Monomer-Dimer Models

1 Basic Models

Firstly, we put some notations here. For a graph $G = (V, E)$ with $|V| = n$, assume that $V = \{v_1, \dots, v_n\}$. For $v \in V$, let $N_G(v)$ denote the sets of neighbors of v, and $\Delta_G(v) := |N_G(v)|$ denote the degree of v in G. The maximum degree of G is defined by $\Delta := \max_{v \in V} \Delta_G(v)$. For $v \in V$, we use the notation $G \setminus \{v\}$ to denote $G[V \setminus \{v\}]$ and $G_k = G \setminus \{v_1, \dots, v_{k-1}\}$ with convention $G_0 = G$.

A matching is a subset of edges such that every pair of two edges share no endpoints. Given a graph $G = (V, E)$ and a fugacity $\lambda > 0$, let Ω be the collection of matchings on G. The Gibbs distribution of the monomer-dimer model on G at fugacity λ is the probability distribution $\mu = \mu_{G,\lambda}$ defined as

$$\mu_{G,\lambda}(S) = \frac{\lambda^{|S|}}{Z_G(\lambda)}, \quad \forall S \in \Omega$$

where $Z_G(\lambda) = \sum_{S \in \Omega} \lambda^{|S|}$ is the normalizing factor named *partition function*. To be consistent with the language of the two-spin systems, we also use $\sigma = \sigma_S \in \{-, +\}^E$ to denote a subset S of edges E, where for every $e \in E$

$$\sigma_S(e) = \begin{cases} + & e \in S \\ - & e \notin S \end{cases}.$$

Our major interests lie in the following three aspects:

1. How to efficiently estimate the partition function $Z_G(\lambda)$ (approximate counting problem).
2. How to efficiently sample a matching from $\mu_{G,\lambda}$ (randomly sampling problem).
3. The zero-free region of $Z_G(\lambda)$ with respect to $\lambda \in \mathbb{C}$ (zero-freeness of the partition function).

2 Deterministically Counting Matchings

In this section we show how to estimate $Z_G(\lambda)$ efficiently.

Theorem 2.1 (Theorem 2.1 in [BGK+07]). *For every $\varepsilon \in (0, 1)$, there exists a deterministic ε-algorithm which provides an* FPTAS *for computing $Z_G(\lambda)$ of a monomer-dimer model on $G = (V, E)$ with $|V| = n$ and constant maximum degree Δ at constant fugacity $\lambda > 0$, running time with $O\left((n/\varepsilon)^{\kappa \log \Delta + 1}\right)$ where $\kappa = -\frac{2}{\log\left(1 - \frac{2}{\sqrt{1+\lambda\Delta}+1}\right)}$.*

To compute $Z_G(\lambda)$, we apply the method of the correlation decay. The following identity is of great significant.

Proposition 2.2 (Proposition 2.2 in [BGK+07]). *Under the settings and notations described above, the following identity holds*

$$Z_G(\lambda) = \frac{1}{\prod_{1 \le k \le |V|} \mathbf{Pr}_{M \sim \mu_{G_k,\lambda}}[v_k \notin M]}.$$

To estimate $Z_G(\lambda)$, we turn our sight to the marginal probability $\mathbf{Pr}_{M\sim\mu_{G,\lambda}}[v \notin M]$. The following recursion is simple but meaningful.

Proposition 2.3 (Proposition 3.1 in [BGK+07]). *For every vertex $v \in V$, it holds that*

$$\mathbf{Pr}_{M\sim\mu_{G,\lambda}}[v \notin M] = \frac{1}{1+\lambda\sum_{u\in N_G(v)}\mathbf{Pr}_{M\sim\mu_{G\setminus\{v\},\lambda}}[u \notin M]}.$$

Proof. For every $v \in V$, by definition, the following identity holds:

$$Z_G(\lambda) = Z_{G\setminus\{v\}}(\lambda) + \lambda \sum_{u\in N_G(v)} Z_{G\setminus\{u,v\}}(\lambda).$$

