Counting Homomorphisms in Bipartite Graphs

Shahab Shams and Nicholas Ruozzi
University of Texas at Dallas
2601 N. Floyd Rd. MS EC31
Richardson, TX 75083-0688
Email: {shahab.shams, nicholas.ruozzi}@utdallas.edu

Péter Csikvári
ELTE Eötvös Loránd University
Department of Computer Science
1117 Budapest, Pázmány P. stny. 1/C, Hungary
Email: csiki@cs.elte.hu

***Abstract*—Graph covers and the Bethe free energy have been useful theoretical tools for producing lower bounds on a variety of counting problems in graphical models, including the permanent and the ferromagnetic Ising model. Here, we propose a new conjecture that the Bethe free energy yields a lower bound on the weighted homomorphism counting problem over bipartite graphs. We show that this conjecture strengthens existing conjectures, and we prove the conjecture in several special cases using a novel reformulation of the graph cover characterization of the Bethe free energy.**

I. Introduction

A homomorphism from a graph $G = (V_G, E_G)$ to a graph $H = (V_H, E_H)$ is a function $h : V_G \to V_H$ such that for each edge $(i, j) \in E_G$ the pair $(h(i), h(j)) \in E_H$. We will be interested in counting the number of homomorphisms from G to H, denoted $\hom(G, H)$. Let M^H be the adjacency matrix of H, we can express $\hom(G, H)$ as follows.

$$\hom(G, H) \triangleq \sum_{x \in V_H^{|V_G|}} \prod_{(i,j) \in E_G} M^H_{x_i, x_j}$$

Many interesting counting problems can be formulated as the problem of counting the number of homomorphisms from a graph G into a specific graph H [1]. For example, if H is the complete graph on n vertices, then $\hom(G, H)$ counts the number of ways in which the vertices of G can be colored with n colors such that no two adjacent vertices receive the same color. If H is allowed to have self loops, then the problem of counting the number of independent sets in G (subsets of the vertices of G such that no two adjacent vertices are in the set) can be expressed as a homomorphism counting problem by choosing H to be the complete graph on two nodes in which exactly one of the nodes has a self loop.

In this work, we examine the following conjectured lower bound on the number of homomorphisms from a bipartite graph into an arbitrary graph. This result was conjectured by Erdös, Simonovits, and Sidorenko and is commonly referred to as Sidorenko's conjecture.

Conjecture I.1 (Sidorenko's Conjecture [2], [3], and [4])**.** *For any bipartite graph $G = (V_G, E_G)$ and any arbitrary graph $H = (V_H, E_H)$,*

$$\hom(G, H) \geq |V_H|^{|V_G|} \left(\frac{2|E_H|}{|V_H|^2} \right)^{|E_G|}.$$

The special cases in which G is a path or an even cycle were initially proven by Sidorenko [3] and Simonovits [4], and subsequent work has expanded the classes of graphs, G, for which the result is known to be true for all H to include cubes [5] and bipartite graphs with one vertex complete to the other side [6], [7]. The known results make use of a variety of information theoretic tools, and many of them can be reduced to standard information theoretic inequalities and clever repeated application of Jensen's inequality [8], [9].

In this work, we conjecture a new lower bound on the weighted homomorphism counting problem and prove several special cases of this conjecture. For a bipartite graph $G = (A_G, B_G, E_G)$ and an arbitrary nonnegative matrix $M \in \mathbb{R}^{m \times m}_{\geq 0}$,

$$\hom(G, M) \triangleq \sum_{x \in \{1,\ldots,m\}^{|V_G|}} \prod_{i \in A_G} \prod_{j \in N_G(i)} M_{x_i, x_j}. \quad (1)$$

The conjectured lower bound is related to the Bethe partition function of the corresponding graphical model. The Bethe partition function is known to yield a lower bound in a variety of settings [10]–[12], and we conjecture that the same is true for the weighted homomorphism counting problem for bipartite graphs G. In this work, We provide a new formulation of a graph cover characterization of Vontobel [13], propose a strengthening of Sidorenko's conjecture using the Bethe partition function, and prove the new conjecture for trees, simple cycles, and complete bipartite graphs.

