Lecture 8: Barvinok's Polynomial Interpolation Algorithm

October 3, 2023

In the next few lectures, we explore a very different type of algorithm based on complex analysis and *zero-freeness* of the partition function, viewed as a polynomial. We'll further see a different perspective on phase transitions based on clustering of the zeros. For this lecture, the flagship application of this technique is the following: Recall that for a complex matrix $A \in \mathbb{C}^{n\times n}$, its *permanent* is defined as

$$\mathrm{per}(A) \stackrel{\mathsf{def}}{=} \sum_{\sigma \in S_n} \prod_{i=1}^{n} A_{i,\sigma(i)},$$

which can be viewed as a multivariate polynomial in the entries of A; here, S_n denotes the group of permutations on $[n]$.

Theorem 0.1 ([Bar16b]; see also [Bar17; Bar16a])**.** *Let $0 < \eta < 1/2$ be arbitrary. If $A \in \mathbb{C}^{n\times n}$ satisfies*

$$|A_{ij} - 1| < \eta, \qquad \forall i, j \in [n],$$

then for every $\epsilon > 0$, there exists a multivariate polynomial $q = q_{n,\eta,\epsilon}$ of degree $O_\eta(\log(n/\epsilon))$ such that $|q(A) - \log \mathrm{per}(A)| \le \epsilon$. Furthermore, q can be computed in deterministic quasipolynomial time, and so there exists a deterministic quasipolynomial time algorithm for estimating $\mathrm{per}(A)$ *up to $(1 \pm \epsilon)$-multiplicative error.*

Here, *quasipolynomial* in n means a function of the form $2^{\mathsf{polylog}(n)}$, and $O_\eta(\cdot)$ means the constant depends on some unspecified function of η.[1] This theorem asserts that $\log \mathrm{per}(A)$ can be approximated by a low-degree polynomial if A is not too far away from the all-ones matrix $\mathbf{1}\mathbf{1}^\top$, for which $\mathrm{per}\left(\mathbf{1}\mathbf{1}^\top\right) = n!$ trivially. Note that an ϵ-additive approximation to $\log \mathrm{per}(A)$ is equivalent to $e^{\pm\epsilon} \approx 1 \pm \epsilon$ multiplicative approximation to $\mathrm{per}(A)$.

The general algorithmic framework we'll use here is the elegant *polynomial interpolation algorithm* pioneered by Barvinok. Theorem 0.1 is essentially the best known algorithm for estimating the permanent of a matrix with possibly negative or complex entries. In the very special case $A \in \mathbb{R}^{n\times n}_{\ge 0}$, we can do much better using Markov chain Monte Carlo methods [JSV04]. Later on, we'll see that for partition functions coming from graphs (e.g. the Ising model), we can modify the algorithm to run in polynomial time with exponent depending on the maximum degree Δ, similar to the running time of the correlation decay algorithm.

1 The Taylor Series Approach to Approximate Counting

The basic idea behind the approach to Theorem 0.1 is to estimate $\log \mathrm{per}(A)$ using Taylor series. Recall that for a real smooth univariate function $f : \mathbb{R} \to \mathbb{R}$ (e.g. some univariate restriction of $\log \mathrm{per}(A)$) with derivatives $f^{(k)}(x) \stackrel{\mathsf{def}}{=} \frac{d^k}{dx^k} f(x)$ and Taylor series $f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \cdot x^k$ converging in a neighborhood $(-r, r)$ of 0, the truncation $\sum_{k=0}^{m} \frac{f^{(k)}(0)}{k!} \cdot x^k$ approximates $f(x)$ uniformly in the interval $[-(r-\eta), r-\eta] \subsetneq (-r, r)$ up to additive error $C_{\eta,m}|x|^{m+1}$ where $C_{\eta,m} = \frac{\sup_{|y|\le r-\eta} \left|f^{(m+1)}(y)\right|}{(m+1)!}$ is some constant depending on η. This error decays exponentially as we use larger and larger degree truncations. If the rate of exponential decay is a constant, then for

[1] Typically, we use this when the function of η is too cumbersome to write down explicitly, or if it is given implicitly via e.g. continuity arguments.