Then,

$$\begin{aligned}\mathbf{Pr}_{M\sim\mu_{G,\lambda}}[v \notin M] &= \frac{Z_{G\setminus\{v\}}(\lambda)}{Z_G(\lambda)}\\ &= \frac{Z_{G\setminus\{v\}}(\lambda)}{Z_{G\setminus\{v\}}(\lambda)+\lambda\sum_{u\in N_G(v)}Z_{G\setminus\{u,v\}}(\lambda)}\\ &= \frac{1}{1+\lambda\sum_{u\in N_G(v)}\frac{Z_{G\setminus\{u,v\}}(\lambda)}{Z_{G\setminus\{v\}}(\lambda)}}\\ &= \frac{1}{1+\lambda\sum_{u\in N_G(v)}\mathbf{Pr}_{M\sim\mu_{G\setminus\{v\},\lambda}}[u \notin M]}.\end{aligned}$$

□

For any subgraph $H \subseteq G$ of the graph G, every vertex $v \in V$ and non-negative integer $t \in \mathbb{N}$, we introduce the quantity $\Phi_H(v,t)$ as:

$$\Phi_H(v,t) = \begin{cases} 0 & t = 0 \\ \frac{1}{1+\lambda\sum_{u\in N_H(v)}\Phi_{H\setminus\{v\}}(u,t-1)} & t \geq 1 \end{cases}.$$

It is easy to observe that for every subgraph $H \subseteq G$, every vertex $v \in V$ and $t \in \mathbb{N}$,

$$\frac{1}{1+\lambda\Delta} \leq \Phi_H(v,t) \leq 1.$$

Theorem 2.4 (Correlation Decay, Theorem 3.2 in [BGK+07]). *For every vertex $v \in V$ and every positive even integer $t \in \mathbb{N}$, it holds that*

$$\left|\mathbf{Pr}_{M\sim\mu_{G,\lambda}}[v \notin M] - \log\Phi_G(v,t)\right| \leq \left(1 - \frac{2}{\sqrt{1+\lambda\Delta}+1}\right)^{t/2}\log(1+\lambda\Delta).$$

Proof. For $v \in G$, let $N_G(v) = \{u_1, \ldots, u_m\}$ and for $i \in [m]$, let $N_{G\setminus\{v\}}(u_i) = \left\{w_1^{(i)}, \ldots, w_{m_i}^{(i)}\right\}$. Furthermore, we use the following notations:

$$\begin{aligned}&x = \log\mathbf{Pr}_{M\sim\mu_{G,\lambda}}[v \notin M], x_i = \log\mathbf{Pr}_{M\sim\mu_{G\setminus\{v\},\lambda}}[u_i \notin M], x_j^{(i)} = \log\mathbf{Pr}_{M\sim\mu_{G\{v,u_i\},\lambda}}\left[w_j^{(i)} \notin M\right]\\ &y = \log\Phi_G(v,t), y_i = \log\Phi_{G\setminus\{v\}}(u_i,t-1), y_j^{(i)} = \log\Phi_{G\setminus\{v,u_i\}}(w_j^{(i)},t-2)\end{aligned}$$

for every $i = 1, \ldots, m$ and $j = 1, \ldots, m_i$.

Let $M = \sum_{i=1}^{m} m_i$, and $\vec{z} = (z_1^{(1)}, \dots, z_{m_1}^{(1)}, \dots, z_1^{(m)}, \dots, z_{m_m}^{(m)})$. Define the function $f : [0,1]^M \to [0,1]$ as

$$f(\vec{z}) = \log\left(1 + \lambda \sum_{i=1}^{m} \frac{1}{1 + \lambda \sum_{j=1}^{m_i} e^{z_j^{(i)}}}\right).$$

Then we know $x = -f(\vec{x})$ and $y = -f(\vec{y})$.

