Approximate Counting and Mixing Time of Markov Chains **Fall 2024**

Lecture 11: Entropic Independence

Lecturer: Shayan Oveis Gharan

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

Recall the following definitions from lecture 9: We have a graph $G = (V, E)$ we want sample from the state space $\{\pm 1\}^V$. For $\sigma \in \{\pm 1\}^V$ let

$$\sigma^{\oplus i}(j) = \begin{cases} -\sigma_j & \text{if } j = i \\ \sigma_j & \text{otherwise} \end{cases}$$

In GD we first we choose a u.r. vertex i then, with probability $\frac{\pi(\sigma^{\oplus i})}{\pi(\sigma)+\pi(\sigma^{\oplus i})}$ we move to $\sigma^{\oplus i}$ and otherwise we stay.

Pinning. Let π be a distribution on $\{\pm 1\}^V$. For a any set of vertices $i_1, \dots, i_k$ (for any $1 \le k < n$) and signs $s_1, \dots, s_k$ we let

$$\pi_{(i_1,s_1),\dots,(i_k,s_k))} := \pi | \sigma_{i_1} = s_1, \dots, \sigma_{i_k} = s_k.$$

In other words, this is the conditional measure on all vertices in $V - \{i_1, \dots, i_k\}$ when we pin $i_1, \dots, i_k$ to signs $s_1, \dots, s_k$ respectively.

Averaging / Projection. Conversely, given π and a set $S \subseteq V$ we let π^S to be distribution π projected onto the set S when we average out all vertices outside of S. In other words,

$$\pi^S(\tau \in \{\pm 1\}^S) = \sum_{\sigma : \sigma_S = \tau} \pi(\sigma).$$

11.1 Entropic Independence

We say μ is η-entropically independent if for any function $f : \{\pm 1\}^n \to \mathbb{R}_{\ge 0}$

$$\left(1 - \frac{1+\eta}{n}\right) \operatorname{Ent}(f) \le \mathbb{E}_i \mathbb{E}_{s\sim\pi^i} \operatorname{Ent}_{\pi_{i,s}} f.$$

where as before,

$$\operatorname{Ent}_\pi f = D_\pi^{x \log x} f = \mathbb{E}_\pi f \log f - \mathbb{E} f \log \mathbb{E} f$$

Theorem 11.1. *Suppose that π and all conditionals of π are η-entropically independent. Then, the modified log-sobolev constant of the Glauber dynamics is at least $M(GD) \ge \Omega(1/n^{1+\eta})$.*

First, by repeated application of the Entropic independent on π and pinnings of π we can write,

$$\begin{aligned}
\mathrm{Ent}(f) &\le \left(1-\frac{1+\eta}{n}\right)^{-1} \mathbb{E}_i \mathbb{E}_{\sigma_i\sim\pi^i} \mathrm{Ent}_{\pi_{i,\sigma_i}}(f) && \text{(entropic Ind of } \pi) \\
&\le \left(1-\frac{1+\eta}{n}\right)^{-1}\left(1-\frac{1+\eta}{n-1}\right)^{-1} \mathbb{E}_{i,j}\mathbb{E}_{\sigma_i,\sigma_j\sim\pi^{i,j}} \mathrm{Ent}_{\pi_{i,\sigma_i,j,\sigma_j}}(f) && \text{(entropic Ind of } \pi_{i,\sigma_i}) \\
&\dots \\
&\le \prod_{j=0}^{k-1}\left(1-\frac{1+\eta}{n-j}\right)^{-1} \mathbb{E}_{S\sim\binom{n}{k}}\mathbb{E}_{\sigma\sim\pi^S} \mathrm{Ent}_{\pi_{S,\sigma}}(f) \\
&\dots \\
&\le \prod_{j=0}^{n-1}\left(1-\frac{1+\eta}{n-j}\right)^{-1} \mathbb{E}_{S\sim\binom{V}{n-1}}\mathbb{E}_{\sigma\sim\pi^S} \mathrm{Ent}_{\pi_{S,\sigma}}(f) \\
&\lesssim 2\exp\left((1+\eta)\sum_{i=0}^{n-1}\frac{1}{n-i}\right)\mathcal{E}(f,\log f) \lesssim 2n^{1+\eta}\mathcal{E}(f,\log f)
\end{aligned}$$