$m \le O_\eta(\log(n/\epsilon))$ where $O_\eta(\cdot)$ depends on $C_{\eta,m}$, we get an ϵ-additive approximation to f uniformly in the interval $[-(r-\eta), r-\eta]$.

However, it could be that the rate of exponential decay is extremely slow e.g. when $\eta \le O(1/n)$ and x is within $O(1/n)$ of the boundary of the region of convergence $(-r, r)$. For such "bad points" x, $f(x)$ isn't "sufficiently smooth" and we need polynomials of much higher degree to maintain the same level of approximation. In our setting, since the function f under consideration is the logarithm of a (multivariate) polynomial, these "bad points" are precisely the points where $\mathrm{per}(A) \approx 0$. Conversely, we'll show that if the input A is bounded away from the zeros of $\mathrm{per}(A)$, then we can indeed obtain accurate approximations via low-degree Taylor series. We emphasize that this is a *totally generic* scheme for approximate counting: If there is a complicated partition function Z we want to compute, then by viewing it as a polynomial in some natural underlying parameter (e.g. inverse temperature β for the Ising model, or fugacity λ for the hardcore model), we can approximate Z using the Taylor series of $\log Z$ whenever the variable is bounded away from the zeros of Z.

Proposition 1.1 ([Bar17])**.** *Let $g : \mathbb{C} \to \mathbb{C}$ be a degree-d univariate complex polynomial with roots $\zeta_1, \ldots, \zeta_d \in \mathbb{C}$, and suppose there exists a radius $R > 1$ such that $|\zeta|_i > R$ for all $i = 1, \ldots, d$. If we write $f(z) = \log g(z)$ and define*

$$T_m(z) \overset{\text{def}}{=} \sum_{k=0}^{m} \frac{f^{(k)}(0)}{k!} z^k, \qquad \forall m \in \mathbb{N},$$

then

$$\frac{f^{(k)}(0)}{k!} = -\frac{1}{k}\sum_{i=1}^{d} \zeta_i^{-k}, \qquad \forall k \in \mathbb{N},$$
$$|f(z) - T_m(z)| \le \frac{d}{(m+1)R^m(R-1)}, \qquad \forall\, |z| \le 1, \forall m \in \mathbb{N}.$$

Remark 1. In Section 3, we will extend this result to much more general types of regions in $\mathbb{C}$, beyond simply scaled disks.

Proof. We may express f, g as $g(z) = g(0)\prod_{i=1}^{d}\left(1 - \frac{z}{\zeta_i}\right)$ and $f(z) = f(0) + \sum_{i=1}^{d} \log\left(1 - \frac{z}{\zeta_i}\right)$. Note this representation is valid since $|\zeta_i| > R > 1$ for all $i = 1, \ldots, d$ by assumption. Applying the Taylor series of $z \mapsto \log(1-z)$, which converges in the entire unit disk $\mathbb{D} = \mathbb{D}(0, 1)$, we have

$$\log\left(1 - \frac{z}{\zeta_i}\right) = -\sum_{k=1}^{m} \frac{1}{k} \cdot \left(\frac{z}{\zeta_i}\right)^k + \epsilon_{i,m},$$

where the error term $\epsilon_{i,m} \in \mathbb{C}$ satisfies

$$\begin{aligned}
|\epsilon_{i,m}| &= \left|\sum_{k=m+1}^{\infty} \frac{1}{k} \cdot \left(\frac{z}{\zeta_i}\right)^k\right| && \text{(By definition)}\\
&\le \sum_{k=m+1}^{\infty} \frac{1}{k} \cdot \left|\frac{z}{\zeta_i}\right|^k && \text{(Triangle Inequality)}\\
&\le \frac{1}{m+1} \sum_{k=m+1}^{\infty} |\zeta_i|^{-k} && \text{(Using } |z| \le 1\text{)}\\
&\le \frac{1}{(m+1)R^m(R-1)}. && \text{(Using } |\zeta_i| > R > 1\text{)}
\end{aligned}$$

