The Glauber dynamics for colourings of bounded degree trees

B. Lucier* M. Molloy†

August 1, 2010

Abstract

We study the Glauber dynamics Markov chain for k-colourings of trees with maximum degree Δ. For $k \geq 3$, we show that the mixing time on the complete tree is $n^{\ (1+\Delta=(k \log \Delta))}$. For $k \geq 4$ we extend our analysis to show that the mixing time on *any* tree is at most $n^{O(1+\Delta=(k \log \Delta))}$. Our proof uses a weighted canonical paths analysis and introduces a variation of the block dynamics that exploits the differing relaxation times of blocks.

1 Introduction

The Glauber dynamics is a Markov chain over configurations of spin systems on graphs, of which k-colourings is a special case. Such chains have generated a great deal of interest in both statistical physics and computer science. In computer science, counting k-colourings is a fundamental #P-hard problem, and Markov chains that sample colourings can be used to obtain an FPRAS to approximately count them. In statistical physics, k-colourings are equivalent to the antiferromagnetic Potts model, and single-site update chains can be used to model how such physical systems arrive at equilibrium. The Glauber dynamics has received a very large part of this interest [12]. It is particularly appealing because it is a natural and simple process that underlies more substantial procedures such as block dynamics and systematic scan [12, 7]. It is also commonly used in practice for physical system simulations.

The focus of this paper will be the performance of the Glauber dynamics on trees. Of course, the task of sampling a k-colouring of a tree is not particularly difficult. Nevertheless, people have studied the Glauber dynamics on trees as a means of understanding its performance on more general graphs, and because the behaviour of the Glauber dynamics on trees is particularly relevant to the understanding of equilibrium concepts such as the uniqueness threshold for the infinite-volume Gibbs distribution [4, 11, 13] and the reconstruction threshold [3]. Such concepts have gained particular interest due to recent proposed connections to algorithmic barriers related to local search algorithms on trees and sparse random graphs [1].

Berger et al. [2] showed that the Glauber dynamics mixes in polynomial time on complete trees of maximum degree Δ, and Martinelli et al. [13] showed that this polynomial is $O(n \log n)$ so long as $k \geq \Delta+2$. Hayes, Vera and Vigoda [9] showed that the Glauber dynamics mixes in polynomial time for all planar graphs if $k \geq C\Delta/\log\Delta$ for a particular constant C. They remarked that this was best possible, up to the value of C: The chain takes superpolynomial time to mix on every tree when $k = o(\Delta/\log n)$, and hence trees with $\Delta \geq n^{\epsilon}$ provide lower-bound examples for any constant ϵ. They asked whether such examples exist for smaller values of Δ; in particular, what is the mixing time for the complete $(\Delta-1)$-ary tree with $k = 3$ and $\Delta = O(1)$?

*Dept of Computer Science, University of Toronto, blucier@cs.toronto.edu.
†Dept of Computer Science, University of Toronto, molloy@cs.toronto.edu.

Proposition 2.5 of Berger et al. [2] shows that the mixing time is polynomial for every constant $k \geq 3$ and $\Delta \geq 2$ (in fact, it shows this for general particle systems on trees for which the Glauber dynamics is ergodic, of which proper colouring is a special case), but did not analyze the degree of this polynomial. Our main result is an upper bound for *every* tree when $k \geq 4$. Our bound is asymptotically tight, matching the lower bound up to a constant factor in the degree.

Theorem 1.1. *For $k \geq 4$, the Glauber dynamics on k-colourings of any tree with maximum degree Δ mixes in time at most $n^{O(1+\Delta/k\log\Delta)}$.*

Thus, for every $k \geq 4$ and $\Delta = O(1)$, we have polytime mixing on every tree. However, if Δ grows with n, no matter how slowly, then on some trees (eg. complete trees) we require the $\Omega(\Delta/\log\Delta)$ colours for polytime mixing that Hayes, Vera and Vigoda noted were required at $\Delta = n^{\epsilon}$.

For the case of $k = 3$ colours, we demonstrate the same mixing time bound for complete trees.

Theorem 1.2. *For $k \geq 3$, the Glauber dynamics on k-colourings of the complete tree with maximum degree Δ mixes in time at most $n^{O(1+\Delta/k\log\Delta)}$.*

We also present a matching lower bound, demonstrating that these results are the best possible up to the constant factor in the exponent. The complete tree with maximum degree Δ provides a lower-bound example:

Theorem 1.3. *For $k \geq 3$, the Glauber dynamics on k-colourings of the complete $(\Delta - 1)$-ary tree mixes in time $n^{\Omega(1+\Delta/k\log\Delta)}$.*

Independently, Goldberg, Jerrum and Karpinski [8] showed a lower bound of $n^{\Omega(1+\Delta/k\log\Delta)}$ and an upper bound of $n^{O(1+\Delta/\log\Delta)}$ on the mixing time of the Glauber dynamics on a complete tree. Subsequently, Tetali et al. [17] extended our results to demonstrate that the mixing time undergoes a phase transition at a threshold $k = (1 \pm o(1))\Delta/\log\Delta$: they show that the mixing time is $O(n^{1+o_\Delta(1)}\ln^2 n)$ when $k > (1 + o(1))\Delta/\log\Delta$, and that it lies in $O(n^{\Delta/k\log\Delta+o_\Delta(1)}\ln^2 n)$ and $\Omega(n^{\Delta/k\log\Delta-o_\Delta(1)})$ when $k < (1 - o(1))\Delta/\log\Delta$. Additionally, Sly [16] has given very tight bounds on the reconstruction threshold for colourings of a tree, which occurs at $k = (1 \pm o(1))\Delta/\log\Delta$.

Our results lie in the regime where k is very small. Let us describe the analytical difficulties that occur when k lies below the reconstruction threshold; that is, when $k = o(\Delta/\log\Delta)$. If $k \geq \Delta + 2$ then no vertex will ever be *frozen*; i.e. there will always be at least one colour that it can switch to. This bound on k also corresponds to the threshold for unique infinite-volume Gibbs distributions [11]. Much of the difficulty in showing rapid mixing for smaller values of k is in dealing with frozen vertices. From this perspective, $k \geq C\Delta/\log\Delta$ for $C > 1$ is another natural threshold: if the neighbours of a vertex are assigned independently random colours then we expect that the vertex will not be frozen. However, if $k < (1 - \epsilon)\Delta/\log\Delta$, then even in the steady state distribution most degree Δ vertices on a tree will be frozen.

While the presence of frozen vertices impedes our analysis, it needn't preclude rapid mixing *a priori*, as the following intuition illustrates. If the children of a vertex u change colours enough times, then eventually u will become unfrozen and change colours. This allows vertices to unfreeze, level by level, much like in the level dynamics of [9]. This is a slow process: the number of times that the children of u have to change colour before u is unfrozen is (roughly) exponential in Δ/k. However, this value is manageable for $\Delta = O(1)$: the running time is a high degree polynomial rather than superpolynomial. For balanced trees, it is very helpful that there are only $O(\log n)$ levels. For taller trees, a more subtle approach is necessary.