Let us explain how to prove the second to last inequality. It can be shown that for a two state system $\{+,-\}$ with $K(+\to-)=p$ and $K(-\to+)=1-p$ (and with remaining probability we stay), the modified log-sobolev constant is at least $1/2$ (see [BT06] for a proof). In other words for such a system and any function $f:\{\pm1\}\to\mathbb{R}$ we have,

$$\frac{\mathcal{E}(f,\log f)}{\mathrm{Ent}_\pi(f)} \ge 1/2 \Rightarrow \mathrm{Ent}_\pi(f) \le 2\mathcal{E}(f,\log f).$$

From this we can write the Dirichlet form as follows:

$$\begin{aligned}
\mathcal{E}(f,\log f) &= \frac{1}{2}\mathbb{E}_{\sigma\sim\{\pm1\}^n}\mathbb{E}_i \frac{\pi(\sigma^{\oplus i})}{\pi(\sigma)+\pi(\sigma^{\oplus i})}(f(\sigma)-f(\sigma^{\oplus i}))(\log f(\sigma)-\log f(\sigma^{\oplus i})) \\
&= \frac{1}{2}\mathbb{E}_i\sum_{\sigma\in\{\pm1\}^V}(\pi(\sigma)+\pi(\sigma^{\oplus i}))\cdot\frac{\pi(\sigma^{\oplus i})}{\pi(\sigma)+\pi(\sigma^{\oplus i})}\cdot\frac{\pi(\sigma)}{\pi(\sigma)+\pi(\sigma^{\oplus i})}\cdot(f(\sigma)-f(\sigma^{\oplus i}))(\log f(\sigma)-\log f(\sigma^{\oplus i})) \\
&= \mathbb{E}_i\sum_{\sigma\in\{\pm1\}^V}(\pi(\sigma)+\pi(\sigma^{\oplus i}))\cdot\mathcal{E}_{\pi_{\sigma-i}}(f,\log f) \\
&\ge \frac{1}{2}\mathbb{E}_i\mathbb{E}_{\sigma\sim\pi^{V-i}}\mathrm{Ent}_{\pi_\sigma}(f).
\end{aligned}$$

where as before $\pi_{\sigma-i}$ is the pinning of π on all vertices in $V-i$ according to σ.

11.2 Main Theorem

The main goal of this lecture is to prove the following theorem.

Theorem 11.2. *Suppose π is a distribution on $\{\pm1\}^n$ such that π and all pinnings of π are η-spectrally independent. If π is B-marginally bounded; namely for any i and $s\in\{\pm1\}$, $\mathbb{P}[\sigma_i=s]\ge B$, then π and all pinnings of π are $O(\eta/B^2)$-entropically independent.*

Having proven that, we can immediately follow the proof technique that we discussed before: Namely use the entropic independence for $k=(1-\theta)n$ for $\theta\ll 1/\Delta$ such that the resulting graph is decomposed into constant size components (in expectation) and then simply use the fact that the MLS constant of a constant size chain is bounded away from 1 to prove the Glauber dynamics mixes in $O_{\eta,\Delta}(n \log n)$. This implies "optimal" mixing result for the Glauber dynamics. This is summarized in the following theorem:

Theorem 11.3. *Let π be a probability measure on $\{\pm 1\}^n$ s.t.,*

- *π and all pinnings of π are η-spectrally independent.*
- *There exists a graph $G = ([n], E)$ of maximum degree Δ such that all pairwise interactions are only defined on edges of G. In other words, For any disjoint sets $S, T \subseteq [n]$, such that there are no edges between S, T, σ_S is independent of σ_T conditioned on $\sigma_{N(S)}$*
- *π is B-marginally bounded.*

Then, $M(GD) \geq \Omega_{\eta,\Delta,B}(1/n)$ so the Glauber Dynamics mixes in $O_{\eta,\Delta,B}(n \log n)$.

