Counting and Sampling **Fall 2017**

Lecture 17: D-Stable Polynomials and Lee-Yang Theorems

Lecturer: Shayan Oveis Gharan *November 29th*

17.1 Course Overview

Let us start this lecture by giving a high-level overview of different techniques that we talked about in this course.

We started the course by studying Markov chains. We introduced the coupling method and used it to give very good bounds on colorings. Later, we introduced the path technology and we used it to estimate the matching polynomial of *any graph* any any t which is a polynomial of n.

In the second part of the course so far we have studied the correlation decay method and the polynomial approximation technique. We used the correlation decay method to estimate the matching polynomial on bounded degree graphs for a bounded value of t. In the last two lectures we also showed that the polynomial approximation technique can be used to prove almost the same thing.

We also discussed the Weitz result on estimating the independent set polynomial of bounded degree graph up to the NP-hardness threshold. In Assignment 2 we will show that

Counting vs Sampling We called this course counting and sampling because of the equivalence theorem that we proved at the beginning of this course. All results that we proved using MCMC technique naturally give (approximate) sampler from our desired probability distributions. On the other hand, all results in the second part of the course naturally estimate the corresponding partition function of the probability distribution that we are interested in. But because of the equivalence theorem we can sample using the counting algorithms of the 2nd part of the course and we can count using the approximate samplers of the first part of the course.

Local Methods: Coupling, Correlation Decay and Polynomial Approximation It turns out that all these three methods are "local" and the set of problems that we can study using them are quite related in the sense that we can typically use these methods in computing partition functions in bounded degree graphs where the probability that a vertex (or an edge) is chosen in the sample is almost independent of any conditioning far away. This was the main idea in correlation decay technique.

In the Polynomial approximation technique at the very high level we only use the first few coefficients of a degree n polynomial to estimate it. Of course these low order coefficients mainly depend on local structures. For example, in the case of matching polynomial we observed that the power sum of the roots of the polynomial correspond to the number of closed tree-like-walks, and more general the method of Patel and Regts exhibit the same phenomenon for a more general class of polynomials.

This pattern may be less obvious to see in the coupling method. But with a little bit of thinking one can see that indeed in the path coupling method we only care about local neighborhood of a vertex independent of any structure in the rest of the current sample.

Problem	Coupling	Path Technology	Correlation Decay	Polynomial Approximation
q-coloring	$\frac{11\Delta}{6}$ [Vig00]			2.58Δ[LY13]
Matching Poly		$t \le \text{poly}(n)$ [JS89]	$t, \Delta \le O(1)$[BGK+07]	$t, \Delta \le O(1)$ [PR17a]
Indepen Poly	$t \le \frac{1}{\Delta-3}$ [DG00]		$t \le \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$[Wei06]	$t \le \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$[PR17a, PR17b]
Perfect Matching		Bipartite [JSV04]		

Although there is no rigorous equivalence theorem between these methods the bounds and the class of problems studied using them are quite related; see the following table for an overview.

Global Methods: Path Technology, Multivariate Polynomials Perhaps, the main "global" technique that we have discussed in this course to this date is the path technology. The main idea is that we find a path between two unrelated states of the Markov chain, and furthermore we design these paths such that no edge is over congested. Due to these reason several of the results based on this method have not been replicated using any other proof technique. In particular, counting perfect matching of the bipartite graph and estimating the matching polynomial of an arbitrary degree graph. In the last part of the course we will discuss another method which is based on studying the structure of the roots of *multivariate* polynomials corresponding to our underlying counting problems. It turns out that this is a very global technique and in some cases it can be used to match or even beat the best known algorithms using the Markov chain technique.

17.2 $\mathbb{D}$-stable Polynomials

As we discussed our last set of techniques follow by studying multivariate polynomials associated with our counting problems.

In general we will study polynomials that do not have roots in a circular open convex part of the complex plane, say Ω. We call these polynomials Ω-stable. There are several families of these polynomials, but they share several basic properties.

