arXiv:0908.2665v2 [math.PR] 30 Dec 2011

Phase Transition for the Mixing Time of the Glauber Dynamics for Coloring Regular Trees

Prasad Tetali * Juan C. Vera † Eric Vigoda ‡ Linji Yang‡

December 30th, 2011

Abstract

We prove that the mixing time of the Glauber dynamics for random k-colorings of the complete tree with branching factor b undergoes a phase transition at $k = b(1 + o_b(1))/\ln b$. Our main result shows nearly sharp bounds on the mixing time of the dynamics on the complete tree with n vertices for $k = Cb/\ln b$ colors with constant C. For $C \geq 1$ we prove the mixing time is $O(n^{1+o_b(1)} \ln n)$. On the other side, for $C < 1$ the mixing time experiences a slowing down, in particular, we prove it is $O(n^{1/C+o_b(1)} \ln n)$ and $\Omega(n^{1/C-o_b(1)})$. The critical point $C = 1$ is interesting since it coincides (at least up to first order) with the so-called reconstruction threshold which was recently established by Sly. The reconstruction threshold has been of considerable interest recently since it appears to have close connections to the efficiency of certain local algorithms, and this work was inspired by our attempt to understand these connections in this particular setting.

1 Introduction

There has been considerable interest in recent years in understanding the mixing time of Markov chains arising from single-site updates (known as Glauber dynamics) for sampling spin systems on finite graphs. The Glauber dynamics is well-studied both for its computational purposes, most immediately its use in Markov chain Monte Carlo (MCMC) algorithms, and for its physical motivation as a model of how physical systems reach equilibrium. Several works in this topic focus on exploring the dynamical and spatial connections between the mixing time and equilibrium properties of the spin system. A notable example of such equilibrium properties is the uniqueness of the infinite volume Gibbs measure, which very roughly speaking corresponds to the influence of a worst-case boundary condition. Recently a related weaker notion known as the *reconstruction threshold* has been the focus of considerable study. Reconstruction considers the influence of a "typical" boundary condition (we define it more precisely momentarily).

Much of the recent interest in reconstruction stems from its conjectured connections to the efficiency of local algorithms on trees and tree-like graphs, such as sparse random graphs. The Glauber dynamics is one particular example of such a local algorithm, another important example is the class of belief propagation algorithms. The work of Achlioptas and Coja-Oghlan [1] gives

*School of Mathematics and School of Computer Science, Georgia Institute of Technology, Atlanta GA 30332. Email: tetali@math.gatech.edu. Research supported in part by NSF grant DMS-0701043.

†Department of Econometrics and Operations Research, Tilburg University, Tilburg, The Netherlands. Email: j.c.veralizcano@uvt.nl.

‡School of Computer Science, Georgia Institute of Technology, Atlanta GA 30332. Email: {vigoda,ljyang}@gatech.edu. Research supported in part by NSF grant CCF-0830298 and CCF-0910584.

strong evidence for the "algorithmic barriers" that arise in the reconstruction phase for several constraint satisfaction problems, including colorings, on sparse random graphs. In this paper we show that the mixing time of the Glauber dynamics for random colorings of the complete tree undergoes a phase transition, and the critical point appears to coincide with the reconstruction threshold, at least up to a first order term.

We study the heat-bath version of the Glauber dynamics on the complete tree with branching factor b for the case of (proper vertex) k-colorings. Proper colorings correspond in the physics community to the zero-temperature limit of the anti-ferromagnetic Potts model, and the infinite complete tree is known as the Bethe lattice. Let $\mathcal{C} = \{1, 2, \ldots, k\}$ denote the set of k colors, and $T_\ell = (V, E)$ denote the complete tree with branching factor b, height ℓ and n vertices. We are looking at the set Ω of proper vertex k-colorings which are assignments $\sigma : V \to \mathcal{C}$ such that for all $(v, w) \in E$ we have $\sigma(v) \neq \sigma(w)$. The Glauber dynamics for colorings is a Markov chain (X_t) whose state space is Ω and transitions $X_t \to X_{t+1}$ are defined as follows:

- Choose a vertex v uniformly at random.
- For all $w \neq v$ set $X_{t+1}(w) = X_t(w)$.
- Choose $X_{t+1}(v)$ uniformly at random from its set of available colors $\mathcal{C} \setminus X_t(N(v))$ where $N(v)$ denotes the neighbors of v.

For the complete tree, when $k \geq 3$ the dynamics is ergodic where the unique stationary distribution is the uniform distribution over Ω. The mixing time is the number of steps, from the worst initial state, to reach within variation distance $\leq 1/2\mathrm{e}$ of the stationary distribution. We also consider the relaxation time which is the inverse of the spectral gap of the transition matrix. We formally define these notions in Section 3.

For general graphs of maximum degree b, the Glauber dynamics is ergodic when $k \geq b + 2$ and the best result for arbitrary graphs proves $O(n^2)$ mixing time when $k > 11b/6$ [31]. There are a variety of improvements for classes of graphs with high degree or girth, see [11] for a survey, and recently, Mossel and Sly [27] proved polynomial mixing time for sparse random graphs $G(n, d/n)$, for constant $d > 1$, for some constant number of colors.

There are two phase transitions of primary interest in the tree T_ℓ – uniqueness and reconstruction. These phase transitions are realized by analyzing the influence of the boundary condition, which in the case of tree corresponds to fixing the coloring of the leaves. We say uniqueness holds if for all boundary conditions, if we consider the uniform distribution conditional on the boundary condition, the influence at the root decays in the limit $\ell \to \infty$ (i.e., the root is uniformly distributed over the set $\mathcal{C}$ in the limit). Jonasson [17] established that the uniqueness threshold is at $k = b + 2$. When $k \leq b + 1$ it is not hard to see that there are boundary conditions which, in fact, "freeze" the root; moreover, the Glauber dynamics is not even ergodic in the case when $k = b + 2$. Martinelli et al [25] analyzed the Glauber dynamics on the tree T_ℓ with a fixed boundary condition. They proved a bound of $O(n \log n)$ on the mixing time when $k \geq b + 3$ for any boundary condition.

The reconstruction threshold corresponds to the influence of a random boundary condition. In particular, we first choose a random coloring of T_ℓ, the colors of the leaves are fixed, and we rechoose a random coloring for the internal tree from this conditional distribution. Reconstruction is said to hold if the leaves have a non-vanishing (as $\ell \to \infty$) influence on the root in expectation. We refer to the reconstruction threshold as the critical point for the transition between the reconstruction and non-reconstruction phases. It was recently established by Sly that the reconstruction threshold occurs at $k = b(1 + o(1))/\ln b$ [30, 4].

A general connection between reconstruction and the convergence time of the Glauber dynamics was shown by Berger et al [3] who showed, for general spin systems, that $O(n)$ relaxation time

on the complete tree (without boundary conditions) implies non-reconstruction. A new work of Ding et al [7] gives very sharp bounds on the mixing time of the Glauber dynamics for the Ising model on the complete tree, and illustrates it undergoes a phase transition at the reconstruction threshold. For the case of colorings, recently Hayes et al [14] proved polynomial mixing time of the Glauber dynamics for any planar graph with maximum degree b when $k > 100b/\ln b$. Subsequently, improved results were established for the tree. In particular, Goldberg et al [12] proved the mixing time is $n^{\Omega(b/(k\ln b))}$ for the complete tree with branching factor b, and Lucier et al [21] proved the mixing time is $n^{O(1+b/(k\ln b))}$ for any tree with maximum degree b and the number of colors $k \geq 4$. In a follow-up paper, Lucier et al [22] further prove the same upper bound for the case when $k = 3$.

Our goal is to understand the relationship between the reconstruction threshold and the mixing time. Thus we want to establish a more precise picture than provided by the results of [12] and [21]. Our main result provides (nearly) sharp bounds on the mixing time and relaxation time of the Glauber dynamics for the complete tree, establishing a phase transition at the critical point $k = b(1 + o_b(1))/\ln b$. Our proofs build upon the approaches used by [12] and [21].

Theorem 1. *For all $C > 0$, there exists b_0 such that, for all $b > b_0$, for $k = Cb/\ln b$, the Glauber dynamics on the complete tree T on n vertices with branching factor b and height $H = \lfloor \log_b n \rfloor$ satisfies the following:*

1. **For $C \geq 1$:**

$$\begin{aligned} \Omega\left(n \ln n/(b \text{ poly}(\log b))\right) &\leq T_{\text{mix}} \leq O(n^{1+o_b(1)} \ln n) \\ \Omega(n) &\leq T_{\text{relax}} \leq O(n^{1+o_b(1)}) \end{aligned}$$

2. **For $C < 1$:**

$$\begin{aligned} \Omega(n^{1/C-o_b(1)}) &\leq T_{\text{mix}} \leq O(n^{1/C+o_b(1)} \ln n) \\ \Omega(n^{1/C-o_b(1)}) &\leq T_{\text{relax}} \leq O(n^{1/C+o_b(1)}) \end{aligned}$$

where the $o_b(1)$ functions are $O(\ln\ln b/\ln b)$ for the upper bounds, $b^{1-1/C}/C$ for the lower bounds when $1/2 < C < 1$ and exactly zero for the lower bounds when $0 < C \leq 1/2$. The constants in the $\Omega()$ and $O()$ are universal constants.

Remark. When $C \geq 1$, the lower bound of the mixing time is proved by Hayes and Sinclair [13] in a more general setting, and for the particular case of the heat-bath version of the Glauber dynamics on the complete tree, we believe it can be improved to $\Omega(n \ln n/\text{poly}(\log b))$ by the same proof. The lower bound of the relaxation time simply follows from the fact that the probability of selecting a specific vertex to recolor in one step of the dynamics is $1/n$. Note, the results of Berger et al [3] imply a lower bound of $T_{\text{relax}} \geq \omega(n)$ for the case $C < 1$ since reconstruction holds in this region.

Our result extends to more general k and b, thereby refining the general picture provided by [12] and [21].

Theorem 2. *There exists b_0 such that, for all k, b satisfying $b/(k\ln b) > 2$ and $b > b_0$, the Glauber dynamics on the complete tree of n vertices with branching factor b satisfies the following:*

$$\begin{aligned} \Omega(n^{b/(k\ln b)}) &\leq T_{\text{mix}} \leq O(n^{b/(k\ln b)+\gamma} \ln n) \\ \Omega(n^{b/(k\ln b)}) &\leq T_{\text{relax}} \leq O(n^{b/(k\ln b)+\gamma}), \end{aligned}$$

where

$$\gamma = \gamma(b) = 1 - \frac{\ln k}{\ln b} + \frac{\ln\ln b}{\ln b} + \frac{O(1)}{\ln b}$$

is at most a small constant.

Remark. The constants in the $\Omega()$ and $O()$ of Theorem 2 are universal constants. Also, note that when $k = b^\alpha$ for constant $\alpha < 1$, then $\lim_{b\to\infty} \gamma = 1 - \alpha$, and when k is constant, then $\lim_{b\to\infty} \gamma = 1$.

2 Proof Overview

We now give an outline of the proofs of Theorem 1. Readers can refer to Section 3 for the definitions and background materials.

2.1 Upper bounds

We first sketch the proof approach for upper bounding the mixing time and relaxation time. Let $G^* = (V, E)$ be the star graph on $b+1$ vertices, i.e., the complete tree T_1 of height 1 with b leaves, and H be the height of the complete tree T_H, i.e., $H = \lfloor \log_b n \rfloor$. Let τ^* be the relaxation time of the Glauber dynamics on the star graph G^* using k colors.

We use the following decomposition result of Lucier and Molloy [21], which is an application of the block dynamics technique (see, Proposition 3.4 in [23]) to the Glauber dynamics on the complete trees combined with an earlier result proved by Berger et al which shows that the relaxation time of this special block dynamics is the same as that of the Glauber dynamics on the star graph (see, Claim 2.9 in [3]).