II. Variational Approximations

The weighted homomorphism counting problem (1) can be reformulated as an optimization problem. To begin, we can view $\hom(G, M)$ as the normalizing constant of a probability distribution over $[m]^{|V_G|}$, where $[m] \triangleq \{1, \ldots, m\}$.

$$p(x_1, \ldots, x_{|V_G|}) = \frac{1}{\hom(G, M)} \prod_{i \in A_G} \prod_{j \in N(i)} M_{x_i, x_j} \quad (2)$$

This observation can be used to construct lower bounds on $\hom(G, M)$ by considering the KL-divergence between some arbitrary probability distribution p' over $[m]^{|V_G|}$ and p. Recall that the KL-divergence between this pair of distributions is given by

$$d(p' || p) \triangleq \sum_{x \in [m]^{|V_G|}} p'(x) \log \frac{p'(x)}{p(x)}$$

$$= \log \hom(G, M) + \sum_{x \in [m]^{|V_G|}} p'(x) \log p'(x) - \sum_{i \in A_G} \sum_{j \in N(i)} \sum_{x \in [m]^{|V_G|}} p'(x) \log M_{x_i, x_j}.$$

As the KL divergence is always nonnegative, this yields

$$\log \hom(G, M) \geq - \sum_{x \in [m]^{|V_G|}} p'(x) \log p'(x) + \sum_{i \in A_G} \sum_{j \in N(i)} \sum_{x_i, x_j} p'_{ij}(x_i, x_j) \log M_{x_i, x_j}$$

where p_{ij} is the marginal distribution over variables x_i and x_j. Equality is achieved when $p'(x) = p(x)$ for all x. Because of this, if $\mathcal{P}$ denotes the space of all probability distributions over $[m]^{|V_G|}$, then $\hom(G, M)$ is the solution to the following optimization problem.

$$\log \hom(G, M) = \max_{p' \in \mathcal{P}} \left[- \sum_{x \in \{1, \ldots, m\}^{|V_G|}} p'(x) \log p'(x) + \sum_{i \in A_G} \sum_{j \in N(i)} \sum_{x_i, x_j} p'_{ij}(x_i, x_j) \log M_{x_i, x_j} \right]$$

A. The Bethe Partition Function

The above optimization problem is computationally intractable. As a result, approximations are often used in practice. One such approximation, known as the Bethe free energy approximation [14], simplifies the optimization problem in two ways. First, it replaces the optimization over $\mathcal{P}$ with an optimization over only local marginal distributions. Second, it approximates the entropy term with the entropy of a tree-structured distribution. Specifically, the *local marginal polytope*, $\mathcal{T}$, consists of vectors of probability distributions. There is exactly one entry in the vector τ entry for each $i \in V_G$ and each edge $(i, j) \in E_G$. The marginals in any given vector should agree on single variable overlaps. More formally,

$$\mathcal{T} \triangleq \{\tau \geq 0 \mid \forall (i, j) \in E_G, x_i \in [m], \sum_{x_j} \tau_{ij}(x_i, x_j) = \tau_i(x_i) \text{ and } \forall i \in V_G, \sum_{x_i} \tau_i(x_i) = 1\}.$$

The negative Bethe free energy approximation for the special case of (2) is given by

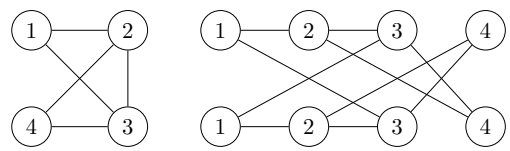
$$F_{\mathrm{B}}(G, \tau; M) \triangleq U_G(\tau; M) + \tilde{H}_G(\tau)$$

where U is the energy,

$$U_G(\tau; M) \triangleq \sum_{i \in A_G} \sum_{j \in N(i)} \sum_{x_i, x_j \in [m]} \tau_{ij}(x_i, x_j) \log M_{x_i, x_j},$$

and $\tilde{H}$ is an entropy approximation,

$$\tilde{H}_G(\tau) \triangleq - \sum_{i \in V_G} \sum_{x_i \in [m]} \tau_i(x_i) \log \tau_i(x_i) - \sum_{i \in A_G} \sum_{j \in N(i)} \sum_{x_i, x_j \in [m]} \tau_{ij}(x_i, x_j) \log \frac{\tau_{ij}(x_i, x_j)}{\tau_i(x_i) \tau_j(x_j)}.$$



(a) A graph, G. (b) One possible cover of G.

