Lecture 20: November 12

Lecturer: Prof. Alistair Sinclair *Scribes: S. Negahban and D. Omidiran*

20.1 Glauber Dynamics for the 2-d Ising model

Recall the fundamental result quoted last time (see, e.g., [M98]):

Theorem 20.1 *The mixing time of the Glauber dynamics for the Ising model on a $\sqrt{n} \times \sqrt{n}$ box in the 2-dimensional square lattice is:*

$$\begin{cases} O(n \log n) & \text{if } \beta < \beta_c; \\ e^{\Omega(\sqrt{n})} & \text{if } \beta > \beta_c, \end{cases}$$

where β_c is the critical (inverse) temperature, $\beta_c = \frac{1}{2} \ln (1 + \sqrt{2})$.

Last time we showed that the mixing time is $e^{\Omega(\sqrt{n})}$ for sufficiently large (but finite) β (though not all the way down to β_c). Now we will show that the mixing time is $O(n \log n)$ for sufficiently small (but finite) β (again, not all the way up to β_c). Getting both of these results to go all the way to β_c is rather challenging and beyond the scope of this course.

We also mention the following interesting conjecture:

Conjecture: For the Ising model with the $+$ (or $-$) boundary condition, the mixing time is poly(n) for *all* $\beta > 0$. [Intuition: The obstacle to rapid mixing for $\beta > \beta_c$ is the bottleneck between the plus-phase and the minus-phase; but one of the phases disappears with such a boundary condition.]

We now turn to the result claimed above, that the mixing time is $O(n \log n)$ for sufficiently small β.

Let $d(X, Y)$ denote the number of disagreements between the configurations X, Y. We use path coupling, meaning that we need consider only pairs X, Y that have one disagreement.

Consider X_t, Y_t with disagreement only at i_0. Define the coupling in which X_t and Y_t always pick the same site i and update it "optimally" (i.e., so as to maximize the probability of agreement).

- Good moves: $i = i_0 \to \Delta d = -1$ with probability 1.
- Bad moves: $i \sim i_0 \to \Delta d = +1$ with probability at most $\frac{\exp(2\beta) - \exp(-2\beta)}{2 + \exp(2\beta) - \exp(-2\beta)}$, where $i \sim i_0$ means that i is a neighbor of i_0.

To see this probability of a bad move, let $\alpha = \alpha^+ - \alpha^-$ where α^+ denotes the number of $+$ neighbors of i excluding i_0, and α^- denotes the number of $-$ neighbors excluding i_0.

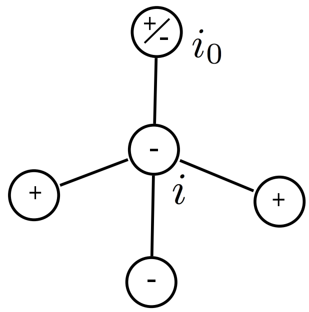

Figure 20.1: Here, X_t, Y_t differ only at node i_0 and are equal everywhere else. Node i is an arbitrary neighbor of i_0. In this example, $\alpha^+ = 2$ and $\alpha^- = 1$.

Letting $\Pr(X_t(i) = +)$ denote the probability that node i is set to $+$ conditioned on its neighbors, then the probability of creating an additional disagreement is $|\Pr(X_t(i) = +) - \Pr(Y_t(i) = +)|$. A simple calculation shows that

$$|\Pr(X_t(i) = +) - \Pr(Y_t(i) = +)| \;=\; \frac{\exp(2\beta)\; - \exp(-2\beta)}{\exp(2\beta\alpha) + \exp(-2\beta) + \exp(2\beta) + \exp(-2\beta\alpha)},$$

which is maximized by letting $\alpha = 0$. This gives the expression claimed above.

Since all other moves leave $d(X_t, Y_t)$ unchanged, the expected change in distance in one move (bearing in mind that each node i_0 has four neighbors), is at most

$$\mathrm{E}[\Delta d] \le \frac{1}{n}\Big[-1 + 4\,\frac{\exp(2\,\beta) - \exp(-2\,\beta)}{2 + \exp(2\,\beta) - \exp(-2\,\beta)}\Big].$$

Therefore, by our usual path coupling analysis, since the maximum distance is at most n, we will get $O(n \log n)$ mixing time if

$$4\,\frac{\exp(2\,\beta) - \exp(-2\,\beta)}{2 + \exp(2\,\beta) - \exp(-2\,\beta)} \;<\; 1.$$

Setting $z = \exp(2\,\beta)$ this becomes $4z - 4z^{-1} \;<\; 2 + z + z^{-1}$. Multiplying out by z and factoring yields $(3z - 5)(z + 1) < 0$, so that we have $O(n \log n)$ mixing time for $\beta < \frac{1}{2}\ln(5/3)$. (Note that this is rather smaller than the critical value $\beta_c = \frac{1}{2}\ln(1 + \sqrt{2})$.)

