Counting matchings in irregular bipartite graphs and random lifts

M. Lelarge*

Abstract

We give a sharp lower bound on the number of matchings of a given size in a bipartite graph. When specialized to regular bipartite graphs, our results imply Friedland's Lower Matching Conjecture and Schrijver's theorem proven by Gurvits and Csikvári. Indeed, our work extends the recent work of Csikvári done for regular and bi-regular bipartite graphs. Moreover, our lower bounds are order optimal as they are attained for a sequence of 2-lifts of the original graph as well as for random n-lifts of the original graph when n tends to infinity.

We then extend our results to permanents and subpermanents sums. For permanents, we are able to recover the lower bound of Schrijver recently proved by Gurvits using stable polynomials. Our proof is algorithmic and borrows ideas from the theory of local weak convergence of graphs, statistical physics and covers of graphs. We provide new lower bounds for subpermanents sums and obtain new results on the number of matching in random n-lifts with some implications for the matching measure and the spectral measure of random n-lifts as well as for the spectral measure of infinite trees.

1 Introduction

Recall that a $n \times n$ matrix A is called doubly stochastic if it is nonnegative entrywise and each of its columns and rows sums to one. Also the permanent of a $n \times n$ matrix A is defined as

$$\mathrm{per}(A) = \sum_{\sigma \in \mathcal{S}_n} \prod_{i=1}^{n} a_{i,\sigma(i)},$$

where the summation extends over all permutation σ of $\{1, \dots, n\}$. The main result proved in [37] is the following theorem:

Theorem 1. *(Schrijver [37]) For any doubly stochastic $n \times n$ matrix $A = (a_{i,j})$, we define $\tilde{A} = (\tilde{a}_{i,j} = a_{i,j}(1 - a_{i,j}))$ and we have*

$$\mathrm{per}(\tilde{A}) \geq \prod_{i,j} (1 - a_{i,j}). \tag{1}$$

It is proved in [21, 22] that this theorem implies:

Theorem 2. *Let A be a non-negative $n \times n$ matrix. Then, we have*

$$\ln \mathrm{per}(A) \geq \max_{x \in M_{n,n}} \sum_{i,j} (1 - x_{i,j}) \ln(1 - x_{i,j}) + x_{i,j} \ln\left(\frac{a_{i,j}}{x_{i,j}}\right), \tag{2}$$

lithe feron

*INRIA-ENS, Paris, France, email: marc.lelarge@ens.fr

Real Stability?

with the convention $\ln \frac{0}{0} = 1$ *and where* $M_{n,n}$ *is the set of* $n \times n$ *doubly stochastic matrices.*

Clearly applying Theorem 2 to $\tilde{A}$ with $x_{i,j} = a_{i,j}$, we get Theorem 1 back, so that both theorems are equivalent. In [19, 20], L. Gurvits provided a new proof of these theorems using stable polynomials, see also [25]. Our main new result is a generalization of Theorem 2 to subpermanent sums, see Theorem 4 below. Our proof is very different from those derived in [37, 19, 20] and borrows ideas from the recent work of Csikvári [14] for matchings in regular bipartite graphs. Permanents and subpermanent sums can be interpreted as weighted sums of matchings in complete bipartite graphs and the extension of the approach in [14] to this general framework is the main technical contribution of our work. Interestingly, the obstacles to overcome are computational in nature and we present a very algorithmic solution. As a byproduct, we also obtain new results on the number of matchings in random lifts, see Theorem 5 below. We compute the limit of the matching generating function (i.e. the partition function of the monomer-dimer model) for a sequence of random n-lifts of a graph G as an explicit function of the original graph G. This result has also some implications for the spectral measure and the matching measure of random lifts, see Theorems 7 and 8.

As a consequence of Theorem 1, Schrijver shows in [37] that any d-regular bipartite graph with $2n$ vertices has at least

row, column sums to d

the same term appears in [Vontobel, 2012 pg 21], but the graph is not d-regular. The term minimizes the BFE of the said graph

$$\left(\frac{(d-1)^{d-1}}{d^{d-2}} \right)^n \quad (3)$$

Pm

perfect matchings (a perfect matching is a set of disjoint edges covering all vertices). For each d, the base $(d-1)^{d-1}/d^{d-2}$ in (3) is best possible [41]. Similarly, our Theorem 4 shows that the right-hand term in (2) is best possible in the following sense: for any $n \times n$ matrix A, we show that there exists a sequence of growing matrices A_ℓ obtained by taking successive 2-lifts of A (see Definition 2) such that its permanent grows exponentially with its size at a rate given by the right-hand term in (2). We refer to Theorem 4 for a precise statement and similar results for subpermanent sums.

In [16], Friedland, Krop and Markström conjectured a possible generalization of (3) which is known as Friedland's lower matching conjecture: for G a d-regular bipartite graph with $2n$ vertices, let $m_k(G)$ denote the number of matchings of size k (see Section 2.1 for a precise definition), then

$$m_k(G) \geq \binom{n}{k}^2 \left(\frac{d-p}{d} \right)^{n(d-p)} (dp)^{np}, \quad (4)$$

K-matchings

where $p = \frac{k}{n}$. An asymptotic version of this conjecture was proved using Theorem 1 in [21, 22]. A slightly stronger statement of the conjecture was proved by Csikvári in [14] and we extend it to cover irregular bipartite graphs, see Theorem 3.

We state our main results in the next section. In Section 3, we summarize the main ideas of the proof and describe related works. We also give some implications of our work for extremal graph theory and for the spectral measure and matching measure of random lifts. Section 4 contains the technical proof. We first summarize the statistical physics results for the monomer dimer model in Section 4.1. Then, we study local recursions associated to this model in Section 4.2 The results in this section build mainly on previous work of the author [28]. In Section 4.3, we show how an idea of Csikvári [14] using 2-lift extends to our framework and connect it to

$G=(V,E)$: Connected multigraph
$B=\{B_e \mid e\in E\}\in\{0,1\}^{|E|}$: incidence vector encoding matching
$\nu(G)=\max_B\{\sum_e B_e\}$: matching number of G
$M(G)$: matching polytope of G (Convex hull of all incidence vectors of matching)
$m_k(G)$: no of matchings of size k in G
$P_G(z)=\sum_{k=0}^{\nu(G)} m_k(G)z^k$: matching generating f^n
$M_k(G)$: Convex hull of all matchings of size

$FM(G)=\{x\in\mathbb{R}^{|E|}, x_e\geq 0, \sum_{e\in\partial v} x_e\leq 1\ \forall v\in V\}$: Fractional matching polytope
$\nu^*(G)=\max_{x\in FM(G)}\{\sum_e x_e\}\geq\nu(G)$: Fractional matching number
$FM_t(G)=\{x\in\mathbb{R}^E, \sum_{e\in\partial v} x_e\leq 1, \sum_{e\in E} x_e=t\}$

G is bipartite $\iff M(G)=FM(G)$ and $M_k(G)=FM_k(G)$

the framework of local weak convergence in Section 4.4. We use probabilistic bounds on the coefficients of polynomials with only real zeros to finish the proof in Section 4.5. In Section 4.6, we provide the details needed for random lifts.

2 Main results

We present our main results in this section. The results concerning lower bounds for the number of matchings given in Section 2.1 are implied by those in Section 2.2 concerning lower bounds for permanents. The reader only interested in the most general result concerning lower bounds might jump directly to Section 2.2 where Theorem 4 is stated in a self-contained manner. Note that the proof will use notations introduced in Section 2.1. Section 2.3 contains our results on random lifts. They are independent from the other results but require some notations from Section 2.1.

2.1 Lower bounds for number of matchings of a given size

We consider a connected multigraph $G=(V,E)$. We denote by $v(G)$ the cardinality of V: $v(G)=|V|$. We denote by the same symbol ∂v the set of neighbors of node $v\in V$ and the set of edges incident to v. A matching is encoded by a binary vector, called its incidence vector, $\mathbf{B}=(B_e,\ e\in E)\in\{0,1\}^E$ defined by $B_e=1$ if and only if the edge e belongs to the matching. We have for all $v\in V$, $\sum_{e\in\partial v}B_e\leq 1$. The size of the matching is given by $\sum_e B_e$. We will also use the following notation $e\in\mathbf{B}$ to mean that $B_e=1$, i.e. that the edge e is in the matching. For a finite graph G, we define the matching number of G as $\nu(G)=\max\{\sum_e B_e\}$ where the maximum is taken over matchings of G.

The matching polytope $M(G)$ of a graph G is defined as the convex hull of incidence vectors of matchings in G. We define the fractional matching polytope as

Belief Polytope

$$FM(G)=\left\{\mathbf{x}\in\mathbb{R}^E,\ x_e\geq 0,\ \sum_{e\in\partial v}x_e\leq 1\right\}. \tag{5}$$

We also define the fractional matching number $\nu^*(G)=\max_{\mathbf{x}\in FM(G)}\sum_e x_e\geq\nu(G)$. It is well-known that: $M(G)=FM(G)$ if and only if G is bipartite and in this case, we have $\nu(G)=\nu^*(G)$.

For a given graph G, we denote by $m_k(G)$ the number of matchings of size k in G ($m_0(G)=1$). For a parameter $z>0$, we define the matching generating function:

$$P_G(z)=\sum_{k=0}^{\nu(G)}m_k(G)z^k.$$

In statistical physics, the function $\ln P_G(z)$ is called the partition function. We define by $M_k(G)$ the convex hull of incidence vectors of matchings in G of size k and similarly for $0\leq t\leq\nu^*(G)$:

$$FM_t(G)=\left\{\mathbf{x}\in\mathbb{R}^E,\ x_e\geq 0,\ \sum_{e\in\partial v}x_e\leq 1,\ \sum_{e\in E}x_e=t\right\}. \tag{6}$$

Convex?

If G is bipartite, we have $M_k(G) = FM_k(G)$. Theorem 3 below deals with bipartite graphs but Theorem 5 deals with general graphs and fractional matching polytopes will be needed there.

For any finite graph G, we define the function $S_G^B : FM(G) \to \mathbb{R}$ by: Bethe Entropy

$$S_G^B(\mathbf{x}) \;=\; \sum_{e\in E} -x_e \ln x_e + (1-x_e)\ln(1-x_e) - \sum_{v\in V}\left(1 - \sum_{e\in\partial v} x_e\right)\ln\left(1-\sum_{e\in\partial v} x_e\right). \quad (7)$$

The function S_G^B is concave on $FM(G)$ by Proposition 13.

Why no loops?

Definition 1. *Let G be a graph with no loops. Then H is a 2-lift of G if $V(H) = V(G)\times\{0,1\}$ and for every $(u,v)\in E(G)$, exactly one of the following two pairs are edges of H: $((u,0),(v,0))$ and $((u,1),(v,1))\in E(H)$ or $((u,0),(v,1))$ and $((u,1),(v,0))\in E(H)$. If $(u,v)\notin E(G)$, then none of $((u,0),(v,0))$,$((u,1),(v,1))$, $((u,0),(v,1))$ and $((u,1),(v,0))$ are edges in H.*

G is Bipartite : M(G) = FM(G)

Note that a bipartite graph G has no loops.

Theorem 3. *For any finite bipartite graph G, we have for $z>0$,*

not necessarily $K_{n,n}$

$$\ln P_G(z) \ge \max_{\mathbf{x}\in M(G)}\left\{\left(\sum_e x_e\right)\ln z + S_G^B(\mathbf{x})\right\}. \quad (8)$$

equality holds for bipartite?

