6.S891 Lecture 10: Lee–Yang Theory and Multivariate Zero-Freeness

Kuikui Liu

October 12, 2023

As we saw in the previous lectures how zero-freeness of the partition function implies the existence of deterministic efficient approximate counting algorithms. This lecture further develops this theory in two important ways. In the first part of this lecture, we study a new method for establishing zero-free regions based on *Asano–Ruelle contractions*. This is another extremely elegant "local method" for proving zero-freeness. We will use this method to prove the famous Lee–Yang Circle Theorem for the ferromagnetic Ising model. In preparation for our discussion of the polymer method and the cluster expansion, the second part of the lecture is devoted to proving *multivariate* zero-freeness in a *polydisk* around 0 for the multivariate independence polynomial. This is related to our previous discussion of the hardcore model, except the emphasis now will be on *nonuniform* fugacities.

Since it will be fundamental throughout, let us first describe the "types" of zero-free regions we will encounter in the setting of multivariate polynomials.

Definition 1 (Stability)**.** *Let $\Gamma_1, \dots, \Gamma_n \subseteq \mathbb{C}$ be subsets of the complex plane. We say a multivariate polynomial $p(z_1, \dots, z_n)$ is $\Gamma_1 \times \cdots \times \Gamma_n$-stable if $p(\boldsymbol{z}) \neq 0$ whenever $z_i \in \Gamma_i$ for all $i = 1, \dots, n$. If $\Gamma_1 = \cdots = \Gamma_n = \Gamma$ for some $\Gamma \subseteq \mathbb{C}$, then we simply say p is Γ-stable.*

Typically, our region Γ will be an open (possibly shifted) half-plane (e.g. $\mathbb{H}_t = \{z \in \mathbb{C} : \operatorname{Re} z > t\}$ for some $t \in \mathbb{R}$), a (scaled) unit disk, or its complement. Such regions are sometimes called *circular*. A *polydisk* is a Cartesian product of (possible scaled and shifted) disks in $\mathbb{C}$.

At first sight, it might seem unnatural to restrict attention to "product-type" sets $\Gamma_1 \times \cdots \times \Gamma_n$, since at maximum generality one could just study the set $\{\boldsymbol{z} \in \mathbb{C}^n : p(\boldsymbol{z}) \neq 0\}$. However, this product structure becomes convenient when we wish to invoke "self-reducibility" arguments, which roughly corresponds to fixing $\boldsymbol{z}_S$ to some $\boldsymbol{a}_S \in \mathbb{C}^S$ and looking at the induced polynomial in the variables $\boldsymbol{z}_{V \setminus S}$.

1 The Lee–Yang Circle Theorem

For a graph $G = (V, E)$, an inverse temperature $\beta \geq 0$, and a vector of external fields $h \in \mathbb{R}^V$, recall the ferromagnetic Ising model has Gibbs distribution on $\{\pm 1\}^V$ given by $\mu(\sigma) \propto \exp\left(\frac{\beta}{2}\sigma^\top A_G \sigma + \langle h, \sigma \rangle\right)$. If we view each $\sigma \in \{\pm 1\}^V$ as being the indicator of the set of vertices $S = \{v \in V : \sigma(v) = +1\}$, then

$$\frac{1}{2}\sigma^\top A_G \sigma = \sum_{uv \in E} \sigma_u \sigma_v = |E| - 2 \cdot |E(S, V \setminus S)|$$

where $E(S, V \setminus S)$ denotes the collection of edges crossing the cut $(S, V \setminus S)$. Hence, using the minor change of variables $\lambda_v = \exp(2h_v)$, we can write rewrite the partition function of the ferromagnetic Ising model as the *multivariate cut polynomial*

$$Z_{G,\beta}(\boldsymbol{\lambda}) \propto \sum_{S \subseteq V} \exp\left(-2\beta \cdot |E(S, V \setminus S)|\right) \cdot \boldsymbol{\lambda}^S$$

up to an easy-to-compute normalization factor; here and throughout, we will use the convenient shorthand $\boldsymbol{\lambda}^S \overset{\mathsf{def}}{=} \prod_{v \in S} \lambda_v$. We also write $Z_{G,\beta}(\lambda) = Z_{G,\beta}(\lambda \mathbf{1})$ for the univariate restriction. The following beautiful theorem of Lee–Yang asserts that all zeros of the $Z_{G,\beta}(\lambda)$ are confined to the unit circle.

Theorem 1.1 (Lee–Yang Circle Theorem; [LY52]). *For every graph $G = (V, E)$ and every $\beta \geq 0$, all zeros of $Z_{G,\beta}(\lambda)$ lie on the unit circle $\{z \in \mathbb{C} : |z| = 1\}$. More generally, the multivariate polynomial $Z_{G,\beta}(\boldsymbol{\lambda})$ is $\mathbb{D}$-stable and $\overline{\mathbb{D}}^{\mathsf{c}}$-stable.*

Remark 1. The zeros of $Z_{G,\beta}(\lambda)$ are sometimes called *Lee–Yang zeros*. Instead of fixing the inverse temperature and reparametrizing the partition function as a polynomial in the external field, one can also fix the external field and reparametrize the partition function as a polynomial in the interaction strength. In this case, the zeros of the resulting polynomial are sometimes called the *Fisher zeros*. In [LSS19a], it is shown that there are no Fisher zeros in a strip around precisely the range of β for which the Gibbs distribution of the ferromagnetic Ising model exhibits strong spatial mixing.

Combined with Barvinok's polynomial interpolation algorithm gives a deterministic quasipolynomial-time algorithm for estimating the partition function as long as the external field h is bounded away from 0 (or equivalently, λ is bounded away from 1). With some more work, one can make this into a polynomial-time algorithm when G has bounded maximum degree; see [LSS19b] for details.

Lee and Yang's interest was originally leveraging the presence/absence of zeros to detect phase transitions in statistical physics models. Their perspective was that a phase transition occurs at some threshold λ_c if the *limiting free energy* $\mathcal{F}_\beta(\lambda) = \lim_{n\to\infty} \frac{1}{n} \log Z_{G,\beta}(\lambda)$ (in the specific setting of the Ising model) is discontinuous at λ_c, signaling a loss of smoothness for $\log Z_{G,\beta}(\lambda)$ as one looks at progressively larger and larger graphs. Since our discussion at present is informal, we won't get into issues of whether or not the large n limit exists. The main intuition is that these discontinuities correspond to sudden changes in the behavior of the Gibbs distribution in large-scale systems as one perturbs λ in a neighborhood of λ_c.

For the ferromagnetic Ising model, the fact that all zeros of the univariate partition function $Z_{G,\beta}(\lambda)$ are restricted to the unit circle means there can be at most one such phase transition in the parameter λ. Furthermore, if there is any such phase transition, then it must occur at $\lambda = 1$. Theorem 1.1 formalizes the intuition that when $\lambda > 1$ is a constant, then all vertices are macroscopically biased towards $+1$ simultaneously, which destroys the natural bottleneck that would otherwise be formed between predominantly -1 configurations and predominantly $+1$ configurations. The same holds for $\lambda < 1$, with the roles of $+1$ and -1 reversed.

In contrast, we mentioned in a previous lecture that for the hardcore model on graphs G of maximum degree Δ, the independence polynomial $Z_G(\lambda)$ is zero-free in a small strip around the interval $[0, \lambda_c(\Delta))$; here, recall $\lambda_c(\Delta)$ is again the uniqueness threshold for the hardcore model. On the other hand, these zeros do cluster around $\lambda_c(\Delta)$, forming a *barrier* between $\lambda < \lambda_c(\Delta)$ where we have efficient approximate counting algorithms (e.g. via correlation decay or polynomial interpolation), and $\lambda > \lambda_c(\Delta)$ where the independence polynomial becomes NP-hard to approximate. We refer interested readers to [Ben+23] for further discussion of the roots of the independence polynomial.