20.2 Spatial and Temporal Mixing

In this section we develop a relationship between "Spatial Mixing" and "Temporal Mixing." That is, roughly speaking, we will prove that

"Spatial Mixing" (decay of correlations) $\Longleftrightarrow$ "Temporal Mixing" (fast mixing time)

For simplicity we will focus on the 2-d Ising model in a $\sqrt{n} \times \sqrt{n}$ box, but our results actually hold for much more general spin systems on more general graphs. We will demonstrate each of the above two implications separately, beginning with the $\Leftarrow$ direction. Let π^+, π^- denote the Gibbs measures (stationary distributions) for the all-plus and all-minus boundary conditions, respectively. Let $\pi^+(\sigma_0 = +)$ denote the probability under π^+ that the origin has spin $+$.

Theorem 20.2 *Suppose, $\tau_{mix} = O(n \log n)$ for all boundary conditions. Then*

$$|\pi^+[\sigma_0 = +] - \pi^-[\sigma_0 = +]| \to 0 \text{ as } n \to \infty.$$

Exercise: With a bit more care in the argument, one can show exponential decay of correlations, i.e., $|\pi^+[\sigma_0 = +] - \pi^-[\sigma_0 = +]| \leq \exp(-c\sqrt{n})$.

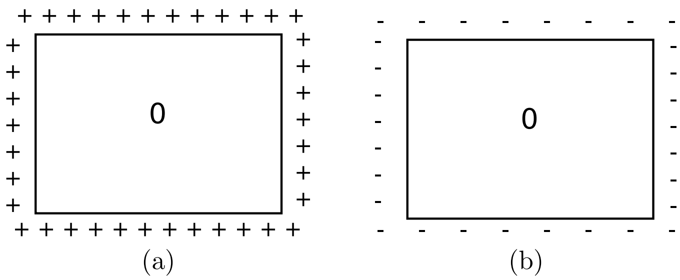

Figure 20.2: (a) Markov chain (X_t) with boundary conditions set to all plus. (b) Markov chain (Y_t) with boundary conditions set to all minus. The initial configurations X_0, Y_0 are identical except on the boundary. The stationary distributions of $(X_t), (Y_t)$ are π^+, π^- respectively.

Proof: Set the initial condition $X_0 = Y_0$ (except on the boundary). Couple (X_t) and (Y_t) such that:

- They choose the same site i to be updated at all times t.
- If the neighborhoods of the i are equal, then they both perform the same update, otherwise they update independently conditional on their neighbors.

We then have

$$\left|\pi^+[\sigma_0 = +] - \pi^-[\sigma_0 = +]\right| \leq \underbrace{\left|\pi^+[\sigma_0 = +] - X_t[\sigma_0 = +]\right|}_{a} + \underbrace{\left|X_t[\sigma_0 = +] - Y_t[\sigma_0 = +]\right|}_{b} + \underbrace{\left|Y_t[\sigma_0 = +] - \pi^-[\sigma_0 = +]\right|}_{c}$$

Choose $t = Cn\log^2 n > \tau_{mix}\log n$. Then, $a \leq \frac{1}{n} \to 0$ and $c \leq \frac{1}{n} \to 0$, by standard properties of the mixing time. To analyze the remaining term b, note that

$$b \leq \Pr[\exists \text{ disagreement at 0 within } t \text{ steps }] = \Pr[\exists \text{ path of disagreement from boundary to 0}].$$

Here, a "path of disagreement" means a contiguous sequence of sites, starting at the boundary and leading to the origin 0, that are chosen to be updated in sequence. Note that only if this happens is it possible for the spins at the origin in X_t, Y_t to differ, since they start out identical except on the boundary.