We have for all $k\le \nu(G)$,

$$m_k(G) \ge b_{\nu(G),k}(k/\nu(G))\exp\left(\max_{\mathbf{x}\in M_k(G)} S_G^B(\mathbf{x})\right),$$

where $b_{n,k}(p)$ is the probability for a binomial random variable $Bin(n,p)$ to take the value k, i.e. $b_{n,k}(p) = \binom{n}{k}p^k(1-p)^{n-k}$. Moreover, there exists a sequence of bipartite graphs $\{G_n = (V_n,E_n)\}_{n\in\mathbb{N}}$ such that $G_0 = G$, G_n is a 2-lift of G_{n-1} for $n\ge 1$ and for all $z>0$,

$$\lim_{n\to\infty}\frac{1}{v(G_n)}\ln P_{G_n}(z) = \frac{1}{v(G)}\max_{\mathbf{x}\in M(G)}\left\{\left(\sum_e x_e\right)\ln z + S_G^B(\mathbf{x})\right\}.$$

Analogous to Thm 39 Pg 14 Vontobel, 2012

Consider the particular case where G is a d-regular bipartite graph on $2n$ vertices. In this case, we have $\nu(G) = n$ and we can take $x_e^* = \frac{k}{nd}$ for all $e\in E$ so that $\mathbf{x}^*\in M_k(G)$ and we have

$$S_G^B(\mathbf{x}^*) = n\left(p\ln\left(\frac{d}{p}\right) + (d-p)\ln\left(1-\frac{p}{d}\right) - 2(1-p)\ln(1-p)\right),$$

for a matching of size k

with $p = \frac{k}{n}$. We see that we recover the first statement in Theorem 1.5 of [14]. In particular, for $k=n$, i.e. $p=1$, we recover (3) and for $k<n$, as explained in [14], we slightly improve upon (4). Note that in this particular case, we have $m_n(G)\le (d!)^{n/d}$ by a result of Bregman [12] (see also [15] for upper bounds for $m_k(G)$ with $k\le n$).

Taking $z=1$ in (8), we obtain the following bound on the total number of matchings:

Proposition 1. *For any bipartite graph G, we have:*

$$\underbrace{\sum_{k=0}^{\nu(G)} m_k(G)}_{\text{Total no of matchings}} \ge \underbrace{\exp\left(\max_{\mathbf{x}\in M(G)} S_G^B(\mathbf{x})\right)}_{\text{Bethe Perm}}.$$

(sum of all sub-perm sums)

M_n : $n \times n$ matrices
$M_{n,n}$: bi-Stoch. matrices
$M_{n,k}$: sub-stoch $n \times n$ matrices with entrywise L_1-norm k
$M_{n,\leq}$: " " "

$\mathrm{per}_k(A)$: Sum of per of all $k \times k$ minors of A

2.2 Lower bounds for permanents

In this section, we extend previous results to weighted graphs. We state our results in term of permanents. Let A be a non-negative $n \times n$ matrix. We denote by M_n the set of such matrices. Recall that the permanent of $A \in M_n$ is defined by

$$\mathrm{per}(A) = \sum_{\sigma \in \mathcal{S}_n} \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

We define by $M_{n,n}$ the set of $n \times n$ doubly stochastic matrices:

$$M_{n,n} = \left\{ A,\ 0 \leq a_{i,j},\ \sum_i a_{i,j} = \sum_j a_{i,j} = 1 \right\} \subset M_n.$$

For $1 \leq k \leq n$, let $\mathrm{per}_k(A)$ be the sum of permanents of all $k \times k$ minors in A and $\mathrm{per}_0(A) = 1$. $\mathrm{per}_k(A)$ is called the k-th subpermanent sum of A. We define $M_{n,k}$ the set of $n \times n$ non-negative sub-stochastic matrices with entrywise L_1-norm k:

$$M_{n,k} = \left\{ A,\ 0 \leq a_{i,j},\ \sum_i a_{i,j} \leq 1,\ \sum_j a_{i,j} \leq 1,\ \sum_{i,j} a_{i,j} = k \right\} \subset M_n.$$

We also define the set of substochastic matrices:

$$M_{n,\leq} = \left\{ A,\ 0 \leq a_{i,j},\ \sum_i a_{i,j} \leq 1,\ \sum_j a_{i,j} \leq 1 \right\} \subset M_n.$$

We define the function $S^B : M_n \times M_{n,\leq} \to \mathbb{R} \cup \{-\infty\}$ by

why sub-stoch : Generalization of $FM(G)$

BFE

$$S^B(A, \mathbf{x}) = \sum_{i,j} x_{i,j} \ln \frac{a_{i,j}}{x_{i,j}} + (1 - x_{i,j}) \ln(1 - x_{i,j}) \tag{9}$$

$$- \sum_i \left(1 - \sum_j x_{i,j}\right) \ln \left(1 - \sum_j x_{i,j}\right) - \sum_j \left(1 - \sum_i x_{i,j}\right) \ln \left(1 - \sum_i x_{i,j}\right) \tag{10}$$

with the convention $\ln \frac{0}{0} = 1$. First note that if A is the incidence matrix of a bipartite graph G, and $\mathbf{x}$ is such that there exists $a_{i,j} = 0$ and $x_{i,j} > 0$, then $S^B(A, \mathbf{x}) = -\infty$. Moreover if $\mathbf{x}$ has only non-negative components corresponding to edges of the graph G, then we have $S^B(A, \mathbf{x}) = S^B_G(\mathbf{x})$ as defined in (7) with a slight abuse of notation: the zero components (on no-edges of G) of $\mathbf{x}$ as argument of $S^B(A, \mathbf{x})$ are removed in the argument of $S^B_G(\mathbf{x})$. Note that $\mathbf{x} \mapsto S^B(A, \mathbf{x})$ is concave on $M_{n,\leq}$ (since S^B_G is concave on $FM(G)$ by Proposition 13).

Definition 2. *Let A be a non-negative $n \times n$ matrix. Then B is a 2-lift of A if B is a $2n \times 2n$ non-negative matrix such that for all $i, j \in \{1, \ldots n\}$, either $b_{i,j} = b_{i+n,j+n} = a_{i,j}$ and $b_{i,j+n} = b_{i+n,j} = 0$ or $b_{i,j+n} = b_{i+n,j} = a_{i,j}$ and $b_{i,j} = b_{i+n,j+n} = 0$.*

A is adj matrix of the graph (not the biadj matrix)
bipartite

Theorem 4. *Let A be a non-negative $n \times n$ matrix. Let $\nu(A) = \max\{k,\ \mathrm{per}_k(A) > 0\}$. For all $k \le \nu(A)$, we have*

$$\mathrm{per}_k(A) \ge b_{\nu(A),k}(k/\nu(A)) \exp\left(\max_{\mathbf{x} \in M_{n,k}} S^B(A, \mathbf{x}) \right), \tag{11}$$

control the size of matching

where $b_{n,k}(p) = \binom{n}{k} p^k (1-p)^{n-k}$. Moreover, there exists a sequence of matrices $\{A_\ell \in M_{2^\ell n}\}_{\ell \in \mathbb{N}}$ such that $A_0 = A$, A_ℓ is a 2-lift of $A_{\ell-1}$ for $\ell \ge 1$ and for all $z > 0$,

$$\lim_{\ell \to \infty} \frac{1}{2^\ell} \ln \left(\sum_{k=0}^{\nu(A_\ell)} \mathrm{per}_k(A_\ell) z^k \right) = \max_{\mathbf{x} \in M_{n,\le}} \left\{ \left(\sum_{i,j} x_{i,j} \right) \ln z + S^B(A, \mathbf{x}) \right\}.$$

Analogous to Thm 39 Pg 14 Vontobel, 2012

If $k = n$ in (11), we recover Theorem 2 which is equivalent to Theorem 1. Note that if $\nu(A) < n$, then $\mathrm{per}(A) = 0$ and the lower bound in Theorem 2 is equal to $-\infty$. Indeed if $\mathrm{per}(A) = 0$, then if $x \in M_{n,n}$ is a permutation matrix then there exists i, j such that $a_{i,j} = 0$ and $x_{i,j} > 0$ so that $\ln\left(\frac{a_{i,j}}{x_{i,j}}\right) = -\infty$. The claim then follows from the Birkhoff-von Neumann Theorem which implies that any doubly stochastic matrix can be written as a convex combination of permutation matrices. Also, results presented in Section 2.1 follow by taking for the matrix A, the incidence matrix of the graph G.

2.3 Number of matchings in random lifts

As in previous section, $G = (V, E)$ is a fixed connected multigraph with no loops. A random n-lift of G is a random graph on vertex set $V_1 \cup V_2 \cup \cdots \cup V_{v(G)}$, where each V_i is a set of n vertices and these sets are pairwise disjoint, obtained by placing a uniformly chosen random perfect matching between V_i and V_j, independently for each edge $e = ij$ of G. We denote the resulting graph $L_n(G)$. → Random n-lift of G

Our main result shows that the lower bounds derived in Section 2.1 are indeed attained by a sequence of random lifts. More precisely, we have

Compare to Thm 3:
- Bounds are in terms of fractional matchings
- G is not bipartite
- Random n-lift instead of sequence of 2-lifts.

Theorem 5. *For any finite graph G, we have*

$$\lim_{n \to \infty} \frac{1}{n} \nu(L_n(G)) = \nu^*(G) \quad a.s.$$

$$\forall z > 0, \quad \lim_{n \to \infty} \frac{1}{n} \ln P_{L_n(G)}(z) = \max_{\mathbf{x} \in FM(G)} \left\{ \left(\sum_e x_e \right) \ln z + S_G^B(\mathbf{x}) \right\} \quad a.s.$$

where $S_G^B(\mathbf{x})$ is defined in (7). We denote by $FPM(G) = FM_{\nu^*(G)}(G) = \{\mathbf{x},\ \sum_{e \in \partial v} x_e = 1\}$ the fractional perfect matching polytope of G, then we have

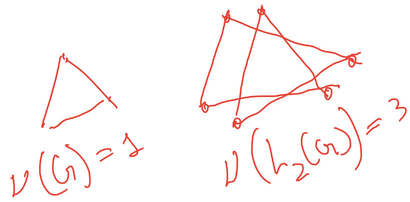
$$\limsup_{n \to \infty} \frac{1}{n} \ln m_{\nu(L_n(G))}(L_n(G)) \le \sup_{\mathbf{x} \in FPM(G)} S_G^B(\mathbf{x}).$$

upper bound for max mat in random n-lift as n→∞

If, in addition, G is bipartite, then the fractional perfect matching polytope is simply the perfect matching polytope $PM(G)$ of G and

$$\lim_{n \to \infty} \frac{1}{n} \ln m_{\nu(L_n(G))}(L_n(G)) = \sup_{\mathbf{x} \in PM(G)} S_G^B(\mathbf{x}).$$

$\vec{x}$ is stochastic in this case



In [31], Linial and Rozenman studied the existence of a perfect matching in $L_n(G)$. Note that if the number of vertices in G is odd, then in order to have a perfect matching in a n-lifts of G, we need to have n even. For n even, they described a large class of graphs G for which $L_n(G)$ contains a perfect matching asymptotically almost surely. This class contains all regular graphs and, in turn, is contained in the class of graphs having a fractional perfect matching, i.e. graphs G such that $2\nu^*(G) = v(G)$. Our result shows that in this case, $L_n(G)$ will contain an almost perfect matching (possibly missing $o(n)$ vertices) almost surely. If in addition the graph G is bipartite, the number of such matchings is exponential in n.

In [18], the number of perfect matchings in $L_n(G)$, denoted by $\mathrm{pm}(L_n(G))$, is studied in the limit $n \to \infty$ (along a subsequence ensuring the existence of a perfect mating), where G is a graph with a fractional perfect matching. Using the small subgraph conditioning method, an asymptotic formula for $\mathbb{E}[\mathrm{pm}(L_n(G))]$ is computed for any connected regular multigraph G with degree at least three. Partial results are also given for $\mathbb{E}[\mathrm{pm}(L_n(G))^2]$ with an explicit formula based on a conjecture (proved only for 3-regular graphs).