The proofs of our main theorems proceed by bounding the second eigenvalue of the update process, using the method of canonical paths (to bound congestion) and the block dynamics (whereby we compare with a process that updates many nodes simultaneously). However, a straightforward application of these techniques is not sufficient to obtain tight bounds on mixing time for non-complete trees. Our main technical contribution is a weighted variation of these tools, which takes account of differing mixing times amongst the blocks. To the best of our knowledge, this is the first time that this variation has been used for the block dynamics.

In order to apply the block dynamics, we need to analyze the mixing time of the Glauber dynamics on subtrees which have colours on their external boundaries fixed. This is equivalent to fixing the colours on some leaves of the tree. Markov chains on trees with fixed leaves are well-studied. When every leaf is fixed, Martinelli, Sinclair and Weitz [13] prove rapid mixing for $k \geq \Delta + 2$; at $k \leq \Delta + 1$ the chain might not be ergodic. In our setting, we consider cases where k may fall below this threshold, but the number of fixed leaves is small. Theorem 1.1 extends to show:

Theorem 1.4. *For any $k \geq 4$, the Glauber dynamics on k-colourings of any tree with maximum degree Δ and with the colours of any $b \leq k-2$ leaves fixed mixes in time $n^{O(1+b+\Delta/k \log \Delta)}$.*

We cover preliminary material in Section 2, then present the weighted block dynamics in Section 3. An upper bound on the mixing time of the Glauber dynamics with 3 colours on the complete tree is given in Section 4; we then extend this to a proof of Theorem 1.1 in Section 5. The lower bound for Theorem 1.3 is presented in Section 7.

2 Preliminaries

2.1 Graph Colourings

Let $G = (V, E)$ be a finite graph, and let $\Sigma = \{0, 1, \ldots, k-1\}$ be a set of k colours. A *proper colouring* of G is an assignment of colours to vertices such that no two vertices connected by an edge are assigned the same colour. Define $\Omega \subset \Sigma^V$ to be the set of proper colourings of G. Given $\sigma \in \Omega$ and $x \in V$, we write $\sigma(x)$ to mean the colour of vertex x in σ. Given $S \subseteq V$, we write $\sigma(S)$ to refer to the assignment of colours to S in σ; that is, $\sigma(S)$ is σ restricted to S.

Given some $S \subseteq V$, Ω_S^σ is the set of proper colourings of G that are fixed to σ at all vertices not in S. We can think of Ω_S^σ as being equivalent to the set of proper colourings of S with boundary configuration σ. However, technically speaking, an element of Ω_S^σ will be viewed as a colouring of the entire graph G.

2.2 Glauber dynamics

The *Glauber dynamics* for k-colourings of G is a Markov process over the space Ω of proper colourings. We make use of the continuous-time Metropolis version of the Glauber dynamics (standard methods, eg. [5, 14], show that our theorems also hold for the heat-bath version). Informally, the behaviour of this process is as follows: each vertex has a (rate 1) Poisson clock. When the clock for vertex v rings, a colour a is chosen uniformly from Σ. The colour of v is set to a if a does not appear on any neighbour of v, otherwise the colouring remains unchanged.

More formally, recall that a continuous-time Markov process is defined by generator $\mathcal{L}$. We can think of $\mathcal{L}$ as a $|\Omega| \times |\Omega|$ matrix, whose non-diagonal entries represent the jump probabilities between colourings (and diagonal entries are such that all rows sum to 0). For $\sigma \neq \eta$, we will write $K[\sigma \to \eta]$ to denote the (σ, η) entry in this matrix. Under this framework, the jump probabilities for the Metropolis version of the Glauber dynamics are given by

$$K[\sigma \to \eta] = \begin{cases} \frac{1}{k} & \text{if } \sigma, \eta \text{ differ on exactly one vertex} \\ 0 & \text{otherwise} \end{cases}$$

Note that this process is symmetric and, for $k \geq 3$, ergodic on trees [2].

In many applications we are interested in the discrete analog of the Glauber dynamics. We then think of $K[\sigma \to \eta]$ as the probability of moving from colouring σ to colouring η, scaled by a factor of n. The mixing time for the discrete chain is approximately n times the mixing time for the corresponding continuous process [2], so our bounds on mixing time apply to the discrete setting.

2.3 Mixing Time

Given probability distributions π and μ over space Ω, the *total variation distance* between π and μ is defined as

$$||\mu - \pi||_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \pi(x)|.$$

Suppose $\mathcal{L}$ is the generator for an ergodic Markov process over Ω. The *stationary distribution* for $\mathcal{L}$ is the unique measure π on Ω that satisfies $\pi\mathcal{L} = \pi$. It is well-known that the Glauber dynamics has uniform stationary distribution when it is ergodic.

Given any $\sigma \in \Omega$, denote by μ_σ^t the measure on Ω given by running the process with generator $\mathcal{L}$ for time t starting from colouring σ. Then the *mixing time* of the process, $\mathcal{M}(\mathcal{L})$, is defined as

$$\mathcal{M}(\mathcal{L}) = \min\left\{ t : \sup_{\sigma \in \Omega} ||\mu_\sigma^t - \pi||_{TV} \leq \frac{1}{4} \right\}.$$

We define the *spectral gap* of $\mathcal{L}$, $\mathrm{Gap}(\mathcal{L})$, to be the second-largest eigenvalue of $-\mathcal{L}$. The *relaxation time* of $\mathcal{L}$, denoted $\tau(\mathcal{L})$, is defined as the inverse of the spectral gap. We will use the following standard bound (see eg. [12]):

$$M(\mathcal{L}) \leq \tau(\mathcal{L}) \log(|\Omega|) \leq (n \log k)\tau(\mathcal{L}) \qquad \text{since } |\Omega| \leq k^n. \tag{1}$$

2.4 Colourings of Trees

Consider a (not necessarily complete) tree $G = (V, E)$ with maximum degree Δ. A subtree T of G is a connected induced subgraph of G. We shall write ∂T and $\overline{\partial} T$ to mean the exterior and interior boundaries of T. That is, writing $N(v)$ for the neighbourhood of vertex v, $\partial T = \{x \in V \backslash T : N(x) \cap T \neq \emptyset\}$ and $\overline{\partial} T = \{x \in T : N(x) \cap \partial T \neq \emptyset\}$. Note that for each $x \in \partial T$ there is a unique $y \in \overline{\partial} T$ adjacent to x.

The following simple Lemma analyzes the ergodicity of the Glauber dynamics on trees.

Lemma 2.1. *Let T be a subtree of G and suppose $k \geq \max\{3, |\partial T| + 2\}$. Then the Glauber dynamics is ergodic over Ω_T^σ for all $\sigma \in \Omega$.*

Proof. It is sufficient to show irreducibility; ergodicity and the uniformity of the stationary distribution then follow since the Glauber dynamics is aperiodic and reversible. Let $\mathcal{L}_T^\sigma$ be the generator for the Glauber dynamics on T with boundary condition σ, with jump probabilities denoted K_T^σ. Take Γ to be the transition graph over Ω_T^σ, where (η, ω) is an edge in Γ if and only if $K_T^\sigma[\eta \to \omega] > 0$. We need to show that Γ is connected. That is, we need to show that for any two colourings η and ω that differ only in T, it is possible to move from η to ω by changing the colour of one vertex of T at a time, so that at each step we have a proper colouring of G.

