Glauber Dynamics on Trees

Manjunath Krishnapur and Gábor Pete

1 Introduction

This paper deals with Glauber dynamics for the Ising model on trees, with [5] as the primary reference. As explained below, Ising model on trees has an interpretation in terms of information flow along the edges, which makes it trivial to sample from the Gibbs distribution. Glauber dynamics is studied, not for sampling purposes, but because similar results are expected to be valid on other non-amenable graphs where there is no alternative interpretation. Indeed, some results generalize (viz. polynomial mixing at all temperatures), whereas some are not known yet (viz. existence of an intermediate phase in nonamenable transitive graphs)

1.1 Ising Model

For $\sigma \in \{+,-\}^V$ define

$$\pi(\sigma) = Z(\beta)^{-1} \exp\Big(\beta \sum_{\{v,w\}\in E_r} \sigma_v \sigma_w\Big)$$

An alternative description of Ising model on trees is as follows. Put +/- u.a.r at the root of the tree. Then send it down the edges of the tree with a probability of error ϵ independently for each edge. We get a configuration in $\{+,-\}^V$. With $\epsilon = (1+e^{2\beta})^{-1}$ we get the Ising model.

In this picture, the problem is to reconstruct the σ_ρ, knowing the spins on ∂T_r (T_r is the subtree of first r generations of T). We say that reconstruction is possible if the probability of correct reconstruction remains bounded away from 1/2 as $r \to \infty$.

2 Phase Transitions

Henceforth we shall assume that T is a b-ary tree. All results stated here are valid for general trees with bounded degree, with b replaced by the branching number of the tree.

In contrast to the situation in $\mathbf{Z}^d$ [6], there are three distinct phases for the Ising model on trees. Also in contrast to $\mathbf{Z}^d$ (where the two sides of phase transition correspond to polynomial v/s exponential mixing time), Glauber dynamics mixes in polynomial time at all temperatures.

High Temperature regime ($1-2\epsilon < \frac{1}{b}$)

Relaxation time is linear; there is a unique Gibbs measure on the infinite tree; the expected value of σ_ρ, given any boundary conditions on ∂T_r, decays exponentially as $r \to \infty$.

Intermediate Temperature regime ($\frac{1}{b} < 1-2\epsilon < \frac{1}{\sqrt{b}}$)

Relaxation time is still linear; there are infinitely many Gibbs measures (three extremal ones) on the infinite tree: the boundary conditions $\sigma(\partial T_r) = +$ and $\sigma(\partial T_r) = -$ induce bias in σ_ρ, so we get two different Gibbs measures in the limit, but for a typical boundary condition on ∂T_r the expected value of σ_ρ decays, which implies that the free Gibbs measure is also extremal [1].

Low Temperature regime ($\frac{1}{\sqrt{b}} < 1-2\epsilon$)

Relaxation time is superlinear; there are infinitely many Gibbs distributions on the infinite tree(two extremal ones) and the expected value of σ_ρ remains bounded away from zero for typical boundary conditions.

3 Results

Definition For a finite graph $G=(V,E)$, the exposure $\mathcal{E}(G)$ is the smallest integer such that there are atmost $\mathcal{E}(G)$ edges from $\{v_1,\ldots,v_k\}$ to $\{v_{k+1},\ldots,v_n\}$, for some ordering of vertices $v_1,\ldots,v_n$.

Proposition 3.1 *Consider the Ising model on a finite graph G with n vertices and maximal degree Δ. Then the relaxation time of the Glauber dynamics is at most $n^2 e^{(4\mathcal{E}(G)+2\Delta)\beta}$.*

Theorem 3.2 *Consider the Ising model on the b-ary tree T_r of height r. Let $\epsilon=(1+e^{2\beta})^{-1}$. The relaxation time τ_2 for Glauber dynamics on T_r can be bounded as follows:*

1. *The relaxation time is polynomial at all temperatures: $\tau_2=n_r^{O(\log(1/\epsilon))}$.*
2. **Low temperature regime**.
   (a) *If $1-2\epsilon\geq 1/\sqrt{b}$ then the relaxation time is superlinear: $\tau_2=\Omega(n_r^{1+\log_b(b(1-2\epsilon)^2)})$.*
   (b) *Moreover, the degree of τ_2 tends to infinity as ϵ tends to zero: $\tau_2=n_r^{\Omega(\log(1/\epsilon))}$.*
3. **Intermediate and high temperature regimes**.
   If $1-2\epsilon<1/\sqrt{b}$ then the relaxation time is linear: $\tau_2=O(n_r)$.

Theorem 3.3 *If G has bounded degree and the relaxation time of the Glauber dynamics satisfies $\tau_2(G_r)=O(n_r)$, then the Gibbs distribution on G_r has the following property. For any fixed finite set of vertices A, there exists $c_A>0$ such that for r large enough*

$$\mathrm{Cov}(f,g)\leq e^{-c_A r}\sqrt{\mathrm{Var}(f)\mathrm{Var}(g)}\,,$$

provided that $f(\sigma)$ depends only on σ_A and $g(\sigma)$ depends only on σ_r. Hence, reconstruction is not possible.