Note that we always have $\nu(L_n(G)) \geq n\nu(G)$ (by lifting a maximum matching of G). In particular, if G has a perfect matching then $\nu(L_n(G)) = nv(G)/2$. Hence, Theorem 5 implies that for any d-regular bipartite graph G, we have

vertices

$$\lim_{n\to\infty} \frac{1}{n} \ln \mathrm{pm}(L_n(G)) \quad = \quad \frac{v(G)}{2} \ln \left(\frac{(d-1)^{d-1}}{d^{d-2}} \right).$$

From Thm 5 & expression of $S_G^B(x)$ from page 2

This result is consistent with [18]. Indeed by Jensen inequality, we always have

$$\mathbb{E}\left[\ln \mathrm{pm}(L_n(G))\right] \leq \ln \mathbb{E}\left[\mathrm{pm}(L_n(G))\right],$$

and our result shows that in the large n limit, the two quantities are asymptotically equal. Note that the fact that $\mathbb{E}\left[\mathrm{pm}(L_n(G))^2\right] \sim \mathbb{E}\left[\mathrm{pm}(L_n(G))\right]^2$ to leading exponential order is not sufficient to prove this asymptotic equality.

A similar result to Theorem 5 was shown for permanent in [39, 40]. It is possible to define a random k-lift for a matrix A (by taking the weighted biadjacency matrix of the random k-lift of the weighted graph associated to the biadjacency matrix A). It is shown in [39, 40] that $\limsup_{k\to\infty} \frac{1}{k}\mathbb{E}[\ln \mathrm{per}(L_k(A))] = \max_{x\in M_{n,n}} \sum_{i,j}(1-x_{i,j})\ln(1-x_{i,j}) + x_{i,j}\ln\left(\frac{a_{i,j}}{x_{i,j}}\right)$. Since [38] recently showed that for A_k any k-lift of A, we have $\mathrm{per}(A_k) \leq \mathrm{per}(A)^k$, they obtained a new proof of (2).

any k-lift, not just for average over all k-lifts

3 Main ideas of the proof and more related works

Recall that $\ln P_G(z) = \ln \sum_k m_k(G) z^k$ is called the partition function. A crucial observation made by Csikvári [14] is that for any 2-lift H of a bipartite graph G, we have $\frac{1}{v(G)} \ln P_G(z) \geq \frac{1}{v(H)} \ln P_H(z)$, see Proposition 8 in the sequel. Note that a 2-lift of a d-regular graph is still d-regular. Starting form any d-regular bipartite graph G, Csikvári builds a sequence of 2-lifts with increasing girth. Then Csikvári uses the framework of local weak convergence in order to define a limiting graph for the sequence of 2-lifts. In this particular case, the limit is the infinite d-regular tree. Using the connection between the local weak convergence and the matching measure of a graph developed in [1], Csikvári computes a limiting partition function associated

girth: length of smallest cycle in the graph to this infinite d-regular tree which is obtained from the Kesten-McKay measure. This limiting partition function is then a lower bound for the original partition function $\ln P_G(z)$ and the lower bounds (3) and (4) are then easily obtained by properly choosing the parameter z as a function of k the size of the matchings that we need to count.

Our work extends the recent work of Csikvári done for regular and bi-regular bipartite graphs to irregular bipartite graphs. We are still building a sequence of 2-lifts with increasing girth and still using the framework of local weak convergence. The limiting object is now the universal cover of the initial graph G, i.e. the tree of non-backtracking walks also called the computation tree (see Section 4.3 for a precise definition). A direct computation of the matching measure of this (possibly infinite) tree seems tedious. Here, we depart significantly from the analysis of Csikvári. Our approach for the computation of the limiting partition function is based on an alternative (more algorithmic) characterization first developed in [11] based on local recursions on the universal cover of G. In order to express the computations done on the universal cover as a function of the original graph G, we rely on results proved by the author in [28] where the local recursions are studied on any finite graph. The solution of these local recursions on the universal cover (hence a possibly infinite tree) is in correspondence with the solution of the local recursions on the initial graph. Since this solution is given by the maximum of a certain "entropy-like" concave function defined by (7) on the fractional matching polytope of the original graph, we obtain an explicit formula for the limiting partition function and the lower bound in Theorem 3 follows. Our approach is then generalized to weighted bipartite graphs in order to get our results for (sub-)permanents (Theorem 4) and to random lifts in order to get Theorem 5.

As explained in the sequel of this section, all the basic ideas used in our proofs were present in a form or another in the literature: using lifts for extremal graph theory was one the main motivation for their introduction in a series of papers by Amit, Linial, Matousek, Rozenman and Bilu [4, 31, 5, 8]; the matching measure and the local recursions already appeared in the seminal work of Heilmann and Lieb [23]; the function S_G^B defined in (7) is known in statistical physics as the Bethe entropy [43]. The main contribution of this paper is a conceptual message showing how known techniques from interdisciplinary areas can lead to new applications in theoretical computer science. In the next subsections, we will try to relate our results to the existing literature and give credit to the many authors who inspired our work.

3.1 Covers, extremal graph theory and message passing algorithms

The idea to use lifts to build graphs with extremal properties is not new. In [8], Bilu and Linial study 2-lift of d-regular graphs in order to construct infinite families of expanders. They showed that the eigenvalues of a 2-lift are the union of the eigenvalues of the original graph and those of the signing associated to the 2-lift. They conjectured that every d-regular graph has a signing with spectral radius at most $2\sqrt{d-1}$. This conjecture was proved by Marcus, Spielman and Srivastava in [32] where they construct bipartite Ramnujan graphs of all degree.

We can informally state our Theorem 3 as an extremal graph theoretic result: among all bipartite graphs G having universal cover T, the universal cover T minimizes the (normalized) partition function $\frac{1}{v(G)} \ln P_G(z)$ for all $z > 0$, in particular it minimizes the (normalized) number of matchings $\frac{1}{v(G)} \sum_k m_k(G)$. Of course, T being infinite, the normalized partition function needs to be defined properly and this can be done thanks to the local weak convergence [11]. Indeed in the proof of Theorems 3 and 5, we will prove:

Proposition 2. *Let G be a finite bipartite graph and T be its universal cover. Let $\mathcal{G}$ be the set of finite bipartite graphs with universal cover T. We have for all $z > 0$,*

$$\inf_{G' \in \mathcal{G}} \frac{1}{v(G')} \ln P_{G'}(z) = \frac{1}{v(G)} \max_{\mathbf{x} \in M(G)} \left\{ \left(\sum_e x_e \right) \ln z + S_G^B(\mathbf{x}) \right\}. \tag{12}$$

Moreover, the sequence (G_n) defined in Theorem 3 or the sequence $L_n(G)$ converge in the local weak sense to T and achieves the bound (12) in the limit when n tends to infinity.

Note that $G_1, G_2 \in \mathcal{G}$ if and only if G_1 and G_2 have a common finite cover which is a result proved in [26]. Also, the right-hand term in (12) is an invariant of $\mathcal{G}$: this expression will be the same for any graph belonging to $\mathcal{G}$. Indeed our proof will proceed by computing its value thanks to the following message passing algorithm: to each edge $uv \in E$ and time step t, we associate two messages $y^t_{u\to v}(z)$ and $y^t_{v\to u}(z)$ obtained by setting $y^0_{u\to v} = y^0_{v\to u} = 0$ and for $t \geq 0$,

Loopy Annealing

B.P. Lelarge 2014

$$\boxed{y^{t+1}_{u\to v}(z) = \frac{z}{1 + \sum_{w\in\partial u\backslash v} y^t_{w\to u}(z)}.} \tag{13}$$

We show that as $t \to \infty$, these iterations will converge to a limit $y_{u\to v}(z)$ and that $x_{uv}(z) = \frac{y_{u\to v}(z) y_{v\to u}(z)}{z + y_{u\to v}(z) y_{v\to u}(z)}$ solves the maximization in (12) (see Propositions 5, 6, 7). Since the recursion (13) is local, the messages obtained after t iterations are the same as the one computed on the computation tree of the graph at depth t. We are able to show that these recursions on infinite trees still have a unique fixed point (see Theorem 9) so that this fixed point should be the limit (as the number of iterations tend to infinity) obtained by our algorithm runned on the original graph G. Note that the recursion (13) is well-known and first appeared in the analysis of the monomer dimer problem [23]. It is also used to define a deterministic approximation algorithm for counting matchings in [6]. Indeed, their analysis directly implies that the convergence of our algorithm is exponentially fast in the number of iterations t. Note however that the recursion used in [6] corresponds to messages sent on the tree of self-avoiding paths. Instead, we use the tree of non-backtracking paths. The tree of self-avoiding paths is finite and depends on the root whereas our tree is the universal cover of the graph. Also, [6] directly implies the convergence of our message passing algorithm, it does not give any indication about the value of the limit.

At this stage, we should recall that computing the number of matchings falls into the class of $\#P$-complete problems as well as the problem of counting the number of perfect matchings in a given bipartite graph, i.e. computing the permanent of an arbitrary $0-1$ matrix. By previous discussion, we see that if the graph is locally tree like, then the tree of self-avoiding paths and the universal cover are locally the same, and one can believe that our algorithm will compute a good approximation for counting matchings. This idea was formalized in [43] and proved rigorously in [11] for random graphs. Our Theorem 5 shows that these results extend to random lifts. The lower bound in (2) is called the (logarithm of the) Bethe permanent in the physics literature [42, 13, 39]. Similar ideas using lifts or covers of graphs have appeared in the literature about message passing algorithms, see [34, 35] and references therein. We refer to [28] for more results connecting Belief Propagation with our setting.

3.2 Matching measure and spectral measure of trees

We now relate our results to the matching measure used by Csikvári in [14] and show how our results allow us to compute spectral measure of infinite trees. The matching polynomial is defined as:

$$\boxed{Q_G(z) = \sum_{k=0}^{\nu(G)} (-1)^k m_k(G) z^{n-2k} = z^n P_G(-z^{-2}).}$$

We define the matching measure of G denoted by ρ_G as the uniform distribution over the roots of the matching polynomial of G:

$$\rho_G = \frac{1}{\nu(G)} \sum_{i=1}^{\nu(G)} \delta_{z_i}, \tag{14}$$

where the z_i's are the roots of Q_G. Note that $Q_G(-z) = (-1)^n Q_G(z)$ so that ρ_G is symmetric.

The fundamental theorem for the matching polynomial is the following.

Theorem 6. *(Heilmann Lieb [23]) The roots of the matching polynomial $Q_G(z)$ are real and in the interval $[-2\sqrt{D_G - 1}, 2\sqrt{D_G - 1}]$, where D_G is the maximal degree in G.*

hefel

In particular, the matching measure of G is a probability measure on $\mathbb{R}$. Of course, the polynomials $P_G(z)$ or $Q_G(z)$ contains the same information as the matching measure ρ_G. We can express the quantity of interest in term of ρ_G (see Lemma 8.5 in [23], [1] or [14]): for $z > 0$,

mean size of the matching under

$$\begin{aligned} \frac{1}{v(G)} \frac{zP'_G(z)}{P_G(z)} &= \frac{1}{2} \int \frac{z\lambda^2}{1 + z\lambda^2} d\rho_G(\lambda), \\ \frac{\nu(G)}{v(G)} &= \frac{1}{2} \left(1 - \rho_G(\{0\})\right). \end{aligned}$$

As explained above, Csikvári [14] uses this representation and the fact that for a sequence of d-regular graphs converging to a d-regular tree, the limiting matching measure is given by the Kesten-MacKay measure, to get an explicit formula for the limiting partition function. Our approach relies on local recursions instead of the connection with the matching measure. Since we are able to solve these recursions, we get the following result for the limiting matching measure.