Choose $\eta, \omega \in \Omega_T^\sigma$; we will generate a path from η to ω in Γ. We begin by choosing a root node $r \in T$. If $|\partial T| \geq 1$, we arbitrarily choose some $v \in \partial T$ and let r be the unique vertex in T adjacent to v. Otherwise, r is chosen arbitrarily. We now proceed by induction on the height of the resulting rooted tree. If the height is 1 then $V(T) = \{r\}$, and hence η and ω are adjacent in Γ (since they differ on at most a single node, namely r). We conclude that η and ω are connected in Γ.

Now suppose tree T has height h. Let z be a child of r and consider the subtree T' of T rooted at z. If $|\partial T| = 0$ then $\partial T' = \{r\}$, and otherwise $|\partial T'| \leq |\partial T|$. We conclude that $k \geq |\partial T'| + 2$. Also, T' has height at most $h - 1$, and its root z is adjacent to $r \in \partial T'$. Thus by induction the Glauber dynamics restricted to T' is ergodic for any boundary condition, and in particular for η. Since $k \geq |\partial T'| + 2$, there is a colouring $\beta \in \Omega^{\eta}_{T'}$ such that $\beta(z) \notin \{\eta(r), \omega(r)\}$. We can find such a β since at most $|\partial T'|$ colours can be forbidden for z due to the boundary configuration η, leaving 2 possible colours; at most one of those colours is $\omega(r)$, leaving one more. Since the Glauber dynamics is ergodic on T' with boundary condition η, there is a path from η to β in Γ.

Proceeding in the same way, we can change the colour of each child of r so that none are $\omega(r)$. We have thus found a colouring $\alpha \in \Omega^{\sigma}_{T}$ in which $\omega(r)$ does not appear in the neighbourhood of r, and there is a path from η to α in Γ. But now note that the colour $\omega(r)$ does not appear in the neighbourhood of r in α, so one can change the colour of r to $\omega(r)$. That is, $(\alpha, \alpha^{\omega(r)}_r) \in \Gamma$. Let $\gamma = \alpha^{\omega(r)}_r$. Finally, repeating the above argument, it is possible to change the colouring of each subtree T' rooted at a child of r from $\gamma(T')$ to $\omega(T')$ without changing any colours outside of T'. This yields a path from γ to ω. Putting this together, we have found a path from η to ω in Γ. □

3 Weighted Canonical Paths and Block Dynamics

In this section we go over two well-known tools for the analysis of local spin systems: canonical paths and the block dynamics. We then present generalizations of these results that allow the addition of weights to the analysis. We prove our results for the Glauber dynamics acting on a finite graph $G = (V, E)$. Our statements actually apply to a more general setting, holding for all local update chains. We avoid a statement in full generality for succinctness. See [12] for a general treatment of local spin systems.

3.1 Weighted Block Dynamics

Suppose $D = \{V_1, \ldots, V_r\}$ is a collection of subsets of V with $V = \cup_i V_i$. For each $1 \leq i \leq r$ and $\sigma \in \Omega$, let $\mathcal{L}^{\sigma}_{V_i}$ be the generator for the Glauber dynamics restricted to V_i with boundary configuration σ. In other words, colours can change only for nodes in V_i.

Suppose that $\mathcal{L}^{\sigma}_{V_i}$ is ergodic for each i and σ. Let $\pi^{\sigma}_{V_i}$ denote the stationary distribution of $\mathcal{L}^{\sigma}_{V_i}$. For each i, define $g_i := \inf_{\sigma \in \Omega} \mathrm{Gap}(\mathcal{L}^{\sigma}_{V_i})$, the minimum spectral gap for $\mathcal{L}^{\sigma}_{V_i}$ over all choices of boundary configurations.

The *block dynamics* is a continuous-time Markov process with generator $\mathcal{L}_D$ defined by

$$K_D[\sigma \to \eta] = \begin{cases} \pi^{\sigma}_{V_i}[\eta] & \text{if there exists } i \text{ such that } \eta \in \Omega^{\sigma}_{V_i} \\ 0 & \text{otherwise.} \end{cases}$$

Note that $K_D[\sigma \to \eta] > 0$ precisely when η and σ differ only within a single block V_i. Informally, we think of the weighted block dynamics as having a Poisson clock of rate 1 for each block V_i. When clock i rings, the colouring of V_i is replaced randomly according to $\pi^{\sigma}_{V_i}$, where σ is the previous colouring.

Using $\tau_{V_i} = 1/g_i$ to denote the maximum relaxation time of $\mathcal{L}^{\sigma}_{V_i}$ over all choices of boundary configurations, Proposition 3.4 of Martinelli [12] is:

Proposition 3.1. $\tau(\mathcal{L}_{\mathcal{V}}) \leq \tau(\mathcal{L}_D) \times (\max_{1 \leq i \leq r} \tau_{V_i}) \times \max_{x \in V} |\{i : x \in V_i\}|$.

We are now ready to define the *weighted block dynamics* corresponding to D. This is a continuous-time Markov process whose generator $\mathcal{L}^*_D$ is given by

$$K^*_D[\sigma \to \eta] = \begin{cases} g_i \pi^{\sigma}_{V_i}[\eta] & \text{for all } \eta, i \text{ such that } \eta \in \Omega^{\sigma}_{V_i} \\ 0 & \text{otherwise.} \end{cases}$$

The weighted block dynamics is similar to the block dynamics, but the transition probabilities for block V_i are scaled by a factor of g_i. The main result for this section is the following variant of Proposition 3.1:

Proposition 3.2. $\tau(\mathcal{L}_{\mathcal{V}}) \leq \tau(\mathcal{L}^*_D) \times \max_{x\in V} |\{i : x \in V_i\}|$.

The proof of Proposition 3.2 is a simple modification to the proof of Proposition 3.1 [12]. We present the proof here for completeness.

Proof of Proposition 3.2. We begin with some necessary background from the field of functional analysis. Recall that we use $K[\sigma \to \eta]$ to denote the entries of $\mathcal{L}$ as a matrix. Then the operation of $\mathcal{L}$ as a generator over functions $f : \Omega \to \mathbb{R}$ can be expressed as

$$\mathcal{L}(f)(\sigma) = \sum_{\eta\in\Omega} K[\sigma \to \eta](f(\eta) - f(\sigma)).$$

Given a function $f : \Omega \to \mathbb{R}$, the *Variance* of f with respect to $\mathcal{L}$ is given by

$$\mathrm{Var}(f) = \sum_{\sigma,\eta\in\Omega} \pi[\sigma]\pi[\eta](f(\sigma) - f(\eta))^2.$$

The *Dirichlet* form of function f with respect to $\mathcal{L}$ is given by

$$\xi(f, f) = \sum_{\sigma,\eta\in\Omega} \pi[\sigma]K[\sigma \to \eta](f(\sigma) - f(\eta))^2.$$

It is known that the spectral gap of the generator $\mathcal{L}$ satisfies

$$\mathrm{gap}(\mathcal{L}) = \inf_f \frac{\xi(f, f)}{\mathrm{Var}(f)}$$

where the infimum is over all non-constant functions $f : \Omega \to \mathbb{R}$.