Fig. 1. An example of a graph cover. The nodes in the cover are labeled for the node that they copy in the base graph.

The Bethe partition function for the weighted homomorphism problem is then expressed in terms of the maximum value achieved by this approximation over $\mathcal{T}$.

$$\log \hom_{\mathrm{B}}(G, M) \triangleq \max_{\tau \in \mathcal{T}} F_{\mathrm{B}}(G, \tau; M) \tag{3}$$

The Bethe approximation is exact, i.e., $\hom_{\mathrm{B}}(G, M) = \hom(G, M)$, whenever G is a tree [14]. However, for general graphs, the behavior of this approximation is less well-understood.

B. Graph Covers

Quite a bit of effort has been expended in order to better understand when the Bethe partition function yields an upper or a lower bound on the corresponding counting problem. In this regard, one of the most useful theoretical tools has been a combinatorial characterization of the Bethe partition function due to Vontobel [13] that expresses the optimization problem (3) as a limit of exact counting problems on graph covers (sometimes called lifts of graphs). Roughly speaking, if a graph G' covers a graph G, then G' looks locally the same as G.

Definition II.1. *A graph G' **covers** a graph $G = (V, E)$ if there exists a graph homomorphism $h : G' \to G$ such that for all vertices $i \in G$ and all $j \in h^{-1}(i)$, h maps the neighborhood ∂j of j in G' bijectively to the neighborhood ∂i of i in G.*

If $h(j) = i$, then we say that $j \in G'$ is a copy of $i \in G$. Further, G' is said to be a k-cover of G if every vertex of G has exactly k copies in G'. We will denote the k copies of $i \in V_G$ as $i_1, \ldots, i_k \in V_{G'}$. An example 2-cover is pictured in Figure 1.

A somewhat surprising result is that the Bethe free energy optimization problem (3) has an equivalent formulation as a limit of exact counting problems on covers of G. For the special case of the weighted homomorphism counting problem, we have the following result.

Theorem II.2 (Special Case of Votobel [13])**.** *For every graph G and every matrix $M \in \mathbb{R}^{m \times m}_{\geq 0}$,*

$$\hom_{\mathrm{B}}(G, M) = \limsup_{k \to \infty} \sqrt[k]{\frac{\sum_{G' \in \mathcal{C}_k(G)} \hom(G', M)}{(k!)^{|E_G|}}}$$

where $\mathcal{C}_k(G)$ is the set of all k-covers of G. Note that $|\mathcal{C}_k(G)| = k!^{|E(G)|}$.

Theorem II.2 translates questions about $\hom_{\mathrm{B}}(G, M)$ into questions about $\hom(G', M)$ for some cover G' of G and expresses the optimization problem (3) as a limit of combinatorial objects.

III. A Stronger Conjecture

We begin by showing that the $\hom_{\mathrm{B}}$ always yields an upper bound on the posited lower bound in Sidorenko's conjecture.

Theorem III.1. *For any bipartite graph G and arbitrary graph H with adjacency matrix M^H,*

$$\hom_{\mathrm{B}}(G, M^H) \geq |V_H|^{|V_G|} \left(\frac{2|E_H|}{|V_H|^2} \right)^{|E_G|}.$$

Proof. We will assume that G is connected with no isolated vertices. If not, a similar argument can be made for each connected component. After which the results for each component can be combined using the observation that $\hom_{\mathrm{B}}(G, H) = \hom_{\mathrm{B}}(G_1, H) \cdot \hom_{\mathrm{B}}(G_2, H)$ whenever G is the disjoint union of G_1 and G_2.

