Lecture 7: *August 10th, 2022*

Lecturer: *Kuikui Liu*

Spectral Independence

These notes are based on the lecture by Kuikui Liu. This is part of the 2022 Summer School on *New tools for optimal mixing of Markov chains: Spectral independence and entropy decay*, which was held at the University of California, Santa Barbara (UCSB) from August 8, 2022 to August 12, 2022. More information on the summer school is available at: https://sites.cs.ucsb.edu/∼vigoda/School/

7.1 Overview

In this lecture we will introduce the notion of spectral independence. For a n vertex graph, the spectral independence condition is a bound on the maximum eigenvalue of the $n \times n$ influence matrix whose entries capture the influence between pairs of vertices, it is closely related to the covariance matrix. We will show that spectral independence implies the mixing time of the Glauber dynamics is polynomial (where the degree of the polynomial depends on certain parameters). The proof utilizes the local-to-global theorems presented in Tali Kaufman's lectures and the ideas from the analysis of the bases-exchange walk for generating a random matroid. In the subsequent lecture we will show that spectral independence with some additional conditions implies an optimal mixing time bound of $O(n \log n)$ for the Glauber dynamics.

7.2 Introduction

7.2.1 Hard-core Model

Let's focus on the following simpler setting. Consider a distribution μ on $\{0, 1\}^n$. An example to keep in mind is the so-called *hard-core model*.

The input to the hard-core model is a graph $G = (V, E)$ and a parameter $\lambda > 0$. The hard-core model is an idealized model of a gas where λ corresponds to the fugacity of the gas. Configurations of the model are defined on independent sets of G; recall, an independent set is a subset S of vertices which does not contain an edge, i.e., for all $\{x, y\} \in E$ either $x \notin S$ and/or $y \notin S$. Let Ω denote the collection of independent sets of G (regardless of their sizes). In the terminology of Pietro Caputo's first lecture, this set Ω corresponds to the set Ω_0 in Caputo's notation for configurations with positive measure in the Gibbs distribution. For an independent set $\sigma \in \Omega$, we can view σ as an n-dimensional vector in $\{0, 1\}^n$ where the i-th coordinate is assigned 1 if the i-th vertex is in σ and is assigned 0 otherwise.

The Gibbs distribution is defined as:

$$\mu(\sigma) = \frac{1}{Z} \lambda^{|\sigma|} \quad \text{for } \sigma \in \Omega,$$

where $|\sigma|$ is the number of occupied vertices in the independent set σ and the partition function $Z = Z_{G,\lambda} = \sum_{\eta \in \Omega} \lambda^{|\eta|}$.

The Glauber dynamics, also known as the Gibbs sampler, is a simple Markov chain designed to sample from the Gibbs distribution. For the case of the hard-core model the chain is defined as follows. From a state $X_t \in \Omega$,

1. Choose a vertex v uniformly at random from V.
2. Let
$$X' = \begin{cases} X_t \cup \{v\} & \text{with probability } \lambda/(1+\lambda) \\ X_t \setminus \{v\} & \text{with probability } 1/(1+\lambda) \end{cases}$$
3. If $X' \in \Omega$ (i.e., it is a valid independent set) then set $X_{t+1} = X_t$ and otherwise set $X_{t+1} = X_t$.

Notice that the Glauber dynamics is irreducible (since all states can reach the empty set) and aperiodic (since there is a self-loop of not changing the state at v) and hence the chain is ergodic where the stationary distribution is the Gibbs distribution.

More generally, the Glauber dynamics is defined as follows. From a state $X_t \in \{0,1\}^n$,

1. Choose a coordinate i uniformly at random from V.
2. For all $j \neq i$, let $X_{t+1}(j) = X_t(j)$.
3. Choose $X_{t+1}(j)$ from the conditional Gibbs distribution $\mu(\sigma(i)|\sigma(j) = X_{t+1}(j)$ for all $j \neq i\})$, i.e., fix the spin at all vertices except i and resample the spin at i conditional on the fixed configuration on the rest of the vertices.

7.3 Spectral Independence

The spectral independence was introduced by Anari, Liu, and Oveis Gharan [ALO20].

Definition 7.1 (Influence Matrix)**.** *Let $G = (V, E)$ be a graph where $V = \{1, \dots, n\}$, and μ be a distribution on $\{0,1\}^n$. Let Ψ be the following real-valued $n \times n$ matrix. For $1 \leq i, j \leq n$,*

$$\Psi(i \to j) = \Psi(i,j) := \Pr_{\sigma\sim\mu}[\sigma(j) = 1 \mid \sigma(i) = 1] - \Pr_{\sigma\sim\mu}[\sigma(j) = 1 \mid \sigma(i) = 0]$$

The matrix Ψ can be asymmetric and the entries of Ψ can be positive or negative. Nevertheless all of its eigenvalues are real.

