Moden Spectral Graph Theory **Winter 2022**

Lecture 11: Approximate Sampling and High Dimensional Complexes

Lecturer: Shayan Oveis Gharan

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

Let U be a ground set of elements and $n \geq 1$ be an integer. Given a weight function $w : \binom{U}{n+1} \to \mathbb{R}_{\geq 0}$ we consider the following tasks:

- Sample a set $S \in \binom{U}{n+1}$ with probability proportional to $w(S)$
- Compute the "normalizing constant" of this distribution, namely $\sum_{S \in \binom{U}{n+1}} w(S)$.

It follows by a classic result of Jerrum, Valiant and Vazirani [JVV86] that the above two problems are equivalent for most interesting probability distributions and more generally even the approximate versions of these problems are equivalent.

So, here, we will mainly address the sampling problem. Broder [Bro86] in his influential paper proposed to design a Markov chain with stationary distribution $\pi_n(S) := \frac{w(S)}{\sum_{T \in \binom{U}{n+1}} w(T)}$ and then bound the *mixing time* of the chain. Recall that the L_1 mixing time for a Markov chain with transition probability matrix P is

$$\tau_{1,\epsilon} = \max_{S \in \binom{U}{n+1}} \min \left\{ t : \left\| \frac{P^t(S,.)}{\pi_n(.)} - 1 \right\|_1 \leq \epsilon \right\},$$

where $P^t(S,.)$ is the distribution of the chain started at S after t steps. Here, we will study the following Markov chain to sample from π_n: Given a state S, first we delete a uniformly random element from S, say i and we go to $S - \{i\}$. Then, from all the sets T that contain $S - \{i\}$ we choose one with probability proportional to $\pi_n(T)$. In other words, consider a weighted bipartite graph $G = (\binom{U}{n+1}, \binom{U}{n}, E)$ where a set $S \in \binom{U}{n+1}$ is connected to $T \in \binom{U}{n}$ iff $T \subseteq S$ and the weight of that edge is equal to $\frac{\pi_n(S)}{n+1}$. The aforementioned chain is the same as running a simple random walk on this bipartite graph where from each vertex we jump to a neighbor with probability proportional to the weight of the edge connecting to the neighbor. Since the (weighted) degree of every vertex $S \in \binom{U}{n+1}$ is exactly $\pi_n(S)$, the stationary distribution on the top vertices is exactly $\pi_n(.)$. See the following diagram for an example. This walk is known as the down-up walk in the high-dimensional expander language [KM17], the Glauber dynamics in statistical physics and basis exchange walk in the matroid language. For a concerete example, let $G = (V, E)$ be a

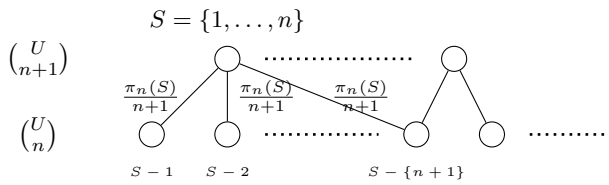


Figure 11.1: An illustration of the Down-Up Walk

graph with $n + 1 := |V|$ vertices that we want to uniformly randomly color with q colors. We define U to be the set of all vertex-color pairs, (v, c), $v \in V, c \in [q]$. A set $S \in \binom{U}{n+1}$ is in support if π_n if it corresponds

to a valid proper color of G and π_n is simply the uniform distribution over all such sets. In such a case the down-up walk corresponds to first choosing a u.r. vertex of G, v; then "un-color" v. Finally, among all valid colors we can assign to v, choose one u.a.r.

Classical Techniques: Classically there are two well-known methods to study mixing time of random walks.

Canonical Path Method This method was proposed in the influential work of Jerrum and Sinclair [JS89]. The high-level idea is to construct a multi-commodity flow on the graph of the Markov chain between each pair of states and then use the connection to the sparsest cut problem to bound the spectral gap of the chain. This method is most famously used to sample a uniformly random perfect matching from a bipartite graph [JSV04]. Unfortunatley, most applications of this method is limited to problems related to matchings.

Path Coupling In this method, one would directly bound the mixing time by designing a "Markovian coupling" between the distribution of the chain and the stationary distribution. This method is widely used in theory but often it does not give the optimal result.

11.1 Spectral Independence and the Local to Global Theorem

Write $P_n^{\vee}$ to denote the transition probability matrix of the down-up walk we defined above. In these notes we explain a new technique to analyze the mixing time of these family of walks called the *spectral independence*.

Definition 11.1 (Spectral Independence)**.** *Given a probability distribution π_n on $\binom{U}{n+1}$, define a matrix $\Psi \in \mathbb{R}^{U\times U}$ where for any $i \neq j \in U$,*

$$\Psi_{\pi_n}(i,j) = \mathbb{P}_{S\sim\pi_n}\left[j \in S | i \in S\right] - \mathbb{P}_{S\sim\pi_n}\left[j \in S\right].$$

If $i = j$ we simply let $\Psi_{\pi_n}(i,i) = -\mathbb{P}\left[i \in S\right]$. We say π_n is η-spectrally independent if $\lambda_{\max}(\Psi_{\pi_n}) \leq \eta$. We say π_n is η^*-spectrally independent if for any sequence $i_1, \ldots, i_{n-1} \in U$, $(\pi_n), (\pi_n|i_1), (\pi_n|i_1, i_2), \ldots, (\pi_n|i_1, \ldots, i_{n-1})$ are η-spectrally independent.

Independent Case. For example, suppose π_n is a product distribution. In that case for any $i, j \in U$, $\mathbb{P}\left[j|i\right] = \mathbb{P}\left[j\right]$. Therefore, all off-diagonal entries of Ψ_{π_n} are zero; since the diagonal entries are non-positive, π_n is 0-spectrally independent.

Negatively Correlated Case. For another example, suppose π_n is a *negatively correlated* distribution, namely for any $i, j \in U$, $\mathbb{P}\left[j|i\right] \leq \mathbb{P}\left[j\right]$. One simple upper-bound on $\lambda_{\max}(\Psi)$ is $\max_i \sum_j |\Psi(i,j)|$. In our example, all off-diagonal entries of Ψ_{π_n} are non-positive; therefore (using the homogeneity of π_n),

$$|Pi| + \sum_{j\neq i} |\mathbb{P}\left[j|i\right] - \mathbb{P}\left[j\right]| = \sum_{j\neq i} \mathbb{P}\left[j\right] - \mathbb{P}\left[j|i\right] + |\mathbb{P}\left[i\right]| = 2(1 - \mathbb{P}\left[i\right]) = 1$$

So, π_n is 1-spectrally independent.

