The Annals of Applied Probability
2017, Vol. 27, No. 5, 2646–2674
DOI: 10.1214/16-AAP1253

THE GLAUBER DYNAMICS OF COLORINGS ON TREES IS RAPIDLY MIXING THROUGHOUT THE NONRECONSTRUCTION REGIME

BY ALLAN SLY[1,*,†] AND YUMENG ZHANG[†]

Australian National University and University of California, Berkeley†*

The mixing time of the Glauber dynamics for spin systems on trees is closely related to the reconstruction problem. Martinelli, Sinclair and Weitz established this correspondence for a class of spin systems with soft constraints bounding the log-Sobolev constant by a comparison with the block dynamics [*Comm. Math. Phys.* **250** (2004) 301–334; *Random Structures Algorithms* **31** (2007) 134–172]. However, when there are hard constraints, the dynamics inside blocks may be reducible.

We introduce a variant of the block dynamics extending these results to a wide class of spin systems with hard constraints. This applies to essentially any spin system that has nonreconstruction provided that on average the root is not locally frozen in a large neighborhood. In particular, we prove that the mixing time of the Glauber dynamics for colorings on the regular tree is $O(n \log n)$ in the entire nonreconstruction regime.

1. Introduction. There has been substantial interest in understanding the mixing times of Markov chains for sampling spin systems, in particular how they relate to the spatial mixing properties of the Gibbs measure. In the case of random colorings on trees, one natural conjecture is that such chains are rapidly mixing [i.e., the mixing time is $O(n \log n)$] whenever the corresponding model is in its nonreconstruction regime. The conjecture was previously verified for the so-called block dynamics by Bhatnagar et al. [2]. In this paper, we establish the result for the original Glauber dynamics and establish the following result.

THEOREM 1.1. *For any $\beta < 1 - \ln 2$, there exists a constant $k(\beta)$, for any $k > k(\beta)$ and*

$$d \le k[\log k + \log\log k + \beta],$$

the mixing time of the Glauber dynamics of the k-coloring model on n-vertex d-ary tree is $O(n \log n)$.

Received January 2015; revised July 2016.

[1]Supported by an Alfred Sloan Fellowship and NSF Grant DMS-12-08338.

MSC2010 subject classifications. 60J10.

Key words and phrases. Mixing time, Glauber dynamics, graph colorings, reconstruction problem.

Our requirement in d is sharp with respect to the best known bound for nonreconstruction established in [17]. In the same paper, it is shown that the k-coloring model is reconstructible for $d \geq k[\log k + \log\log k + 1 + o(1)]$. In a forthcoming work, we will give an improved upper bound on the reconstruction threshold from which it will follow that the $O(n \log n)$ mixing holds throughout the whole nonreconstruction regime [18].

1.1. *Previous work.* There have been intensive studies on the mixing times of Markov chains for sampling spin systems in both theoretical computer science and statistical physics. Many results have shown that the mixing time of the Glauber dynamics, both for the k-coloring model and general spin systems, are related to the spatial properties of the Gibbs measure. Two properties of primary interest are the uniqueness of the infinite volume Gibbs measure and reconstructability, which corresponds to the extremality of the infinite Gibbs measure induced by free boundary conditions.

In a sequence of results by Martinelli, Sinclair and Weitz [12, 13, 21], it was shown under quite general settings that the Glauber dynamics exhibits rapid mixing on d-regular trees regardless of the boundary condition, when the corresponding spin system admits an unique infinite-volume Gibbs measure. Their method uses the decay of correlation between the root and the leaves to bound the log-Soblev constant of the block dynamics. Less general results are known beyond the uniqueness threshold. The main obstacle, as in the case of graph colorings, is that the chain might be reducible under certain boundary conditions. Thus, one can not hope to get a meaningful bound for all boundary conditions.

Notwithstanding this, it is still interesting to consider the mixing time under the free boundary condition. The correlation between the roots and leaves in the absence of boundary conditions is closely related to the problem of reconstructions on trees. Roughly speaking, a model on trees is reconstructible if, given the leaves of a randomly chosen configuration, one's best guess for the root is "strictly better" than the stationary distribution, as the number of levels goes to infinity. In other words, reconstruction corresponds to a nonvanishing influence of an average boundary configurations on the root. It is natural to hope that the rapid mixing of the Glauber dynamics holds throughout the nonreconstruction regime.

We restrict our attention to the k-coloring problem on d-ary trees for the moment. The uniqueness of the Gibbs measure is shown to hold for $k \geq d + 2$ by Jonasson [9] and the results of [12, 13] imply an $O(n \log n)$ mixing time in the same region. For the reconstruction threshold, Mossel and Peres [15] proved reconstruction for $k \geq (1 + o(1))d/\log d$ by considering the probability of having a frozen boundary condition, that is, boundary condition that uniquely determines the root. In the other direction, Bhatnagar et al. [2] and Sly [17] independently proved that the model is nonreconstructible for $k \leq (1 + o(1))d/\log d$. It is also shown in [2] that the block dynamics for k-coloring model mixes in $O(n \log n)$ time in the same region using nonreconstruction and following the methods of

[12]. However, their result cannot be easily extended to the Glauber dynamics due to the failure of Markov chain comparison between the two dynamics. Namely, one step of the block dynamics might not be replaced by steps of the Glauber dynamics of bounded length.

For more results in the nonuniqueness regime, Berger et al. [1] showed for general models that the mixing time on trees is at most polynomial whenever the dynamics is ergodic, which in the case of coloring corresponds to $k \geq 3$ and $d \geq 2$. Goldberg et al. [7] proved an upper bound of $n^{O(d/\log d)}$ for the complete tree with branching factor d. Lucier et al. [10] showed $n^{O(1+d/k\log d)}$ mixing time for all d and $k \geq 3$. Recently, Tetali et al. [19] proved that the mixing time undergoes a phase transition at the reconstruction threshold $k = (1+o(1))d/\log d$, where their upper bound for the mixing time when $k \geq (1+o(1))d/\log d$ is $O(n^{1+o_k(1)})$. They also showed that the mixing time is $\Omega(n^{d/k\log d - o_k(1)})$ for $k \leq (1-o(1))d/\log d$, that is, rapid mixing does not hold in the reconstruction regime.

The main purpose of this paper is to reduce the mixing time in the nonreconstruction regime from the polynomial upper-bound of $n^{1+o(1)}$ to the sharp bound of $O(n\log n)$. Our proof is based on a modification of the techniques used in [12]. The main obstacle, as hinted above, is the reducibility of the Glauber dynamics on subtrees under fixed boundary condition. Heuristically, below the uniqueness threshold, there will be vertices whose states are "frozen" by their neighbors. While the block dynamics can update "frozen" vertices together with their neighbors in one single move, extra efforts are needed for the Glauber dynamics to pass around the barrier and "defreeze" the vertices, leading to the failure of the standard Markov chain comparison result between the two dynamics. With that in mind, we introduce a new variant of the block dynamics that focuses on the connected component on the state space of the usual block dynamics induced by valid moves of the Glauber dynamics. By carefully examining the portion of "frozen" vertices and their influences on nearby sites, we will show rapid mixing of our new version of the block dynamics which implies the final result.

We conclude this section by describing the literature on the mixing times on general graphs. For k-colorings on graphs with n vertices and maximal degree d, the Glauber dynamics is not in general irreducible if $k \leq d+1$. A long-standing conjectured is that the chain exhibits rapid mixing whenever $k \geq d+2$. So far, the best result on general graphs is given by Vigoda in [20], where he showed $O(n^2\log n)$ mixing time for $k \geq \frac{11}{6}d$. A series of improvements on the constant $\frac{11}{6}$ for rapid mixing have been made with extra conditions on the degree or the girth. See the survey [5] for more results toward this direction.

1.2. *General spin system.* The correspondence between rapid mixing and spatial correlation decay is not restricted to the coloring model alone, but is a common phenomenon that extends to general spin systems. For instance, Weitz conjectured in [21] that for any k-state spin system on regular trees, the system mixes in $O(n\log n)$ time whenever it admits an unique Gibbs measure and the Glauber dynamics is connected under given boundary condition. He proved the statement for $k = 2$ and for the ferromagnetic Potts model and colorings as two special cases of $k > 2$. He also provided a sufficient condition that applies to a wide range of other models.

As suggested by the case of the coloring model, the mixing time under free boundary condition is more closely related to the reconstruction threshold. In fact, Berger et al. [1] showed that for general spin systems on trees, $O(n)$ relaxation time under free-boundary condition implies nonreconstruction. Our methods for the coloring model can also be extended to general k-state spin systems provided that the spin system satisfies certain mild connectivity conditions. Therefore, as an intermediate result, we provide a sufficient condition for spin systems to exhibit rapid mixing in the nonreconstruction region.

In Section 2.1, we specify a spin system by its Markov chain kernel M, where $M(c, c') = \mu(\sigma_y = c' \mid \sigma_x = c)$ for any $(x, y) \in E$, and restrict our discussion to kernels that are ergodic and reversible (see also [6] for more details). Let λ be the second largest eigenvalue of M. We show that the Glauber dynamics is rapidly mixing for spin systems M assuming a certain *connectivity condition* $\mathcal{C}$ that will be specified in Section 2.3.

THEOREM 1.2. *Let M be a k-state spin system on the n-vertex d-ary tree T with second eigenvalue λ. If M satisfies the connectivity condition $\mathcal{C}$, is nonreconstructible on T, and $d\lambda^2 < 1$ then the mixing time of Glauber dynamics on T under free boundary condition is $O(n \log n)$.*

In the statement of Theorem 1.2, the connectivity condition C mainly concerns about the hard constraints. Roughly speaking, it requires the root to be able to "change freely" between all k states with high probability as the size of the tree grows. In particular, it includes all models without hard constraints or models with a permissive state, a state that can occur next to all other states (e.g., the hardcore model).

The requirement of $d\lambda^2 < 1$ comes from the Kesten–Stigum bound $d\lambda^2 = 1$ in reconstruction problems: Whenever $d\lambda^2 > 1$, the system is reconstructible by simply counting the number of leaves in each state [14]. Hence, nonreconstruction implies $d\lambda^2 \le 1$. The Kestin–Stigum bound is known to be tight for models including the Ising model (symmetric binary channel) and near-symmetric binary channels [3]. For other models such as hardcore model and graph colorings, it is strictly larger than the true threshold. Nonetheless, it has been suggested that the speed of decay of correlation undergoes a phase transition at the critical value $d\lambda^2 = 1$ with different scalings for $d\lambda^2 = 1$ and $d\lambda^2 < 1$. Indeed, a recent work of Ding, Lubetzky and Peres [4] showed that the mixing time for the Ising model is at least of order $n \log^3 n$ when $d\lambda^2 = 1$. Therefore, we can only hope to prove Theorem 1.2 for $d\lambda^2$ strictly smaller than 1.

2. Preliminaries.

2.1. *Definition of model.* *General spin systems*: Throughout the paper, we will write $T = (V, E)$ for the d-ary tree (i.e., every vertex have d offspring) with root ρ and $|V| = n$ vertices and denote the lth level of T by L_l. In particular, we have $L_0 = \{\rho\}$. Given vertex $x \in T$, we will use T_x to represent the subtree rooted at x and let $B_{x,l}$, $L_{x,l}$ denote the first l levels and the lth level of T_x, respectively.

Let $[k] = \{1, \dots, k\}$ denote the set of possible spin values. We are interested in general k-state spin systems specified by potentials U and W, where U is a symmetric function from $[k] \times [k] \to \mathbb{R} \cup \{\infty\}$ and W is a function from $[k] \to \mathbb{R}$. Given U and W, the (free-boundary) Gibbs measure on T is the probability measure on configurations $\sigma \in [k]^V$ defined as

$$\mu(\sigma) = \frac{1}{Z} \exp\left[-\left(\sum_{(x,y)\in E} U(\sigma_x, \sigma_y) + \sum_{x \in V} W(\sigma_x)\right)\right],$$

where Z, also known as the partition function, is the normalizing constant such that $\sum_{\sigma \in [k]^V} \mu(\sigma) = 1$. We say that a configuration σ is *proper* if $\mu(\sigma) > 0$ and denote the set of proper configurations on T by $\Omega_T = \{\sigma : \mu(\sigma) > 0\}$. For each pair of states $(i, j) \in [k]^2$, we say that (i, j) is a hard constraint if $U(i, j) = \infty$, otherwise we say that i and j are *compatible*. For each subset of vertices $A \subseteq T$, we will write σ_A for the restriction of σ to A and use superscript for boundary conditions. In particular, $\Omega_A^\eta = \{\sigma : \sigma \in \Omega_T, \sigma_{T\setminus A} = \eta_{T \setminus A}\}$ is the set of configurations compatible with boundary condition η and we denote the conditional law on $\Omega^\eta(A)$ as $\mu_A^\eta(\sigma) = \mu(\sigma \mid \sigma \in \Omega_A^\eta)$.