We are now ready to proceed with the proof.

Note that $\mathcal{L}^*_D$ is ergodic and reversible with respect to distribution π_V. Let Var^*_D and ξ^*_D denote the variance and Dirichlet form for $\mathcal{L}^*_D$. Note that since $\mathcal{L}^*_D$ and $\mathcal{L}_V$ have the same stationary distributions, $\mathrm{Var}^*_D(f) = \mathrm{Var}_V(f)$ for all functions f.

For each $x \in V$, let $N_x = |\{i : x \in V_i\}|$ and let $N = \max_{x\in V} N_x$. We now bound $\xi^*_D(f, f)$ with respect to N and $\xi_V(f, f)$, as follows.

$$
\begin{aligned}
\xi^*_D(f,f) &= \frac{1}{2}\sum_{\sigma,\eta\in\Omega}\pi[\sigma]K^*_D[\sigma\to\eta](f(\sigma)-f(\eta))^2\\
&= \frac{1}{2}\sum_{\sigma\in\Omega}\pi[\sigma]\sum_{i=1}^{r}g_i\sum_{\eta\in\Omega^\sigma_{V_i}}\pi^\sigma_{V_i}[\eta](f(\sigma)-f(\eta))^2\\
&= \frac{1}{2}\sum_{\sigma\in\Omega}\pi[\sigma]\sum_{i=1}^{r}g_i\mathrm{Var}^\sigma_{V_i}(f)\\
&\le \frac{1}{2}\sum_{\sigma\in\Omega}\pi[\sigma]\sum_{i=1}^{r}\xi^\sigma_{V_i}(f,f)\\
&= \frac{1}{2}\sum_{\sigma\in\Omega}\pi[\sigma]\sum_{i=1}^{r}\sum_{\eta\in\Omega^\sigma_{V_i}}\pi^\sigma_{V_i}[\eta]\sum_{x\in V_i}\sum_{a\in A}K[\eta\to\eta^a_x](f(\eta)-f(\eta^a_x))^2\\
&\le \frac{1}{2}\sum_{\eta\in\Omega}\pi[\eta]\sum_{x\in V}N_x\sum_{a\in A}K[\eta\to\eta^a_x](f(\eta)-f(\eta^a_x))^2\\
&\le N\xi_V(f,f)
\end{aligned}
$$

for all functions f. Note that in the second-last inequality we used the fact that choosing $\sigma\in\Omega$ and then choosing $\eta\in\Omega^\sigma_{V_i}$ is equivalent to choosing $\eta\in\Omega$. But now

$$
\mathrm{gap}(\mathcal{L}_V) = \inf_f \frac{\xi_V(f,f)}{\mathrm{Var}_V(f)} \ge \inf_f \frac{\xi^*_D(f,f)}{\mathrm{Var}^*_D(f)}N^{-1} = \mathrm{gap}(\mathcal{L}^*_D)N^{-1}
$$

as required.

□

It is worth noting the difference between Proposition 3.2 and the original block dynamics, Proposition 3.1. In the original version, the block dynamics Markov process can be thought of as having a Poisson clock of rate g for each block, where g is the minimum over all g_i. In other words, each block is chosen with the same rate, that being the worst case over all blocks. On the other hand, in the modified version each block is chosen with the rate corresponding to that block. The original version yields a simpler Markov process, but a looser bound on the gap of the original process. In particular, applying the original block dynamics to our main result yields a mixing time of $n^{O(1+\Delta/k)}$, while the modified block dynamics tightens the bound to $n^{O(1+\Delta/k\log\Delta)}$ (see Remark 5.7).

We next show that the weighted block dynamics is equivalent to a related process. Informally, we wish to "collapse" each block to its set of internal boundary nodes. We will assign colours to these boundary nodes according to the probability such a boundary configuration would occur in the block dynamics. More formally, suppose $D=\{V_1,\dots,V_m\}$ is a set of blocks of vertices of T. Let $B=\cup_{i=1}^m\overline{\partial}V_i$. That is, B contains all internal boundary nodes for the blocks in D. Note $B\cap V_i=\overline{\partial}V_i$. We define a Markov process $\mathcal{L}_B$ on Ω_B, which simulates the behaviour of $\mathcal{L}_D$ restricted to the nodes in B. Given distribution π over Ω_T, $S\subseteq T$, and $\eta\in\Omega_S$, write $\pi_T[\eta':\eta'(S)=\eta(S)]$ to denote $\sum_{\eta':\eta'(S)=\eta(S)}\pi_T[\eta']$, the probability that the configuration of S agrees with η. Then $\mathcal{L}_B$ is defined by

$$
K_B[\sigma\to\eta] = \begin{cases} g_i\pi^\sigma_{V_i}[\eta':\eta'(\overline{\partial}V_i)=\eta(\overline{\partial}V_i)] & \text{if } \sigma \text{ and } \eta \text{ differ only on } \overline{\partial}V_i\\ 0 & \text{otherwise.}\end{cases} \tag{2}
$$

In other words, η is chosen according to the probability that η is the configuration on B after a step of the block dynamics. Our claim is that the relaxation times of $\mathcal{L}^*_D$ and $\mathcal{L}_B$ are the same; this is similar to Claim 2.9 due to Berger et al [2].

Proposition 3.3. $\tau(\mathcal{L}^*_D) = \tau(\mathcal{L}_B)$.

Proof. Given dynamics $\mathcal{L}$ on configuration space Ω and function $f : \Omega_V \to \mathbb{R}$, we will write $R(\mathcal{L}, f) = \frac{Var(f)}{\xi(f,f)}$. We recall that

$$\tau(\mathcal{L}) = \sup\{R(\mathcal{L}, f) : \pi[f] = 0\} \tag{3}$$

where the maximum is over non-constant functions. From the definition of $\mathcal{L}_B$, we have that

$$\mathcal{L}_B(\sigma', \eta') = \sum_{\sigma:\sigma'=\sigma|_B} \sum_{\eta:\eta'=\eta|_B} \mathcal{L}^*_D(\sigma, \eta). \tag{4}$$

Now suppose we have functions f on Ω_B and g on Ω_T. Suppose further that

$$f(\sigma) = g(\eta) \text{ for all } \sigma, \eta \text{ such that } \eta|_B = \sigma. \tag{5}$$

Then we will have

$$\begin{aligned} R(\mathcal{L}^*_D, g) &= \frac{Var^*_D(g)}{\xi^*_D(g,g)} \\ &= \frac{\sum_{\sigma,\eta\in\Omega_T} \pi^*_D(\sigma)\pi^*_D(\eta)(g(\sigma) - g(\eta))^2}{\sum_{\sigma,\eta\in\Omega_T} \pi^*_D(\sigma)K^*_D(\sigma \to \eta)(g(\sigma) - g(\eta))^2} \\ &= \frac{\sum_{\sigma',\eta'\in\Omega_B} \pi_B(\sigma')\pi_B(\eta')(f(\sigma') - f(\eta'))^2}{\sum_{\sigma',\eta'\in\Omega_B} \pi_B(\sigma')K_B(\sigma' \to \eta')(f(\sigma') - f(\eta'))^2} \\ &= \frac{Var_B(f)}{\xi_B(f,f)} \\ &= R(\mathcal{L}_B, f) \end{aligned} \tag{6}$$

where we used (2) and (4) in the third equality.