Note that $\frac{f^{(k)}(0)}{k!} = -\frac{1}{k}\sum_{i=1}^{d} \zeta_i^{-k}$ for all $k \in \mathbb{N}$ by linearity. Furthermore, $T_m(z) = f(0) - \sum_{i=1}^{d}\sum_{k=1}^{m} \frac{1}{k} \cdot \left(\frac{z}{\zeta_i}\right)^k$, and $|f(z) - T_m(z)| \le \left|\sum_{i=1}^{d} \epsilon_{i,m}\right| \le \frac{d}{(m+1)R^m(R-1)}$. □

Note that to use Proposition 1.1, we need to be able to compute the derivatives $f^{(k)}(0)$ of $f = \log g$. It turns out we can do this easily assuming we have access to the derivatives of g itself. This is the content of the following lemma. The latter is still a somewhat nontrivial task, but we'll see later how the derivatives of g itself have nice combinatorial interpretations.

Lemma 1.2 (Derivatives of $\log g$)**.** *Let $z \in \mathbb{C}$ be arbitrary. Given the first $m+1$ derivatives $\{g^{(k)}(z)\}_{k=0}^{m}$ of g, we can compute the first $m+1$ derivatives $\{f^{(k)}(z)\}_{k=0}^{m}$ of $f = \log g$ in time $O(m^2)$.*

Proof. Observe that for any $f^{(1)}(z) = \frac{g^{(1)}(z)}{g^{(0)}(z)}$, or equivalently, $g^{(1)}(z) = f^{(1)}(z) \cdot g^{(0)}(z)$. By differentiating the latter identity, we have inductively that

$$g^{(k)}(z) = \sum_{j=0}^{k-1} \binom{k-1}{j} \cdot f^{(k-j)}(z) \cdot g^{(j)}(z), \qquad \forall k \in \mathbb{N}.$$

After rearranging appropriately, we see that each $f^{(k)}(z)$ can be computed from the previously computed coefficients $f^{(0)}(z), \ldots, f^{(k-1)}(z)$ and the given coefficients $g^{(0)}(z), \ldots, g^{(k)}(z)$. Each of these computations takes roughly $O(k)$-time, and so summing over all $0 \le k \le m$ yields an $O(m^2)$-time algorithm. $\square$

2 Zero-Freeness of the Permanent

In this section, we prove Theorem 0.1 using Proposition 1.1 and Lemma 1.2. The most nontrivial task is establishing an appropriate zero-free region.

Theorem 2.1 ([Bar17])**.** *There exists a universal constant $\eta_0 \ge 1/2$ such that if $A \in \mathbb{C}^{n\times n}$ satisfies $|A_{ij} - 1| \le \eta_0$ for all $i, j \in [n]$, then $\mathrm{per}(A) \neq 0$.*

Before we prove this theorem, we use it to finish the proof of Theorem 0.1.

Proof of Theorem 0.1. Fix $A \in \mathbb{C}^{n\times n}$ such that $|A_{ij} - 1| \le \eta$ for all $i, j \in [n]$, where $\eta < 1/2 < \eta_0$ is some constant. Define the univariate polynomial

$$g(z) \stackrel{\mathsf{def}}{=} \mathrm{per}\left(zA + (1-z)\mathbf{1}\mathbf{1}^\top\right) = \mathrm{per}\left(\mathbf{1}\mathbf{1}^\top + z\left(A - \mathbf{1}\mathbf{1}^\top\right)\right).$$

By Theorem 2.1, the polynomial g satisfies $g(z) \neq 0$ for all $|z| \le R \stackrel{\mathsf{def}}{=} \frac{\eta_0}{\eta}$. Hence, by Proposition 1.1, we have that for $m \le O\left(\frac{1}{1-(\eta/\eta_0)} \log(n/\epsilon)\right)$, the order-$m$ Taylor approximation T_m to $f = \log g$ satisfies $|T_m(z) - \log g(z)| \le O(\epsilon)$, or equivalently, $(1-\epsilon)g(z) \le \exp(T_m(z)) \le (1+\epsilon)g(z)$.