Theorem 7. *Let G be a finite graph and $T(G)$ be its universal cover. For any sequence of graphs (G_i) with maximal degree D_G and with local weak limit $T(G)$, the matching measure ρ_{G_i} of the graph G_i is weakly convergent to some measure $\mu_{T(G)}$ defined by the formula for $z > 0$,*

$$\int \frac{z\lambda^2}{1 + z\lambda^2} d\mu_{T(G)}(\lambda) = \frac{2}{v(G)} \left(\sum_e x_e(z) \right),$$

where the vector $\mathbf{x}(z)$ is the unique maximizer of $\Phi_G^B(\mathbf{x}, z)$ in $FM(G)$. Moreover, we have $\mu_{T(G)}(\{0\}) = 1 - 2\frac{\nu^*(G)}{v(G)}$.

Note that our theorem gives a generating function of the moments of $\mu_{T(G)}$ since for $z > 0$ sufficiently small, we have:

$$\int \frac{z\lambda^2}{1 + z\lambda^2} d\mu_{T(G)}(\lambda) = \sum_{i=1}^{\infty} (-1)^{i+1} z^i \int \lambda^{2i} d\mu_{T(G)}(\lambda),$$

and the series is convergent since the support of all the ρ_{G_i} and hence of $\mu_{T(G)}$ is contained in $[-2\sqrt{D_G-1}, 2\sqrt{D_G-1}]$. As shown by Godsil in [17], for finite trees, the spectral measure and the matching measure coincide, this is still true for infinite trees [9, 10, 11]. In particular, the moments of the matching measure $\int \lambda^{2i} d\mu_{T(G)}(\lambda)$ can be interpreted as the average number of closed walks on $T(G)$ where the average is taken over the starting point of the walk (see Proposition 11 for a precise definition of the random root as the starting point of the walk).

To be more precise, for a finite graph G, we denote by $\lambda_1 \leq \cdots \leq \lambda_{v(G)}$ the real eigenvalues of its adjacency matrix and we define the empirical spectral measure of the graph G as the probability measure on $\mathbb{R}$:

$$\mu_G = \frac{1}{v(G)} \sum_{i=1}^{v(G)} \delta_{\lambda_i}.$$

The following theorem follows from [9, 10, 11] (see also Chapter 2 in [29])

Theorem 8. *Let G be a finite graph and $T(G)$ be its universal cover. For any sequence of graphs (G_i) with maximal degree D_G and with local weak limit $T(G)$, the spectral measure μ_{G_i} of the graph G_i is weakly convergent to the measure $\mu_{T(G)}$ defined in Theorem 7. Moreover for all $x \in \mathbb{R}$, we have $\lim_{i\to\infty} \mu_{G_i}(\{x\}) = \mu_{T(G)}(\{x\})$.*

In particular, we can apply this theorem to characterize the limiting spectral measure of random lifts $L_n(G)$ as a function of the original graph G. For the atom at zero, we have:

$$\lim_{n\to\infty} \mu_{L_n(G)}(\{0\}) = 1 - 2\frac{\nu^*(G)}{v(G)}, \tag{15}$$

and Theorem 7 allows us to get the moments of the limiting measure of $\mu_{L_n(G)}$.

4 Proofs

4.1 Statistical physics

To ease the notation, we will consider a setting with a weighted graph $G = (V, E)$ with positive weights on edges $\{\theta_e\}_{e\in E}$. Taking a bipartite graph G and $\theta_e = 1$ for all $e \in E$, we recover the framework of Section 2.1. To recover the more general framework of Section 2.2, consider the bipartite graph described by the support of A seen as an incidence matrix and for each $e = (ij) \in E$, define $\theta_e = a_{i,j}$.

We introduce the family of probability distributions on the set of matchings in G parametrised by a parameter $z > 0$:

$$\mu_G^z(\mathbf{B}) = \frac{z^{\sum_e B_e} \prod_{e\in\mathbf{B}} \theta_e}{P_G(z)}, \tag{16}$$

where $P_G(z) = \sum_{\mathbf{B}} z^{\sum_e B_e} \prod_{e\in\mathbf{B}} \theta_e \prod_{v\in V} \mathbf{1}\left(\sum_{e\in\partial v} B_e \leq 1\right) = \sum_{k=0}^{\nu(G)} w_k(G) z^k$, with

$$w_k(G) = \sum_{\{\mathbf{B}:\, \sum_e B_e = k\}} \prod_{e\in\mathbf{B}} \theta_e,$$ total weight of matchings of size k

$$U_G^\theta(z) = \sum_K z^K \sum_{B:\, |B|=K} w(B)\Big(\sum_{e\in B} \log \theta_e\Big) = \sum_K z^K \sum_{B,\, |B|=K} w(B) \log w(B)$$

where the sum is over matchings of size k. Note that we have $w_k(G) = \text{per}_k(A)$. Note also that when z tends to infinity, the measure μ_G^z converges to the measure:

sub-Perm sum

$$\mu_G^\infty(\mathbf{B}) = \frac{\prod_{e\in E}\theta_e}{\text{per}_{\nu(G)}(\theta)},$$

which is simply the uniform measure on maximum matchings when $\theta_e = 1$ for all edges. In statistical physics, this model is known as the monomer-dimer model and its analysis goes back to the work of Heilmann and Lieb [23].

We define the following functions:

$\mu_G^z(B_e=1)$: sum over all matchings that include the edge e

$$\boxed{\begin{aligned} U_G^s(z) &= -\sum_{e\in E} \mu_G^z(B_e = 1), \\ U_G^\theta(z) &= \sum_{e\in E} \mu_G^z(B_e=1)\ln\theta_e, \\ S_G(z) &= -\sum_{\mathbf{B}} \mu_G^z(\mathbf{B})\ln\mu_G^z(\mathbf{B}). \end{aligned}}$$

Gibbs Free Energy

main diff from vanilla etc is that G is NOT bipartite

Note that when $\theta_e = 1$, we have $U_G^\theta(z) = 0$ and U_G^s is called the internal energy while S_G is the canonical entropy. We now define the partition function $\Phi_G(z)$ by

$$\boxed{\Phi_G(z) = -U_G^s(z)\ln z + U_G^\theta(z) + S_G(z).}$$

A more conventional notation in the statistical physics literature corresponds to an inverse temperature $\beta = \ln z$. Note that with our definitions, the internal energy $U_G^s(z)$ is negative, equals to minus the average size of a matching sampled from μ_G^z. This convention is consistent with standard models in statistical physics where the low temperature regime minimizes the internal energy, i.e. in our context maximizes the size of the matching. A simple computation shows that:

→ from these

$$\Phi_G(z) = \ln P_G(z) \text{ and, } \Phi_G'(z) = \frac{-U_G^s(z)}{z}.$$

Lemma 1. *The function $U_G^s(z)$ is strictly decreasing and mapping $[0,\infty)$ to $(-\nu(G), 0]$.*

Proof. We have $-U_G^s(z) = \sum_k kw_k(G)z^k/P_G(z)$ so that taking the derivative and multiplying by z, we get:

every matching of size K occurs K times (one time for each edge)

$$\begin{aligned} -z(U_G^s)'(z) &= \frac{\sum_k k^2 w_k(G)z^k}{P_G(z)} - \left(\frac{\sum_k kw_k(G)z^k}{P_G(z)}\right)^2 \\ &= \sum_k \left(k - \frac{\sum_\ell \ell w_\ell(G)z^\ell}{P_G(z)}\right)^2 \frac{w_k(G)z^k}{P_G(z)} > 0. \end{aligned}$$

□

We define $\tau = \tau(G) = 2\nu(G)/v(G)$ which is the maximum fraction of nodes covered by a matching in G. Note that $\tau(G) \le 1$ and $\tau(G) = 1$ if and only if the graph G has a perfect matching. For $t \in [0,\tau)$, we define $z_t(G) \in [0,\infty)$ as the unique root to $U_G^s(z_t(G)) = -tv(G)/2$.

$z_t(G)$ is the value of z so that $-U_G^s(z)$ equals t times the size of perfect matching

Note that $t \mapsto z_t(G)$ is an increasing function which maps $[0,\tau)$ to $[0,\infty)$. The function $\Sigma_G(t)$ is then defined for $t \in [0,\tau)$ by:

$$\Sigma_G(t) = \frac{S_G(z_t(G)) + U_G^\theta(z_t(G))}{v(G)}, \tag{17}$$

significance

and $\Sigma_G(t) = -\infty$ for $t > \tau$.

Proposition 3. *For $t < \tau$, we have $\Sigma'_G(t) = -\frac{1}{2}\ln z_t(G)$. The limit $\lim_{t\to\tau}\Sigma_G(t)$ exists and we define $\Sigma_G(\tau) = \lim_{t\to\tau}\Sigma_G(t) = \frac{1}{v(G)}\ln w_{\nu(G)}(G)$.*

$\rightarrow$ sum of maximum matchings

Proof. We have for $t < \tau$, $\Sigma_G(t) = \frac{1}{v(G)}\ln P_G(z_t) - t/2\ln z_t$, so that taking the derivative with respect to t, we get:

$\rightarrow z^{\nu(G)}$ term will dominate

$$\Sigma'_G(t) \;=\; z'_t\left(\frac{-t}{2z_t} + \frac{P'_G(z_t)}{v(G)P_G(z_t)}\right) - \frac{\ln z_t}{2}.$$

$\Sigma_G(t) = \frac{1}{v(G)} \ln z_t^{\nu(G)} w_{\nu(G)}(G) - \frac{\tau}{2}\ln z_t$, $t \to \tau$

$= \frac{1}{v(G)} \ln w_{\nu(G)}(G)$ $\because \frac{\nu(G)}{v(G)}\ln z_t = \frac{\tau}{2}\ln z_t$

Since $U_G^s(z) = -z\frac{P'_G(z)}{P_G(z)}$ and $U_G^s(z_t) = -tv(G)/2$, we get $\Sigma'_G(t) = -\frac{1}{2}\ln z_t$. For t large enough, we have $z_t \geq 1$ and the proposition follows. ∎

The following proposition is proved in [14] for unweighted graphs (see Proposition 2.1(g)) but the proof is the same in the weighted case. We include it her for convenience.

Proposition 4. *If for some graphs G_1 and G_2, we have for every $z \geq 0$,*

$$\frac{\Phi_{G_1}(z)}{v(G_1)} \geq \frac{\Phi_{G_2}(z)}{v(G_2)},$$

$\Rightarrow \frac{\ln P_{G_1}(z)}{v(G_1)} \geq \frac{\ln P_{G_2}(z)}{v(G_2)} \Rightarrow \frac{\deg P_{G_1}}{v} \geq \frac{\deg P_{G_2}}{v}$

$\Rightarrow \frac{\nu(G_1)}{v} \geq \cdots$

then

$$\Sigma_{G_1}(t) \geq \Sigma_{G_2}(t)$$

for all $0 \leq t \leq 1$.