Now we consider the function $g(\alpha) = f(\alpha\vec{x} + (1-\alpha)\vec{y})$ for $\alpha \in [0,1]$. By the mean-value theorem and Hölder's inequality

$$|x - y| = \left|\nabla f(\vec{z}_\alpha)^\top (\vec{x} - \vec{y})\right| \le \|\nabla f(\vec{z}_\alpha)\|_1 \cdot \|\vec{x} - \vec{y}\|_\infty.$$

By calculation,

$$\|\nabla f(z)\|_1 = \frac{1}{1 + \lambda \sum_{i=1}^{m} \frac{1}{1+\lambda \sum_{j=1}^{m_i} e^{z_j^{(i)}}}} \sum_{i=1}^{m} \lambda \left(\frac{1}{1 + \lambda \sum_{j=1}^{m_i} e^{z_j^{(i)}}}\right)^2 \lambda \sum_{j=1}^{m_i} e^{z_j^{(i)}}.$$

For convenience, let $A_i = 1 + \lambda \sum_{j=1}^{m_i} e^{z_j^{(i)}}$. Then we show that

$$\|\nabla f(z)\|_1 = \frac{\sum_{i=1}^{m} \frac{\lambda(A_i - 1)}{A_i^2}}{1 + \lambda \sum_{i=1}^{m} \frac{1}{A_i}} = 1 - \frac{1 + \lambda \sum_{i=1}^{m} \frac{1}{A_i^2}}{1 + \lambda \sum_{i=1}^{m} \frac{1}{A_i}}.$$

The maximal value of $\|\nabla f(z)\|_1$ takes at the point for every $1/A_i = \frac{\sqrt{1+\lambda m} - 1}{\lambda m}$. Then

$$\|\nabla f(z)\|_1 \le 1 - \frac{2}{\sqrt{\lambda m} + 1} \le 1 - \frac{2}{\sqrt{\lambda \Delta} + 1}.$$

Then we obtain

$$\begin{aligned} &\left|\log \mathbf{Pr}_{M \sim \mu_{G,\lambda}}\left[v \notin M\right] - \log \Phi_G(v,t)\right| \\ \le &\left(1 - \frac{2}{\sqrt{\lambda\Delta} + 1}\right) \max_{i,j} \left|\log \mathbf{Pr}_{M \sim \mu_{G \setminus \{v,u_i\},\lambda}}\left[w_j^{(i)} \notin M\right] - \log \Phi_{G\setminus\{v,u_i\}}(w_j^{(i)}, t-2)\right|. \end{aligned}$$

Then the inequality holds by the simple calculation when $t = 0$ or 1 by Proposition 2.3. □

To estimate $Z_G(\lambda)$, it suffices to estimate $\Phi_G(v,t)$ when t is not large. Simply by definition we can compute $\Phi_G(v,t)$ in time $O(\Delta^t)$, and with proper choice of t, the error can be bounded. The core is the following algorithm:

Algorithm 1: estimating $Z_G(\lambda)$

input : a graph $G = (V, E)$ with $|V| = n$ and maximum degree Δ, a fugacity $\lambda > 0$ and an tolerance error $\varepsilon \in (0, 1)$

output: an ε-approximation $\widehat{Z}$ for the partition function $Z_G(\lambda)$

```
Z ← 1, H ← G;
Set δ ← −log(1 − 2/(√(1+λΔ)+1)) and t ← 2⌈(log n + log log(1 + λΔ) − log ε)/δ⌉;
while H ≠ ∅ do
    choose an arbitrary vertex v ∈ H;
    compute Φ_H(v, t);
    Set Z ← 1/Φ_H(v,t) and H ← H \ {v};
return Z.
```

Equipped with Algorithm 1, we are ready to prove Theorem 2.1.

Proof of Theorem 2.1. We apply Algorithm 1. Note that, under the choice of δ and t, by Theorem 2.4, it holds that

$$e^{-\varepsilon/n} \le \frac{\Phi_H(v,t)}{\mathbf{Pr}_{M\sim\mu_{H,\lambda}}\left[v \notin M\right]} \le e^{\varepsilon/n}$$

and the running time is $O(n\Delta^t) = O\left((n/\varepsilon)^{\kappa \log \Delta + 1}\right)$ where $\kappa = -\frac{2}{\log\left(1-\frac{2}{\sqrt{1+\lambda\Delta}+1}\right)}$. Then we know

$$e^{-\varepsilon} \le \frac{\widehat{Z}}{Z} \le e^{\varepsilon}$$

by Proposition 2.2. □

3 Sampling Matchings

To sample from $\mu_{G,\lambda}$, we consider the following chain introduced in [JS96]. Suppose that we are now at $M \in \Omega$. The update rule is described as following:

1. with probability $1/2$ let $M' = M$; otherwise
2. pick an edge $e = \{u, v\} \in E$ uniformly at random and let
$$M' = \begin{cases} M \setminus \{e\} & e \in M; \\ M \cup \{e\} & u, v \in M; \\ M \cup \{e\} \setminus \{e'\} & \text{exactly one of } u \text{ and } v \text{ is in } M \text{ and } e' \text{ is the matching edge;} \\ M & \text{otherwise;} \end{cases}$$
3. go to M' with probability $\min\left\{1, \frac{\mu_{G,\lambda}(M')}{\mu_{G,\lambda}(M)}\right\}$.