$$\begin{aligned}
&\log \hom_{\mathrm{B}}(G, M^H) \\
&\quad = \sup_{\tau \in \mathcal{T}} \Bigg[\sum_{i \in A_G} \sum_{j \in N(i)} \sum_{x_i, x_j} \tau_{ij}(x_i, x_j) \log M^H_{x_i, x_j} \\
&\qquad - \sum_{i \in V_G} \sum_{x_i} \tau_i(x_i) \log \tau_i(x_i) \\
&\qquad - \sum_{i \in A_G} \sum_{j \in N(i)} \sum_{x_i, x_j} \tau_{ij}(x_i, x_j) \log \frac{\tau_{ij}(x_i, x_j)}{\tau_i(x_i) \tau_j(x_j)} \Bigg]
\end{aligned}$$

For all $(i, j) \in E_G$, let $\tau'_{i,j}(x_i, x_j) \triangleq \frac{M^H_{x_i, x_j}}{2|E_H|}$, and for all $i \in V_G$, let $\tau'_i(x_i) \triangleq \frac{\deg_H(x_i)}{2|E_H|}$ for all x_i. As τ' is in the local marginal polytope (i.e., it represents local probability distributions that satisfy the marginalization conditions), we have that

$$\begin{aligned}
&\log \hom_{\mathrm{B}}(G, M^H) \\
&\quad \geq - \sum_{i \in V_G} \sum_{x_i} \tau'_i(x_i) \log \tau'_i(x_i) \\
&\qquad - \sum_{i \in A_G} \sum_{j \in N(i)} \sum_{x_i, x_j} \tau'_{ij}(x_i, x_j) \log \frac{\tau'_{ij}(x_i, x_j)}{\tau'_i(x_i) \tau'_j(x_j)} \\
&\quad = \sum_{i \in V_G} \sum_{x_i} \left[(\deg_G(i) - 1) \frac{\deg_H(x_i)}{2|E_H|} \log \frac{\deg_H(x_i)}{2|E_H|} \right] \\
&\qquad - \sum_{i \in A_G} \sum_{j \in N(i)} \sum_{x_i, x_j} \frac{M_{x_i, x_j}}{2|E_H|} \log \frac{M_{x_i, x_j}}{2|E_H|} \\
&\quad \geq \sum_{i \in V_G} \left[(\deg_G(i) - 1) \log \frac{1}{|V_H|} \right] - |E_G| \log \frac{1}{2|E_H|} \\
&\quad = (|V_G| - 2|E_G|) \log |V_H| + |E_G| \log 2|E_H|
\end{aligned}$$

as desired. Note that the last inequality follows from the observation that the entropy is maximized by the uniform distribution. Hence, the negative entropy is minimized there. □

While there exist situations in which equality is achieved, $\hom_{\mathrm{B}}$ can be significantly larger than the lower bound given by Sidorenko's conjecture. For example, $\hom(K_{3,3}, C_4) = 164$. The lower bound given by Sidorenko's conjecture for these graphs is 8 while $\hom_{\mathrm{B}}(K_{3,3}, C_4) = 64$. In the remainder of this work, we aim to provide support for a new conjecture from which Sidorenko's conjecture would follow.

Conjecture III.2. *For any bipartite graph G and every $M \in \mathbb{R}^{m \times m}_{\geq 0}$, $\hom(G, M) \geq \hom_{\mathrm{B}}(G, M)$.*

Note that the conjecture is easily seen to be false if we remove the bipartite requirement on the graph G. That is, in general, the Bethe approximation need not yield a lower bound on the desired counting problem. Also note that, in the special case that the weight matrix M is positive semidefinite, we can extend the conjecture to general graphs by letting $M = LL^T$ and observing that $\hom(G, M) = \hom(G', L)$, where G' is the bipartite graph obtained from G by subdividing each edge of G with a single vertex.

Conjecture III.3. *For any graph G and every positive semidefinite matrix $M \in \mathbb{R}^{m \times m}_{\geq 0}$, $\hom(G, M) \geq \hom_{\mathrm{B}}(G, M)$.*

A variety of recent work has shown that the Bethe partition function yields lower bounds for a number of counting problems including matrix permanents [10], the ferromagnetic Ising model with arbitrary external field [12], the ferromagnetic Potts model with uniform external field [11], weight enumerators of linear codes [11], the weighted homomorphism counting problem when $\mathrm{rank}(M) \leq 2$ [11], etc.