Theorem 3. *The relaxation time T_{relax} of the Glauber dynamics on the complete tree of height H with branching factor b satisfies*

$$T_{\text{relax}} \le (\tau^*)^H.$$

Therefore, proving the upper bounds in Theorem 1 reduces to the problem of getting tight upper bounds of the relaxation time τ^* of the Glauber dynamics on G^*. In [21], the authors used a canonical path argument to bound $\tau^* = O(b^{2+1/C}k)$ for any $C > 0$. Instead, here we use two different coupling arguments to show the following two theorems for τ^*.

Theorem 4. *For any $C < 1$, there exists $b_0 > 0$ such that, for any $b > b_0$, the mixing and relaxation times of the Glauber dynamics on G^* using $k = Cb/\ln b$ colors are $O(b^{1/C} \ln^2 b)$. When $C = 1$, the mixing and relaxation times are $O(b \ln^4 b)$.*

Theorem 5. *For any $C > 1$, there exists $b_0 > 0$ such that, for any $b > b_0$, the mixing and relaxation times of the Glauber dynamics on G^* using $k \ge Cb/\ln b$ colors are $O(b \ln b)$.*

Remark. It can be shown that the relaxation time is actually $O(b)$ when $C > 1$, from our analysis. However, unless we can also eliminate the constant factors and thereby show a very sharp bound of at most b, the extra $\ln b$ factor makes little difference to the relaxation time of the dynamics on the whole tree.

The most difficult (and also interesting) case turns out to be when $C \le 1$. We will prove Theorem 4 in Section 4 and Theorem 5 in Section 5. We sketch the high-level idea of the proof of Theorem 4 in Section 4.1. Having Theorems 4 and 5 in hand, we can then apply Theorem 3 to get the upper bounds on the relaxation time as stated in Theorem 1. We get

$$T_{\text{relax}} = \begin{cases} O(b \ln b)^H = O\left(n^{1+(\ln\ln b + O(1))/\ln b}\right), & \text{if } C > 1; \\ O(b \ln^4 b)^H = O\left(n^{1+(4\ln\ln b + O(1))/\ln b}\right), & \text{if } C = 1; \\ O(b^{1/C} \ln^2 b)^H = O\left(n^{1/C+(2\ln\ln b + O(1))/\ln b}\right), & \text{if } C < 1. \end{cases}$$

To then get the desired upper bounds on the mixing time of the whole tree we need a slightly more advanced tool, the logarithmic Sobolev constant of the Markov chain; we define the log-Sobolev constant formally in the next section along with the other technical preliminaries. By adapting Theorem 5.7 in Martinelli, Sinclair and Weitz [24] to our setting of colorings, we establish and improve (in Section 8) the following relationship between the inverse of the log-Sobolev constant c_{sob}^{-1} and the relaxation time T_{relax} of the Glauber dynamics on trees.

Theorem 6.

$$c_{sob}^{-1} \leq T_{\text{relax}} \cdot 2b \ln(k).$$

Since the inverse of the log-Sobolev constant gives a relatively tight upper bound on the mixing time (see Inequality (2) in Section 3), using Theorem 6 we are able to complete the proofs of the upper bounds in Theorem 1.

2.2 Lower bounds

Our proof of the lower bound in Theorem 1 when $C < 1$ builds upon the approach used in [12]. They lower bounded the relaxation time by upper bounding the conductance of the Glauber dynamics on the subset $S \subseteq \Omega$ where the root is *frozen* (meaning that the configuration at the leaves uniquely determine the color of the root) to some color in $\{1, 2, ..., \lfloor k/2 \rfloor\}$. They showed the conductance of S satisfies $\Phi_S = O(n^{-1/6C})$ when $0 < C < 1/2$, which implies (by (1) and (3) in Section 3) that $T_{\text{mix}} \geq T_{\text{relax}} - 1 = \Omega(n^{1/6C})$.

We improve their bound on the conductance of S by analyzing the probability that for a given leaf z, in a random coloring σ of the complete tree, the root is frozen and changing the color of z in σ to some other color unfreezes the root. We prove that the number of such leaves in most colorings that freeze the root is $O(n^{-1/C+1+o_b(1)})$. Since the probability of recoloring a specific leaf is $1/n$, then intuitively we have $\Phi_S = O(n^{-1/C+o_b(1)})$, and hence $T_{\text{mix}} \geq T_{\text{relax}} - 1 = \Omega(n^{1/C-o_b(1)})$. A complete analysis of the lower bound is in Section 6, and in the analysis we will see that the $o_b(1)$ error term is $b^{1-1/C}/C$ when $1/2 < C < 1$ and zero when $C \leq 1/2$.

Finally, we will show in Section 7 how all of the proofs generalize for $k = o(b/\ln b)$, and thus prove Theorem 2.

3 Technical Preliminaries

Let $P(\cdot, \cdot)$ denote the transition matrix of the Glauber dynamics, and $P^t(\cdot, \cdot)$ denote the t-step transition probability. The total variation distance at time t from initial state σ is defined as

$$\|P^t(\sigma, \cdot) - \pi\|_{TV} := \frac{1}{2} \sum_{\eta} |P^t(\sigma, \eta) - \pi(\eta)|.$$

The mixing time T_{mix} for a Markov chain is then defined as

$$T_{\text{mix}} = \min_t \{\max_\sigma \{\|P^t(\sigma, \cdot) - \pi\|_{TV}\} \leq 1/2\text{e}\}.$$

Given two copies, (X_t) and (Y_t), of the Markov chain at time $t > 0$, recall that a (one-step) coupling of (X_t) and (Y_t), is a joint distribution whose left and right marginals are identical to the (one-step) evolution of (X_t) and (Y_t), respectively. The Coupling Lemma [2] (c.f., Theorem 5.2 in [20]) guarantees that if, there is a coupling and time $t > 0$, so that for every pair (X_0, Y_0) of initial states, $\Pr[X_t \neq Y_t] \leq 1/2\text{e}$ under the coupling, then $T_{\text{mix}} \leq t$.

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|\Omega|}$ be the eigenvalues of the transition matrix P. The spectral gap c_{gap} is defined as $1 - \lambda$ where $\lambda = \max\{\lambda_2, |\lambda_{|\Omega|}|\}$ denotes the second largest eigenvalue in absolute value. The relaxation time T_{relax} of the Markov chain is then defined as c_{gap}^{-1}, the inverse of the spectral gap. It is an elementary fact that the mixing time gives a good upper bound on the relaxation time (see, e.g., [10] for the following bound), which we will use in our analysis:

$$T_{\text{relax}} \leq T_{\text{mix}} + 1. \tag{1}$$

Since we will also work with the logarithmic Sobolev constant of a (finite) Markov chain, we briefly recall here the variational definition of both the spectral gap and the log-Sobolev constant.

Let f be a function (vector) from Ω to R, π be the uniform distribution over Ω and μ be any probability distribution over Ω. Let $\mathcal{D}(f)$ be the standard *Dirichlet form* of the heat-bath Glauber dynamics defined as:

$$\mathcal{D}(f) = \frac{1}{2} \sum_{\sigma} \sum_{\sigma'} (f(\sigma) - f(\sigma'))^2 \pi(\sigma) P(\sigma, \sigma').$$

Let $E_\mu(f)$ be the average of f under the distribution μ, and let $\text{Var}_\mu(f) := E_\mu(f^2) - E_\mu^2 f$ be the corresponding variance, which can also be written as:

$$\text{Var}_\mu(f) = \frac{1}{2} \sum_{\sigma} \sum_{\sigma'} (f(\sigma) - f(\sigma'))^2 \mu(\sigma) \mu(\sigma')$$

Let $\text{Ent}_\mu(f) := E_\mu(f \log f) - E_\mu(f) \log(E_\mu(f))$. When it is clear what the underlying distribution is we will drop the subscript μ in the notation $\text{Ent}(f)$.

The spectral gap c_{gap} is equivalently defined as (see, e.g., Chapter 13 in [20])

$$c_{gap} = \inf_f \frac{\mathcal{D}(f)}{\text{Var}(f)}$$

and the log-Sobolev constant c_{sob} is defined as (see, e.g., [6]),

$$c_{sob} = \inf_{f \geq 0} \frac{\mathcal{D}(\sqrt{f})}{\text{Ent}(f)}$$

where the infimum in both equations is over non-constant functions f.

For the upper bounds on the mixing time of the dynamics on the whole tree, we also use the following well-known relationship between the mixing time and the inverse of the log-Sobolev constant (see e.g. [6] for more details):

$$T_{\text{mix}} \quad = \quad O\left(c_{sob}^{-1} \ln \ln \frac{1}{\min_{\sigma \in \Omega}\{\pi(\sigma)\}} \right). \tag{2}$$

To lower bound the mixing and relaxation times we analyze the conductance. The conductance of the Markov chain on Ω with transition matrix P is given by $\Phi = \min_{S \subseteq \Omega}\{\Phi_S\}$, where Φ_S is the conductance of a specific set $S \subseteq \Omega$ defined as

$$\Phi_S = \frac{\sum_{\sigma \in S} \sum_{\eta \in \bar{S}} \pi(\sigma) P(\sigma, \eta)}{\pi(S)\pi(\bar{S})}.$$

Thus, a general way to find a good upper bound on the conductance is to find a set S such that the probability of escaping from S is relatively small. The well-known relationship between the relaxation time and the conductance is established in [18] and [29] and we will use the form

$$T_{\text{relax}} = \Omega(1/\Phi), \tag{3}$$

for proving the lower bounds.

4 Upper Bound on Mixing Time for $C \leq 1$: Proof of Theorem 4

In this section, we upper bound the mixing time of the Glauber dynamics on the star graph $G^* = (V, E)$ when $k = Cb/\ln b$ for any $C \leq 1$. To be more precise, let $V = \{r, \ell_1, ..., \ell_b\}$, where r refers to the root and $\ell_1, ..., \ell_b$ are the b leaves and $E = \{(r, \ell_1), ..., (r, \ell_b)\}$. For convenience, here we let

$$\epsilon := 1/C - 1,$$

and hence $k = b/((1+\epsilon)\ln b)$.

We use the maximal one-step coupling, originally studied for colorings by Jerrum [16] to upper bound the mixing time of the Glauber dynamics on general graphs. For a coloring $X \in \Omega$, let $A_X(v)$ denote the set of available colors of v in the coloring X, i.e., $A_\sigma(v) = \{c \in \mathcal{C} : \forall u \in N(v), \sigma(u) \neq c\}$. The coupling (X_t, Y_t) of the two chains is done by choosing the same random vertex v_t for recoloring at step t and maximizing the probability of the two chains choosing the same update for the color of v_t. Thus, for each color $c \in A_{X_t}(v) \cap A_{Y_t}(v)$, with probability $1/\max\{|A_{X_t}(v)|, |A_{Y_t}(v)|\}$ we set $X_{t+1}(v) = Y_{t+1}(v) = c$. With the remaining probability, the color choices for $X_{t+1}(v)$ and $Y_{t+1}(v)$ are coupled arbitrarily.

We prove the theorem by analyzing the coupling in rounds, where each round consists of $T := 20b\ln b$ steps. Our main result is the following lemma which says that in each round we have a good probability of coalescing (i.e., achieving $X_t = Y_t$).

Lemma 7. *For all $\epsilon \geq 0$, there exists $b_0(\epsilon)$ such that for all $b > b_0(\epsilon)$ if $k = b/((1+\epsilon)\ln b)$ and $T = 20b\ln b$ for all $(x_0, y_0) \in \Omega \times \Omega$, the following holds:*

$$\Pr\left[X_T = Y_T \mid X_0 = x_0, Y_0 = y_0\right] \geq \begin{cases} (20(1+\epsilon)b^\epsilon \ln b)^{-1}, & \text{if } \epsilon > 0; \\ (20\ln^3 b)^{-1}, & \text{if } \epsilon = 0. \end{cases}$$

It is then straightforward to prove Theorem 4.