We denote this Markov chain by P. Note that when $M \neq M'$, the ratio $\frac{\mu_{G,\lambda}(M')}{\mu_{G,\lambda}(M)}$ takes values in $\left\{\lambda^{-1}, 1, \lambda\right\}$, corresponding to three kinds of transitions:

- (Type 1) An edge is removed from M.
- (Type 2) An edge is added to M.

- (Type 3) A new edge is exchanged with an edge in M.

Proposition 3.1 (Proposition 12.4 in [JS96]). *For every graph $G = (V, E)$ with $|V| = n$ and fugacity $\lambda > 0$, let P be the Jerrum and Sinclair's chain with respect to the Gibbs distribution $\mu_{G,\lambda}$ of the monomer-dimer model on G at fugacity $\lambda > 0$. Then the mixing time of Jerrum and Sinclair's chain satisfies*

$$t_{\text{mix}}(\varepsilon) \leq 4|E|n\overline{\lambda}\left(n\left(\log n + \log \overline{\lambda}\right) - \log \varepsilon\right)$$

where $\overline{\lambda} = \max\{1, \lambda\}$.

We prove Proposition 3.1 by canonical paths. For simplicity of the analysis we consider the following definition of congestion:

$$\overline{\rho}(\Gamma) := \max_{M,M':P(M,M')\neq 0} \frac{1}{\mu_{G,\lambda}(M)P(M,M')} \sum_{\gamma\in\Gamma:\gamma\ni(M,M')} \mu_{G,\lambda}(M)\mu_{G,\lambda}(M')|\gamma|$$

where $|\gamma|$ is the length of γ.

Proposition 3.2 (Proposition 12.1 in [JS96]). *Let P be a finite, reversible and ergodic lazy Markov chain with respect to the stationary distribution μ over Ω. Let Γ be a set of canonical paths from Ω to Ω. Then*

$$t_{\text{mix}}(\varepsilon) \leq \overline{\rho}(\Gamma)\left(\log \frac{1}{\mu(x)} + \log \frac{1}{\varepsilon}\right)$$

for any initial state $x \in \Omega$.

Then to show the rapid mixing of the Jerrum and Sinclair's chain, it suffices to construct canonical paths Γ with low congestion.

Lemma 3.3 ([JS96]). *For every graph $G = (V, E)$ with $|V| = n$ and fugacity $\lambda > 0$, let P be the Jerrum and Sinclair's chain with respect to the Gibbs distribution $\mu_{G,\lambda}$ of the monomer-dimer model on G at fugacity $\lambda > 0$. Then there exists a family of canonical paths Γ from Ω to Ω such that*

$$\overline{\rho}(\Gamma) \leq 4|E|n\overline{\lambda}$$

where $\overline{\lambda} = \max\{1, \lambda\}$.

3.1 Construction and analysis of canonical paths

Now we construct Γ to prove Lemma 3.3. For a pair of matchings $X, Y \in \Omega$, we consider the symmetric difference $X \oplus Y$. It is not hard to observe that it consists of a disjoint collection of paths or cycles in G, each of which has edges that belong alternately to X and Y. Let $\mathcal{P}(G)$ be the collection of all simple paths and cycles in G. Now suppose that there exists an arbitrary order of $\mathcal{P}(G)$ and we designate each of them a 'start vertex', which is arbitrary if it is a cycle and must be an endpoint otherwise. Then it induces a unique order $P_1, \ldots, P_m$ on the paths and cycles over $X \oplus Y$. Then the canonical path from X to Y involves 'unwinding' each of the P_i in turn as follows:

1. P_i is a simple path. Let P_i consist of the sequence $(v_0, v_1, \ldots, v_\ell)$ where v_0 is the start vertex. If $(v_0, v_1) \in Y$, we perform a sequence of (Type 3) transitions replacing (v_{2j+1}, v_{2j+2}) with (v_{2j}, v_{2j+1}) for $j = 0, 1, \ldots$ and finish with a (Type 2) transition if necessary. If $(v_0, v_1) \in X$, we firstly perform a (Type 1) transition removing (v_0, v_1) and proceed as before for the reduced path $(v_1, \ldots, v_\ell)$.

