Counting and Sampling **Fall 2017**

Lecture 11: Lower Bounds on Mixing

Lecturer: Shayan Oveis Gharan *Nov 1st*

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

We will finish the discussion on Markov chain by introducing a technique to lower bound mixing. First, we prove a lower bound on mixing using the conductance. Note that when the chain is reversible, the inverse of the Poincaré constant also gives a lower bound on mixing time. The importance of the following lemma is that it is significantly easier to study. To lower bound the mixing time of the chain, all we need to do is to find a cut with small conductance.

Lemma 11.1. *For a (not necessarily reversible) kernel K let $Q(x,y) = \pi(x)K(x,y)$. Then, for any set $S \subseteq \Omega$ where $\pi(S) \leq 1/2$,*

$$\tau_{\text{mix}} \geq \frac{1}{4\phi(S)}.$$

Proof. Consider the following starting distribution:

$$\mu^0(x) = \begin{cases} \frac{\pi(x)}{\pi(S)} & \text{if } x \in S \\ 0 & \text{otherwise} \end{cases}$$

Also, let $\mu^t = \mu K^t$ be the distribution at time t. Then, observe that

$$\begin{aligned}
\left\|\mu^1 - \mu^0\right\|_{TV} &= \frac{1}{2}\sum_x |\mu^1(x) - \mu^0(x)| && \text{(Def of TV distance)} \\
&= \frac{1}{2}\sum_x |\sum_y \mu^0(y)K(y,x) - \mu^0(x)| \\
&= \sum_{x\notin S} |\sum_y \mu^0(y)K(y,x) - \mu^0(x)| && (\mu^1(x) \geq \mu^1(x') \text{ only when } x \notin S) \\
&= \sum_{x\notin S}\sum_y \mu^0(y)K(y,x) && (\mu^0(x) = 0 \text{ for } x \notin S) \\
&= \sum_{y\in S}\sum_{x\notin S} \mu^0(y)K(y,x) = \phi(S). && (11.1)
\end{aligned}$$

Now, recall that we used the coupling lemma to show the monotonicity of the total variation distance; so, for all t,

$$\left\|\mu^t - \mu^{t-1}\right\|_{TV} \geq \left\|\mu^{t+1} - \mu^t\right\|_{TV}.$$

Therefore, by triangle inequality,

$$\left\|\mu^t - \mu^0\right\|_{TV} \leq \sum_{i=0}^{t-1} \left\|\mu^{i+1} - \mu^i\right\|_{TV} \leq t\left\|\mu^1 - \mu^0\right\|_{TV}.$$

But,

$$\left\|\mu^t - \pi\right\|_{TV} \geq \left\|\mu^0 - \pi\right\| - \left\|\mu^t - \mu^0\right\|_{TV} \geq 1/2 - t\phi(S)$$

So, for $t < 1/4\phi(S)$, the total variation distance is at least $1/4$. □

The above lemma gives a useful technique, known as the *bottleneck ratio*, to bound the mixing time. Namely, all we need to do is to find a large set S of size $\pi(S) \geq \Omega(1)$, and we need to show that it is very hard for the random walk to enter S (or equivalently to leave S) because it has small conductance.

11.1 An Exponential Lower bound for Mixing time of Glauber dynamics in Ising Model

Consider the Glauber dynamics, a.k.a., Heat-Bath chain, that we studied in the last problem of HW1. There we showed that for a constant β_0, if $\beta < \beta_0$, then the Heat-Bath chain mixes rapidly. Here we show that there is another constant β_1 such that if $\beta > \beta_1$, then the mixing time is exponential in n. The more general theorem is as follows, but the proof is technical and we don't give all details.

