1}{\Delta(d_v-1)} = 1,$$

we upper bound the left-hand side of (4.8) by

$$\int_{\Omega_w} \prod_{v\in I_1} \|f_{wv}(x_w,\cdot)\|_1 \prod_{vu\in E_{11}} \|f_{vu}^{x_w,x_w}\|_{K_{d_u-1,d_v-1}} \prod_{vu\in E_{12}} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}}\, dx_w$$
$$\le \prod_{v\in I_1} \left(\int_{\Omega_w} \|f_{wv}(x_w,\cdot)\|_1^{\Delta}\, dx_w\right)^{\frac{1}{\Delta}} \prod_{vu\in E_{11}} \left(\int_{\Omega_w} \|f_{vu}^{x_w,x_w}\|_{K_{d_u-1,d_v-1}}^{\frac{\Delta(d_v-1)(d_u-1)}{d_v+d_u-2}}\, dx_w\right)^{\frac{d_v+d_u-2}{\Delta(d_v-1)(d_u-1)}}$$
$$\cdot \prod_{vu\in E_{12}} \left(\int_{\Omega_w} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}}^{\Delta(d_v-1)}\, dx_w\right)^{\frac{1}{\Delta(d_v-1)}}. \quad (4.9)$$

We have

$$\int_{\Omega_w} \|f_{wv}(x_w,\cdot)\|_1^{\Delta}\, dx_w = \|f_{wv}\|_{K_{1,\Delta}}^{\Delta}.$$

For every $vu \in E_{12}$, applying Lemma 3.3 yields

$$\int_{\Omega_w} \|f_{vu}^{x_w}\|_{K_{d_u,d_v-1}}^{\Delta(d_v-1)}\, dx_w \le \|f_{wv}\|_{K_{d_v,\Delta}}^{\Delta} \|f_{vu}\|_{K_{d_u,d_v}}^{\Delta(d_v-1)}.$$

For every $vu \in E_{11}$, applying Lemma 4.9 with $a=\Delta$, $b=d_v$, $c=d_u$, noting that $\|f_{uv}\|_{K_{s,t}}^{st} = \left\langle \begin{smallmatrix} \Omega_u & \Omega_v \\ s & t \end{smallmatrix} \right\rangle$ (and likewise with other edges), we obtain

$$\int_{\Omega_w} \|f_{vu}^{x_w,x_w}\|_{K_{d_u-1,d_v-1}}^{\frac{\Delta(d_v-1)(d_u-1)}{d_v+d_u-2}}\, dx_w \le \|f_{wv}\|_{K_{d_v,\Delta}}^{\frac{\Delta(d_u-1)}{d_v+d_u-2}} \|f_{wu}\|_{K_{d_u,\Delta}}^{\frac{\Delta(d_v-1)}{d_v+d_u-2}} \|f_{vu}\|_{K_{d_u,d_v}}^{\frac{\Delta(d_v-1)(d_u-1)}{d_v+d_u-2}}. \quad (4.10)$$

Applying the above three inequalities to upper bound the right-hand side of (4.9), we obtain (4.8), and thereby complete the proof of Theorem 4.1. □

5. Clique maximizers for positive semidefinite models

In this section we prove Theorem 1.14 that for every positive semidefinite (i.e., ferromagnetic) $H\colon \Omega\times\Omega\to\mathbb{R}_{\ge 0}$, one has

$$\hom(G,H) \le \prod_{v\in V(G)} \hom(K_{d_v+1},H)^{1/(d_v+1)}.$$

We introduce some notation that allows us to alter the vertex weights of H separately for each vertex of G. It can be thought of as a weighted version of list coloring, where each vertex of G has a different vector of weights on the set of "colors" (the vertices of H). Given graph $G=(V,E)$ and weighted graph $H\colon \Omega\times\Omega\to\mathbb{R}_{\ge 0}$, and vector $\boldsymbol{\lambda}=(\lambda_v)_{v\in V}$ whose coordinates $\lambda_v\colon \Omega\to\mathbb{R}_{\ge 0}$ are measurable functions, write

$$\hom_{\boldsymbol{\lambda}}(G,H) := \int_{\Omega^V} \prod_{uv\in E} H(x_u,x_v) \prod_{v\in V} \lambda_v(x_v)dx_v.$$

For a single $\lambda\colon \Omega\to\mathbb{R}_{\ge 0}$, we write $\hom_{\lambda}(G,H)$ to mean $\hom_{\boldsymbol{\lambda}}(G,H)$ with $\lambda_v=\lambda$ for all $v\in G$. Observe that when $H=K_q$ and every λ_v is the indicator function on a subset of $V(H)$, the above quantity is precisely the usual list coloring.