For the reconstruction problem, it is easy to work with the Markov chain construction of the Gibbs measure on trees, which can be taken as a special case of the broadcast model on trees. Information is sent on tree T from the root ρ downwards and each edge acts as a noisy channel. For each input $c_1 \in [k]$, the output of c_2 is chosen randomly according to some probability kernel $M(c_1, c_2)$. If the input at the root is chosen according to the stationary distribution of M, denoted by π, then the law of a random configuration on T is given by

$$\mu(\sigma) = \pi(\sigma_\rho) \prod_{(x,y)\in E} M(\sigma_x, \sigma_y).$$

It is easy to check that for any reversible probability kernel M, the aformentioned probablity measure corresponds to the spin system with potentials U, W given by

$$U(c_1, c_2) = -\ln\left(\frac{M(c_1, c_2)}{\pi(c_2)}\right), \qquad W(c) = -\ln \pi(c). \tag{2.1}$$

Note that not all potential pairs U, W can be expressed this way. A necessary condition for (2.1) is that

$$\sum_{c' \in [k]} \exp[-(U(c, c') + W(c'))] \equiv C \qquad \forall c \in [k]$$

for some constant C. We will henceforth restrict our attention to spin systems that can be expressed as (2.1) and refer such systems by their probability kernel M. We will also assume that M is ergodic and reversible.

The principal example of spin systems for this paper is the graph coloring model, where for each $c, c \in [k]$ $W(c) \equiv 0, U(c, c') = \infty \cdot 1(c = c')$, or equivalently $M(c, c') = \frac{1}{k-1} 1(c = c'), \pi(c) \equiv \frac{1}{k}$. A proper k-coloring of the graph $G = (V, E)$ is an assignment $\sigma : V \to [k]$ such that for each edge $(x, y) \in E$, $\sigma_x \neq \sigma_y$ and the Gibbs measure is the uniform distribution over all proper colorings. In the statistical physics literature, this corresponds to zero-temperature anti-ferromagnetic Potts model.

Uniqueness and reconstruction: Two key notions of spatial decay of correlation for spin systems on trees are the uniqueness and reconstruction thresholds. Recalling that L_l is the set of the vertices at level l in T, we have the following definition.

DEFINITION (Reconstruction). For $k \geq 2$, we say that a k-state system M is reconstructible on tree T if there exist two states $c, c' \in [k]$ such that

$$\limsup_{l \to \infty} d_{\mathrm{TV}}\big(\mu(\sigma_{L_l} = \cdot \mid \sigma_\rho = c), \mu\big(\sigma_{L_l} = \cdot \mid \sigma_\rho = c'\big)\big) > 0.$$

Otherwise, we say that the system has nonreconstruction on T.

Nonreconstruction is equivalent to the extremality of infinite volume Gibbs measure under free boundary conditions. More equivalent definition and an extensive literature review are given in the survey [14]. A strictly stronger condition is the uniqueness property.

DEFINITION (Uniqueness). For $k \geq 2$, we say that a k-state system M has uniqueness on tree T if

$$\limsup_{l \to \infty} \sup_{\eta, \eta' \in \Omega_{L_l}} d_{\mathrm{TV}}\big(\mu(\sigma_\rho = \cdot \mid \sigma_{L_l} = \eta), \mu\big(\sigma_\rho = \cdot \mid \sigma_{L_l} = \eta'\big)\big) > 0,$$

where Ω_{L_l} is the set of configurations restricted to level l.

Glauber dynamics and mixing time: The Glauber dynamics of a k-state spin system M is a Markov chain X_t on state space Ω_T. A step of the Markov chain from X_t to X_{t+1} is defined as follows:

1. Pick a vertex x uniformly at random from T;
2. Pick a state $c \in [k]$ according to the conditional distribution of the spin value of x given the rest of configuration, that is, state c is picked with probability $\mu^{\sigma}_{\{x\}}(c) = \mu(\sigma'_x = c \mid \sigma'_y = \sigma_y, \forall y \neq x)$;
3. Set $X_{t+1}(x) = c$ and $X_{t+1}(y) = X_t(y)$, for all $y \neq x$.

In the case of graph coloring, the second step corresponds to picking uniformly at random colors that do not appear in the neighborhood of x.

To justify our study of the Glauber dynamics under free boundary condition, we first show that the Markov chain is irreducible, and hence ergodic on the set of all proper configurations. For the sake of recursive analysis on subtrees later, we also prove irreducibility in a related case where the root of T is connected to one more vertex, namely its "parent", and the value of its parent is fixed. For each $c \in [k]$, let Ω_T^c denote the set of configurations with the parent of root ρ fixed to state c and let μ_T^c be the corresponding conditional Gibbs measure.

LEMMA 2.1. *For any k-state system M on d-ary tree T, if M is reversible and ergodic, then Ω_T is irreducible under the Glauber dynamics and so is Ω_T^c for each $c \in [k]$.*

PROOF. We first prove the irreducibility of Ω_T^c by induction on the number of levels l in T. For $l = 0$, it is trivially true since Ω_T^c is simply the set of states compatible of c. Suppose that the Glauber dynamics is irreducible on the $(l-1)$-level tree. For the l-level tree T, we need to show that for any two configurations $\sigma, \sigma' \in \Omega_T^c$, there exists a path of valid moves of the dynamics connecting σ to σ'. To construct such a path, one can first change every vertex $x \in L_1$ to state c by a sequence of moves in the tree T_x. This is possible since alternating layers of states c and σ_ρ is a proper configuration in $\Omega_{T_x}^{\sigma_\rho}$ and any two configurations in $\Omega_{T_x}^{\sigma_\rho}$ are connected by the inductive hypothesis. One may then change the spin of the root from σ_ρ to σ'_ρ, since both states are compatible with c. Finally, we may change the configuration of every subtree T_x to σ'_{T_x} using the inductive hypothesis, ending in the configuration σ'.

To show the irreducibility of Ω_T, we choose $\sigma, \sigma' \in \Omega_T$. By the ergodicity of M, there exists a sequence of states $c_0, \dots, c_{2m} \in [k]$ such that $c_0 = \sigma_\rho$, $c_{2m} = \sigma'_\rho$ and for each $0 \le i \le 2m-1$, c_i is compatible with c_{i+1}. For each $0 \le i \le m$, let $\tau_i \in \Omega_T$ be the configuration with alternating layers of c_{2i} and c_{2i+1}(let c_{2m+1} be an arbitrary state compatible with $c_{2m} = \sigma'_\rho$). One can first change σ to τ_0 using the irreducibility of the Glauber dynamics on $\Omega_{T_x}^{\sigma_\rho}$ for each $x \in L_1$, then for each $1 \le i \le m$ change from τ_{i-1} to τ_i by first changing all vertices on even levels to c_{2i} then vertices on odd levels to c_{2i+1}, and finally change each $(\tau_m)_{T_x}$ to σ'_{T_x}. $\square$

Lemma 2.1 implies that the Glauber dynamics with free boundary conditions will always converges to the Gibbs measure μ. The *mixing time* of the Glauber dynamics is defined as

$$t_{\text{mix}} = \max_{\sigma \in \Omega_T} \min\left\{t : d_{\text{TV}}\big(P^t(\sigma, \cdot), \mu\big) \le \frac{1}{2e}\right\},$$

where P is the probability kernel of X_t and $d_{\text{TV}}(\eta, \mu) = \frac{1}{2}\sum_\sigma |\eta(\sigma) - \mu(\sigma)|$ is the total variance distance. To bound the mixing time, we will make use of the *log-Sobolev constant*. For a nonnegative function $f : \Omega_T \to \mathbb{R}$, let $\mu(f) = \sum_\sigma \mu(\sigma) f(\sigma)$ be the expectation of f and $\mathrm{Ent}(f) = \mu(f \log f) - \mu(f) \log \mu(f)$ be its entropy. The Dirichlet form of f is defined as

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

And the log-Sobolev constant is defined as $\gamma = \inf_{f \geq 0} \frac{\mathcal{D}(\sqrt{f})}{\mathrm{Ent}(f)}$. Applying results in functional analysis to the Glauber dynamics yields the following bound (see, e.g., Theorem 2.2.5 of [16]).

THEOREM. *For k-state system M on n-vertex d-ary tree T, there exists a constant $C > 0$ such that $t_{\mathrm{mix}} \leq \frac{1}{\gamma} \cdot Cn \log n$.*

Therefore, to show rapid mixing it is enough to show that γ is uniformly bounded away from zero as $n \to \infty$.

2.2. *Component dynamics.* Next, we define a new variant of block dynamics on T, which we call the "component dynamics". Each step of the new dynamics updates a block of vertices each step, but only chooses configurations within the connected component of the Glauber dynamic. In this way, we can utilize the techniques in [12] while bypassing the problem that one step of the block dynamics may not be connected in the Glauber dynamics when $k \leq d + 1$. To give a formal definition, for $A \subset T$, we say that $\sigma' \sim_A \sigma$ if ${\sigma'}_{T \setminus A} = \sigma_{T \setminus A}$ and σ'_A, σ_A are connected by valid moves of the Glauber dynamics on A with fixed boundary condition $\sigma_{T \setminus A}$. We will omit the A in $\sigma \sim_A \sigma'$ when it is clear from context. Let $\Omega_A^{*,\sigma} = \{\sigma' \in \Omega_A^\sigma, \sigma' \sim_A \sigma\}$ denote the connected component of σ in Ω_A^σ, and $\mu_A^{*,\sigma}(\sigma') = \mu(\sigma' \mid \Omega_A^{*,\sigma})$ be the Gibbs distribution conditioned on both configuration outside A and the connected component within A.

For $l \geq 1$, recall $B_{x,l}$ is the block of l levels rooted at x and $L_{x,l}$ be the lth level of $B_{x,l}$. If x is within distance l of the leaves, let $B_{x,l} = T_x$. We define the *component dynamics* to be the Markov chain on Ω_T with the following update rule: In each step:

1. Pick a vertex x uniformly randomly from T.
2. Replace σ by σ' drawn from conditional distribution $\mu_{B_{x,l}}^{*,\sigma}$.

The dynamics is reversible with respect to the Gibbs distribution. For test functions $f : \Omega_T \to \mathbb{R}$, let $\mu_A^{*,\sigma}(f) = \sum_{\sigma' \in \Omega_A^{*,\sigma}} f(\sigma') \mu_A^{*,\sigma}(\sigma')$ be the conditional expectation of f on $\Omega_A^{*,\sigma}$ and for $f \geq 0$, let

$$\mathrm{Ent}_A^{*,\sigma}(f) = \mathrm{Ent}\big(f \mid \Omega_A^{*,\sigma}\big) = \mu_A^{*,\sigma}(f \log f) - \mu_A^{*,\sigma}(f) \log \mu_A^{*,\sigma}(f)$$

be the conditional entropy of f. We write the sum of local entropies of block size l as $\mathcal{E}_l^* = \sum_{x \in T} \mu_T(\mathrm{Ent}_{B_{x,l}}^{*,\sigma}(f))$. With minor modification, the comparison result of block dynamics also works for component dynamics (see, e.g., Proposition 3.4 of [11]; in the proof substitute $\mathcal{E}_D(f, f)$ by $\mathcal{E}_l^*$ and note $\sum_{\sigma'} \mu_T^\tau(\sigma')\mu_{B_{x,l}}^{*,\sigma'}(\sigma) = \mu_T^\tau(\sigma)$):

$$\gamma \geq \frac{1}{l} \cdot \inf_{f \geq 0} \frac{\mathcal{E}_l^*}{\operatorname{Ent}(f)} \cdot \min_{\sigma,x} \gamma_{B_{x,l}}^{*,\sigma},$$

where $\gamma_{B_{x,l}}^{*,\sigma}$ is the log-Soblev constant of the Glauber dynamics on $\Omega_{B_{x,l}}^{*,\sigma}$ with the boundary condition on $\partial B_{x,l}$ given by σ. From our definition of $\Omega_{B_{x,l}}^{*,\sigma}$, it is easy to see that $\min_{\sigma,x} \gamma_{B_{x,l}}^{*,\sigma}$ is a constant only depending on the branching number d, block size l and M itself and is strictly greater than 0 independent of T. Thus, to show $O(n \log n)$ mixing time for the Glauber dynamics, it is enough to show $\mathcal{E}_l^* \geq \text{const} \times \operatorname{Ent}(f)$ for all $f \geq 0$ and some choice of block size l independent of tree size $|T| = n$.

2.3. *Connectivity condition.* In this section, we specify the connectivity condition $\mathcal{C}$ mentioned in Theorem 1.2. First, we will define the notion of free vertices. Let T be a tree of l levels. Given configuration $\sigma \in \Omega_T$ with $\sigma_\rho = c$, $\sigma_{L_l} = \eta$, we say that the root can change (from c) to state c' *in one step* if and only if there exists a path $\sigma = \sigma^0, \sigma^1, \dots, \sigma^n \in \Omega_T$ such that:

1. $\sigma_{L_l}^i \equiv \eta$ for each $0 \leq i \leq n$. $\sigma_\rho^i = c$, for each $0 \leq i \leq n-1$ and $\sigma_\rho^n = c'$.
2. For each $0 \leq i \leq n-1$, configuration σ^i differs from σ^{i+1} at exactly one vertex.

Put another way, the path is a valid trajectory of the Glauber dynamics with fixed boundary condition which changes the state of ρ only once in the final step. For $x \in T$, we say x is *free* (in σ) if, considered as the root of T_x, x can change to all the other $(k-1)$-states in one step. We are interested in the probability that the root of an l-level tree is free and we denote it by $p_l^{\text{free}} = \mu(\sigma : \rho \text{ is free in } \sigma)$.