11.3 Entropic Independence from Local Entropy Contraction

In this section we prove Theorem 11.2. In previous lectures, for a function $f : \{\pm 1\}^n \to \mathbb{R}$ we defined $\pi^1 = \mathbb{E}_i \pi^i$ and $f^1 = \mathbb{E}_{\pi^1}[f]$. This definition naturally generalizes to k:

$$\pi^k = \mathbb{E}_{S \in \binom{n}{k}} \pi^S, \quad f^k(S, \tau) = \mathbb{E}_{\sigma \sim \pi}[f(\sigma) | \sigma_S = \tau], S \in \binom{n}{k}, \tau \in \{\pm 1\}^S.$$

We also discussed the law of total variance which naturally generalize to the law of total entropy proved below.

Lemma 11.4 (Law of Total Entropy)**.** *Let X, Y be random variable jointly defined. Then,*

$$\text{Ent}[Y] = \text{Ent}[\mathbb{E}[Y|X]] + \mathbb{E}[\text{Ent}[Y|X]]$$

Proof. First, we write

$$\mathbb{E}Y \log Y = \mathbb{E}_X[\mathbb{E}[Y \log Y | X]] = \mathbb{E}_X[\text{Ent}[Y|X] + \mathbb{E}[Y|X] \cdot \log \mathbb{E}[Y|X]]$$

On the other hand,

$$\text{Ent}\,\mathbb{E}[Y|X] = \mathbb{E}_X \mathbb{E}[Y|X] \log \mathbb{E}[Y|X] - \mathbb{E}_X \mathbb{E}[Y|X] \log \mathbb{E}_X \mathbb{E}[Y|X] = \mathbb{E}_X[\mathbb{E}[Y|X] \log \mathbb{E}[Y|X]] - \mathbb{E}Y \log \mathbb{E}Y.$$

Therefore,

$$\begin{aligned}
\text{Ent}[Y] &= \mathbb{E}Y \log Y - \mathbb{E}[Y] \log \mathbb{E}[Y] \\
&= \mathbb{E}Y \log Y + \text{Ent}\,\mathbb{E}[Y|X] - \mathbb{E}_X[\mathbb{E}[Y|X] \log \mathbb{E}[Y|X]] \\
&= \text{Ent}\,\mathbb{E}[Y|X] + \mathbb{E}_X \,\text{Ent}[Y|X]
\end{aligned}$$

as desired. □

The following is an immediate consequence of the law of total entropy:

Corollary 11.5. *For $\ell > k$*

$$\text{Ent}_{\pi^\ell} f^\ell = \text{Ent}_{\pi^k} f^k + \mathbb{E}_{S,\sigma \sim \pi^k} \text{Ent}_{\pi_{S,\sigma}} f^{\ell-k}_{S,\sigma}$$

To see this we let X be the random variable which chooses a set $S \in \binom{[n]}{k}$ u.a.r. and a random signing of S (and $Y = f^\ell(S \cup T, \sigma_S \cup \sigma_T)$ for $T \in \binom{[n]-S}{\ell}$).

Definition 11.6. *We say π is satisfies α-local entropy contraction if for any function $f : \{\pm 1\}^n \to \mathbb{R}_{\geq 0}$,*

$$\mathrm{Ent}_{\pi^1} f^1 \leq \frac{1}{2}(1 - \frac{\alpha}{n})^{-1} \mathrm{Ent}_{\pi^2} f^2.$$

Lemma 11.7. *Suppose π and all pinnings of π are η-spectrally and B-marginally bounded. Then π satisfies α-local entropy contraction for $\alpha = O(\eta/B^2)$.*

Proof of Theorem 11.2. Our main goal is to prove that

$$\mathrm{Ent}_{\pi^1}(f^1) \leq \frac{1 + O(\eta/B^2)}{n} \mathrm{Ent}_\pi f. \tag{11.1}$$

Then, by an immediate application of law of total entropy we get $O(\eta/B^2)$-entropic independence.

