Counting matchings in irregular bipartite graphs

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. It 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.

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. We provide new lower bounds for subpermanents sums.

Our proof borrows ideas from the theory of local weak convergence of graphs, statistical physics and covers of graphs.

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 \( A \) is defined as

\[
\text{per}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^n a_{i,\sigma(i)},
\]

where the summation extends over all permutation \( \sigma \) of \( \{1, \ldots, n\} \). The main result proved in \cite{Schrijver} is the following theorem:

**Theorem 1.** (Schrijver \cite{Schrijver}) 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

\[
\text{per}(\tilde{A}) \geq \prod_{i,j} (1 - a_{i,j}).
\]

It is proved in \cite{Laporte, Lovasz} that this theorem implies:

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

\[
\ln \text{per}(A) \geq \max_{x \in M_{n,n}} \sum_{i,j} (1 - x_{i,j}) \ln(1 - x_{i,j}) + x_{i,j} \ln \left( \frac{a_{i,j}}{x_{i,j}} \right),
\]

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

\*INRIA-ENS, Paris, France, email: marc.lelarge@ens.fr
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. Our first main contribution is a new proof of Theorem 2. L. Gurvits provided another proof using stable polynomials in [10, 11], see also [10].

As a consequence of Theorem 1, Schrijver shows in [22] that any $d$-regular bipartite graph with $2n$ vertices has at least

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

(3)

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 [25]. Similarly, our second main contribution shows that the right-hand term in (2) is best possible. More precisely for any $n \times n 0,1$ matrix $A$, we show that there exists a sequence of growing matrices $A_\ell$ with the same row and column sums as $A$ such that its permanent grows exponentially with its size at a rate given by the right-hand term in (2). We refer to Theorem 5 for a precise and more general statement.

In [9], 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

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

(4)

where $p = \frac{k}{n}$. An asymptotic version of this conjecture was proved using Theorem 1 in [12, 13]. A slightly stronger statement of the conjecture was proved in [8]. Our third main contribution is an extension of the results in [8] to cover irregular bipartite graphs, see Theorem 3. Our last main contribution shows that these lower bounds extend beyond counting matchings to a more general notion of permanent called $k$-th subpermanent sum, see Theorem 5. To the best of our knowledge, Theorem 5 is new at this level of generality.

The lower bound in (2) is also called the (logarithm of the) Bethe permanent [26, 7, 24]. A very recent proof of (2) using results from [24] on $k$-lifts is given in [23]. Similar ideas using lifts or covers of graphs have appeared in the literature about message passing algorithms, see [21] and references therein. We refer to [18] for more results connecting Belief Propagation with our setting.

We state our main results in the next section. Section 3 contains the technical proof. We first summarize the statistical physics results for the monomer dimer model in Section 3.1. Then, we study local recursions associated to this model in Section 3.2. The results in this section build mainly on previous work of the author [18]. In Section 3.3, we show how an idea of Csikvári [8] using 2-lift extends to our framework and connect it to the framework of local weak convergence in Section 3.4. Finally, we use probabilistic bounds on the coefficients of polynomials with only real zeros to finish the proof in Section 3.5.

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. Indeed all our results (on number of matchings and permanents) are implied by Theorem 4 below.

2.1 Lower bounds for number of matchings of a given size

We consider a graph $G = (V, E)$. We denote by $v(G)$ the cardinality of $V$: $v(G) = |V|$. We denote by the same symbol $\partial v$ 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, $B = (B_e, e \in E) \in \{0, 1\}^E$ 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 $\sum B_e$. We will also use the following notation $e \in B$ 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 $\nu(G) = \max\{\sum B_e\}$ 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 $G$. It is well-known that:

$$M(G) = \left\{ x \in \mathbb{R}^E, \ x_e \geq 0, \sum_{e \in \partial v} x_e \leq 1 \right\} \text{ if and only if } G \text{ is bipartite.} \quad (5)$$

For a given graph $G$, we denote by $m_k(G)$ the number of matchings of size $k$ in $G$ ($m_0(G) = 1$). We define by $M_k(G)$ the convex hull of incidence vectors of matchings in $G$ of size $k$. If $G$ is bipartite, we have:

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

We define the function $S^B_G : M(G) \to \mathbb{R}$ by:

$$S^B_G(x) = \sum_{e \in E} -x_e \ln x_e + (1 - x_e) \ln(1 - x_e) - \sum_{v \in V} \left(1 - \sum_{e \in \partial v} x_e\right) \ln \left(1 - \sum_{e \in \partial v} x_e\right). \quad (6)$$

**Definition 1.** Let $G$ be a graph. 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 $H$.

**Theorem 3.** For any finite bipartite graph $G$, we have for all $k \leq \nu(G)$,

$$m_k(G) \geq b_{\nu(G), k}(k/\nu(G)) \exp \left( \max_{x \in M_k(G)} S^B_G(x) \right),$$

where $b_{n, k}(p)$ is the probability for a binomial random variable $\text{Bin}(n, p)$ 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 $n \geq 1$ and for all $k \leq \nu(G)$,

$$\lim_{n \to \infty} \frac{1}{\nu(G_n)} \ln m_k(G_n) = \frac{1}{\nu(G)} \max_{x \in M_k(G)} S^B_G(x).$$
Note that the function $S_B^G$ is concave on $M(G)$ by Proposition 11.

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 $x^* = k \frac{t}{m}$ for all $e \in E$ so that $x^* \in M_k(G)$ and we have

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

with $p = \frac{k}{n}$. We see that we recover the first statement in Theorem 1.5 of [8]. In particular, for $k = n$, i.e. $p = 1$, we recover (3) and for $k < n$, we slightly improve upon (4).

We will also prove the following bound on the total number of matchings:

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

$$\sum_{k=0}^{\nu(G)} m_k(G) \geq \exp \left( \max_{x \in M(G)} S_B^G(x) \right).$$

### 2.2 Lower bounds for permanents

In this section, we extend previous results to weighted graphs. We state our results in terms 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

$$\text{per}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{\sigma(i),i}.$$ 

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

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

We restate here Theorem 2.

**Theorem 4.** Let $A$ be a non-negative $n \times n$ matrix. Then, we have

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

with the convention $\ln 0 = 1$.

Note that if $\text{per}(A) = 0$, then the lower bound above is equal to $-\infty$. Indeed if $\text{per}(A) = 0$, 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.
For $1 \le k \le n$, let $\text{per}_k(A)$ be the sum of permanents of all $k \times k$ minors in $A$. $\text{per}_k(A)$ 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 $k$:

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

We also define the set of substochastic matrices:

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

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

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

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 (6) 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 $S^B_G(\mathbf{x})$.

**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 $b_{i,j} = b_{i+n,j+n} = 0$.

**Theorem 5.** Let $A$ be a non-negative $n \times n$ matrix. Let $\nu(A) = \max\{k, \text{per}_k(A) > 0\}$. For all $k \le \nu(A)$, we have

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

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 $A_0 = A$, $A_\ell$ is a 2-lift of $A_{\ell-1}$ for $\ell \ge 1$ and for all $k \le \nu(A)$,

$$\lim_{\ell \to \infty} \frac{1}{2\ell n} \ln \text{per}_k(A_\ell) = \frac{1}{n} \max_{x \in M_{n,k}} S^B(A, \mathbf{x}).$$

Note that $\mathbf{x} \mapsto S^B(A, \mathbf{x})$ is concave on $M_{n,\le}$. If $k = n$, we recover Theorem 2 which is equivalent to Theorem 1. Note that if $\nu(A) < n$, then $\text{per}(A) = 0$ and as noted above the lower bound (2) in Theorem 2 is equal to $-\infty$. Also, results presented in Section 2.1 follow by taking for the matrix $A$, the incidence matrix of the graph $G$. 

5
3 Proof

3.1 Statistical physics

To ease the notation, we will consider a setting with a weighted bipartite graph $G = (V, E)$ with positive weights on edges $\{\theta_e\}_{e \in E}$. Taking $\theta_e = 1$ 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 $z > 0$:

$$
\mu_z^G(B) = \frac{z^{\sum_e B_e} \prod_{e \in B} \theta_e}{P_G(z)},
$$

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

$$
w_k(G) = \sum_{B, \sum_e B_e = k} \prod_{e \in B} \theta_e,
$$

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 $\mu_z^G$ converges to the measure:

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

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 [14].