Lemma 7.2. *All eigenvalues of Ψ are non-negative real numbers.*

Proof. For every $i, j \in [n]$, the covariance of $\mathbb{1}_i, \mathbb{1}_j$ is given by

$$\begin{aligned} \mathrm{Cov}_\mu(i,j) &= \mathbb{E}_\mu[\mathbb{1}_i \mathbb{1}_j] - \mathbb{E}_\mu[\mathbb{1}_i] \cdot \mathbb{E}_\mu[\mathbb{1}_j] \\ &= \Pr_{\sigma\sim\mu}[\sigma(i) = \sigma(j) = 1] - \Pr_{\sigma\sim\mu}[\sigma(i) = 1] \cdot \Pr_{\sigma\sim\mu}[\sigma(j) = 1] \\ &= \Pr[\sigma(i) = 1] \times \Big(\Pr[\sigma(j) = 1 \mid \sigma(i) = 1] - \Pr[\sigma(j) = 1] \Big) \end{aligned}$$

Plugging in

$$\Pr[\sigma(j) = 1] = \Pr[\sigma(j) = 1 \mid \sigma(i) = 1] \cdot \Pr[\sigma(i) = 1] + \Pr[\sigma(j) = 1 \mid \sigma(i) = 0] \cdot \Pr[\sigma(i) = 0],$$

we obtain that

$$\begin{aligned}\mathrm{Cov}_\mu(i,j) &= \Pr[\sigma(i)=1] \times \Big(\Pr[\sigma(j)=1 \mid \sigma(i)=1] - \Pr[\sigma(j)=1 \mid \sigma(i)=1] \cdot \Pr[\sigma(i)=1] \\ &\qquad - \Pr[\sigma(j)=1 \mid \sigma(i)=0] \cdot \Pr[\sigma(i)=0] \Big) \\ &= \Pr[\sigma(i)=1] \times \Big(\Pr[\sigma(j)=1 \mid \sigma(i)=1](1-\Pr[\sigma(i)=1]) \\ &\qquad - \Pr[\sigma(j)=1 \mid \sigma(i)=0] \cdot \Pr[\sigma(i)=0] \Big) \\ &= \Pr[\sigma(i)=1] \cdot \Pr[\sigma(i)=0]) \times \Big(\Pr[\sigma(j)=1 \mid \sigma(i)=1] - \Pr[\sigma(j)=1 \mid \sigma(i)=0] \Big).\end{aligned}$$

From the definition of the influence matrix Ψ we have:

$$\mathrm{Cov}_\mu(i,j) = \Pr[\sigma(i)=1] \cdot \Pr[\sigma(i)=0] \cdot \Psi_\mu(i,j).$$

Since this holds for all $i,j \in [n]$, if we let D denote the diagonal matrix with $D(i,i) = \Pr[\sigma(i) = 1] \cdot \Pr[\sigma(i)=0]$ for all $i \in [n]$, then we have the matrix identity

$$\Psi_\mu = D^{-1}\mathrm{Cov}_\mu.$$

Since Cov_μ is symmetric positive semidefinite and D is a diagonal matrix with positive diagonal entries, it follows that Ψ_μ has nonnegative real eigenvalues. □

Since all eigenvalues of Ψ are real numbers we can denote the maximum eigenvalue by $\lambda_{\max}(\Psi)$. Now we can define spectral independence.

Definition 7.3 (Spectral Independence)**.** *For $\eta > 0$, we say that μ is η-spectrally independent if $\lambda_{\max}(\Psi) \le 1+\eta$.*

Note that the spectral independence condition only depends on the distribution μ, it is not a function of the Glauber dynamics. Moreover, the definition does not require that μ is a Gibbs distribution. The definition was extended to non-binary spin spaces, such as the Potts model and colorings, in [FGYZ21, CGŠV21] (see also [CLV21, BCC+22] for a general formulation).

When μ is a product distribution then $\eta = 0$. Our goal is to show that η is constant.

Remark 7.4. Note the diagonals of the influence matrix Ψ are 1 since if $i = j$ then conditioning on i prescribes j. We could have defined the influence matrix so that the off-diagonal entries remain the same and the diagonals are 0; this would decrease all of the eigenvalues by 1, and hence with this alternative definition we would change the spectral independence requirement from $1+\eta$ to η.

7.4 Rapid mixing

The main result of this lecture is that if μ is η-spectrally independent for a constant η and for all pinnings the conditional distribution is also η-spectrally independent, then the Glauber dynamics has polynomial mixing time. Recall from Pietro Caputo's Lecture 1 that a pinning is a fixed assignment of spins to a subset of vertices. Hence, for $S \subset [n] = \{1,\dots,n\}$ and for a pinning $\tau : S \to \{0,1\}$ then let μ^τ denote the conditional Gibbs distribution, i.e., the distribution μ conditional on the fixed assignment τ on S.