Now, we may bound the latter term by

$$\Pr[\exists \text{ path of disagreement from boundary to 0}] \leq \sum_{k \geq \frac{1}{2}\sqrt{n}} (4\sqrt{n}\,4^k)\binom{t}{k}\left(\frac{1}{n}\right)^k,$$

where $(4\sqrt{n}\,4^k)$ is the number of paths, $\binom{t}{k}$ counts the number of ways of choosing the sequence of update times along the path, and $(\frac{1}{n})^k$ is the probability that this sequence of updates is actually chosen. The above quantity is then bounded by

$$4\sqrt{n}\sum_{k\geq\frac{1}{2}\sqrt{n}}\left(\frac{4et}{kn}\right)^k \;\leq\; \underbrace{4\sqrt{n}}_{\text{"noise"}}\sum_{k\geq\frac{1}{2}\sqrt{n}}\left(\frac{4ec\log^2 n}{k}\right)^k \to 0 \text{ as } n\to\infty,$$

■

We turn now to the ⇒ direction of the above equivalence: i.e., we will show that exponential decay of correlations implies $O(n\log n)$ mixing time for all boundary conditions. We will actually prove the mixing time bound not for the standard Glauber (heat-bath) dynamics, but for the *block* version of this dynamics, which is the same except that, instead of choosing a random site and updating it, we choose a random $L\times L$ block of sites and update the entire block (again, conditional on the spins on the neighbors of the block). By a result of Peres and Winkler [PW05], this implies a similar mixing time for the single-site version of the dynamics, at least for monotone systems such as the Ising model.

Theorem 20.3 *Suppose we have exponential decay of correlations, i.e., for all pairs of sites i,j,*

$$\left|\pi^{i=+}[\sigma_j=+]-\pi^{i=-}[\sigma_j=+]\right| \;\leq\; \exp\big(-\alpha d(i,j)\big),$$

where $d(i,j)$ is the distance between i and j and $\alpha>0$ is a constant. Then the mixing time of the $L\times L$-block dynamics (for sufficiently large L) is $O(n\log n)$ for all boundary conditions.

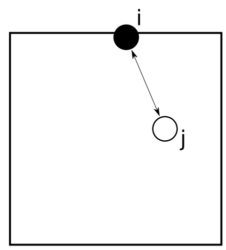

Figure 20.3: The effect of changing the spin at i on the distribution of the spin at j decays exponentially with the distance $d(i,j)$.

Proof: We use path coupling. Consider two copies $(X_t),(Y_t)$ of the Markov Chain with same boundary conditions and arbitrary initial states X_0 and Y_0. Consider the case where X_t and Y_t differ at one site i.

- Good moves: We select a point in the $L\times L$ block centered around i. This implies that $\Delta d=-1$ with probability 1. There are L^2 of these choices.
- Bad moves: We select a point j such the node i is on the edge of the $L\times L$ box centered around the point j. In such a case, $\mathrm{E}[\Delta d]=c'L+L^2\exp(-\alpha c\sqrt{L})\leq c''L$ for arbitrarily small $c''>0$ by taking L large enough. (The first term here comes from the sites in the box that are at distance at most $c\sqrt{L}$ from i; clearly there are at most cL of these, for some constant c. The second term comes from the remaining sites in the box, for which we apply the decay of correlations in the hypothesis. By choosing c appropriately small, and then L appropriately large, we can make c'' as small as we like.) There are $4L$ of these choices, corresponding to the size of the boundary of an $L\times L$ block.

  Since all other moves do not change the distance d, we have

$$E[\Delta d]\leq\frac{1}{n}[-L^2+4c''L^2]\leq-\frac{C}{n},$$

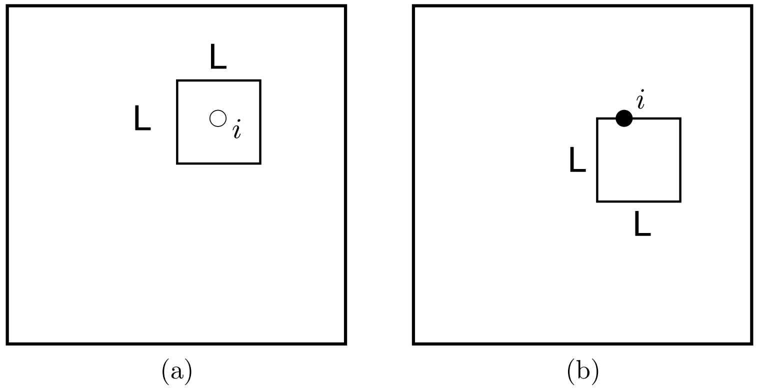

Figure 20.4: (a) Good moves. (b) Bad moves

provided we make c'' small enough. Hence, since the maximum distance is n, by our usual path coupling analysis we get that $\tau_{mix} = O(n \log n)$. ■

References

[M98] F. Martinelli. Lectures on Glauber dynamics for discrete spin models. *Lectures on Probability Theory and Statistics*, Springer Lecture Notes in Mathematics **1717**, 1998, pp. 93–191.

[PW05] Y. Peres, *Mixing for Markov chains and spin systems.* Lecture Notes for PIMS Summer School at UBC, August 2005. www.stat.berkeley.edu/∼peres/ubc.pdf