Proof of Theorem 4. For $\epsilon > 0$, let $p_T := (20(1+\epsilon)b^\epsilon \ln b)^{-1}$; and for $\epsilon = 0$ let $p_T := (20\ln^3 b)^{-1}$. By repeatedly applying Lemma 7 we have, for all (x_0, y_0),

$$\Pr\left[X_{2iT} \neq Y_{2iT} \mid X_0 = x_0, Y_0 = y_0\right] \leq (1 - p_T)^{2i} \leq 1/2\mathrm{e}$$

for $i = 1/p_T$. Therefore, by applying the Coupling Lemma, mentioned in Section 3, the mixing time is $O((1+\epsilon)b^{1+\epsilon}\ln^2 b)$ for $\epsilon > 0$ and $O(b\ln^4 b)$ for $\epsilon = 0$. □

4.1 Overview of the Coupling Argument

Before formally proving Lemma 7 we give a high-level overview of its proof. We will analyze the maximal one-step coupling on the star graph G^*. We say a vertex v "disagrees" at time t if $X_t(v) \neq Y_t(v)$, otherwise we say the vertex v "agrees". We denote the set of disagreeing vertices at time t of our coupled chains by

$$D_t = \{v \in V : X_t(v) \neq Y_t(v)\},$$

and we use $D_t^L = D_t \setminus \{r\}$ to represent the set of disagreeing leaves. When we use the term "with high probability" in this section, it means that the probability goes to 1 as b goes to infinity.

If the coupling selects a leaf ℓ to recolor at time t, then the probability that ℓ disagrees in X_t and Y_t is at most $1/(k-1)$, and with probability at least $(k-2)/(k-1)$, the leaf will use the same color that is chosen uniformly at random from $\mathcal{C} \setminus \{X_t(r), Y_t(r)\}$. We also know that if we simply

assign a random color from $\mathcal{C}$ to each leaf, with probability at least $\Omega(1/(b^\epsilon \ln b))$, there is a color in $\mathcal{C}$ that is unused in any leaf. This last point hints at the success probability in the statement of Lemma 7.

We analyze the T-step epoch in three stages. The warm-up round is of length $T_{\mathrm{w}} := 8(b+1)\ln b$ steps. We will show in Lemmas 10 and 11 that with good probability, after the warm up, all of the leaf disagreements will be of the same form in the sense that they will have the same pair of colors.

The next stage is of a random length T_1, which is defined as the first time (after T_{w}) where we are recoloring the root and the root has a common available color in (X_t) and (Y_t). We prove in Lemma 12, that with probability $\Omega(1/b^\epsilon \ln b)$, $T_1 < 4(b+1)\ln b$. We then have probability at least $1/2$ of the root agreeing after it is updated, and then after at most $T_2 := 4(b+1)\ln b$ further steps we are likely to coalesce since we just need to recolor each leaf at least once before the root changes back to a disagreement.

4.2 Coupling Argument: Proof of Lemma 7

We begin with a basic observation about the maximal one-step coupling.

Observation 8. *Let $\mathcal{C}(D_t^L) := \bigcup_{\ell \in D_t^L} \{X_t(\ell), Y_t(\ell)\}$ denote the set of colors that appear in the disagreeing leaves at time t. Then, $A_{X_t}(r) \oplus A_{Y_t}(r) \subseteq \mathcal{C}(D_t^L)$.*

This is simply because those colors that appear on the leaves with agreements are both unavailable in X_t and Y_t for the root. We now analyze the first stage of the T-step epoch.

Proposition 9. *The probability that in $T_0 = 4(b+1)\ln b$ steps, the coupling (X_t, Y_t) (or the Glauber dynamics (X_t)) will recolor the root at most $20\ln b$ times and recolor every leaf at least once is at least $1 - 2b^{-3}$.*

Proof. Using the union bound the probability that there is a leaf which is not recolored in T_0 steps is at most

$$b\left(1 - \frac{1}{b+1}\right)^{4(b+1)\ln b} \le b^{-3}.$$

Now, let N be the number of times the root is recolored in T_0 steps. The expectation $\mathrm{E}[N]$ is $4\ln b$. Then, by the Chernoff bound (see, e.g., Theorem 4.5 Part 2 in [26])

$$\Pr[N \ge 20\ln b] \le \Pr[N \ge (1+4)\mathrm{E}[N]] \le b^{-3}.$$

Therefore the lemma holds by the union bound. □

Then we will prove that in $T_{\mathrm{w}} = 2T_0$ steps, with high probability all of the leaf disagreements are of the same type when $\epsilon > 0$.

Lemma 10. *For any $\epsilon > 0$ and $k > (1+\epsilon)b/\ln b$, for any pair of initial states (x_0, y_0),*

$$\Pr\left[\forall \ell \in D_{T_w}^L, X_{T_w}(\ell) = Y_{T_w}(r) \wedge Y_{T_w}(\ell) = X_{T_w}(r) \mid x_0, y_0\right] \ge 1 - O\left(\frac{1}{b^\epsilon}\right).$$

Proof. The idea is that if we just look at one chain, say (X_t), then after T_0 steps, with high probability the root is frozen. Moreover, the root is likely to continue to be frozen for the remainder of the T_{w} steps since we recolor the root at most $O(\ln b)$ times. In the worst case the root is frozen to a disagreement, say $X_t(r) = 2$ and $Y_t(r) = 1$. Then after recoloring a leaf ℓ at time t' where

$t < t' < T_{\mathrm{w}}$, the only possible disagreement is $X_{t'}(\ell) = 1, Y_{t'}(\ell) = 2$. Hence, it suffices to recolor each leaf at least once.

Let $\mathcal{E}$ be the event that in the first T_0 steps, every leaf is recolored at least once and in another $4(b+1)\ln b$ steps, every leaf is recolored again at least once and the root is recolored at most $20\ln b$ times. We are first going to bound that for $t > T_0$,

$$\Pr\left[|A_{X_t}(r)| > 1 \mid \mathcal{E}\right] \leq \frac{1}{(1+\epsilon)b^{\epsilon}\ln b} := p_0, \tag{4}$$

and the same thing happens for Y_t.

Let G_W be the graph with b isolated vertices $\{v_1, ..., v_b\}$, corresponding to the leaves $\{\ell_1, ..., \ell_b\}$. Let (W_t) be a Glauber process on G_W using $k-1$ colors from another color set $\mathcal{C}_W$. We are going to define W_0 and couple (W_t) with (X_t) such that $|A_{X_t}(r)| = |A_{W_t}| + 1$ at any time t, where $A_{W_t} := \{c \in \mathcal{C}_W : \forall v_i, W_t(v_i) \neq c\}$. To do this, for every t we are going to define a bijection $f_t : \mathcal{C} \setminus \{X_t(r)\} \to \mathcal{C}_W$ such that $f_t(X_t(\ell_i)) = W_t(v_i)$ for all i. Notice that if such a bijection exists then $|A_{X_t}(r)| = |A_{W_t}| + 1$.

At time $t = 0$, pick any bijection f_0 from $\mathcal{C}_W$ to $\mathcal{C} \setminus \{X_0(r)\}$. Define W_0 by $W_0(v_i) = f(X_0(\ell_i))$ for all i. We will update f_t only when we choose the root to recolor at time t in the coupling of (W_t) and (X_t). To do the coupling at time $t+1$, we first choose a vertex v in G^* to recolor:

- if $v = \ell_i$, then we choose a random color c different from $X_t(r)$ to recolor v. Correspondingly, we choose the vertex v_i in G_W to recolor using color $f_t(c)$.
- if $v = r$, then we choose a random color c from $A_{X_t}(r)$ to recolor the root in G^*. Correspondingly, we update the mapping f_t in the following natural way: $f_t(X_{t-1}(r)) = f_{t-1}(c)$, (and $f_t(c)$ is undefined).

Since (W_t) itself is a Glauber process that recolors the vertices of G_W uniformly at random from C_W, conditioning on $\mathcal{E}$, simple calculations yield that for any $t > T_0$,

$$\Pr\left[|A_{W_t}| \geq 1 \mid \mathcal{E}\right] \leq \frac{1}{(1+\epsilon)b^{\epsilon}\ln b}\,.$$

Then (4) follows by coupling.

Since the same thing happens for (Y_t) and the root is recolored at most $20\ln b$ times, then by the union bound, conditioning on $\mathcal{E}$, the probability that at each time we try to recolor the root after T_0 steps, the root is always frozen in both copies is at least $1 - (40\ln b)(p_0) = 1 - 40/((1+\epsilon)b^{\epsilon})$. Finally, by Proposition 9, $\mathcal{E}$ happens with high probability, and hence the lemma holds. □

Note that for the warm-up stage, we need to show, with probability at least $1/\mathrm{poly}(\log b)$, that for $\epsilon \geq 0$, all of the leaf disagreements are of the same type in $O(b\ln b)$ steps. This is easier to prove for the $\epsilon > 0$ case – that this happens with high probability, if we run the dynamics for $T_{\mathrm{w}} = 8(b+1)\ln b$ steps. For the threshold case when $\epsilon = 0$, we will prove a slightly weaker lemma, in the sense that the successful probability will be at least $\Omega(1/\ln^2 b)$.

Lemma 11. *Let $T'_w = T_0 + 2b\ln\ln b$. For $k = b/\ln b$, for any pair of initial states (x_0, y_0),*

$$\Pr\left[\forall \ell \in D^L_{T'_w}, X_{T'_w}(\ell) = Y_{T'_w}(r) \wedge Y_{T'_w}(\ell) = X_{T'_w}(r) \mid x_0, y_0\right] \geq 1/(2\ln^2 b).$$

Proof. We use a different approach to prove this lemma, since it is not true that the root will still always be frozen during T'_{w} steps with high probability.

Let $T_0 = 4(b+1)\ln b$. We first prove that after T_0 steps, with high probability, the number of disagreeing leaves is at most $O(\ln b)$, namely:

$$\Pr\left[|D^L_{T_0}| \geq 4\ln b \mid X_0 = x_0, Y_0 = y_0\right] \leq \frac{2}{b^2}. \tag{5}$$

To prove (5), we construct a simpler process that stochastically upper bounds the number of disagreements. We define the following Markov chain (U_t) on 2-colorings of the graph G_U which consists of b isolated vertices $\{v_1, ..., v_b\}$. We view the set of colors as $\{0, 1\}$. In each step, a random vertex v_i is chosen, then with probability $1/(k-1)$, v_i is recolored to 1, and with probability $1 - 1/(k-1)$, v_i is recolored to 0. Let $D^U_t = \{v \in \{v_1, v_2, ..., v_b\} : U_t(v) = 1\}$. The initial state U_0 is constructed in the following way: for any $i > 0$, $U_0(v_i) = 1$ if and only if $x_0(\ell_i) \neq y_0(\ell_i)$. By associating the b vertices of G_U with the leaves of G^*, we can easily couple the process (U_t) with (X_t, Y_t) such that $|D^U_t| \geq |D^L_t|$.

Let $\mathcal{E}$ denote the event that all of the vertices of G_U are recolored at least once in T_0 steps. Note, $\Pr[\mathcal{E}] \geq 1 - 1/b^2$. Conditioned on $\mathcal{E}$, the expected size of $|D^U_{T_0}|$ is $b/(k-1) \approx \ln b$. Then we have

$$\begin{aligned}\Pr\left[|D^U_{T_0}| \geq 4\ln b\right] &\leq \Pr\left[|D^U_{T_0}| \geq 4\ln b \mid \mathcal{E}\right] + \Pr[\mathcal{E}] \\ &\leq \frac{2}{b^2}.\end{aligned}$$

Where, for the last inequality, we have used the Chernoff bounds (see, e.g., Theorem 4.5 Part 2 in [26]). Since $|D^U_t| \geq |D^L_t|$, this proves (5).

Hence, with high probability there are $O(\ln b)$ disagreeing leaves in G^* at time T_0. Notice that from time T_0, if we recolor all of the disagree leaves before we recolor the root again, then all of the remaining disagreements in the leaves will be of the same type (more precisely, for such a leaf ℓ that becomes a disagreement at time t we will have that $X_t(\ell) = Y_{T_0}(r)$ and $Y_t(\ell) = X_{T_0}(r)$), and this implies the desired conclusion of the lemma. To this end, let $\mathcal{E}_2$ be the event that the root is not chosen from recoloring from time T_0 to T'_{w}. Let $\mathcal{E}_3$ be the event that each leaf in $D^L_{T_0}$ is recolored at least once in the interval of times $[T_0, T'_{\mathrm{w}}]$. By simple calculations, we have that:

$$\Pr[\mathcal{E}_2] \geq \ln^{-2} b, \quad \Pr[\mathcal{E}_3 \mid \mathcal{E}_2] \geq 1 - \frac{O(1)}{\ln b}. \tag{6}$$

Therefore, conditioned on $|D^L_{T_0}| \leq 4\ln b$, from time T_0 to T'_{w} with probability at least $2/(3\ln^2 b)$, both $\mathcal{E}_2$ and $\mathcal{E}_3$ happen, which implies all of the leaf disagreements will be of the same type at time T'_{w}.