First, by law of total entropy we can write,

$$\begin{aligned} \mathrm{Ent}_{\pi^k} f^k - \mathrm{Ent}_{\pi^{k-1}} f^{k-1} &= \mathbb{E}_{S,\sigma\sim\pi^{k-1}} \mathrm{Ent}_{\pi^1_{S,\sigma}} f^1_{S,\sigma} \\ \mathrm{Ent}_{\pi^{k+1}} f^{k+1} - \mathrm{Ent}_{\pi^{k-1}} f^{k-1} &= \mathbb{E}_{S,\sigma\sim\pi^{k-1}} \mathrm{Ent}_{\pi^2_{S,\sigma}} f^2_{S,\sigma} \end{aligned}$$

Now, by Lemma 11.7 we can write

$$\begin{aligned} \mathrm{Ent}_{\pi^{k+1}} f^{k+1} - \mathrm{Ent}_{\pi^{k-1}} f^{k-1} &= \mathbb{E}_{S,\sigma\sim\pi^{k-1}} \mathrm{Ent}_{\pi^2_{S,\sigma}} f^2_{S,\sigma} \\ &\geq 2\left(1 - \frac{\alpha}{n-k+1}\right) \mathbb{E}_{S,\sigma\sim\pi^k} \mathrm{Ent}_{\pi^1_{S,\sigma}} f^1 \\ &= 2\left(1 - \frac{\alpha}{n-k+1}\right) (\mathrm{Ent}_{\pi^k} f^k - \mathrm{Ent}_{\pi^{k-1}} f^{k-1}) \end{aligned}$$

The rest of the proof is a simple induction with the right induction hypothesis. In particular suppose for $\beta_k := \sum_{j=0}^{k-1} \prod_{i=0}^{j-1} (1 - \frac{2\alpha}{n-i})$. Then, we show by induction that

$$\frac{\mathrm{Ent}_{\pi^{k-1}} f^{k-1}}{\beta_{k-1}} \leq \frac{\mathrm{Ent}_{\pi^k} f^k}{\beta_k}. \tag{11.2}$$

Re-arranging the above equation and using IH we can write

$$\begin{aligned} \mathrm{Ent}_{\pi^{k+1}} f^{k+1} &\geq 2\left(1 - \frac{\alpha}{n-k+1}\right) \mathrm{Ent}_{\pi^k} f^k - \left(1 - \frac{2\alpha}{n-k+1}\right) \mathrm{Ent}_{\pi^{k-1}} f^{k-1} \\ &\geq \left(2\left(1 - \frac{\alpha}{n-k+1}\right) - \left(1 - \frac{2\alpha}{n-k+1}\right)\frac{\beta_{k-1}}{\beta_k}\right) \mathrm{Ent}_{\pi^k} f^k \qquad \text{(IH)} \\ &= \left(1 - \left(1 - \frac{2\alpha}{n-k+1}\right) - \left(\frac{\beta_{k-1}}{\beta_k} - 1\right)\right) \mathrm{Ent}_{\pi^k} f^k \\ &= \frac{\beta_{k+1}}{\beta_k} \mathrm{Ent}_{\pi^k} f^k \end{aligned}$$

The last equation follow simply by the definition of $\beta_{k-1}, \beta_k, \beta k + 1$. This proves (11.2). Now, using (11.2) repeatedly we can write

$$\mathrm{Ent}_{\pi^1} f^1 \leq \frac{1}{\beta_n} \mathrm{Ent}_\pi f, \tag{11.3}$$

and by definition:

$$\beta_n := \sum_{j=0}^{n-1}\prod_{i=0}^{j-1}(1-\frac{2\alpha}{n-i}) \gtrsim \sum_{j=0}^{n-1}\left(\frac{n}{n-j}\right)^{-2\alpha}$$

We can easily see $\beta_n \geq \Omega(\eta n/B^2)$ for $\alpha = O(\eta/B^2)$. □

11.4 Local Entropy Contraction

In this section we explain the main ideas in the proof of Lemma 11.7.