All that remains to show is that we can indeed compute $T_m(z)$ in deterministic quasipolynomial time. By Lemma 1.2, it suffices to show that we can compute all coefficients $\{g^{(k)}(0)\}_{k=0}^{m}$ up to $m \le O\left(\frac{1}{1-(\eta/\eta_0)} \log(n/\epsilon)\right)$ in deterministic quasipolynomial time. We do this by brute force. Observe that

$$\begin{aligned} g^{(k)}(z) &= \sum_{\sigma \in S_n} \frac{d^k}{dz^k} \prod_{i=1}^{n} \left(1 + z(A_{i,\sigma(i)} - 1)\right) \\ &= \sum_{\sigma \in S_n} \sum_{S \subseteq [n]} \prod_{i \in S} (A_{i,\sigma(i)} - 1) \cdot \frac{d^k}{dz^k} z^{|S|}. \end{aligned}$$

It follows that

$$\begin{aligned} g^{(k)}(0) &= k! \cdot \sum_{\sigma \in S_n} \sum_{S \in \binom{[n]}{k}} \prod_{i \in S} (A_{i,\sigma(i)} - 1) \\ &= k!(n-k)! \cdot \sum_{S \in \binom{[n]}{k}} \sum_{\substack{\sigma : S \to [n] \\ \text{injective}}} \prod_{i \in S} (A_{i,\sigma(i)} - 1). \end{aligned}$$

Note the second identity follows from the fact that we only care about the image of $S \in \binom{[n]}{k}$ under σ. Hence, we can just sum over the possible images, i.e. injective maps $\sigma : S \to [n]$, picking up a multiplicative factor of $(n-k)!$ for the number of possible extensions of σ into a full permutation of $[n]$. For each k, the final expression above is a summation of at most $n^{O(k)}$ terms, and so we can compute $g^{(k)}(0)$ deterministically in $n^{O(k)}$-time via brute force enumeration. Since we're only asking for the first $m \le O_\eta(\log(n/\epsilon))$ terms, this amounts to a quasipolynomial time algorithm. $\square$

All that remains is to establish the zero-free region guaranteed by Theorem 2.1. Before we do this, we make a few remarks. The logarithmic bound on the number of coefficients needed is essentially the best possible. Hence, the main bottleneck preventing us from getting a truly polynomial-time algorithm is computation of the coefficients $\{g^{(k)}(0)\}_{k=0}^{m}$. In most applications like the permanent, we can do this in quasipolynomial time. Later, we will see that for graph polynomials, we can reduce the computation cost to a polynomial with exponent depending on the maximum degree Δ of the input graph. We emphasize that in this framework, the most challenging part is establishing a suitable zero-free region.

2.1 Zero-Freeness via Barvinok's Inductive Angle Method

For the proof of Theorem 2.1, we use a clever inductive argument due to Barvinok [Bar17]. For each $j \in [n]$, we write $A_j \in \mathbb{C}^{(n-1)\times(n-1)}$ for the submatrix of A obtained by deleting the *first* row and the jth column of A. Then

$$\mathrm{per}(A) = \sum_{j=1}^{n} A_{1,j} \cdot \mathrm{per}(A_j).$$

Clearly, each A_j inherits the property $\left\|A_j - \mathbf{1}_{n-1}\mathbf{1}_{n-1}^\top\right\|_\infty \le \eta_0$ from $\left\|A - \mathbf{1}_n\mathbf{1}_n^\top\right\|_\infty \le \eta_0$, where for convenience we write $\|\cdot\|_\infty$ for the maximum entry in absolute value.[2] Now, roughly speaking, there are two reasons that $\mathrm{per}(A)$ could be close to 0. Either

- the numbers $\mathrm{per}(A_j)$ themselves are close to 0, or
- $\{\mathrm{per}(A_j)\}_{j=1}^n$ as complex numbers (or vectors in $\mathbb{R}^2$) all point in different directions which "cancel" each other, so that $\sum_{i=1}^n A_{1,j} \cdot \mathrm{per}(A_j) \approx 0$ even if individually each $\mathrm{per}(A_j)$ is far from 0.