In conclusion, combining the above bounds with (5), we proved that with probability at least $1/(2\ln^2 b)$, all of the uncoupled leaves are of the same at time T'_{w}. □

After we succeed in the warm-up stage meaning that all of the leaf disagreements are of the same type, we enter the root-coupling stage, where we try to couple the root. Let T_1 be the first time that there is a common available color in the root and the coupling chain selects the root to recolor, that is

$$T_1 := T_1^{XY} = \min\{t : A_{X_t}(r) \bigcap A_{Y_t}(r) \neq \emptyset \text{ and the root } r \text{ is selected at step } t\}.$$

Lemma 12. *For $\epsilon \geq 0$, for any pair of initial states (x_0, y_0) where all of the leaf disagreements are of the same type (i.e., there is a pair of colors c_1, c_2 such that for all $\ell \in D^L_0$, we have $x_0(\ell) = c_1$ and $y_0(\ell) = c_2$), we have*

$$\Pr\left[T_1^{XY} < 4(b+1)\ln b \mid (X_0, Y_0) = (x_0, y_0)\right] > \frac{1}{4(1+\epsilon)b^\epsilon \ln b}.$$

Proof. First of all, by Proposition 8, $|A_{X_0}(r) \oplus A_{Y_0}(r)| \leq 2$. We are interested in the time t when there is a common color available for the root in (X_t, Y_t).

Let (Z_t) be a Glauber process on the graph G_Z of $b+1$ isolated vertices $\{v_0, v_1, v_2, ..., v_b\}$ in which v_0 corresponds to the root and v_i corresponds to the leaves ℓ_i for any $i > 0$. The color set used in the process (Z_t) is $\mathcal{C}_Z = [k] \setminus \{c_1, c_2\}$. Each step, (Z_t) chooses a random vertex and recolors it with a random color from the set $\mathcal{C}_Z$. Let T_Z be the stopping time on Z satisfying:

$$T_1^Z = \min\{t > 2(b+1)\ln b : |A_{Z_t}| \geq 1 \text{ and } v_0 \text{ is selected at the step } t\},$$

where $A_{Z_t} = \{c \in \mathcal{C}_Z : \forall i \in [1, .., b], Z_t(v_i) \neq c\}$ is the set of unused colors in the vertices $\{v_1, v_2, ...v_b\}$. We want to couple (Z_t) with (X_t, Y_t) in such a way that $T_1^Z \geq T_1^{XY}$ for all the runs, and then if we show that for any initial state z_0, we have

$$\Pr\left[T_1^Z < 4(b+1)\ln b \mid Z_0 = z_0\right] > \frac{1}{4(1+\epsilon)b^\epsilon \ln b}. \tag{7}$$

Then by the coupling, we know that the lemma is also true.

Now we are going to construct the coupling between (Z_t) and (X_t, Y_t) for $t \leq T_1^{XY}$. Let z_0 be the initial state satisfying that for any $i \in [1, .., b]$, if $x_0(\ell_i) = y_0(\ell_i) \in \mathcal{C}_Z$ then $z_0(v_i) = x_0(\ell_i)$, otherwise we give an arbitrary color to the vertex v_i. On each step t, we first randomly select a vertex in G^* to update in (X_t, Y_t) and accordingly we select the corresponding vertex in G_Z to update in Z_t:

- If the vertex is a leaf ℓ_i:

  (X_t, Y_t) selects a random color c or a disagreement to update. If $c \in \mathcal{C}_Z$ then we give the same color to v_i in Z_t, otherwise we give a random color to v_i.

- If the vertex is the root r:

  Recolor the root on (X_t, Y_t) according to the maximal one-step coupling and pick a random color in $\mathcal{C}_Z$ to recolor v_0 in Z.

Observe that, $A_{Z_t} \subseteq A_{X_t}(r) \bigcap A_{Y_t}(r)$ for any $0 \leq t \leq T_1^{XY}$, which implies that $T_1^Z \geq T_1^{XY}$ holds with probability 1. Now we will show that (7) holds. Let $\mathcal{E}$ be the event that, in (Z_t), every vertex in the graph G_Z will be recolored at least once within the first $2(b+1)\ln b$ steps. Let t_z be the first time after time $2(b+1)\ln b$ when the dynamics (Z_t) recolors the root. For each color $c \in \mathcal{C}_Z$, define the indicator function $\mathbf{1}_c := \mathbf{1}\{c \neq Z_{t_z}(v_i), \forall 1 \leq i \leq b\}$. These indicator functions are negatively associated to each other (c.f., Theorem 14 in [8]). It follows by elementary calculations that conditioned on $t_z = t$ for some $t > 2(b+1)\ln b$ and for large enough b, we have

$$\begin{aligned}
&\Pr\left[A_{Z_t} \neq \emptyset \mid t_z = t\right] \\
&\quad\geq \Pr[\mathcal{E}] \cdot \Pr\left[A_{Z_t} \neq \emptyset \mid t_z = t, \mathcal{E}\right] \\
&\quad\geq 0.99 \Pr\left[A_{Z_t} \neq \emptyset \mid t_z = t, \mathcal{E}\right] \qquad (\text{since } \Pr[\mathcal{E}] > 1 - 1/b^2) \\
&\quad\geq 0.99\left(1 - \prod_{c \in \mathcal{C}_z} \Pr\left[\mathbf{1}_c = 0 \mid t_z = t, \mathcal{E}\right]\right) \qquad (\text{negative association}) \\
&\quad\geq 0.99\left(1 - \left(1 - \left(1 - \frac{1}{|\mathcal{C}_Z|}\right)^b\right)^{|\mathcal{C}_Z|}\right) \\
&\quad\geq \frac{1}{3(1+\epsilon)b^\epsilon \ln b}.
\end{aligned} \tag{8}$$

Since $\Pr[t_z \le 4(b+1)\ln b] > 1 - 1/b^2$, by applying (8) we have

$$\begin{aligned}\Pr\left[T_1^Z < 4(b+1)\ln b \mid Z_0 = z_0\right] &\ge \sum_{t=2(b+1)\ln b}^{4(b+1)\ln b} \Pr[A_{Z_t} \neq \emptyset \mid t_z = t] \cdot \Pr[t_z = t] \\ &\ge \frac{\Pr[t_z \le 4(b+1)\ln b]}{3(1+\epsilon)b^\epsilon \ln b} \\ &\ge \frac{1}{4(1+\epsilon)b^\epsilon \ln b}.\end{aligned}$$

This completes the proof of Lemma 12. □

We also know that when the root is recolored, if $|A_X(r) \oplus A_Y(r)| \le 2$ and $|A_X(r) \bigcap A_Y(r)| \ge 1$ holds, then the probability that the root will be recolored to the same color in both X and Y is at least $1/2$. Hence, at time $T_1 = T_1^{XY}$, with probability at least $1/2$ the root will become an agreement. Combining with Lemma 10, we proved that with probability at least $1/O((1+\epsilon)b^\epsilon \ln b)$ when $\epsilon > 0$, starting from arbitrary initial states (x_0, y_0), the root will couple in at most $12(b+1)\ln b$ steps and by that time all the disagreements (if there is any) in the leaves are of the same type. When $\epsilon = 0$, combining with Lemma 11, we get that the probability of the same event happening is at least $1/O(\ln^3 b)$.

The last step is to let all of the disagreements in the leaves go away without changing the root to a disagreement, again with constant probability, after $T_2 = 4(b+1)\ln b$ more steps. Here is the precise statement of the lemma.

Lemma 13. *For $\epsilon \ge 0$, consider a pair of initial states (x_0, y_0) where the root r agrees (i.e., $x_0(r) = y_0(r)$) and all of the leaf disagreements are of the same type (i.e., there is a pair of colors c_1, c_2 such that for all $\ell \in D_0^L$, we have $x_0(\ell) = c_1$ and $y_0(\ell) = c_2$). Then, with probability at least $1/2$ after $T_2 = 4(b+1)\ln b$ steps, we have $X_{T_2} = Y_{T_2}$.*

Proof. First, observe that with high probability after T_2 steps all of the leaves will be recolored at least once. Assuming all of the leaves are recolored at least once, if the root does not become a disagreement within these T_2 steps, then all of the leaves will be agreements. Therefore, we just need to show that the root will not change to a disagreement in T_2 steps with probability at least $3/5$. This is done by a coupling argument.

Let t_2 be the first time when the root becomes a disagreement, that is, $X_{t_2}(r) \neq Y_{t_2}(r)$. Note, since any disagreements on the leaves are colored c_1 in X_0 and c_2 in Y_0, either $X_{t_2}(r) = c_2$ and/or $Y_{t_2}(r) = c_1$. Therefore, we define the stopping times T_2^X and T_2^Y as follows:

$$T_2^X = \min\{t : X_t(r) = c_2\}, \quad T_2^Y = \min\{t : Y_t(r) = c_1\}.$$

We can assume without loss of generality that $X_0(r)$ (and hence $Y_0(r)$) does not equal either c_1 or c_2. Otherwise, by the hypothesis of the lemma, there are no disagreements in the leaves and hence $X_0 = Y_0$. Hence, our goal is to show that

$$\Pr\left[T_2^X \le T_2 \text{ or } T_2^Y \le T_2\right] < \frac{2}{5}.$$

And the main step is to show that

$$\Pr\left[T_2^X \le T_2\right] < \frac{1}{5}. \tag{9}$$

Let (S_t) be a random subset process on $V(G^*)$. Each time it picks a vertex v,

- If $v \neq r$, with probability $1/(k-1)$, $S_{t+1} = S_t \cup \{v\}$ and with probability $1 - 1/(k-1)$, $S_{t+1} = S_t \setminus \{v\}$;
- If $v = r$, if $S_t = \emptyset$ then $S_{t+1} = \{r\}$ otherwise $S_{t+1} = S_t$.

Let us define $T^S = \min_t\{t : r \in S_t\}$. We are going to couple (S_t) with (X_t) such that $\{v \in V(G^*) : X_t(v) = c_2\} \subseteq S_t$. This implies $T^S \leq T_2^X$. And if we can show that $\Pr\left[T^S \leq T_2\right] \leq 1/5$ then we have proved inequality (9).

The coupling (X_t, S_t) is defined as follows. We start with $S_0 = X_0^{-1}(c_2)$, the set of vertices of color c_2 in the initial coloring. Each time both processes picks the same vertex v to update.

- If $v = r$, X_t and S_t act independently at this time.
- If $v \neq r$ and $X_t(r) \neq c_2$, then X_t chooses a random color different from the root to recolor v and if that color is not c_2, $S_{t+1} = S_t \setminus \{v\}$ otherwise $S_{t+1} = S_t \cup \{v\}$.
- If $v \neq r$ and $X_t(r) = c_2$, then X_t chooses a random color different from c_2 to recolor v and if that color is not c_1, $S_{t+1} = S_t \setminus \{v\}$ otherwise $S_{t+1} = S_t \cup \{v\}$.

It is easy to see that this is a valid coupling. More importantly, it satisfies $X_t^{-1}(c_2) \subseteq S_t$.