DEFINITION. We say that a k-state system M on the d-ary tree satisfies the connectivity condition $\mathcal{C}$ if M is ergodic, reversible and satisfies the following conditions:

1. If $k \geq 3$, then for any $c_1, c_2, c_3 \in [k]$, there exists $c \in [k]$ such that c is compatible with c_1, c_2, c_3.
2. The probability of being free tends to 1 as l tends to infinity, that is, $\lim_{l\to\infty} p_l^{\text{free}} = 1$.

Roughly speaking, the connectivity condition controls the behavior of "frozen" vertices in a typical configuration. As will be shown in Section 4, under the connectivity condition the probability that a vertex is "frozen" by the boundary condition is extremely small and the extra restriction of the component dynamics is negligible for vertices faraway from the bottom (see the remark after Claim 4.2 for more discussions).

2.4. *Outline of proof.* A key ingredient in [12] is that a certain strong concentration property implies "entropy mixing" in space which in turn implies the fast mixing of the block dynamics. The following Theorem 2.2 can be seen as the combination of Theorems 3.4 and 5.3 of [12] adapted to the component dynamics (the notation here is closer to Theorem 5.1 of [2]). For completeness, we include an outline of the proof in Section 5, highlighting the differences from [12].

THEOREM 2.2. *There exists some constant* $\alpha > 0$ *such that for every* $\delta > 0$ *and* $l \geq 1$ *the following statement holds*: *If for all* $x \in T$ *that is at least* l *levels from the leaves and all compatible pairs of states* $c, c' \in [k]$, *the conditional measure* $\mu^c = \mu^c_{T_x}$ *satisfies*

$$\Pr_{\tau \sim \mu^c}\left(\left|\frac{\mu^c(\sigma_x = c' \mid \sigma \sim_{B_{x,l}} \tau)}{\mu^c(\sigma_x = c')} - 1\right| \geq \frac{(1-\delta)^2}{\alpha(l+1-\delta)^2}\right) \leq e^{\frac{-2\alpha(l+1-\delta)^2}{(1-\delta)^2}}, \tag{2.2}$$

then for every function $f \geq 0$, $\mathrm{Ent}(f) \leq \frac{2}{\delta}\mathcal{E}^*_l$.

To prove Theorem 1.2, it suffices to verify (2.2) for some choice of l and δ. We first show a weaker inequality (2.3) in the following theorem. Note that the same inequality is proved in Theorem 5.3 of [12] or Theorem 5.1 of [2] for specific models such as the coloring model. Here, we provide a different proof that works for general models using only nonreconstruction.

THEOREM 2.3. *For a* k-*state system* M, *if* M *is nonreconstructible and* $d\lambda^2 < 1$, *then for any* $\alpha > 0, 0 < \delta < 1$, *there exist* $l_0 \geq 1$ *such that for all* $l \geq l_0$, *every* $x \in T$ *that is at least* l *levels from the leaves*, *and any pair of compatible states* $c, c' \in [k]$, $\mu^c = \mu^c_{T_x}$ *satisfies*

$$\Pr_{\tau \sim \mu^c}\left(\left|\frac{\mu^c(\sigma_x = c' \mid \sigma_{L_{x,l}} = \tau_{L_{x,l}})}{\mu^c(\sigma_x = c')} - 1\right| \geq \frac{(1-\delta)^2}{\alpha(l+1-\delta)^2}\right) \leq e^{\frac{-2\alpha(l+1-\delta)^2}{(1-\delta)^2}}. \tag{2.3}$$

The difference between (2.2) and (2.3) is that in equation (2.2), the inner measure μ^c conditions not only on the boundary condition $\sigma_{L_{x,l}} = \tau_{L_{x,l}}$, but also the connected component of τ. We will show that under connectivity condition $\mathcal{C}$, the difference between $\sigma \sim_{B_{x,l}} \tau$ and $\sigma_{L_{x,l}} = \tau_{L_{x,l}}$ is negligible in the upper half of a large block, hence (2.2) holds.

LEMMA 2.4. *Let* M *be a* k-*state system satisfying* $\mathcal{C}$ *such that* (2.3) *holds for* $l \geq l_0$ *and* $\delta = \delta_0$. *Then there exist constants* $l_1 \geq 2l_0$ *and* $\delta_1 \geq \delta_0$ *such that for all* $l \geq l_1$, *equation* (2.2) *holds with* $\delta = \delta_1$.

Theorems 2.2 and 2.3 and Lemma 2.4 together imply Theorem 1.2. The rest of the paper is structured as follows: We will prove Theorem 2.3 in Section 3 and Lemma 2.4 in Section 4, and we will include a sketch of Theorem 2.2 in Section 5. After that, we will apply the result to the k-coloring model and prove Theorem 1.1 in Section 6.

3. Proof of Theorem 2.3. In this section, we prove Theorem 2.3. The result for the k-coloring model was proved in [2], which used the specific structure of coloring model. Here, we will give a different proof for general systems M using only nonreconstruction and $d\lambda^2 < 1$. We first introduce some notation. Recall that the stationary distribution of M is π. For $x \in T$, let

$$\tilde{R}_{x,l}(\tau)(c) = \frac{1}{\pi(c)}\mu_{T_x}(\sigma_x = c \mid \sigma_{L_{x,l}} = \tau_{L_{x,l}})$$

denote the ratio of conditional and unconditional distribution at x and write $R_{x,l}(\tau) = \|\tilde{R}_{x,l}(\tau) - 1\|_\infty = \max_{c\in[k]} |\tilde{R}_{x,l}(\tau)(c) - 1|$. We will omit τ when it is clear from context. In the proof, we will work with the unconditional Gibbs measure $\mu = \mu_{T_x}$ and π instead of $\mu^c_{T_x}$ and $\mu^c_{T_x}(\sigma_x = c')$ and show the following stronger inequality.

THEOREM 3.1. *Under the assumptions of Theorem 2.3, there exists constant $\xi > 0$ and $l_0 > 0$, such that for all $l \ge l_0$, every $x \in T$ that is at least l levels from the leaves, $\mu = \mu_{T_x}$ satisfies*

$$\Pr_{\tau\sim\mu}\bigl(R_{x,l}(\tau) \ge e^{-\xi l}\bigr) \le \exp\bigl(-e^{\xi l}\bigr). \tag{3.1}$$

PROOF OF THEOREM 2.3. To see that (3.1) implies (2.3), consider the Markov chain construction of σ. Let E_x be the edge set of T_x, we have

$$\mu(\sigma) = \pi(\sigma_x) \prod_{(y,z)\in E_x} M(\sigma_y, \sigma_z), \qquad \mu^c(\sigma) = M(c, \sigma_x) \prod_{(y,z)\in E_x} M(\sigma_y, \sigma_z).$$

Hence, for any event $A \subseteq \Omega_{T_x}$,

$$\Pr_{\tau\sim\mu^c}(A) = \sum_{\tau\in A} \mu^c(\tau) = \sum_{\tau\in A} \frac{M(c, \tau_x)}{\pi(\tau_x)}\mu(\tau) \le \pi_{\min}^{-1}\mu(A) = \pi_{\min}^{-1} \Pr_{\tau\sim\mu}(A),$$

where $\pi_{\min} = \min_{c\in[k]} \pi(c) > 0$. Note that

$$\left|\frac{\mu^c(\sigma_x = c' \mid \sigma_{L_{x,l}} = \tau_{L_{x,l}})}{\mu^c(\sigma_x = c')} - 1\right| = \left|\frac{\mu(\sigma_x = c' \mid \sigma_{L_{x,l}} = \tau_{L_{x,l}})}{\pi(c')} - 1\right| \le R_{x,l}(\tau).$$

It follows that

$$\Pr_{\tau\sim\mu^c}\left(\left|\frac{\mu^c(\sigma_x = c' \mid \sigma_{L_{x,l}} = \tau_{L_{x,l}})}{\mu^c(\sigma_x = c')} - 1\right| \ge e^{-\xi l}\right) \le \pi_{\min}^{-1} \exp\bigl(-e^{\xi l}\bigr).$$

Theorem 2.3 then follows by taking l_0 large enough such that $\exp(-\xi l_0) \le (1 - \delta)^2/\alpha(l_0 - 1 + \delta)^2$. $\square$

In the rest of the section, we assume that M satisfies the assumptions of Theorem 3.1. The following lemma gives the recursive relation of $\tilde{R}_{x,l}(c)$.

LEMMA 3.2. *Fix $\tau \in \Omega_T$ and $x \in T$ and let $x_1, \dots, x_d$ denote the d children of x. $\tilde{R}_{x,l}$ can be written as a rational function of $\tilde{R}_{x_i,l-1}$:*

$$\begin{aligned}\tilde{R}_{x,l}(c) &= \frac{\prod_{i=1}^d M\tilde{R}_{x_i,l-1}}{\pi \prod_{i=1}^d M\tilde{R}_{x_i,l-1}}(c) \\ &= \frac{\prod_{i=1}^d \sum_{c_i \in [k]} M(c,c_i)\tilde{R}_{x_i,l-1}(c_i)}{\sum_{c' \in [k]} \pi(c') \prod_{i=1}^d \sum_{c_i \in [k]} M(c',c_i)\tilde{R}_{x_i,l-1}(c_i)}.\end{aligned} \tag{3.2}$$

PROOF. Let E_x and E_{x_i} denote the edge set of T_x and T_{x_i}, they satisfy that $E_x = \bigcup_i (E_{x_i} \cup \{(x,x_i)\})$. Let $\Omega_x(c) = \{\sigma : \sigma_x = c, \sigma_{L_{x,l}} = \tau_{L_{x,l}}\}$ and $\Omega_{x_i}(c) = \{\sigma : \sigma_{x_i} = c, \sigma_{L_{x_i,l-1}} = \tau_{L_{x_i,l-1}}\}$ be the set of configurations on T_x and T_{x_i} with boundary condition τ. By the Markov chain construction, we have

$$\begin{aligned}\mu\bigl(\Omega_x(c)\bigr) = \mu(\sigma_x = c, \sigma_{L_{x,l}} = \tau_{L_{x,l}}) &= \sum_{\sigma \in \Omega_x(c)} \pi(c) \prod_{(y,z)\in E} M(\sigma_y, \sigma_z) \\ &= \sum_{c_1,\dots,c_d \in [k]} \pi(c) \prod_{i=1}^d M(c,c_i) \sum_{\sigma^i \in \Omega_i(c_i)} \prod_{(y,z)\in E_i} M\bigl(\sigma_y^i, \sigma_z^i\bigr) \\ &= \sum_{c_1,\dots,c_d \in [k]} \pi(c) \prod_{i=1}^d \frac{M(c,c_i)}{\pi_{c_i}} \mu\bigl(\Omega_i(c_i)\bigr) \\ &= \pi(c) \prod_{i=1}^d \sum_{c_i \in [k]} \frac{M(c,c_i)}{\pi_{c_i}} \mu\bigl(\Omega_i(c_i)\bigr).\end{aligned}$$

Therefore, by Bayes' formula,

$$\begin{aligned}\tilde{R}_{x,l}(c) &= \frac{1}{\pi(c)} \mu(\sigma_x = c \mid \sigma_{L_{x,l}} = \tau_{L_{x,l}}) = \frac{1}{\pi(c)} \frac{\mu(\Omega_x(c))}{\sum_{c' \in [k]} \mu(\Omega(c'))} \\ &= \frac{\prod_{i=1}^d \sum_{c_i \in [k]} \frac{M(c,c_i)}{\pi_{c_i}} \mu(\Omega_i(c_i))}{\sum_{c' \in [k]} \pi(c') \prod_{i=1}^d \sum_{c_i \in [k]} \frac{M(c',c_i)}{\pi_{c_i}} \mu(\Omega_i(c_i))} \\ &= \frac{\prod_{i=1}^d \sum_{c_i \in [k]} M(c,c_i)\tilde{R}_{x_i,l-1}(c_i)}{\sum_{c' \in [k]} \pi(c') \prod_{i=1}^d \sum_{c_i \in [k]} M(c',c_i)\tilde{R}_{x_i,l-1}(c_i)},\end{aligned}$$

where the last step followed by dividing both the numerator and denominator by $\prod_{i=1}^d \sum_{c_i' \in [k]} \mu(\Omega_i(c_i'))$. □

Observe that in the recursive relationship of (3.2), $\tilde{R}_{x,l}(c)$ is a rational function of $\tilde{R}_{x_i,l-1}, i = 1, \dots, d$, where $\tilde{R}_{x,l}$ takes values from the k dimensional simplex $\Delta_{[k]} = \{R \in \mathbb{R}^k : \pi R = 1, R_i \geq 0, i = 1, \dots, k\}$. The next lemma establishes a contraction property of $R_{x,l}$, using the continuity of (3.2) and the ergodicity of M.