Suppose the supremum in (3) for $\mathcal{L}^*_D$ occurs at a function g_1. That is, $g_1 : \Omega_V \to \mathbb{R}$ satisfies $\pi[g_1] = 0$ and $\tau(\mathcal{L}^*_D) = R(\mathcal{L}^*_D, g_1))$. Then g_1 must be an eigenfunction of $\mathcal{L}^*_D$, so $g_1 = \mathcal{L}^*_D(g_1)$. Choose $\sigma, \eta \in \Omega_V$ such that $\sigma(B) = \eta(B)$; then $(\mathcal{L}^*_D(g_1))(\sigma) = (\mathcal{L}^*_D(g_1))(\eta)$ from the definition of $\mathcal{L}^*_D$, and hence $g_1(\sigma) = g_1(\eta)$. We can therefore define function $f_1 : \Omega_B \to \mathbb{R}$ as follows: for each $\alpha \in \Omega_B$, $f_1(\alpha)$ will be the (unique) value of $g_1(\eta)$ for all η with $\eta|_B = \alpha$. Thus f_1 and g_1 satisfy (5), so (6) implies

$$\tau(\mathcal{L}^*_D) = R(\mathcal{L}^*_D, g_1) = R(\mathcal{L}_B, f_1) \le \tau(\mathcal{L}_B). \tag{7}$$

Next suppose that the supremum in (3) for $\mathcal{L}_B$ occurs at a function f_2. Then we can define function g_2 by $g_2(\sigma) = f_2(\sigma|_B)$, from which (6) and (3) imply

$$\tau(\mathcal{L}_B) = R(\mathcal{L}_B, f_2) = R(\mathcal{L}^*_D, g_2) \le \tau(\mathcal{L}^*_D). \tag{8}$$

Equations (7) and (8) imply that $\tau(\mathcal{L}_B) = \tau(\mathcal{L}^*_D)$, as required.

□

3.2 Weighted Canonical Paths

We begin by recalling a standard statement of the method of canonical paths. Fix graph G and Markov process $\mathcal{L}$ over Ω_G. Let Γ be the transition graph of $\mathcal{L}$; that is, $\Gamma = (\Omega_G, E)$ where $(\alpha, \beta) \in E$ if and only if $K[\alpha \to \beta] > 0$.

For each $\alpha, \beta \in \Omega_G$, we will choose a simple path $\gamma(\alpha, \beta)$ from α to β in Γ. We will write $|\gamma(\alpha, \beta)|$ for the length of path $\gamma(\alpha, \beta)$. The *congestion* of this choice of paths is defined to be

$$\rho = \max_{(\sigma,\eta)\in\Gamma} \sum_{\gamma(\alpha,\beta)\ni(\sigma,\eta)} \frac{\pi(\alpha)\pi(\beta)}{\pi[\sigma]K[\sigma \to \eta]}.$$

The following is a standard bound on the relaxation time of $\mathcal{L}$, generally referred to as the *canonical paths bound* [10, 15].

Proposition 3.4. $\tau \leq \rho \times \max_{\alpha,\beta\in\Omega_G} |\gamma(\alpha, \beta)|$

We wish to extend Proposition 3.4 to allow us to add weights to the edges of graph Γ. For each edge (σ, η) in Γ we define an arbitrary real-valued weight $w(\sigma, \eta) \geq 0$. Given a simple path $\gamma(\alpha, \beta)$ from α to β, we define the weighted length of $\gamma(\alpha, \beta)$ to be $|\gamma(\alpha, \beta)|_w = \sum_{(\sigma,\eta)\in\gamma(\alpha,\beta)} w(\sigma, \eta)$. The weighted congestion for a given choice of paths is then defined to be

$$\rho_w = \max_{(\sigma,\eta)\in\Gamma} \frac{1}{w(\sigma,\eta)} \sum_{\gamma(\alpha,\beta)\ni(\sigma,\eta)} \frac{\pi(\alpha)\pi(\beta)}{\pi[\sigma]K[\sigma \to \eta]}.$$

Proposition 3.5. $\tau \leq \rho_w \times \max_{\alpha,\beta\in\Omega_G} |\gamma(\alpha, \beta)|_w$

We note that while we have not seen a proof of this result precisely as stated here, it has been pointed out elsewhere that alternative weighting choices can be used to obtain variants of the canonical paths bound (see, for example, the remarks preceeding Proposition 1' in Diaconis and Stroock [6]). We present a proof of Proposition 3.5 for completeness.

Proof of Proposition 3.5. Recall the definitions of variance $\mathrm{Var}(f)$ and Dirichlet form $\xi(f, f)$ of a function $f : \Omega \to \mathbb{R}$ with respect to $\mathcal{L}$, as described at the beginning of the proof of Proposition 3.2. We recall that

$$\tau = \frac{1}{\mathrm{gap}(\mathcal{L})} = \sup_f \frac{\mathrm{Var}(f)}{\xi(f, f)}$$

where the supremum is over all non-constant functions $f : \Omega \to \mathbb{R}$.

For all non-constant functions $f : \Omega \to \mathbb{R}$,

$$\begin{aligned}
\mathrm{Var}(f) &= \sum_{\sigma,\eta\in\Omega_G} \pi(\sigma)\pi(\eta)(f(\eta) - f(\sigma))^2 \\
&= \sum_{\sigma,\eta\in\Omega_G} \sum_{(\alpha,\beta)\in\gamma(\sigma,\eta)} \frac{w(\alpha,\beta)}{w(\alpha,\beta)}(f(\beta) - f(\alpha))^2\pi(\sigma)\pi(\eta) \\
&\leq \sum_{\sigma,\eta\in\Omega_G} |\gamma(\sigma,\eta)|_w \pi(\sigma)\pi(\eta) \sum_{(\alpha,\beta)\in\gamma(\sigma,\eta)} \frac{1}{w(\alpha,\beta)}(f(\alpha) - f(\beta))^2 \\
&= \sum_{(\alpha,\beta)\in\Gamma} \frac{\pi(\alpha)K[\alpha \to \beta]}{w(\alpha,\beta)}(f(\alpha) - f(\beta))^2 \sum_{\substack{\sigma,\eta: \\ \gamma(\sigma,\eta)\ni(\alpha,\beta)}} |\gamma(\sigma,\eta)|_w \frac{\pi(\sigma)\pi(\eta)}{\pi(\alpha)K[\alpha \to \beta]} \\
&\leq \xi(f, f) \times \max_{\alpha,\beta\in\Omega_G} |\gamma(\alpha,\beta)|_w \times \rho_w
\end{aligned}$$

which implies that

$$\tau = \sup_f \frac{\mathrm{Var}(f)}{\xi(f,f)} \leq \max_{\alpha,\beta\in\Omega_G} |\gamma(\alpha,\beta)| \times \rho_w$$

as required.

□

4 3-Colourings of the Complete Tree

Before proving Theorem 1.1, we will warm up by bounding the relaxation time of the Glauber dynamics on the complete tree in the special case $k = 3$. In the next section we will extend the argument to bound the relaxation time on arbitrary trees for $k \geq 4$; Theorem 1.2 follows from the combination of these two results.