Proof. The assumption ensures that $\frac{\nu(G_1)}{v(G_1)} \geq \frac{\nu(G_2)}{v(G_2)}$. Moreover if $\frac{\nu(G_1)}{v(G_1)} = \frac{\nu(G_2)}{v(G_2)}$, then

$$\frac{\ln w_{\nu(G_1)}(G_1)}{v(G_1)} \geq \frac{\ln w_{\nu(G_2)}(G_2)}{v(G_2)}.$$

Hence the statement is trivial for $t \geq 2\nu(G_2)/v(G_2)$. We consider now $t \in [0, 2\nu(G_2)/v(G_2))$. Note that $\Sigma_{G_1}(0) = \Sigma_{G_2}(0) = 0$. The derivative of $\Sigma_{G_1}(t) - \Sigma_{G_2}(t)$ for $t < 2\nu(G_2)/v(G_2)$ is

$$-\frac{1}{2}\left(\ln z_t(G_1) - \ln z_t(G_2)\right)$$

Assume this derivative is 0 at t_0, then we have $z_{t_0}(G_1) = z_{t_0}(G_1) = z_0$ and then

$$\frac{S_{G_1}(z_0)}{v(G_1)} = \frac{\ln P_{G_1}(z_0)}{v(G_1)} - \frac{t_0}{2}\ln z_0 \geq \frac{\ln P_{G_2}(z_0)}{v(G_2)} - \frac{t_0}{2}\ln z_0 = \frac{S_{G_2}(z_0)}{v(G_2)}$$

Hence the minimums of $\Sigma_{G_1}(t) - \Sigma_{G_2}(t)$ on $[0, 2\nu(G_2)/v(G_2))$ are non-negative. ∎

4.2 Local recursions on finite graphs and infinite trees

Let $G = (V, E)$ be a (possibly infinite) graph with bounded degree and weights on edges $\{\theta_e\}_{e \in E}$. We introduce the set $\overrightarrow{E}$ of directed edges of G comprising two directed edges $u \to v$ and $v \to u$ for each undirected edge $(uv) \in E$. For $\overrightarrow{e} \in \overrightarrow{E}$, we denote by $-\overrightarrow{e}$ the edge with opposite direction. With a slight abuse of notation, we denote by ∂v the set of incident edges to $v \in V$ directed towards v. We also denote by $\partial v \backslash u$ the set of neighbors of v from which we removed u. We also use this notation to denote the set of incident edges to v directed towards v from which we removed $u \to v$.

Given G, we define the map $\mathcal{R}_G : (0, \infty)^{\overrightarrow{E}} \to (0, \infty)^{\overrightarrow{E}}$ by $\mathcal{R}_G(\mathbf{a}) = \mathbf{b}$ with

↳ contains message for each edge

- similar to BP given that we normalize the messages and beliefs
- no div by 0 occurs

$$b_{u \to v} = \frac{1}{1 + \sum_{w \in \partial u \backslash v} \theta_{wu} a_{w \to u}},$$

LABP

with the convention that the sum over the empty set equals zero. We also denote by $\mathcal{R}_{u \to v} : (0, \infty)^{\partial u \backslash v} \to (0, \infty)$ the local mapping defined by: $b_{u \to v} = \mathcal{R}_{u \to v}(\mathbf{a})$ (note that only the coordinates of $\mathbf{a}$ in $\partial u \backslash v$ are taken as input of $\mathcal{R}_{u \to v}$). Comparisons between vectors are always componentwise.

Proposition 5. *Let G be a finite graph. For any $z > 0$, the fixed point equation $\mathbf{y}(z) = z\mathcal{R}_G(\mathbf{y}(z))$ has a unique attractive solution $\mathbf{y}(z) \in (0, +\infty)^{\overrightarrow{E}}$. The function $z \mapsto \mathbf{y}(z)$ is increasing and the function $z \mapsto \frac{\mathbf{y}(z)}{z}$ is decreasing for $z > 0$.*

Note that the mapping $z\mathcal{R}_G$ defined in this proposition is simply the mapping multiplying by z each component of the output of the mapping $\mathcal{R}_G$ (making the notation consistent).

Proof. This result is proved for the case $\theta_e = 1$ for all edges in [28] (see also [36]) and the proof extends to this setting. □

We define for all $v \in V$, the following function of the vector $(y_{\overrightarrow{e}}, \overrightarrow{e} \in \partial v)$,

Prob of v being covered by a matching distributed acc to μ_G^z

$$\mathcal{D}_v(\mathbf{y}) = \sum_{\overrightarrow{e} \in \partial v} \frac{\theta_e y_{\overrightarrow{e}} \mathcal{R}_{-\overrightarrow{e}}(\mathbf{y})}{1 + \theta_e y_{\overrightarrow{e}} \mathcal{R}_{-\overrightarrow{e}}(\mathbf{y})} \tag{18}$$

$$= \frac{\sum_{\overrightarrow{e} \in \partial v} \theta_e y_{\overrightarrow{e}}}{1 + \sum_{\overrightarrow{e} \in \partial v} \theta_e y_{\overrightarrow{e}}}. \tag{19}$$

Clearly from (19), we see that $\mathcal{D}_v$ is an increasing function of $\mathbf{y}$ and the proposition below follows directly from the monotonicity of $\mathbf{y}(z)$ proved in Proposition 5:

Proposition 6. *Let $G = (V, E)$ be a finite graph and $\mathbf{y}(z)$ be the solution to $\mathbf{y}(z) = z\mathcal{R}_G(\mathbf{y}(z))$. For any $v \in V$, the mapping $z \mapsto \mathcal{D}_v(\mathbf{y}(z))$ is increasing and $\mathcal{D}_v(\mathbf{y}(z)) = \sum_{e \in \partial v} x_e(z)$, where*

belongs to FM polytope
≅ belief polytope

$$x_e(z) = \frac{\theta_e y_{\overrightarrow{e}}(z) y_{-\overrightarrow{e}}(z)}{z + \theta_e y_{\overrightarrow{e}}(z) y_{-\overrightarrow{e}}(z)} \in (0, 1). \tag{20}$$

We denote by $\mathbf{x}(z) = (x_e(z), e \in E)$ the vector defined by (20), then $\mathbf{x}(z) \in FM(G)$ and we have:

$$\lim_{z \to \infty} \sum_{v \in V} \mathcal{D}_v(\mathbf{y}(z)) = 2\nu^*(G). \tag{21}$$

} similar to eq (13) in LABP, '14

S_G^B encodes both Bethe Avg energy and bethe entropy as defined in vontobel

Proof. The only non-trivial statement in the above proposition is the value of the limit in (21). In the case $\theta_e = 1$, it follows from Theorem 1 in [28] and the proof carries over to the case $\theta_e > 0$. $\square$

For a finite graph $G = (V, E)$ with weights on edges $\{\theta_e\}_{e\in E}$, we define for $\mathbf{x} \in FM(G)$ defined by (5) and $z > 0$,

Belief polytope

$S_G^B(x)$ contains θ terms unlike $S_G(z)$

$$\boxed{\begin{aligned} U_G^B(\mathbf{x}) &= -\sum_{e\in E} x_e, \\ S_G^B(\mathbf{x}) &= \sum_{e\in E} x_e \ln \frac{\theta_e}{x_e} + (1-x_e)\ln(1-x_e) - \sum_{v\in V}\left(1 - \sum_{e\in\partial v} x_e\right)\ln\left(1 - \sum_{e\in\partial v} x_e\right), \\ \Phi_G^B(\mathbf{x}, z) &= -U_G^B(\mathbf{x})\ln z + S_G^B(\mathbf{x}). \end{aligned}}$$

not imp for m/a in bipartite

BFE for NON-bipartite graphs

We denote by $\mathbf{x}(z)$ the vector defined by (20) in Proposition 6 where $\mathbf{y}(z) = z\mathcal{R}_G(\mathbf{y}(z))$. Note that

$$U_G^B(\mathbf{x}(z)) = \frac{-1}{2}\sum_{v\in V} \mathcal{D}_v(\mathbf{y}(z)), \qquad (22)$$

by def of D & X

so that by Proposition 6, the mapping $z \mapsto U_G^B(\mathbf{x}(z))$ is decreasing from $[0,\infty)$ to $(-\nu^*(G), 0]$. Thus, we can define z_t^B as the unique solution in $[0,\infty)$ to

analog. to $U_G^S(z)$

$$U_G^B(\mathbf{x}(z_t^B)) = -\frac{tv(G)}{2}, \text{ for } t < \tau^*(G) = \frac{2\nu^*(G)}{v(G)}.$$

Similarly as in (17), we define

S contains θ terms also

$$\Sigma_G^B(t) = \frac{S_G^B(\mathbf{x}(z_t^B))}{v(G)} \text{ for } t < \tau^*(G).$$

fractional

Similar to $\Sigma_G(t)$

Note that we have $\tau^*(G) \geq \tau(G)$ with equality if G is bipartite.

Proposition 7. *Recall that $\mathbf{x}(z) \in FM(G)$ is defined by (20). We have for any $z > 0$,*

$$\sup_{\mathbf{x}\in FM(G)} \Phi_G^B(\mathbf{x}; z) = \Phi_G^B(\mathbf{x}(z); z),$$

For a fixed z in ϕ_G^B, supremum is achieved at $x(z)$

corr. to the fixed point $y(z)$

and for $t < \tau^*(G)$,

$$\Sigma_G^B(t) = \frac{1}{v(G)} \max_{\mathbf{x}\in FM_{tv(G)/2}(G)} S_G^B(\mathbf{x}),$$

maximized at $x(z_t^B)$

where FM_t is defined in (6) and where the maximum taken over an empty set is equal to $-\infty$.

Proof. The first statement is proved in [28] for the case where $\theta_e = 1$ but extends easily to the current framework. For the second statement, note that for any $\mathbf{x} \in FM_{tv(G)/2}(G)$ with $t < \tau^*(G)$, we have

$$\Phi_G^B(\mathbf{x}, z_t^B) = \frac{tv(G)}{2}\ln z_t^B + S_G^B(\mathbf{x}) \leq \Phi_G^B(\mathbf{x}(z_t^B), z_t^B) = \frac{tv(G)}{2}\ln z_t^B + S_G^B(\mathbf{x}(z_t^B)).$$

By definition, we have $\mathbf{x}(z_t^B) \in M_{tv(G)/2}(G)$, so that $\max_{\mathbf{x}\in M_{tv(G)/2}(G)} S_G^B(\mathbf{x}) = S_G^B(\mathbf{x}(z_t^B))$. $\square$

M_k or FM_k?

We now extend Proposition 5 to infinite trees:

Theorem 9. *Let $T = (V, E)$ be a (possibly infinite) tree with bounded degree. For each $z > 0$, there exists a unique solution in $(0, \infty)^{\vec{E}}$ to the fixed point equation $\mathbf{y}(z) = z\mathcal{R}_T(\mathbf{y}(z))$, i.e. such that*

$$y_{u\to v}(z) = \frac{z}{1 + \sum_{w\in\partial u\setminus v} \theta_{wu} y_{w\to u}(z)}. \tag{23}$$

Proof. First note that any non-negative solution must satisfy $y_{u\to v}(z) \le z$ for all $(uv) \in E$. The compactness of $[0, z]^{\vec{E}}$ (as a countable product of compact spaces) guarantees the existence of a solution by Schauder fixed point theorem.

To prove the uniqueness, we follow the approach in [6]. First, we define the change of variable: $h_{u\to v} = -\ln \frac{y_{u\to v}(z)}{z}$ so that (23) becomes:

$$h_{u\to v} = \ln\left(1 + z \sum_{w\in\partial u\setminus v} \theta_{wu} e^{-h_{w\to u}}\right). \tag{24}$$

We define the function $f : [0, +\infty)^d \mapsto [0, \infty)$ as:

$$f(\mathbf{h}) = \ln\left(1 + z\sum_{i=1}^{k} \frac{\theta_i}{1 + z\sum_{j=1}^{k_i} \theta_j^i e^{-h_j^i}}\right),$$

where the parameters k, k_i, θ_i, θ_j^i and z are fixed and $d = \sum_{i=1}^{k} k_i$.

Iterating the recursion (24), we can rewrite it using such a function f so that uniqueness would be implied if we show that f is contracting.