2. P_i is a cycle. Let P_i consist of the sequence $(v_0, v_1, \ldots, v_{2\ell+1})$ where $\ell \geq 1$, v_0 is the start vertex and $(v_{2j}, v_{2j+1}) \in X$ for $0 \leq j \leq \ell$. Then we firstly perform a (Type 1) transition to remove (v_0, v_1), and leave an open path O with endpoints v_0, v_1. Since one of v_0, v_1 must be the start vertex of O, suppose that v_k is not the start vertex. Then we proceed as 1 but treat v_k as the start vertex, in order to distinguish paths from cycles.

Now we bound $\overline{\rho}(\Gamma)$. Let $e = (M, M')$ be a transition edge in the Markov chain and $\text{pass}(e) = \{(X, Y) : \gamma_{XY} \ni e\}$. Now we consider the injective mapping

$$\eta_e : \text{pass}(e) \to \Omega.$$

Intuitively we want $\eta_e(X, Y) = X \oplus Y \oplus (M \cup M')$. However, $\eta_t(X, Y)$ might not be a matching. To ensure that it is a matching, we might remove the edge of X which is adjacent to the start vertex of the path currently unwound: we call this edge f^e_{XY}. Then we define

$$\eta_e(X, Y) = \begin{cases} (X \oplus Y \oplus (M \cup M')) \setminus f^e_{XY}, & e \text{ is (Type 3) and the current path is a cycle;} \\ X \oplus Y \oplus (M \cup M'), & \text{otherwise.} \end{cases}$$

It is not hard to see η_e is a injective function. Now under the mapping η_e, we show the low congestion of canonical paths.

Proof of Lemma 3.3. We construct Γ and injective mapping η_e for transition $e = (M, M')$ as above. Then firstly we show

$$\mu_{G,\lambda}(X)\mu_{G,\lambda}(Y) \le 2|E|\overline{\lambda}^2 \mu_{G,\lambda}(M)P(M, M')\mu_{G,\lambda}(\eta_e(X, Y)). \tag{1}$$

This bound is enough to show the congestion, since the following equality holds

$$\begin{aligned} \frac{1}{\mu_{G,\lambda}(M)P(M, M')} \sum_{\gamma_{XY} \ni e} \mu_{G,\lambda}(X)\mu_{G,\lambda}(Y)|\gamma_{XY}| &\le 2|E|\overline{\lambda}^2 \sum_{\gamma_{XY} \ni e} \mu_{G,\lambda}(\eta_e(X, Y))|\gamma_{XY}| \\ &\le 4|E|n\overline{\lambda}^2 \sum_{\gamma_{XY} \ni e} \mu_{G,\lambda}(\eta_e(X, Y)) \\ &\le 4|E|n\overline{\lambda}^2. \end{aligned}$$

Now we prove (1). Observe that

$$\mu_{G,\lambda}(M)P(M, M') = \frac{\min\left\{\mu_{G,\lambda}(M), \mu_{G,\lambda}(M')\right\}}{2|E|}.$$

We separate the remaining parts into four cases:

1. e is a (Type 1) transition. Suppose that $M' = M \setminus \{f\}$. Then $\eta_e(X, Y) = X \oplus Y \oplus M$. Then we have

$$\begin{aligned} \mu_{G,\lambda}(X)\mu_{G,\lambda}(Y) &= \mu_{G,\lambda}(M)\mu_{G,\lambda}(\eta_e(X, Y)) \\ &= \frac{2|E|\mu_{G,\lambda}(M)P(M, M')}{\min\left\{\mu_{G,\lambda}(M), \mu_{G,\lambda}(M')\right\}}\mu_{G,\lambda}(M)\mu_{G,\lambda}(\eta_e(X, Y)) \\ &= 2|E|\mu_{G,\lambda}(M)P(M, M') \max\left\{1, \frac{\mu_{G,\lambda}(M)}{\mu_{G,\lambda}(M')}\right\} \mu_{G,\lambda}(\eta_e(X, Y)) \\ &\le 2|E|\overline{\lambda}\mu_{G,\lambda}(M)P(M, M')\mu_{G,\lambda}(\eta_e(X, Y)). \end{aligned}$$

2. e is a (Type 2) transition. The analysis is similar to the last one.