Let T be a complete rooted tree of maximum degree Δ, possibly with a single external boundary node adjacent to its root.[1] Let $n = |T|$, let v be the root of T, and let μ be a boundary configuration for T (i.e. the fixed colour of the external boundary node, if it exists).

We consider the Glauber dynamics with 3 colours on T with boundary configuration μ, which is ergodic by Lemma 2.1. Note that, up to a possible relabelling of the boundary colour, the behaviour of the dynamics is completely determined by the height of T, say h. We therefore write $\tau(h) := \tau_T^\mu$ to be the relaxation time of this process.

Lemma 4.1. *For some fixed constant c, and for all $h > 0$,*

$$\tau(h) \leq c\Delta 2^\Delta \tau(h-1). \tag{9}$$

Before proving Lemma 4.1, let us show how it implies the upper bound for Theorem 1.1. For the case $h = 0$, we have that T is a single vertex, and hence $\tau(0) = 1$. This plus Lemma 4.1 implies that $\tau(h) \leq \left(c\Delta 2^\Delta\right)^h$. But then, using (1) and the fact that T has height at most $\lfloor \log_\Delta n \rfloor$, we have that the mixing time for the Glauber dynamics on T is

$$\begin{aligned}\mathcal{M} &\leq (n\log 3)\tau(\lfloor \log_\Delta n \rfloor) \\ &\leq (n\log 3)\left(c\Delta 2^\Delta\right)^{\log n/\log\Delta} \\ &= n^{O(1+\Delta/\log\Delta)}\end{aligned}$$

which is the upper bound from Theorem 1.2. It remains to prove Lemma 4.1, to which we devote the rest of this section.

Proof of Lemma 4.1. We note that the general structure of our proof is very similar to the proof of Lemma 2.8 in [2]. We include the argument for completeness, and because we extend it in Section 5.1.

Suppose $h > 0$ and let $u_1, \ldots, u_{\Delta-1}$ be the children of v in T. Let V_i be the subtree of T rooted at u_i, for each $1 \leq i \leq \Delta - 1$; note that each V_i is a complete tree of height $h - 1$. Let $D = \{\{v\}, V_1, \ldots, V_{\Delta-1}\}$. Consider the block dynamics $\mathcal{L}_D$ on subtree T with blocks D. Let K_D denote the transition probabilities for $\mathcal{L}_D$, and let τ_D be the relaxation time of this Markov process. Then since no vertex in V lies in multiple blocks in D, Proposition 3.1 implies that

$$\tau(h) \leq \tau_D \max_{\mu\in\Omega}\{\tau^\mu_{\{v\}}, \tau^\mu_{V_1}, \ldots, \tau^\mu_{V_{\Delta-1}}\} = \tau_D\tau(h-1) \tag{10}$$

[1] We take the convention that each node in the complete tree of degree Δ has $\Delta - 1$ children, including the root.

since $\tau^{\mu}_{\{v\}} = 1$, the relaxation time over a single vertex. It therefore remains to show that

$$\tau_D \leq c\Delta 2^{\Delta}. \tag{11}$$

Recall the definition of graph B and dynamics $\mathcal{L}_B$ from Proposition 3.3. In our case, B is a star with root v and leaves $u_1, \ldots, u_{\Delta-1}$, and $\mathcal{L}_B$ is precisely the Glauber dynamics on B (possibly with a boundary condition effecting v, corresponding to any boundary condition for T). Proposition 3.3 then implies $\tau_D = \tau_B$. Thus, to prove (11), it is sufficient to bound τ_B.

First let us give some intuition into the bound in (11) as it applies to τ_B. We would expect the mixing time to be at least the expected time for v to change colour starting from a configuration chosen uniformly at random. There are 2 colours that v might change to, and the probability that a particular colour is not present on the leaves at some point of time is $2^{-(\Delta-1)}$. Thus we expect it to take roughly $2^{\Delta-1}$ time before the colour of the root can change. The bound in (11) states that the mixing time is not much more than this.

We now proceed to bound τ_B using Proposition 3.4, the method of canonical paths. We note that this is not the simplest way to obtain the desired bound, but it introduces techniques that will be useful when proving Theorem 1.1.

Choose two colourings $\sigma, \eta \in \Omega_B$. Our goal is to define a sequence of steps of the Glauber dynamics that begins in state σ and ends in state η. If $\sigma(v) = \eta(v)$ this sequence is simple: the colours of nodes $u_1, \ldots, u_{\Delta-1}$ are changed from σ to η one at a time. If $\sigma(v) \neq \eta(v)$, our strategy is to first change the colours of $u_1, \ldots, u_{\Delta-1}$ so that none have colour $\eta(v)$, then change the colour of v to $\eta(v)$, and finally set the colours of the u_i nodes to match η.

Let Γ be the transition graph over Ω_B, with $(\alpha, \beta) \in \Gamma$ if and only if $K_B[\alpha \to \beta] > 0$. That is, $(\alpha, \beta) \in \Gamma$ if and only if colourings α and β differ on exactly one vertex of G. For each $\sigma, \eta \in \Omega_G$ we will define a simple path in Γ, denoted $\gamma(\sigma, \eta)$. If $\sigma(v) = \eta(v)$, our path changes the colour of each u_i from $\sigma(u_i)$ to $\eta(u_i)$, one at a time. If $\sigma(v) \neq \eta(v)$, then (writing ζ for the unique colour not in $\{\alpha(v), \eta(v)\}$) our path $\gamma(\sigma, \eta)$ is as follows:

1. For each u_i in increasing order: recolour from $\sigma(u_i)$ to ζ.
2. Recolour v from $\sigma(v)$ to $\eta(v)$.
3. For each u_i in decreasing order: recolour from ζ to $\eta(u_i)$.

Let L be the maximum length of any such path $\gamma(\sigma, \eta)$. We note that $L \leq 2\Delta - 1$.

For each edge $(\alpha, \beta) \in \Gamma$, define the congestion of that edge, $\rho(\alpha, \beta)$, as

$$\rho(\alpha, \beta) := \sum_{\substack{\sigma,\eta: \\ \gamma(\sigma,\eta) \ni (\alpha,\beta)}} \frac{\pi[\sigma]\pi[\eta]}{\pi[\alpha]K[\alpha \to \beta]}.$$

The congestion for our set of paths is

$$\rho := \sup_{\alpha, \beta} \rho(\alpha, \beta).$$

Proposition 3.4 gives us that

$$\tau_B \leq L\rho \leq (2\Delta - 1)\rho \tag{12}$$

Our goal now is to bound ρ. From the definition of the Glauber dynamics,

$$K[\alpha \to \beta] = \frac{1}{3}. \tag{13}$$

We also note that

$$|\Omega_B| \geq 2^{\Delta}, \tag{14}$$

which follows by choosing a colouring for B in a top-down manner starting at v, where the colour chosen for each vertex must avoid the colour chosen for its parent (including v, if B has an external boundary).