- **Multiplication:** If p, q are Ω-stable, then so is $p \cdot q$.
- **Substitution:** If $p(z_1, \ldots, z_n)$ is Ω-stable, then for any $a \in \Omega$ or in boundary of Ω, $p(a, z_2, \ldots, z_n)$ is either identically zero or is Ω-stable. Note that the claim is obvious if $a \in \Omega$. To prove it for the boundary of Ω we have to use a limiting theorem of Hurwitz that we do not discuss here.
- **Symmetrization:** If $p(z_1, \ldots, z_n)$ is Ω-stable, then so is $p(z_1, z_1, z_3, \ldots, z_n)$.

Now, let us discuss a specific example called $\mathbb{D}$-stable and see its applications in estimating the partition function of Ising model. Let $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. We say a multivariate polynomial $p \in \mathbb{C}[z_1, \ldots, z_n]$ is $\mathbb{D}$-stable if $p(z_1, \ldots, z_n) \neq 0$ whenever $|z_i| \in \mathbb{D}$ for all i. In other words, p has no roots in $\mathbb{D}^n$.

Fact 17.1. *For any $a \in \mathbb{C}$ such that $|a| \le 1$, $1 + az$ is $\mathbb{D}$-stable.*

Lemma 17.2. *For any $a \in \mathbb{C}$ such that $|a| \le 1$, $1 + az_1 + \overline{a}z_2 + z_1z_2$ is $\mathbb{D}$-stable.*

Proof. If $|a| = 1$ then $a\overline{a} = 1$ and

$$1 + az_1 + \overline{a}z_2 + z_1z_2 = (1 + az_1)(1 + \overline{a}z_2).$$

But by the above fact each of these two terms is $\mathbb{D}$-stable and so is the product.

Now, suppose $|a| < 1$. Then, solving the equation $1 + az_1 + \overline{a}z_2 + z_1 z_2 = 0$ for z_2 we obtain

$$z_1 = -\frac{1 + az_1}{\overline{a} + z_1}.$$

Now observe that for any complex z such that $|z| = 1$ we have

$$\begin{aligned} |1 + az| &= |1 + \overline{a}\overline{z}| \\ |\overline{a} + z| &= |\overline{a} + z||\overline{z}| = |\overline{a}\overline{z} + 1|. \end{aligned}$$

Therefore, the function $z \to -\frac{1+az}{\overline{a}+z}$ maps the unit circle $|z| = 1$ onto itself and the disc $|z| < 1$ onto its complement $|z| > 1$ (here we used that $|a| < 1$). Therefore, if $|z_1| < 1$, then $|z_2| > 1$. ☐

Let $p(z_1, \dots, z_n) = \sum_{S \subset [n]} a_S z^S$ and $q(z_1, \dots, z_n) = \sum_S b_S z^S$ where $z^S = \prod_{i \in S} z_i$. The G-product of p, q is defined as follows:

$$p * q = \sum_{S \subseteq [n]} a_S b_S z^S.$$

In the following theorem we show that if p, q are $\mathbb{D}$-stable, then so is $p * q$.

Theorem 17.3 (G-product)**.** *For any two multilinear $\mathbb{D}$-stable polynomial p, q, we have $p * q$ is $\mathbb{D}$-stable.*

Proof. We prove by induction. First of all, if $n = 1$ then we can write

$$\begin{aligned} p &= a + bz, \\ q &= c + dz, \\ p * q &= ac + bdz. \end{aligned}$$

Since p is $\mathbb{D}$-stable, we have $a \neq 0$ and $|a| \geq |b|$. Similarly, we have $c \neq 0$ and $|c| \geq |d|$. Therefore, $ac \neq 0$ and $|ac| \geq |bd|$.