In the rest of the section, we prove Theorem 1.1. As a sanity check, let us start by proving the "base case".

Lemma 1.2. *For any $a \in \mathbb{C}$ with $|a| \leq 1$, the bivariate polynomial $p(z_1, z_2) = 1 + az_1 + \overline{a}z_2 + z_1z_2$ is $\mathbb{D}$-stable.*

If $a = \exp(-2\beta)$ for $\beta \geq 0$, the polynomial $p(z_1, z_2)$ is exactly the cut polynomial for the graph consisting of two vertices connected by an edge. Hence, Lemma 1.2 genuinely captures a very special (but nonetheless important) case of Theorem 1.1.

Proof of Lemma 1.2. If $|a| = 1$, then $|\overline{a}| = 1$ and $a \cdot \overline{a} = 1$ and so we can factor p as $p(z_1, z_2) = (1 + az)(1 + \overline{a}z)$. This is clearly $\mathbb{D}$-stable since each of the factors are $\mathbb{D}$-stable.

Now suppose $|a| < 1$. For every $z_2 \in \mathbb{C}$, there is a unique z_1 such that $p(z_1, z_2) = 0$ given by $z_1 = f(z_2)$, where f is the *Möbius transformation*

$$f(z) = -\frac{1 + \overline{a}z}{a + z}.$$

To prove $\mathbb{D}$-stability, we must show that for any z satisfying $|z| < 1$, we have $|f(z)| > 1$. For this, a routine computation shows that $f^{-1}(z) = -\frac{1+az}{\overline{a}+z}$; essentially f^{-1} is the same as f except we have exchanged the roles of a and $\overline{a}$.

We prove that f (and f^{-1}) map the unit circle $\partial\mathbb{D} = \{z \in \mathbb{C} : |z| = 1\}$ into itself. Clearly, $|f(0)| > 1$ by the fact that $|a| < 1$. These two facts imply $|f(z)| > 1$ for all $z \in \mathbb{D}$ because if

$|f(z)| < 1$ for some $z \in \mathbb{D}$, then by continuity there must exist $z' \in \mathbb{D}$ along the line segment (or any smooth curve contained in $\mathbb{D}$) connecting 0 and z such that with $|f(z')| = 1$, contradicting $f^{-1}(\partial\mathbb{D}) = \partial\mathbb{D}$.

To show $f(\partial\mathbb{D}) = \partial\mathbb{D}$, observe that if $|z| = 1$, then

$$\begin{aligned} |1 + \overline{a}z| &= |1 + \overline{a}z| \cdot |\overline{z}| && (\text{Using } |z| = |\overline{z}| = 1) \\ &= |\overline{z} + \overline{a}| && (\text{Using } |z|^2 = 1) \\ &= |a + z|\,. \end{aligned}$$

In particular, $|f(z)| = 1$ for every $|z| = 1$. The same holds for f^{-1} be replacing a with $\overline{a}$. □

1.1 On Hadamard Products of Polynomials

The main technical theorem towards the full proof of Theorem 1.1 is the following result showing $\mathbb{D}$-stability is preserved under taking *Hadamard/Schur products* of polynomials. For two multiaffine polynomials $p(\boldsymbol{z}) = \sum_{S\subseteq[n]} a_S \boldsymbol{z}^S$ and $q(\boldsymbol{z}) = \sum_{S\subseteq[n]} b_S \boldsymbol{z}^S$, we define their Hadamard product to be the polynomial $(p * q)(z) = \sum_{S\subseteq[n]} a_S b_S \boldsymbol{z}^S$ given by coefficient-wise products; as previously mentioned, we use the shorthand $\boldsymbol{z}^S = \prod_{i\in S} z_i$. This can obviously be defined much more generally, but we concentrate on the multiaffine case.

Theorem 1.3 (Hadamard Products and $\mathbb{D}$-Stability)**.** *Suppose the multiaffine polynomials $p(\boldsymbol{z}) = \sum_{S\subseteq[n]} a_S \boldsymbol{z}^S$ and $q(\boldsymbol{z}) = \sum_{S\subseteq[n]} b_S \boldsymbol{z}^S$ are $\mathbb{D}$-stable. Then so is their Hadamard product $p * q$.*

We will prove Theorem 1.3 via a technique called *Asano–Ruelle contractions* in the next subsection. For now, we combine it with Lemma 1.2 and an inductive argument to complete the proof of the Lee–Yang Theorem. While we could have directly used Asano–Ruelle contractions and Lemma 1.2 to do this, we believe Theorem 1.3 and theorems like it are interesting in their own right.

Proof of Theorem 1.1. For each edge $e = \{u, v\} \in E$, define the polynomial

$$\begin{aligned} p_e(\boldsymbol{\lambda}) &\stackrel{\text{def}}{=} \sum_{S\subseteq V : E(S, V\setminus S)\ni e} \exp(-2\beta) \cdot \boldsymbol{\lambda}^S + \sum_{S\subseteq V : E(S,V\setminus S)\not\ni e} \boldsymbol{\lambda}^S \\ &= \left(1 + e^{-2\beta}\lambda_u + e^{-2\beta}\lambda_v + \lambda_u\lambda_v\right) \cdot \prod_{w\in V\setminus\{u,v\}} (1 + \lambda_w). \end{aligned}$$

The second representation of p_e plus Lemma 1.2 shows that p_e is $\mathbb{D}$-stable. The first representation of p_e shows that our final partition function $Z_{G,\beta}(\boldsymbol{\lambda})$, the full multivariate cut polynomial of G, is a Hadamard product of the $\{p_e(\boldsymbol{\lambda})\}_{e\in E}$. A straightforward inductive application of Theorem 1.3 shows that $Z_{G,\beta}(\boldsymbol{\lambda})$ is $\mathbb{D}$-stable. Since $Z_{G,\beta}(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^V \cdot Z_{G,\beta}(1/\boldsymbol{\lambda})$ just from the fact that $E(S, V\setminus S)$ is invariant if you replace S with $V \setminus S$, it also follows that $Z_{G,\beta}(\boldsymbol{\lambda})$ is $\overline{\mathbb{D}}^{\mathsf{c}}$-stable. By restricting $\boldsymbol{\lambda}$ to $\lambda\boldsymbol{1}$, it immediately follows that the univariate polynomial $Z_{G,\beta}(\lambda)$ is simultaneously zero-free in $\mathbb{D}$ and in $\overline{\mathbb{D}}^{\mathsf{c}}$; in particular, its roots are forced to lie on the unit circle $\{z \in \mathbb{C} : |z| = 1\}$. □

1.2 Asano–Ruelle Contractions

In this subsection, we introduce an intriguing local operation for proving zero-freeness where we merge two variables into a single variable and throw away some terms. The goal is to understand the evolution of the zeros under these operations. We will then use this to give an inductive proof of Theorem 1.3.

Lemma 1.4 ([Asa70; Rue71])**.** *Let $K_1, K_2 \subseteq \mathbb{C}$ be closed subsets which do not contain 0. Suppose $a, b, c, d \in \mathbb{C}$ are complex numbers such that the bivariate polynomial $p(z_1, z_2) = a + bz_1 + cz_2 + dz_1z_2$ is $K_1^{\mathsf{c}} \times K_2^{\mathsf{c}}$-stable. Then the univariate polynomial $q(z) = a + dz$ is $(-K_1 \cdot K_2)^{\mathsf{c}}$-stable.*

Proof. We prove this lemma in the special case $K_1 = K_2 = \mathbb{D}^{\mathsf{c}}$. In this case, the lemma asserts that if $a + bz_1 + cz_2 + dz_1z_2$ is $\mathbb{D}$-stable, then so is $a + dz$. For the general case, we refer the reader to [Rue71].