Positively Correlated Case. For a bad example, suppose there are only two sets in the support of π_n; namely $\pi_n(\{1, \ldots, n+1\}) = \pi_n(\{n+2, \ldots, 2n+2\} = 1/2$. In this case the distribution is very positively

correlated (and in fact the down-up walk explained before is not even connected). It follows that $\Psi_{\pi_n} = \frac{1}{2}\begin{pmatrix} J_{n+1} & -J_{n+1} \\ -J_{n+1} & J_{n+1} \end{pmatrix}$. So, π_n is $n+1$-spectrally independent.

The following "local-to-global" theorem follows from a long line of works in theory of high dimensional expanders and extensions to the field of analysis of random walks:

Theorem 11.2 ([**AL20**; DK17; KO20; ALO21a])**.** *If π_n is η^* spectrally independent then the down-up walk $P_n^\vee$ has spectral gap at least $\frac{1}{O(n^{1+\eta})}$ and thus it mixes in polynomial time, assuming $\eta \le O(1)$.*

In other words, the above theorem shows that even if π_n is positively correlated, but the positive correlations are "limited" then still the simple down-up walk mixes rapidly.

11.2 Background on Simplicial Complexes

Similar to lectures on locally testable code, we use X to denote a simplical complex. We use $X(0)$ to denote, vertices, or faces of dimension 0; $X(1)$ to denote edges. In our case we let X contain all subsets S in the support of π_n and all of their subsets. Note that, by definition, X is a simplicial complex if it is downward closed, namely for any $S \in X$, and $T \subseteq S$ we have $T \in X$.

For a face $\sigma \in X$, we write $\dim(\sigma) = |\sigma| - 1$ to denote the dimension of σ and we write $\dim(X) := \max_{\sigma \in X} \dim(\sigma)$. So, if X is defined based on π_n as above, it will be an n-dimensional complex.

We equip this complex with a probability distribution over its maximal faces (namely sets in the support of π_n. So, we use π_n to denote the distribution over $X(n)$. Now, we can naturally extend this distribution to π_i for any $0 \le i < n$. Say we have defined π_i, to define π_{i-1} we choose $\tau \sim \pi_i$ randomly and we delete a uniformly random element of τ. So, for any $\sigma \in X(i-1)$ we have

$$\pi_{i-1}(\sigma) = \sum_{\tau \in X(i): \sigma \subseteq \tau} \frac{\pi_i(\tau)}{i+1}$$

This is a natural extension of the π_1 to π_0 definition we had in the first lecture. Similarly, we can go from π_{i-1} to π_i as follows: First we sample $\sigma \sim \pi_{i-1}$ then we sample $\tau \sim \pi_i|\sigma$. Pictorially, these two conditionals is the same as moving down/up in the bipartite graph illustrated in Figure 11.1.

Links. For the complex X and a face σ define the link X_σ as the local view of σ:

$$X_\sigma = \{\tau : \sigma \cap \tau = \emptyset, \sigma \cup \tau \in X\}.$$

In the special case that X is a 1-dimensional complex, namely a graph G, for any vertex $v \in X(0)$, X_v is the set of neighbors of v in G. For another example, see the use of links in the construction of locally testable codes.

Observe that for any $\sigma \in X$, X_σ is also a simplicial complex. Furthermore, given a distribution π_n (on top faces of X) it natural extends to a distribution on top faces of X_σ. Say, X_σ is k-dimensional; then for any $\tau \in X_\sigma(k)$ we have

$$\pi_{\sigma,k}(\tau) := \frac{\pi_n(\tau \cup \sigma)}{\sum_{\tau' \in X_\sigma(k)} \pi_n(\tau' \cup \sigma)} = \mathbb{P}_{X \sim \pi_n}\left[X = \sigma \cup \tau | \sigma \subset X\right].$$

Given $\pi_{\sigma,k}$ naturally, we extend this to distributions $\pi_{\sigma,i}$ for all $0 \le i < k$.

For a concrete example, suppose $\pi_2(\{1,3,4\}) = 1/2$, $\pi_2(\{1,2,4\}) = 1/4$ and $\pi_2(\{2,3,4\} = 1/4$. Then,

$$X_1 = \{\{3,4\},\{2,4\},\{2\},\{3\},\{4\}\}.$$

So it has dimension 1. Then, $\pi_{\{1\},1}(\{3,4\}) = 2/3$ and $\pi_{\{1\},1}(\{2,4\}) = 1/3$. So,

$$\pi_{\{1\},0}(4) = 1/2, \pi_{\{1\},0}(2) = \pi_{\{1\},0}(2) = 1/4.$$

Inner Products Having defined π_i's, we define an inner-product, i.e., a Hilbert space, for functions on $\mathbb{R}^{X(i)}$. Namely, for $f, g \in \mathbb{R}^{X(i)}$, define

$$\langle f, g\rangle_{\pi_i} = \mathbb{E}_{\sigma\sim\pi_i} f(\sigma)g(\sigma).$$

For any $\sigma \in X_\tau$, define

$$f_\tau(\sigma) = f(\tau \cup \sigma)$$

to denote the *restriction* of a function to X_τ. The following fact is fundamental to our analysis:

Fact 11.3 (Garland's Method)**.** *For any two functions $f, g \in \mathbb{R}^{X(i)}$, and $k < i$,*

$$\begin{aligned}\langle f, g\rangle_{\pi_i} &= \mathbb{E}_{\sigma\sim\pi_i} f(\sigma)g(\sigma) = \mathbb{E}_{\tau\sim\pi_k}\mathbb{E}_{\sigma\sim\pi_i|_\tau} f(\sigma)g(\sigma)\\ &= \mathbb{E}_{\tau\sim\pi_k}\mathbb{E}_{\sigma\sim\pi_{\tau,i-k-1}} f(\sigma\cup\tau)g(\sigma\cup_\tau) = \mathbb{E}_{\tau\sim\pi_k}\langle f_\tau, g_\tau\rangle_{\pi_{\tau,i-k-1}}.\end{aligned}$$

Down and Up Operators. We define $P^{\downarrow}_{i\to i-1}$ as follow: For $f \in \mathbb{R}^{X(i)}$ and any $\sigma \in X(i-1)$,

$$P^{\downarrow}_{i\to i-1} f(\sigma) = \mathbb{E}_{i\sim\pi_{\sigma,0}} f(\sigma + i).$$

It turns out that the up-operator $P^{\uparrow}_{i-1\to i}$ is the adjoint of $P^{\downarrow}_{i\to i-1}$ with respect to these inner products: Namely for any $f \in \mathbb{R}^{X(i)}, g \in \mathbb{R}^{X(i-1)}$,

$$\langle P^{\downarrow}_{i\to i-1} f, g\rangle_{\pi_{i-1}} = \langle f, P^{\uparrow}_{i-1\to i} g\rangle_{\pi_i}.$$

