MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.265/15.070J Lecture 20 May 1, SP17
Lecturer: Guy Bresler Scribe notes by Jason Altschuler

Mixing Times for Glauber Dynamics on the Ising Model

Content.

1. Review of definitions
2. Fast mixing at high temperature
3. 2D Ising model

1 Review

Let us first review the definition of an Ising model and its associated Glauber dynamics. (For further reading, see Section 3 of the textbook [LPW].)

Definition 1. *The **Ising model** on a graph $G = (V, E)$ with temperature β is the probability distribution on $\{-1, 1\}^V$ defined by:*

$$\pi(x) = \frac{1}{Z(\beta)} \exp\left(\beta \sum_{(i,j)\in E} x_i x_j\right) \tag{1}$$

We often denote $w(x) := \exp\left(\beta \sum_{(i,j)\in E} x_i x_j\right)$. The normalization constant $Z(\beta)$ is then equal to $\sum_{x\in\{-1,1\}^V} w(x)$.

Remark.

$$\sum_{(i,j)\in E} x_i x_j = (\textit{\# agreeing neighbors}) - (\textit{\# disagreeing neighbors})$$

Definition 2. *The **Glauber dynamics** for the above Ising model is the Markov chain that evolves by selecting a vector $u \in V$ at random and updating the spin at u according to the distribution π conditioned to agree with the spins at all* *vertices not equal to u. That is, if the current configuration is x and vertex u is selected, then the probability that the spin at u is updated to $+1$ is equal to*

$$P(x,u) = \frac{\exp\left(\beta \sum_{v \in N(u)} x_v\right)}{\exp\left(\beta \sum_{v \in N(u)} x_v\right) + \exp\left(-\beta \sum_{v \in N(u)} x_v\right)}$$

2 Fast mixing of Glauber dynamics of Ising models at high temperatures

(For further reading, see Section 15.1 of the textbook [LPW].)

The main result of this section is the following.

Theorem 1. *Consider Glauber dynamics for the Ising model on a graph with n vertices and max degree Δ. Define $c(\beta) := 1 - \Delta \tanh(\beta)$. If $\Delta \tanh \beta < 1$ then:*

$$t_{mix}(\varepsilon) \leq \left\lceil \frac{n(\log n + \log \frac{1}{\varepsilon})}{c(\beta)} \right\rceil$$

Remark. *$\Delta \tanh \beta < 1$ whenever $\beta < \frac{1}{\Delta}$. The latter condition is typically easier to check.*

Remark. *Theorem 1 is tight e.g. for d-regular graphs.*

Before we prove Theorem 1, let us first recall the method of path-coupling, which will be our main proof ingredient. For us, the path metric $\rho(\cdot,\cdot)$ can be thought of as Hamming distance.

Theorem 2. *Suppose for all pairs of neighboring states $x, y \in \Omega$, there exists a coupling (X_1, Y_1) where $X_1 \sim P(x,\cdot)$, $Y_1 \sim P(y,\cdot)$, and $\mathbb{E}[\rho(X_1, Y_1)] \leq e^{-\alpha}\rho(x,y)$. Then:*

$$t_{mix}(\varepsilon) = O\left(\alpha^{-1} \log diam(\Omega)\right)$$

We are now ready to prove Theorem 1.

Proof of Theorem 1. To use the path-coupling theorem, we will need to, for each pair of neighboring configurations, create a coupling that contracts in expectation w.r.t. our path metric. It will be convenient to use Hamming distance as our path metric $\rho(\cdot,\cdot)$. Let $\sigma, \tau \in \{-1,1\}^V$ be neighbors (equal except one vertex); that is, $\sigma(u) = \tau(u)$ for all $u \in V \setminus \{v\}$. Then define the following coupling:

- Draw $U \sim \text{unif}[0,1]$
- Pick $w \in V$ u.a.r.

- Update both $\sigma(w) = \begin{cases} +1 & U \leq P(\sigma, w) \\ -1 & U > P(\sigma, w) \end{cases}$ and $\tau(w) = \begin{cases} +1 & U \leq P(\tau, w) \\ -1 & U > P(\tau, w) \end{cases}$

Denote by σ' and τ' the updated configurations after one step of this coupling. We wish to upper bound $\mathbb{E}[\rho(\sigma', \tau')]$. Let us evaluate this by conditioning on which $w \in V$ was chosen. There are three cases.