For any $\mathbf{h}$ and $\mathbf{h}'$, we apply the mean value theorem to the function $f(\alpha\mathbf{h} + (1-\alpha)\mathbf{h}')$ so that there exists $\alpha \in [0, 1]$ such that for $\mathbf{h}_\alpha = \alpha\mathbf{h} + (1-\alpha)\mathbf{h}'$,

$$|f(\mathbf{h}) - f(\mathbf{h}')| = |\nabla f(\mathbf{h}_\alpha)(\mathbf{h} - \mathbf{h}')| \le \|\nabla f(\mathbf{h}_\alpha)\|_{L_1} \|\mathbf{h} - \mathbf{h}'\|_\infty.$$

A simple computation shows that:

$$\|\nabla f(\mathbf{h})\|_{L_1} = \frac{z\sum_{i=1}^{k} \theta_i \frac{z\sum_{j=1}^{k_i} \theta_j^i e^{-h_j^i}}{\left(1+z\sum_{j=1}^{k_i} \theta_j^i e^{-h_j^i}\right)^2}}{1 + z\sum_{i=1}^{k} \frac{\theta_i}{1+z\sum_{j=1}^{k_i} \theta_j^i e^{-h_j^i}}}.$$

Let $A_i = \left(1 + z\sum_{j=1}^{k_i} \theta_j^i e^{-h_j^i}\right)^{-1}$, then we get

$$\|\nabla f(\mathbf{h})\|_{L_1} = \frac{z\sum_{i=1}^{k} \theta_i (A_i - A_i^2)}{1 + z\sum_{i=1}^{k} \theta_i A_i} = 1 - \frac{1 + z\sum_{i=1}^{k} \theta_i A_i^2}{1 + z\sum_{i=1}^{k} \theta_i A_i}.$$

By taking the partial derivatives, we note that this last expression is maximized when all A_i are equal. Then the solution for the optimal A_i reduces to a quadratic equation with solution in

$[0,+\infty)$ equals to $A_i = \frac{\sqrt{1+z\Theta}-1}{z\Theta}$, where $\Theta = \sum_{i=1}^{k} \theta_i$. Substituting for the maximum value, we get for any real vector $\mathbf{h}$,

$$\|\nabla f(\mathbf{h})\|_{L_1} \leq 1 - \frac{2}{\sqrt{1+z\Theta}+1}.$$

$\square$

4.3 2-lifts

If G is a graph and $v \in V(G)$, the 1-neighbourhood of v is the subgraph consisting of all edges incident upon v. A graph homomorphism $\pi : G' \to G$ is a covering map if for each $v' \in V(G')$, π gives a bijection of the edges of the 1-neighbourhood of v' with those of $v = \pi(v')$. G' is a cover or a lift of G. If edges of $G = (V, E)$ have weights θ_e then the edges of $G' = (V', E')$ will also have weights with $\theta_{e'} = \theta_{\pi(e')}$. Note that the definition of 2-lift for matrices given in Section 2.2 is consistent with the definition of 2-lift for graphs by identifying the matrix A as the weighted incidence matrix of the bipartite graph.

Proposition 8. *Let G be a bipartite graph and H be a 2-lift of G. Then $P_G(z)^2 \geq P_H(z)$ for $z > 0$, $\Sigma_G(t) \geq \Sigma_H(t)$ for $t \in [0,1]$ and $\nu(H) = 2\nu(G)$.*

Proof. The proof follows from an argument of Csikvári [14]. Note that $G \cup G$ is a particular 2-lift of G with $P_{G\cup G}(z) = P_G(z)^2$. To prove the first statement of the proposition, we need to show that for any 2-lift H of G, we have: $w_k(G \cup G) \geq w_k(H)$. Consider the projection of a matching of a 2-lift of G to G. It will consist of disjoint union of cycles of even lengths (since G is bipartite), paths and double-edges when two edges project to the same edge. For such a projection $R = R_1 \cup R_2 \subset E$ where R_2 is the set of double edges, its weight is $\prod_{e\in R_1} \theta_e \prod_{e\in R_2} \theta_e^2$. Now for such a projection, we count the number of possible matchings in $G \cup G$: $n_R(G \cup G) = 2^{k(R)}$, where $k(R)$ is the number of connected components of R_1. The number of possible matchings in H is $n_R(H) \leq 2^{k(R)}$ since in each component if the inverse image of one edge is fixed then the inverse images of all other edges is also determined. There is no equality as in general not every cycle can be obtained as a projection of a matching of a 2-lift. For example, if one considers a 8-cycle as a 2-lift of a 4-cycle, then no matching will project on the whole 4-cycle.

Hence we proved that $w_k(G \cup G) \geq w_k(H)$ so that $P_G(z)^2 \geq P_H(z)$ for $z > 0$ and the second statement follows from Proposition 4. For the last statement, since $P_G(z)^2 \geq P_H(z)$, we have $2\nu(G) \geq \nu(H)$ but the opposite inequality is true for any graph G since a maximum matching in G can be lifted to a matching in H with size twice the size of the original matching. $\square$

Given a graph G with a distinguished vertex $v \in V$, we construct the (infinite) rooted tree $(T(G), v)$ of non-backtracking walks at v as follows: its vertices correspond to the finite non-backtracking walks in G starting in v, and we connect two walks if one of them is a one-step extension of the other. With a slight abuse of notation, we denote by v the root of the tree of non-backtracking walks started at v. Note that also we constructed $T(G)$ from a particular vertex v, this choice is irrelevant. It is easy to see that $T(G)$ is a cover of G, indeed it is the (unique up to isomorphism) cover of G that is also a cover of every other cover of G. $T(G)$ is called the universal cover of G.

Since the local recursions are the same for both $\mathcal{R}_{T(G)}$ and $\mathcal{R}_G$ and since there is a unique fixed point for both $z\mathcal{R}_{T(G)}$ and $z\mathcal{R}_G$, the proposition below follows:

Proposition 9. *Let G be a finite graph and $T(G)$ be its universal cover and associated cover $\pi : T(G) \to G$. By Propositions 6 and 5, we can define:*

$$\tilde{\mathbf{y}}(z) = z\mathcal{R}_{T(G)}(\tilde{\mathbf{y}}(z)) \text{ , } \textit{and, } \mathbf{y}(z) = z\mathcal{R}_G(\mathbf{y}(z)).$$

We have $\pi(\tilde{\mathbf{y}}(z)) = \mathbf{y}(z)$, i.e. $\tilde{y}_{\vec{e}}(z) = y_{\pi(\vec{e})}(z)$.

4.4 The framework of local weak convergence

This section gives a brief account of the framework of local weak convergence. For more details, we refer to the surveys [3, 2].

Rooted graphs. A *rooted graph* (G, o) is a graph $G = (V, E)$ together with a distinguished vertex $o \in V$, called the *root*. We let $\mathcal{G}_\star$ denote the set of all locally finite connected rooted graphs considered up to *rooted isomorphism*, i.e. $(G, o) \equiv (G', o')$ if there exists a bijection $\gamma\colon V \to V'$ that preserves roots ($\gamma(o) = o'$) and adjacency ($\{i, j\} \in E \iff \{\gamma(i), \gamma(j)\} \in E'$). We write $[G, o]_h$ for the (finite) rooted subgraph induced by the vertices lying at graph-distance at most $h \in \mathbb{N}$ from o. The distance

$$\text{DIST}\left((G, o), (G', o')\right) := \frac{1}{1 + r} \quad \text{where} \quad r = \sup\left\{h \in \mathbb{N}\colon [G, o]_h \equiv [G', o']_h\right\},$$

→ max h for which both graphs stay isomorphic

turns $\mathcal{G}_\star$ into a complete separable metric space, see [2]. How is this separable?

With a slight abuse of notation, (G, o) will denote an equivalence class of rooted graph also called unlabeled rooted graph in graph theory terminology. Note that if two rooted graphs are isomorphic, then their rooted trees of non-backtracking walks are also isomorphic. It thus makes sense to define $(T(G), o)$ for elements $(G, o) \in \mathcal{G}_\star$.

Proposition 10. *For any graph $G = (V, E)$, there exists a graph sequence $\{G_n\}_{n \in \mathbb{N}}$ such that $G_0 = G$, G_n is a 2-lift of G_{n-1} for $n \geq 1$. Hence G_n is a 2^n-lift of G and we denote by $\pi_n : G_n \to G$ the corresponding covering. For any $v \in V$, if $v_n \in \pi_n^{-1}(v)$, we have $(G_n, v_n) \to (T(G), v)$ in $\mathcal{G}_\star$.*

Proof. The proof follows from an argument of Nathan Linial [30], see also [14].

A random 2-lift H of a base graph G is the random graph obtained by choosing between the two pairs of edges $((u, 0), (v, 0))$ and $((u, 1), (v, 1)) \in E(H)$ or $((u, 0), (v, 1))$ and $((u, 1), (v, 0)) \in E(H)$ with probability $1/2$ and each choice being made independently.

Let G be a graph with girth γ and let k be the number of cycles in G with size γ. Let X be the number of γ-cycles in H a random 2-lift of G. The girth of H must be at least γ and a γ-cycle in H must be a lift of a γ-cycle in G. A γ-cycle in G yields: a 2γ-cycle in H with probability $1/2$; or two γ-cycles in H with probability $1/2$. Hence we have $\mathbb{E}[X] = k$. But $G \cup G$ (the trivial lift) has $2k$ γ-cycles. Hence there exists a 2-lift with strictly less than k γ-cycles. By iterating this step, we see that there exists a sequence $\{G_n\}$ of 2-lifts such that for any γ, there exists a $n(\gamma)$ such that for $j \geq n(\gamma)$, the graph G_j has no cycle of length at most γ. This implies that for any $v \in V$ and $v_j \in \pi_j^{-1}(v)$, we have $\text{DIST}\left((G_j, v_j), (T(G), v)\right) \leq \frac{2}{\gamma}$ and the proposition follows. $\square$

Local weak limits. Let $\mathcal{P}(\mathcal{G}_\star)$ denote the set of Borel probability measures on $\mathcal{G}_\star$, equipped with the usual topology of weak convergence (see e.g. [7]). Given a finite graph $G = (V, E)$, we construct a random element of $\mathcal{G}_\star$ by choosing uniformly at random a vertex $o \in V$ to be the root, and restricting G to the connected component of o. The resulting law is denoted by $\mathcal{U}(G)$. If $\{G_n\}_{n\geq 1}$ is a sequence of finite graphs such that $\{\mathcal{U}(G_n)\}_{n\geq 1}$ admits a weak limit $\mathcal{L} \in \mathcal{P}(\mathcal{G}_\star)$, we call $\mathcal{L}$ the *local weak limit* of $\{G_n\}_{n\geq 1}$. If (G, o) denotes a random element of $\mathcal{G}_\star$ with law $\mathcal{L}$, we shall use the following slightly abusive notation : $G_n \rightsquigarrow (G, o)$ and for $f : \mathcal{G}_\star \to \mathbb{R}$:

$$\mathbb{E}_{(G,o)}\left[f(G, o)\right] = \int_{\mathcal{G}_\star} f(G, o) d\mathcal{L}(G, o).$$

As a direct consequence of Proposition 10, we get:

Proposition 11. *For $G = (V, E)$, let $\{G_n\}_{n\in\mathbb{N}}$ be the sequence of 2-lifts defined in Proposition 10. Then $G_n \rightsquigarrow (T(G), o)$ where $T(G)$ is the universal cover of G with associated cover $\pi : T(G) \to G$ and o is the inverse image of a uniform vertex v of G, $o = \pi^{-1}(v)$.*

We now state the corresponding well-known result for random lifts:

→ not a seq of 2-lifts

Proposition 12. *For $G = (V, E)$, let $L_n(G)$ be a random n-lift of G. Then $L_n(G) \rightsquigarrow (T(G), o)$ a.s. where $T(G)$ is the universal cover of G with associated cover $\pi : T(G) \to G$ and o is the inverse image of a uniform vertex v of G, $o = \pi^{-1}(v)$.*

We are now ready to use the results of the above sections. The existence of the limits for the partition function, the internal energy of the monomer-dimer model is known to be continuous for the local weak convergence (in a much more general setting than here) [23, 11, 27, 1] but the explicit expressions given in the right-hand side below are new.

use of $T(G)$?