LEMMA 3.3. *There exist an integer $m \geq 1$ and constant $\varepsilon > 0$ such that for all d^m vertices $y_1, \dots, y_{d^m} \in L_{x,m}$, if at most one y_i has $R_{y_i,l-m} > \varepsilon$ then*

$$R_{x,l} \leq \frac{1}{2} \sum_{i=1}^{d^m} R_{y_i,l-m}. \tag{3.3}$$

PROOF. Let $f : \Delta_{[k]}^d \to \Delta_{[k]}$ be the function on the RHS of (3.2) such that $\tilde{R}_{x,l} = f(\tilde{R}_{x_1,l-1}, \dots, \tilde{R}_{x_d,l-1})$. Observe from (3.2) that f is a rational function with $f(1, \dots, 1) = 1$. When $\tilde{R}_{x_2,l-1} = \cdots = \tilde{R}_{x_d,l-1} = 1$, f can be simplified as

$$\tilde{R}_{x,l} = f(\tilde{R}_{x_1,l-1}, 1, \dots, 1) = \frac{M\tilde{R}_{x_1,l-1}}{\pi M \tilde{R}_{x_1,l-1}} = M\tilde{R}_{x_1,l-1}.$$

Iterating the function m times, we get $\tilde{R}_{x,l} = f^{(m)}(\tilde{R}_{y_1,l-m}, \dots, \tilde{R}_{y_{d^m},l-m})$ where $f^{(m)} : \Delta_{[k]}^{d^m} \to \Delta_{[k]}$ is another rational function. A similar calculation shows when $\tilde{R}_{y_2,l-m} = \cdots = \tilde{R}_{y_{d^m},l-m} = 1$,

$$\tilde{R}_{x,l} = f^{(m)}(\tilde{R}_{y_1,l-m}, 1, \dots, 1) = M^m \tilde{R}_{y_1,l-m}.$$

Since $f^{(m)}$ is a smooth function in any regions without poles, there exists constant $C_1 = C_1(d, m, M)$ such that in the local neighborhood of $(1, \dots, 1)$,

$$\begin{aligned} \left\| \tilde{R}_{x,l} - 1 - \sum_{i=1}^{d^m} \bigl(M^m \tilde{R}_{y_i,l-m} - 1\bigr) \right\| &\leq C_1 \sum_{i=1}^{d^m} \|\tilde{R}_{y_i,l-m} - 1\|^2 \\ &\leq C_1 k \sum_{i=1}^{d^m} \|\tilde{R}_{y_i,l-m} - 1\|_\infty^2. \end{aligned}$$

By the ergodicity of M, for sufficiently large m and all $\tilde{R} \in \Delta_{[k]}$ we have $\|M^m \tilde{R} - 1\|_\infty \leq \frac{1}{4} \|\tilde{R} - 1\|_\infty$. Therefore, there exists $\varepsilon_1 = \varepsilon_1(C_1, k)$ such that if $R_{y_i,l-m} \leq \varepsilon_1$ for all vertices $y_i \in L_{x,m}$ then

$$\|\tilde{R}_{x,l} - 1\|_\infty \leq \left(\frac{1}{4} + C_1 k \varepsilon_1\right) \sum_{i=1}^{d^m} \|\tilde{R}_{y_i,l-m} - 1\|_\infty \leq \frac{1}{2} \sum_{i=1}^{d^m} R_{y_i,l-m}. \tag{3.4}$$

This suffices provided that there are no large $R_{y_i,l-m}$.

We now consider the case when there is one large $R_{y_i,l-m}$, which we can without loss of generality assume is $i = 1$. Again since $f^{(m)}$ is smooth, there exists $C_2, \varepsilon_2 > 0$ such that for all $\tilde{R}_{y_1,l-m} > \varepsilon_1$, if $\sup_{i \geq 2} R_{y_i,l-m} \leq \varepsilon_2$ then

$$\|\tilde{R}_{x,l} - M^m \tilde{R}_{y_1,l-m}\| \leq C_2 \sum_{i=2}^{d^m} \|\tilde{R}_{y_i,l-m} - 1\|.$$

Let $\varepsilon = \varepsilon_2 \wedge (4C_2 d^m k)^{-1}\varepsilon_1$, if we moreover have $\sup_{i\ge 2} R_{y_i,l-m} \le \varepsilon$, then

$$\begin{aligned}\|\tilde{R}_{x,l} - 1\|_\infty &\le \frac{1}{4}\|\tilde{R}_{y_1,l-m} - 1\|_\infty + C_2 d^m k\varepsilon \\ &\le \frac{1}{4}R_{y_1,l-m} + \frac{1}{4}\varepsilon_1 \le \frac{1}{2}R_{y_1,l-m}.\end{aligned} \tag{3.5}$$

Combining equations (3.4) and (3.5) and noting that $\varepsilon < \varepsilon_1$ completes the proof. □

So far, we have not used the assumption of nonreconstruction and $d\lambda^2 < 1$. In [8], Janson and Mossel introduced the notion of "robust reconstruction" and showed the following result (rephrased to the notation here).

THEOREM 3.4 (Lemma 2.7 and Lemma 2.8 of [8]). *If M is ergodic and $d\lambda^2 < 1$, then there exist constants $C_1 = C_1(d) > 0$ and $\delta = \delta(d) > 0$ such that for any $l \ge 1$ if $d_{\mathrm{TV}}(\mu^c_{L_l}, \mu_{L_l}) \le \delta$ for all $c \in [k]$, then*

$$d_{\mathrm{TV}}(\mu^c_{L_{l+1}}, \mu_{L_{l+1}}) \le e^{-C_1} d_{\mathrm{TV}}(\mu^c_{L_l}, \mu_{L_l}) \qquad \textit{for all } c \in [k].$$

Theorem 3.4 combined with nonreconstruction implies the following weaker concentration inequality.

COROLLARY 3.5. *Under the assumptions of Theorem 3.1, there exist constants $C_1, C_2 > 0$ such that*

$$\Pr_{\tau\sim\mu}(R_{x,l} > z) \le \frac{C_2}{z} e^{-C_1 l}. \tag{3.6}$$

PROOF. By the definition of nonreconstruction, $\lim_{l\to\infty} d_{\mathrm{TV}}(\mu^c_{L_l}, \mu_{L_l}) = 0$. Hence, for sufficiently large l, $d_{\mathrm{TV}}(\mu^c_{L_l}, \mu_{L_l}) \le \delta$ and by induction there exists constant $C_2 > 0$ that

$$d_{\mathrm{TV}}(\mu^c_{L_l}, \mu_{L_l}) \le C_2 e^{-C_1 l}.$$

A duality argument then shows that

$$\begin{aligned}&\mathbb{E}_{\tau\sim\mu}|\tilde{R}_{x,l}(c) - 1| \\ &\quad = \mathbb{E}_{\tau\sim\mu}\left|\frac{1}{\pi(c)}\mu(\sigma_x = c \mid \sigma_{L_{x,l}} = \tau_{L_{x,l}}) - 1\right| = \mathbb{E}_{\tau\sim\mu}\left|\frac{\mu^c(\sigma_{L_{x,l}} = \tau_{L_{x,l}})}{\mu(\sigma_{L_{x,l}} = \tau_{L_{x,l}})} - 1\right| \\ &\quad = \sum_\tau |\mu^c(\sigma_{L_{x,l}} = \tau_{L_{x,l}}) - \mu(\sigma_{L_{x,l}} = \tau_{L_{x,l}})| = 2d_{\mathrm{TV}}(\mu^c_{L_l}, \mu_{L_l}) \le 2C_2 e^{-C_1 l}.\end{aligned}$$

Maximizing over $c \in [k]$, we get $\mathbb{E}_{\tau\sim\mu} R_{x,l} \le C_2 e^{-C_1 l}$ for some (different) constant $C_1, C_2 > 0$ and (3.6) follows by Markov's inequality. □

Finally, we improve the concentration bound of (3.6) using Lemma 3.3.

PROOF OF THEOREM 3.1. By Lemma 3.3, the event $R_{x,l} > z$ implies that either there exist two $i \in [d^m]$ such that $R_{y_i,l-m} > \varepsilon$ or $\sum_{i=1}^{d^m} R_{y_i,l-m} > 2z$. In the second case if the event $\sum_{i=1}^{d^m} R_{y_i,l-m} > 2z$ holds and for every y_i, $R_{y_i,l-m} \le \frac{3}{2}z$, then there must exist at least two i such that $R_{y_i,l-m} > \frac{1}{2d^m}z$, otherwise $\sum_{i=1}^{d^m} R_{y_i,l-m} \le \frac{3}{2}z + \frac{d^m-1}{2d^m}z < 2z$. Therefore, we can write

$$\begin{aligned}\Pr_{\tau\sim\mu}(R_{x,l} > z) \le{}& \Pr_{\tau\sim\mu}(\exists \text{ two } y_i \in L_{x,m}, R_{y_i,l-m} > \varepsilon)\\ &+ \Pr_{\tau\sim\mu}\left(\exists y_i \in L_{x,m}, R_{y_i,l-m} > \frac{3}{2}z\right)\\ &+ \Pr_{\tau\sim\mu}\left(\exists \text{ two } y_i \in L_{x,m}, R_{y_i,l-m} > \frac{1}{2d^m}z\right).\end{aligned}$$

Let $g(z,l) = \Pr_{\tau\sim\mu}(R_{x,l} > z)$ and $C = \max\{2d^m, \frac{1}{\varepsilon\pi_{\min}}\}$, note $g(z,l)$ is a decreasing function in z, the equation above become

$$\begin{aligned}g(z,l) &\le d^{2m}g^2(\varepsilon, l-m) + d^m g\left(\frac{3}{2}z, l-m\right) + d^{2m}g^2\left(\frac{1}{2d^m}z, l-m\right)\\ &\le d^m g\left(\frac{3}{2}z, l-m\right) + 2d^{2m}g^2\left(\frac{1}{C}z, l-m\right).\end{aligned}$$

Iterating this estimation h times, we have

$$g(z,l) \le \sum_{i=0}^{h} (2d^{2m})^{2^{h-i}(1+i)} g^{2^{h-i}}\left(\left(\frac{3}{2}\right)^i \left(\frac{1}{C}\right)^{h-i} z, l-hm\right), \tag{3.7}$$

where the coefficient can be shown by induction on h using inequality $(a+b)^2 \le 2(a^2+b^2)$.

Since for all $z > \pi_{\min}^{-1}$, we have $g(z,l) = 0$, the summand on the RHS of (3.7) is zero for large i. Fix $\kappa = \log\frac{4}{3}C/\log\frac{3}{2}C < 1$, for $h \ge \log(\frac{1}{z\pi_{\min}})/\log(\frac{4}{3})$ and $i > \kappa h$, we have $(\frac{3}{2})^i(\frac{1}{C})^{h-i}z > \pi_{\min}^{-1}$. Therefore

$$\begin{aligned}g(z,l) &\le \sum_{i=0}^{\kappa h} (2d^{2m})^{2^{h-i}(1+i)} g^{2^{h-i}}\left(\left(\frac{3}{2}\right)^i \left(\frac{1}{C}\right)^{h-i} z, l-hm\right)\\ &\le \kappa h\left[(2d^{2m})^h g(C^{-h}z, l-hm)\right]^{2^{(1-\kappa)h}}.\end{aligned}$$

Now apply (3.6) and let $h = rl/m$ for small $r > 0$ such that $(1-r)C_1 - r\cdot\frac{1}{m}\log(2Cd^{2m}) > \frac{1}{2}C_1 > 0$. For large enough l such that $\log l \le 2^{\frac{(1-\kappa)r}{m}l}$, we have

$$\begin{aligned}g(z,l) &\le \kappa h\left((2d^{2m})^h \frac{C_2 C^h}{z} e^{-C_1(l-hm)}\right)^{2^{(1-\kappa)h}}\\ &\le \frac{\kappa r}{m} l\left(\frac{C_2}{z}(2Cd^{2m})^{\frac{r}{m}l} e^{-C_1(1-r)l}\right)^{2^{\frac{(1-\kappa)r}{m}l}} \le \frac{\kappa r}{m}\left(\frac{2C_2}{z} e^{-\frac{1}{2}C_1 l}\right)^{2^{\frac{(1-\kappa)r}{m}l}}.\end{aligned}$$

Let $C_3 = \frac{2}{C_1}$, $C_4 = \frac{\kappa r}{m}$, $C_5 = \frac{(1-\kappa)r}{m}\log 2$. For $l > C_3(1+\log 2C_2 - \log z)$, we have $g(z,l) \le C_4 \exp\{-\exp(C_5 l)\}$.