Theorem 11.2 ([MO94])**.** *There is a constant β_c such that for any $\beta < \beta_c$, the heat-bath chain on the Ising model in a $\sqrt{n} \times \sqrt{n}$ grid mixes in time $O(n \log n)$ and for $\beta > \beta_c$ the mixing time is at least $e^{c\sqrt{n}}$ for some universal constant $c > 0$.*

The high-level intuition is clear: Clearly for large enough β, the all-plus state and the all-minus state would have the largest probabilities. But this Markov chain has exponentially many states. So, these two states do not contain all of the probability mass. Instead, the idea is to divide the states into those close to begin all-plus, say S_+, and those which are close to being all-minus, say S_-, and show that $\phi(S_+)$ is very small. See the following figure.

Figure 11.1: A bottleneck between S_+ and S_-.

A *path* is sequence of adjacent sites. A *line* is a sequence of line segments where each segment is a side of sub-square of $\sqrt{n} \times n$ square. The blue curve in Definition 11.1 is a line.

Definition 11.3 (Fault Line)**.** *For a configuration σ, a fault line is a line where for each segment the two sites on the opposite side of the segment have different spins, i.e., signs.*

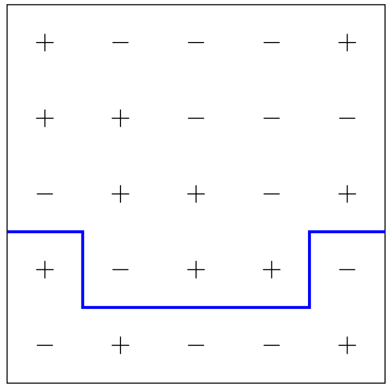
For example, the blue curve in Definition 11.1 is a fault line.

Lemma 11.4. *Let F be the set of configurations that contain a fault line. The for some universal constant β_1, and any $\beta > \beta_1$,*

$$\pi(F) \leq e^{-c\sqrt{n}}.$$

where c is a universal constant.

Proof. First observe that the number of fault lines of length $\ell \geq \sqrt{n}$ is at most $2\sqrt{n}3^\ell$. This is because there are at most $2\sqrt{n}$ starting position and each time there are at most 3 options (because the line is self-avoiding).



Now, fix a fault line L and let $F(L)$ be all configurations where L is a fault line. For a $\sigma \in F(L)$ flip the spin of all sites on one side of L. Observe that the weight of σ goes up by $e^{2\beta\ell}$. Moreover, this mapping is one-to-one. Therefore, $\pi(F(L)) \leq e^{-2\beta\ell}$. Now, brute forcing over all possible L we have

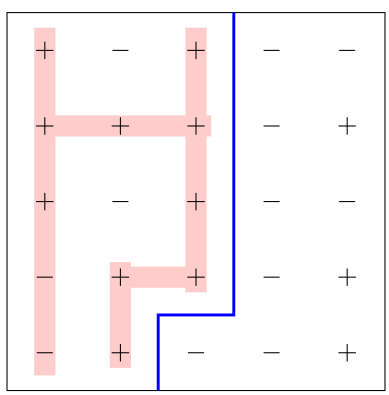
$$\pi(F) \leq \sum_L \pi(F(L)) \leq 2\sqrt{n} \sum_\ell 3^\ell e^{-2\beta\ell} \leq e^{-c\sqrt{n}},$$

for $\beta > \frac{1}{2}\ln 3$. □

In order to make sure that the fault lines that we consider have length at least $\sqrt{n}$ from now on we only focus on left-right and bottom-up fault lines. It follows by the above lemma that most of the probability mass in π is on states which do not have a fault line. In the following lemma we characterize these states.

Lemma 11.5. *Let σ be a configuration which has no monochromatic left-right path, i.e., no red/blue left-right path. Then, σ has a bottom-up fault line.*

Proof. Think of the following picture for intuition. Let A be the set of sites i where there is a path from



the left side to i by sites with spin the same as σ_i. For example, in the above figure the set A are all sites marked in red. The assumption of the lemma implies that A does not have any site in the right boundary. The definition of the set A implies that for any site i which is not in A but adjacent to a site of A, say j, we have $\sigma_i \neq \sigma_j$. It follows that the line at the right boundary of A is a bottom-up fault line. □

Lemma 11.6. *Suppose there is a site i such that there is a plus-path from i to the top and there is a minus-path from i to the top. Then there is a fault line from x to the top.*

We defer the proof of this lemma as an exercise.