Using (14) and the fact that π is the uniform distribution, we have

$$\begin{aligned} \rho(\alpha, \beta) &= \left(|\{\sigma, \eta : \gamma(\sigma, \eta) \ni (\alpha, \beta)\}| \times \frac{1}{|\Omega_G| K[\alpha \to \beta]} \right) \\ &\leq \left(|\{\sigma, \eta : \gamma(\sigma, \eta) \ni (\alpha, \beta)\}| \times \frac{3}{2^{\Delta+1}} \right). \end{aligned} \tag{15}$$

To bound ρ, it remains to compute the number of paths $\gamma(\sigma, \eta)$ that include (α, β) for each $(\alpha, \beta) \in \Gamma$. If $\alpha(v) \neq \beta(v)$, then (α, β) will appear in $\gamma(\sigma, \eta)$ for each σ and η for which $\sigma(v) = \alpha(v)$ and $\eta(v) = \beta(v)$. There are at most $2^{2(\Delta-1)}$ such σ and η, corresponding to the choices of $\sigma(u_1), \ldots, \sigma(u_{\Delta-1})$ and $\eta(u_1), \ldots, \eta(u_{\Delta-1})$.

On the other hand, if α and β differ on the colour of u_i for some i, then if (α, β) occurs on a path $\gamma(\sigma, \eta)$ then one of the following is true:

- $\sigma(v) = \alpha(v)$, $\sigma(u_i) = \alpha(u_i)$, and $\eta(v) \neq \alpha(u_i)$, or
- $\eta(v) = \beta(v)$, $\eta(u_i) = \beta(u_i)$ and $\sigma(v) \neq \eta(u_i)$.

There are at most $2^{2\Delta-1}$ choices for σ and η in either case, for a total of $2 \times 2^{2(\Delta-1)}$. We conclude that at most $2^{2\Delta-1}$ paths include any given transition (α, β).

Substituting the value $2^{2\Delta-1}$ into (15), we have

$$\begin{aligned} \rho(\alpha, \beta) &\leq 2^{2\Delta-1} \frac{3}{2^{\Delta}} \\ &\leq 3 \times 2^{\Delta} \end{aligned}$$

Applying (12), we conclude that

$$\tau_D \leq L\rho \leq 3(2\Delta - 1)2^{\Delta} \leq 6\Delta 2^{\Delta}$$

which is (11) with $c = 6$. This completes the proof of Lemma 4.1.

□

5 An Upper Bound for General Trees

We now begin our proof of Theorem 1.1. Our approach is to decompose a tree into smaller subtrees, apply the block dynamics to the resulting subgraphs, and then use induction to bound the mixing time of the entire tree. Implicitly, this yields an iterative decomposition of the tree into smaller and smaller subtrees. How should we decompose a tree? A first idea is to use the same decomposition that we applied to the complete tree in the previous section: root the tree at a vertex v, then take each subtree rooted at a child of v as a block (and v itself as a block of size 1). A nice property of this decomposition is that each subtree has at most one boundary node, adjacent to its root. In this case there will be h levels of recursion in the induction, where h is the height of tree T, and we will obtain a bound of the form c^h, where $c = c(\Delta, k)$ is the mixing time for an instance of the block dynamics. This method worked for complete trees since they have logarithmic height. However, the height of a general tree could be much greater, leading to a super-polynomial bound.

Instead, we will partition the tree in a manner that guarantees each block has size at most half the size of the tree. This ensures that our recursion halts after logarithmically many steps, and yields a polynomial mixing time. To obtain such a partition, we choose a central node x and conceptually split the tree by removing x, obtaining at most Δ subtrees plus $\{x\}$.

There are difficulties with the above approach that must be overcome. First, a subtree T may have multiple boundary nodes, which complicates the behaviour of the block dynamics. We therefore make our choice of x carefully, so that boundaries are of size at most 2. It is this point in the analysis that we require that there be at least 4 colours, so that Lemma 2.1 implies that the Glauber dynamics is ergodic on each subtree T. Second, for non-complete trees we might have blocks of vastly differing sizes, which makes a tight analysis of the block dynamics more difficult. We therefore use the weighted version of the block dynamics.

In this section we describe our choice of blocks for the block dynamics. We then show that the upper bound of Theorem 1.1 holds, given a bound on the relaxation time of the block dynamics. The details of analyzing the block dynamics are encapsulated in Lemma 5.1, which is proved in Section 5.1.

Let T be any tree with maximum degree Δ. Suppose $|T| = n$ and $|\partial T| \leq 2$ (that is, T has at most two external boundary nodes). Suppose also that $k \geq 4$ (i.e. there are at least 4 colours). Let μ be a boundary configuration for T. Let $\mathcal{L}_T^\mu$ denote the Glauber dynamics on T with k colours and boundary configuration μ. Since $|\partial T| \leq 2$, Lemma 2.1 implies that $\mathcal{L}_T^\mu$ is ergodic for all μ. Let τ_T^μ denote the relaxation time for $\mathcal{L}_T^\mu$. We wish to consider the maximum relaxation time over all boundary configurations and trees of a certain size. To this end, we define

$$\tau_T := \max_{\mu \in \Omega} \tau_T^\mu \qquad \text{and} \qquad \tau_i(n) := \max_{T : |T| \leq n,\ |\partial T| \leq i} \tau_T.$$

We will prove Theorem 1.1 by showing the slightly stronger result that $\tau_2(n) = n^{O(1+\Delta/k \log \Delta)}$. We will show that, for some fixed constant c and some $2 \leq i \leq \Delta$,

$$\tau_2(n) \leq c i^2 \left(\frac{k-1}{k-2} \right)^{i+1} \tau_2\left(\lfloor n/i \rfloor \right). \tag{16}$$

First let us show how (16) implies Theorem 1.1.By induction on n, (16) implies that $\tau_2(n) \leq n^{d(1+\Delta/k \log \Delta)}$ for some constant d (since we can assume $k \leq 2\Delta$, as otherwise the result is known [9]). By (1), the mixing time of the Glauber dynamics on a tree T with n nodes satisfies $\mathcal{M}(\mathcal{L}) \leq (n \log k)\tau_T \leq (n \log k)\tau_2(n) = n^{O(1+\Delta/k \log \Delta)}$ as required.

We now turn to proving (16). The following Lemma will be our main tool.

Lemma 5.1. *Suppose $k \geq 4$ and let T be a subtree of a tree G with $|\partial T| \leq 2$ and let $\mu \in \Omega$ be a boundary condition for T. Choose $v \in T$ and let $D_v = \{\{v\}, V_1, \ldots, V_t\}$ be a partition of T into disjoint connected subtrees, where $1 \leq t \leq \Delta$. Suppose $|\partial V_i| \leq 2$ for each V_i. Then there exists constant c such that*

$$\tau_T^\mu \leq c \max_{1 \leq i \leq t} i^2 \left(\frac{k-1}{k-2} \right)^i \tau_{V_i}.$$

We prove Lemma 5.1 in Section 5.1. Let us show how it implies (16). We first consider trees with boundaries of size one, then size two.