- Case 1: $w \in N(v) \cup \{v\}$. Then $\rho(\sigma', \tau') = 1$.
- Case 2: $w = v$. Then $\rho(\sigma', \tau') = 0$.
- Case 3: $w \in N(v)$. WLOG let $\tau(v) = +1$ and $\sigma(v) = -1$. Then $\rho(\sigma', \tau') = 2$ occurs w.p. $P(\tau, w) - P(\sigma, w)$ and is otherwise equal to 1.

In conclusion, we have:

$$\mathbb{E}[\rho(\sigma', \tau')] = 1 - \frac{1}{n} + \frac{1}{n} \sum_{w \in N(v)} \left(P(\tau, w) - P(\sigma, w)\right)$$

Let us analyze the final term. Let $S := \sum_{u \in N(w)} \sigma(u)$. Then $S + 2 = \sum_{u \in N(w)} \tau(u)$, and so by definition of Glauber updates:

$$P(\tau, w) - P(\sigma, w) = \frac{e^{\beta(S+2)}}{e^{\beta(S+2)} + e^{-\beta(S+2)}} - \frac{e^{\beta S}}{e^{\beta S} + e^{-\beta S}} = \frac{1}{2}\left(\tanh\left(\beta(S+2)\right) - \tanh\left(\beta S\right)\right)$$

which can be bounded above by $\tanh(\beta)$ (details omitted in lecture; see instead e.g. Section 15.1 of the textbook [LPW]). Therefore we conclude:

$$\mathbb{E}[\rho(\sigma', \tau')] \leq 1 - \frac{1}{n}\left(1 - \Delta \tanh(\beta)\right) = 1 - \frac{c(\beta)}{n} \leq e^{-c(\beta)/n}$$

The desired claim then follows directly from an application of Theorem 2 with $\alpha := c(\beta)/n$ and $\mathrm{diam}(\{-1, +1\}^V) = n$. $\square$

3 2D Ising Models

(For further reading, see Section 15.6 of the textbook [LPW].)

The setup. The graph will be the grid of $n \times n$ squares $G = \mathbb{L}^2 \cap [0, n]^2$. Formally, the graph is defined with $V = \{(i, j) \ : \ 0 \leq i, j \leq n - 1\}$ and edges connecting vertices at unit Euclidean distance.

In what follows, we will study the associated Glauber dynamics for the Ising model on V. Mixing times of these Glauber dynamics undergo the following phase transition:

Theorem 3. *Let* $\beta_c := \frac{1}{2}\log\left(1+\sqrt{2}\right)$. *Then:*

$$t_{mix} = \begin{cases} O(n^2 \log n) & \beta < \beta_c \\ e^{cn} & \beta > \beta_c \end{cases}$$

We will not prove Theorem 3, since getting all the way down to the critical threshold β_c requires significant work. Instead we will prove the following much weaker result.

Theorem 4. *For large enough constant* $\beta > 0$*, mixing time is exponentially slow:*

$$t_{mix} = e^{\Omega(n)}$$

The path to proving Theorem 4 will require developing some definitions and small lemmas, which we do presently. (Note that in what follows we distinguish between lattice paths (paths on edges) and square paths (paths on the squares of the grid).)

Definition 3. *A* ***fault line*** *is a self-avoiding lattice path from* LEFT *to* RIGHT *or* TOP *to* BOTTOM *of* $[0, n]^2$*, such that for each of the edges of the path, the two squares adjacent to it have different spins. A* ***fault line with at most*** k ***defects*** *is defined similarly with all but at most* k *edges satisfying the property of: the two squares adjacent to it have different spins.*

Further, denote by $F := \{x \in \{-1,1\}^V \ : \ \exists$ LEFT-RIGHT *fault line*$\}$ *the set of all Ising configurations that contain at least one* LEFT $-$ RIGHT *fault line. Define* F_k *analogously for configurations that contain at least one* LEFT$-$RIGHT *fault line with at most* k *defects.*

Lemma 5. *For* $\beta > \frac{1}{2}\log 3$ *and* $k \leq 3$*:*

$$\pi(F) \leq e^{-cn}$$
$$\pi(F_k) \leq e^{-c(k)n}$$

where $c(k)$ *is a constant depending only on* k*.*

Proof. Fix a fault line L of length ℓ from LEFT to RIGHT (or from TOP to BOTTOM). Define $F(L) := \{x \in \{-1,1\}^V \ : \ L \text{ is a fault line}\}$. Consider the operation $x \mapsto x'$ that flips all signs of the configuration above L. Then:

$$\pi(F(L)) = \sum_{x \in F(L)} \frac{w(x)}{Z} = e^{-2\beta\ell} \sum_{x \in F(L)} \frac{w(x')}{Z} \leq e^{-2\beta\ell}$$