Since p is $\mathbb{D}$-stable, $a \neq 0$. We can assume $d \neq 0$ since otherwise, the claim is vacuous. The unique root of q is $-\frac{a}{d}$. Suppose for contradiction that it lies inside the unit disk $\mathbb{D}$, i.e. $|a| < |d|$.

We may also assume without loss of generality that $|b| \geq |c|$. We use this to construct a pair $z_1, z_2 \in \mathbb{D}$ such that $p(z_1, z_2) = 0$. For this, we exhibit some $z_2 \in \mathbb{D}$ such that $|b + dz_2| > |a| + |c|$. This is sufficient because for each $z_2 \in \mathbb{C}$, there is a unique choice of z_1 such that $p(z_1, z_2) = 0$ given by $z_1 = -\frac{a+cz_2}{b+dz_2}$, whence

$$|z_1| = \frac{|a + cz_2|}{|b + dz_2|} < \frac{|a| + |c|}{|b + dz_2|} < 1,$$

again assuming z_2 satisfies $|z_2| < 1$ and $|b + dz_2| > |a| + |c|$. To construct such a $z_2 \in \mathbb{D}$, first observe that $|a| < |d|$ and $|b| \geq |c|$ imply $|b| + |d| > |a| + |c|$. With this in mind, choose the angle of z_2 such that dz_2 and b point in the same direction (viewed as vectors in $\mathbb{R}^2$). Then $|b + dz_2| = |b| + |z_2| \cdot |d|$. Clearly, we can choose the length of z_2 to be sufficiently close to 1 such that $|b| + |z_2| \cdot |d| > |a| + |c|$ (since $|b| + |d| > |a| + |c|$ is a strict inequality). The chosen angle and length completely determine z_2, and so we're done. □

Corollary 1.5. *Let $p(z_1, \ldots, z_n) = \sum_{S \subseteq [n]} a_S \boldsymbol{z}^S$ be a multivariate polynomial. If p is $\mathbb{D}$-stable, then so is the polynomial q obtained by applying Asano–Ruelle contraction to the two variables (z_i, z_j), for any $i \neq j$.*

Proof. We may write

$$p(\boldsymbol{z}) = \underbrace{\sum_{S \not\ni i,j} a_S \boldsymbol{z}^S}_{=a(\boldsymbol{z}_{-ij})} + z_i \underbrace{\sum_{S \ni i, S \not\ni j} a_S \boldsymbol{z}^{S-i}}_{=b(\boldsymbol{z}_{-ij})} + z_j \underbrace{\sum_{S \not\ni i, S \ni j} a_S \boldsymbol{z}^{S-j}}_{=c(\boldsymbol{z}_{-ij})} + z_i z_j \underbrace{\sum_{S \ni i,j} a_S \boldsymbol{z}^{S-i-j}}_{=d(\boldsymbol{z}_{-ij})}$$

$$q(\boldsymbol{z}_{-ij}, w) = \sum_{S \not\ni i,j} a_S \boldsymbol{z}^S + w \sum_{S \ni i,j} a_S \boldsymbol{z}^{S-i-j} = a(\boldsymbol{z}_{-ij}) + w \cdot d(\boldsymbol{z}_{-ij}).$$

Fix all variables $\boldsymbol{z}_{-ij}$ to be arbitrarily chosen complex numbers in $\mathbb{D}$. Then with this specialization, p becomes a bivariate polynomial in z_i, z_j, and q becomes a univariate polynomial in w. Since p must remain $\mathbb{D}$-stable, by Lemma 1.4, q is $\mathbb{D}$-stable as a univariate polynomial in w. Since we chose arbitrary complex numbers in $\mathbb{D}$ for $\boldsymbol{z}_{-ij}$, it follows that $q(\boldsymbol{z}_{-ij}, w) \neq 0$ whenever all of its inputs variables are in $\mathbb{D}$, i.e. q is $\mathbb{D}$-stable. □

Proof of Theorem 1.3. Consider the product polynomial

$$r(\boldsymbol{x}, \boldsymbol{y}) \stackrel{\text{def}}{=} p(\boldsymbol{x}) \cdot q(\boldsymbol{y}) = \sum_{S,T \subseteq [n]} a_S b_T \boldsymbol{x}^S \boldsymbol{y}^T$$

on $2n$ variables $x_1, \ldots, x_n, y_1, \ldots, y_n$. Since p, q are each individually $\mathbb{D}$-stable, so is r. We iteratively apply Asano–Ruelle contractions to the pairs of variables (x_i, y_i) for each $i = 1, \ldots, n$. In each contraction, we keep all other variables fixed as if they are constants, and so Corollary 1.5 and $\mathbb{D}$-stability of r imply that all polynomials encountered throughout this process are all $\mathbb{D}$-stable. At the end of all n contractions, we obtain the Hadamard product $p * q$, since we kept only pairs of subsets $S, T \subseteq [n]$ such that for all $i \in [n]$, either i is in both S, T or is in neither; in particular, $S = T$, and we replaced each $x_i y_i$ in $\boldsymbol{x}^S \boldsymbol{y}^T$ with the single variable z_i. □

1.3 A Remark on Half-Plane Stability for Even Subgraphs

Previously, we saw how to design an FPRAS for the ferromagnetic Ising partition function by running Glauber dynamics on a transformation of the configuration space into the collection of *even subgraphs* of G. At the time, the transformation was obtained by a clever change of variables: $\rho_v = \tanh(h_v)$ for all $v \in V$, and $w = \tanh(\beta)$. This was achieved using the rather mysterious identity $e^x = \cosh(x) \cdot (1 + \tanh(x))$ followed by brute force calculations. It turns out, we can provide some post hoc justification for this based on complex analysis, assuming one already knows the Lee–Yang Circle Theorem. If we write $\lambda_v = e^{2h_v}$ as we did above, then $\rho_v = f(\lambda_v)$ where f is the following *Möbius transformation*

$$f(z) = \frac{z - 1}{z + 1}.$$

These basic transformations are extremely fundamental and well-studied in complex analysis. In particular, our f here maps

- the unit circle $\partial\mathbb{D}$ to the imaginary line $\{z \in \mathbb{C} : \operatorname{Re} z = 0\}$,
- the interior of the unit disk $\mathbb{D}$ to the open left half-plane $\{z \in \mathbb{C} : \operatorname{Re} z < 0\}$, and
- the open exterior of the unit disk $\overline{\mathbb{D}}^{\mathsf{c}}$ to the open right half-plane $\mathbb{H}_0 = \{z \in \mathbb{C} : \operatorname{Re} z > 0\}$.

This Möbius transformation is (almost) the unique one with these properties.[1] In particular, the corresponding partition function for the even subgraphs model admits a huge zero-free region similar to what we saw for the matching polynomial/monomer-dimer model.

Corollary 1.6. *For every $\beta \geq 0$, the multivariate polynomial*

$$\widehat{Z}_{G,\beta}(\boldsymbol{\rho}) \stackrel{\mathsf{def}}{=} \sum_{F \subseteq E} \tanh(\beta)^{|F|} \prod_{v \in \mathsf{odd}(F)} \rho_v$$

is $\mathbb{H}_0$-stable.