The proof of the induction step basically follows from the following lemma. Firstly we can write

$$\begin{aligned} p(z_1, \dots, z_n) &= \sum_{S \subseteq [n-1]} (a_S + z_n a_{S \cup n}) z^S \\ q(w_1, \dots, w_n) &= \sum_{S \subseteq [n-1]} (b_S \cup w_n b_{S \cup n}) w^S. \end{aligned}$$

Now, fix $z = z_n, w = w_n$ both in $\mathbb{D}$. The resulting polynomials are $\mathbb{D}$-stable by the above fact so by induction hypothesis the following polynomial is $\mathbb{D}$-stable:

$$r_{z,w}(z_1, \dots, z_{n-1}) = \sum_{S \subseteq [n-1]} (a_S + z a_{S \cup n})(b_S + w b_{S \cup n}) z^S.$$

Now, fix $z_1, \dots, z_{n-1} \in \mathbb{D}$. It follows that the polynomial

$$f_{z_1,\dots,z_{n-1}}(z, w) = \sum_{S \subseteq [n-1]} a_S b_S z^S + z \sum_{S \subseteq [n-1]} a_{S \cup n} b_S z^S + w \sum_{S \subseteq [n-1]} a_S b_{S \cup n} z^S + zw \sum_{S \subseteq [n-1]} a_{S \cup n} b_{S \cup n} z^S$$

is $\mathbb{D}$-stable. Therefore, by Lemma 17.4

$$g_{z_1,\dots,z_{n-1}}(z, w) = \sum_{S \subseteq [n-1]} a_S b_S z^S + z \sum_{S \subseteq [n-1]} a_{S \cup n} b_{S \cup n} z^S$$

is $\mathbb{D}$-stable for any $z_1, \dots, z_{n-1}, z \in D$. Therefore, the polynomial $p * q(z_1, \dots, z_n)$ has no root in $\mathbb{D}^n$ and it is $\mathbb{D}$-stable. ☐

Lemma 17.4. *Suppose that the bivariate polynomial $p(z_1, z_2) = a + bz_1 + cz_2 + dz_1z_2$ is $\mathbb{D}$-stable. Then, the univariate polynomial $q(z) = a + dz$ is $\mathbb{D}$-stable.*

Proof. Since p is $\mathbb{D}$-stable we must have $a \neq 0$. For the sake of contradiction suppose that q is not $\mathbb{D}$-stable. Then, we must have $|a|/|d| < 1$ so $|d| > |a|$. Without loss of generality assume that $|b| \geq |c|$. It follows that $|d| + |b| > |a| + |c|$.

Now, choose $z_2 < 1$ such that

$$|b + dz_2| = |b| + |z_2||d| > |a| + |c|.$$

Note that this is always possible. Choose z_2 such that dz_2 points to the same direction as b and choose its norm to be very close to 1.

Now, consider the open disc

$$K = \{(b + dz_2)z_1 : |z_1| < 1\}.$$

This gives a disc of radius $|b + dz_2| > |a| + |c|$ around the origin. On the other hand, by triangle inequality

$$|a + cz_2| \leq |a| + |cz_2| < |a| + |c|.$$

Therefore, for the above chosen z_2 and $|z_1| < 1$, $a + cz_2 + bz_1 + dz_2z_1$ is the set of points in K translated by the vector $a + cz_2$. But because $|a + cz_2|$ is less than the radius of K, the origin is still a feasible point in $a + cz_2 + bz_1 + dz_2z_1$. So, we get a contradiction. Therefore, $|d| \leq |a|$ and q is $\mathbb{D}$-stable. ☐

17.3 Application: Cut Polynomial

Let G be a graph with vertices $V = \{1, \ldots, n\}$. Let A be the adjacency matrix, where each edge i, j has weight $A_{i,j}$. For simplicity we assume that we have a complete graph; so if there is no edge between i, j assume $A_{i,j} = 1$. The cut polynomial of G is defined as follows:

$$\text{cut}_A(z_1, \ldots, z_n) = \sum_{S \subset V} z^S \prod_{i \in S, j \notin S} A_{i,j},$$

where as usual $z^S = \prod_{i \in S} z_i$. The following spectacular result due to Lee and Yang [LY52] says that all of the roots of the symmetrized univariate cut polynomial lie on the unit circle around the origin as long as $|a_{i,j}| \leq 1$ for all i, j.