Now we are going to show that $\Pr\left[T^S \leq T_2\right] < 1/5$ holds. It is not hard to show that with probability at least 0.9, the first time when the root is updated is later than $0.1b$ steps. We now condition on this event. The indicators of whether each leaf is in S_t or not during those $0.1b$ steps are negatively associated (c.f., Theorem 14 in [8]). Then by using the Chernoff bound with negative association among the random variables (c.f., Proposition 7 in [8]), it can be shown that with high probability at least $\geq 0.01b$ many different leaves are recolored before the first time we recolor the root. Thus, together with the proof of Proposition 9, we can claim that with probability at least 0.85, before the first t such that $r \in S_t$, at least $0.01b$ many leaves have been recolored and root will be recolored at most $20 \ln b$ times before T_2. Denote this event as $\mathcal{E}$. We have

$$\Pr\left[T^S \leq T_2\right] \leq \Pr\left[T^S \leq T_2 \mid \mathcal{E}\right] + \Pr\left[\bar{\mathcal{E}}\right] \leq \Pr\left[T^S \leq T_2 \mid \mathcal{E}\right] + 0.15.$$

In fact $\Pr\left[T^S \leq T_2 \mid \mathcal{E}\right]$ can be arbitrarily small when b grows, since at each time t we update the root in (S_t), we know that the probability of $S_{t-1} = \emptyset$ is at most $b^{-0.01(1+\epsilon)}$, and we know that the root updates at most $20 \ln b$ times.

In conclusion, we proved inequality (9) and hence the lemma. □

Finally, by combining Lemmas 10, 12 and 13 together, we can conclude that: when $\epsilon > 0$, with probability at least $1/(20(1+\epsilon)b^\epsilon \ln b)$ after $t = T_{\mathrm{w}} + T_1 + T_2 < T$ steps of the coupling, we have $X_t = Y_t$; when $\epsilon = 0$, from Lemmas 11, 12 and 13, we have that with probability at least $1/(20 \ln^3 b)$ after $t = T'_{\mathrm{w}} + T_1 + T_2 < T$ steps of the coupling, we have $X_t = Y_t$, which proves Lemma 7.

5 Upper Bound on Mixing Time for $C > 1$: Proof of Theorem 5

In this section we analyze the upper bound of the mixing time of the Glauber dynamics on the star graph G^* when $k = Cb/\ln b$ for $C > 1$. Here, let

$$\delta := C - 1,$$

and hence, $k = (1+\delta)b/\ln b$.

We will analyze the maximal one-step coupling using a weighted Hamming distance. The root r will have weight $w(r) = b^{\delta/2} > 1$ and the leaves will have weight $w(v) = 1$. For a set of vertices S, let $w(S) = \sum_{v \in S} w(v)$. Let D_t^r denote whether there is a disagreement at the root.

We want to show that the coupling decreases the distance in expectation. Hence, we say a pair of colorings (X_0, Y_0) are η-distance-decreasing if there exists a coupling $(X_0, Y_0) \to (X_1, Y_1)$ such that:

$$\mathrm{E}\left[w(D_1) \mid X_0, Y_0\right] < (1-\eta) w(D_0).$$

To simplify the analysis of the coupling, we will use the following theorem of Hayes and Vigoda [15] to utilize properties of the stationary distribution. The quantity $\mathrm{diam}(\Omega)$ is the diameter of Ω with respect to the Glauber dynamics. In our case, a trivial bound is $\mathrm{diam}(\Omega) \le 2b$.

Theorem 14. *[15, Theorem 1.2] Let $\eta > 0$. Suppose $S \subseteq \Omega$ such that every $(X_0, Y_0) \in S \times \Omega$ is η-distance-decreasing, and*

$$\pi(S) \ge 1 - \frac{\eta}{16 \ \mathit{diam}(\Omega)},$$

then the mixing time is

$$T_{\mathrm{mix}} \le 3\eta^{-1} \lceil \ln(32 \ \mathit{diam}(\Omega)) \rceil.$$

We use S as the set of colorings where the root has many available colors. Along the lines of the Dyer-Frieze [9] local uniformity results, we will prove the following statement about the available colors for the root r in a random coloring.

Lemma 15. *Let X be a random coloring of the star graph on b vertices. For every $\delta > 0$, there exists b_0, such that for all $b > b_0$ and $k = (1+\delta)b/\ln b$,*

$$\Pr\left[|A_X(r)| > b^{.9\delta}\right] > 1 - \exp(-b^{.99\delta}/10).$$

Hence, we let the set S be those colorings $X \in \Omega$ where $|A_X(r)| \ge b^{.9\delta}$.

5.1 Analyzing the coupling

We need to analyze $\mathrm{E}\left[w(D_1) \mid X_0, Y_0\right]$. Note, when a leaf v is recolored, if the root is a disagreement (i.e., $X_0(r) \neq Y_0(r)$) then with probability $1/(k-1)$ we have $X_1(v) \neq Y_1(v)$. Hence,

$$\begin{aligned}
&\mathrm{E}\left[w(D_1^L) \mid X_0, Y_0\right] \\
&= \sum_{v \in V \setminus \{r\}} w(v) \Big[\Pr\left[v \text{ is recolored}\right] \cdot \Pr\left[X_1(v) \neq Y_1(v) \mid v \text{ is recolored}, \ X_0, Y_0\right] \\
&\quad + \ (1 - \Pr\left[v \text{ is recolored}\right]) \mathbf{1}[X_0(v) \neq Y_0(v)] \Big] \\
&= \frac{b}{b+1} \frac{\mathbf{1}[r \in D_0]}{k-1} + \left(1 - \frac{1}{b+1}\right) w(D_0^L).
\end{aligned}$$

There is probability at most $|D_0^L| / \max\{|A_{X_0}(r)|, |A_{Y_0}(r)|\}$ that $X_1(r) \neq Y_1(r)$, when the root r is recolored. Hence, for $X_0 \in S$, we have:

$$\begin{aligned}
&\mathrm{E}\left[w(D_1^r) \mid X_0, Y_0\right] \\
&\le w(r) \frac{1}{b+1} \frac{|D_0^L|}{\max\{|A_{X_0}(r)|, |A_{Y_0}(r)|\}} + \left(1 - \frac{1}{b+1}\right) w(D_0^r) \\
&\le \frac{|D_0^L| b^{-\delta/3}}{b+1} + \left(1 - \frac{1}{b+1}\right) w(D_0^r).
\end{aligned}$$

Therefore, for $(X_0, Y_0) \in S \times \Omega$, we have:

$$\begin{aligned}
&\mathrm{E}\left[w(D_1) \mid X_0, Y_0\right] \\
&\leq \frac{1}{b+1}\left(\mathbf{1}[r \in D_0]\frac{b}{k-1} + b^{-\delta/3}|D_0^L|\right) + \left(1 - \frac{1}{b+1}\right) w(D_0) \\
&\leq w(D_0) + \frac{1}{b+1}\left(-w(D_0) + \mathbf{1}[r \in D_0]w(r)b^{-\delta/3} + b^{-\delta/3}|D_0^L|\right) \\
&\leq w(D_0) + \frac{1}{b}\left(-1 + b^{-\delta/4}\right) w(D_0).
\end{aligned}$$

Thus, they are η-distance-decreasing for $\eta = (1 - b^{-\delta/4})/b$.

Now applying Theorem 14, by Lemma 15 we have the necessary bound on $\pi(S)$, and thus conclude for b sufficiently large we have:

$$T_{\mathrm{mix}} \leq (6b \ln b)/(1 - b^{-\delta/4}) \leq 12b \ln b.$$

This completes the proof of Theorem 5, except for the proof of Lemma 15.

Proof of Lemma 15. Fix the color of the root to be c. Let σ be a random coloring conditional on the root receiving color c. We are going to prove that

$$\Pr\left[|A_\sigma(r)| \leq b^{0.9\delta} \mid \sigma(r) = c\right] < \exp(-b^{0.99\delta}/10).$$

For each color $i \in C \setminus \{c\}$, let Z_i be the indicator function that $c \in A_\sigma(r)$. $|A_\sigma(r)| = \sum_{i \in C} Z_i$. By Theorem 14 in [8], the Z_i's are negatively associated with each other once the root is fixed. Note that for b sufficiently large,

$$\begin{aligned}
\mathrm{E}\left[|A_\sigma(r)|\right] &\geq k \exp(-b/(k-1)) \\
&\geq b^{0.99\delta}.
\end{aligned}$$

Now applying the Chernoff bound, which hold for negatively associated random variables (c.f., Proposition 7 in [8]), we have:

$$\Pr\left[|A_\sigma(r)| \leq b^{0.9\delta} \mid \sigma(r) = c\right] < \exp(-b^{0.99\delta}/10).$$

□

6 Proof of the Lower Bounds Below the Threshold in Theorem 1

In this section we prove that when $C < 1$:

$$T_{\mathrm{relax}} = \Omega(n^{1/C - o(1)}). \tag{10}$$

In the remainder of this section, let $L(T)$ or simply L denote the leaves of T and the root is denoted by r. For a vertex v of T, let T_v denote the subtree of T rooted at v and T_v^* denote $T_v \setminus \{v\}$. For convenience, in this section, let $\epsilon := 1/C - 1$, and hence $k = b/(1+\epsilon)\ln b$.

In coloring $\sigma \in \Omega(T)$, we say a vertex v is *frozen* in σ if in the subtree T_v the coloring $\sigma(L(T_v))$ of the leaves of T_v forces the color for v. In other words, v is frozen in σ if: for all $\eta \in \Omega$ where $\eta(L(T_v)) = \sigma(L(T_v))$, we have $\eta(v) = \sigma(v)$. Note, by definition, the leaves are always frozen. Observe that for a vertex to be frozen, its frozen children must "block" all other color choices. This is formalized in the following observation as in [12].

Observation 16. *A vertex v where $h(v) > 0$ is frozen in coloring σ if and only if, for every color $c \neq \sigma(v)$, there is a child w of v where $\sigma(w) = c$ and w is frozen.*

Using this inductional way of defining a vertex being "frozen" in a coloring, we can further show the following lemma. It is a generalization of Lemma 8 in [12], which only applied to the case $\epsilon \geq 1$, i.e., $C \leq 1/2$.

Lemma 17. *For any $\epsilon \in (0, 1)$, in a random coloring of tree T, the probability that a vertex of T is not frozen is at most $b^{-\epsilon}$. For the leaves in T, by definition, they are always frozen.*

6.1 Upper Bound on the Conductance

Let $S_c = S_c(T)$ denote those colorings in $\Omega(T)$ where the root of T is frozen to color c. Let $S = \bigcup_{1 \leq c \leq k/2} S_c$. We will analyze the conductance of S to lower bound the mixing time.

To upper bound the conductance of S we need to bound the number of colorings $\sigma \in S$ which can leave S with one transition, and also the total number of transitions leaving S. To unfreeze the root, we need to recolor a leaf. Thus, we need to bound the number of colorings frozen at the root which can become unfrozen by one recoloring, and in that case, we need to bound the number of leaves which can be recolored to unfreeze the root. For a coloring σ, vertex v and color c, let $\sigma^{v \to c}$ denote the coloring obtained by recoloring v to c.

We capture the colorings on the "frontier" of S as follows. For tree T, coloring $\sigma \in \Omega(T)$, a vertex v and a leaf z of T_v, let $\mathcal{E}^{\sigma}_{v,z}$ denote the event that the coloring σ is frozen at the vertex v of T and there exists a color c where the coloring $\sigma^{z \to c}$ is not frozen at the vertex v. By definition, this event only depends on the configurations at the leaves of the subtree T_v. In particular, for the root of the tree, let $\mathcal{E}(\sigma, z) := \mathcal{E}^{\sigma}_{r,z}$ and $\mathbf{1}_{\sigma,z}$ be the indicator of it.

We can convert the above intuition into the following upper bound on conductance of S (similar to Lemma 10 in [12]):

Lemma 18.

$$\Phi_S \leq \frac{6}{n} \sum_{z \in L(T)} \Pr_{\sigma \in \Omega} [\mathcal{E}(\sigma, z)]$$

Now if we can prove that

$$\Pr_{\sigma \in \Omega} [\mathcal{E}(\sigma, z)] \leq b^{-(1+\epsilon-o(1))H}, \tag{11}$$

where $o(1)$ is an inverse polynomial of b when $\epsilon < 1$ and equals to zero when $\epsilon \geq 1$. This will be clarified later in the proof of Lemma 19. Then by plugging this back into the upper bound (11) we get

$$\Phi_S \leq \frac{6}{n} \cdot b^H \cdot b^{-(1+\epsilon-o(1))H} \leq 20 n^{-1-\epsilon+o(1)}.$$

Therefore, we can conclude that the conductance of this Glauber dynamics is $O(n^{-1-\epsilon+o(1)})$, and hence by (1) and (3), the mixing time and the relaxation time is $\Omega(n^{1/C-o(1)})$.