IV. Special Cases of the Conjecture

Before proving several special cases of the conjecture, we show that Theorem II.2 can be converted from a counting problem over graph covers to a new counting problem over the original graph. Every k-cover, G', of a bipartite graph $G = (A_G, B_G, E_G)$ can be obtained in the following way. For each edge $i \in A_G$ and $j \in N_G(i)$, select a permutation $\sigma_{ij} \in \mathcal{S}_k$, the group of all permutations on k elements. $V_{G'}$ consists of k copies of each vertex $i \in V_G$ (recall that we will denote them as $i_1, \ldots, i_k$). For $(i, j) \in E_G$ add the edge $(i_a, j_{\sigma_{ij}(a)})$ to G' for each $a \in \{1, \ldots, k\}$. With this notation, observe that

$$\begin{aligned}
\hom(G', M) &= \sum_{x \in [m]^{|V_{G'}|}} \prod_{i \in A_{G'}} \prod_{j \in N_{G'}(i)} M_{x_i, x_j} \\
&= \sum_{x \in [m]^{|V_{G'}|}} \prod_{i \in A_G} \prod_{j \in N_G(i)} \prod_{a=1}^{k} M_{x_{i_a}, x_{j_{\sigma_{ij}(a)}}}.
\end{aligned}$$

The average number of weighted homorphisms from a k-cover of G to M is then given by the sum over all possible $\sigma_{ij} \in \mathcal{S}_k$ for each $(i, j) \in E_G$.

$$\sum_{G' \in \mathcal{C}_k(G)} \frac{\hom(G', M)}{(k!)^{|E|}}$$

$$= \sum_{x\in[m]^{|V_{G'}|}} \prod_{i\in A_G} \prod_{j\in N_G(i)} \left[\frac{1}{k!}\sum_{\sigma\in\mathcal{S}_k}\prod_{a=1}^{k} M_{x_{i_a},x_{j_{\sigma(a)}}}\right]$$

Let $\phi : [m]^k \to [m^k]$ be the bijection that sends an element of $[m]^k$ to its position in lexicographical order among all vectors in $[m]^k$. For each $I, J \in \{1,\dots,m\}^k$, we can define a matrix

$$R_k(M)_{\phi(I),\phi(J)} \triangleq \frac{1}{k!}\sum_{\sigma\in\mathcal{S}_k}\prod_{a=1}^{k} M_{I_a,J_{\sigma(a)}}.$$

This reduces the counting problem over all k-covers to a counting problem over the original graph G into $R_k(M)$.

The matrix $R_k(M)$ can be equivalently expressed as the product of a matrix depending only on M and a symmetrizing matrix that we will denote $T^{m,k}$. To see this, let $D \in \mathbb{R}^{m\times m}$ be the identity matrix and define $D^{\sigma,k} \in \mathbb{R}_{\geq 0}^{m^k\times m^k}$ such that for $I, J \in [m]^k$,

$$D^{\sigma,k}_{\phi(I),\phi(J)} \triangleq \prod_{a=1}^{k} D_{I_a,J_{\sigma(a)}}.$$

Notice that $D^{\sigma,k} \in \mathbb{R}^{m^k\times m^k}$ is a permutation matrix. If $\sigma \in \mathcal{S}_k$ is the identity permutation, then $D^{\sigma,k} = D^{\otimes k}$, the standard k-fold Kronecker product. Define $T^{m,k} \triangleq \frac{1}{k!}\sum_{\sigma\in\mathcal{S}_k} D^{\sigma,k}$. With this definition,

$$\begin{aligned} R_k(M)_{\phi(I),\phi(J)} &= \frac{1}{k!}\sum_{\sigma\in\mathcal{S}_k}\prod_{a=1}^{k} M_{I_a,J_{\sigma(a)}} \\ &= \frac{1}{k!}\sum_{\sigma\in\mathcal{S}_k}\left(M^{\otimes k}D^{\sigma,k}\right)_{\phi(I),\phi(J)} \\ &= \left(M^{\otimes k}T^{m,k}\right)_{\phi(I),\phi(J)}. \end{aligned}$$

As a consequence of the above arguments, we can reduce the counting problem over graph covers to a counting problem over the original graph with a different matrix on each edge.

$$\begin{aligned} \sum_{G'\in\mathcal{C}_k(G)} \frac{\hom(G',M)}{(k!)^{|E|}} &= \hom(G, R_k(M)) \\ &= \hom(G, M^{\otimes k}T^{m,k}) \end{aligned}$$

Theorem II.2 now implies the following lemma.