2 Stability for Multivariate Independence Polynomials

In this section, we switch gears and study the zeros of the multivariate independence polynomial, which we recall is given by

$$Z_G(\boldsymbol{z}) \stackrel{\mathsf{def}}{=} \sum_{I \subseteq V \text{ independent}} \boldsymbol{z}^I$$

for an arbitrary graph $G = (V, E)$. The emphasis here really is on the multivariate case. We will study stability w.r.t. a polydisk $\prod_{v \in V} \mathbb{D}(0, r_v)$, where the $\{r_v\}_{v \in V}$ can be highly nonuniform. This will be essential for applications to polymer models, where our graphs have highly nonuniform degrees, and each vertex has nontrivial combinatorial meaning.

2.1 Shearer's Condition for Stability

We begin with the sharpest possible condition for polydisk stability of Z_G. Later, we will see stronger sufficient conditions (resulting in weaker theorems) which are easier to check in practice. The following result says that the closest zeros to Z_G to $\mathbf{0}$ lie on the negative real axis. Hence, to check stability, it suffices to look at evaluations of Z_G on the negative real axis. This is known as *Shearer's Condition* [She85]. In Appendix A, we discuss connections with the Lovász Local Lemma (LLL). This connection between polydisk stability of the multivariate independence polynomial and the LLL was first elucidated in a paper of Scott–Sokal [SS05].

To state the result, recall that for a subset of vertices $S \subseteq V$, we write $N[S] = \bigcup_{v \in S} N[v] = S \sqcup \{v : v \sim u \text{ for some } u \in S\}$ for the closed neighborhood of S. We also write $G[S]$ for the induced subgraph on S.

Theorem 2.1 ([SS05])**.** *Let $G = (V, E)$ be a graph, and let $\boldsymbol{p} \in \mathbb{R}^V_{\geq 0}$. Then the following are equivalent:*

- ***Polydisk Stability:*** *The multivariate independence polynomial $Z_G(\boldsymbol{z})$ is $\prod_{v \in V} \overline{\mathbb{D}(0, p_v)}$-stable.*
- ***Shearer's Condition:*** *For every $S \subseteq V$, $Z_{G[S]}(-\boldsymbol{p}_S) > 0$.*

Furthermore, if either of these conditions hold, then the Taylor expansion for $\log Z_G(\boldsymbol{z})$ around $\mathbf{0}$ converges absolutely for all $\boldsymbol{z} \in \prod_{v \in V} \overline{\mathbb{D}(0, p_v)}$.

Proof. We first prove Shearer's Condition implies polydisk stability. Let $\boldsymbol{z} \in \mathbb{C}^V$ be such that $|z_v| \leq p_v$ for all $v \in V$. We prove that

$$\left| \frac{Z_{G[S]}(\boldsymbol{z})}{Z_{G[S-v]}(\boldsymbol{z})} \right| \geq \frac{Z_{G[S]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})}, \qquad \forall v \in S \subseteq V. \tag{1}$$

[1]Any nondegenerate Möbius transformation mapping $\partial\mathbb{D}$ to the imaginary line must have the form $\frac{\overline{b}z + be^{i\theta}}{z - e^{i\theta}}$ for some angle θ and complex number $b \in \mathbb{C}$. In our case, take $\theta = \pi$ so that $e^{i\theta} = -1$, and take $b = 1$.

Once we have Eq. (1), the zero-freeness claim follows. Indeed, $\left|Z_{G[S]}(\boldsymbol{z})\right| \geq \left|Z_{G[S-v]}(\boldsymbol{z})\right| \cdot \frac{Z_{G[S]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})}$, which is positive since $Z_{G[S-v]}(\boldsymbol{z}) \neq 0$ by induction, and $\frac{Z_{G[S]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})} > 0$ by assumption.

We inductively prove Eq. (1). Using what essentially is the tree recursion for the hardcore model, we may decompose the left-hand side as

$$\begin{aligned}
\left|\frac{Z_{G[S]}(\boldsymbol{z})}{Z_{G[S-v]}(\boldsymbol{z})}\right| &= \left|\frac{Z_{G[S-v]}(\boldsymbol{z}) + z_v \cdot Z_{G[S-N[v]]}(\boldsymbol{z})}{Z_{G[S-v]}(\boldsymbol{z})}\right| \\
&\geq 1 - |z_v| \cdot \left|\frac{Z_{G[S-N[v]]}(\boldsymbol{z})}{Z_{G[S-v]}(\boldsymbol{z})}\right| && \text{(Triangle Inequality)} \\
&\geq 1 - p_v \cdot \left|\frac{Z_{G[S-N[v]]}(\boldsymbol{z})}{Z_{G[S-v]}(\boldsymbol{z})}\right|. && \text{(Using } |z_v| \leq p_v\text{)}
\end{aligned}$$

On the other hand, by the same decomposition, the right-hand side is

$$\frac{Z_{G[S]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})} = 1 - p_v \cdot \frac{Z_{G[S-N[v]]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})}.$$

Hence, to show Eq. (1), it suffices to use the inductive hypothesis to establish

$$\frac{Z_{G[S-N[v]]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})} \leq \left|\frac{Z_{G[S-N[v]]}(\boldsymbol{z})}{Z_{G[S-v]}(\boldsymbol{z})}\right|.$$

We use the classic telescoping trick, similar to how we developed the correlation decay algorithm for the hardcore model. If we order the vertices of $N(v) \cap S$ arbitrarily as $u_1, \ldots, u_k$, then defining $S_0 = S - v$ and $S_i = S_{i-1} - u_i$ for all $i = 1, \ldots, k$, we obtain

$$\begin{aligned}
\left|\frac{Z_{G[S-N[v]]}(\boldsymbol{z})}{Z_{G[S-v]}(\boldsymbol{z})}\right| &= \prod_{i=1}^{k} \left|\frac{Z_{G[S_i]}(\boldsymbol{z})}{Z_{G[S_{i-1}]}(\boldsymbol{z})}\right| \\
&\geq \prod_{i=1}^{k} \frac{Z_{G[S_i]}(-\boldsymbol{p})}{Z_{G[S_{i-1}]}(-\boldsymbol{p})} && \text{(Induction)} \\
&= \frac{Z_{G[S-N[v]]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})}.
\end{aligned}$$

This completes the induction and the proof of stability.

Now, suppose Shearer's Condition fails, i.e. there exists $S \subseteq V$ such that $Z_{G[S]}(-\boldsymbol{p}_S) \leq 0$. Let $\boldsymbol{z} \in \mathbb{C}^V$ be defined by $z_v = -p_v$ for $v \in S$, and $z_v = 0$ otherwise. Then $Z_G(\boldsymbol{z}) = Z_{G[S]}(-\boldsymbol{p}_S) \leq 0$. On the other hand, $Z_G(\boldsymbol{0}) = 1$. By the Intermediate Value Theorem, there exists $t \in [0, 1]$ such that $Z_G(t\boldsymbol{z}) = 0$. By construction $t\boldsymbol{z} \in \prod_{v \in V} \overline{\mathbb{D}(0, p_v)}$, and so stability is violated.

To establish absolute convergence of the Taylor series, fix $\boldsymbol{\lambda} \in \prod_{v \in V} \overline{\mathbb{D}(0, p_v)}$, and consider the univariate restriction $\boldsymbol{z} = t\boldsymbol{\lambda}$ for a new complex-valued variable t. Since Z_G is $\prod_{v \in V} \overline{\mathbb{D}(0, p_v)}$-stable, the univariate polynomial $Z_G(t\boldsymbol{\lambda})$ is zero-free in $[-1, 1]$. Since $Z_G(t\boldsymbol{\lambda})$ only has finitely many zeros and $[-1, 1]$ is a closed set, $Z_G(t\boldsymbol{\lambda})$ is zero-free in $(-R, R)$ for some $R > 1$.[2] Convergence of the Taylor series then follows via similar arguments to the one used in the analysis of Barvinok's polynomial interpolation algorithm. □

2.2 Stronger but Easier-to-Check Conditions

While Shearer's Condition is sharp, it is unwieldy to use. The goal of this subsection is to derive stronger sufficient conditions which are easier to apply in practice.