The following theorem generalizes Theorem 1.14 by taking $\lambda_v=1$ for all $v\in V$.

Theorem 5.1. *Let $G=(V,E)$ be a graph and $H\colon \Omega\times\Omega\to\mathbb{R}_{\ge 0}$ a positive semidefinite weighted graph. Let $\boldsymbol{\lambda}=(\lambda_v)_{v\in V}$ where $\lambda_v\colon \Omega\to\mathbb{R}_{\ge 0}$ is a measurable function. Then*

$$\hom_{\boldsymbol{\lambda}}(G,H) \le \prod_{v\in V} \hom_{\lambda_v}(K_{d_v+1},H)^{1/(d_v+1)}.$$

Note that the positive semidefiniteness hypothesis is optimal, since for $G = K_2$ the above inequality is just the Cauchy–Schwarz inequality for the bilinear form in $\mathbb{R}^\Omega$ given by H.

Our proof of Theorem 5.1 proceeds by induction on the maximum degree of G. We begin with a lemma that is essentially an application of the Cauchy–Schwarz inequality. Let $G^\bullet$ be the graph obtained from G by adding a new vertex adjacent to all other vertices. Let $G^{\bullet\bullet}$ be the graph obtained from G by adding two new vertices adjacent to all vertices of G but not to each other.

Lemma 5.2. *Let $G = (V, E)$ be a graph and $H\colon \Omega\times\Omega \to \mathbb{R}_{\geq 0}$ a weighted graph. Let $\boldsymbol{\lambda} = (\lambda_v)_{v\in V(G^{\bullet\bullet})}$, $\boldsymbol{\mu} = (\mu_v)_{v\in V(G)}$, and $\boldsymbol{\nu} = (\nu_v)_{v\in V(G^\bullet)}$ be vectors of measurable functions $\Omega \to \mathbb{R}_{\geq 0}$ such that for every $v \in V(G)$, one has $\lambda_v\mu_v = \nu_v^2$. Furthermore, assume that the entries of $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ associated to the new vertices (i.e., the vertices not in V) are all identical. Then*

$$\hom_{\boldsymbol{\lambda}}(G^{\bullet\bullet}, H)\hom_{\boldsymbol{\mu}}(G, H) \geq \hom_{\boldsymbol{\nu}}(G^\bullet, H)^2.$$

Proof. For $\boldsymbol{x}_V = (x_v)_{v\in V} \in \Omega^V$, we write

$$H^G(\boldsymbol{x}_V) := \prod_{uv\in E} H(x_u, x_v),$$

and

$$\boldsymbol{\lambda}(\boldsymbol{x}_V) = \prod_{v\in V}\lambda_v(x_v), \quad \boldsymbol{\mu}(\boldsymbol{x}_V) = \prod_{v\in V}\mu_v(x_v), \quad \text{and} \quad \boldsymbol{\nu}(\boldsymbol{x}_V) = \prod_{v\in V}\nu_v(x_v).$$

Recall that the entries of $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ associated to the new vertices are all identical, which we call $\nu_\bullet\colon \Omega \to \mathbb{R}_{\geq 0}$. We have

$$\hom_{\boldsymbol{\lambda}}(G^{\bullet\bullet}, H) = \int_{\Omega^V}\left(\int_\Omega \prod_{v\in V} H(x_v, y)\,\nu_\bullet(y)dy\right)^2 H^G(\boldsymbol{x}_V)\boldsymbol{\lambda}(\boldsymbol{x}_V)\,d\boldsymbol{x}_V,$$

$$\hom_{\boldsymbol{\mu}}(G, H) = \int_{\Omega^V} H^G(\boldsymbol{x}_V)\boldsymbol{\mu}(\boldsymbol{x}_V)\,d\boldsymbol{x}_V, \quad \text{and}$$

$$\hom_{\boldsymbol{\nu}}(G^\bullet, H) = \int_{\Omega^V}\left(\int_\Omega \prod_{v\in V} H(x_v, y)\,\nu_\bullet(y)dy\right) H^G(\boldsymbol{x}_V)\boldsymbol{\nu}(\boldsymbol{x}_V)\,d\boldsymbol{x}_V.$$