More formally, the up-operator is defined as follows: For any $f \in \mathbb{R}^{X(i-1)}$ and $\sigma \in X(i)$ we have

$$P^{\uparrow}_{i-1\to i} f(\sigma) = \sum_{i\in\sigma} \frac{f(\sigma - i)}{|\sigma + 1|}.$$

If we think of these operators as a matrix, (so we allow for left multiplication), then we can write

$$\pi_{i-1} = \pi_i^T P^{\uparrow}_{i-1\to i} \qquad \text{and} \qquad \pi_i = \pi_{i-1}^T P^{\downarrow}_{i\to i-1}.$$

11.3 Proof of the Local to Global Theorem

Recall that for a function $f \in \mathbb{R}^{X(n)}$,

$$\mathrm{Var}_{\pi_n}(f) = \mathbb{E}_{\sigma\sim\pi_n} f(\sigma)^2 - (\mathbb{E} f)^2.$$

Having this, let $f^{(n)} \in \mathbb{R}^{X(n)}$; to prove the claim it is enough to show that

$$\mathrm{Var}(P^{\vee} f^{(n)}) \le (1-\epsilon)\,\mathrm{Var}(f^{(n)})$$

for some $\epsilon \geq \frac{1}{O(n^{1+\eta}}$. Let

$$f^{(n-1)} = P^{\downarrow}_{n\to n-1} f^{(n)}$$

And, similarly, we write $f^{(i)}$ for every $i \leq n-1$ where $f^{(i)} = P^{\downarrow}_{i+1\to i} f^{(i+1)}$.

To upper bound variance of $P^{\vee} f^{(n)}$ it is enough to show that

$$\mathrm{Var}(P^{\vee} f^{(n)}) = \mathrm{Var}(P^{\uparrow}_{n-1\to n} P^{\downarrow}_{n\to n-1} f^{(n)}) = \mathrm{Var}(P^{\uparrow}_{n-1\to n} f^{(n-1)}) \leq \mathrm{Var}_{\pi_{n-1}}(f^{(n-1)}) \underset{?}{\leq} (1-\epsilon)\,\mathrm{Var}_{\pi_n}(f^{(n)}). \tag{11.1}$$

To see the first inequality notice that for any $f \in \mathbb{R}^{X(n-1)}$,

$$\mathrm{Var}_{\pi_n}(P^{\uparrow}_{n-1\to n} f) = \left\| P^{\uparrow}_{n-1\to n} f \right\|^2_{\pi_n} - \langle P^{\uparrow}_{n-1\to n} f, \mathbf{1}\rangle_{\pi_n} \leq \|f\|^2_{\pi_{n-1}} - \langle f, \mathbf{1}\rangle_{\pi_n} = \mathrm{Var}_{\pi_n}(f).$$

where we leave it an exercise to prove the inequality.

The main tool that we used in the proof of inequality (11.1) is the following version of the law of total variance.

Lemma 11.4 (Law of Total Variance)**.** *For any $1 \leq i \leq n-1$, and $f^i \in \mathbb{R}^{X(i)}$, we have*

$$\mathrm{Var}_{\pi_i}(f^{(i)}) = \mathrm{Var}_{\pi_{i-1}}(f^{(i-1)}) + \mathbb{E}_{\tau\sim\pi_{i-1}} \mathrm{Var}_{\pi_{\tau,0}}(f^{(i)}_{\tau})$$

Proof. By Garland's method Fact 11.3, we can write,

$$\begin{aligned}
\mathrm{Var}_{\pi_i}(f^{(i)}) &= \mathbb{E}_{\sigma\sim\pi_i} f^{(i)}(\sigma)^2 - (\mathbb{E} f^{(i)})^2 \\
&= \mathbb{E}_{\tau\sim\pi_{i-1}} \langle f^{(i)}_{\tau}, f^{(i)}_{\tau}\rangle - \mathbb{E}_{\tau\sim\pi_{i-1}} \left(\mathbb{E}_{\pi_{\tau,0}} f^{(i)}_{\tau}\right)^2 + \mathbb{E}_{\tau\sim\pi_{i-1}} \left(\mathbb{E}_{\pi_{\tau,0}} f^{(i)}_{\tau}\right)^2 - \left(\mathbb{E} f^{(i)}\right)^2 \\
&= \mathbb{E}_{\tau\sim\pi_{i-1}} \mathrm{Var}(f^{(i)}_{\tau}) + \mathbb{E}_{\tau\sim\pi_{i-1}} f^{(i-1)}(\tau)^2 - (\mathbb{E} f^{(i-1)})^2
\end{aligned}$$

as desired. □

The following lemma is the second important tool of the proof.

Lemma 11.5. *Let $f^{(k+2)} \in \mathbb{R}^{X(k+2)}$. For every $\tau \in X(k)$, If π_τ is η-spectrally independent, then we have*

$$\mathrm{Var}_{\pi_{\tau,0}}(f^{(k+1)}_{\tau}) \leq \left(1 + \frac{2\eta}{n - |\tau|}\right) \mathbb{E}_{i\sim\pi_{\tau,0}} \mathrm{Var}(f^{(k+2)}_{\tau+i}).$$

Proof. We prove the lemma for the special case that $\tau = \emptyset$, and it naturally generalizes to all τ's. Construct a weighted (complete) graph G with vertices U such that for any $i, j \in U$, $w(\{i,j\}) = \mathbb{P}_{S\sim\pi_n}[i,j \in S]$. Let P be transition probability operator of this walk and notice that its at stationarity i has probability

$$\frac{\sum_j \mathbb{P}[i,j]}{\sum_{k\in U}\sum_j \mathbb{P}[k,j]} = \frac{n\mathbb{P}[i]}{\sum_{k\in U} n\mathbb{P}[k]} = \frac{\mathbb{P}[i]}{n+1}$$

Call this stationary vector μ, i.e., $\mu(i) = \frac{\mathbb{P}[i]}{n+1}$. The observation is that

$$\Psi_{\pi_n} = n(P - \frac{n+1}{n}\mathbf{1}^T\mu).$$

Since $\lambda_{\max}(\Psi_{\pi_n}) \leq \eta$ we have that $\lambda_2(P) \leq \eta/n$.