We will bound the mixing time of the Glauber dynamics by considering the spectral gap. Let $\gamma = 1-\lambda_2$ where $1 > \lambda_2 > \cdots \ge 0$ are the eigenvalues of the (lazy version of the) Glauber dynamics. The following result was proved by Anari, Liu, and Oveis Gharan [ALO20].

Theorem 7.5 ([ALO20])**.** *Let μ be a probability distribution on $\{0,1\}^n$. Suppose there exists $\eta > 0$ such that for all $S \subset [n]$, all $\tau : S \to \{0,1\}$, the conditional distribution μ^τ is η-spectrally independent, then the spectral gap of the Glauber dynamics satisfies:*

$$\gamma(P_{\text{Glauber}}) \geq \Omega\left(n^{-(1+\eta)}\right).$$

Therefore, the mixing time satisfies $T_{\text{mix}} = O(n^{2+\eta} \log n)$.

Therefore, the goal is to prove that $\eta = O(1)$. This has been established in a variety of settings, including the hard-core model in the tree uniqueness region; this will be explained in Zongchen Chen's lecture.

7.5 Proof of Theorem 7.5

The proof of Theorem 7.5 will follow from the random walk theorem of Alev and Lau [AL20], which improved upon Kaufman and Oppenheim [KO20]; this result was presented and proved in Tali Kaufman's lectures.

Consider the following simplicial complex X. The ground set are the (vertex, spin) pairs $\{(i, s_i) : 1 \leq i \leq n, s_i \in \{0,1\}\}$. The faces are "consistent" sets of (vertex, spin) pairs (in particular, we have the following: every vertex occurs in at most one pair).

Let's begin by defining the local walks Q_μ introduced in Tali Kaufman's lectures. The matrix Q_μ is a real-valued $2n \times 2n$ matrix. For $1 \leq i \neq j \leq n$, $s_i, s_j \in \{0,1\}$, let

$$Q_\mu((i, s_i), (j, s_j)) = \frac{1}{n-1} \Pr_{\sigma \sim \mu}[\sigma(j) = s_j \mid \sigma(i) = s_i].$$

The Glauber dynamics is the down-up walk on levels $(n, n-1)$. Hence the random walk theorem implies:

$$\gamma(P_{\text{Glauber}}) \geq \frac{1}{n} \prod_{k=0}^{n-2} (1 - \lambda_k), \tag{7.1}$$

where

$$\lambda_k = \max_{S \subset [n] : |S| = k} \max_{\tau \in \{0,1\}^S} \lambda_2(Q_{\mu^\tau})$$

is the second largest eigenvalue for the local walk with a worst-case link on level k, i.e., pinning k vertices.

We begin by bounding λ_2 of the local walk Q_μ in terms of the influence matrix Ψ_μ for the spectral independence technique. In fact, we can relate the entire spectrum of Q_μ with the spectrum of Ψ_μ in the following manner.

Lemma 7.6. *We have*

$$\lambda_2(Q_\mu) = \frac{1}{n-1} \left(\lambda_{\max}(\Psi_\mu) - 1\right).$$

Moreover,

$$\textit{spectrum}(Q_\mu) = \textit{spectrum}\left(\frac{1}{n-1}(\Psi_\mu - I)\right) \cup \{1\} \cup \left\{n-1 \textit{ copies of } \frac{-1}{n-1}\right\}. \tag{7.2}$$

Proof. To prove this lemma, we take the local random walk Q_μ and we "zero-out" the trivial eigenvalues of 1 and $-1/(n-1)$; the latter come from the n-partite structure of the graph corresponding to Q_μ, one part for each coordinate/vertex. This is obtained as follows. Let

$$M_\mu = Q_\mu - \frac{n}{n-1}\mathbf{1}\mu^T + \frac{1}{n-1}\sum_{i=1}^{n}\mathbf{1}_i\mu_i^T,$$

where μ is the stationary distribution of the local walk (that is, $\mu((j,s_j)) = (1/n)\Pr_{\sigma\sim\mu}[\sigma(j)=s_j]$) and μ_i is the marginal distribution at vertex i (that is, $\mu_j((j,s_j)) = \Pr_{\sigma\sim\mu}[\sigma(j)=s_j]$, with the remaining entries 0).