Theorem 17.5. *For any graph G if $|a_{i,j}| \leq 1$ for all i, j then all roots of the cut polynomial $cut(z, \ldots, z)$ lie on the unit circle around the origin.*

Proof. To prove the theorem we will write the cut polynomial as the G-product of many $\mathbb{D}$-stable polynomials. For any $1 \leq i < j \leq n$ let

$$\text{cut}_{i,j}(z_1, \ldots, z_n) = \sum_{\substack{S \subset V, \\ i,j \notin S \text{ or } i,j \in S}} z^S + \sum_{\substack{S \subset V, \\ i \in S, j \notin S \text{ or } i \in S, j \notin S}} z^S a_{i,j}$$

Now, observe that the cut polynomial is exactly the G-product of all polynomials $\text{cut}_{i,j}$ over all i, j. On the other hand, we claim that $\text{cut}_{i,j}$ is $\mathbb{D}$-stable. This is because we can write

$$\begin{aligned}
\text{cut}_{i,j}(z_1, \ldots, z_n) &= (1 + a_{i,j}z_i + a_{i,j}z_j + z_iz_j) \sum_{\substack{S \subset V, \\ i,j \notin S}} z^S \\
&= (1 + a_{i,j}z_i + a_{i,j}z_j + z_iz_j) \prod_{k \in \{1, \ldots, n\} - \{i,j\}} (1 + z_k).
\end{aligned}$$

Therefore, by Lemma 17.2 $\text{cut}_{i,j}$ is $\mathbb{D}$-stable for all i, j. It follows by Theorem 17.3 that $\text{cut}(z_1, \ldots, z_n)$ is $\mathbb{D}$-stable.

Therefore by symmetrization operator $\text{cut}(z, \ldots, z)$ is also $\mathbb{D}$-stable and for any z where $|z| < 1$, $\text{cut}(z, \ldots, z) \neq 0$. On the other hand, observe that

$$z^n \,\text{cut}(1/z, \ldots, 1/z) = \text{cut}(z, \ldots, z).$$

Therefore, $\text{cut}(z, \ldots, z)$ has no roots z with $|z| > 1$. □

We remark that the above theorem also holds true even if the weights on the edges of G are complex numbers and A is a Hermitian matrix.

Furthermore, observe that by Barvinok's polynomial approximation technique we can estimate the cut polynomial within $1 + \epsilon$ approximation at any z where $|z| < 1/\beta$ in time quasipolynomial in n, ϵ and exponential in β.

17.4 Connection to Ferromagnetic Ising Model

In this section we see application of the cut polynomial in computing the partition function of the ferromagnetic Ising model. We prove the following theorem:

Theorem 17.6. *Given a graph G consider the (ferromagnetic) Ising model where for every configuration $\sigma : V \to \{+1, -1\}$,*

$$\pi(\sigma) \propto e^{\gamma \sum_{i\sim j} \sigma_i \sigma_j} \prod_i z^{\sigma_i}$$

where $\gamma > 0$ corresponds to the inverse temperature and λ represents vertex activities. Then, for any $\lambda \neq 1$ and $\epsilon > 0$ we can estimate the partition function $\sum_\sigma \pi(\sigma)$ in a quasi-polynomial time in n and ϵ and exponential in $1/|1 - \lambda|$.

The idea is to use a change variable and write the partition function of the Ising model as a cut polynomial. In particular, observe that

$$\sum_\sigma e^{\gamma \sum_{i\sim j} \sigma_i \sigma_j} \prod_i z^{\sigma_i} = Z \sum_\sigma e^{-\gamma |\{i\sim j : \sigma_i \neq \sigma_j\}|} \lambda^{-|\{i:\sigma_i=-1\}|}$$

where Z is a normalizing constant.