Therefore,

$$\left\|f^{(0)}\right\|_{\pi_0}^2 = \left\|P_{1\to 0}^{\downarrow} f^{(1)}\right\|_{\pi_0}^2 = \langle P_1^{\vee} f^{(1)}, f^{(1)}\rangle_{\pi_1} \le \lambda_2(P_1^{\vee}) \left\|f^{(1)}\right\|^2$$
$$= \lambda_2(P_0^{\wedge}) \left\|f^{(1)}\right\|^2 = (1+\lambda_2(P))/2 \left\|f^{(1)}\right\|^2 \le \frac{\eta/n+1}{2} \left\|f^{(1)}\right\|^2.$$

The second to last equality uses that non-zero eigenvalues of $P_1^{\wedge} = P_{1\to 0}^{\downarrow} P_{0\to 1}^{\uparrow}$ are the same as the non-zero eigenvalues of $P_1^{\vee} = P_{1\to 0}^{\downarrow} P_{0\to 1}^{\uparrow}$. The last equality uses that $P_0\wedge$ is simply a (half)-lazy version of P. This is because from a vertex i we choose an edge $\{i,j\}$ proportional to its weight and we delete one endpoint, so with probability $1/2$ we stay at i. So, by the law of total variance we get

$$\mathrm{Var}(f^{(0)}) \le \frac{\eta/n+1}{2} \mathrm{Var}(f^{(1)}) = \frac{\eta/n+1}{2} \left(\mathrm{Var}(f^{(0)}) + \mathbb{E}_{i\sim\pi_0} \mathrm{Var}(f_i^{(1)})\right) \le .$$

Switching around we get

$$\mathrm{Var}(f^{(0)}) \le \frac{1+\eta/n}{1-\eta/n} \mathbb{E}_{i\sim\pi_0} \mathrm{Var}(f_i^{(1)}) \approx (1+2\eta/n) \mathbb{E}_{i\sim\pi_0} \mathrm{Var}(f_i^{(1)})$$

as desired. □

First, using repeated application of Lemma 11.5, for any $k \le n-1$ we have

$$\mathbb{E}_{\tau\sim\pi_k} \mathrm{Var}_{\pi_{\tau,0}}(f_\tau^{(k+1)}) \le \left(1 + \frac{2\eta}{n-(k+1)}\right) \mathbb{E}_{\tau\sim\pi_{k+1}} \mathrm{Var}_{\pi_{\tau,0}}(f_\tau^{(k+2)}) \le \dots$$
$$\le \underbrace{\prod_{i=k}^{n-2} \left(1 + \frac{2\eta}{n-i+1}\right)}_{\approx n^{2\eta}} \mathbb{E}_{\tau\sim\pi_{n-1}} \mathrm{Var}_{\pi_{\tau,0}}(f_\tau^{(n)})$$

On the other hand, using repeated applications of the law of total variance we can write,

$$\mathrm{Var}(f^{(n-1)}) = \sum_{j=0}^{n-2} \mathbb{E}_{\tau\sim\pi_j} \mathrm{Var}_{\pi_{\tau,0}}(f_\tau^{(j+1)}) \le \underbrace{\sum_j n^{2\eta}}_{n^{2\eta+1}} \mathbb{E}_{\tau\sim\pi_{n-1}} \mathrm{Var}(f_\tau^{(n)}). \tag{11.2}$$

So, to put it differently,

$$\frac{1}{n^{2\eta+1}} \mathrm{Var}(f^{(n-1)}) \le \mathbb{E}_{\tau\sim\pi_{n-1}} \mathrm{Var}(f_\tau^{(n)})$$

Adding $\mathrm{Var}(f^{(n-1)})$ to both sides, by law of total variance, we get

$$(1+\frac{1}{n^{2\eta+1}}) \mathrm{Var}(f^{(n-1)}) \le \mathrm{Var}(f^{(n)}) \Rightarrow \mathrm{Var}(f^{(n-1)}) \le (1-\frac{1}{n^{2\eta+1}}) \mathrm{Var}(f^{(n-1)})$$

This finishes the proof of (11.1) and Theorem 11.2.

Remarks Here are some remarks about this proof: The spectral independence technique have found numerous applications in approximate counting and sampling since its introduction in [ALO21b]. It can be seen as (perhaps) the strongest technique to bound the mixing time as it is shown that any approximate counting/sampling algorithm based on the polynomial interpolation technique, correlation decay or path coupling leads to a bound on the spectral independence of the natural Markov chain. One of the most

important open problems in this regard is to relate the Canonical path method which is famously used by Jerrum, Sinclair and Vigoda [JSV04] to sample perfect matchings in a bipartite graph to spectral independence. Unfortunately, it can be seen that the chain on perfect and near perfect matchings defined in [JSV04] is not spectrally independent when the underlying graph is a cycle. So, a fascinating question is to prove an "average" variant of the above local to global theorem and use it to sampling matchings from bipartite or general graphs.

Another active area of research is to obtain optimal (near linear) mixing time for the up-down chain. To do that, one needs to study how the entropy of the distribution changes when running a down/up step (as opposed to tracking down the Variance). See [CLV21; AJKPV21] and references therein.

11.4 Application of Local to Global Theorem to the Hardcore Model

Building on the above theorem over the last couple of years it was shown that a number of well-known probability distributions are indeed spectrally independent; this has lead to the resolution of several long standing open problems.

The focus of these lectures is application to the hard-core model: Given a graph $G = (V, E)$ with $n+1 = |V|$ vertices, define $U = \{\{v, in\}, \{v, out\} : v \in V\}$. For $S \in \binom{U}{n+1}$, S is in the support of π_n if for every $v \in V$ either v is in or out, i.e., exactly one of $\{v, in\}, \{v, out\}$ are in S; furthermore, the in-vertices form an independent set in G. In such a case, we define $w(S) := \lambda^{\#\text{in vertices}}$ and we define

$$\pi_n(S) := \frac{w(S)}{\sum_{T \in \binom{U}{n+1}} w(T)}.$$

This probability distribution is called the *hardcore model*. Originally it is rooted in statistical physics and it used to model the arrangement of gas molecules. The *activity parameter* λ corresponds to an inverse of the ambient temperature in the environment.

Having that the down-up walk will correspond to the following natural Glauber dynamics: Given an independent set corresponding to S, choose a vertex $v \in V$ uniformly at random and forget whether it is in or out. If any of v's neighbor are in, set it to out. Otherwise, set it in with probability $\lambda/(1+\lambda)$ and out otherwise. The main goal of these lectures is to use the spectral independence technique to show that the Glauber dynamics mixes rapidly up to the *tree uniqueness threshold*, $\lambda_c(\Delta) \approx \frac{e}{\Delta}$. The latter is a threshold above which it is provably NP-hard to sample from the hardcore model in graphs of maximum degree Δ.