Then we notice that M_μ has the following block structure:

$$M_\mu = \begin{pmatrix} A_\mu & -A_\mu \\ B_\mu & -B_\mu \end{pmatrix}, \tag{7.3}$$

where

$$A_\mu - B_\mu = \frac{1}{n-1}(\Psi_\mu - I). \tag{7.4}$$

To see (7.3) and (7.4) note that for $i \neq j$ we have

$$M_\mu((i,s_i),(j,s_j)) = \frac{1}{n-1}\Big(\Pr_{\sigma\sim\mu}[\sigma(j)=s_j \mid \sigma(i)=s_i] - \Pr_{\sigma\sim\mu}[\sigma(j)=s_j]\Big),$$

and

$$M_\mu((i,s_i),(i,s_i)) = M_\mu((i,s_i),(i,1-s_i)) = 0.$$

Note that if w is a left-eigenvector of $A_\mu - B_\mu$ then $\binom{w}{-w}$ is a left-eigenvector of M_μ with the same eigenvalue. Note that the vectors of the form $(v^T v^T)$ (a space of dimension n) are 1) right-eigenvectors of M_μ with eigenvalue 0 and 2) are perpendicular to the left-eigenvectors of the form $\binom{w}{-w}$. The vectors of the form $(v^T v^T)$ yield right-eigenvectors of $\frac{n}{n-1}\mathbf{1}\mu^T - \frac{1}{n-1}\sum_{i=1}^{n}\mathbf{1}_i\mu_i^T$, where $\mathbf{1}$ has eigenvalue 1 and the subspace perpendicular to μ (a space of dimension $n-1$) has eigenvectors with eigenvalue $-\frac{1}{n-1}$. This implies (7.2). □

We can now utilize Lemma 7.6 with (7.1) to conclude the main theorem about the spectral gap of the Glauber dynamics.

Proof of Theorem 7.5. Recall, the definition of spectral independence in Defintion 7.3. First of all it states that the maximum eigenvalue is at most $1+\eta$. This means that, from Lemma 7.6, we get that $\lambda_2(Q_\mu) \leq \eta/(n-1)$. Moreover, Definition 7.3 is for the worst-case pinning and hence we can apply Lemma 7.6 to any pinning. Hence, we have that $\lambda_k \leq \eta/(n-k-1)$ (where λ_k is defined below (7.1)). Note, when $n-k = O(1)$ then $\lambda_k < C$ for some constant $C < 1$.

Now we can apply (7.1) and we obtain the following:

$$\begin{aligned}
\gamma(P_{\text{Glauber}}) &\geq \frac{1}{n}\prod_{k=0}^{n-2}(1-\lambda_k) \\
&\geq (1-C)^{\eta} \times \frac{1}{n}\prod_{k=0}^{n-2-\eta}(1-\lambda_k) \\
&\geq (1-C)^{\eta} \times \frac{1}{n}\prod_{k=0}^{n-2-\eta}\left(1-\frac{\eta}{n-k-1}\right) \\
&\geq \frac{C'}{n}\prod_{k=0}^{n-2-\eta}\left(1-\frac{\eta}{n-k-1}\right) \\
&\geq \frac{C'}{n}\exp\left(-\sum_{k=0}^{n-2-\eta}\frac{\eta}{n-k-1-\eta}\right) && \text{since } 1-x \geq \exp\left(-\tfrac{x}{1-x}\right) \\
&\geq \frac{C''}{n}\exp(-\eta\ln(n-1-\eta)) \\
&= \frac{C''}{n}(n-1-\eta)^{-\eta} \\
&= \Omega\left(n^{-1-\eta}\right).
\end{aligned}$$

□

References

[AL20] Vedat Levi Alev and Lap Chi Lau. Improved analysis of higher order random walks and applications. In *Proceedings of the 52nd Annual ACM Symposium on Theory of Computing (STOC)*, pages 1198–1211, 2020.

[ALO20] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. In *Proceedings of the 61st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 1319–1330, 2020.

[BCC$^+$22] Antonio Blanca, Pietro Caputo, Zongchen Chen, Daniel Parisi, Daniel Štefankovič, and Eric Vigoda. On Mixing of Markov Chains: Coupling, Spectral Independence, and Entropy Factorization. In *Proceedings of the 33rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3670–3692, 2022.

[CGŠV21] Zongchen Chen, Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Rapid mixing for colorings via spectral independence. In *Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1548–1557, 2021.

[CLV21] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of Glauber dynamics: Entropy factorization via high-dimensional expansion. In *Proceedings of the 53rd Annual ACM Symposium on Theory of Computing (STOC)*, pages 1537–1550, 2021.

[FGYZ21] Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. Rapid mixing from spectral independence beyond the Boolean domain. In *Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1558–1577, 2021.

[KO20] Tali Kaufman and Izhar Oppenheim. High Order Random Walks: Beyond Spectral Gap. *Combinatorica*, 40(2):245–281, 2020.