6.2 Proof of (11)

Let $\Omega^* = \{\sigma \in \Omega : \sigma(r) = c^*\}$ be the set of colors where the root is colored c^*. By symmetry, it is easy to see that $\Pr_{\sigma \in \Omega} [\mathcal{E}(\sigma, z)] = \Pr_{\sigma \in \Omega^*} [\mathcal{E}(\sigma, z)]$. Therefore, for the remainder of the proof we condition on the root being colored c^*. To simplify the notation, we denote $B := b^{-(1+\epsilon-o(1))}$. Let

$w_0 = r, w_1, \ldots, w_{H-1}, w_H = z$ denote the path in T from the root r down to the leaf z. We will show by induction that,

$$\begin{aligned}\Pr_{\sigma\in\Omega^*}\left[\mathcal{E}^\sigma_{r,z}\right] &\le B\,\Pr_{\sigma\in\Omega^*}\left[\mathcal{E}^\sigma_{w_1,z}\right] \le B^2\,\Pr_{\sigma\in\Omega^*}\left[\mathcal{E}^\sigma_{w_2,z}\right] \\ &\le \cdots \le B^H\,\Pr_{\sigma\in\Omega^*}\left[\mathcal{E}^\sigma_{w_H,z}\right] = B^H.\end{aligned}$$

For the event $\mathcal{E}(\sigma, z)$ to occur we need that along the path from the leaf z to the root r, unfreezing each of these vertices will "free" a color for their parent. More precisely, for σ to be in $\mathcal{E}(\sigma, z)$, w_1 has to be frozen because the color of z only affects the root through w_1, and if w_1 is not frozen then it cannot affect the root becoming unfrozen. In order for the root to become unfrozen by changing the color of the leaf z, it must also occur that w_1 becomes unfrozen at the same time, hence $\sigma \in \mathcal{E}^\sigma_{w_1,z}$, i.e., $\mathcal{E}(\sigma, z) \subseteq \mathcal{E}^\sigma_{w_1,z}$ and more generally, $\mathcal{E}^\sigma_{w_i,z} \subseteq \mathcal{E}^\sigma_{w_{i+1},z}$.
For each $1 \le i \le H$, let $\mathcal{A}^\sigma_{w_i,z}$ denote the event that no sibling y of w_i satisfies both of the following: $\sigma(y) = \sigma(w_i)$ and σ is frozen at y. By the siblings of w_i, as usual we mean the children (other than w_i) of w_{i-1}. The event $\mathcal{E}^\sigma_{w_i,z}$ implies the fact that w_{i+1} is the only child that causes w_i simultaneously being frozen and being blocked from using color $\sigma(w_{i+1})$, which means $\mathcal{E}^\sigma_{w_i,z} \subseteq \mathcal{A}^\sigma_{w_{i+1},z}$. We will show the following lemma for bounding the probability of $\mathcal{A}^\sigma_{w_1,z}$.

Lemma 19. *Let $\mathcal{C}^* = \mathcal{C} - c^*$. For a fixed color $c_1 \in \mathcal{C}^*$,*

$$\Pr_{\sigma\in\Omega^*}\left[\mathcal{A}^\sigma_{w_1,z} \mid \sigma(w_1) = c_1\right] \le B = b^{-(1+\epsilon-o(1))}.$$

Observe that the events $\mathcal{A}^\sigma_{1,z}$ and $\mathcal{E}^\sigma_{w_1,z}$ are independent, conditioned on the fixed colors of the root and w_1, because they depend on the configurations of different parts of leaves. Then we have that for each $c_1 \in \mathcal{C}^*$,

$$\begin{aligned}&\Pr_{\sigma\in\Omega^*}\left[(\sigma(w_1) = c_1)\bigcap \mathcal{E}^\sigma_{w_1,z}\bigcap \mathcal{A}^\sigma_{1,z}\right] \\ &= \Pr_{\sigma\in\Omega^*}\left[\mathcal{E}^\sigma_{w_1,z} \mid \sigma(w_1) = c_1\right] \cdot \Pr_{\sigma\in\Omega^*}\left[\mathcal{A}^\sigma_{1,z} \mid \sigma(w_1) = c_1\right] \cdot \frac{1}{k-1} \\ &\le \frac{B^{H-1}\cdot B}{k-1}, \qquad (12)\end{aligned}$$

where the last inequality is by the inductive hypothesis applied on the complete tree T_{w_1} of height $H-1$ and Lemma 19.
Finally, by the fact that $\mathcal{E}^\sigma_{r,z} \subseteq \mathcal{E}^\sigma_{w_1,z}\bigcap \mathcal{A}^\sigma_{1,z}$ and (12) above, we have

$$\begin{aligned}\Pr_{\sigma\in\Omega^*}\left[\mathcal{E}^\sigma_{r,z}\right] &\le \Pr_{\sigma\in\Omega^*}\left[\mathcal{E}^\sigma_{w_1,z}\bigcap \mathcal{A}^\sigma_{w_1,z}\right] \\ &= \sum_{c_1\in\mathcal{C}^*} \Pr_{\sigma\in\Omega^*}\left[(\sigma(w_1) = c_1)\bigcap \mathcal{E}^\sigma_{w_1,z}\bigcap \mathcal{A}^\sigma_{w_1,z}\right] \\ &\le B^H.\end{aligned}$$

This completes the proof of (11). To complete the proof of the lower bounds when $C < 1$ in Theorem 1 we need to prove Lemmas 17, 18, and 19.

6.3 Proofs of Lemmas

Proof of Lemma 17. The proof is very similar to the proof of Lemma 8 in [12]. We include it here for completeness.

Let U_ℓ be the probability that a vertex at the height ℓ is not frozen. We are going to prove that $U_\ell < b^{-\epsilon}$ by induction.

First of all, by definition, $U_0 = 0$ since they are leaves. Let v be a vertex at height $\ell > 0$. Since the probability that the color of v equals c is independent from the probability that v is frozen, therefore we can just fixed the color of v to some $c^* \in \mathcal{C}$, and hence

$$U_\ell = \Pr\left[v \text{ is not frozen in } \sigma \mid \sigma(v) = c^*\right].$$

Let w be a child of v. Again by the same argument using the independency, the probability that w is frozen to color c equals $\frac{1-U_{\ell-1}}{k-1}$. Thus, the probability that all the children of v are either not frozen or not colored by using c is $(1-(1-U_{\ell-1})/(k-1))^b$.

By the union bound and induction, U_ℓ is bounded by:

$$(k-1)\Big(1-\frac{1-U_{\ell-1}}{k-1}\Big)^b \le (k-1)\exp\Big(-\frac{b(1-b^{-\epsilon})}{k-1}\Big) \le b^{-\epsilon},$$

where the last inequality holds for large b. □

Proof of Lemma 18. Let $F := \bigcup_{c\in\mathcal{C}} S_c$ be the set of colorings that freeze the root. As we discussed before, by symmetry, $\pi(S_{c_1}) = \pi(S_{c_2})$ for $c_1, c_2 \in \mathcal{C}$. Then $\pi(\bar{S}) \ge 1/2$. Also, by Lemma 17, we know that $\pi(F) \ge 1 - b^{-\epsilon}$. Therefore for any $\epsilon > 0$, there exists a b_0 such that for all $b > b_0$, $\pi(S)\pi(\bar{S}) \ge 1/6$. From the definition of Φ_S, we know that

$$\Phi_S \le 6\Big(\sum_{\sigma\in S}\sum_{\eta\in\bar{S}} \pi(\sigma)P(\sigma,\eta)\Big).$$

Notice that, for any $\sigma \in S_{c_1}$, $\eta \in S_{c_2}$ and $c_1 \neq c_2$, we have $P(\sigma,\eta) = 0$, because it is impossible to change the color of the frozen root by just one move. Further, in order to unfreeze the root in one step, the Glauber dynamics has to first recolor a leaf and change the color of the leaf so as to unfreeze the root. That is η can only be $\sigma^{z\to c}$ for some $z \in L(T)$ and $c \in \mathcal{C}^*$, where $\mathcal{C}^* = \mathcal{C} - \{$ the color of the parent of z in $\sigma\}$. Therefore,

$$\begin{aligned}\Phi_S &\le 6\sum_{\sigma\in F}\sum_{\eta\in\bar{F}} \pi(\sigma)P(\sigma,\eta)\\ &\le 6\sum_{\sigma\in F}\sum_{z\in L(T)} \Big(\mathbf{1}_{\sigma,z}\pi(\sigma)P(\sigma,\eta)\Big), \qquad (13)\end{aligned}$$

where $\mathbf{1}_{\sigma,z}$ is the indicator for the event that the root in coloring σ is frozen and there exists a color c where the root in the coloring $\sigma^{z\to c}$ is not frozen.

By the definition of the Glauber dynamics, we know that $\pi(\sigma) = 1/|\Omega|$ and $P(\sigma,\eta) = 1/(n(k-1))$ for the case that the change of color happens at a leaf. Therefore, from (13), we have:

$$\Phi_S \le \frac{6}{n}\sum_{\sigma\in\Omega}\sum_{z\in L(T)}\frac{\mathbf{1}_{\sigma,z}}{|\Omega|} = \frac{6}{n}\sum_{z\in L(T)}\sum_{\sigma\in\Omega(T)}\frac{\mathbf{1}_{\sigma,z}}{|\Omega(T)|}.$$

□

Proof of Lemma 19. When $\epsilon < 1$, the probability that all the siblings of w_1 are either not frozen or not colored with c_1 is upper bounded by

$$\begin{aligned}\Big(1-\frac{1-U_{H-1}}{k-1}\Big)^{b-1} &\le \exp\Big(-\frac{(b-1)(1-b^{-\epsilon})}{k-1}\Big)\\ &\le b^{-(1+\epsilon)(1-b^{-\epsilon})}.\end{aligned}$$

Now we can see that $o(1)$ is actually $(1+\epsilon)/b^{\epsilon}$ when $\epsilon < 1$.

Note that, when $\epsilon \geq 1$, by the same way it is easy to see that

$$\Big(1 - \frac{1 - U_{H-1}}{k-1}\Big)^{b-1} \leq b^{-(1+\epsilon)}.$$

□

7 A Simple Generalization to $k = o(b/\ln b)$: Proof of Theorem 2

In all of the previous sections, we assumed $k = Cb/\ln b$ where C is constant. But we are also interested in the case when k is constant, say a hundred colors, and what the mixing time of the Glauber dynamics will be in this case. Let $\alpha = \alpha(k,b) := b/(k \ln b)$. We would also like to see how to generalize the upper bound and lower bound analysis assuming α is any function growing with b, that is when k is $o(b/\ln b)$. Actually, all of our proofs will be the same and we just need to slightly modify the statements.

For the upper bound, we change Lemma 7 and Lemma 12 into the following ones.

Lemma 20. *Let $T = 20b \ln b$. There exists b_0, for all $(x_0, y_0) \in \Omega \times \Omega$, all $\alpha(k,b) \geq 2$, and all $b > b_0$ the following holds:*

$$\Pr\left[X_T = Y_T \mid X_0 = x_0, Y_0 = y_0\right] \geq 1/(20\alpha(k,b) b^{\alpha(k,b)} \ln b).$$

Lemma 21. *For any pair of initial states (x_0, y_0) where all of the leaf disagreements are of the same type, then*

$$\Pr\left[T_1^{XY} < 4b \ln b \mid (X_0, Y_0) = (x_0, y_0)\right] \geq 1/(4\alpha(k,b) b^{\alpha(k,b)-1} \ln b).$$

Then by the same argument as in Section 4, we are able to show that the relaxation time of the Glauber dynamics on G^* is upper bounded by $O(\alpha b^{\alpha} \ln b)$. Thus, the mixing time of the Glauber dynamics on the complete tree is bounded by

$$T_{\text{mix}} = O\big(n^{\alpha + (\ln\alpha + 2\ln\ln b + 20)/\ln b} \ln n\big),$$

and the relaxation time is bounded by

$$T_{\text{relax}} = O\big(n^{\alpha + (\ln\alpha + 2\ln\ln b + 20)/\ln b}\big).$$

For the lower bound, we replace Lemma 17 and Lemma 19 into the following lemmas.