Theorem 2.2 (Dobrushin's Condition; [Dob96a; Dob96b]). *Let $G = (V, E)$ be a graph and $\boldsymbol{z} \in \mathbb{C}^V$. If there exists a nonnegative vector $\boldsymbol{y} \in \mathbb{R}^V_{\geq 0}$ such that*

$$|z_v| \leq \frac{y_v}{\prod_{u \in N[v]}(1 + y_u)}, \qquad \forall v \in V, \tag{2}$$

then $Z_G(\boldsymbol{z}) \neq 0$, and the Taylor expansion for $\log Z_G(\boldsymbol{z})$ around $\boldsymbol{0}$ converges absolutely.

[2] Note that if $\boldsymbol{\lambda}$ is on the boundary of $\prod_{v \in V} \overline{\mathbb{D}(0, p_v)}$, then in principle this $R > 1$ could be extremely close to 1, with proximity depending on n.

Remark 2 (The Univariate Case). Consider the univariate independence polynomial $Z_G(\lambda) = Z_G(\lambda\mathbf{1})$ of a graph $G = (V, E)$ of maximum degree Δ. Since $\boldsymbol{z} = \lambda \cdot \mathbf{1}$, if we take $y_v = \frac{1}{\Delta}$ for all $v \in V$ (which maximizes the function $y \mapsto \frac{y}{(1+y)^{\Delta+1}}$), then Eq. (2) becomes

$$|\lambda| \leq \frac{\Delta^\Delta}{(\Delta+1)^{\Delta+1}}.$$

This is close but slightly better than the bound $\frac{1}{e(\Delta+1)}$ appearing in the Lovász Local Lemma. However, the correct bound using Shearer's Condition is $\lambda^*(\Delta) = \frac{(\Delta-1)^{\Delta-1}}{\Delta^\Delta}$, which is tight by considering $(\Delta-1)$-ary trees of increasingly larger depth (see e.g. [SS05]). Notably this is much *smaller* than the uniqueness threshold $\lambda_c(\Delta) = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta} \approx \frac{e}{\Delta-1}$ for the hardcore Gibbs measure on graphs of maximum degree Δ.

Remark 3 (An Intermediate Version). It was shown in [Bis+11] that one can replace $\prod_{u\in N[v]}(1+y_u)$ in the denominator with $Z_{N[v]}(\boldsymbol{y}_{N[v]})$, the independence polynomial of the closed neighborhood of v. This is a weaker condition than what is required in Eq. (2). In the context of the Lovász Local Lemma and its many variants, this is sometimes called the "Cluster Expansion version", while Eq. (2) is sometimes called the "asymmetric version" (or the "lopsided version").

Proof. Define $\boldsymbol{p} \in \mathbb{R}^V_{\geq}$ by $p_v = \frac{y_v}{\prod_{u\in N[v]}(1+y_u)}$ for all $v \in V$. We verify Shearer's Condition, i.e. that $Z_{G[S]}(-\boldsymbol{p}) > 0$ for all $S \subseteq V$, and then apply Theorem 2.1. We do this by (strong) induction. For each $1 \leq k \leq n$, let $\mathrm{IH}(k)$ denote the following claim: For every $S \subseteq V$ with $|S| = k$,

- $Z_{G[S]}(-\boldsymbol{p}) > 0$, and
- $\frac{Z_{G[S]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})} \geq \frac{1}{1+y_v}$ for every $v \in S$.

The base case $k = 1$ is trivial since $G[S]$ consists of a single isolated vertex v, whence

$$Z_{G[S]}(-\boldsymbol{p}) = 1 - p_v \geq 1 - \frac{y_v}{1+y_v} = \frac{1}{1+y_v} \cdot Z_\emptyset(-\boldsymbol{p}) > 0.$$

Now suppose we have established $\mathrm{IH}(k)$ for some $1 \leq k \leq n-1$. We use this to deduce $\mathrm{IH}(k-1)$. Note that the second claim of $\mathrm{IH}(k)$ says that $Z_{G[S]}(-\boldsymbol{p}) \geq \frac{1}{1+y_v} \cdot Z_{G[S-v]}(-\boldsymbol{p})$, which combined with the first claim of $\mathrm{IH}(k-1)$ (i.e. $Z_{G[S-v]}(-\boldsymbol{p}) > 0$), already implies $Z_{G[S]}(-\boldsymbol{p}) > 0$ for all $S \subseteq V$ with $|S| = k$. Hence, the key is to establish the second claim of $\mathrm{IH}(k)$; note that $Z_{G[S-v]}(-\boldsymbol{p}) > 0$ makes the left-hand ratio well-defined. Using what essentially is the tree recursion for the hardcore model, we may decompose the left-hand side as

$$\begin{aligned}\frac{Z_{G[S]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})} &= \frac{Z_{G[S-v]}(-\boldsymbol{p}) - p_v \cdot Z_{G[S-N[v]]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})} \\ &= 1 - \frac{y_v}{1+y_v} \cdot \frac{1}{\prod_{u\sim v}(1+y_u)} \cdot \frac{Z_{G[S-N[v]]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})}.\end{aligned}$$

Hence, to establish $\frac{Z_{G[S]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})} \geq \frac{1}{1+y_v}$, it suffices to use $\mathrm{IH}(1), \ldots, \mathrm{IH}(k-1)$ to show

$$\frac{Z_{G[S-N[v]]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})} \leq \prod_{u\sim v}(1+y_u).$$

We use the classic telescoping trick, similar to how we developed the correlation decay algorithm for the hardcore model. If we order the vertices of $N(v) \cap S$ arbitrarily as $u_1, \ldots, u_k$, then defining $S_0 = S - v$ and $S_i = S_{i-1} - u_i$ for all $i = 1, \ldots, k$, we obtain

$$\begin{aligned}\frac{Z_{G[S-N[v]]}(-\boldsymbol{p})}{Z_{G[S-v]}(-\boldsymbol{p})} &= \prod_{i=1}^k \frac{Z_{G[S_i]}(-\boldsymbol{p})}{Z_{G[S_{i-1}]}(-\boldsymbol{p})} \\ &\leq \prod_{i=1}^k (1+y_{u_i}) && \text{(Inductive Hypothesis)} \\ &\leq \prod_{u\in N(v)} (1+y_u). && (N(v)\cap S \subseteq N(v))\end{aligned}$$

This completes the induction and the proof. □

Dobrushin's Condition is nice but still not so easy to check. We will see a number of important applications of the following theorem in the next lecture.

Theorem 2.3 (Kotecký–Preiss Condition; [KP86])**.** *Let $G = (V, E)$ be a graph and $\boldsymbol{z} \in \mathbb{C}^V$. If there exists a nonnegative function $a : V \to \mathbb{R}_{\geq 0}$ such that*

$$\sum_{u \in N[v]} |z_u| \cdot e^{a(u)} \leq a(v), \qquad \forall v \in V, \tag{3}$$

then $Z_G(\boldsymbol{z}) \neq 0$, and the Taylor expansion for $\log Z_G(\boldsymbol{z})$ around $\mathbf{0}$ converges absolutely.