4 Sketch of proofs

Proof of Proposition 3.1 The proof is by constructing a good flow f. Take the ordering $v_1,\ldots,v_n$ which gives the exposure. Given two configurations σ and τ, send an amount $\pi(\sigma)\pi(\tau)$ from σ to τ via the path obtained by altering the spins of the disagreeing vertices in the above order. It is easy to bound $\rho(f)$ above via the obvious bijection between the set of flow-carrying paths containing a particular edge and the set of all configurations. ∎

Proof of Theorem 3.2 Part 1 is a direct consequence of Proposition 3.1.
Part 2a and Part 2b: we have the variational principle

$$\tau_2=\sup\left\{\frac{2\sum_\sigma \pi[\sigma](g(\sigma))^2}{\sum_{\tau\neq\sigma}\pi[\sigma]\mathbf{P}[\sigma\to\tau](g(\sigma)-g(\tau))^2}\ :\ \pi(g)=0,\ g\neq\mathrm{const}\right\},$$

For Part 2a close to $\frac{1}{\sqrt{b}}$ we use "Global Majority":

$$g(\sigma)=\sum_{v\in\partial T_r}\sigma_v$$

For Part 2b we use "Recursive Majority": starting from ∂T_r, inductively define m_v to be the sum of the spins of v's children. Let $g(\sigma)=m_\rho$ (There is some subtlety in defining this function when b is even.)

These functions give the lower bounds claimed.

Part 3: First we show fast mixing (linear relaxation time) for the following **block dynamics**:

We view our tree T_r as part of a larger b-ary tree T_* of height $r+2h$, where the root ρ of T_r is at level h in T_*. For each vertex v of T_*, consider the subtree of height h rooted at v. A **block** is by definition the intersection of T_r with such a subtree. In this way, each vertex of the tree T_r is inside exactly h blocks. At each step of the block dynamics, we pick a block at random, erase all the spins belonging to the block, and put new spins in, according to the Gibbs distribution conditional on the spins in the rest of T.

Let us define a weighted Hamming metric on configurations,

$$d(\sigma, \eta) = \sum_v \lambda^{|v|} 1(\sigma_v \neq \eta_v).$$

Let $\theta = 1 - 2\epsilon$ and $\lambda = 1/\sqrt{b}$. Note that $b\lambda\theta < 1$ and $\theta < \lambda$.

We will use **path coupling**: for every pair of configurations σ and η which differ by a single spin, $\sigma_v \neq \eta_v$, $d(\sigma, \eta) = \lambda^{|v|}$, we want to design a coupled updating $(\sigma, \eta) \mapsto (\sigma', \eta')$ such that the expected distance reduces by a factor uniformly smaller than 1. Then the general path coupling argument [2] gives a good coupling for each pair, and fast mixing will follow.

In the coupling we will always choose the same block B. There are four situations to consider.

Case 1. If B contains neither v nor any vertex adjacent to v, then $d(\sigma', \eta') = d(\sigma, \eta)$.

Case 2. If B contains v, then $\sigma' = \eta'$ and $d(\sigma', \eta') = 0 = d(\sigma, \eta) - \lambda^{|v|}$. There are h such blocks, corresponding to the h ancestors of v at $1, 2, \dots, h$ generations above v.

Case 3. If B is rooted at one of v's children, then the boundary of B in σ and η differ, but only at one vertex, v. Lemma 4.1 below shows that forgetting about the agreeing spins in the boundary condition can only increase the expectation of $d(\sigma', \eta') - d(\sigma, \eta)$, so we may assume a free boundary. But then we can use the broadcasting interpretation of the free Gibbs measure [4]: along an edge downwards from v, keep the sign faithfully with probability θ, and update to a uniform $\pm$ sign with probability $1 - \theta$. These updates can be coupled together, hence we will have different spins in σ' and η' exactly along the θ-percolation connected cluster of v. Thus the expectation of $d(\sigma', \eta') - d(\sigma, \eta)$ is at most $\sum_j \lambda^{|v|+j} b^j \theta^j \leq \lambda^{|v|}/(1 - b\lambda\theta)$. There are b such blocks, corresponding to the b children of v.

Case 4. If B is rooted at v's ancestor exactly $h+1$ generations above v, then there is a differing spin at the lower boundary of B. Again by Lemma 4.1, the expected distance between σ' and η' is dominated by the size of the θ-cluster of v. The expected weight of v's cluster is bounded by summing over the ancestors w of v: $\sum_w \theta^{|v|-|w|} \sum_j \lambda^{|w|+j} b^j \theta^j = \lambda^{|v|}/\big((1 - \theta\lambda^{-1})(1 - b\lambda\theta)\big)$.