Having the above lemmas, we are ready to prove the theorem.

Theorem 11.7. *There is a universal constant $\beta_1 > 0$ such that for all $\beta > \beta_1$, the mixing time of the heat-bath chain on a $\sqrt{n} \times \sqrt{n}$ grid is at least $e^{c\sqrt{n}}$.*

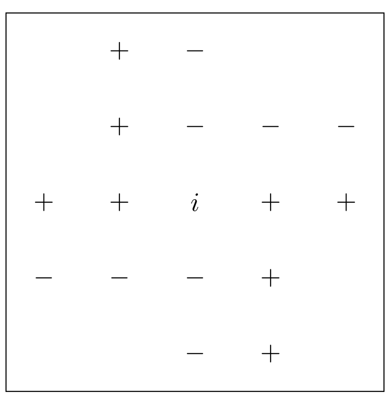
Proof. Let S_+ be all states with a left-right plus-path and a bottom-up plus-path. Similarly, let S_- be all states with a left-right minus-path and a bottom-up minus-path. Observe that by Lemma 11.5 any state that is not in S_+, S_- has a fault line. This is because say a state σ does not have a left-right plus path. If σ does not have a left-right minus-path as well then by Lemma 11.5 it has a bottom-top fault line. So, say σ has a left-right minus path. Then, it cannot have a bottom-up plus-path. Futhermore, since $\sigma \notin S_-$ it does not have bottom-up minus-path. So, σ must have a left-right fault line. Now, by Lemma 11.4 $\pi(\overline{S_+ \cup S_-}) \le e^{-c\sqrt{n}}$. Therefore, by symmetry, as $n \to \infty$, $\pi(S_+), \pi(S_-) \to 1/2$.

Now, let $N(S_+)$ be states which are not in S_+ such that they adjacent to some state in S_+. In other words, every state in $N(S_+)$ differ in the sign of one site with respect to a state in S_+. To prove the claim it is enough to show that $\pi(N(S_+)) \le e^{-c'\sqrt{n}}$. This is because by Lemma 11.1 we have

$$\tau_{\text{mix}} \ge \frac{1}{4\phi(S_+)} \ge \Omega(\pi(N(S_+))),$$

where we used that $\pi(S_+) \approx 1/2$.

It remains to upper bound $\pi(N(S_+))$. We divide states $\sigma \in N(S_+)$ in two groups: (i) States $\sigma \in N(S_+) \cap \overline{S_- \cup S_+}$. But, as we discussed earlier such a set has probability at most $e^{-c\sqrt{n}}$ because it has a fault line. (ii) States $\sigma \in N(S_+) \cap S_-$. Fix such a σ. Since σ is a neighbor of S_+ there is a site i such that by flipping the spin of i the new state σ' is in S_+. Observe that since $\sigma \in S_-$, there are left-right and bottom-up minus-paths in σ. Furthermore, since $\sigma' \in S_+$ there are left-right and bottom-up plus-paths in σ'. But σ, σ' differ in the sign of a single site. Therefore, site i must have plus and minus paths to top/bottom/left/right (see the following figure). Therefore, by Lemma 11.6 there are fault lines from i to top and bottom. So, there is an "almost" fault line from bottom to top in σ. Such a line may have agreeing signs in the neighborhood of i. But that is a constant length part of the path. We can adopt the proof of Lemma 11.4 to argue that probability of states with an almost fault line is also $e^{-c'\sqrt{n}}$. Therefore, $\pi(N(S_+)) \le e^{-2c'\sqrt{n}}$.



□

References

[MO94] F. Martinelli and E. Olivieri. Approach to equilibrium of glauber dynamics in the one phase region. ii. the general case. *Comm. Math. Phys.*, 161(3):487–514, 1994. 11-2