Remark 4. The astute reader might notice that Eq. (3) looks awfully similar to various versions of *Dobrushin/Dobrushin–Shlosman uniqueness criterion* [Dob70; DS85; Hay06; DGJ09]. For this, recall that for a q-spin system, we define the Dobrushin influence matrix $\mathscr{R} \in \mathbb{R}^{V \times V}$ via

$$\mathscr{R}(u \to v) \stackrel{\mathsf{def}}{=} \max_{\tau} \max_{\mathfrak{b}, \mathfrak{c} \in [q]} d_{\mathsf{TV}}\left(\mu_v^{\tau, u \leftarrow \mathfrak{b}}, \mu_v^{\tau, u \leftarrow \mathfrak{c}}\right),$$

where $\tau : V \setminus \{u, v\} \to [q]$ is a partial configuration. For the Gibbs measure of the hardcore model with heterogeneous fugacities $\boldsymbol{\lambda} = \{\lambda_v\}_{v \in V} \in \mathbb{R}_{\geq 0}^V$, we have $\mathscr{R}(u \to v) = \frac{\lambda_v}{1+\lambda_v}$ for all pairs of distinct neighboring vertices $\{u, v\} \in E$; all other entries are 0. To certify rapid mixing of Glauber dynamics via path coupling, we typically impose some condition of the following form: For some diagonal matrix $D \in \mathbb{R}_{\geq 0}^{V \times V}$, we have $\left\|D^{-1} \mathscr{R} D\right\|_{\ell_1 \to \ell_1} < 1$, or equivalently,

$$\sum_{u \sim v} \frac{\lambda_u}{1 + \lambda_u} \cdot D(u, u) < D(v, v), \qquad \forall v \in V. \tag{4}$$

Cosmetically, this is almost the same as what we are asking for in Eq. (3). However, Eq. (4) is weaker, which is made possible by the fact that we only consider nonnegative inputs to the partition function; see also the discussion on univariate zero-freeness in Remark 2.

Proof. We verify the conditions of Theorem 2.2. Let $y_v = |z_v| \cdot e^{a(v)}$ for all $v \in V$. Rearranging Eq. (2), we need

$$\prod_{u \in N[v]} \left(1 + |z_u| \cdot e^{a(u)}\right) \leq e^{a(v)}, \qquad \forall v \in V,$$

which after taking logarithms reads

$$\sum_{u \in N[v]} \log \left(1 + |z_u| \cdot e^{a(u)}\right) \leq a(v), \qquad \forall v \in V.$$

This is clearly implied by Eq. (3) due to the inequality $\log(1 + x) \leq x$, which holds for all x. $\square$

References

[Asa70] Taro Asano. "Theorems on the Partition Functions of the Heisenberg Ferromagnets". In: *Journal of the Physical Society of Japan* 29.2 (1970), pp. 350–359 (cit. on p. 3).

[Ben+23] Ferenc Bencs, Péter Csikvári, Piyush Srivastava, and Jan Vondrák. "On complex roots of the independence polynomial". In: *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2023, pp. 675–699. DOI: 10.1137/1.9781611977554.ch29. eprint: https://epubs.siam.org/doi/pdf/10.1137/1.9781611977554.ch29 (cit. on p. 2).

[Bis+11] Rodrigo Bissacot, Roberto Fernández, Aldo Procacci, and Benedetto Scoppola. "An Improvement of the Lovász Local Lemma via Cluster Expansion". In: *Combinatorics, Probability and Computing* 20.5 (2011), pp. 709–719 (cit. on p. 7).

[DGJ09] Martin Dyer, Leslie Ann Goldberg, and Mark Jerrum. "Matrix Norms and Rapid Mixing for Spin Systems". In: *The Annals of Applied Probability* 19.1 (Feb. 2009), pp. 71–107 (cit. on p. 8).

[Dob70] Roland Lvovich Dobrushin. "Prescribing a System of Random Variables by Conditional Distributions". In: *Theory of Probability and its Applications* 15.3 (1970), pp. 458–486 (cit. on p. 8).

[Dob96a] Roland Lvovich Dobrushin. "Estimates of Semiinvariants for the Ising Model at Low Temperatures". In: *American Mathematical Society Translations* 177.32 (1996). Topics in statistical and theoretical physics. F. A. Berezin memorial volume, pp. 59–81 (cit. on p. 6).

[Dob96b] Roland Lvovich Dobrushin. "Perturbation methods of the theory of Gibbsian fields". In: *Lectures on Probability Theory and Statistics: Ecole d'Eté de Probabilités de Saint-Flour XXIV—1994*. Ed. by Pierre Bernard. Berlin, Heidelberg: Springer Berlin Heidelberg, 1996, pp. 1–66. ISBN: 978-3-540-49635-9. DOI: 10.1007/BFb0095674 (cit. on p. 6).

[DS85] R. L. Dobrushin and S. B. Shlosman. "Completely Analytical Gibbs Fields". In: *Statistical Physics and Dynamical Systems: Rigorous Results*. Boston, MA: Birkhäuser Boston, 1985, pp. 371–403. ISBN: 978-1-4899-6653-7 (cit. on p. 8).

[Hay06] Thomas P. Hayes. "A simple condition implying rapid mixing of single-site dynamics on spin systems". In: *FOCS* (2006), pp. 39–46 (cit. on p. 8).

[KP86] R. Kotecký and D. Preiss. "Cluster expansion for abstract polymer models". In: *Communications in Mathematical Physics* 103.3 (1986), pp. 491–498 (cit. on p. 8).

[LSS19a] Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. "Fisher zeros and correlation decay in the Ising model". In: *Journal of Mathematical Physics* 60.10 (2019), p. 103304 (cit. on p. 2).

[LSS19b] Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. "The Ising partition function: Zeros and deterministic approximation". In: *Journal of Statistical Physics* 174.2 (2019), pp. 287–315 (cit. on p. 2).

[LY52] Tsung-Dao Lee and Chen-Ning Yang. "Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model". In: *Physical Review* 87.3 (1952), pp. 410–419 (cit. on p. 2).

[Rue71] David Ruelle. "Extension of the Lee-Yang circle theorem". In: *Physical Review Letters* 26.6 (1971), pp. 303–304 (cit. on p. 3).

[She85] James B Shearer. "On a Problem of Spencer". In: *Combinatorica* 5.3 (July 1985), pp. 241–245. ISSN: 0209-9683. DOI: 10.1007/BF02579368 (cit. on pp. 5, 10).

[SS05] Alexander D. Scott and Alan D. Sokal. "The Repulsive Lattice Gas, the Independent-Set Polynomial, and the Lovász Local Lemma". In: *Journal of Statistical Physics* 118.5 (2005), pp. 1151–1261 (cit. on pp. 5, 7, 10).

A Connections Between Independence Polynomials and the Lovász Local Lemma

Suppose we have a collection of events $A_1, \dots, A_n$, and we wish to certify that with positive probability, none of the events occur (i.e. $\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right] > 0$). We typically view $A_1, \dots, A_n$ as "bad events" we want to avoid, e.g. some desirable property fails for some random object of interest. If the events were all jointly independent, then of course $\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right] = \prod_{i=1}^n (1 - \Pr[A_i])$, which is positive if (and only if) none of the events individually happen with probability 1. However, in most applications, the events $A_1, \dots, A_n$ are not jointly independent. One could use the Union Bound to say that $\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right] = 1 - \Pr\left[\bigvee_{i=1}^n A_i\right] \geq 1 - \sum_{i=1}^n \Pr[A_i]$, but this requires $\sum_{i=1}^n \Pr[A_i] < 1$ which is often too restrictive. The famous *Lovász Local Lemma (LLL)* allows us to significantly "beat the Union Bound" assuming the collection of events $\{A_i\}_{i=1}^n$ are "not too dependent" on each other. It is a fundamental tool in discrete probability and theoretical computer science, and often goes hand-in-hand with the *probabilistic method*.