3. e is a (Type 3) transition and the current path is a cycle. Suppose that $M' = M \cup \{f\} \setminus \{f'\}$. Then

$$\eta_e(X, Y) = X \oplus Y \oplus (M \cup \{f\}) - f^e_{XY}.$$

   Then we know $M \cup \eta_e(X, Y)$ differs from $X \cup Y$ only in f and f^e_{XY}. Thus we have

$$\mu_{G,\lambda}(X)\mu_{G,\lambda}(Y) \le 2|E|\overline{\lambda}^2 \mu_{G,\lambda}(M)P(M, M')\mu_{G,\lambda}(\eta_e(X, Y)).$$

4. e is a (Type 3) transition and the current path is not a cycle. The analysis is identical to 3 with no necessity of the consideration of f_{XY}^e. Then it holds that

$$\mu_{G,\lambda}(X)\mu_{G,\lambda}(Y) \le 2|E|\overline{\lambda}\mu_{G,\lambda}(M)P(M,M')\mu_{G,\lambda}(\eta_e(X,Y)).$$

Then we can conclude the inequality except the term with respect to $\overline{\lambda}$ is $\overline{\lambda}^2$. To reduce this order, firstly observe that only the analysis of the third case produces this term. In this case we will show $\sum_{\gamma_{XY}\ni e}\mu_{G,\lambda}(\eta_e(X,Y))$ is upper bounded by $\overline{\lambda}^{-1}$.

To see this, note that $\eta_e(X,Y)$ has at least two unmatched vertices, the start vertex of the current cycle and the common vertex adjacent to f and f'. Moreover, in $\eta_e(X,Y)\oplus M$ the two vertices are linked by an alternating path. Then we argument this path and produce a new matching. Note that different $\eta_e(X,Y)$ produce different matchings. Then we can show the upper bound. □

4 Zeros of Partition Functions

In this section we investigate the zeros of $Z_G(\lambda)$ when $\lambda\in\mathbb{C}$. We rewrite it as

$$Z_G(\lambda)=\sum_{k=0}^{\infty}m_k\lambda^k$$

where m_k denotes the number of matchings on G with size k for all $k\in\mathbb{N}$.

Theorem 4.1 (Theorem 1.2 in [PR17], a restatement of Theorem 2.1)**.** *For any graph $G=(V,E)$ with maximum degree Δ and any $\lambda\in\mathbb{C}$ which is not a non-negative real number, there exists a deterministic algorithm for $(1+\varepsilon)$-approximation to $Z_G(\lambda)$ with running time polynomial in n and ε^{-1} and exponential in λ and Δ.*

By Riemann's Mapping Theorem it suffices to show the partition function $Z_G(\lambda)$ is zero-free outside of a disk centered at the origin with radius $1/\Omega(\Delta)$, and following the methodology of Patel and Regts [PR17] we can prove Theorem 4.1. It is equivalent to consider the following monomer-dimer polynomial which is fully investigated in Heilmann and Lieb [HL72]:

$$P(G,x)=\sum_{k=0}^{\infty}(-1)^k m_k x^{n-2k},\quad \forall x\in\mathbb{C}.$$

Lemma 4.2. *For every graph $G=(V,E)$ with maximum degree $\Delta>0$, the largest root of the polynomial $P(G,x)$ is at most $2\sqrt{\Delta-1}$.*

Before we prove Lemma 4.2, the following identities are of great significance.