Lemma 22. *In a random coloring of the tree T, the probability that a vertex of T is not frozen is at most b^{-1}.*

Lemma 23.

$$\Pr_{\sigma \in \Omega^*}\left[\mathcal{A}^{\sigma}_{w_1,z} \mid \sigma(w_1) = c_1\right] \leq b^{-\alpha(k,b)}.$$

Then, by exactly the same way as in Section 6, we can show that the mixing time and the relaxation time of the Glauber dynamics on the complete tree T when $\alpha \geq 2$ is lower bounded by $\Omega(n^{\alpha}) = \Omega(n^{b/(k \ln b)})$.

8 Bounding the Log-Sobolev Constant: Proof of Theorem 6

In this section we will analyze the log-Sobolev constant c_{sob} of the heat-bath Glauber dynamics on the complete tree by comparing it with the spectral gap c_{gap}. For completeness, we prove Theorem 6, which is an improvement over the proof of Theorem 5.7 in Martinelli, Sinclar and Weitz [24]. In their paper, they proved it for the case of the Ising model on the complete tree with a fixed boundary condition, although they observed that it holds more generally. For convenience, we will use the same notation for the complete tree and its vertices, that is, T_ℓ stands for both the complete tree of height ℓ and its vertices $V(T_\ell)$.

Let $B \subseteq A \subseteq T$ be two subsets of the vertices on tree T. Let $\eta \in \Omega$ be a configuration. Let $E_A^\eta(f)$ be the expectation of f under a prefixed distribution μ in the region A with boundary condition η. That is

$$E_A^\eta(f) = \sum_\sigma \frac{\mu(\sigma)}{Z} f(\sigma)$$

where σ ranges over the configurations that are the same as η outside A (denoted as $\sigma \sim_A \eta$) and Z is the normalizing factor. The quantities Var_A^η and Ent_A^η are defined similarly. If we drop η then $E_A(f), \mathrm{Var}_A(f), \mathrm{Ent}_A(f)$ become functions from Ω to R. The following are standard facts concerning variance and entropy; the first being the chain rule, and the second follows from the so-called tensoring property over a product distribution – see e.g., Proposition 5.6 of [19]. In the following, we will use the fact that the distribution on configurations over the tree with the root removed, has a product form over the subtrees rooted at the children of the root, to satisfy the hypothesis for the tensoring property.

Proposition 24.

$$Var_A^\eta(f) = E_A^\eta(Var_B(f)) + Var_A^\eta(E_B(f)).$$
$$Ent_A^\eta(f) = E_A^\eta(Ent_B(f)) + Ent_A^\eta(E_B(f)).$$

Proposition 25. *Let $A = \bigcup A_i$ where A_i are disjoint, and suppose that conditioning on the boundary being η, the probability of A_i's being in any configuration for different i's is completely independent. Then*

$$Var_A^\eta(f) \le \sum_i E_A^\eta(Var_{A_i}(f)),$$

and

$$Ent_A^\eta(f) \le \sum_i E_A^\eta(Ent_{A_i}(f)).$$

Lemma 26. *Let $c_{sob}(\ell)$ be the log-Sobolev constant of the heat-bath Glauber dynamics on the complete tree of height $\ell > 0$ with the root being attached to an external vertex with a fixed color, then*

$$c_{sob}(\ell)^{-1} \le c_{sob}(\ell-1)^{-1} + \alpha \cdot c_{gap}(\ell)^{-1},$$

where $\alpha = \frac{\log(k-2)}{1-2/(k-1)} = c_{sob}(0)^{-1}$.

Proof. Let f be any non-negative function. Let I be the set of vertices in the complete tree T_ℓ without the root, i.e., $I = T_\ell \setminus \{\text{root}\}$. Let us first use Proposition 24 to analyze the $\mathrm{Ent}(f)$:

$$\mathrm{Ent}(f) = E(\mathrm{Ent}_I(f)) + \mathrm{Ent}(E_I(f))$$

We will bound $E(\mathrm{Ent}_I(f))$ and $\mathrm{Ent}(E_I(f))$ separately. For $E(\mathrm{Ent}_I(f))$, by Proposition 25, it can be upper bounded as

$$E(\mathrm{Ent}_I(f)) \leq \sum_v E(\mathrm{Ent}_{T_v}(f)), \tag{14}$$

where v ranges over all the children of the root of T_ℓ and T_v denotes the subtree of T_ℓ rooted at the vertex v. Let $\eta \in \Omega(T_\ell)$, then for a specific $\mathrm{Ent}^\eta_{T_v}(f)$, we then have

$$\mathrm{Ent}^\eta_{T_v}(f) \leq c_{sob}(\ell-1)^{-1}\mathcal{D}_{T_v}(\sqrt{f}), \tag{15}$$

where $\mathcal{D}_{T_v}(\sqrt{f})$ is the corresponding Dirichlet form for the dynamics on the subtree T_v. For the heat-bath Glauber dynamics, since $P(\sigma,\tau) \neq 0$ only if they differ at a single vertex, we can further derive that

$$\begin{aligned} \mathcal{D}_{T_v}(f) &= \frac{1}{2}\sum_{\sigma,\tau}(f(\sigma)-f(\tau))^2\mu(\sigma)P(\sigma,\tau) \\ &= \frac{1}{2}\sum_{x\in T_v} E^\eta_{T_v}(\mathrm{Var}_{\{x\}}(f)) \end{aligned} \tag{16}$$

where $\mu(\sigma)$ is the marginal distribution with respect to η.

Then, from (14), (15) and above we have

$$\begin{aligned} E(\mathrm{Ent}_I(f)) &\leq \sum_v E(\mathrm{Ent}_{T_v}(f)) \quad \text{(by (14))} \\ &\leq \sum_v c_{sob}(\ell-1)^{-1}E(\mathcal{D}_{T_v}(\sqrt{f})) \quad \text{(by (15))} \\ &= \sum_v c_{sob}(\ell-1)^{-1}E\big(\sum_{x\in T_v} E^\eta_{T_v}[\mathrm{Var}_{\{x\}}(f)]\big) \quad \text{(by (16))} \\ &= c_{sob}(\ell-1)^{-1}\sum_{x\in I} E(\mathrm{Var}_{\{x\}}(f)) \\ &\leq c_{sob}(\ell-1)^{-1}\mathcal{D}(\sqrt{f}) \quad \text{(by applying (16) again)} \end{aligned}$$

For $\mathrm{Ent}(E_I(f))$, $E_I(f)$ can be viewed as a function from $\{1,2,...,k-1\}$ to R since those $k-1$ values can represent the colors of the root (boundary). Therefore $\mathrm{Ent}(E_I(f))$ is the entropy of the random variable $E_I(f)$ taking $k-1$ values uniformly at random. It is well-known (see e.g. Appendix of [6]) that $\frac{\log(k-2)}{1-2/(k-1)}$ is the inverse of the log-Sobolev constant of the random walk $\mathcal{R}$ on the complete graph K_{k-1}, which jumps to stationarity in one step. Thus, letting $\alpha = \frac{\log(k-2)}{1-2/(k-1)}$, we may upper bound $\mathrm{Ent}(E_I(f))$ as follows.

$$\begin{aligned} \mathrm{Ent}(E_I(f)) &\leq \alpha\mathcal{D}_{\mathcal{R}}(\sqrt{E_I(f)}) \quad \text{(by the log-Sobolev Inequality)} \\ &= \alpha\mathrm{Var}_{\mathcal{R}}(\sqrt{E_I(f)}) \quad \text{(for the complete graph } P(x,y)=\pi_{\mathcal{R}}(y)) \\ &= \alpha\mathrm{Var}_T(\sqrt{E_I(f)}) \\ &\leq \alpha\big(E[E_I(f)] - E^2(\sqrt{E_I(f)})\big) \quad \text{(by the definition of the variance)} \\ &\leq \alpha E(\sqrt{f})^2 - E^2(\sqrt{f}) \quad \text{(by the concavity of } \sqrt{x}) \\ &\leq \alpha c_{gap}(\ell)^{-1}\mathcal{D}(\sqrt{f}) \quad \text{(by the definition of the spectral gap)} \end{aligned}$$

Putting everything together, we proved

$$\begin{aligned}\mathrm{Ent}(f) &= E(\mathrm{Ent}_I(f)) + \mathrm{Ent}(E_I(f)) \\ &\leq c_{sob}(\ell-1)^{-1}\mathcal{D}(\sqrt{f}) + \alpha c_{gap}(\ell)^{-1}\mathcal{D}(\sqrt{f}),\end{aligned}$$

then by the definition of c_{sob}, we get

$$c_{sob}(\ell)^{-1} \leq c_{sob}(\ell-1)^{-1} + \alpha c_{gap}(\ell)^{-1}\,.$$

□

Lemma 27. *Let $c_{sob}(\ell)$ be the spectral gap of the heat-bath Glauber dynamics on the complete tree of height $\ell > 0$ with the root being attached to an external vertex with a fixed color, then for $\ell > 0$, we have $c_{gap}(\ell) \leq c_{gap}(\ell-1)/b$.*

Proof. Let $\mathcal{D}_\ell(f)$ and $\mathrm{Var}_\ell(f)$ be the Dirichlet form and the variance of function $f : \Omega(T_\ell) \to \mathcal{R}$ for the Glauber dynamics on the complete tree of height ℓ with the root attached to an external vertex with a fixed color. Let P_ℓ denote the probability transition of the dynamics, and let π_ℓ denote its unique stationary distribution.

Let g be the eigenfunction such that $c_{gap}(\ell-1) = \mathcal{D}_{\ell-1}(g)/\mathrm{Var}_{\ell-1}(g)$. Now we are going to construct a function $f : \Omega(T_\ell) \to \mathcal{R}$, such that $\mathcal{D}_\ell(f) \leq \mathcal{D}_{\ell-1}(g)$ and $\mathrm{Var}_\ell(f) = \mathrm{Var}_{\ell-1}(g)$. Then, since

$$c_{gap}(\ell) \leq \frac{\mathcal{D}_\ell(f)}{\mathrm{Var}_\ell(f)} \leq \frac{\mathcal{D}_{\ell-1}(g)}{\mathrm{Var}_{\ell-1}(g)} = c_{gap}(\ell-1),$$

we prove the lemma.

Let $A \subseteq T_\ell$ be the set of non-leaf vertices of T_ℓ, i.e., $A = T_\ell \setminus L(T_\ell)$, where $L(T_\ell)$ is the set of leaves in the tree T_ℓ. There is a natural correspondence between vertices in A and in $T_{\ell-1}$. The function f is then defined as: for $\sigma \in \Omega(T_\ell)$ and $\sigma' \in \Omega(T_{\ell-1})$, $f(\sigma) = g(\sigma')$ if the configuration σ agrees with σ' on the subset A.