This second case obstructs us from using $\mathrm{per}(A) \ne 0$ directly as the inductive hypothesis, even though this is the most "obvious" choice. To get around this, we strengthen the inductive hypothesis. We will impose that for any two matrices A, B which differ in at most one row or column, and which satisfy $\left\|A - \mathbf{1}_n\mathbf{1}_n^\top\right\|_\infty, \left\|B - \mathbf{1}_n\mathbf{1}_n^\top\right\|_\infty \le \eta_0$, we have that the angle between $\mathrm{per}(A), \mathrm{per}(B) \in \mathbb{C}$ does not exceed some threshold θ; later, we will see that we can take $\theta = \pi/2$. The key technical lemma which will make this argument go through is the following.

Lemma 2.2 ([Bar17]). *Let $u_1, \dots, u_n \in \mathbb{C}$ be nonzero complex numbers such that for some $0 < \theta < 2\pi/3$, we have $\angle(u_i, u_j) \le \theta$ for all $i, j \in [n]$. Let $a_1, \dots, a_n$ be complex numbers satisfying $\|\boldsymbol{a} - \mathbf{1}_n\|_\infty \le \eta_0$. Then we have:*

- $\sum_{i=1}^n a_i u_i \ne 0$*, and*
- $\angle\left(\sum_{i=1}^n u_i, \sum_{i=1}^n a_i u_i\right) \le \arcsin\left(\frac{\eta_0}{\cos(\theta/2)}\right)$.

Before we prove this lemma, we first use it to complete the proof of the zero-free region for the permanent.

Proof of Theorem 2.1. We prove the claim by induction on n. Let $0 < \theta < 2\pi/3$ be a parameter to be determined later. For each $n \in \mathbb{N}$, let $\mathrm{IH}(n)$ denote the following statement:

(1) For every complex matrix $A \in \mathbb{C}^{n\times n}$ satisfying $\left\|A - \mathbf{1}_n\mathbf{1}_n^\top\right\|_\infty \le \eta_0$, $\mathrm{per}(A) \ne 0$.

(2) For every pair of complex matrices $A, B \in \mathbb{C}^{n\times n}$ which differ in at most one row/column, and which satisfy $\left\|A - \mathbf{1}_n\mathbf{1}_n^\top\right\|_\infty \le \eta_0$, we have $\angle(\mathrm{per}(A), \mathrm{per}(B)) \le \theta$.

We first prove $\mathrm{IH}(n)$ assuming $\mathrm{IH}(n-1)$, since this induction step is the most interesting and nontrivial. Let $A \in \mathbb{C}^{n\times n}$ satisfy $\left\|A - \mathbf{1}_n\mathbf{1}_n^\top\right\|_\infty \le \eta_0$. By (2) for $\mathrm{IH}(n-1)$, we may apply the first conclusion of Lemma 2.2 with $\{u_j\}_{j=1}^n = \{\mathrm{per}(A_j)\}_{j=1}^n$ and $\{a_j\}_{j=1}^n = \{A_{1,j}\}_{j=1}^n$ to deduce $\mathrm{per}(A) = \sum_{j=1}^n A_{1,j} \cdot \mathrm{per}(A_j) \ne 0$.

[2] In other words, we view the input matrix as a vector with n^2-dimensions, and take its ℓ_∞-norm. Note this is not the same as the matrix norm induced by the ℓ_∞ vector norm.

Now, let $A, B \in \mathbb{C}^{n\times n}$ satisfy $\left\|A - \mathbf{1}_n \mathbf{1}_n^\top\right\|_\infty, \left\|B - \mathbf{1}_n \mathbf{1}_n^\top\right\|_\infty \leq \eta_0$ and differ in at most one row/column; without loss of generality, we can assume A, B differ in the first row. Then

$$\begin{aligned}\angle\left(\mathrm{per}(A), \mathrm{per}(B)\right) \leq \angle &\left(\sum_{i=1}^n A_{1,j} \cdot \mathrm{per}(A_j), \sum_{i=1}^n \mathrm{per}(A_j)\right) \\ &+ \angle\left(\sum_{i=1}^n \mathrm{per}(A_j), \sum_{i=1}^n \mathrm{per}(B_j)\right) \\ &+ \angle\left(\sum_{i=1}^n \mathrm{per}(B_j), \sum_{i=1}^n B_{1,j} \cdot \mathrm{per}(B_j)\right).\end{aligned}$$