We define the following functions:

$$
U^z_\theta G(z) = -\sum_{e \in E} \mu^z_e(B_e = 1),
$$

$$
U^\theta G(z) = \sum_{e \in E} \mu^z_e(B_e = 1) \ln \theta_e,
$$

$$
S_G(z) = -\sum_B \mu^z_e(B) \ln \mu^z_e(B).
$$

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

$$
\Phi_G(z) = -U^z_\theta G(z) \ln z + U^\theta G(z) + S_G(z).
$$

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 $U^z_\theta G(z)$ is negative, equals to minus the average size of a matching sampled from $\mu^z_e$. 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:

\[ \Phi_G(z) = \ln P_G(z) \text{ and, } \Phi_G'(z) = \frac{-U^*_G(z)}{z}. \]

**Lemma 1.** The function \( U^*_G(z) \) is strictly decreasing and mapping \([0, \infty)\) to \((-v(G), 0]\).

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

\[
-z(U^*_G)'(z) = \frac{\sum_k k^2 w_k(G)z^k}{P_G(z)} - \left( \frac{\sum_k k w_k(G)z^k}{P_G(z)} \right)^2
= \sum_k \left( k - \frac{\sum \ell w_\ell(G)z^\ell}{P_G(z)} \right) \frac{w_k(G)z^k}{P_G(z)} > 0.
\]

\( \Box \)

We define \( t^* = t^*(G) = 2v(G)/v(G) \) which is the maximum fraction of nodes covered by a matching in \( G \). Note that \( t^*(G) \leq 1 \) and \( t^*(G) = 1 \) if and only if the graph \( G \) has a perfect matching. For \( t \in [0, t^*) \), we define \( z_t(G) \in [0, \infty) \) as the unique root to \( U^*_G(z_t(G)) = -tv(G)/2 \). Note that \( t \mapsto z_t(G) \) is an increasing function which maps \([0, t^*)\) to \([0, \infty)\). The function \( \Sigma_G(t) \) is then defined for \( t \in [0, t^*) \) by:

\[
\Sigma_G(t) = \frac{S_G(z_t(G)) + U^*_G(z_t(G))}{v(G)}, \tag{10}
\]

and \( \Sigma_G(t) = -\infty \) for \( t > t^* \).

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

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

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

Since \( U^*_G(z) = -zP'_G(z)/P_G(z) \) and \( U^*_G(z_t) = -tv(G)/2 \), 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. \( \Box \)

**Proposition 3.** If for some graphs \( G_1 \) and \( G_2 \), we have for every \( z \geq 0 \),

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

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

\[
-\frac{1}{2} (\ln z_t(G_1) - \ln z_t(G_2))
\]

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

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

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

### 3.2 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 \( \vec{E} \) of directed edges of \( G \) comprising two directed edges \( u \to v \) and \( v \to u \) for each undirected edge \( (uv) \in E \). For \( \vec{v} \in \vec{E} \), we denote by \( -\vec{v} \) the edge with opposite direction. With a slight abuse of notation, we denote by \( \partial v \) the set of incident edges to \( v \in V \) directed towards \( v \). We also denote by \( \partial v \setminus u \) 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 \( \mathcal{R}_G : (0, \infty)^{\vec{E}} \to (0, \infty)^{\vec{E}} \) by \( \mathcal{R}_G(a) = b \) with

\[
b_{u \to v} = \frac{1}{1 + \sum_{w \in \partial u \setminus v} \theta_{wu} a_{w \to u}},
\]

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

**Proposition 4.** Let \( G \) be a finite graph. For any \( z > 0 \), the fixed point equation \( y(z) = z\mathcal{R}_G(y(z)) \) has a unique attractive solution \( y(z) \in (0, +\infty)^{\vec{E}} \). The function \( z \mapsto y(z) \) is increasing and the function \( z \mapsto \frac{y(z)}{z} \) is decreasing for \( z > 0 \).