Fact 4.3. *For any pair of disjoint graphs G,H, it holds that*

$$P(G\cup H,x)=P(G,x)\cdot P(H,x).$$

For any graph $G=(V,E)$ and every vertex $v\in V$, it holds that

$$P(G,x)=xP(G\setminus\{v\},x)-\sum_{u\in N_G(v)}P(G\setminus\{v,u\},x).$$

4.1 Zeros of the monomer-dimer polynomials on trees

When the graph is a tree $T = (V, E)$, it is much simpler to show the properties of the monomer-dimer polynomial $P(T, x)$.

Lemma 4.4. *For any tree $T = (V, E)$, let A be the adjacent matrix of T. Then the following identity holds*

$$P(T, x) = \det(xI - A).$$

Proof. To show the identity, we consider the coefficients of x^{n-2k} in the two polynomial for every $k \in \mathbb{N}$. Firstly we show that they are the same in x^0. The coefficient of x^0 in $P(T, x)$ is just the number of perfect matchings of T. On the other hand, the coefficient of x^0 in the characteristic polynomial of A is just the determinant of A, *i.e.*,

$$\sum_{\sigma} \operatorname{sgn}(\sigma) \prod_{i} A_{i,\sigma(i)}.$$

We claim that every σ satisfies $\prod_i A_{i,\sigma(i)} \neq 0$ if and only if σ corresponds to a perfect matching. In fact, it holds that σ is a collection of cycles if for all i, $A_{i,\sigma(i)} = 1$. However, since T is a tree, the cycles in σ must have length 2. Then that is to say, $\sigma(\sigma(i)) = i$. Then this produces a unique perfect matching. On the other hand when σ corresponds to a perfect matching it is trivial that the term is non-zero. In this case, it is easy to see $\operatorname{sgn}(\sigma) = (-1)^{n/2}$.

Now for other $k \in \mathbb{N}$, note that the coefficient of x^{n-2k} in $\det(xI - A)$ is the sum of determinant of all principal $2k \times 2k$ minors of A. Then we know each such a determinant is equal to the number of perfect matchings in the corresponding induced subgraph of T with $2k$ vertices. □

Lemma 4.5. *For any tree $T = (V, E)$ with maximum degree Δ, the largest root of the monomer-dimer polynomial $P(T, x)$ is at most $2\sqrt{\Delta - 1}$.*

Proof. By Lemma 4.4 it suffices to bound the largest eigenvalue of the adjacent matrix A of T. We apply the trace method. Note that for every real symmetric matrix $M \in \mathbb{R}^{n\times n}$

$$\operatorname{Tr}(M) = \sum_{i=1}^{n} \lambda_i$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Then it holds that

$$\lim_{k\to\infty} \operatorname{Tr}(A^k)^{1/k} = \lambda_{\max}(A).$$

Then it suffices to show for every $u \in V$, it holds that when $k \to \infty$,

$$A^k_{u,u} \leq 2^k (\Delta - 1)^{k/2} \tag{2}$$

and plugging into the trace we know

$$\lambda_{\max}(A) = \lim_{k\to\infty} \operatorname{Tr}(A^k)^{1/k} \leq \lim_{k\to\infty} n^{1/k} 2\sqrt{\Delta - 1} = 2\sqrt{\Delta - 1}.$$

To prove (2), observe that $A^k_{u,u}$ is the number of closed walks of length k starting at u. Now we think of the tree $T = (V, E)$ rooted at vertex u. Then we know that there are $k/2$ 'down walks' (each of which has at most $\Delta - 1$ choices) and $k/2$ 'up walks' (each of which has one choice). Then we know

$$A^k_{u,u} \leq 1^{k/2} (\Delta - 1)^{k/2} \binom{k}{k/2} \leq 2^k (\Delta - 1)^{k/2}$$

where the last inequality holds from $\binom{k}{k/2} \leq 2^k$. □

4.2 Zeros of the monomer-dimer polynomials on general graphs

To prove Lemma 4.2 on general graphs, we compare the graph to a tree and show the connection between the monomer-dimer polynomials of them. For a graph $G = (V, E)$ and every vertex $v \in V$, define the *path tree* $T = T(G, v)$ as the tree rooted at v and for each simple path in G starting at v, T has a node corresponding to it and two paths are adjacent if their length differ by 1 and one is a prefix of another.