The Lemma then follows from the Cauchy–Schwarz inequality applied with respect to $d\boldsymbol{x}_V$, noting that $\boldsymbol{\lambda}(\boldsymbol{x}_V)\boldsymbol{\mu}(\boldsymbol{x}_V) = \boldsymbol{\nu}(\boldsymbol{x}_V)^2$. □

Set

$$h_a(\lambda) := \hom_\lambda(K_a, H).$$

In particular, $h_0(\lambda) = 1$.

Lemma 5.3. *Suppose that Theorem 5.1 holds for all G with maximum degree less than Δ. Let $2 \leq t < \Delta$ be a positive integer. Let $\lambda, \mu, \nu\colon \Omega \to \mathbb{R}_{\geq 0}$ be measurable functions satisfying $\lambda\mu = \nu^2$ pointwise. Then*

$$h_{t+1}(\lambda)^{1/(t+1)}h_{t-1}(\mu)^{1/(t-1)} \geq h_t(\nu)^{2/t}.$$

Proof. Define $\boldsymbol{\lambda}' = (\lambda'_v)_{v\in V(K_{t-1}^{\bullet\bullet})}$ by setting $\lambda'_v = \lambda$ for each vertex v in the original K_{t-1} of $K_{t-1}^{\bullet\bullet}$ and $\lambda'_v = \nu$ for the two other vertices v (each of degree $t-1$). By Lemma 5.2,

$$\hom_{\boldsymbol{\lambda}'}(K_{t-1}^{\bullet\bullet}, H)\hom_\mu(K_{t-1}, H) \geq \hom_\nu(K_{t-1}^\bullet, H)^2.$$

Since the maximum degree of $K_{t-1}^{\bullet\bullet}$ is $t < \Delta$, using the hypothesis of the Lemma to apply Theorem 5.1 to $K_{t-1}^{\bullet\bullet}$, we have

$$\hom_{\boldsymbol{\lambda}'}(K_{t-1}^{\bullet\bullet}, H) \leq h_{t+1}(\lambda)^{\frac{t-1}{t+1}}h_t(\nu)^{\frac{2}{t}},$$

so that the previous inequality implies that

$$h_{t+1}(\lambda)^{\frac{t-1}{t+1}}h_t(\nu)^{\frac{2}{t}}h_{t-1}(\mu) \geq h_t(\nu)^2.$$

The Lemma follows after rearranging. □

Remark. The hypothesis $t < \Delta$ in Lemma 5.3 is important for applying the induction hypothesis.

Lemma 5.4. *Let $H\colon \Omega \to \mathbb{R}_{\geq 0}$ be positive semidefinite. Let a be a positive integer. Let $\mu, \nu\colon \Omega \to \mathbb{R}_{\geq 0}$ be measurable functions. For each $0 \leq i \leq a$, let*

$$F_i = \hom_{\boldsymbol{\lambda}}(K_a, H), \quad \text{where } \boldsymbol{\lambda} = (\underbrace{\mu, \dots, \mu}_{i \text{ times}}, \underbrace{\nu, \dots, \nu}_{a-i \text{ times}}).$$

Then $F_0, \dots, F_a$ is log-convex ($F_s F_{s+2} \geq F_{s+1}^2$ for each $0 \leq s \leq a-2$), and consequently, $F_0^{a-1} F_a \geq F_1^a$.