Comparisons between vectors are always componentwise. 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 \( [18] \) and the proof extends to this setting. \( \square \)
We define for all $v \in V$, the following function of the vector $(y_{\partial^+}, z_\partial) \in \partial v$,

$$D_v(y) = \sum_{\partial^+ \in \partial v} \frac{\theta_e y_{\partial^+} R_{-\partial^+}(y)}{1 + \theta_e y_{\partial^+} R_{-\partial^+}(y)}$$

(11)

$$= \sum_{\partial^+ \in \partial v} \frac{\theta_e y_{\partial^+}}{1 + \sum_{\partial^+ \in \partial v} \theta_e y_{\partial^+}} \quad \quad (12)$$

Clearly from (12), we see that $D_v$ is an increasing function of $y$ and the proposition below follows directly from the monotonicity of $y(z)$ proved in Proposition 4.

**Proposition 5.** Let $G = (V, E)$ be a finite graph and $y(z)$ be the solution to $y(z) = zR_G(y(z))$. For any $v \in V$, the mapping $z \mapsto D_v(y(z))$ is increasing and $D_v(y(z)) = \sum_{e \in \partial v} x_e(z)$, where

$$x_e(z) = \frac{\theta_e y_{\partial^+}(z) y_{\partial^-}(z)}{z + \theta_e y_{\partial^+}(z) y_{\partial^-}(z)} = (0, 1).$$

(13)

We denote by $x(z) = (x_e(z), e \in E)$ the vector defined by (13). If $G$ is a bipartite graph, then we have:

$$\lim_{z \to \infty} \sum_{v \in V} D_v(y(z)) = 2\nu(G).$$

(14)

**Proof.** The only non-trivial statement in the above proposition is the value of the limit in (14). In the case $\theta_e = 1$, it follows from Theorem 1 in [18] and the proof carries over to the case $\theta_e > 0$.

For a finite bipartite graph $G = (V, E)$ with weights on edges $\{\theta_e\}_{e \in E}$, we define for $x \in M(G)$ defined by [10] and $z > 0$,

$$U_B^G(x) = -\sum_{e \in E} x_e,$n$$

$$S_B^G(x) = \sum_{e \in E} x_e \ln \frac{\theta_e}{x_e} + (1 - x_e) \ln(1 - x_e) - \sum_{e \in \partial v} \left(1 - \sum_{e \in \partial v} x_e\right) \ln \left(1 - \sum_{e \in \partial v} x_e\right),$$

$$\Phi_B^G(x, z) = -U_B^G(x) \ln z + S_B^G(x).$$

We denote by $x(z)$ the vector defined by (13) in Proposition 5 where $y(z) = zR_G(y(z))$. Note that

$$U_B^G(x(z)) = -\frac{1}{2} \sum_{v \in V} D_v(y(z)),$n$$

so that by Proposition 5, the mapping $z \mapsto U_B^G(x(z))$ is decreasing from $[0, \infty)$ to $(-\nu(G), 0]$ provided $G$ is bipartite. Thus, we can define $z^*_B$ as the unique solution in $[0, \infty)$ to

$$U_B^G(x(z^*_B)) = -\frac{\nu(G)}{2}$$

for $t < t^*(G) = \frac{2\nu(G)}{v(G)}$.

Similarly as in (13), we define

$$\Sigma_B^G(t) = \frac{S_B^G(x(z^*_B))}{v(G)}$$

for $t < t^*(G)$.
Proposition 6. Recall that $x(z) \in \mathbb{R}^E$ is defined by (13). If $G$ is bipartite, then we have for any $z > 0$,

$$\sup_{x \in M(G)} \Phi^B_G(x; z) = \Phi^B_G(x(z); z),$$

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

$$\Sigma^B_G(t) = \frac{1}{v(G)} \max_{x \in M_t(G)} S^B_G(x),$$

where for $s \geq 0$, $M_s(G) = \{ x \in \mathbb{R}^E, x_e \geq 0, \sum_{e \in \partial u} x_e \leq 1, \sum_{e \in E} x_e = sv(G) \}$ and where the maximum taken over an empty set is equal to $-\infty$.