Theorem 10. *Let G be a finite graph and $T(G)$ be its universal cover. Let $(G_n)_{n\geq 1}$ be a sequence such that $G_n \rightsquigarrow (T(G), o)$. We denote by $\mathbf{x}(z)$ the vector defined by (20) in Proposition 6 where $\mathbf{y}(z) = z\mathcal{R}_G(\mathbf{y}(z))$. Then we have as $n \to \infty$, for $z > 0$,*

↳ of FM polytope

$$\lim_{n\to\infty} \frac{1}{|V_n|}\nu(G_n) = \frac{1}{v(G)}\nu^*(G), \tag{25}$$

$$\lim_{n\to\infty} \frac{1}{|V_n|}\ln P_{G_n}(z) = \frac{1}{v(G)}\Phi^B_G(\mathbf{x}(z), z), \tag{26}$$

$$\lim_{n\to\infty} \frac{1}{|V_n|}U^s_{G_n}(z) = \frac{1}{v(G)}U^B_G(\mathbf{x}(z)), \tag{27}$$

$$\lim_{n\to\infty} \frac{1}{|V_n|}\left(S_{G_n}(z) + U^\theta_G(z)\right) = \frac{1}{v(G)}S^B_G(\mathbf{x}(z)), \tag{28}$$

$$\lim_{n\to\infty} \Sigma_{G_n}(t) = \Sigma^B_G(t), \text{ for } t < \tau^*(G). \tag{29}$$

kind-of same

Proof. In [23, 11], it is shown that the root exposure probability satisfies (with our notation):

$$r_{u\to v}(z) = \frac{1}{1 + z\sum_{w\in\partial u\backslash v}\theta_{wu}r_{w\to u}(z)}.$$

Hence we can use directly results from [11] by the simple change of variable: $y_{u\to v}(z) = zr_{u\to v}(z)$. In particular Theorem 6 in [11] implies that

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{|V_n|} U_{G_n}(z) &= \frac{1}{2}\mathbb{E}_{(T(G),o)}\left[1 - \frac{1}{1 + z\sum_{\overrightarrow{e}\in\partial o}\theta_e r_{\overrightarrow{e}}(z)}\right] \\
&= \frac{1}{2}\mathbb{E}_{(T(G),o)}\left[\frac{\sum_{\overrightarrow{e}\in\partial o}\theta_e y_{\overrightarrow{e}}(z)}{1 + \sum_{\overrightarrow{e}\in\partial o}\theta_e y_{\overrightarrow{e}}(z)}\right] \\
&= \frac{1}{2}\mathbb{E}_{(T(G),o)}\left[\mathcal{D}_o(\mathbf{y}(z))\right],
\end{aligned}$$

and (27) follows from Propositions 9 and 6. (25) follows by taking the limit $z\to\infty$ as shown in Theorem 11 in [11] and (21) in Proposition 6.

We now prove (26). We start by noting that $\Phi'_G(z) = \frac{U_G^{s}(z)}{z}$ so that the convergence of $\frac{1}{|V_n|}\ln P_{G_n}(z)$ follows from (27) and Lebesgue dominated convergence theorem (see Corollary 7 in [11]). We only need to check the validity of the right-hand side expression in (26).

Note that, we have with $\theta_{\min} = \min_e \theta_e > 0$ and $\theta_{\max} = \max_e \theta_e > 0$ $\sum_e \mu^z \cdots$

$$\frac{1}{|V_n|}\frac{\ln P_{G_n}(z)}{\ln z} \geq -U^s_{G_n}(z) + \frac{|E_n|}{|V_n|}\frac{\ln\theta_{\min}}{\ln z}, \quad ?\,?\;0$$

and since the number of matching is upper bounded by $2^{|E_n|}$, we have

$$\frac{1}{|V_n|}\frac{\ln P_{G_n}(z)}{\ln z} \leq -U^s_{G_n}(z) + \frac{|E_n|}{|V_n|}\frac{\ln\theta_{\max}}{\ln z} + \frac{|E_n|\ln 2}{|V_n|\ln z}. \quad ?\;0$$

Hence, taking first the limit $n\to\infty$ and then the limit $z\to\infty$, we have

$$\lim_{z\to\infty}\lim_{n\to\infty}\frac{1}{|V_n|}\frac{\ln P_{G_n}(z)}{\ln z} = \frac{\nu^*(G)}{v(G)}.$$

Since $\frac{1}{v(G)}\Phi^B_G(\mathbf{x}(z), z) \sim \frac{\nu(G)}{v(G)}\ln z$ by Proposition 6 (note that $S^B_G(\mathbf{x})$ is bounded), we only need to check that the derivative with respect to z of the right-hand term in (26) is $\frac{U^B_G(\mathbf{x}(z))}{z}$.

Lemma 2. *In the setting of Proposition 6, we have*

$$\frac{x_e(z)(1-x_e(z))}{z} = \theta_e\left(1 - \sum_{e'\in\partial u} x_{e'}(z)\right)\left(1 - \sum_{e'\in\partial v} x_{e'}(z)\right) \tag{30}$$

Proof. Note that $\sum_{f\in\partial v} x_f(z) = \mathcal{D}_v(\mathbf{y}(z))$, so that we have by (19)

$$\begin{aligned}
\left(1 - \sum_{f\in\partial v} x_f(z)\right) &= \left(1 - \frac{\sum_{\overrightarrow{e}\in\partial v}\theta_e y_{\overrightarrow{e}}(z)}{1 + \sum_{\overrightarrow{e}\in\partial v}\theta_e y_{\overrightarrow{e}}(z)}\right) \\
&= \left(1 + \sum_{\overrightarrow{e}\in\partial v}\theta_e y_{\overrightarrow{e}}(z)\right)^{-1}
\end{aligned}$$

We have for $e = (uv) \in E$,

$$x_e(z) = \frac{\theta_e y_{u\to v}(z)}{\frac{z}{y_{v\to u}(z)} + \theta_e y_{u\to v}(z)},$$

and using the fact that $\mathbf{y}(z) = z\mathcal{R}_G(\mathbf{y}(z))$, we get

$$\begin{aligned} x_e(z) &= \frac{\theta_e y_{u\to v}(z)}{1+\sum_{w\in\partial v}\theta_{wv} y_{w\to v}(z)} = \theta_e y_{u\to v}(z)\left(1-\sum_{f\in\partial v} x_f(z)\right) \\ 1 - x_e(z) &= \frac{1+\sum_{w\in\partial u\setminus v}\theta_{wu} y_{w\to u}(z)}{1+\sum_{w\in\partial u}\theta_{wu} y_{w\to u}(z)} = \frac{z}{y_{u\to v}(z)}\left(1-\sum_{f\in\partial u} x_f(z)\right), \end{aligned}$$

and the lemma follows. ∎

Note that for $e = (uv)$, we have

$$\frac{\partial \Phi_G^B}{\partial x_e} = \ln z + \ln\left(\theta_e \frac{\left(1-\sum_{f\in\partial u} x_f\right)\left(1-\sum_{f\in\partial v} x_f\right)}{x_e(1-x_e)}\right).$$

In particular, we have $\frac{\partial \Phi_G^B}{\partial x_e}(\mathbf{x}(z)) = 0$ by Lemma 2 and then $\frac{d\Phi_G^B}{dz}(z) = -U_G^B(\mathbf{x}(z))/z$ and (26) follows. Moreover (28) follows from (26) and (27).

We now prove (29). Assume that there exists an infinite sequence of indices n such that $z_t(G_n) \geq z_t^B + \epsilon$. We denote $z_1 = z_t^B$ and $z_2 = z_t^B + \epsilon$. We have for those indices:

$$-\frac{1}{|V_n|}U_{G_n}^s(z_1) \leq -\frac{1}{|V_n|}U_{G_n}^s(z_2) \leq -\frac{1}{|V_n|}U_{G_n}^s(z_t(G_n)) = \frac{t}{2}.$$

Then by the first part of the proof, we have $-\frac{1}{|V_n|}U_{G_n}^s(z_1) \to -\frac{1}{v(G)}U_G^B(\mathbf{x}(z_1)) = \frac{t}{2}$ and $-\frac{1}{|V_n|}U_{G_n}^s(z_2) \to -\frac{1}{v(G)}U_G^B(\mathbf{x}(z_2)) > \frac{t}{2}$ by the strict monotonicity of $z \mapsto U_G^B(\mathbf{x}(z))$. Hence we obtain a contradiction. We can do a similar argument for indices such that $z_t(G_n) \leq z_t^B - \epsilon$, so that we proved that $z_t(G_n) \to z_t^B$. Then (29) follows from the continuity of the mappings $z \mapsto \mathbf{y}(z)$ and $\mathbf{x} \mapsto S_G^B(\mathbf{x})$. ∎

Proposition 13. *The function $S_G^B(\mathbf{x})$ is non-negative and concave on $FM(G)$.*

Proof. From Theorem 20 in [39], we know that the function

$$\begin{aligned} h(\mathbf{x}) &= -\sum_i x_i \ln x_i + \sum_i (1-x_i)\ln(1-x_i) \\ &\quad - \left(1-\sum_i x_i\right)\ln\left(1-\sum_i x_i\right) + \left(\sum_i x_i\right)\ln\left(\sum_i x_i\right) \end{aligned}$$

is non-negative and concave on $\Delta^k = \{\mathbf{x} \in \mathbb{R}^k,\ x_i \geq 0, \sum_{i=1}^k x_i \leq 1\}$. Hence the function

$$g(\mathbf{x}) = -\sum_i x_i \ln x_i + \sum_i (1-x_i)\ln(1-x_i) - 2\left(1-\sum_i x_i\right)\ln\left(1-\sum_i x_i\right)$$

is concave and non-negative on Δ^k since

$$g(\mathbf{x}) \quad = \quad h(\mathbf{x}) + H\left(\sum_i x_i\right),$$

where $H(p) = -p\ln p - (1-p)\ln(1-p)$ is the entropy of a Bernoulli random variable and is concave in p. The proposition follows by decomposing the sum in $S_G^B(\mathbf{x})$ vertex by vertex. $\square$

4.5 Proof of Theorem 4

Corollary 1. *Let G be a bipartite graph, then for any $z > 0$,*

$$\Phi_G(z) = \ln P_G(z) \geq \max_{\mathbf{x}\in FM(G)} \Phi_G^B(\mathbf{x}; z)$$

and for $t < t^*(G)$, we have

$$\Sigma_G(t) \geq \frac{1}{v(G)} \max_{\mathbf{x}\in FM_{tv(G)/2}(G)} S_G^B(\mathbf{x}).$$

Proof. We consider the sequence of graphs defined in Theorem 10. By Proposition 8, the sequence $\{\frac{1}{|V_n|}\Phi_{G_n}(z)\}_{n\in\mathbb{N}}$ is non-increasing in n and converges to $\frac{1}{v(G)}\Phi_G^B(\mathbf{x}(z), z)$ by Theorem 10. Hence the first statement follows from Proposition 7.

The second statement of Proposition 8 implies that the sequence $\{\Sigma_{G_n}(t)\}_{n\in\mathbb{N}}$ is non-increasing in n and converges to $\Sigma_G^B(t)$ by Theorem 10 and the last statement follows from Proposition 7. $\square$

The final step for the proof of Theorem 4 is now a standard application of probabilistic bounds on the coefficients of polynomials with only real zeros [33].