Lemma IV.1. *For every bipartite graph G and nonnegative matrix $M \in \mathbb{R}_{\geq 0}^{m\times m}$,*

$$\hom_{\mathrm{B}}(G,M) = \limsup_{k\to\infty} \sqrt[k]{\hom(G, M^{\otimes k}T^{m,k})}.$$

We note that a similar statement can be made for general so-called pairwise graphical models over arbitrary graphs, though we omit the general statement for space.

The "symmetrizing" matrix $T^{m,k}$ arises in a variety of mathematical applications and has a number of interesting properties that follow from simple algebraic manipulations.

Lemma IV.2. *For each $m, k \in \mathbb{Z}_{>0}$,*

- $T^{m,k}\cdot T^{m,k} = T^{m,k}$
- *For every $M \in \mathbb{R}^{m\times m}$, $T^{m,k}\cdot M^{\otimes k} = M^{\otimes k}\cdot T^{m,k}$*
- *$T^{m,k}$ is symmetric and doubly stochastic*
- *For any k vectors $v_1,\dots,v_k \in \mathbb{R}^m$,*

$$T^{m,k}\cdot(v_1\otimes\cdots\otimes v_k) = \frac{1}{k!}\sum_{\sigma\in\mathcal{S}_k} v_{\sigma(1)}\otimes\cdots\otimes v_{\sigma(k)}.$$

One approach to using Theorem II.2 to show that the Bethe approximation yields a lower bound on $\hom(G,\cdot)$ would be to show that $\hom(G,M) \geq \sqrt[k]{\hom(G',M)}$ for any k-cover G' of G. While this was the approach taken in previous work, e.g., [12], here we will use Lemma IV.1 to argue only that the k-covers yield a lower bound on average. In the following sections, we will prove a few special cases of the conjecture using the characterization of Lemma IV.1.

A. Trees and Cycles

As discussed above, the Bethe partition function is exact whenever G is a tree. This is, essentially, by construction, but it also follows from the observation that any k-cover of a tree, G, is isomorphic to the disjoint union of k copies of G, hence $\hom(G',M) = \hom(G,M)^k$ for any k-cover G' of G. As a result, we have that $\hom_{\mathrm{B}}(G,M) = \hom(G,M)$ for any tree G. The result for even cycles is also straightforward.

Theorem IV.3. *For every $n > 1$ and every $M \in \mathbb{R}_{\geq 0}^{m\times m}$, $\hom(C_{2n},M) \geq \hom_{\mathrm{B}}(C_{2n},M)$, where C_{2n} is the simple cycle on $2n$ nodes.*

Proof. Observe that, for a single cycle, $\hom(C_{2n},M) = \mathrm{tr}((MM^T)^n)$. From Lemma IV.1, it suffices to show that $\hom(C_{2n}, M^{\otimes k}T^{m,k}) \leq \hom(C_{2n},M)^k$ for each $k > 1$. Consider,

$$\begin{aligned} \hom(C_{2n}, M^{\otimes k}T^{m,k}) &\overset{(a)}{=} \mathrm{tr}\left(\left(M^{\otimes k}M^{\otimes k^T}\right)^n T^{m,k}\right) \\ &\overset{(b)}{\leq} \mathrm{tr}\left(\left(\left(MM^T\right)^{\otimes k}\right)^n\right) \\ &\overset{(c)}{=} \mathrm{tr}\left(\left(MM^T\right)^n\right)^k \\ &= \hom(C_{2n},M)^k, \end{aligned}$$

where (a) follows from Lemma IV.2, (b) follows from a standard majorization argument on the eigenvalues of positive semidefinite matrices, see H.1.g. Marshall and Olkin [15], and (c) is a consequence of the observation that $\hom(G,M^{\otimes k}) = \hom(G,M)^k$ for all graphs G. □