Finally, define $\xi = \frac{1}{2}\min\{C_3^{-1}, C_5\}$, plug in $z_l = \exp(-\xi l)$. When l is large enough, we have $C_3(1 + \log 2C_2 - \log z) \le C_3(1 + \log 2C_2) + \frac{1}{2}l < l$ and $\exp(\exp(\frac{1}{2}C_5 l)) > C_4$, therefore,

$$\Pr_{\tau\sim\mu}\bigl(R_{x,l}(\tau) \ge e^{-\xi l}\bigr) = g(z_l, l) \le C_4 \exp\bigl(-e^{C_5 l}\bigr) \le \exp\bigl(-e^{\xi l}\bigr),$$

completing the proof. □

4. Proof of Lemma 2.4. The proof of Lemma 2.4 contains two steps. First, for block $B_{x,l}$ with sufficiently large l, we study the measure $\mu^{*,\tau}_{B_{x,l}}$ induced on the upper half of block $B_{x,l/2}$ (here and throughout the section, we choose l to be even) and consider the following subset of $\Omega^{\tau}_{B_{x,l}}$:

$$A_\tau = \bigl\{\sigma \in \Omega^{\tau}_{B_{x,l}} : \forall x \in L_{x,l/2+2}, x \text{ is free w.r.t. } \sigma\bigr\}.$$

A_τ can be considered as the set of "good" configurations with boundary condition τ. As we will show later, under connectivity condition $\mathcal{C}$, $\mu^{*,\tau}_{B_{x,l}}(A_\tau)$ is close to 1 with high probability. And as the following lemma claims, conditioning on A_τ and the configuration on $L_{x,l/2}$, the boundary of $B_{x,l/2}$, the marginal of x induced by $\mu^{*,\tau}_{B_{x,l}}$ equals to the marginal induced by μ^c. Therefore, as a second step we can apply the result of Theorem 2.3 to $B_{x,l/2}$. Let $\Omega_{L_{x,l/2}}$ be the set of configuration on $L_{x,l/2}$. Throughout the section, we assume that M satisfies the connectivity condition $\mathcal{C}$.

LEMMA 4.1. *For arbitrary $\tau \in \Omega^{c}_{T_x}$, $\eta \in \Omega_{L_{x,l/2}}$ and state $c' \in [k]$ that is compatible with c,*

$$\mu^{*,\tau}_{B_{x,l}}\bigl(\sigma_x = c' \mid \sigma_{L_{x,l/2}} = \eta, \sigma \in A_\tau\bigr) = \mu^c\bigl(\sigma_x = c' \mid \sigma_{L_{x,l/2}} = \eta\bigr). \tag{4.1}$$

PROOF. For convenience of notation, abbreviate $\sigma_{(1)} = \sigma_{B_{x,l/2-1}}$, $\sigma_{(2)} = \sigma_{B_{x,l}\setminus B_{x,l/2}}$, so every configuration $\sigma \in \Omega_{B_{x,l}}$ can be written as a three tuple $(\sigma_{(1)}, \eta, \sigma_{(2)})$. We of course have that $\sigma_{(1)}, \sigma_{(2)}$ are conditionally independent given $\sigma_{L_{x,l/2}} = \eta$. By the definition of A_τ, $\{\sigma \in A_\tau\}$ only depends on $\sigma_{(2)}$. Therefore, to show (4.1), it is enough to show that conditioned on $\sigma_{L_{x,l/2}}$ and $\sigma \in A_\tau$, $\sigma \sim \tau$ is independent of $\sigma_{(1)}$. From there we have

$$\begin{aligned}
&\mu^{*,\tau}_{B_{x,l}}\bigl(\sigma_x = c' \mid \sigma_{L_{x,l/2}} = \eta, \sigma \in A_\tau\bigr)\\
&\quad= \mu^c\bigl(\sigma_x = c' \mid \sigma_{L_{x,l/2}} = \eta, \sigma \sim \tau, \sigma \in A_\tau\bigr)\\
&\quad= \mu^c\bigl(\sigma_x = c' \mid \sigma_{L_{x,l/2}} = \eta, \sigma \in A_\tau\bigr) = \mu^c\bigl(\sigma_x = c' \mid \sigma_{L_{x,l/2}} = \eta\bigr).
\end{aligned}$$

Since "$\sim$" is a transitive relation, the conditional independence of $\sigma \sim \tau$ and $\sigma_{(1)}$ follows from the following claim. □

CLAIM 4.2. *For each $\tau \in \Omega^c_{T_x}$, $\eta \in \Omega_{L_{x,l/2}}$ and for all $\sigma = (\sigma_{(1)}, \eta, \sigma_{(2)})$, $\sigma' = (\sigma'_{(1)}, \eta, \sigma_{(2)}) \in \Omega^\tau_{B_{x,l}}$ if $\sigma, \sigma' \in A_\tau$, then $\sigma \sim \sigma'$.*

PROOF. For each $x \in T$, let $p(x)$ denote the parent of x. By Lemma 2.1, there exists a path Γ connecting $\sigma_{(1)}$ to $\sigma'_{(1)}$ in $\Omega^c_{B_{x,l/2}}$ via valid moves of the Glauber dynamics on $B_{x,l/2}$ with $\sigma_{p(x)} = c$ and free boundary condition on $L_{x,l/2}$. We will construct a path Γ' in $\Omega^\tau_{B_{x,l}}$ connecting σ to σ' by adding steps between steps of Γ which only changes the configuration on $B_{x,l} \setminus B_{x,l/2}$, such that vertices in $L_{x,l/2+1}$ won't block the moves in Γ and after finishing Γ, we can change the configuration on $B_{x,l} \setminus B_{x,l/2}$ back to the original $\sigma_{(2)}$. The construction of Γ' is specified below:

(1) *Before starting* Γ. For each $y \in L_{x,l/2+2}$, $\sigma \in A_\tau$ implies that there exists a path Γ_y in T_y changing y from σ_y to $\sigma_{p(p(y))} = \eta_{p(p(y))}$ in one step. To see Γ_y is also a connected path in $B_{x,l}$, we have to show that the parent of y won't block Γ_y. The only neighbor of $p(y)$ in T_y is y and the only move involving y in Γ_y is the last step changing y from σ_y to $\sigma_{p(p(y))}$. The value of $p(y)$ will not block this last step because $\sigma_{p(y)}$ is compatible with both σ_y and $\sigma_{p(p(y))}$ (they are states of neighboring vertices in σ). Now we will concatenate the Γ_y's for each $y \in L_{x+l/2+2}$ and change σ_y to $\sigma_{p(p(y))}$. After that, for each $w \in L_{x,l/2}$, all vertices in $L_{w,2}$ are in state $\sigma_w = \eta_w$. The configuration on and below $L_{x,l/2+2}$ will henceforth remain fixed until we finish Γ.

(2) *Performing* Γ. For each step in Γ, the existence of $B_{x,l-1} \setminus B_{x,l/2}$ might block this move only if it changes the state of some vertex $w \in L_{x,l/2}$. Suppose it changes w from c_1 to c_2. Remember in the construction above, all vertices in $L_{w,2}$ have states η_w. By part 1 of $\mathcal{C}$, we can find $c_3 \in [k]$ which is compatible with c_1, c_2 and η_w. Now in order to change w from c_1 to c_2, it suffices to first change the state of every vertex $z \in L_{w,1}$ to c_3, and then change w from c_1 to c_2. This construction keeps the configuration on and below $L_{x,l/2+2}$ unchanged.

(3) *After* Γ. After the moves in Γ, the configuration in $B_{x,l/2}$ is $(\sigma'_{(1)}, \eta)$. We can change every vertex $z \in L_{x,l/2+1}$ back to $\sigma'_z = \sigma_z$ because at this moment its parent $p(z) \in L_{x,l/2}$ and all children of z in $L_{z,1}$ have state $\eta_{p(z)} = \sigma_{p(z)}$, which is compatible with σ_z. From there, we can reverse the path Γ_y for each $y \in L_{x,l/2+2}$ and change the configuration on and below $L_{x,l/2+2}$ back to the original configuration $\sigma_{(2)}$. This completes the construction achieving $\sigma'_{(2)} = \sigma_{(2)}$. $\square$

LEMMA 4.3. *There exist constants $C_1 > 1$, $C_2 > 0$ such that for all $l \geq 1$,*

$$1 - p_l^{\text{free}} \leq C_2 \exp(-C_1^l). \tag{4.2}$$

PROOF. Fix $x \in T$ and $\sigma \in \Omega_{T_x}$. First if for all $1 \leq i \leq d$, $z_i \in L_{x,1}$ is free, then x is also free. To see that, for any $c \in [k]$, by connectivity condition there exists $c' \in [k]$ such that c' is compatible with both c and σ_x, we can first change all z_i to c' in one step and then change x from σ_x to c as the final step.

Now consider the set of y_{ij}'s where $y_{ij} \in L_{z_i,1} \subset L_{x,2}$ for $1 \le i, j \le d$. If at most one of the y_{ij}'s is not free, say $y_{11} \in L_{z_1,1}$, then for each $i \neq 1$, z_i is free and z_1 can change in one step to all states compatible with $\sigma_{y_{11}}$. Again by $\mathcal{C}$, for all $c \in [k]$ there exists $c' \in [k]$ such that c' is compatible with c, σ_x and $\sigma_{y_{11}}$. By the construction above, we can change x from σ_x to c in one step, hence x is also free.

This implies if x is not free, then there exist at least two $y_{ij} \in L_{x,2}$ that are not free. By part 2 of $\mathcal{C}$, there exists $l_0 > 0$, such that for all $l > l_0$ we have $1 - p_l^{\text{free}} < 1/d^8$, and hence

$$1 - p_l^{\text{free}} \le \binom{d^2}{2}(1 - p_{l-2}^{\text{free}})^2 \le d^4(1 - p_{l-2}^{\text{free}})^2 \le (1 - p_{l-2}^{\text{free}})^{1.5}.$$

By induction, $1 - p_l^{\text{free}} \le (1 - p_{l_0}^{\text{free}})^{(1.5)^{(l-l_0)/2}}$ which completes the proof. $\square$

REMARK. Claim 4.2 and Lemma 4.3 are the two main places where connectivity conditions are used: The first part of condition $\mathcal{C}$ is used in the construction of Γ'. It might be possible circumvented the assumption by using more carefully constructed paths. However, this would be purely technical and not the main interest of this paper. The second part of condition $\mathcal{C}$ is used to show that A_τ happens with high probability.

Note that Claim 4.2 implies that when restricted to A_τ, the fixed boundary Glauber dynamics on $B_{x,l/2}$ is irreducible as a subgraph of the Glauber dynamic on the larger block $B_{x,l}$. It is possible to replace the current connectivity condition by general assumptions bounding the probability of the later events directly.

Now we can complete the proof of Lemma 2.4, from which Theorem 1.2 follows immediately.

PROOF OF LEMMA 2.4. Let $C = \alpha(l/2 + 1 - \delta)^2/[(1-\delta)^2\mu^c(\sigma_x = c')]$ be the quantity on the left-hand side of (2.3). It is enough to show that there exist constants $l_1 \ge 2l_0$, $K \ge 1$ such that for all $l \ge l_1$

$$\Pr_{\tau\sim\mu^c}\left(\left|\mu^c(\sigma_x = c' \mid \sigma \sim \tau) - \mu^c(\sigma_x = c')\right| \ge \frac{K}{C}\right) \le e^{-2C/K}.$$

To see the sufficiency, note that this is just equation (2.2) with δ_1 satisfying $1-\delta_1 = \frac{1}{4K}(1-\delta)$.

Recall $A_\tau = \{\sigma \in \Omega^\tau_{B_{x,l}} : \forall x \in L_{x,l/2+2}, x \text{ is free in } \sigma\}$. Lemma 4.3 implies that for some constant $C_1 > 1$, $C_2 > 0$ and $l \ge 1$,

$$\begin{aligned}\mathbb{E}_{\tau\sim\mu^c}\big(\mu^{*,\tau}_{B_{x,l}}(A^c_\tau)\big) &= \mathbb{E}_{\tau\sim\mu^c}\big(\mu^c(\sigma \notin A_\tau \mid \sigma\sim\tau)\big)\\ &= \Pr_{\sigma\sim\mu^c}(\exists y \in L_{x,l/2+2}, y \text{ is not free})\\ &\le d^{l/2+2}(1 - p^{\text{free}}_{l/2-2}) \le C_2 d^{l/2+2}\exp(-C_1^{l/2-2}).\end{aligned}$$

By Markov's inequality,

$$\begin{aligned}\Pr_{\tau\sim\mu^c}\left(\mu^{*,\tau}_{B_{x,l}}(A^c_\tau) > \frac{1}{2C}\right) &\le 2C\mathbb{E}_{\tau\sim\mu^c}\big(\mu^{*,\tau}_{B_{x,l}}(A^c_\tau)\big)\\ &\le Cd^{l/2+2}C_2\exp\big(-C_1^{l/2-2}\big)\to 0\end{aligned}\tag{4.3}$$

as $l\to\infty$. In the event $\{\tau:\mu^{*,\tau}_{B_{x,l}}(A^c_\tau)\le\frac{1}{2C}\}$,

$$\begin{aligned}\mu^{*,\tau}_{B_{x,l}}(\sigma_x=c'\mid\sigma\in A_\tau) &\le \frac{\mu^{*,\tau}_{B_{x,l}}(\sigma_x=c')}{\mu^{*,\tau}_{B_{x,l}}(\sigma\in A_\tau)}\le\mu^{*,\tau}_{B_{x,l}}(\sigma_x=c')+\frac{1}{C},\\ \mu^{*,\tau}_{B_{x,l}}(\sigma_x=c'\mid\sigma\in A_\tau) &\ge \mu^{*,\tau}_{B_{x,l}}(\sigma_x=c',\sigma\in A_\tau)\ge\mu^{*,\tau}_{B_{x,l}}(\sigma_x=c')-\frac{1}{C}.\end{aligned}\tag{4.4}$$