Proof. The first statement is proved in [18] for the case where $\theta_e = 1$ but extends easily to the current framework. For the second statement, note that for any $x \in M_t(G)$ with $t < t^*(G)$, we have

$$\Phi^B_G(x, z) = \frac{tv(G)}{2} \ln z^B + S^B_G(x) \leq \frac{tv(G)}{2} \ln z^B + S^B_G(x(z))).$$

By definition, we have $x(z) \in M_t(G)$, so that $\max_{x \in M_t(G)} S^B_G(x) = S^B_G(x(z)))$.

We now extend Proposition 4 to infinite trees:

Theorem 6. Let $T = (V, E)$ be a (possibly infinite) tree with bounded degree. For each $z > 0$, there exists a unique solution in $(0, \infty)$ to the fixed point equation $y(z) = zR_T(y(z))$, i.e. such that

$$y_{u \rightarrow v}(z) = \frac{z}{1 + \sum_{w \in \partial u \setminus v} \theta_{wu} y_{w \rightarrow u}(z)} . \quad (15)$$

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

To prove the uniqueness, we follow the approach in [4]. First, we define the change of variable: $h_{u \rightarrow v} = -\ln \frac{y_{u \rightarrow v}(z)}{z}$ so that (15) becomes:

$$h_{u \rightarrow v} = \ln \left( 1 + z \sum_{w \in \partial u \setminus v} \theta_{wu} e^{-h_{w \rightarrow u}} \right) . \quad (16)$$

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

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

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

10
Iterating the recursion (16), we can rewrite it using such a function \( f \) so that uniqueness would be implied if we show that \( f \) is contracting.

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

\[
|f(h) - f(h')| = |\nabla f(h_\alpha)(h-h')| \leq \|\nabla f(h_\alpha)\|_1 \|h-h'\|_\infty.
\]

A simple computation shows that:

\[
\|\nabla f(h)\|_1 = \frac{z \sum_{i=1}^k \theta_i (1+z \sum_{j=1}^{k_i} \theta_j e^{-h_j^i})}{1 + z \sum_{i=1}^k \theta_i A_i}.
\]

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

\[
\|\nabla f(h)\|_1 = \frac{z \sum_{i=1}^k \theta_i (A_i - A_i^2)}{1 + z \sum_{i=1}^k \theta_i A_i} = 1 - \frac{1 + z \sum_{i=1}^k \theta_i A_i^2}{1 + z \sum_{i=1}^k \theta_i A_i}.
\]

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

\[
\|\nabla f(h)\|_1 \leq 1 - \frac{2}{\sqrt{1+z\Theta} + 1}.
\]

\[\blacksquare\]

### 3.3 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 \( \pi : G' \rightarrow G \) is a covering map if for each \( v' \in V(G') \), \( \pi \) 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 \( \theta_e \) 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 7.** Let \( G \) be a bipartite graph and \( H \) be a 2-lift of \( G \). Then \( P_G(z)^2 \geq P_H(z) \) for \( z > 0 \) and \( \Sigma_G(t)^2 \geq \Sigma_H(t) \) for \( t \in [0, 1] \).

**Proof.** The proof follows from an argument of Csikvári [8]. 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 \( R = \)
Proof. We refer to the surveys \cite{3,2}.

We have \( \pi: \tilde{T}(G) \to \tilde{T}(G) \) so that \( \tilde{P}_T(z)^2 \geq \tilde{P}_H(z) \) for \( z > 0 \) and the second statement follows from Proposition 8. \( \square \)

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 at \( v \), 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 \( R_{T(G)} \) and \( R_G \) and since there is a unique fixed point for both \( zR_{T(G)} \) and \( zR_G \), the proposition below follows:

**Proposition 8.** Let \( G \) be a finite graph and \( T(G) \) be its universal cover and associated cover \( \pi: T(G) \to G \). By Propositions 5 and 4 we can define:

\[
\tilde{y}(z) = zR_{T(G)}(y(z)), \quad \text{and,} \quad y(z) = zR_{G}(y(z)).
\]

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

### 3.4 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 \cite{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 \( \mathcal{G}_* \) 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 \( \gamma: V \to V' \) 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

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

turns \( \mathcal{G}_* \) into a complete separable metric space, see \cite{2}.