Theorem 11.6 ([ALO21b; CLV20; CLV21; AJKPV21])**.** *If $\lambda \leq (1-\delta)\lambda_c(\Delta)$ then the Glauber dynamics mixes in time $n^{O}_{\delta,\Delta}(1)$.*

To use the spectral independence technique for the hardcore model, we need to show π_n and all conditionals of π_n are η-spectrally independent (assuming $\lambda < \lambda_c(\Delta)$). First observe that any conditional of $\pi_n|\{v, in\}$ or $\pi_n|\{v, out\}$ is another instance of the hard-core model: In particular, if we condition v to be in, then all neighbors of v are out so we can delete v and all of its neighbors and study the spectral independence in the resulting graph and similarly, if v is in, then we can simply delete it and study spectral independence of the resulting graph. Furthermore, if we start with a graph of max-degree Δ, its maximum degree remain at most Δ after these deletions. So, this makes our job significantly easier: instead of bounding max-eigenvalue of exponentially many matrices as suggested in Theorem 11.2 all we need to prove is to show that $\lambda_{\max}(\Psi_{\pi_n}) \leq O_\delta(1)$ assuming $\lambda \leq (1-\delta)\lambda_c(\Delta)$ where $O_\delta(1)$ hides constants depending on δ.

Now, define a new matrix $\mathcal{I} \in \mathbb{R}^{V \times V}$, called the influence matrix, where

$$\mathcal{I}(u,v) := \mathcal{I}_G(u \to v) := \mathbb{P}[v|u] - \mathbb{P}[v|\overline{u}].$$

Lemma 11.7. *The non-zero eigenvalues of $\mathcal{I}$ are the same as the non-zero eigenvalues of Ψ_{π_n}.*

Proof. Define $A, B \in \mathbb{R}^{V \times V}$ where for any $u \neq v$,

$$A(u,v) = (\mathbb{P}[v|u] - \mathbb{P}[v])\mathbb{I}[u \neq v] \qquad B(u,v) = (\mathbb{P}[v|\overline{u}] - \mathbb{P}[v])\mathbb{I}[u \neq v],$$

and we let $A(u,u) = B(u,u) = -\mathbb{P}[u]$. Then, notice that

$$\mathcal{I} = A - B$$

On the other hand, we claim that

$$\Psi_{\pi_n} = \begin{bmatrix} A & -A \\ B & -B \end{bmatrix}$$

Here we are assuming that rows/columns of Ψ_n are arranged such that the first $n+1$ columns are for vertices to be in and the next $n+1$ columns are for vertices to be out. Now notice that

$$\Psi_{\pi_n}(u, \overline{v}) = \mathbb{P}[\overline{v}|u] - \mathbb{P}[\overline{v}] = (1 - \mathbb{P}[v|u]) - (1 - \mathbb{P}[v]) = -A(u,v).$$

Similarly,

$$\Psi(\overline{u}, v) = \mathbb{P}[v|\overline{u}] - \mathbb{P}[v] = B(u,v)$$

The rest can be checked similarly.

Now, we write the characteristic polynomial of Ψ_{π_n}.

$$\begin{aligned} \det(xI - \Psi_{\pi_n}) &= \det \begin{bmatrix} xI - A & -A \\ B & xI + B \end{bmatrix} = \det((xI - A)(xI + B) + AB) \\ &= \det(x^2 I - xA + xB) \\ &= x^n \det(xI - (A - B)). \end{aligned}$$

So, the $A - B$ matrix has n extra zero eigenvalues. □

Using the above lemma together with Theorem 11.2, to prove Theorem 11.6 it is enough to prove the following theorem.

Theorem 11.8. *If $\lambda < (1-\delta)\lambda_c(\Delta)$ then for any vertex r,*

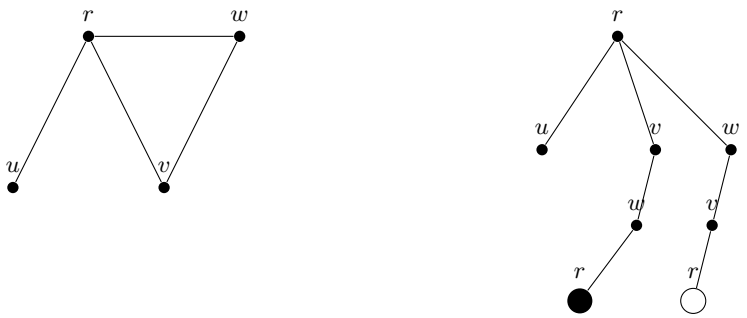
$$\sum_{v \neq r} |\mathcal{I}(r \to v)| \leq O_\delta(1).$$

This will be our focus for the rest of these lectures. The proof will be mostly based on [CLV20].

11.5 Self-avoiding Walk Tree

The main fundamental step in proving Theorem 11.8 is to reduce the theorem from arbitrary graphs G (with maximum degree Δ) to *trees* (with maximum degree Δ) in which we want to bound the maximum influence of the root to the rest of the vertices. This builds on Weitz's influential correlation decay technique [Wei06]

We start by defining the self-avoiding walk trees. Given a connected graph $G = (V, E)$ be a connected graph, and a specific vertex $r \in V$, and a total ordering of the vertex set V, the self-avoiding walk (SAW) tree rooted at r, $T_{SAW}(G, r)$ is defined as follows: It is a tree rooted at r of all paths starting at r in G except that whenever a path closes a cycle, say $r = v_0, v_1, \ldots, v_k, v_i$ where $0 \leq i \leq k-1$, the copy (in the tree) of of v_i (in G) is fixed to be occupied if $v_{i+1} < v_k$ in the total order and un-occupied otherwise. See the following picture for an example. So, observe that there are multiple copies of every vertex of G in the tree. For each $v \in V$ we denote the set of all unfixed copies of v in $T_{SAW}(G, r)$ by $\mathcal{C}_v$.



For the sake of the proof we assume that every vertex v has a distinct activity parameter, λ_v. In that case, all copies of v from $\mathcal{C}_v$ will have the *same* activity parameter λ_v in the SAW tree. As alluded to above, we will show that for any vertex $v \neq r$, $\mathcal{I}_G(r \to v) = \sum_{\hat{v} \in \mathcal{C}_v} \mathcal{I}_T(r \to \hat{v})$.

To establish that, the idea is to look the generating polynomial of the hardcore model as a multivariate polynomial in terms of vertex activities $\{\lambda_v\}_{v \in V}$ and relate the generating polynomial of G to the generating polynomial of T. Let $\lambda = \{\lambda_v\}_{v \in V}$ denote the vector of vertex activities. We define the partition function,

$$g(\lambda) = \sum_{I \text{ independent set}} \prod_{v \in I} \lambda_v$$

Theorem 11.9. *Let $G = (V, E)$ be a connected graph, $r \in V$ be a vertex such that G is connected. Let $T = T_{SAW}(G, r)$ be the self-avoiding walk tree of G rooted at r. Then, $g_G(\lambda)$ divides $g_T(\lambda)$. More precisely, there exists a polynomial $p_{G(r)} = p_{G(r)}(\lambda_{-r})$ such that*

$$g_T = g_G \cdot p_{G(r)}$$

For a vertex u we write $g_{G,u}$ to denote the generating polynomial of all independent sets that contain u and similarly we write $g_{G,\overline{u}}$ to denote the polynomial that u is out. First, we use the above theorem to prove the following lemma.