Proof. Let $0 \leq s \leq a-1$. Write $T = [a] \setminus \{s+1, s+2\}$, $\boldsymbol{x}_T = (x_i)_{i \in T} \in \Omega^{a-2}$,

$$\ell_1(x; \boldsymbol{x}_T) = \nu(x) \prod_{i \in T} H(x, x_i),$$

$$\ell_2(x; \boldsymbol{x}_T) = \mu(x) \prod_{i \in T} H(x, x_i),$$

and

$$g = g(\boldsymbol{x}_T) = \prod_{i,j \in T : i \neq j} H(x_i, x_j) \prod_{i=1}^{s} \mu(x_i) \prod_{i=s+3}^{a} \nu(x_i).$$

For $i, j \in \{1, 2\}$, write

$$L_{ij} = L_{ij}(\boldsymbol{x}_T) = \int_{\Omega^2} H(x, y) \ell_i(x; \boldsymbol{x}_T) \ell_j(y; \boldsymbol{x}_T) \, dx dy.$$

By the positive semidefiniteness of H, we have, for every $\boldsymbol{x}_T \in \Omega^{a-2}$,

$$L_{11} L_{22} \geq L_{12}^2.$$

We can write

$$F_s = \int_{\Omega^{a-2}} g L_{11} \, d\boldsymbol{x}_T, \quad F_{s+1} = \int_{\Omega^{a-2}} g L_{12} \, d\boldsymbol{x}_T, \quad \text{and} \quad F_{s+2} = \int_{\Omega^{a-2}} g L_{22} \, d\boldsymbol{x}_T.$$

By the Cauchy–Schwarz inequality, we have

$$F_s F_{s+2} \geq \left(\int_{\Omega^{a-2}} g \sqrt{L_{11} L_{22}} \, \boldsymbol{x}_T \right)^2 \geq \left(\int_{\Omega^{a-2}} g L_{12} \, \boldsymbol{x}_T \right)^2 = F_{s+1}^2. \qquad \square$$

Lemma 5.5. *Let $b \leq a \leq \Delta$ be positive integers. Suppose that Theorem 5.1 holds for all graphs G with maximum degree less than Δ. Let $H\colon \Omega \times \Omega \to \mathbb{R}_{\geq 0}$ be positive semidefinite. Let $\lambda, \mu\colon \Omega \to \mathbb{R}_{\geq 0}$. Define $\eta\colon \Omega \to \mathbb{R}_{\geq 0}$ by*

$$\eta(x) := h_b(\mu H(x, \cdot))^{1/b}$$

(here $\mu H(x, \cdot)$ is the pointwise product of two functions $\Omega \to \mathbb{R}_{\geq 0}$). For each $0 \leq s \leq a+1$, set

$$M_s := h_s(\lambda \eta^{a+1-s}) h_{b+1}(\mu)^{\frac{s(s-1)}{b+1}}.$$

Then

(a) *$M_{b+1} \geq M_b M_1$ provided that $b < \Delta$;*
(b) *$M_s M_{s+2} \geq M_{s+1}^2$ for all $b \leq s \leq a-1$ provided that $b < \Delta$;*
(c) *$M_{a+1} \geq M_1^{a+1}$.*

Proof. (a) Rewriting the desired inequality, we wish to prove

$$h_{b+1}(\lambda\eta^{a-b})h_{b+1}(\mu)^b \ge h_b(\lambda\eta^{a+1-b})h_{b+1}(\mu)^{\frac{b(b-1)}{b+1}}h_1(\lambda\eta^a).$$

Let $\lambda' = \lambda\eta^{a-b}$. The desired inequality can be rewritten as

$$h_{b+1}(\lambda')h_{b+1}(\mu)^{\frac{2b}{b+1}} \ge h_b(\lambda'\eta)h_1(\lambda'\eta^b). \tag{5.1}$$

Note that

$$h_1(\lambda'\eta^b) = \int_\Omega \lambda'(x)h_b(\mu H(x,\cdot))\,dx = \hom_{(\lambda',\mu,\mu,\dots,\mu)}(K_{b+1}, H). \tag{5.2}$$

Applying Lemma 5.4, we have

$$h_{b+1}(\lambda')^{\frac{1}{b+1}}h_{b+1}(\mu)^{\frac{b}{b+1}} \ge h_1(\lambda'\eta^b). \tag{5.3}$$

Thus (5.1) reduces to (after eliminating $h_{b+1}(\mu)^{\frac{2b}{b+1}}$)

$$h_{b+1}(\lambda')^{\frac{b-1}{b+1}}h_1(\lambda'\eta^b) \ge h_b(\lambda'\eta). \tag{5.4}$$

Let $N_s := h_{b+1-s}(\lambda'\eta^s)^{\frac{1}{b+1-s}}$. Lemma 5.3 implies that N_s is log-convex, i.e., $N_sN_{s+2} \ge N_{s+1}^2$ for all $0 \le s \le b-2$ (the hypothesis of Lemma 5.3 is satisfied since $b < \Delta$), and consequently, $N_0^{b-1}N_b \ge N_1^b$, which proves (5.4).