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 \( (G, o) \in \mathcal{G}_* \).
Proposition 9. 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 $\mathcal{G}_\ast$.

**Proof.** The proof follows from an argument of Nathan Linial [19], see also [8].

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

**Local weak limits.** Let $\mathcal{P}(\mathcal{G}_\ast)$ denote the set of Borel probability measures on $\mathcal{G}_\ast$, equipped with the usual topology of weak convergence (see e.g. [5]). Given a finite graph $G = (V, E)$, we construct a random element of $\mathcal{G}_\ast$ 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 $\mathcal{L} \in \mathcal{P}(\mathcal{G}_\ast)$, we call $\mathcal{L}$ the local weak limit of $\{G_n\}_{n \geq 1}$. If $(G, o)$ denotes a random element of $\mathcal{G}_\ast$ with law $\mathcal{L}$, we shall use the following slightly abusive notation: $G_n \leadsto (G, o)$ and for $f : \mathcal{G}_\ast \to \mathbb{R}$:

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

As a direct consequence of Proposition 9, we get:

**Proposition 10.** For $G = (V, E)$, let $\{G_n\}_{n \in \mathbb{N}}$ be the sequence of 2-lifts defined in Proposition 9. Then $G_n \leadsto (T(G), o)$ 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 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) [14] but the explicit expressions given in the right-hand side below are new.

**Theorem 7.** Let $G$ be a finite bipartite graph and $T(G)$ be its universal cover. Let $(G_n)_{n \geq 1}$ be a sequence of 2-lifts as defined in Proposition 9 such that $G_n \leadsto (T(G), o)$. We denote by $x(z)$ the vector defined by (13) in Proposition 5 where $y(z) = zR_G(y(z))$. Then we have as $n \to \infty$,
In particular Theorem 6 in [6] implies that

$$
l_{\text{exposure}}(\theta, z) = \frac{1}{\nu(G)} \varPhi(z), \quad (17)$$

$$
\nu(G) \varPhi(z) \quad (18)
$$

$$
\nu(G) \varPhi(z) \quad (19)
$$

$$
\nu(G) \varPhi(z) \quad (20)
$$

Proof. In [14] [6], it is shown that the root exposure probability satisfies (with our notation):

$$
r_{u,v}(z) = \frac{1}{1 + z \sum_{w \in \partial u \setminus v} \theta_w r_{w,u}(z)}. \quad (18)
$$

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

$$
\lim_{n \to \infty} \frac{1}{|V_n|} P_{G_n}(z) = \frac{1}{\nu(G)} \varPhi(z), \quad (17)
$$

$$
\lim_{n \to \infty} \frac{1}{|V_n|} U_{G_n}(z) = \frac{1}{\nu(G)} \varPhi(z), \quad (18)
$$

$$
\lim_{n \to \infty} \frac{1}{|V_n|} \left( S_{G_n}(z) + U_{G_n}(z) \right) = \frac{1}{\nu(G)} \varPhi(z), \quad (19)
$$

$$
\lim_{n \to \infty} S_{G_n}(t) = \varPhi(t), \quad (20)
$$

and (18) follows from Propositions 8 and 5.

We now prove (17). We start by noting that $\varPhi(z) = \varUpsilon(z)$ so that the convergence of

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

Since \( \frac{1}{\nu(G)} \varPhi(z) \sim \frac{\nu(G)}{\nu(G)} \ln z \) by Proposition 3 (note that $S^B_G(x)$ is bounded), we only need to check that the derivative with respect to $z$ of the right-hand term in (17) is $\frac{\nu(G)}{z} \varPhi(z)$.