Definition 2 (Mutual Independence)**.** *Let $A_1, \dots, A_n$ be a collection of events within the same probability space. For each $i \in [n]$ and $J \subseteq [n] \setminus \{i\}$, we say A_i is* mutually independent *of*

$\{A_j : j \in J\}$ *if*

$$\Pr\left[A_i \wedge \bigwedge_{j\in J'} A_j\right] = \Pr[A_i] \cdot \Pr\left[\bigwedge_{j\in J'} A_j\right], \qquad \forall J' \subseteq J.$$

Definition 3 (Dependency Graph)**.** *Let $A_1, \ldots, A_n$ be a collection of events within the same probability space. We say an (undirected) graph $G = (V, E)$ with $V \cong [n]$ is a* dependency graph *for the events $\{A_i\}_{i=1}^n$ if for every $v \in V$, the event A_v is mutually independent of the events $\{A_u : u \neq v, u \not\sim v\}$.*

Remark 5. We emphasize that a collection of events $\{A_i\}_{i=1}^n$ need not have a *unique* dependency graph. For instance, the complete graph K_n is always a valid dependency graph w.r.t. any collection of events $\{A_i\}_{i=1}^n$, although this is not very useful since its vertices have large degree. One can also consider directed dependency graphs, but we will not do so here. For further discussion of these points, see e.g. [SS05].

The usual version of the Lovász Local Lemma states that if the events $\{A_i\}_{i=1}^n$

- admit a dependency graph of maximum degree $\leq d$, and
- $\Pr[A_i] \leq \frac{1}{e(d+1)}$ for all $i = 1, \ldots, n$,

then $\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right] > 0$. Since d could be much smaller than the total number of events n (e.g. $d \leq O(1)$), the second condition $\Pr[A_i] \leq \frac{1}{e(d+1)}$ is considered much less stringent than $\sum_{i=1}^n \Pr[A_i] < 1$, which is what the Union Bound requires. For us, we will be interested in the sharpest condition under which we can guarantee $\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right] > 0$, and more importantly, its connection with zero-freeness of the multivariate independence polynomial and Shearer's Condition.

Theorem A.1 (Shearer's Lemma; [She85])**.** *Let $G = (V, E)$ be a graph on n vertices, and let $\boldsymbol{p} \in (0, 1)^n$. Then the following are equivalent:*

(1) $Z_{G-N[I]}(-\boldsymbol{p}) > 0$ for every independent set $I \subseteq V$.

(2) $Z_{G[S]}(-\boldsymbol{p}) > 0$ for every $S \subseteq V$.

(3) $Z_G(-\lambda\boldsymbol{p}) > 0$ for every $\lambda \in [0, 1]$.

(4) For every collection of events $\{A_i\}_{i=1}^n$ in some common probability space such that

- *G is a valid dependency graph for $\{A_i\}_{i=1}^n$, and*
- *$\Pr[A_i] \leq p_i$ for all $i = 1, \ldots, n$,*

the conclusion of the Lovász Local Lemma holds, i.e. $\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right] > 0$.

Furthermore, if any of these conditions are satisfied, then for every such collection of events $\{A_i\}_{i=1}^n$, we have the lower bound

$$\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right] \geq Z_G(-\boldsymbol{p}) > 0.$$

Condition (2) is just Shearer's Condition from Theorem 2.1. Conditions (1) and (3) are essentially reformulations of (2) and stability of Z_G. The most interesting part is how these three zero-free conditions are all equivalent to (4), a purely probabilistic statement.

A.1 The Easy Equivalences Between (1), (2), and (3)

We begin by showing the equivalence between (1) and (2), which is straightforward. Clearly (2) implies (1) just by taking $S = V - N[I]$. The following lemma immediately shows that (1) implies (2).

Lemma A.2 (Inclusion-Exclusion for Independence Polynomials)**.** *For every $\boldsymbol{p} \in (0, 1)^V$ and $S \subseteq V$, we have the identity*

$$Z_{G[S]}(-\boldsymbol{p}) = \sum_{\substack{I \subseteq V \text{ indep.} \\ I \cap S = \emptyset}} \boldsymbol{p}^I \cdot Z_{G-N[I]}(-\boldsymbol{p}).$$

Proof.

$$\begin{aligned}
\sum_{\substack{I\subseteq V \text{ indep.}\\ I\cap S=\emptyset}} \boldsymbol{p}^I \cdot Z_{G-N[I]}(-\boldsymbol{p}) &= \sum_{\substack{I\subseteq V \text{ indep.}\\ I\cap S=\emptyset}} \sum_{\substack{J\subseteq V \text{ indep.}\\ I\subseteq J}} (-1)^{|J\setminus I|}\boldsymbol{p}^J \\
&= \sum_{J\subseteq V \text{ indep.}} \boldsymbol{p}^J \sum_{I\subseteq J\setminus S} (-1)^{|J\setminus I|} && \text{(Exchange order of summation)} \\
&= \sum_{J\subseteq S \text{ indep.}} (-1)^{|J|}\boldsymbol{p}^J && \text{(Nonzero contribution only if } J\setminus S=\emptyset) \\
&= Z_{G[S]}(-\boldsymbol{p}). && \text{(Definition of } Z_{G[S]})
\end{aligned}$$

□

This shows that (1) is equivalent to (2). (2) implies (3) simply by combining Shearer's Condition with Theorem 2.1. Indeed a special case of $\prod_{v\in V}\overline{\mathbb{D}(0,p_v)}$-stability is that $Z_G(-\lambda\boldsymbol{p}) \neq 0$ for all $\lambda \in [0,1]$. Since $Z_G(-\lambda\boldsymbol{p}) = 1$ at $\lambda = 0$, by continuity, it must be that $Z_G(-\lambda\boldsymbol{p}) > 0$ for all $\lambda \in [0,1]$. One could also have directly proven (2) implies (3) by adapting the inductive proof of Theorem 2.1.

Now we show that (3) implies (2). We establish the contrapositive. Define the *Shearer region* to be the set

$$\mathcal{S} \stackrel{\mathsf{def}}{=} \left\{\boldsymbol{q} \in \mathbb{R}^V_{\geq 0} : Z_{G[S]}(-\boldsymbol{q}) > 0, \forall S \subseteq V\right\}. \tag{5}$$

Since $\mathcal{S}$ is the preimage of an open set under a continuous function, it is open. Now, suppose (2) does not hold, i.e. $\boldsymbol{p} \notin \mathcal{S}$. Define

$$\tilde{\lambda} = \tilde{\lambda}(\boldsymbol{p}) \stackrel{\mathsf{def}}{=} \sup\left\{\lambda \in [0,1] : \lambda\boldsymbol{p} \in \mathcal{S}\right\}.$$

We show that $Z_G(-\tilde{\lambda}\boldsymbol{p}) = 0$. Since $\tilde{\lambda}\boldsymbol{p} \in \partial\mathcal{S}$, $Z_{G[S]}(-\tilde{\lambda}\boldsymbol{p}) \geq 0$ for all $S \subseteq V$ by continuity. It follows that

$$\begin{aligned}
Z_G(-\tilde{\lambda}\boldsymbol{p}) &= Z_{G-v}(-\tilde{\lambda}\boldsymbol{p}) - \tilde{\lambda}p_v Z_{G-N[v]}(-\tilde{\lambda}\boldsymbol{p}) && \text{(For all } v \in V) \\
&\leq Z_{G-v}(-\tilde{\lambda}\boldsymbol{p}) \\
&\leq \cdots && \text{(Induction)} \\
&\leq Z_{G[S]}(-\tilde{\lambda}\boldsymbol{p}),
\end{aligned}$$

for any $S \subseteq V$. Furthermore, because $\mathcal{S}$ is open, $\tilde{\lambda}\boldsymbol{p} \notin \mathcal{S}$, and so there exists $S \subseteq V$ such that $Z_{G[S]}(-\tilde{\lambda}\boldsymbol{p}) = 0$, whence $0 \leq Z_G(-\tilde{\lambda}\boldsymbol{p}) \leq Z_{G[S]}(-\tilde{\lambda}\boldsymbol{p}) = 0$. So, we indeed have $Z_G(-\tilde{\lambda}\boldsymbol{p}) = 0$ and (3) fails.