(b) The desired inequality $M_sM_{s+2} \ge M_{s+1}^2$ is equivalent, upon expanding and simplying, to

$$\sqrt{h_s(\nu_1)h_{s+2}(\nu_2)}h_{b+1}(\mu)^{\frac{1}{b+1}} \ge h_{s+1}(\sqrt{\nu_1\nu_2}). \tag{5.5}$$

where $\nu_1 = \lambda\eta^{a+1-s}$ and $\nu_2 = \lambda\eta^{a-1-s}$.

In this proof, we write, for any $\boldsymbol{x} = (x_1, \dots, x_\ell) \in \Omega^\ell$,

$$H(\boldsymbol{x}; y) = \prod_{i=1}^{\ell} H(x_i, y), \quad H(\boldsymbol{x}) = \prod_{1\le i<j\le \ell} H(x_i, x_j), \quad \nu(\boldsymbol{x}) = \prod_{i=1}^{\ell}\nu(x_i), \quad \text{and} \quad d\boldsymbol{x} = \prod_{i=1}^{\ell} dx_i.$$

For any measure μ on Ω, by expanding $h_s(\nu) = \hom_\nu(K_s, H)$ via fixing all but $b-1$ vertices of K_s,

$$h_s(\nu) = \int_{\Omega^{s-b+1}} h_{b-1}(\nu H(\boldsymbol{x};\cdot))H(\boldsymbol{x})\nu(\boldsymbol{x})d\boldsymbol{x}.$$

By the Cauchy–Schwarz inequality,

$$\sqrt{h_s(\nu_1)h_{s+2}(\nu_2)} \ge \int_{\Omega^{s-b+1}} \sqrt{h_{b-1}(\nu_1 H(\boldsymbol{x};\cdot))h_{b+1}(\nu_2 H(\boldsymbol{x};\cdot))(\nu_1\nu_2)(\boldsymbol{x})}H(\boldsymbol{x})d\boldsymbol{x}.$$

Also,

$$h_{s+1}(\sqrt{\nu_1\nu_2}) = \int_{\Omega^{s-b+1}} h_b(\sqrt{\nu_1\nu_2}H(\boldsymbol{x};\cdot))\sqrt{(\nu_1\nu_2)(\boldsymbol{x})}H(\boldsymbol{x})d\boldsymbol{x}.$$

Thus, to establish (5.5), it suffices to show that for every fixed $\boldsymbol{x} \in \Omega^{s-b+1}$,

$$\sqrt{h_{b-1}(\nu_1 H(\boldsymbol{x};\cdot))h_{b+1}(\nu_2 H(\boldsymbol{x};\cdot))}h_{b+1}(\mu)^{\frac{1}{b+1}} \ge h_b(\sqrt{\nu_1\nu_2}H(\boldsymbol{x};\cdot)),$$

as integrating it against $\sqrt{(\nu_1\nu_2)(\boldsymbol{x})}H(\boldsymbol{x})$ over $\boldsymbol{x} \in \Omega^{s-b+1}$ would yield (5.5).

Writing $\lambda' = \nu_2 H(\boldsymbol{x};\cdot)$, and recall that $\nu_2 = \nu_1\eta^2$, it remains to prove that

$$h_{b-1}(\lambda'\eta^2)h_{b+1}(\lambda')h_{b+1}(\mu)^{\frac{2}{b+1}} \ge h_b(\lambda'\eta)^2,$$

which follows from multiplying together the following three inequalities (our earlier proofs establish their validity for all measures λ' on Ω):

$$\begin{aligned} h_{b-1}(\lambda'\eta^2)^{\frac{1}{2}} h_{b+1}(\lambda')^{\frac{b-1}{2(b+1)}} &\geq h_b(\lambda'\eta)^{\frac{b-1}{b}}, \text{ and} && \text{[by Lemma 5.3]} \\ h_{b+1}(\lambda')^{\frac{1}{b(b+1)}} h_{b+1}(\mu)^{\frac{1}{b+1}} &\geq h_1(\lambda'\eta^b)^{\frac{1}{b}}, && \text{[by (5.3)]} \\ h_{b+1}(\lambda')^{\frac{b-1}{b(b+1)}} h_1(\lambda'\eta^b)^{\frac{1}{b}} &\geq h_b(\lambda'\eta)^{\frac{1}{b}}. && \text{[by (5.4)]} \end{aligned}$$