Lemma 2. In the setting of Proposition 3, we have

$$
x_e(z) \frac{1 - x_e(z)}{z} = \theta_e \left( 1 - \sum_{e' \in \partial u} x_{e'}(z) \right) \left( 1 - \sum_{e' \in \partial v} x_{e'}(z) \right) \quad (21)
$$

Proof. Note that $\sum_{f \in \partial v} x_f(z) = D_{\partial v}(y(z))$, so that we have by (12)

$$
\left( 1 - \sum_{f \in \partial v} x_f(z) \right) = \left( 1 - \sum_{e \in \partial v} \theta_e y_{e}(z) \right) \left( 1 + \sum_{e' \in \partial v} \theta_{e'} y_{e'}(z) \right) \quad (22)
$$

$$
\left( 1 - \sum_{f \in \partial v} x_f(z) \right) = \left( 1 + \sum_{e \in \partial v} \theta_e y_{e}(z) \right)^{-1} \quad (23)
$$
We have for $e = (uv) \in E$,

$$x_e(z) = \frac{\theta_ey_{uv}(z)}{y_{uv}(z) + \theta_ey_{uv}(z)},$$

and using the fact that $y(z) = zR_G(y(z))$, we get

$$x_e(z) = \frac{\theta_ey_{uv}(z)}{1 + \sum_{w \in \partial u} \theta_{vw}y_{wu}(z)} = \theta_ey_{uv}(z) \left(1 - \sum_{f \in \partial v} x_f(z)\right),$$

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

and the lemma follows.

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

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

In particular, we have $\frac{\partial \Phi^B}{\partial x_e}(x(z)) = 0$ by Lemma 2 and then $\frac{\partial \Phi^B}{\partial x_e}(z) = -U^B_G(x(z))/z$ and (17) follows. Moreover (19) follows from (17) and (18).

We now prove (20). Assume that there exists an infinite sequence of indices $n$ such that $z_{t(G_n)} \geq z^B_t + \epsilon$. We denote $z_1 = z^B_t$ and $z_2 = z^B_t + \epsilon$. We have for those indices:

$$-\frac{1}{|V_n|}U^a_G(z_1) \leq -\frac{1}{|V_n|}U^a_G(z_2) \leq -\frac{1}{|V_n|}U^a_G(z_{t(G_n)}) = \frac{t}{2}. $$

Then by the first part of the proof, we have $-\frac{1}{|V_n|}U^a_G(z_1) \to -\frac{1}{|V_n|}U^a_G(x(z_1)) = \frac{t}{2}$ and $-\frac{1}{|V_n|}U^a_G(z_2) \to -\frac{1}{|V_n|}U^B_G(x(z_2)) > \frac{t}{2}$ by the strict monotonicity of $z \mapsto U^B_G(x(z))$. Hence we obtain a contradiction. We can do a similar argument for indices such that $z_{t(G_n)} \leq z^B_t - \epsilon$, so that we proved that $z_{t(G_n)} \to z^B_t$. Then (20) follows from the continuity of the mappings $z \mapsto y(z)$ and $x \mapsto S^B_G(x)$.

**Proposition 11.** The function $S^B_G(x)$ is non-negative and concave on $M(G)$.

**Proof.** From Theorem 20 in [21], we know that the function

$$h(x) = -\sum_i x_i \ln x_i + \sum_i (1 - x_i) \ln(1 - x_i)$$

$$- \left(1 - \sum_i x_i\right) \ln \left(1 - \sum_i x_i\right) + \left(\sum_i x_i\right) \ln \left(\sum_i x_i\right)$$

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

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

is non-negative and concave on $\Delta^k$. Therefore, $S^B_G(x)$ is non-negative and concave on $M(G)$. So Proposition 11 is proved.

15
is concave and non-negative on $\Delta^k$ since

$$g(x) = h(x) + H\left(\sum_i x_i\right),$$

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 $S^B_G(x)$ vertex by vertex. \qed

3.5 Proof of Theorem 5

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

$$\Phi_G(z) = \ln P_G(z) \geq \max_{x \in M(G)} \Phi^B_G(x; z)$$

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

$$\Sigma_G(t) \geq \frac{1}{v(G)} \max_{x \in M_i(G)} S^B_G(x),$$

where $M_i(G) = \{ x \in \mathbb{R}^E, x_e \geq 0, \sum_{e \in \partial v} x_e \leq 1, \sum_{e \in E} x_e = \frac{sv(G)}{2}\}$.

**Proof.** We consider the sequence of graphs defined in Theorem 7. By Proposition 7, the sequence $\{1, |V_n|, \Phi_G(z)\} \in \mathbb{N}$ is non-increasing in $n$ and converges to $\Sigma_B G(x; z)$ by Theorem 7. Hence the first statement follows from Proposition 6.