Lemma 11.10. *For any vertex $v \neq r$, $\mathcal{I}_G(r \to v) = \sum_{\hat{v} \in \mathcal{C}_v} \mathcal{I}_T(r \to \hat{v})$*

Proof. First, notice

$$\begin{aligned}
\lambda_v \partial_{\lambda_v} \log \frac{g_{G,r}(\lambda)}{g_{G,\overline{r}}(\lambda)} &= \frac{g_{G,\overline{r}}(\lambda)}{g_{G,r}(\lambda)} \cdot \lambda_v \partial_{\lambda_v} \frac{g_{G,r}(\lambda)}{g_{G,\overline{r}}(\lambda)} \\
&= \frac{g_{G,r,v}(\lambda) g_{G,\overline{r}}(\lambda) - g_{G,\overline{r},v}(\lambda) \cdot g_{G,r}(\lambda)}{g_{G,\overline{r}}(\lambda) \cdot g_{G,r}(\lambda)} \\
&= \mathbb{P}[v|r] - \mathbb{P}[v|\overline{r}] = \mathcal{I}_G(r \to v). \qquad (11.3)
\end{aligned}$$

In other words, the above calculations follows by a simple fact that if $g(z_1, \ldots, z_n)$ is a generating polynomial of a probability distribution over n items, then for any i, the marginal of i is exactly equal to $z_i \partial_{z_i} \log g$.

On the other hand, recall that for the SAW tree T , every free copy $\hat{v}$ of v has the same activity $\lambda_{\hat{v}} = \lambda_v$. So, by the above theorem,

$$\begin{aligned}
\lambda_v \partial_{\lambda_v} \log \frac{g_{G,r}(\lambda)}{g_{G,\overline{r}}(\lambda)} &= \lambda_v \partial_{\lambda_v} \log \frac{g_{T,r}(\lambda)}{g_{T,\overline{r}}(\lambda)} && \text{(Theorem 11.9)} \\
&= \sum_{\hat{v} \in \mathcal{C}_v} \lambda_{\hat{v}} \partial_{\lambda_{\hat{v}}} \log \frac{g_{T,r}(\lambda)}{g_{T,\overline{r}}(\lambda)} \cdot \frac{\partial \lambda_{\hat{v}}}{\partial \lambda_v} \lambda_{\hat{v}}(\lambda_v) && \text{(Chain Rule)} \\
&= \sum_{\hat{v} \in \mathcal{C}_v} \mathcal{I}_T(r \to \hat{v})
\end{aligned}$$

This completes the proof of the lemma. □

11.6 Reduction to Self Avoiding Walk Tree

In this section we prove Theorem 11.9. The proof is an inductive argument in which we condition on additional vertices of the graph G to be in/out. Therefore, we will need a stronger inductive hypothesis. For $\Lambda \subseteq V$ and a partial configuration $\sigma_\Lambda \in \{0,1\}^\Lambda$, we define the SAW tree with conditioning σ_Λ by assigning the configuring σ_v to every copy $\hat{v}$ of v from $\mathcal{C}_v$ and removing all descendants of $\hat{v}$ (from the tree), for each $v \in \Lambda$. Recall that in general, different copies of v from $\mathcal{C}_v$ can receive different in/out assignments. We define the generating polynomial $g^{\sigma_\Lambda}(.)$ to denote the generating polynomial of all independent sets consistent with the status of the set Λ of vertices.

We inductively prove that, there is a polynomial $p^{\sigma_\Lambda}_{G(r)}(\lambda)$ (that is independent of λ_r such that

$$g^{\sigma_\Lambda}_{T,r} = g^{\sigma_\Lambda}_{G,r} \cdot p^{\sigma_\Lambda}_{G(r)} \qquad \text{and} \qquad g^{\sigma_\Lambda}_{T,\overline{r}} = g^{\sigma_\Lambda}_{G,\overline{r}} \cdot p^{\sigma_\Lambda}_{G(r)}$$

We induct on the number of edges with (at least) one endpoint in the set $V \setminus \Lambda$.

Suppose that the root r has d neighbors $v_1, \ldots, v_d$ in G. Define G' to be the graph obtained by replacing the vertex r with d vertices $r_1, \ldots, r_d$ and then connecting $\{r_i, d_i\}$ for $1 \le i \le d$. For simplicity, we assume that $(G \setminus \{r\}) \setminus \Lambda$ is still connected. For each i, let $G_i = G' - r_i$. Consider the hardcore model on $G_i^{\sigma_\Lambda}$ together with an additional conditioning that the vertices $r_1, \ldots, r_{i-1}$ are fixed to be **out** while $r_{i+1}, \ldots, r_d$ are fixed to be **in**; we denote this conditioning by σ_{U_i} with $U_i := \{v_1, \ldots, v_d\} \setminus \{v_i\}$. Then, $T = T_{SAW}(G, r)$ can be generated by the following recursive procedure.

Step 1) For each i, let $T_i = T_{SAW}(G_i, v_i)$ plus the conditioning σ_{U_i};

Step 2) Let $T = T_{SAW}(G, r)$ be the union of r and $T_1, \ldots, T_d$ by connecting $\{r, v_i\}$ for $1 \le i \le d$; output T.

Observe that this algorithm exactly corresponds to the definition of the self-avoiding walk tree we gave in the previous section.