(c) We apply induction on $a - b$. When $a = b$, we have

$$M_{b+1} = h_{b+1}(\lambda)h_{b+1}(\mu)^b \quad \text{and} \quad M_1 = h_1(\lambda\eta^b).$$

We have $h_{b+1}(\lambda)h_{b+1}(\mu)^b \geq h_1(\lambda\eta^b)^{b+1}$ by Lemma 5.4 (noting (5.3)), and thus $M_{b+1} \geq M_1^b$, as claimed.

Now assume $b < a$. By the induction hypothesis, we have $M_a \geq M_1^a$. From (a) and (b) above (note that $b < \Delta$ now), we have $M_{a+1}/M_a \geq M_a/M_{a-1} \geq \cdots \geq M_{b+1}/M_b \geq M_1$, and thus $M_{a+1} \geq M_1 M_a \geq M_1^{a+1}$, as claimed. □

Now we are ready to prove Theorem 5.1, which, as a reminder, says that for $G = (V, E)$, $\boldsymbol{\lambda} = (\lambda_v : \Omega \to \mathbb{R}_{\geq 0})_{v \in V}$, and positive semidefinite $H \colon \Omega \times \Omega \to \mathbb{R}_{\geq 0}$, one has

$$\hom_{\boldsymbol{\lambda}}(G, H) \leq \prod_{v \in V} h_{d_v+1}(\lambda_v)^{1/(d_v+1)}. \tag{5.6}$$

Proof of Theorem 5.1. We apply induction first on Δ, an upper bound on the maximum degree of G, and then on the number of vertices of G. The base case, for each Δ, is when G has no vertices, in which case the statement is trivial. The only non-trivial case is if G contains a vertex w of degree exactly Δ. As earlier, let V_i, $i \in \{0, 1, \dots\} \cup \{\infty\}$, denote the set of vertices at distance exactly i from w.

First choosing the color on w, we obtain

$$\hom_{\boldsymbol{\lambda}}(G, H) = \int_\Omega \lambda_w(x_w) \hom_{\boldsymbol{\mu}^{x_w}}(G - w, H)\, dx_w,$$

where $\boldsymbol{\mu}^{x_w} = (\mu_v)_{v \in V(G-w)}$ is defined by $\mu_v = \lambda_v H(x_w, \cdot)$ for $v \in V_1$, and $\mu_v = \lambda_v$ for all $v \in V_{\geq 2}$. Now, by applying the induction hypothesis on $G - w$ to upper bound the integrand, we have,

$$\hom_{\boldsymbol{\lambda}}(G, H) \leq \int_\Omega \lambda_w(x_w) \prod_{v \in V_1} h_{d_v}(\lambda_v H(x_w, \cdot))^{1/d_v} \prod_{v \in V_{\geq 2}} h_{d_v+1}(\lambda_v)^{1/(d_v+1)}\, dx_w.$$

Comparing with the right-hand side of (5.6), we see that it suffices to prove that

$$\int_\Omega \lambda_w(x_w) \prod_{v \in V_1} h_{d_v}(\lambda_v H(x_w, \cdot))^{1/d_v}\, dx_w \leq h_{d_w+1}(\lambda_w)^{1/(d_w+1)} \prod_{v \in V_1} h_{d_v+1}(\lambda_v)^{1/(d_v+1)}.$$