B. Complete Bipartite Graphs

In this section, we show that Conjecture III.2 is true for the special case in which G is a complete bipartite graph. The result makes use of majorization inequalities. A vector $v \in \mathbb{R}^n$ is majorized by a vector $w \in \mathbb{R}^n$, written $v \prec w$ if

$$\sum_{i=1}^{k} v_{[i]} \leq \sum_{i=1}^{k} w_{[i]}, \text{for all } k \in \{1,\dots,n\} \tag{4}$$

and

$$\sum_{i=1}^{n} v_{[i]} = \sum_{i=1}^{n} w_{[i]},$$

where $w_{[i]}$ denotes the ith largest entry of w. Equivalently, $v \prec w$ if and only if there exists a doubly stochastic matrix S such that $v = Sw$. The vector v is said to be weakly majorized by w, denoted $v \prec_w w$, if only condition (4) holds. Finally, if $f : \mathbb{R} \to \mathbb{R}$ is a convex function, then $v \prec w$ implies that $f(v) \prec_w f(w)$, where $f(v)$ denotes the vector obtained by applying f to each component of v [15].

Theorem IV.4. *For every complete bipartite graph $K_{a,b}$ and every $M \in \mathbb{R}^{m\times m}_{\geq 0}$, $\hom(K_{a,b}, M) \geq \hom_{\mathrm{B}}(K_{a,b}, M)$.*

Proof. The proof considers a reformulation of the counting problem in which the variables in one of the partitions of $K_{a,b} = (A, B, E)$, say A, have been summed out. For $N \in \mathbb{R}^{m\times m}_{\geq 0}$,

$$\begin{aligned}\hom(K_{a,b}, N) &= \sum_{x_B\in[m]^b} \sum_{x_A\in[m]^a} \prod_{i\in A}\prod_{j\in B} N_{x_i,x_j} \\ &= \sum_{x_B\in[m]^b} \prod_{i\in A}\left(\sum_{x_i\in[m]}\prod_{j\in B} N_{x_i,x_j}\right) \\ &= \sum_{x_B\in[m]^b}\left(\sum_{y\in[m]}\prod_{j\in B} N_{y,x_j}\right)^a.\end{aligned}$$

We can think of the product $\prod_{j\in B} N_{y,x_j}$ as a vector indexed by assignments to the variables x_B, i.e.,

$$\prod_{j\in B} N_{y,x_j} = \left[\left(N_{y,:}\right)^{\otimes b}\right]_{x_B},$$

where $N_{y,:}$ denotes the yth row of N. Now, consider $\hom(K_{a,b}, M^{\otimes k}T^{m,k})$ for some $k > 1$. Following the above argument and substituting $M^{\otimes k}T^{m,k}$ for N yields

$$\begin{aligned}\sum_{y\in[m^k]}\prod_{j\in B}(M^{\otimes k}T^{m,k})_{y,x_j} &= \sum_{y\in[m^k]}\left[\left(M^{\otimes k}_{y,:}T^{m,k}\right)^{\otimes b}\right]_{x_B} \\ &= \sum_{y\in[m^k]}\left[\left(M^{\otimes k}_{y,:}\right)^{\otimes b}\left(T^{m,k}\right)^{\otimes b}\right]_{x_B} \\ &= \left[\sum_{y\in[m^k]}\left(M^{\otimes k}_{y,:}\right)^{\otimes b}\left(T^{m,k}\right)^{\otimes b}\right]_{x_B}\end{aligned}$$

for each $x_B \in [m^k]^b$. Since $T^{m,k}$ is doubly stochastic, so is $\left(T^{m,k}\right)^{\otimes b}$, and we have

$$\sum_{y\in[m^k]}\left(M^{\otimes k}_{y,:}\right)^{\otimes b}\left(T^{m,k}\right)^{\otimes b} \prec \sum_{y\in[m^k]}\left(M^{\otimes k}_{y,:}\right)^{\otimes b}.$$

Finally, as $f(x) = x^a$ is a convex function for all $a > 1$, $x \geq 0$, raising each component of the vectors to the ath power preserves weak majorization.