Note the second term is zero since $\mathrm{per}(A_j) = \mathrm{per}(B_j)$ for all j by the assumption that A, B differ only in the first row. Furthermore, the third term is the same as the first term except A is replaced by B. By the second conclusion of Lemma 2.2, we get the first term is upper bounded by $\arcsin\left(\frac{\eta_0}{\cos(\theta/2)}\right)$ and so we get

$$\angle\left(\mathrm{per}(A), \mathrm{per}(B)\right) \leq 2 \arcsin\left(\frac{\eta_0}{\cos(\theta/2)}\right).$$

Choosing $\theta = \pi/2$ and $\eta_0 = 1/2$, we get the right-hand side is exactly θ. This establishes $\mathrm{IH}(n)$ given $\mathrm{IH}(n-1)$.

For this choice of $\theta = \pi/2$ and $\eta_0 = 1/2$, we finish the proof by establishing the base case $\mathrm{IH}(1)$. When $n = 1$, the first claim (1) is obvious. (2) says that for any $a, b \in \mathbb{C}$ satisfying $|z-1|, |w-1| \leq 1/2$, we have $\angle(z, w) \leq \pi/2$. For this, it suffices to show that for any $z \in \mathbb{C}$ satisfying $|z - 1| \leq 1/2$, we have $|\arg(z)| \leq \pi/4$. If we write $z = a + bi$ for $a, b \in \mathbb{R}$, then $\tan(\arg(z)) = \frac{b}{a}$, so $|\arg(z)| \leq \pi/4$ is equivalent to requiring $|b| \leq |a|$. For this, observe that $|z-1| \leq 1/2$ translates to $(a-1)^2 + b^2 \leq 1/4$; in particular, we must have $|a - 1| \leq 1/2$ and $|b| \leq 1/2$ so $|b| \geq |a|$ as desired. $\square$

We finally prove the angle lemma.

Proof of Lemma 2.2. For convenience, write $u = \sum_{i=1}^n u_i$ and $v = \sum_{i=1}^n a_i u_i$. First, we claim the following "reverse Triangle Inequality"

$$|u| \geq \cos\left(\frac{\theta}{2}\right) \sum_{i=1}^n |u_i|. \tag{1}$$

If we have this, then

$$|u - v| = \left|\sum_{i=1}^n (a_i - 1) \cdot u_i\right| \leq \eta_0 \sum_{i=1}^n |u_i| < |u|,$$

where in the last step, we used Eq. (1) and the assumption $\eta_0 \leq \cos(\theta/2)$. It immediately follows by the Triangle Inequality that

$$|v| \geq |u| - |u - v| > 0,$$

i.e. $v \neq 0$, establishing the first item. Furthermore, if we view $u, v \in \mathbb{C}$ instead as vectors in $\mathbb{R}^2$, then since $|u - v| < |u|$,

$$\sin\left(\angle(u, v)\right) = \frac{\left|u - \frac{\langle u, v\rangle}{|v|^2} \cdot v\right|}{|u|} \leq \frac{|u - v|}{|u|} \leq \frac{\eta_0}{\cos(\theta/2)},$$

from which the second item follows. Hence, all that remains is to prove Eq. (1). To do this, we will construct a suitable vector $y \in \mathbb{C}$ onto which we will project $u_1, \ldots, u_n$.

First, we claim that there is a convex cone $K \subseteq \mathbb{C} \cong \mathbb{R}^2$ of angle θ and vertex at 0, i.e. $K = \mathsf{cone}(z, w)$ where its extremal rays z, w satisfy $\angle(z, w) \leq \theta$, such that $u_1, \ldots, u_n \in K$. This crucially uses the fact that all pairwise angles between the $u_1, \ldots, u_n$ are at most θ, and that $\theta < 2\pi/3$; if $\theta = 2\pi/3$, we can violate this containment by a small-aperture convex cone since we could take $u_1, \dots, u_n$ to be a bunch of copies of the third primitive roots of unity $1, e^{2\pi i/3}, e^{4\pi i/3}$. To see this, we just need to show that $0 \notin \mathsf{conv}(u_1, \dots, u_n)$. Suppose for contradiction that $0 \in \mathsf{conv}(u_1, \dots, u_n)$. Then by Carathéodory's Theorem, there exists distinct i, j, k such that $0 \in \mathrm{conv}(u_i, u_j, u_k)$. But then it must be that two of the vectors u_i, u_j, u_k have angle at least $2\pi/3$ from each other, a contradiction.