Let $k < \nu(G) = \nu$, $t = \frac{2k}{v(G)}$ and $z = z_t(G)$ such that $U_G^s(z) = -tv(G)/2 = -k$. For $i \leq \nu$, we define

$$a_i = \frac{w_i(G)z^i}{P_G(z)}.$$

By the Heilmann-Lieb theorem [23], the polynomial $A(x) = \sum_{i=0}^{\nu} a_i x^i$ has only real zeros, i.e. $(a_0, \dots, a_\nu)$ is a Pólya Frequency (PF) sequence. Note that $A(1) = 1 = \sum_i a_i$. By Proposition 1 in [33], the sequence $(a_0, \dots, a_\nu)$ is the distribution of the number S of successes in ν independent trials with probability p_i of success on the i-th trial, where the roots of $A(x)$ are given by $-(1-p_i)/p_i$ for i with $p_i > 0$. Note that $\mathbb{E}[S] = \sum_i ia_i = -U_G^s(z) = k$.

We can now use Hoeffding's inequality see Theorem 5 in [24]: let S be a random variable with probability distribution of the number of successes in ν independent trials. Assume that $\mathbb{E}[S] = \nu p \in [b, c]$. Then

$$\mathbb{P}\left(S \in [b, c]\right) \geq \sum_{i=b}^{c} \binom{\nu}{i} p^i (1-p)^{\nu-i}.$$

Hence, we have in our setting with $b = c = k$ and $p = \frac{k}{\nu}$:

$$\begin{aligned} a_k &\geq \binom{\nu}{k} p^k (1-p)^{\nu(1-p)} \\ w_k(G) &\geq b_{\nu,k}(p) \exp\left(v(G)\Sigma_G(t)\right) \\ &\geq b_{\nu,k}(k/\nu) \exp\left(\max_{\mathbf{x} \in FM_{tv(G)/2}(G)} S_G^B(\mathbf{x}) \right), \end{aligned}$$

where the last inequality follows from Corollary 1.

The case $k = \nu$ is easy. Take $t = \frac{2\nu(1-\epsilon)}{v(G)}$ with $\epsilon > 0$ and $z = z_t(G)$ so that $U_G^s(z) = -tv(G)/2 = -\nu(1-\epsilon)$. We define the sequence of a_i's as above. We now have $\mathbb{E}[S] = \nu(1-\epsilon)$. We then have $\mathbb{E}[S] = \sum_i i a_i \leq \nu a_\nu + (1 - a_\nu)(\nu - 1) = a_\nu + \nu - 1$, so that $a_\nu \geq 1 - \nu\epsilon$ and

$$w_{\nu(G)}(G) \geq (1-\nu\epsilon)\exp\left(v(G)\Sigma_G(t)\right) \geq (1-\nu\epsilon)\exp\left(\max_{\mathbf{x} \in FM_{tv(G)/2}(G)} S^B(\mathbf{x}) \right).$$

Letting $\epsilon \to 0$ concludes the proof.

4.6 Proof of Theorem 5

We start with a definition: the perfect matching corresponding to the edge e in G is called the fibre corresponding to e, which we denote by F_e.

We denote $\nu_n = \nu(L_n(G))$. If G is bipartite, we have by Theorem 3 for all $k \leq \nu_n$,

$$\ln m_k(L_n(G)) \geq \ln b_{\nu_n,k}\left(\frac{k}{\nu_n}\right) + \max_{\mathbf{x} \in M_k(L_n(G))} S^B_{L_n(G)}(\mathbf{x}).$$

It is easy to see that

$$\max_{\mathbf{x} \in M_k(L_n(G))} S^B_{L_n(G)}(\mathbf{x}) \geq n \max_{\mathbf{x} \in M_{k/n}(G)} S_G^B(\mathbf{x}),$$

since to any $\mathbf{x} \in M_{k/n}$, we can associate $\mathbf{y} \in M_k(L_n(G))$ by taking $y_{e'} = x_e$ for all $e' \in F_e$. Hence, we get

$$\frac{1}{n}\ln m_k(L_n(G)) \geq \max_{\mathbf{x} \in M_{k/n}(G)} S_G^B(\mathbf{x}) + \frac{1}{n}\ln b_{\nu_n,k}\left(\frac{k}{\nu_n}\right)$$

Taking $k = \nu_n$ and letting $n \to \infty$, we get

$$\liminf_{n\to\infty} \frac{1}{n}\ln m_{\nu_n}(L_n(G)) \geq \max_{\mathbf{x} \in FM_{\nu^*(G)}(G)} S_G^B(\mathbf{x})$$

For the upper bound, we do not need to assume that G is bipartite as we have for all $z > 0$,

$$\frac{1}{n}\ln m_{\nu_n}(L_n(G)) \leq \frac{1}{n}\ln P_{L_n(G)}(z) - \frac{\nu_n}{n}\ln z.$$

Hence letting $n \to \infty$, we get

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \ln m_{\nu_n}(L_n(G)) &\leq \Phi_G^B(\mathbf{x}(z), z) - \nu^*(G) \ln z \\ &= \sup_{\mathbf{x}\in FM(G)} \left\{ \ln z \left(\sum_e x_e - \nu^*(G) \right) + S_G^B(\mathbf{x}) \right\}\end{aligned}$$

Taking now $z \to \infty$ and noting that $\sum_e x_e - \nu^*(G) \leq 0$, we have:

$$\limsup_{n\to\infty} \frac{1}{n} \ln m_{\nu_n}(L_n(G)) \leq \sup_{\mathbf{x}\in PM(G)} S_G^B(\mathbf{x}).$$

References

[1] M. Abért, P. Csikvári, and T. Hubai. Matching measure, Benjamini-Schramm convergence and the monomer-dimer free energy. *arXiv preprint arXiv:1405.6740*, 2014.

[2] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electronic Journal of Probability*, 12:1454–1508, 2007.

[3] D. Aldous and J. M. Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 1–72. Springer, Berlin, 2004.

[4] A. Amit and N. Linial. Random graph coverings. I. General theory and graph connectivity. *Combinatorica*, 22(1):1–18, 2002.

[5] A. Amit, N. Linial, and J. Matoušek. Random lifts of graphs: independence and chromatic number. *Random Structures Algorithms*, 20(1):1–22, 2002.

[6] M. Bayati, D. Gamarnik, D. Katz, C. Nair, and P. Tetali. Simple deterministic approximation algorithms for counting matchings. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, page 127. ACM, 2007.

[7] P. Billingsley. *Convergence of probability measures.* John Wiley & Sons, Inc., New York-London-Sydney, 1968.

[8] Y. Bilu and N. Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, 2006.

[9] C. Bordenave and M. Lelarge. Resolvent of large random graphs. *Random Structures Algorithms*, 37(3):332–352, 2010.

[10] C. Bordenave, M. Lelarge, and J. Salez. The rank of diluted random graphs. *Ann. Probab.*, 39(3):1097–1121, 2011.

[11] C. Bordenave, M. Lelarge, and J. Salez. Matchings on infinite graphs. *Probability Theory and Related Fields*, pages 1–26, 2012.

[12] L. M. Brègman. Certain properties of nonnegative matrices and their permanents. *Dokl. Akad. Nauk SSSR*, 211:27–30, 1973.

[13] M. Chertkov and A. B. Yedidia. Approximating the permanent with fractional belief propagation. *The Journal of Machine Learning Research*, 14(1):2029–2066, 2013.

[14] P. Csikvári. Lower matching conjecture, and a new proof of Schrijver's and Gurvits's theorems. *arXiv preprint arXiv:1406.0766*, 2014.

[15] E. Davies, M. Jenssen, W. Perkins, and B. Roberts. Independent sets, matchings, and occupancy fractions. *arXiv preprint arXiv:1508.04675*, 2015.

[16] S. Friedland, E. Krop, and K. Markström. On the number of matchings in regular graphs. *Electron. J. Combin.*, 15(1):Research Paper 110, 28, 2008.

[17] C. D. Godsil. Matchings and walks in graphs. *J. Graph Theory*, 5(3):285–297, 1981.

[18] C. Greenhill, S. Janson, and A. Ruciński. On the number of perfect matchings in random lifts. *Combin. Probab. Comput.*, 19(5-6):791–817, 2010.

[19] L. Gurvits. Hyperbolic polynomials approach to Van der Waerden/Schrijver-Valiant like conjectures: sharper bounds, simpler proofs and algorithmic applications. In *STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 417–426. ACM, New York, 2006.

[20] L. Gurvits. Van der Waerden/Schrijver-Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: one theorem for all. *Electron. J. Combin.*, 15(1):Research Paper 66, 26, 2008. With a corrigendum.

[21] L. Gurvits. Unleashing the power of Schrijver's permanental inequality with the help of the Bethe Approximation. *arXiv preprint arXiv:1106.2844*, 2011.

[22] L. Gurvits and A. Samorodnitsky. Bounds on the permanent and some applications. In *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*, pages 90–99, Oct 2014.

[23] O. J. Heilmann and E. H. Lieb. Theory of monomer-dimer systems. *Comm. Math. Phys.*, 25:190–232, 1972.

[24] W. Hoeffding. On the distribution of the number of successes in independent trials. *Ann. Math. Statist.*, 27:713–721, 1956.

[25] M. Laurent and A. Schrijver. On Leonid Gurvits's proof for permanents. *Amer. Math. Monthly*, 117(10):903–911, 2010.

[26] F. T. Leighton. Finite common coverings of graphs. *J. Combin. Theory Ser. B*, 33(3):231–238, 1982.

[27] M. Lelarge. A new approach to the orientation of random hypergraphs. In Y. Rabani, editor, *SODA*, pages 251–264. SIAM, 2012.

[28] M. Lelarge. Loopy annealing belief propagation for vertex cover and matching: convergence, lp relaxation, correctness and Bethe approximation. *arXiv preprint arXiv:1401.7923*, 2014.

[29] M. Lelarge. Topics in random graphs, combinatorial optimization, and statistical inference. *HDR thesis*, 2015.

[30] N. Linial. Lifts of graphs (slides).

[31] N. Linial and E. Rozenman. Random lifts of graphs: perfect matchings. *Combinatorica*, 25(4):407–424, 2005.

[32] A. W. Marcus, D. A. Spielman, and N. Srivastava. Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Ann. of Math. (2)*, 182(1):307–325, 2015.

[33] J. Pitman. Probabilistic bounds on the coefficients of polynomials with only real zeros. *J. Combin. Theory Ser. A*, 77(2):279–303, 1997.

[34] N. Ruozzi. The Bethe partition function of log-supermodular graphical models. In F. Pereira, C. Burges, L. Bottou, and K. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 117–125. Curran Associates, Inc., 2012.

[35] N. Ruozzi. Beyond Log-Supermodularity: Lower Bounds and the Bethe Partition Function. *ArXiv e-prints*, Sept. 2013.

[36] J. Salez. Weighted enumeration of spanning subgraphs in locally tree-like graphs. *Random Structures & Algorithms*, 2012.

[37] A. Schrijver. Counting 1-factors in regular bipartite graphs. *J. Combin. Theory Ser. B*, 72(1):122–135, 1998.

[38] R. Smarandache and M. Haenggi. Bounding the bethe and the degree-m bethe permanents. *arXiv preprint arXiv:1503.02217*, 2015.

[39] P. O. Vontobel. The Bethe permanent of a nonnegative matrix. *IEEE Trans. Inform. Theory*, 59(3):1866–1901, 2013.

[40] P. O. Vontobel. Counting in graph covers: a combinatorial characterization of the Bethe entropy function. *IEEE Trans. Inform. Theory*, 59(9):6018–6048, 2013.

[41] I. M. Wanless. Addendum to Schrijver's work on minimum permanents. *Combinatorica*, 26(6):743–745, 2006.

[42] Y. Watanabe and M. Chertkov. Belief propagation and loop calculus for the permanent of a non-negative matrix. *Journal of Physics A: Mathematical and Theoretical*, 43(24):242002, 2010.

[43] L. Zdeborová and M. Mézard. The number of matchings in random graph. *Journal of Statistical Mechanics*, 2006(5):P05003, 2006.