Over all four cases, the expected change in distance is

$$\mathbf{E}(d(\sigma', \eta') - d(\sigma, \eta)) \leq \left(\frac{b\lambda^{|v|}}{1 - b\lambda\theta} + \frac{\lambda^{|v|}}{(1 - \theta\lambda^{-1})(1 - b\lambda\theta)} - h\lambda^{|v|} \right) \frac{1}{n + h - 1}.$$

If the block height h is a sufficiently large constant, we get that for some positive constant c,

$$\mathbf{E}(d(\sigma', \eta') - d(\sigma, \eta)) \leq \frac{-c\lambda^{|v|}}{n} \leq \frac{-c}{n} d(\sigma, \eta).$$

This implies a mixing time of at most $O(n \log n)$ and a relaxation time of at most $O(n)$ for the block dynamics — we don't go into these standard arguments here. See [2] and [3].

Since each block update can be simulated by doing a constant number of single-site updates inside the block, and each tree vertex belongs only to a bounded number of blocks, it follows from proposition 3.4 of [6] that the relaxation time of the single-site Glauber dynamics is also $O(n)$. $\blacksquare$

Note that the similar transition from the mixing time of the block dynamics to that of the single site dynamics is not automatic — see Open Problem 1.

Finally, we state the lemma which was used in the coupling analysis. The crucial importance of this result will be explained at Open Problem 2.

Lemma 4.1 *Let T be a finite tree and π a Gibbs measure for the Ising model on this tree. For a fixed set of vertices A of T, $v \notin A$, and some boundary conditions τ, we consider the following conditional Gibbs expectations:*

$$\pi_+[\sigma_w] = \mathbf{E}[\sigma_w \mid \sigma_v = +], \qquad \pi_{+,\tau}[\sigma_w] = \mathbf{E}[\sigma_w \mid \sigma_v = +, \sigma_A = \tau]$$

for any $w \in T$. Similarly with the condition $\sigma_v = -$. Then

$$\pi_{+,\tau}[\sigma_w] - \pi_{-,\tau}[\sigma_w] \leq \pi_+[\sigma_w] - \pi_-[\sigma_w].$$

5 Open questions and future directions

1. Is the mixing time $O(n \log n)$ at high and intermediate temperatures? (Affirmative for $1 - 2\epsilon < \frac{1}{2\sqrt{b}}$.)

2. Note that Lemma 4.1 is intuitively obvious even for an arbitrary finite graph G instead of T: adding more boundary conditions makes the influence of one differing spin less significant. But it is only a conjecture that the lemma holds for general graphs. For trees, it is proven by explicit calculations, heavily using the simple tree structure.

Suppose one can prove the general lemma. The quantitative bounds in the four cases of the path coupling argument seem to rely mainly on growth conditions for T, and not too much on the tree structure, so linear relaxation time might follow for general nonamenable transitive graphs, maybe already for intermediate temperatures. By Theorem 3.3, this would imply a double phase transition for the number of extremal Gibbs measures on these graphs, a major open problem in this field [8].

3. In the case of a binary symmetric channel dealt with above, reconstruction by Global majority and general reconstruction have the same critical temperature. For a q-ary/asymmetric binary channel this is false. Critical value for majority reconstruction is known but not that for general reconstruction [7].

References

[1] Bleher, P. M., Ruiz, J. and Zagrebnov V. A. (1995) On the purity of limiting Gibbs state for the Ising model on the Bethe lattice, *J. Stat. Phys* **79**, 473–482.

[2] Bubley, R. and Dyer, M. (1997) Path coupling: a technique for proving rapid mixing in Markov chains. In Proceedings of the 38th Annual Symposium on Foundations of Computer Science (FOCS), 223–231.

[3] Chen, M. F. (1998) Trilogy of couplings and general formulas for lower bound of spectral gap. *Probability towards 2000* Lecture Notes in Statist., **128**, Springer, New York, 123–136.

[4] Evans, W., Kenyon, C., Peres, Y. and Schulman, L. J. (2000) Broadcasting on trees and the Ising Model, *Ann. Appl. Prob.*, **10**, 410–433.

[5] Kenyon, C., Mossel, E., Peres, Y. (2001) Glauber dynamics on trees and hyperbolic graphs. To appear in Proceedings of FOCS 2001.

[6] Martinelli, F. (1998) Lectures on Glauber dynamics for discrete spin models. *Lectures on probability theory and statistics (Saint-Flour, 1997)* 93–191, *Lecture Notes in Math.* **1717**, Springer, Berlin.

[7] Mossel, E. and Peres, Y. (2002) Information flow on trees. Preprint.

[8] Schonmann, R. H. (2001) Multiplicity of phase transitions and mean-field criticality on highly non-amenable graphs. *Comm. Math. Phys.* **219**, no. 2, 271–322.