Lemma 4.6 ([GG81]). *Let $G = (V, E)$ be a graph and $v \in V$ be an arbitrary vertex in G. Let $T = T(G, v)$ be the path tree of G from v. Then it holds that*

$$\frac{P(G,x)}{P(G\setminus\{v\},x)} = \frac{P(T,x)}{P(T\setminus\{v\},x)}.$$

Furthermore, the polynomial $P(G, x)$ divides $P(T, x)$.

Note that Lemma 4.6 shows that the roots of $P(G, x)$ are the subset of the roots of $P(T, x)$. Then by Lemma 4.4, we conclude Lemma 4.2.

Proof of Lemma 4.6. When G is a tree, the identity holds since $G = T(G, v)$. Then inductively we suppose that the identity holds for any proper subgraph of G. Let $H = G \setminus \{v\}$. By Fact 4.3, it holds that

$$\begin{aligned}\frac{P(G,x)}{P(H,x)} &= \frac{xP(H,x) - \sum_{u\in N_G(v)} P(H\setminus\{u\},x)}{P(H,x)} \\ &= x - \sum_{u\in N_G(v)} \frac{P(H\setminus\{u\},x)}{P(H,x)} \\ &= x - \sum_{u\in N_G(v)} \frac{P(T(H,u)\setminus\{u\},x)}{P(T(H,u),x)}.\end{aligned}$$

Observe that, the tree $T(H, u) = T(G \setminus \{v\}, u)$ is isomorphic to the component of $T(G, v) \setminus \{v\}$ which contains the point corresponding to the path $v \to u$. Therefore,

$$\frac{P(T(H,u)\setminus\{u\},x)}{P(T(H,u),x)} = \frac{P(T(G,v)\setminus\{u,v\},x)}{P(T(G,v)\setminus\{v\},x)}.$$

Then we know

$$\begin{aligned}x - \sum_{u\in N_G(v)} \frac{P(T(H,u)\setminus\{u\},x)}{P(T(H,u),x)} &= x - \sum_{u\in N_G(v)} \frac{P(T(G,v)\setminus\{u,v\},x)}{P(T(G,v)\setminus\{v\},x)} \\ &= \frac{xP(T(G,v)\setminus\{v\},x) - \sum_{u\in N_G(v)} P(T(G,v)\setminus\{u,v\},x)}{P(T(G,v)\setminus\{v\},x)} \\ &= \frac{P(T,x)}{P(T\setminus\{v\},x)}.\end{aligned}$$

To see the second conclusion, firstly by the first identity we know

$$P(T,x) = P(G,x)\frac{P(T\setminus\{v\},x)}{P(G\setminus\{v\},x)}.$$

Then we need to show $P(T \setminus \{v\}, x)$ is divisible by $P(G \setminus \{v\}, x)$. To show this firstly note that $P(T \setminus \{v\}, x)$ is divisible by $P(T(G \setminus \{v\}, u), x)$, since the latter is isomorphic to one of the connected component of the previous one. Then by induction we know $P(T(G\setminus\{v\}, u), x)$ is divisible by $P(G\setminus\{v\}, x)$. Thus we conclude the results. □

References

[BGK+07] Mohsen Bayati, David Gamarnik, Dimitriy Katz, Chandra Nair, and Prasad Tetali. Simple Deterministic Approximation Algorithms for Counting Matchings. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '07, pages 122–127, New York, NY, USA, 2007. Association for Computing Machinery. 1, 2

[GG81] C. D. Godsil and I. Gutman. On the Theory of the Matching Polynomial. *Journal of Graph Theory*, 5(2):137–144, 1981. 9

[HL72] Ole J. Heilmann and Elliott H. Lieb. Theory of Monomer-Dimer Systems. *Communications in Mathematical Physics*, 25(3):190–232, 1972. 7

[JS96] Mark Jerrum and Alistair Sinclair. *The Markov Chain Monte Carlo Method: an Approach to Approximate Counting and Integration*, pages 482–520. PWS Publishing Co., USA, 1996. 4, 5

[PR17] Viresh Patel and Guus Regts. Deterministic Polynomial-Time Approximation Algorithms for Partition Functions and Graph Polynomials. *Electronic Notes in Discrete Mathematics*, 61:971–977, 2017. The European Conference on Combinatorics, Graph Theory and Applications (EUROCOMB'17). 7