For the purpose of the proof we set $\lambda_{r_i} = 1$ for all $1 \le i \le d$ instead of λ_r (this is basically how we will avoid λ_r in as a parameter of $p^{\sigma_\Lambda}_{G,r}$). Observe that by definition

$$g^{\sigma_\Lambda}_{G,r} = \lambda_r g^{\sigma_\Lambda}_{G', r_1, \ldots, r_d} \qquad g^{\sigma_\Lambda}_{G,\overline{r}} = g^{\sigma_\Lambda}_{G', \overline{r}_1, \ldots, \overline{r}_d} \tag{11.4}$$

The main observation is that the graph G_i has one edge less than G, so by induction hypothesis, its generating

polynomial divides the generating polynomial of a tree. Define $\Lambda_i = \Lambda \cup U_i$. We write

$$\begin{aligned} g^{\sigma_\Lambda}_{T,r} &= \lambda_r \prod_{i=1}^{d} g^{\sigma_{\Lambda_i}}_{T_i,\overline{v}_i} && \text{(recursion of a tree)} \\ &= \lambda_r \prod_{i=1}^{d} g^{\sigma_{\Lambda_i}}_{G_i,\overline{v}_i} \cdot p^{\sigma_{\Lambda_i}}_{G_i(v_i)} && \text{(Induction Hypothesis)} \\ &= \lambda_r \prod_{i=1}^{d} g^{\sigma_\Lambda}_{G',\overline{r}_1,\ldots,\overline{r}_{i-1},r_i,\ldots,r_d} \cdot \prod_{i=1}^{d} p^{\sigma_{\Lambda_i}}_{G_i(v_i)} \\ &= g^{\sigma_\Lambda}_{G,r} \prod_{i=2}^{d} g^{\sigma_\Lambda}_{G',\overline{r}_1,\ldots,\overline{r}_{i-1},r_i,\ldots,r_d} \cdot \prod_{i=1}^{d} p^{\sigma_{\Lambda_i}}_{G_i(v_i)} && \text{(by (11.4))} \end{aligned}$$

Similarly, we can write

$$\begin{aligned} g^{\sigma_\Lambda}_{T,\overline{r}} &= \prod_{i=1}^{d} (g^{\sigma_{\Lambda_i}}_{T_i,v_i} + g^{\sigma_{\Lambda_i}}_{T_i,\overline{v}_i}) = \prod_{i=1}^{d} (g^{\sigma_{\Lambda_i}}_{G_i,v_i} \cdot p^{\sigma_{\Lambda_i}}_{G_i(v_i)} + g^{\sigma_{\Lambda_i}}_{T_i,\overline{v}_i} \cdot p^{\sigma_{\Lambda_i}}_{G_i(v_i)}) \\ &= \prod_{i=1}^{d} g^{\sigma_\Lambda}_{G',\overline{r}_1,\ldots,\overline{r}_i,r_{i+1},\ldots,r_d} \cdot \prod_{i=1}^{d} p^{\sigma_{\Lambda_i}}_{G_i(v_i)} = g^{\sigma_\Lambda}_{G,\overline{r}} \prod_{i=1}^{d-1} g^{\sigma_\Lambda}_{G',\overline{r}_1,\ldots,\overline{r}_i,r_{i+1},\ldots,r_d} \cdot \prod_{i=1}^{d} p^{\sigma_{\Lambda_i}}_{G_i(v_i)} \end{aligned}$$

The inductive step simply follows by letting $g^{\sigma_\Lambda}_{G(r)} = \prod_{i=2}^{d} g^{\sigma_\Lambda}_{G',\overline{r}_1,\ldots,\overline{r}_i,r_{i+1},\ldots,r_d} \cdot \prod_{i=1}^{d} p^{\sigma_{\Lambda_i}}_{G_i(v_i)}$

This completes the proof of Theorem 11.9. Using Theorem 11.9 and Lemma 11.10 to prove Theorem 11.8 it is enough to prove the following theorem:

Theorem 11.11. *For any Δ-ary tree T rooted at a vertex r and any $\lambda \le (1-\delta)\lambda_c(\Delta)$, we have*

$$\sum_{v} \mathcal{I}(r \to v) \le O_\delta(1).$$

11.7 Bounding Influences on a Tree

Given a tree T (where every vertex has at most $\Delta - 1$ many children (note that root can really have Δ children but we ignore that for simplicity let $L_r(k)$ be the number of vertices at distance k of the root. [CLV20] proved that if the activity parameter $\lambda \le (1-\delta)\lambda_c(\Delta)$, then we have the following bound: For any $k \ge 1$,

$$\sum_{v \in L_r(k)} \mathcal{I}(r \to v) \le 4(1-\delta/2)^{k-1}$$

Summing this up for $k = 1 \to \infty$, even if T has infinitely many vertices, we get

$$\sum_{v} \mathcal{I}(r \to v) \le 8/\delta.$$

Next, we will explain the main ideas to prove the above bound. First, for a vertex $v \in T$, let T_v be the sub-tree of T rooted at v; thus $T_r = T$. Let $R_v := \frac{g_{T_v,v}(\lambda)}{g_{T_v,\overline{v}}(\lambda)} = \frac{\mathbb{P}[v \text{ in}]}{\mathbb{P}[v \text{ out}]}$. Say a vertex u has d children $v_1, \ldots, v_d$ in the tree; the tree recursion is a formula that computes R_u given $R_{v_1}, \ldots, R_{v_d}$ due to the independence of T_{v_i}'s. More specifically, there is a function $F_d : [0,\infty]^d \to [0,\infty]$ such that

$$R_u = F_d(R_{v_1}, \ldots, R_{v_d}) := \lambda \prod_{i=1}^{d} \frac{1}{R_{v_i} + 1}.$$

We leave it as an exercise to verify the above formula.

Recall that by equation (11.3), the influence of r to a vertex u is the derivative of $\log R_r$ with respect to the external field at u. So, it is natural to define an analogue of the F_d function for the log ratio quantity. More specifically, let $H_d := [-\infty, +\infty]^d \to [-\infty, +\infty]$ defined as follows:

$$\log R_u = H_d(\log R_{v_1}, \ldots, \log R_{v_d}) := \log \lambda + \sum_{i=1}^{d} \log \frac{1}{1 + e^{\log R_{v_i}}}$$

To put it differently, $H_d = \log \circ F_d \circ \exp$.

The following lemma's are simple observations that we leave as an exercise. The first lemma follows from the fact that we are analyzing influences in a tree.

Lemma 11.12. *Suppose that $u, v, w \in T$ are three distinct vertices such that u is on the unique path from v to w. Then*

$$\mathcal{I}(v \to w) = \mathcal{I}(v \to u) \cdot \mathcal{I}(u \to w)$$

For the second lemma we need another notation: For $y \in [-\infty, \infty]$ define

$$h(y) := -\frac{e^y}{1 + e^y} = \frac{\partial}{\partial_y} H_d(y_1, \ldots, y_{i-1}, y, y_{i+1}, \ldots, y_d). \tag{11.5}$$

It follows by (11.3) that

Lemma 11.13. *For any vertex $v \in T$ and andy child u of v we have*

$$\mathcal{I}(v \to u) = h(\log R_u).$$

Having the above two lemmas we can simply write the influence of r to vertices in $L_k(r)$ inductively. Now, the main issue is that the straightforward recursion gives us terms of the form $\prod_{i=0}^{k-1} h(\log R_{u_i})$ for any path $r = u_0, \ldots, u_{k-1}, u_k$. And, in principal we can have as many as $(\Delta - 1)^k$ many such paths. A direct upper bound on such a product does not give a tight bound on the influence (that is independent of Δ) as we have to multiply the upper-bound by $(\Delta - 1)^k$.