It is straightforward to show that $\mathrm{Var}_\ell(f) = \mathrm{Var}_{\ell-1}(g)$. We will show $\mathcal{D}_\ell(f) \leq \mathcal{D}_{\ell-1}(g)/b$. By definition,

$$\mathcal{D}_\ell(f) = \sum_{\sigma,\eta\in\Omega(T_\ell)} \pi_\ell(\sigma)P_\ell(\sigma,\eta)(f(\sigma)-f(\eta))^2.$$

For a subset of vertices $S \subset T_\ell$, let

$$\Omega(S) = \{\sigma' \in [k]^S : \text{ there exists } \sigma \in \Omega(T_\ell) \text{ where } \sigma(A) = \sigma'\}.$$

Let $\sigma', \eta' \in \Omega(A)$ be colorings of the internal vertices. Let $\phi, \psi \in \Omega(L(T_\ell))$ be colorings of the leaves. Finally, let $\circ$ be the concatenation operator, thus $\sigma' \circ \psi = \sigma \in \Omega(T_\ell)$ where $\sigma(A) = \sigma'$ and $\sigma(L(T_\ell)) = \phi$. Then we can rewrite the Dirichlet form as:

$$\mathcal{D}_\ell(f) = \sum_{\sigma',\eta'\in\Omega(A)} \sum_{\phi,\psi\in\Omega(L(T_\ell))} \pi_\ell(\sigma'\circ\phi)P_\ell(\sigma'\circ\phi,\eta'\circ\psi)(f(\sigma'\circ\phi)-f(\eta'\circ\psi))^2.$$

According to the definition of the Glauber dynamics, for configurations $\sigma, \eta \in \Omega(T_\ell)$ which differ at more than one vertex, we have $P_\ell(\sigma,\eta) = 0$. Let $\oplus$ denote the symmetric difference. Now we

can rewrite the Dirichlet form as:

$$\begin{aligned}
&\mathcal{D}_\ell(f)\\
&= \sum_{v\in A} \sum_{\substack{\sigma',\eta'\in\Omega(A):\\ \sigma'\oplus\eta'=\{v\}}} \sum_{\phi\in\Omega(L(T_\ell))} \left[(f(\sigma'\circ\phi)-f(\eta'\circ\phi))^2\pi_\ell(\sigma'\circ\phi)P_\ell(\sigma'\circ\phi,\eta'\circ\phi)\right]\\
&+ \sum_{v\in L(T_\ell)} \sum_{\sigma'\in\Omega(A)} \sum_{\substack{\phi,\psi\in\Omega(L(T_\ell)):\\ \phi\oplus\psi=\{v\}}} \left[(f(\sigma'\circ\phi)-f(\sigma'\circ\psi))^2\pi_\ell(\sigma'\circ\phi)P_\ell(\sigma'\circ\phi,\sigma'\circ\psi)\right]\\
&= \sum_{v\in A} \sum_{\substack{\sigma',\eta'\in\Omega(A):\\ \sigma'\oplus\eta'=\{v\}}} \left[(g(\sigma')-g(\eta'))^2 \sum_{\phi\in\Omega(L(T_\ell))} \pi_\ell(\sigma'\circ\phi)P_\ell(\sigma'\circ\phi,\eta'\circ\phi)\right],
\end{aligned}$$

since $g(\sigma'\circ\phi)=g(\sigma'\circ\psi)=f(\sigma')$.

Thus we only need to consider when the sole disagreement is at an internal vertex. We can further decompose based on whether the disagreement is an internal vertex of the tree $T_{\ell-1}$, which we denote as I, or a leaf of $T_{\ell-1}$.

For $v\in L(T_{\ell-1})$, the goal is to bound the sum $\sum_\phi \pi_\ell(\sigma'\circ\phi)P_\ell(\sigma'\circ\phi,\eta'\circ\phi)$ by $\pi_{\ell-1}(\sigma')/(|T_{\ell-1}|(k-1)b)$, i.e., $\pi_{\ell-1}(\sigma')P_{\ell-1}(\sigma',\eta')/b$. We have the following observation: Fix the vertex v, for each color c such that $\sigma'\oplus\eta'=\{v\}$ and $\eta'(v)=c$, the quantity $Q(c):=\sum_\phi \pi_\ell(\sigma'\circ\phi)P_\ell(\sigma'\circ\phi,\eta'\circ\phi)$ are the same, i.e., $Q(c)=Q(c')$ for any two colors $c\neq c'$ because of the symmetry. Therefore, in order to bound $Q(c)$, it is easier to bound $\sum_{c\neq\sigma'(v)}Q(c)$ by $\pi_{\ell-1}(\sigma')/(|T_{\ell-1}|b)$. Then, by taking the average over $k-1$ colors, we are done. It is a straightforward calculation to upper bound the sum of $Q(c)$:

$$\begin{aligned}
\sum_{c\neq\sigma'(v)} Q(c) &= \pi_{\ell-1}(\sigma')\sum_\phi \frac{\pi_\ell(\sigma'\circ\phi)}{\pi_{\ell-1}(\sigma')} \sum_{c\neq\sigma'(v)} \mathbf{1}\{c\in A_{\sigma'\circ\phi}(v)\}P_\ell(\sigma'\circ\phi,\eta'\circ\phi)\\
&= \pi_{\ell-1}(\sigma')\sum_\phi \frac{\pi_\ell(\sigma'\circ\phi)}{\pi_{\ell-1}(\sigma')}\frac{|A_{\sigma'\circ\phi}(v)|-1}{|T_\ell||A_{\sigma'\circ\phi}(v)|}\\
&\le \pi_{\ell-1}(\sigma')\frac{1}{|T_{\ell-1}|b}, \qquad (17)
\end{aligned}$$

where by definition, $A_{\sigma'\circ\phi}(v)$ is the set of available colors for vertex v in the configuration $\sigma'\circ\phi$.

Recall, I denotes the internal vertices of $T_{\ell-1}$, i.e., $I=V(T_{\ell-1})\setminus L(T_{\ell-1})$. Similarly, for $v\in I$ we have:

$$\begin{aligned}
&\sum_{\sigma',\eta'\in\Omega(A):\sigma'\oplus\eta'=\{v\}} [(g(\sigma')-g(\eta'))^2 \sum_{\phi\in\Omega(L(T_\ell))} \pi_\ell(\sigma'\circ\phi)P_\ell(\sigma'\circ\phi,\eta'\circ\phi)]\\
&\qquad = \sum_{\substack{\sigma',\eta'\in\Omega(T_{\ell-1}):\\ \sigma'\oplus\eta'=\{v\}}} (g(\sigma')-g(\eta'))^2\pi_{\ell-1}(\sigma')P_{\ell-1}(\sigma',\eta')/b. \qquad (18)
\end{aligned}$$

Combining (17) and (18), and summing over $v\in T_{\ell-1}$ we have shown that $\mathcal{D}_\ell(f)\le\mathcal{D}_{\ell-1}(g)/b$, which implies the lemma. □

Proof of Theorem 6. Now we apply Lemma 26 inductively, and we get

$$c_{sob}^{-1}=c_{sob}^{-1}(H)\le\alpha\big(1+c_{gap}^{-1}(1)+\cdots+c_{gap}^{-1}(\lfloor\log_b n\rfloor)\big).$$

Then by applying Lemma 27 on the spectral gaps, we can conclude that

$$c_{sob}^{-1} \leq b\alpha c_{gap}^{-1}(H) \leq c_{gap}^{-1} \cdot 2b \log k.$$

□

9 Conclusions

Recently, Restrepo et al [28] studied the analogous problem for the hard-core model which is defined on independent sets weighted by an activity $\lambda > 0$. In contrast to the picture we have shown for colorings, Martinelli et al [25] had shown that on the complete tree with branching factor b, the Glauber dynamics has $O(n \log n)$ mixing time for all λ. Thus, there is no slow-down at the reconstruction threshold. However, Restrepo et al [28] show that there is a boundary condition for the complete tree so that the Glauber dynamics has a slow-down that appears to coincide with the reconstruction threshold.

References

[1] D. Achlioptas and A. Coja-Oghlan. Algorithmic barriers from phase transitions. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science* (FOCS), 793-802, 2008.

[2] D. Aldous. Random walks on finite groups and rapidly mixing Markov chains. *Séminaire de Probabilités XVII,* Springer Lecture Notes in Mathematics 986, 243-297, 1983.

[3] N. Berger, C. Kenyon, E. Mossel, and Y. Peres. Glauber dynamics on trees and hyperbolic graphs. *Probability Theory and Related Fields*, 131(3):311–340, 2005.

[4] N. Bhatnagar, J. Vera, E. Vigoda and D. Weitz. Reconstruction for Colorings on Trees. *SIAM J. Discrete Math.*, 25(2):809-826, 2011.

[5] J. Chayes, F. Martinelli, M. Molloy, and P. Tetali. Workshop report: Phase transitions, hard combinatorial optimization problems and message passing algorithms. Available at: http://www.birs.ca/workshops/2008/08w5109/report08w5109.pdf

[6] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for Markov chains. *Annals of Applied Probability*, 6(3):695–750, 1996.

[7] J. Ding, E. Lubetzky, and Y. Peres. Mixing time of critical Ising model on trees is polynomial in the height. *Communications in Mathematical Physics*, 295(1):161-207, 2010.

[8] D. Dubhashi and D. Ranjan. Balls and bins: A study in negative dependence. *Random Struct. Algorithms*, 13(2):99–124, 1998.

[9] M. Dyer and A. Frieze. Randomly coloring graphs with lower bounds on girth and maximum degree. *Random Struct. Algorithms*, 23(2):167–179, 2003.

[10] M. Dyer, L. Goldberg, M. Jerrum, and R. Martin. Markov chain comparison. *Probability Surveys*, 3:89–111, 2006.

[11] A. Frieze and E. Vigoda. A survey on the use of Markov chains to randomly sample colourings. In *Combinatorics, Complexity and Chance*, G. Grimmett and C. McDiarmid eds., Oxford University Press, 2007.

[12] L. Goldberg, M. Jerrum, and M. Karpinski. The mixing time of Glauber dynamics for colouring regular trees. *Random Struct. Algorithms*, 36(4), 464-476, 2010.

[13] T. Hayes and A. Sinclair. A general lower bound for mixing of single-site dynamics on graphs. *Annals of Applied Probability*, 17(3):931–952, 2007.

[14] T. Hayes, J. Vera, and E. Vigoda. Randomly coloring planar graphs with fewer colors than the maximum degree. In *Proceeding of the 39th ACM Symposium on Theory of Computing* (STOC), 450–458, 2007.

[15] T. Hayes and E. Vigoda. Coupling with the stationary distribution and improved sampling for colorings and independent sets. *Annals of Applied Probability*, 6(4):1297–1318, 2006.

[16] M. Jerrum. A very simple algorithm for estimating the number of k-colorings of a low-degree graph. *Random Struct. Algorithms*, 7(2):157–166, 1995.

[17] J. Jonasson. Uniqueness of uniform random colorings of regular trees. *Statistics and Probability Letters*, 57:243–248, 2002.

[18] G. Lawler and A. Sokal. Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality. *Transactions of the American Mathematical Society*, 309:557–580, 1988.

[19] M. Ledoux. *The Concentration of Measure Phenomenon*. Mathematical Surveys and Monographs (Vol. 89), American Mathematical Society, 2001.

[20] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2009.

[21] B. Lucier and M. Molloy. The Glauber dynamics for colourings of bounded degree trees. *SIAM J. Discrete Math.*, 25(2):827-853, 2011.

[22] B. Lucier, M. Molloy, and Y. Peres. The Glauber dynamics for colourings of bounded degree trees. In *Proceedings of the 13th Intl. Workshop on Randomization and Computation* (RANDOM), 631–645, 2009.

[23] F. Martinelli. *Lectures on Glauber dynamics for discrete spin models*. Springer Lecture Notes in Mathematics (Vol. 1717), 2000.

[24] F. Martinelli, A. Sinclair, and D. Weitz. The Ising model on trees: Boundary conditions and mixing time. *Communications in Mathematical Physics*, 250(2):301–334, 2004.

[25] F. Martinelli, A. Sinclair, and D. Weitz. Fast mixing for independent sets, colorings, and other models on trees. *Random Struct. Algorithms*, 31(2):134–172, 2007.

[26] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

[27] E. Mossel and A. Sly. Gibbs rapidly samples colorings of $G(n, d/n)$. *Probability Theory and Related Fields*, 147:37–69, 2010.

[28] R. Restrepo, D. Stefankovic, J. C. Vera, E. Vigoda, L. Yang. Phase Transition for Glauber Dynamics for Independent Sets on Regular Trees. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 945-956, 2011.

[29] A. Sinclair and M. Jerrum. Approximate counting, uniform generation and rapidly mixing Markov chains. *Information and Computation*, 82:93–133, 1989.

[30] A. Sly. Reconstruction of random colourings. *Communications in Mathematical Physics*, 288(3):943–961, 2009.

[31] E. Vigoda. Improved bounds for sampling colorings. *Journal of Mathematical Physics*, 41:155–1569, 2000.