Combining the two results together, we have

$$\big|\mu^{*,\tau}_{B_{x,l}}(\sigma_x=c')-\mu^{*,\tau}_{B_{x,l}}(\sigma_x=c'\mid\sigma\in A_\tau)\big|\le\frac{1}{C}.\tag{4.5}$$

Now splitting $\mu^{*,\tau}_{B_{x,l}}(\sigma_x=c'\mid\sigma\in A_\tau)$ according to $\sigma_{L_{x,l/2}}$ and applying Lemma 4.1, we have

$$\begin{aligned}&\mu^{*,\tau}_{B_{x,l}}(\sigma_x=c'\mid\sigma\in A_\tau)\\ &\quad=\sum_\eta\mu^{*,\tau}_{B_{x,l}}(\sigma_x=c'\mid\sigma\in A_\tau,\sigma_{L_{x,l/2}}=\eta)\mu^{*,\tau}_{B_{x,l}}(\sigma_{L_{x,l/2}}=\eta\mid\sigma\in A_\tau)\\ &\quad=\sum_\eta\mu^c(\sigma_x=c'\mid\sigma_{L_{x,l/2}}=\eta)\mu^{*,\tau}_{B_{x,l}}(\sigma_{L_{x,l/2}}=\eta\mid\sigma\in A_\tau).\end{aligned}\tag{4.6}$$

We would like to estimate the set of η such that $\mu^c(\sigma_x=c'\mid\sigma_{L_{x,l/2}}=\eta)$ has a large bias. Let

$$B=\left\{\eta:\big|\mu^c(\sigma_x=c'\mid\sigma_{L_{x,l/2}}=\eta)-\mu^c(\sigma_x=c')\big|\ge\frac{1}{C}\right\}.$$

Theorem 2.3 implies that for $l/2\ge l_0$ and some $\delta>0$, we have $\Pr_{\eta\sim\mu^c}(B)\le e^{-2C}$, where $\eta\sim\mu^c$ denotes the measure μ^c induced on $L_{x,l/2}$. Again by Markov's inequality,

$$\begin{aligned}\Pr_{\tau\sim\mu^c}\left(\mu^{*,\tau}_{B_{x,l}}(\sigma_{L_{x,l/2}}\in B)>\frac{1}{C}\right)&\le C\mathbb{E}_{\tau\sim\mu^c}\mu^{*,\tau}_{B_{x,l}}(\sigma_{L_{x,l/2}}\in B)\\ &=C\mu^c(B)\le Ce^{-2C}.\end{aligned}\tag{4.7}$$

On the event $\{\tau:\mu^{*,\tau}_{B_{x,l}}(\sigma_{L_{l/2}}\in B)\le\frac{1}{C}\}\cap\{\tau:\mu^{*,\tau}_{B_{x,l}}(A^c_\tau)\le\frac{1}{2C}\}$, from (4.6) we have

$$\begin{aligned}&\big|\mu^{*,\tau}_{B_{x,l}}(\sigma_x=c'\mid\sigma\in A_\tau)-\mu^c(\sigma_x=c')\big|\\ &\quad\le\sum_\eta\big|\mu^c(\sigma_x=c'\mid\sigma_{L_{l/2}}=\eta)-\mu^c(\sigma_x=c')\big|\mu^{*,\tau}_{B_{x,l}}(\sigma_{L_{l/2}}=\eta\mid\sigma\in A_\tau)\end{aligned}\tag{4.8}$$

$$\leq \sum_{\eta \in B^c} \frac{1}{C} \mu^{*,\tau}_{B_{x,l}}(\sigma_{L_{l/2}} = \eta \mid \sigma \in A_\tau) + \mu^{*,\tau}_{B_{x,l}}(\sigma_{L_{l/2}} \in B \mid \sigma \in A_\tau)$$

$$\leq \frac{1}{C} \cdot 1 + \frac{1}{C} + \frac{1}{C} = \frac{3}{C},$$

where the last inequality follows from similar argument to (4.4).

Combining the result of equations (4.5) and (4.8), on the event $\{\tau : \mu^{*,\tau}_{B_{x,l}}(\sigma_{L_{l/2}} \in B) \leq \frac{1}{C}\} \cap \{\tau : \mu^{*,\tau}_{B_{x,l}}(A^c_\tau) \leq \frac{1}{2C}\}$, we have

$$|\mu^{*,\tau}_{B_{x,l}}(\sigma_x = c') - \mu^c(\sigma_x = c')| \leq \frac{3}{C} + \frac{3}{C} = \frac{6}{C}.$$

Therefore, using the bounds from (4.3) and (4.7), for all $l \geq 2l_0$,

$$\Pr_{\tau \sim \mu^c}\left(|\mu^c(\sigma_x = c' \mid \sigma \sim \tau) - \mu^c(\sigma_x = c')| > \frac{6}{C}\right)$$

$$\leq \Pr\left(\mu^{*,\tau}_{B_{x,l}}(\sigma_{L_{l/2}} \in B) \leq \frac{1}{C}\right) + \Pr\left(\mu^{*,\tau}_{B_{x,l}}(A^c_\tau) \leq \frac{1}{2C}\right)$$

$$\leq C d^{l/2+2} C_2 \exp(-C_1^{l/2-2}) + C e^{-2C} \leq e^{-16C},$$

where the last step is true for large enough constant $\tilde{l}$ depending on d, C_1, C_2 and C'. This means that the strong concentration inequality (2.2) holds for $K = 6$, $\delta_1 = 1 - \frac{1}{4K}(1-\delta)$ and $l_1 = \max\{2l_0, \tilde{l}\}$. Moreover, by taking l large enough and changing the constant C to $6C$ in (4.5) and (4.8), we can make K arbitrarily close to 1. $\square$

5. Component dynamics version of fast mixing results. In this section, we prove Theorem 2.2. The theorem was originally proved for block dynamics in [12]. Here, we give a modification of their theorem adapted to the component dynamics by roughly "adding stars" at all occurrence of $B_{x,l}$. We will only state the key steps and refer the details to [12]. For the remainder of this section, we let $\mu = \mu^c_T$, $\Omega = \Omega^c_T$. Recall that $\tilde{T}_x = T_x \setminus \{x\}$. First, we define the entropy mixing condition for Gibbs measure to be the following.

DEFINITION (Entropy mixing). We say that μ satisfies $\mathrm{EM}^*(l, \varepsilon)$ if for every $x \in T$, $\eta \in \Omega$ and any $f \geq 0$ that does not depend on the connected component of $B_{x,l}$, that is, $f(\sigma) = \mu^{*,\sigma}_{B_{x,l}}(f)$, $\forall \sigma \in \Omega$, we have $\mathrm{Ent}^\eta_{T_x}[\mu_{\tilde{T}_x}(f)] \leq \varepsilon \cdot \mathrm{Ent}^\eta_{T_x}(f)$ where $\mathrm{Ent}^\eta_{T_x}$ means the entropy w.r.t. $\mu^\eta_{T_x}$.

Let $p_{\min} = \min_{c,c' \in [k]}\{M(c,c') : M(c,c') > 0\}$. By the Markov chain construction of configurations, it satisfies that $p_{\min} = \min_{x,c,c'}\{\mu^c_{T_x}(\sigma_x = c') : c, c' \text{ are compatible}\}$. The following theorem relates the entropy mixing condition to the log-Soblev constant.

THEOREM 5.1. *For any* l *and* $\delta > 0$, *if* μ *satisfies* $\mathrm{EM}^*(l, [(1-\delta)p_{\min}/(l+1-\delta)]^2)$ *then* $\mathrm{Ent}(f) \le \frac{2}{\delta} \cdot \mathcal{E}_l^*(f)$.

To prove Theorem 5.1, we need the following modification of Lemma 3.5(ii) of [12]. The proof follows from its analog in [12] immediately once we replace $\nu_A, \mathrm{Ent}_A, \nu_B, \mathrm{Ent}_B$ there with $\nu_{\tilde{T}_x}, \mathrm{Ent}_{\tilde{T}_x}, \nu^*_{B_{x,l}}, \mathrm{Ent}^*_{B_{x,l}}$, respectively.

LEMMA 5.2. *For any* $\varepsilon < p_{\min}^2$, *if* μ *satisfies* $\mathrm{EM}^*(l, \varepsilon)$ *then for every* $x \in T$, *any* $\eta \in \Omega$ *and any* $f \ge 0$ *we have* $\mathrm{Ent}^\eta_{T_x}[\mu_{\tilde{T}_x}(f)] \le \frac{1}{1-\varepsilon'} \cdot \mu^\eta_{T_x}[\mathrm{Ent}^*_{B_{x,l}}(f)] + \frac{\varepsilon'}{1-\varepsilon'} \cdot \mu^\eta_{T_x}[\mathrm{Ent}_{\tilde{T}_x}(f)]$ *with* $\varepsilon' = \sqrt{\varepsilon}/p_{\min}$.

Now plugging $\varepsilon = [(1-\delta)p_{\min}/(l+1-\delta)]^2$ into Lemma 5.2 verifies the hypothesis of the following claim, which then implies Theorem 5.1:

CLAIM 5.3. *If for every* $x \in T$, $\eta \in \Omega$ *and any* $f \ge 0$,

$$\mathrm{Ent}^\eta_{T_x}[\mu_{\tilde{T}_x}(f)] \le c \cdot \mu^\eta_{T_x}[\mathrm{Ent}^*_{B_{x,l}}(f)] + \frac{1-\delta}{l} \cdot \mu^\eta_{T_x}[\mathrm{Ent}_{\tilde{T}_x}(f)], \tag{5.1}$$

then $\mathrm{Ent}(f) \le \frac{c}{\delta} \cdot \mathcal{E}_l^*(f)$ *for all* $f \ge 0$.

PROOF. First, we decompose $\mathrm{Ent}(f)$ as a sum of $\mathrm{Ent}^\eta_{T_x}[\mu_{\tilde{T}_x}(f)]$. Suppose T have m levels, consider $\varnothing = F_0 \subset F_1 \subset \cdots \subset F_{m+1} = T$, where F_i is the lowest i levels of T. By basic properties of conditional entropy (equation (3), (4), (5) of [12]) and Markov's property of Gibbs measure, we have

$$\begin{aligned} \mathrm{Ent}(f) &= \cdots = \sum_{i=1}^{m+1} \mu[\mathrm{Ent}_{F_i}(\mu_{F_{i-1}}(f))] \\ &\le \sum_{i=1}^{m+1} \sum_{x \in F_i \setminus F_{i-1}} \mu[\mathrm{Ent}_{T_x}(\mu_{F_{i-1}}(f))] \le \sum_{x \in T} \mu[\mathrm{Ent}_{T_x}(\mu_{\tilde{T}_x}(f))]. \end{aligned} \tag{5.2}$$

Denote the final sum by $\mathrm{PEnt}(f)$. For each term in the sum of $\mathrm{PEnt}(f)$, apply (5.1) to $g = \mu_{T_x \setminus B_{x,l} \cup \partial B_{x,l}}(f)$ and perform the decomposition trick of (5.2) again, we have for every $x \in T$ and $\eta \in \Omega$ that

$$\begin{aligned} &\mathrm{Ent}^\eta_{T_x}[\mu_{\tilde{T}_x}(f)] \\ &\quad = \mathrm{Ent}^\eta_{T_x}[\mu_{\tilde{T}_x}(g)] \\ &\quad \le c \cdot \mu^\eta_{T_x}[\mathrm{Ent}^*_{B_{x,l}}(g)] + \frac{1-\delta}{l} \cdot \mu^\eta_{T_x}[\mathrm{Ent}_{\tilde{T}_x}(g)] \\ &\quad \le c \cdot \mu^\eta_{T_x}[\mathrm{Ent}^*_{B_{x,l}}(f)] + \frac{1-\delta}{l} \cdot \sum_{y \in B_{x,l} \cup \partial B_{x,l} \setminus \{x\}} \mu^\eta_{T_x}[\mathrm{Ent}_{T_y}(\mu_{\tilde{T}_y}(f))]. \end{aligned}$$

Now sum over $x \in T$ and take expectation w.r.t. μ for $\eta \in \Omega$. Note that the first term of the last line sums up to $\mathcal{E}_l^* = \sum_{x\in T} \mu(\mathrm{Ent}^*_{B_{x,l}}(f))$ and each y in second term appears in at most l blocks, we have

$$\begin{aligned}
\mathrm{PEnt}(f) &\le c \cdot \mathcal{E}_l^*(f) + \frac{1-\delta}{l} \cdot \sum_{x\in T} \sum_{y \in B_{x,l} \cup \partial B_{x,l} \setminus \{x\}} \mu\big[\mathrm{Ent}_{T_y}\big(\mu_{\tilde{T}_y}(f)\big)\big] \\
&\le c \cdot \mathcal{E}_l^*(f) + \frac{1-\delta}{l} \cdot l \cdot \sum_{y\in T} \mu\big[\mathrm{Ent}_{T_y}\big(\mu_{\tilde{T}_y}(f)\big)\big] \\
&= c \cdot \mathcal{E}_l^*(f) + (1-\delta) \cdot \mathrm{PEnt}(f),
\end{aligned}$$

and hence $\mathrm{Ent}(f) \le \mathrm{PEnt}(f) \le \frac{c}{\delta} \cdot \mathcal{E}_l^*$. $\square$