Lemma 11.8. *If π is η-spectrally independent then,*

$$\mathrm{Ent}_{\pi^2} f^2 - 2\,\mathrm{Ent}_{\pi^1} f^1 \geq -\frac{\eta}{n-1}\frac{\mathrm{Var}(f_1)}{\mathbb{E}f_1}$$

Proof. First, since this equation is scale invariant, i.e., entropy remains invariant under re-scaling f, we assume $\mathbb{E}f = 1$. Recall that

$$\pi^2((i,s_1),(j,s_2)) = \frac{1}{\binom{n}{2}}\mathbb{P}\left[\sigma_i = s_1, \sigma_j = s_2\right].$$

Note that for any $x \in [n]\times\{\pm\}$,

$$\sum_{y\in[n]\times\{\pm\}} \pi^2(x,y) = 2\cdot\pi^1(x)$$

This is mainly because we define π^2 on unordered **sets** of size 2.

We write,

$$\begin{aligned}\mathrm{Ent}_{\pi^1} f^1 &= \sum_{x\in[n]\times\{\pm\}} \pi^1(x) f^1(x)\log f^1(x)\\ &\underset{f_1(x)=\frac{1}{2}\mathbb{E}_y f_2(x,y)}{=} \frac{1}{2}\mathbb{E}_{\{x,y\}\sim\pi^2} f^2(x,y)(\log f^1(x) + \log f^1(y))\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{Ent}_{\pi^2} f^2 - 2\,\mathrm{Ent}_{\pi^1} f^1 &= \mathbb{E}_{\{x,y\}\sim\pi^2} f^2(x,y)(\log f^2(x,y) - \log(f^1(x)\cdot f^1(y)))\\ &\geq \mathbb{E}_{\{x,y\}\sim\pi^2}(f^2(x,y) - f^1(x)f^1(y)) \qquad \text{(since for } a,b\geq 0 \text{ we have } a\log\tfrac{a}{b}\geq a-b)\\ &= 1 - \mathbb{E}_{(x,y)\sim\pi^2} f^1(x)f^1(y) \qquad \text{(since } \mathbb{E}f^2 = 1.)\end{aligned}$$

The next observation is that we can write $f^1 = 1 + g^1$ where $\langle g^1, 1\rangle_{\pi^1} = 0$. Recall that $M \in \mathbb{R}^{2n\times 2n}$ where $M((i,s_1),(j,s_2)) = \mathbb{P}\left[\sigma_j = s_2|\sigma_i = s_1\right]$ that we defined as few lectures ago. Then,

$$\begin{aligned}\mathrm{Ent}_{\pi^2} f^2 - 2\,\mathrm{Ent}_{\pi^1} f^1 &\geq 1 - \mathbb{E}_{(x,y)\sim\pi^2} f^1(x)f^1(y)\\ &= \mathbb{E}_{(x,y)\sim\pi^2} g^1(x)g^1(y) \qquad \text{(since } \mathbb{E}f^1 = 1)\\ &= \left\langle g^1, \frac{n}{n-1}(M-I)g^1\right\rangle_{\pi^1} \qquad \text{(we subtract } I \text{ since there is no term } g^1(x)^2)\\ &\leq -\lambda_2(\frac{n}{n-1}(M-I))\cdot\mathrm{Var}(f^1) \leq -\frac{\eta}{n-1}\mathrm{Var}(f^1)\end{aligned}$$

as desired. IN the last inequality we used that π is η-spectrally independent which implies that $\lambda_2(M) \leq \eta/n$. □

Finally to finish the proof of Lemma 11.7 we use the following fact which we leave as an exercise:

Fact 11.9. *Suppose π is B-marginally bounded. Then for any function $f : \{\pm 1\}^n \to \mathbb{R}_{\geq 0}$,*

$$f^1(i, s) \leq \frac{1}{B} \mathbb{E} f, \quad \forall i \in [n], s \in \{\pm 1\}.$$

As a consequence of this we have, $\frac{\mathrm{Var}(f^1)}{\mathbb{E} f^1} \leq \frac{4}{B^2} \mathrm{Ent}_{\pi^1} f^1$

The proof uses that the function $x \log x$ behaves quadratically in a neighborhood of 1.

References

[BT06] S. Bobkov and P. Tetali. "Modified Logarithmic Sobolev Inequalities in Discrete Settings". In: *Journal of Theoretical Probability* 19.2 (2006) (cit. on p. 11-2).