Applying Hölder's inequality to the left-hand side (noting that $|V_1| = d_w$), we have

$$\begin{aligned} \int_\Omega \lambda_w(x_w) \prod_{v \in V_1} h_{d_v}(\lambda_v H(x_w, \cdot))^{1/d_v}\, dx_w &\leq \prod_{v \in V_1} \left(\int_\Omega \lambda_w(x_w) h_{d_v}(\lambda_v H(x_w, \cdot))^{d_w/d_v}\, dx_w \right)^{1/d_w} \\ &= \prod_{v \in V_1} h_1(\lambda_w \eta_v^{d_w})^{1/d_w}, \end{aligned}$$

where $\eta_v(x) = h_{d_v}(\lambda_v H(x, \cdot))^{1/d_v}$. Thus it suffices to prove that, for each $v \in V_1$,

$$h_1(\lambda_w \eta_v^{d_w}) \leq h_{d_w+1}(\lambda_w)^{1/(d_w+1)} h_{d_v+1}(\lambda_v)^{d_w/(d_v+1)}.$$

But this is exactly Lemma 5.5(c) with $a = d_w$, $b = d_v$, $\lambda = \lambda_w$, $\mu = \lambda_v$, $\eta = \eta_v$. Note that this is a valid application of Lemma 5.5(c) as $d_v \leq d_w \leq \Delta$, and Theorem 5.1 is assumed to hold for all graphs with maximum degree less than Δ by inductive hypothesis. □

Acknowledgments. We thank Suvrit Sra for suggesting the proof of Lemma 3.1. We would also like to thank the anonymous referees for detailed comments which helped to improve the paper.

References

[1] Noga Alon, *Independent sets in regular graphs and sum-free subsets of finite groups*, Israel J. Math. **73** (1991), 247–256.
[2] Ivona Bezáková, Daniel Štefankovič, Vijay V. Vazirani, and Eric Vigoda, *Accelerating simulated annealing for the permanent and combinatorial counting problems*, SIAM J. Comput. **37** (2008), 1429–1454.
[3] G. R. Blakley and Prabir Roy, *A Hölder type inequality for symmetric matrices with nonnegative entries*, Proc. Amer. Math. Soc. **16** (1965), 1244–1245.
[4] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, *Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing*, Adv. Math. **219** (2008), 1801–1851.
[5] Herm Jan Brascamp and Elliott H. Lieb, *Best constants in Young's inequality, its converse, and its generalization to more than three functions*, Advances in Math. **20** (1976), 151–173.
[6] Emma Cohen, Péter Csikvári, Will Perkins, and Prasad Tetali, *The Widom-Rowlinson model, the hard-core model and the extremality of the complete graph*, European J. Combin. **62** (2017), 70–76.
[7] Emma Cohen, Will Perkins, and Prasad Tetali, *On the Widom-Rowlinson occupancy fraction in regular graphs*, Combin. Probab. Comput. **26** (2017), 183–194.
[8] David Conlon, Jacob Fox, and Benny Sudakov, *An approximate version of Sidorenko's conjecture*, Geom. Funct. Anal. **20** (2010), 1354–1366.
[9] David Conlon, Jeong Han Kim, Choongbum Lee, and Joonkyung Lee, *Some advances on Sidorenko's conjecture*, J. Lond. Math. Soc. (2) **98** (2018), 593–608.
[10] David Conlon and Joonkyung Lee, *Sidorenko's conjecture for blow-ups*, arXiv:1809.01259.
[11] David Conlon and Joonkyung Lee, *Finite reflection groups and graph norms*, Adv. Math. **315** (2017), 130–165.
[12] Péter Csikvári, *Extremal regular graphs: the case of the infinite regular tree*, arXiv:1612.01295.
[13] Jonathan Cutler and A. J. Radcliffe, *The maximum number of complete subgraphs in a graph with given maximum degree*, J. Combin. Theory Ser. B **104** (2014), 60–71.
[14] Ewan Davies, *Counting proper colourings in 4-regular graphs via the Potts model*, Electron. J. Combin. **25** (2018), Paper 4.7, 17.
[15] Ewan Davies, Matthew Jenssen, Will Perkins, and Barnaby Roberts, *Independent sets, matchings, and occupancy fractions*, J. Lond. Math. Soc. (2) **96** (2017), 47–66.
[16] Ewan Davies, Matthew Jenssen, Will Perkins, and Barnaby Roberts, *Extremes of the internal energy of the Potts model on cubic graphs*, Random Structures Algorithms **53** (2018), 59–75.
[17] Ewan Davies, Matthew Jenssen, Will Perkins, and Barnaby Roberts, *On the average size of independent sets in triangle-free graphs*, Proc. Amer. Math. Soc. **146** (2018), 111–124.
[18] Paul Erdős and Miklós Simonovits, *Supersaturated graphs and hypergraphs*, Combinatorica **3** (1983), 181–192.
[19] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda, *Inapproximability for antiferromagnetic spin systems in the tree non-uniqueness region*, STOC'14—Proceedings of the 2014 ACM Symposium on Theory of Computing, ACM, New York, 2014, pp. 823–831.
[20] David Galvin, *Bounding the partition function of spin-systems*, Electron. J. Combin. **13** (2006), Research Paper 72, 11.
[21] David Galvin, *Maximizing H-colorings of a regular graph*, J. Graph Theory **73** (2013), 66–84.
[22] David Galvin and Prasad Tetali, *On weighted graph homomorphisms*, Graphs, morphisms and statistical physics, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 63, Amer. Math. Soc., Providence, RI, 2004, pp. 97–104.
[23] David Galvin and Yufei Zhao, *The number of independent sets in a graph with small maximum degree*, Graphs Combin. **27** (2011), 177–186.
[24] Hamed Hatami, *Graph norms and Sidorenko's conjecture*, Israel J. Math. **175** (2010), 125–150.
[25] Matthew Jenssen, Felix Joos, and Will Perkins, *On kissing numbers and spherical codes in high dimensions*, Adv. Math. **335** (2018), 307–321.
[26] Matthew Jenssen, Felix Joos, and Will Perkins, *On the hard sphere model and sphere packings in high dimensions*, Forum Math. Sigma **7** (2019), e1, 19.
[27] Jeff Kahn, *An entropy approach to the hard-core model on bipartite graphs*, Combin. Probab. Comput. **10** (2001), 219–237.