Given the result of Theorem 5.1, it is enough to show that for some constant α, concentration inequality of (2.2) implies $\mathrm{EM}^*(l, [(1-\delta)p_{\min}/(l+1-\delta)]^2)$. For convenience of notation, we define the two following functions for each $c' \in [k]$:

$$g_{c'}(\sigma) = \frac{\mu(\sigma \mid \sigma_\rho = c')}{\mu(\sigma)} = \frac{1}{\mu(\sigma_\rho = c')} \cdot 1\{\sigma_\rho = c'\}, \qquad g_{c'}^{*(l)} = \mu^*_{B_{\rho,l}}(g_{c'}).$$

Letting $\delta' = (1-\delta)^2/\alpha(l+1-\delta)^2$, we can rewrite (2.2) as

$$\mu\big(\big|g_{c'}^{*(l)} - 1\big| > \delta'\big) \le e^{-2/\delta'}. \tag{5.3}$$

THEOREM 5.4. *There exists a constant C such that if (5.3) holds for some $\delta' \ge 0$ and all pairs of states $c, c' \in [k]$, we have $\mathrm{Ent}[\mu_{\tilde{T}}(f)] \le C\delta' \,\mathrm{Ent}(f)$ for any $f \ge 0$ satisfying $f(\sigma) = \mu^{*,\sigma}_{B_{\rho,l}}(f)$, $\forall \sigma \in \Omega^c$, that is, $\mathrm{EM}^*(l, C\delta')$ holds.*

PROOF. Since for any $f' \ge 0$, $\mathrm{Ent}(f') \le \mathrm{Var}(f')/\mu(f)$, we can write

$$\begin{aligned}
\mathrm{Ent}\big[\mu_{\tilde{T}}(f)\big] &\le \frac{\mathrm{Var}[\mu_{\tilde{T}}(f)]}{\mu(\mu_{\tilde{T}}(f))} \\
&= \frac{1}{\mu(f)} \sum_{c'\in[k]} \mu(\sigma_\rho = c')\big(\mu(f \mid \sigma_\rho = c') - \mu(f)\big)^2 \\
&= \frac{1}{\mu(f)} \sum_{c'\in[k]} \mu(\sigma_\rho = c') \,\mathrm{Cov}(g_{c'}, f)^2 \\
&\le \max_{c'\in[k]} \frac{\mathrm{Cov}(g_{c'}, f)^2}{\mu(f)} = \max_{c'\in[k]} \frac{\mathrm{Cov}(g_{c'}^{*(l)}, f)^2}{\mu(f)},
\end{aligned} \tag{5.4}$$

where covariance is taken w.r.t. μ and the last step is because $f(\sigma) = \mu^{*,\sigma}_{B_{\rho,l}}(f)$. Now using Lemma 5.4 of [12] (cited below) with

$$f_1 = \big(g_{c'}^{*(l)} - 1\big)/\big\|g_{c'}^{*(l)}\big\|_\infty, \qquad f_2 = f/\mu(f)$$

and noting that $\|g_{c'}^{*(l)}\|_\infty \le \|g_{c'}\|_\infty \le p_{\min}$, we have

$$\mathrm{Cov}\big(g_{c'}^{*(l)}, f\big)^2 \le C\delta'\mu(f)\,\mathrm{Ent}(f)$$

for some constant $C = C'/p_{\min}^2$. Plug it into (5.4), we get $\mathrm{Ent}[\mu_{\tilde{T}}(f)] \le C\delta'\,\mathrm{Ent}(f)$. $\square$

LEMMA 5.5 (Lemma 5.4 of [12]). *Let $\{\Omega, \mathcal{F}, \nu\}$ be a probability space and let f_1 be a mean-zero random variable such that $\|f\|_\infty \le 1$ and $\nu[|f_1| > \delta] \le e^{-2/\delta}$ for some $\delta \in (0,1)$. Let f_2 be a probability density w.r.t. ν, that is, $f_2 \ge 0$ and $\nu(f_2) = 1$. Then there exists a numerical constant $C' > 0$ independent of ν, f_1, f_2 and δ, such that $\nu(f_1 f_2) \le C'\delta\,\mathrm{Ent}_\nu(f_2)$.*

PROOF OF THEOREM 2.2. Fix $\alpha = C/p_{\min}^2$ where C is the constant in Theorem 5.4. The desired result follows the combination of Theorems 5.1 and 5.4. $\square$

6. Results for *k*-coloring. In this section, we prove Theorem 1.1, for which it is enough to verify the connectivity condition $\mathcal{C}$, in particular to show that $p_l^{\mathrm{free}} \to 1$, as $l \to \infty$. In fact for the coloring model, as we will show in a moment, a vertex can change to all k states in one step if all its children can change to 2 or 3 states in one step. We will first formalize this idea by defining the "types" of vertices and then analyze the recursion of this new definition.

Recall the definition that for given configuration $\sigma \in \Omega_T$ with $\sigma_\rho = c$, we say that the root can change to color c' in one step if and only if there exists a path $\sigma = \sigma^0, \sigma^1, \dots, \sigma^n \in \Omega_T^\sigma$ such that for each i, σ^i, σ^{i+1} differs by only one vertex and

$$\sigma_\rho^i = \begin{cases} c, & 0 \le i \le n-1, \\ c', & i = n. \end{cases}$$

Let $C(\rho)$ denote the set of colors the root can change to in one step (including its original color). We define the type of root to be *rigid* (*type 2*, *type 3*, resp.) if $|C(\rho)| = 1$ (=2, $\ge$3, resp.). For general vertex $x \in T$, not necessarily the root, we can similarly define $C(x)$ and rigid, type 2, type 3 by treating x as the root of subtree T_x and considering $\sigma|_{T_x}$. Set $C(x)$ is a function of σ_{T_x} and is independent of the rest of the tree.

Let $p_l^r = \mu_l$(the root is rigid), where μ_l is the Gibbs measure on l-level tree with free boundary condition. Define $p_l^{(2)}$, $p_l^{(3)}$ and similarly we have $p_l^r + p_l^{(2)} + p_l^{(3)} = 1$. For tree T with $l' > l$ levels and vertex $x \in T$, that is, l levels above the bottom boundary, noting that $\mu_{l'}|_{T_x} = \mu_l$, we have

$$\mu_{l'}(x \text{ is rigid}) = \mu_{l'}|_{T_x}(x \text{ is rigid}) = p_l^r,$$

and the same goes for type 2, type 3 and p_l^2, p_l^3, respectively.

Observe that the definition above is independent of the parent of x. In order to analyze these probabilities recursively, we introduce one further definition describing how the type of one vertex affects the type of its parent. Recall that $p(x)$ denotes the parent of x. Fix a configuration $\sigma \in \Omega_T$. For any $x \in \tilde{T} = T \setminus \{\rho\}$, we say x is *bad* if $C(x) \setminus \{\sigma_{p(x)}\} = \{\sigma_x\}$ and *good* other wise. Observe that σ_x is always an element of $C(x)$. If x is good, then $|C(x) \setminus \{\sigma_{p(x)}\}| \geq 2$, that is, x has at least one more choice other than $\sigma_{p(x)}$. Note that the event that x is bad depends only on $\sigma|_{T_{p(x)}}$ and given σ_x, for $x_i \in L_{x,1}$, events $\{x_i \text{ is bad}\}$ are conditionally i.i.d. and independent of the configurations outside T_x. Hence, by similar argument, we can define $p_l^b = 1 - p_l^g = \mu_{l'}(x \text{ is bad})$. The relation between the type of a vertex and its goodness/badness is given in the following lemma.

LEMMA 6.1. *For $l' > l > 0$ and $x \in T$ l levels above the bottom boundary,*

$$\begin{aligned} &\mu_{l'}(x \textit{ is bad} \mid x \textit{ is rigid}) = 1, \\ &\mu_{l'}(x \textit{ is bad} \mid x \textit{ is type } 2) = \frac{1}{k-1}, \qquad (6.1) \\ &\mu_{l'}(x \textit{ is bad} \mid x \textit{ is type } 3) = 0. \end{aligned}$$

Hence, $p_l^b = p_l^r + \frac{1}{k-1} p_l^{(2)}$, $p_l^g = p_l^{(3)} + \frac{k-2}{k-1} p_l^{(2)}$.

PROOF. The first and third equations of (6.1) is obvious as $|C(x)|$ and $|C(x) \setminus \{\sigma_{p(x)}\}|$ differs at most by one, and the equality about p_l^b and p_l^g follows immediately from (6.1). Hence, it lefts to show the second equation. Given $|C(x)| = 2$, x is bad if and only if $\sigma_{p(x)} \in C(x)$. Therefore, the conditional probability on the left-hand side of the second equation equals to $\Pr(C(x) = \{\sigma_{p(x)}, \sigma_x\} \mid |C(x)| = 2)$.

Note that $C(x)$ is a function of σ_{T_x}, in particular it is conditionally independent of $\sigma_{p(x)}$ given σ_x. By symmetry, the distribution of $C(x) \setminus \{\sigma_x\}$ given $|C(x)|$ and σ_x is the uniformly distribution on the $\binom{k-1}{|C(x)|-1}$ ways of choosing $|C(x)| - 1$ elements from $[k] \setminus \{\sigma_x\}$. Hence,

$$\Pr(C(x) = \{\sigma_{p(x)}, \sigma_x\} \mid |C(x)| = 2) = \frac{1}{\binom{k-1}{1}} = \frac{1}{k-1}. \qquad \square$$

The next lemma follows a similar argument to Claim 4.2 and Lemma 4.3, and shows that in order to bound the probability of a vertex being free, it is enough to bound the probability of being bad.

LEMMA 6.2. *Suppose $k \geq 4$. For any $\sigma \in \Omega_T$ and $x \in T$, if every child of x is good, then x is free.*

PROOF. Fix $c \in [k]$. Since all children of x are god, for each child y_i there exists $c_i \in C(y_i) \setminus \{c, \sigma_x\}$. Therefore, to change x from σ_x to c in one step, we can first change the color of every y_i to c_i in one step and then in the final step change x from σ_x to c. Since this is true for all $c \in [l]$, we conclude that x is free. $\square$

Now we will show that for large enough k, in the nonreconstruction regime, the probability of seeing a bad vertex l levels above bottom decays double exponentially fast in l. In fact, we will prove the result for a region slightly larger than the known nonreconstruction region, which is $d \le k[\log k + \log\log k + \beta]$, for any $\beta < 1 - \ln 2$ (see [17]).

THEOREM 6.3. *Suppose $\beta < 1$, For sufficiently large k and $d \le k[\log k + \log\log k + \beta]$, there exists a constant l_0 depending only on k and d, such that for $l \ge l_0$,*

$$p_l^b \le \exp\bigl(-(k/2)^{l-l_0}\bigr). \tag{6.2}$$

We first finish the proof of Theorem 1.1 using Lemma 6.2 and Theorem 6.3.

PROOF OF THEOREM 1.1. It has been shown in [17] that for any $\beta < 1 - \ln 2$, there exist $k_0 = k_0(\beta)$ such that for any $k \ge k_0$ and $d \le k(\log k + \log\log k + \beta)$, the k-coloring model is nonreconstructible on d-ary trees. Therefore by Theorem 1.2, it is enough to show that the connectivity condition holds. The first part of the condition is obviously true for $k \ge 4$. For the second condition,

$$\begin{aligned} 1 - p_l^{\mathrm{free}} &= \Pr_{\sigma\sim\mu_l}(\text{root is not free}) \le \Pr(\exists x \in L_1, x \text{ is bad}) \\ &\le d p_{l-1}^b \le d\exp\biggl(-\Bigl(\frac{1}{2}k\Bigr)^{l-l_0}\biggr). \end{aligned}$$

The last term in the equation above tends to 0 as l tends to infinity, which completes the proof. $\square$

The proof of Theorem 6.3 is split into two phases: when p_l^b is close to 1 and when p_l^b is smaller than $\frac{1}{ed}$.

LEMMA 6.4. *Under the assumption of Theorem 6.3, there exist a constant l_0 depending only on k and d such that $p_{l_0}^b < \frac{1}{ed}$.*

PROOF. This proof is similar to Lemma 2 and Lemma 4 of [17]. We recursively analyze the probabilities as a function of the depth of the tree l. For $l = 0$, T consist only the bottom boundary, and hence $p_0^r = 1$, $p_0^{(2)} = p_0^{(3)} = 0$, $p_0^b = p_0^r + \frac{1}{k-1}p_0^{(2)} = 1$.

For $l \ge 1$, suppose without loss of generality that the color of the root is 1 and its children are $x_1, \dots, x_d \in L_1$. Let $\mathcal{F}$ denote the sigma-field generated by $(\sigma_{x_i})_{i=1}^d$ and let $d_c = |\{i, \sigma_{x_i} = c\}|$ be the number of children with color c for $2 \le c \le k$.

By definition, the sizes of $C(x_i)$'s, and hence the types of x_i's are independent of $\mathcal{F}$ and i.i.d. distributed. Conditioning on $\mathcal{F}$ and $(|C(x_i)|)_{i=1}^k$, set $C(x_i) \setminus \{\sigma_{x_i}\}$ is uniformly randomly chosen among all subsets of $[k] \setminus \{\sigma_{x_i}\}$ with $(|C(x_i)| - 1)$ elements. Therefore, the number of bad vertices of color c given $\mathcal{F}$ is follows the binomial distribution with parameter $\mathrm{Bin}(d_c, p_{l-1}^b)$.