Thus, we have shown that $u_1, \dots, u_n \in K = \mathsf{cone}(z, w)$ for $z, w \in \mathbb{C}$ satisfying $\angle(z, w) \leq \theta$. Take y to be the *bisector* of K, i.e. the (unit) vector in $\mathbb{C}$ with the same angle as that of $\frac{z+w}{2}$. Since $u_1, \dots, u_n \in K$, we have $\angle(u_i, y) \leq \theta/2$, whence $\cos(\theta/2) \leq \cos(\angle(u_i, y)) = \frac{\langle u_i, y\rangle}{|u_i|}$. It follows that $\langle u_i, y\rangle \geq \cos(\theta/2) \cdot |u_i|$ for all $i \in [n]$ and

$$|u| \geq \langle u, y\rangle = \sum_{i=1}^{n} \langle u_i, y\rangle \geq \cos\left(\frac{\theta}{2}\right) \sum_{i=1}^{n} |u_i|$$

as desired. □

3 Beyond Zero-Free Disks

The method from Proposition 1.1 requires that the input polynomial g is zero-free within an entire disk of radius $R > 1$ around 0, assuming we can easily compute $g(0)$. In practice, we oftentimes cannot hope for such a zero-free disk containing both the point at which we wish to (approximately) compute g, and the point at which evaluating g is easy. For example, for the hardcore model, we have the following theorem due to Peters–Regts [PR19], originally conjectured by Sokal [Sok01].

Theorem 3.1 ([PR19])**.** *Let $Z_G(\lambda) = \sum_{I \subseteq V} \lambda^{|I|}$ denote the univariate independence polynomial of a graph G, where the summation is only over independent sets $I \subseteq V$. Then for every $\Delta \in \mathbb{N}$, there exists $\delta > 0$ such that $Z_G(\lambda) \neq 0$ for all λ in the strip*

$$\{z \in \mathbb{C} : \exists w \in [0, \lambda_c(\Delta)) \text{ s.t. } |z - w| < \delta\}.$$

We refer interested readers to [Ben+23] for the current state-of-the-art on the zeros of the independence polynomial on bounded-degree graphs. We will not prove Theorem 3.1, only using it as an (important) illustrative example. The following result allows us to handle zero-free regions which are strips.

Proposition 3.2. *Let $g : \mathbb{C} \to \mathbb{C}$ be a degree-d univariate complex polynomial, and suppose there exists $\delta > 0$ such that g is nonzero in a δ-strip around the interval $[0, 1]$, i.e. $g(z) \neq 0$ for all $z \in \mathbb{C}$ satisfying $-\delta \leq \operatorname{Re} z \leq 1 + \delta$ and $|\operatorname{Im} z| \leq \delta$. Then for every $\epsilon > 0$, there exists a degree-m polynomial T_m with $m \leq O_\delta(\log(n/\epsilon))$ such that*

$$|T_m(z) - \log g(z)| \leq \epsilon$$

uniformly in the $\frac{\delta}{2}$-strip around $[0, 1]$. Furthermore, the $m + 1$ coefficients of T_m can be efficiently computed given the numbers $\{g^{(k)}(0)\}_{k=0}^{m}$.

The key technical lemma we will need is the following, which constructs a polynomial map from the unit disk to a strip around $[0, 1]$.