The second statement of Proposition 7 implies that the sequence $\{\Sigma_G(t)\} \in \mathbb{N}$ is non-increasing in $n$ and converges to $\Sigma_B G(t)$ by Theorem 7 and the last statement follows from Proposition 6. \qed

Note that taking $\theta_e = 1$ for all edges and $z = 1$, we get $P_G(1) = \sum_k m_k(G)$ and $\Phi^B_G(x, 1) = S^B_G(x)$ so that Proposition 1 follows directly from the first statement of Corollary 1.

The final step for the proof of Theorem 5 is now a standard application of probabilistic bounds on the coefficients of polynomials with only real zeros [20].

Let $k < \nu(G) = \nu, t = \frac{2k}{v(G)}$ and $z = z_1(G)$ such that $U^G(z) = -tv(G)/2 = -k$. For $i \leq \nu$, we define

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

By the Heilmann–Lieb theorem [14], the polynomial $A(x) = \sum_{i=0}^\nu a_i x^i$ has only real zeros, i.e. $(a_0, \ldots, a_\nu)$ is a Pólya Frequency (PF) sequence. Note that $A(1) = 1 = \sum_i a_i$. By Proposition 1 in [20], the sequence $(a_0, \ldots, a_\nu)$ is the distribution of the number $S$ of successes in $\nu$ 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 $\mathbb{E}[S] = \sum_i ia_i = -U^G(z) = k$. 16
We can now use Hoeffding’s inequality see Theorem 5 in [15]: let $S$ be a random variable with probability distribution of the number of successes in $\nu$ independent trials. Assume that $E[S] = \nu p \in [b, c]$. Then
\[
P(S \in [b, c]) \geq \sum_{i=b}^{c} \binom{\nu}{i} p^i (1-p)^{\nu-i}.
\]

Hence, we have in our setting with $b = c = k$ and $p = \frac{k}{\nu}$:
\[
a_k \geq \binom{\nu}{k} p^k (1-p)^{\nu-\nu} \exp \left(\max_{x \in M_t(G)} S^B(x)\right),
\]
where the last inequality follows from Corollary [11].

The case $k = \nu$ is even simpler. Take $t = \frac{2(1-\epsilon)}{|V|}$ with $\epsilon > 0$ and $z = z_t(G)$ so that $U_G^*(z) = -t|V|/2 = -\nu(1-\epsilon)$. We define the sequence of $a_i$’s as above. We now have $E[S] = \nu(1-\epsilon)$. We then have $\mu = \sum_i i a_i \leq \nu a_{\nu} + (1 - a_{\nu})(\nu - 1) = a_{\nu} + \nu - 1$, so that $a_{\nu} \geq 1 - \nu \epsilon$ and
\[
w_{\nu(G)}(G) \geq (1 - \nu \epsilon) \exp (v(G)\Sigma_G(t)) \geq (1 - \nu \epsilon) \exp \left(\max_{x \in M_t(G)} S^B(x)\right).
\]

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

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] 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.

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

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

[7] M. Chertkov and A. B. Yedidia. Approximating the permanent with fractional belief propagation. The Journal of Machine Learning Research, 14(1):2029–2066, 2013.
[8] P. Csikvári. Lower matching conjecture, and a new proof of schrijver’s and gurvits’s theo-
rems. arXiv preprint arXiv:1406.0766, 2014.

[9] 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.

[10] 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.

[11] 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.

[12] L. Gurvits. Unleashing the power of Schrijver’s permanental inequality with the help of the 
Bethe Approximation. arXiv preprint arXiv:1106.2844 2011.

[13] 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.

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

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

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

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

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

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

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

[21] 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.

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

[23] R. Smarandache and M. Haenggi. Bounding the bethe and the degree-m bethe permanents. 
arXiv preprint arXiv:1503.02217, 2015.
[24] P. O. Vontobel. The Bethe permanent of a nonnegative matrix. *IEEE Trans. Inform. Theory*, 59(3):1866–1901, 2013.

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

[26] 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.