Following the similar argument of Lemma 6.2, the root can change to color c in one step if and only if none of the x_i's with color c is bad, which happens with probability $(1 - p_{l-1}^b)^{d_c}$. Therefore, we have

$$\begin{aligned}
p_l^b &= p_l^r + \frac{1}{k-1} p_l^{(2)} \\
&= \prod_{c=2}^{k} \mathbb{E}[1 - (1 - p_{l-1}^b)^{d_c}] \\
&\quad + \frac{1}{k-1} \sum_{c'=2}^{k} \mathbb{E}\Bigl[(1 - p_{l-1}^b)^{d_{c'}} \prod_{c \neq c'} (1 - (1 - p_{l-1}^b)^{d_c})\Bigr].
\end{aligned}$$

Viewing the right-hand side as a function of $(d_2, \dots, d_k)$, increasing d_c means adding more vertices of color c, which increases the probability of blocking the move of the root. Therefore, p_l^b is an increasing function w.r.t. every d_c. By symmetry, $(d_2, \dots, d_k)$ follows a multi-nominal distribution. Fix $\beta < \beta^* < 1$ and let $\tilde{d}_c$ be i.i.d. Poisson(D) random variables where $D = \log k + \log\log k + \beta^*$. We can couple $(d_2, \dots, d_k)$ and $(\tilde{d}_2, \dots, \tilde{d}_k)$ such that $(d_2, \dots, d_k) \le (\tilde{d}_2, \dots, \tilde{d}_k)$ whenever $\sum_{c=2}^k \tilde{d}_c \ge d$. Letting $p = \Pr(\mathrm{Poisson}((k-1)D) < d)$, the recursion relationship satisfies

$$\begin{aligned}
p_l^b &= p_l^r + \frac{1}{k-1} p_l^{(2)} \\
&= \prod_{c=2}^{k} \mathbb{E}[1 - (1 - p_{l-1}^b)^{d_c}] \\
&\quad + \frac{1}{k-1} \sum_{c'=2}^{k} \mathbb{E}\Bigl[(1 - p_{l-1}^b)^{d_{c'}} \prod_{c \neq c'} [1 - (1 - p_{l-1}^b)^{d_c}]\Bigr] \\
&\le \prod_{c=2}^{k} \mathbb{E}[1 - (1 - p_{l-1}^b)^{\tilde{d}_c}] \\
&\quad + \frac{1}{k-1} \sum_{c'=2}^{k} \mathbb{E}(1 - p_{l-1}^b)^{\tilde{d}_{c'}} \prod_{c \neq c'} \mathbb{E}[1 - (1 - p_{l-1}^b)^{\tilde{d}_c}] + p \\
&= (1 - \exp(-p_{l-1}^b D))^{k-1}
\end{aligned}$$

$$+\frac{k-1}{k-1}\exp(-p^b_{l-1}D)\bigl(1-\exp(-p^b_{l-1}D)\bigr)^{k-2}+p$$
$$=\bigl(1-\exp(-p^b_{l-1}D)\bigr)^{k-2}+p\le\exp\bigl(-(k-2)\exp(-p^b_{l-1}D)\bigr)+p,$$

where the last step follows from the fact that $(1-r)^k\le e^{-kr}$ for $0<r<1$.

The rest of the proof resembles the argument of Lemma 3 of [17]. Let $f(x)=\exp(-(k-2)\exp(-xD))+p$, $y_0=p^b_0=1$ and recursively define $y_l=f(y_{l-1})$. Since $f(x)$ is an increasing function of x, we have that $p^b_l\le y_l$ for any $l\ge 0$. Hence, it is enough to show the existence of l_0 such that $y_{l_0}\le\frac{1}{ed}$.

Note that $\frac{d}{dx}\exp(-x)\mid_{x=0}=-1$. For any sufficiently small $\varepsilon>0$, there exists $\delta>0$ such that for any $0<x<\delta$, $e^{-x}\le 1-(1-\varepsilon)x$. Let k be large enough such that $(k-2)\exp(-D)=\frac{k-2}{k\log k}e^{-\beta^*}<\delta$. We have

$$y_1=f(1)\le 1-(1-\varepsilon)\frac{k-2}{k\log k}e^{-\beta^*}+p.$$

Recall our choice of $\beta<\beta^*<1$ and $(k-1)D-d\ge(\beta^*-\beta)k+o(k)$, by Hoeffding's inequality, the error term p satisfies that $p=\exp(-\Omega(\frac{k}{\sqrt{d}}))=o(k^{-2})=o(d^{-1})$. Therefore, for large enough k,

$$y_1\le 1-\frac{1-\varepsilon}{2e\log k}+o\bigl(k^{-1}\bigr)\le 1=y_0.$$

Repeating the arguments above shows that y_l is decreasing in l as long as $(k-2)\exp(-y_lD)<\delta$. Pick ε small enough such that $(1-\varepsilon)e^{-\beta^*}>e^{-1}$ and choose $r'>r>0$ such that $(1-\varepsilon)e^{-\beta^*}>e^{-1}(1+r')$. It follows that

$$\begin{aligned}1-y_{l+1}&\ge 1-\bigl(p+1-(1-\varepsilon)(k-2)\exp(-y_lD)\bigr)\\&\ge(1-\varepsilon)\frac{(k-2)e^{-\beta^*}}{k\log k}\exp\bigl((1-y_l)\log k\bigr)-p\\&\ge\frac{k-2}{k}(1-\varepsilon)e^{1-\beta^*}(1-y_l)-p\\&\ge\bigl(1+r'\bigr)(1-y_l)-p\ge(1+r)(1-y_l),\end{aligned}$$

where the second-last inequality follows from inequality $e^x>ex$, and the last inequality follows from that $1-y_l\ge 1-y_1=O(\frac{1}{\log k})$ while $p=o(k^{-2})$. Therefore, after a constant number of steps, there must exist some l such that $(k-2)\exp(-y_lD)\ge\delta$. Now choose α,α' such that $e^{-\delta}<\alpha'<\alpha<1$. When k is large enough, $y_{l+1}\le p+e^{-\delta}<\alpha'<1$. Then again for k large enough, $\exp(-y_{l+1}D)\ge\exp(-\alpha'D)\ge\exp(-\alpha\log k)=k^{-\alpha}$. Therefore, for k large enough,

$$y_{l+2}\le p+\exp\bigl(-(k-2)\exp(-y_{l+1}D)\bigr)\le p+\exp\biggl(-\frac{1}{2}k^{1-\alpha}\biggr)\le\frac{1}{ed}.\qquad\square$$

After first l_0 levels, we cannot use the same method because the error of Poisson coupling becomes nonnegligible; but meanwhile, p_l^b is small enough such that bounding the total number of bad children is enough to complete the proof.

PROOF OF THEOREM 6.3. In order for a vertex to be bad, there must be at least $k-2$ of its children which are bad. Therefore,

$$p_l^b \le \binom{d}{k-2} (p_{l-1}^b)^{k-2} \le (d p_{l-1}^b)^{k-2}.$$

Let l_0 be the constant in Lemma 6.4. We complete the proof by inducting on l for $l \ge l_0$: If $l = l_0$, then $p_{l_0}^b \le \frac{1}{ed} \le \frac{1}{e}$. If for $l > l_0$, p_l^b satisfies (6.2), then for k large enough such that $\log(2k \log k) \le \frac{1}{6}k$ and $k-2 \ge \frac{3}{4}k$,

$$\begin{aligned} p_{l+1}^b &\le (d p_l^b)^{k-2} \le [2k \log k \exp(-(k/2)^{l-l_0})]^{k-2} \\ &= \exp\{(k-2)[-(k/2)^{l-l_0} + \log(2k \log k)]\} \\ &\le \exp\left\{-\frac{3}{4}k \cdot \frac{2}{3}(k/2)^{l-l_0}\right\} = \exp\{-(k/2)^{l+1-l_0}\}. \end{aligned}$$

Therefore, (6.2) holds for all $l \ge l_0$. □

REFERENCES

[1] BERGER, N., KENYON, C., MOSSEL, E. and PERES, Y. (2005). Glauber dynamics on trees and hyperbolic graphs. *Probab. Theory Related Fields* **131** 311–340. MR2123248

[2] BHATNAGAR, N., BHATNAGAR, N., VERA, J., VERA, J., VIGODA, E., VIGODA, E., WEITZ, D. and WEITZ, D. (2011). Reconstruction for colorings on trees. *SIAM J. Discrete Math*. **25** 809–826. MR2817532

[3] BORGS, C., CHAYES, J., MOSSEL, E. and ROCH, S. (2006). The Kesten–Stigum reconstruction bound is tight for roughly symmetric binary channels. In *Proceedings of the* 47*th Annual IEEE Symposium on Foundations of Computer Science* 518–530. IEEE, New York.

[4] DING, J., LUBETZKY, E. and PERES, Y. (2010). Mixing time of critical Ising model on trees is polynomial in the height. *Comm. Math. Phys*. **295** 161–207. MR2585995

[5] FRIEZE, A. and VIGODA, E. (2007). A survey on the use of Markov chains to randomly sample colourings. In *Combinatorics, Complexity, and Chance*. *Oxford Lecture Ser. Math. Appl*. **34** 53–71. Oxford Univ. Press, Oxford. MR2314561

[6] GEORGII, H.-O. (2011). *Gibbs Measures and Phase Transitions*, 2nd ed. *De Gruyter Studies in Mathematics* **9**. de Gruyter, Berlin. MR2807681

[7] GOLDBERG, L. A., JERRUM, M. and KARPINSKI, M. (2010). The mixing time of Glauber dynamics for coloring regular trees. *Random Structures Algorithms* **36** 464–476. MR2666764

[8] JANSON, S. and MOSSEL, E. (2004). Robust reconstruction on trees is determined by the second eigenvalue. *Ann. Probab*. **32** 2630–2649. MR2078553

[9] JONASSON, J. (2002). Uniqueness of uniform random colorings of regular trees. *Statist. Probab. Lett*. **57** 243–248. MR1912082

[10] LUCIER, B., MOLLOY, M. and PERES, Y. (2009). The Glauber dynamics for colourings of bounded degree trees. In *Approximation, Randomization, and Combinatorial Optimization*. *Lecture Notes in Computer Science* **5687** 631–645. Springer, Berlin. MR2551034

[11] MARTINELLI, F. (1999). Lectures on Glauber dynamics for discrete spin models. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1997). *Lecture Notes in Math*. **1717** 93–191. Springer, Berlin. MR1746301

[12] MARTINELLI, F., SINCLAIR, A. and WEITZ, D. (2004). Glauber dynamics on trees: Boundary conditions and mixing time. *Comm*. *Math*. *Phys*. **250** 301–334. MR2094519

[13] MARTINELLI, F., SINCLAIR, A. and WEITZ, D. (2007). Fast mixing for independent sets, colorings, and other models on trees. *Random Structures Algorithms* **31** 134–172. MR2343716

[14] MOSSEL, E. (2004). Survey: Information flow on trees. In *Graphs*, *Morphisms and Statistical Physics*. *DIMACS Ser. Discrete Math*. *Theoret*. *Comput*. *Sci*. **63** 155–170. Amer. Math. Soc., Providence, RI. MR2056226

[15] MOSSEL, E. and PERES, Y. (2003). Information flow on trees. *Ann*. *Appl*. *Probab.* **13** 817–844. MR1994038

[16] SALOFF-COSTE, L. (1997). Lectures on finite Markov chains. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1996). *Lecture Notes in Math*. **1665** 301–413. Springer, Berlin. MR1490046

[17] SLY, A. (2009). Reconstruction of random colourings. *Comm*. *Math*. *Phys*. **288** 943–961. MR2504861

[18] SLY, A. and ZHANG, Y. (2016). The reconstruction threshold for graph coloring is strictly below the rigidity threshold. Preprint. Available at arXiv:1610.02770.

[19] TETALI, P., VERA, J. C., VIGODA, E. and YANG, L. (2012). Phase transition for the mixing time of the Glauber dynamics for coloring regular trees. *Ann*. *Appl*. *Probab.* **22** 2210–2239. MR3024967

[20] VIGODA, E. (2000). Improved bounds for sampling colorings. *J*. *Math*. *Phys*. **41** 1555–1569. MR1757969

[21] WEITZ, D. (2004). *Mixing in Time and Space for Discrete Spin Systems*. Ph.D. thesis, Univ. California, Berkeley. ProQuest LLC, Ann Arbor, MI. MR2706464

MATHEMATICAL SCIENCES INSTITUTE
AUSTRALIAN NATIONAL UNIVERSITY
JOHN DEDMAN BUILDING 27
UNION LANE
CANBERRA
AUSTRALIAN CAPITAL TERRITORY 2601
AUSTRALIA
AND
DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA, BERKELEY
367 EVANS HALL
BERKELEY, CALIFORNIA 94720
USA
E-MAIL: sly@stat.berkeley.edu

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA, BERKELEY
367 EVANS HALL
BERKELEY, CALIFORNIA 94720
USA
E-MAIL: ymzhang@berkeley.edu