[28] Jeong Han Kim, Choongbum Lee, and Joonkyung Lee, *Two approaches to Sidorenko's conjecture*, Trans. Amer. Math. Soc. **368** (2016), 5057–5074.
[29] J.L. Xiang Li and Balázs Szegedy, *On the logarithmic calculus and Sidorenko's conjecture*, Combinatorica, to appear.
[30] Elliott H. Lieb, *Gaussian kernels have only Gaussian maximizers*, Invent. Math. **102** (1990), 179–208.
[31] László Lovász, *Large networks and graph limits*, American Mathematical Society Colloquium Publications, vol. 60, American Mathematical Society, Providence, RI, 2012.
[32] Eyal Lubetzky and Yufei Zhao, *On replica symmetry of large deviations in random graphs*, Random Structures Algorithms **47** (2015), 109–146.
[33] Guillem Perarnau and Will Perkins, *Counting independent sets in cubic graphs of given girth*, J. Combin. Theory Ser. B **133** (2018), 211–242.
[34] Ashwin Sah, Mehtaab Sawhney, David Stoner, and Yufei Zhao, *The number of independent sets in an irregular graph*, J. Combin. Theory Ser. B **138** (2019), 172–195.
[35] Luke Sernau, *Graph operations and upper bounds on graph homomorphism counts*, J. Graph Theory **87** (2018), 149–163.
[36] Alexander Sidorenko, *A correlation inequality for bipartite graphs*, Graphs Combin. **9** (1993), 201–204.
[37] Balázs Szegedy, *An information theoretic approach to Sidorenko's conjecture*, arXiv:1406.6738.
[38] Yufei Zhao, *The number of independent sets in a regular graph*, Combin. Probab. Comput. **19** (2010), 315–320.
[39] Yufei Zhao, *The bipartite swapping trick on graph homomorphisms*, SIAM J. Discrete Math. **25** (2011), 660–680.
[40] Yufei Zhao, *Extremal regular graphs: independent sets and graph homomorphisms*, Amer. Math. Monthly **124** (2017), 827–843.

Massachusetts Institute of Technology, Cambridge, MA, USA
E-mail address: asah@mit.edu

Massachusetts Institute of Technology, Cambridge, MA, USA
E-mail address: msawhney@mit.edu

Department of Matheamtics, Stanford University, Stanford, CA, USA
E-mail address: dwstoner@stanford.edu

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA
E-mail address: yufeiz@mit.edu