Lemma 3.3. *For every $0 < \delta < 1$, define*

$$\varphi_\delta(z) \stackrel{\mathsf{def}}{=} \frac{1}{\sigma} \sum_{k=1}^{N} \frac{(\alpha z)^k}{k},$$

where $N \leq O(e^{1/\delta})$ is the degree, $\alpha \stackrel{\mathsf{def}}{=} 1 - \exp(-1/\delta) < 1$, and $\sigma \stackrel{\mathsf{def}}{=} \sum_{k=1}^{N} \frac{\alpha^k}{k}$ ensures $\varphi(1) = 1$. Then for every $z \in \mathbb{C}$ satisfying $|z| \leq \frac{1 - e^{-1-(1/\delta)}}{1 - e^{-1/\delta}} \stackrel{\mathsf{def}}{=} R$, we have

$$-\delta \leq \operatorname{Re} \varphi_\delta(z) \leq 1 + 2\delta$$
$$|\operatorname{Im} \varphi_\delta(z)| \leq 2\delta.$$

Note that $R > 1$ and $\varphi_\delta(0) = 0$.

Proof Sketch. We only provide a sketch of the proof; see [Bar17] for more details. Consider the function $f_\delta(z) \stackrel{\text{def}}{=} -\delta \log(1-\alpha z)$ defined on the closed unit disk $\overline{\mathbb{D}(0,1)}$. Clearly, $f_\delta(0) = 0$ and $f_\delta(1) = 1$. It is straightforward to directly check that for all z satisfying $|z| \leq 1$,

$$\begin{aligned} -\delta \log\left(2 \operatorname{Re} f_\delta(z)\right) &\leq 1+\delta \\ |\operatorname{Im} f_\delta(z)| &\leq \pi\delta/2. \end{aligned}$$

Since φ_δ is the degree-N Taylor polynomial for f_δ (suitably renormalized by σ) an argument similar to Proposition 1.1 which accounts for the additional approximation error completes the proof. $\square$

Proof Sketch of Proposition 3.2. Let $G(z) = g(\varphi_\delta(z))$, where φ_δ is furnished by Lemma 3.3. Then G is a degree-d polynomial with roots have magnitude at least $R > 1$, where R is as in Lemma 3.3 and $d = \deg(g) \cdot \deg(\varphi_\delta) \leq O(e^{1/\delta}) \cdot \deg(g)$. Let T_m denote the degree-m Taylor approximation to $\log G$ with $m \leq O_\delta(\log(n/\epsilon))$. Proposition 1.1 gives us the desired approximation result. The $m+1$ coefficients of T_m can be computed in $O(m^2)$-time from the numbers $\{G^{(k)}(0)\}_{k=0}^m$ by Lemma 1.2. Furthermore, since g, φ_δ are polynomials, so is G and $k! \cdot G^{(k)}(0)$ gives the coefficient of z^k in G. Clearly this can be computed given access to the first m coefficients of g and φ_δ as polynomials, which are given by $\{k! \cdot g^{(k)}(0)\}_{k=0}^m$ and $\{k! \cdot \varphi_\delta^{(k)}(0)\}_{k=0}^m$, respectively. $\square$

References

[Bar16a] Alexander Barvinok. *Combinatorics and Complexity of Partition Functions.* Vol. 30. Springer Algorithms and Combinatorics, 2016 (cit. on p. 1).

[Bar16b] Alexander Barvinok. "Computing the Permanent of (Some) Complex Matrices". In: *Found. Comput. Math.* 16.2 (Apr. 2016), pp. 329–342. ISSN: 1615-3375 (cit. on p. 1).

[Bar17] Alexander Barvinok. "Approximating permanents and hafnians". In: *Discrete Analysis* 2 (2017) (cit. on pp. 1–4, 7).

[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. 6).

[JSV04] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. "A Polynomial-Time Approximation Algorithm for the Permanent of a Matrix with Nonnegative Entries". In: *Journal of the ACM* 51.4 (July 2004), pp. 671–697 (cit. on p. 1).

[PR19] Han Peters and Guus Regts. "On a conjecture of Sokal concerning roots of the independence polynomial". In: *The Michigan Mathematical Journal* 68.1 (2019), pp. 33–55 (cit. on p. 6).

[Sok01] Alan D. Sokal. "A Personal List of Unsolved Problems Concerning Lattice Gases and Antiferromagnetic Potts Models". In: *Markov Processes And Related Fields* 7 (1 2001), pp. 21–38 (cit. on p. 6).