The trick is to use a method called the potential method: Instead of tracking log ratios in the tree recursion we apply a potential function Ψ and study how $\Psi(\log R_u)$ evolves in the tree. We also let $\psi := \Psi'$ be the derivative of the potential. More precisely define

$$H_d^{\Psi} := \Psi \circ H_d \circ \Psi^{-1}.$$

We prove inductively that for any vertex $u \in T$,

$$\sum_{v \in L_u(k)} \psi(\log R_u) |\mathcal{I}(u \to v)| \leq \max_{v \in L_u(k)} \{\psi(\log R_v)\} \cdot (1 - \alpha)^k$$

where $L_u(k)$ is the set of vertices at distance k of u and α is a parameter that we choose later. The base case can be checked easily. Now, suppose the claim is checked for $k - 1$. Say u has d children $w_1, \ldots, w_d$.

We write,

$$\begin{aligned}
\sum_{v\in L_u(k)} \psi(\log R_u)|\mathcal{I}(u\to v) &= \sum_{i=1}^{d} \psi(\log R_u)|\mathcal{I}(u\to w_i)| \sum_{v\in L_{w_i}(k-1)} |\mathcal{I}(w_i\to v)| && \text{(Lemma 11.12)}\\
&= \sum_{i=1}^{d} \frac{\psi(\log R_u)}{\psi(\log R_{w_i})}|h(\log R_{w_i})| \sum_{v\in L_{w_i}(k-1)} \psi(\log R_{w_i})\mathcal{I}(w_i\to v)| && \text{(Lemma 11.13)}\\
&\le \sum_{i=1}^{d} \frac{\psi(\log R_u)}{\psi(\log R_{w_i})}|h(\log R_{w_i})| \max_{v\in L_{w_i}(k-1)} \psi(\log R_v)\cdot(1-\alpha)^{k-1} && \text{(IH)}\\
&\le \max_{v\in L_u(k)} \psi(\log R_v)(1-\alpha)^{k-1}\cdot \sum_{i=1}^{d} \frac{\psi(\log R_u)}{\psi(\log R_{w_i})}|h(\log R_{w_i})|
\end{aligned}$$

Finally, the last observation is that the quantity in the sum is exactly $\left\|\nabla H_d^{\Psi}(\Psi(\log R_{w_1}),\dots,\Psi(\log R_{w_d}))\right\|+1$. So, the main property of the potential function is that for any $y_1,\dots,y_d$ in the range of Ψ we have

$$\left\|\nabla H_d^{\Psi}(y_1,\dots,y_d)\right\|_1 \le 1-\alpha.$$

It turns out that this can be achieved for $\psi(y)=\sqrt{|h(y)|}$ and Ψ defined accordingly and for $\alpha\ge\delta/2$. In particular, we can write

$$\begin{aligned}
\sum_{i=1}^{d} \frac{\psi(\log R_u)}{\psi(\log R_{w_i})}|h(\log R_{w_i})| &= \sum_{i=1}^{d} \frac{\sqrt{h(\log R_u)}}{\sqrt{|h(\log R_{w_i})}}|h(\log R_{w_i})|\\
&= \sum_{i=1}^{d} \sqrt{\frac{\lambda\prod_{j=1}^{d}\frac{1}{1+R_{w_j}}}{1+\lambda\prod_{j=1}^{d}\frac{1}{1+R_{w_j}}}}\sqrt{\frac{R_{w_i}}{1+R_{w_i}}}
\end{aligned}$$

We leave it as an exercise to bound the RHS by $1-\delta/2$ assuming $\lambda\le(1-\delta)\lambda_c(\Delta)$. Note that $d\le\Delta-1$.

References

[AJKPV21] N. Anari, V. Jain, F. Koehler, H. T. Pham, and T.-D. Vuong. "Entropic Independence II: Optimal Sampling and Concentration via Restricted Modified Log-Sobolev Inequalities". arXiv. 2021. URL: https://arxiv.org/abs/2111.03247 (cit. on p. 11-7).

[ALO21a] D. Abdolazimi, K. Liu, and S. Oveis Gharan. "A Matrix Trickle-Down Theorem on Simplicial Complexes and Applications to Sampling Colorings". to appear. 2021 (cit. on p. 11-3).

[ALO21b] N. Anari, K. Liu, and S. Oveis Gharan. "Spectral Independence in High-Dimensional Expanders and Applications to the Hardcore Model". In: *Siam Journal of Computing* (July 2021). Special Issue of FOCS 2020, FOCS20–1–FOCS20–37 (cit. on pp. 11-6, 11-7).

[Bro86] A. Z. Broder. "How hard is it to marry at random? (On the approximation of the permanent)". In: *STOC*. 1986, pp. 50–58 (cit. on p. 11-1).

[CLV20] Z. Chen, K. Liu, and E. Vigoda. "Rapid Mixing of Glauber Dynamics up to Uniqueness via Contraction". In: *FOCS*. 2020, pp. 1307–1318 (cit. on pp. 11-7, 11-8, 11-11).

[CLV21] Z. Chen, K. Liu, and E. Vigoda. "Optimal mixing of Glauber dynamics: entropy factorization via high-dimensional expansion". In: *STOC*. Ed. by S. Khuller and V. V. Williams. ACM, 2021, pp. 1537–1550 (cit. on p. 11-7).

[DK17] I. Dinur and T. Kaufman. "High Dimensional Expanders Imply Agreement Expanders". In: *FOCS*. 2017, pp. 974–985 (cit. on p. 11-3).

[JS89] M. Jerrum and A. Sinclair. "Approximating the Permanent". In: *Siam J. of Computing* 18.6 (1989), pp. 1149–1178 (cit. on p. 11-2).

[JSV04] M. Jerrum, A. Sinclair, and E. Vigoda. "A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries". In: *Journal of the ACM* 51.4 (2004), 671–697 (cit. on pp. 11-2, 11-7).

[JVV86] M. Jerrum, L. Valiant, and V. Vazirani. *Random Generation of Combinatorial Structures from a Uniform Distribution*. 1986 (cit. on p. 11-1).

[KM17] T. Kaufman and D. Mass. "High dimensional random walks and colorful expansion". In: *ITCS*. 2017, 14–27 (cit. on p. 11-1).

[KO20] T. Kaufman and I. Oppenheim. "High Order Random Walks: Beyond Spectral Gap". In: *Combinatorica* 40 (2020), 245–281 (cit. on p. 11-3).

[Wei06] D. Weitz. "Counting independent sets up to the tree threshold". In: *STOC*. ACM, 2006, pp. 140–149 (cit. on p. 11-8).