Now, let $z = \lambda^{-1}$, and $A_{i,j} = e^{-\gamma}$ if $i \sim j$ and $A_{i,j} = 1$ otherwise. Note that since we assumed that $\gamma \geq 0$ (i.e., the ferromagnetic case), we have that $|A_{i,j}| \leq 1$ and this satisfies the necessary condition of Theorem 17.5

Then, up to a normalization the partition function is exactly:

$$\sum_{\sigma, S=\{i:\sigma_i=-1\}} z^S \prod_{i\in S, j\notin S} A_{i,j}.$$

So, we can estimate this polynomial at any z where $|z| < 1$ or $|z| > 1$ given the first $O(\log n/\epsilon)$ coefficients. Those coefficients correspond to computing k-wise correlation between the spins of vertices and can be done in quasi-polynomial time. If G has bounded degree then this can be done in polynomial time by a recent work of Liu, Sinclair and Srivastava [LSS17].

Upshot. The above method shows that we can compute the partition function of the (ferromagnetic) Ising model deterministically in quasi-polynomial time assuming that the vertex activities are different from 1. Recall that we proved in the middle of the course that the natural Heat-Bath chain mixes in exponential time if the temperature is larger than a constant even in the case of 2d lattice. So, in some sense this above machinery is stronger.

On the other hand, Jerrum and Sinclair in 1993 designed an FPRAS to generate random samples at any (positive) temperature and for any vertex activities [JS93]. They study the mixing time of a very different chain that instead of walking over the spin configurations, moves over the spanning subgraphs of the input graph. So, at a high-level Markov chain techniques still give stronger result for estimating the partition function of the (ferromagnetic) Ising model in the sense that the algorithm is polynomial and it works even for $\lambda = 1$. Nonetheless, the deterministic approach based on geometry of polynomials is a more direct approach.

References

[BGK+07] Mohsen Bayati, David Gamarnik, Dimitriy A. Katz, Chandra Nair, and Prasad Tetali. Simple deterministic approximation algorithms for counting matchings. In *STOC*, pages 122–127, 2007. 17-2

[DG00] Martin Dyer and Catherine Greenhill. On markov chains for independent sets. *Journal of Algorithms*, 35(1):17–49, 2000. 17-2

[JS89] M.R. Jerrum and A.J. Sinclair. Approximate counting, uniform generation and rapidly mixing markov chains. *Information and Computation*, 82:93–133, 1989. 17-2

[JS93] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the ising model. *SIAM J. Comput.*, 22(5):1087–1116, 1993. 17-6

[JSV04] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *J. ACM*, 51(4):671–697, July 2004. 17-2

[LSS17] Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. The ising partition function: Zeros and deterministic approximation. In *FOCS*, pages 986–997, 2017. 17-5

[LY52] T. D. Lee and C. N. Yang. Statistical theory of equations of state and phase transitions. ii. lattice gas and ising model. *Phys. Rev.*, 87:410–419, Aug 1952. 17-4

[LY13] P. Lu and Y. Yin. Improved fptas for multi-spin systems. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 639–654. Springer Berlin Heidelberg, 2013. 17-2

[PR17a] Viresh Patel and Guus Regts. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *Electronic Notes in Discrete Mathematics*, 61(Supplement C):971 – 977, 2017. The European Conference on Combinatorics, Graph Theory and Applications (EUROCOMB). 17-2

[PR17b] H. Peters and G. Regts. On a conjecture of sokal concerning roots of the independence polynomial. January 2017. 17-2

[Vig00] Eric Vigoda. Improved bounds for sampling colorings. *Journal of Mathematical Physics*, 41(3):1555–1569, 2000. 17-2

[Wei06] Dror Weitz. Counting independent sets up to the tree threshold. In *STOC*, pages 140–149, New York, NY, USA, 2006. ACM. 17-2