A.2 Shearer's Condition Implies LLL Conclusion

Proof of Theorem A.1: (2) $\implies$ (4). For convenience, we write $\mathcal{P}_S \stackrel{\mathsf{def}}{=} \Pr\left[\bigwedge_{i\in S}\overline{A_i}\right]$ for any $S \subseteq V \cong [n]$. We will prove that for every $S \subseteq V$, $\mathcal{P}_S \geq Z_{G[S]}(-\boldsymbol{p})$, the latter of which is positive by assumption. For this, it suffices to show that for every $S \subseteq V$ and every $r \in S$,

$$\frac{\mathcal{P}_S}{\mathcal{P}_{S-r}} \geq \frac{Z_{G[S]}(-\boldsymbol{p})}{Z_{G[S-r]}(-\boldsymbol{p})}. \tag{6}$$

Indeed, this implies $\mathcal{P}_S \geq Z_{G[S]}(-\boldsymbol{p}) \cdot \frac{\mathcal{P}_{S-r}}{Z_{G[S-r]}(-\boldsymbol{p})}$, which is at least $Z_{G[S]}(-\boldsymbol{p})$ by invoking the inductive hypothesis $\frac{\mathcal{P}_{S-r}}{Z_{G[S-r]}(-\boldsymbol{p})} \geq 1$. We prove Eq. (6) inductively by leveraging what is essentially the tree recursion for the hardcore model. Observe that

$$\begin{aligned}
\mathcal{P}_S &= \Pr\left[\bigwedge_{i\in S} \overline{A}_i\right] \\
&= \Pr\left[\bigwedge_{i\in S-r} \overline{A}_i\right] - \Pr\left[A_r \wedge \bigwedge_{i\in S-r} \overline{A}_i\right] \\
&\geq \Pr\left[\bigwedge_{i\in S-r} \overline{A}_i\right] - \Pr\left[A_r \wedge \bigwedge_{i\in S-N[r]} \overline{A}_i\right] \\
&= \Pr\left[\bigwedge_{i\in S-r} \overline{A}_i\right] - \Pr[A_r]\cdot \Pr\left[\bigwedge_{i\in S-N[r]} \overline{A}_i\right] && \text{(Mutual Independence)} \\
&\geq \mathcal{P}_{S-r} - p_r \cdot \mathcal{P}_{S-N[r]} && \text{(Using } \Pr[A_r] \leq p_r\text{)}
\end{aligned}$$

Hence, $\frac{\mathcal{P}_S}{\mathcal{P}_{S-r}} \geq 1 - p_r \cdot \frac{\mathcal{P}_{S-N[r]}}{\mathcal{P}_{S-r}}$. At the same time, we already saw that $\frac{Z_{G[S]}(-\boldsymbol{p})}{Z_{G[S-r]}(-\boldsymbol{p})} = 1 - p_r \cdot \frac{Z_{G[S-N[r]]}(-\boldsymbol{p})}{Z_{G[S-r]}(-\boldsymbol{p})}$. Hence, it suffices to use the inductive hypothesis to certify

$$\frac{Z_{G[S-N[r]]}(-\boldsymbol{p})}{Z_{G[S-r]}(-\boldsymbol{p})} \geq \frac{\mathcal{P}_{S-N[r]}}{\mathcal{P}_{S-r}}.$$

This inequality follows by combining the inductive hypothesis and the same telescoping trick as in the proof of Theorem 2.1. □

A.3 Tightness of Shearer's Condition for the LLL

Proof of Theorem A.1: (4) $\implies$ (2). We prove the contrapositive. Assume (2) fails for $\boldsymbol{p} \in \mathbb{R}^V_{\geq 0}$. We will construct a finite probability space (Ω, ν) and a collection of events $A_1, \ldots, A_n \subseteq \Omega$ with dependency graph G, $\Pr[A_v] \leq p_v$ for all $v \in V$, and $\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right] = 0$. As in the proof of the equivalence between (2) and (3), let $\tilde{\lambda} = \sup\{\lambda \in [0,1] : \lambda\boldsymbol{p} \in \mathcal{S}\}$ where $\mathcal{S}$ is Shearer's region defined as in Eq. (5). As we argued in that proof, we have $Z_{G[S]}(-\tilde{\lambda}\boldsymbol{p}) \geq 0$ for all $S \subseteq V$, with equality for $S = V$. For the rest of the proof, we replace $\tilde{\lambda}\boldsymbol{p}$ by $\boldsymbol{p}$ for notational convenience.

Applying Lemma A.2 with $S = \emptyset$, we see that

$$1 = \sum_{I\subseteq V \text{ indep.}} \boldsymbol{p}^I \cdot Z_{G-N[I]}(-\boldsymbol{p}).$$

Since $Z_{G-N[I]}(-\boldsymbol{p}) \geq 0$ for all independent sets $I \subseteq V$, we have a probability distribution over all subsets of vertices $\Omega = 2^V$ given by

$$\nu(S) = \begin{cases} \boldsymbol{p}^S \cdot Z_{G-N[S]}(-\boldsymbol{p}), & \text{if } S \text{ is independent} \\ 0, & \text{otherwise} \end{cases}.$$

We then take A_v to be the event that v is in a random independent set sampled from ν.

Claim A.3. *For every $S \subseteq V$, the* marginal probability *of S is given by*

$$\Pr_{I\sim\nu}[S \subseteq I] = \begin{cases} \boldsymbol{p}^S, & \textit{if } S \textit{ is independent} \\ 0, & \textit{otherwise} \end{cases}.$$

This follows from the same inclusion-exclusion argument as in the proof of Lemma A.2. Claim A.3 directly implies mutual independence of v from $V \setminus N[v]$ for all $v \in V$ in the sense of Definition 2. This establishes that G is a valid dependency graph for $\{A_v\}_{v\in V}$. We also immediately get the upper bounds $\Pr[A_v] \leq p_v$ for all $v \in V$ (additionally using $\tilde{\lambda} \in [0,1]$). All that remains is to show

$\Pr\left[\bigwedge_{i=1}^{n} \overline{A}_i\right] = 0$. For this, observe that for every $S \subseteq V$,

$$\begin{aligned}
\Pr\left[\bigwedge_{i\in S} \overline{A}_i\right] &= \sum_{I\subseteq S} (-1)^{|I|} \cdot \Pr\left[\bigwedge_{i\in I} A_i\right] && \text{(Inclusion-Exclusion)}\\
&= \sum_{I\subseteq S \text{ indep.}} (-1)^{|I|} \boldsymbol{p}^I && \text{(Claim A.3)}\\
&= Z_{G[S]}(-\boldsymbol{p}). && \text{(Definition of } Z_{G[S]}\text{)}
\end{aligned}$$

Taking $S = V$ and using our assumption that $Z_G(-\boldsymbol{p}) = 0$ completes the proof. $\square$