Now summing over all self-avoiding paths in F, we obtain:

$$\pi(F) \leq 2n \sum_{\ell \geq n} 3^{\ell} e^{-2\beta\ell}$$

where the 2 is because it could go from TOP$-$BOTTOM or LEFT$-$RIGHT, and the 3^{ℓ} is because it is self-avoiding. This bound is of the desired form $\leq e^{-cn}$ since we assumed $\beta > \frac{1}{2}\ln 3$. The same argument applies to $\pi(F_k)$ with only a few changes (e.g. the flipping operation now amplifies the probability by $e^{2\beta(\ell-2k)}$ instead of $e^{2\beta\ell}$). □

Lemma 6. *Fix some Ising configuration $x \in \{-1, 1\}^V$.*

- *If there is neither an all positive nor an all negative* LEFT $-$ RIGHT *square path, then there exists a* BOTTOM $-$ TOP *fault line.*
- *If $\exists$ a square $v \in V$ with both a positive path and a negative path of squares to* TOP*, then there exists a "fault line" from v to* TOP*.*

Proof. For the first part, the key observation is just to define the set A of squares reachable by a monochromatic path from left, and then to go along the boundary of A from top to bottom. The remaining details are simple and were ommitted in lecture.

The second part is slightly more nuanced. Denote by Γ_+ (resp. Γ_-) a positive (resp. negative) path of squares from v to the top. Suppose WLOG that the square that Γ_+ terminates on (in the TOP) is to the left of the square that Γ_- terminates on. Define A_+ to be the set of squares reachable from T_+ with all positive squares, and let A_+^* be the set of negative squares separated from the boundary of $[0, n]^2$ by A_+. We can then create a "fault line" from v to the top inductively. Start with a lattice edge that has v on the right, and an element of Γ_+ on the left. Inductively move to an edge that has on its left a positive square of A_+ and a negative square not in A_+^* on its right. It is simple to see that we can always find such a next edge, and that the path will never contain a cycle. □

We are now ready to conclude the proof of the main result of this section.

Proof of Theorem 4. Define S_+ (resp. S_-) to be the set of configurations with positive (resp. negative) paths both from LEFT $-$ RIGHT and TOP $-$ BOTTOM. Consider an element $x \in (S_+ \cup S_-)^C$. Then either x contains no monochromatic path from LEFT $-$ RIGHT (in which case the first part of Lemma 6 implies there is a TOP $-$ BOTTOM fault line) or from TOP $-$ BOTTOM (in which case there is a LEFT $-$ RIGHT fault line). Therefore every configuration of $(S_+ \cup S_-)^C$

contains a fault line, thus $\pi((S_+ \cup S_+)^C) \leq e^{-cn}$ by Lemma 5, which implies that $\pi(S_+) \geq \frac{1}{2} - \frac{1}{2}e^{-cn}$.

Now, define the "external boundary" $\partial S_+ := \{x \in \{-1,1\}^V \quad : \quad x \notin S_+, d_H(S_+, x) = 1\}$; By the bottleneck-ratio method:

$$\mathrm{t}_{\mathrm{mix}} \geq \frac{1}{4}\frac{\pi(S_+)}{C(S_+, S_+^C)} \geq \frac{1}{4}\frac{\pi(S_+)}{\pi(\partial S_+)}$$

Therefore since we already showed that $\pi(S_+)$ is bounded below by roughly $\frac{1}{2}$, it suffices to show that $\pi(\partial S_+)$ decays exponentially in n. By Lemma 5, it suffices to show that every configuration $x \in \partial S_+$ contains a fault line with at most 3 defects. We do this is as follows.

First, the case when $x \notin S_-$. Then by a nearly identical argument to what we did in the first part of this proof, we have by the first part of Lemma 6 that x contains a fault line. Next, the case where $x \in S_-$, i.e. $x \in \partial S_+ \cap S_-$. Then by definition there exists some square v in the grid such that our configuration would be in S_+ if we flipped $x(v)$. Therefore by the second part of Lemma 6 that there exists a fault line from (e.g.) LEFT to RIGHT that contains at most 3 edges of q which might be defects. This completes the proof. $\square$