$$\begin{aligned}\hom(K_{a,b}, M^{\otimes k}T^{m,k}) &= \sum_{x_B\in[m^k]^b}\left[\sum_{y\in[m^k]}\left(M^{\otimes k}_{y,:}\right)^{\otimes b}\left(T^{m,k}\right)^{\otimes b}\right]^a_{x_B} \\ &\leq \sum_{x_B\in[m^k]^b}\left[\sum_{y\in[m^k]}\left(M^{\otimes k}_{y,:}\right)^{\otimes b}\right]^a_{x_B} \\ &= \hom(K_{a,b}, M^{\otimes k}),\end{aligned}$$

which, in conjunction with Lemma IV.1, yields the desired result. □

Note that a stronger version of Theorem IV.4 form the case in which M is an adjacency matrix was proven by Galvin and Tetali [16]. In particular, their results imply that $\hom(G', M^H) \leq \hom(K_{a,b}, M^H)^k$ for any k-cover G' of $K_{a,b}$ and any graph H with adjacency matrix M^H. The above proof argues the result for the k-covers on average but is somewhat simpler than that of [16]. Also note that the proof of Theorems IV.3 and IV.4 could be generalized to show that the results hold under multiplication by any doubly stochastic matrix, e.g., $\hom(K_{a,b}, M) \geq \hom(K_{a,b}, MS)$ for any doubly stochastic matrix S. While this is true for $K_{a,b}$ and C_{2n}, it is not true for any arbitrary bipartite graph and an arbitrary doubly stochastic matrix.

Acknowledgments

This work was supported by the DARPA Explainable Artificial Intelligence (XAI) program under contract number N66001-17-2-4032 and NSF grant III-1527312.

References

[1] L. Lovász, *Large Networks and Graph Limits.*, ser. Colloquium Publications. American Mathematical Society, 2012, vol. 60.
[2] A. Sidorenko, "Inequalities for functionals generated by bipartite graphs," *(Russian) Diskret. Mat.*, vol. 3, p. 50–65, 1991.
[3] ——, "A correlation inequality for bipartite graphs," *Graphs and Combin.*, vol. 9, p. 201–204, 1993.
[4] M. Simonovits, "Extremal graph problems, degenerate extremal problems, and supersaturated graphs," in *Progress in Graph Theory*. NY Academy Press, 1984, p. 419–437.
[5] H. Hatami, "Graph norms and sidorenko's conjecture," *Israel Journal of Mathematics*, vol. 175, no. 1, pp. 125–150, 2010.
[6] D. Conlon, J. Fox, and B. Sudakov, "An approximate version of sidorenko's conjecture," *Geometric and Functional Analysis*, vol. 20, no. 6, pp. 1354–1366, 2010.
[7] J. Li and B. Szegedy, "On the logarithimic calculus and sidorenko's conjecture," *arXiv preprint arXiv:1107.1153*, 2011.
[8] J. H. Kim, C. Lee, and J. Lee, "Two Approaches to Sidorenko's Conjecture," *Trans. Amer. Math. Soc.*, vol. 368, no. 7, pp. 5057–5074, 2016.
[9] B. Szegedy, "An information theoretic approach to Sidorenko's conjecture," *arXiv preprint arXiv:1406.6738*, 2014.
[10] L. Gurvits, "Unleashing the power of Schrijver's permanental inequality with the help of the Bethe approximation," *ArXiv e-prints*, Jun. 2011.
[11] N. Ruozzi, "Beyond log-supermodularity: Lower bounds and the bethe partition function," in *Proceedings of the Twenty-Ninth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-13)*. Corvallis, Oregon: AUAI Press, 2013, pp. 546–555.
[12] ——, "The Bethe partition function of log-supermodular graphical models," in *Neural Information Processing Systems (NIPS)*, Lake Tahoe, NV, Dec. 2012.
[13] P. O. Vontobel, "Counting in graph covers: A combinatorial characterization of the Bethe entropy function," *Information Theory, IEEE Transactions on*, Jan. 2013.
[14] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *Information Theory, IEEE Transactions on*, vol. 51, no. 7, pp. 2282 – 2312, July 2005.
[15] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*. New York: Academic Press, 1979.
[16] D. Galvin and P. Tetali, "On weighted graph homomorphisms," *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, vol. 63, pp. 97–104, 2004.