Lemma 5.2. *For some $2 \leq i \leq \Delta$, we have $\tau_1(n) \leq c i^2 \left(\frac{k-1}{k-2} \right)^i \tau_2\left(\lfloor n/i \rfloor \right)$.*

Proof. Choose some T with $|\partial T| \leq 1$. It is a classical result that we can find a vertex $x \in T$ such that if $D_x = \{\{x\}, V_1, \ldots, V_t\}$ is the set of disjoint connected subtrees obtained by disconnecting x from the rest of T, then we will have $|V_i| \leq \lfloor n/2 \rfloor$ for all $1 \leq i \leq t$. Indeed, if we root T arbitrarily and define $T(y)$ to be the subtree rooted at y for each $y \in T$, then we can choose x to be the vertex such that $|T(x)| > \lfloor n/2 \rfloor$ and $|T(x)|$ is as small as possible. Then $|T \backslash T(x)| \leq \lfloor n/2 \rfloor$ and $|T(z)| \leq \lfloor n/2 \rfloor$ for all children z of x, so x satisfies the desired property.

Fix such a node x and choose indices so that $|V_1| \geq |V_2| \geq \ldots \geq |V_t|$. Since $|\partial T| \leq 1$, we have $|\partial V_i| \leq 2$ for all i. By Lemma 5.1, $\tau_T \leq ci^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}$ for some $1 \leq i \leq t$.

If $i \geq 2$, we get $\tau_{V_i} \leq \tau_2(|V_i|) \leq \tau_2(\lfloor n/i \rfloor)$, since the V_i are given by increasing size. Thus $\tau_T \leq ci^2 \left(\frac{k-1}{k-2}\right)^i \tau_2(\lfloor n/i \rfloor)$ for some $2 \leq i \leq t$ as required. If $i = 1$, then we recall that $|V_1| \leq \lfloor n/2 \rfloor$ by our choice of x. Hence $\tau_T \leq c\left(\frac{k-1}{k-2}\right) \tau_{V_1} < c(2)^2 \left(\frac{k-1}{k-2}\right)^2 \tau_2(\lfloor n/2 \rfloor)$ as required. □

Lemma 5.3. *For some* $2 \leq i \leq \Delta$, $\tau_2(n) \leq c^2 i^2 \left(\frac{k-1}{k-2}\right)^{i+1} \tau_2(\lfloor n/i \rfloor)$.

Proof. Let T be a subtree with $|T| = n$ and $|\partial T| = 2$, say $\partial T = \{z_1, z_2\}$. Choose x as in Lemma 5.2, with x separating T into subtrees of size at most $\lfloor n/2 \rfloor$. We will call the unique path in T from z_1 to z_2 the *boundary path* for T. Suppose x is on the boundary path for T. Let $D_x = \{\{x\}, V_1, \ldots, V_t\}$ be a partition into disjoint connected subtrees, indexed so that $|V_1| \geq \ldots \geq |V_t|$; note that $|\partial V_i| \leq 2$ for all i. We then apply Lemma 5.1 as in Lemma 5.2 and obtain the desired result.

Now suppose that x is not on the boundary path for T. Consider T to be rooted at some $r \in \overline{\partial} T$. Let y be the least ancestor of x that lies on the boundary path. Consider $D_y = \{\{y\}, V_1, \ldots, V_t\}$. Since x separates T into subtrees of size at most $\lfloor n/2 \rfloor$, in particular the subtree containing y must have size at most $\lfloor n/2 \rfloor$. This implies that the subtree separated by y that contains x must contain at least $\lfloor n/2 \rfloor$ nodes, and is therefore V_1, the largest subtree separated by y. Also, $|\overline{\partial} V_i| \leq 2$ for all i, since y is on the boundary path for T. Lemma 5.1 implies

$$\tau_T \leq ci^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}$$

for some i. If $i > 1$ then we obtain the desired result since $|V_i| \leq \lfloor n/i \rfloor$. If $i = 1$, then since $|V_1| < n$ and $|\overline{\partial} V_1| = 1$ (by our choice of y), Lemma 5.2 implies

$$\begin{aligned} \tau_T &\leq c\left(\frac{k-1}{k-2}\right) \tau_1(|V_1|) \leq c\left(\frac{k-1}{k-2}\right) \tau_1(n) \\ &\leq c^2 i^2 \left(\frac{k-1}{k-2}\right)^{i+1} \tau_2(\lfloor n/i \rfloor) \quad \text{for some } 2 \leq i \leq \Delta. \end{aligned}$$

□

We have now derived (16), completing the proof of Theorem 1.1.

5.1 Proof of Lemma 5.1

We now proceed with the proof of Lemma 5.1, which bounds the relaxation time on a tree with respect to the relaxation times for subtrees. Our approach is to use a canonical paths argument to bound the behaviour of the block dynamics. Indeed, there is a simple canonical path to move between any two configurations σ and η of T: modify the configuration of each V_i to an intermediate state so that v is free to change colour to $\eta(v)$, change the colour of v to $\eta(v)$, then set the configuration of each V_i to $\eta(V_i)$. The block dynamics paired with this set of canonical paths yields a bound on the relaxation time. However, that bound is not tight enough to imply the mixing rate we desire: it only implies a mixing time of $n^{O(\Delta)}$. We therefore apply the following sequence of improvements to the above approach.

1. We explicitly describe an intermediate configuration for the neighbours of v, in order to balance congestion over all start and end configurations. This improves the bound on the mixing time to $n^{O(\log \Delta + \log k + \Delta/k)}$.

2. Our path shifts between 3 different intermediate configurations to maximize the dependency on the start and end configurations at each step. This improves our bound to $n^{O(\log \Delta + \Delta/k)}$.

3. We apply the weighted block dynamics, to differentiate between large and small subtrees. We always change configurations of blocks in order of subtree size. This improves our bound to $n^{O(\log \Delta + \Delta/k \log \Delta)}$. See Remark 5.7.

4. We apply weights to our canonical path to discount the congestion on smaller subtrees. The net effect is that the presence of many small subtrees does not influence the congestion of our paths. This improves our bound to $n^{O(1 + \Delta/k \log \Delta)}$. See Remark 5.6.

5.1.1 The Block Dynamics

Recall the conditions of Lemma 5.1. Suppose $k \geq 4$ and let T be a tree with $|\partial T| \leq 2$ and let $\mu \in \Omega$ be a boundary condition for T. We will root T at some vertex $v \in T$. Let $u_1, \ldots, u_t$ be the children of v, where $1 \leq t \leq \Delta$. Let V_i denote the subtree rooted at u_i, so that $D = \{\{v\}, V_1, \ldots, V_t\}$ is a partition of T. Suppose we chose v so that $|\partial V_i| \leq 2$ for each V_i. For improved clarity of exposition, we will prove Lemma 5.1 under the assumption that $u_i \notin \overline{\partial} T$ for all i. The (simple) extension to remove this assumption is discussed at the conclusion of this section; see Section 5.1.5.

Let $\mathcal{L}_D^*$ be the generator for the weighted block dynamics corresponding to D and boundary configuration μ. Let τ_D^μ denote the relaxation time of $\mathcal{L}_D^*$. Since no vertex lies in more than one block, Proposition 3.2 implies $\tau_T^\mu \leq \tau_D^\mu$.

Next recall the definition of graph B and dynamics $\mathcal{L}_B$ from Proposition 3.3. In this context, we can view $\mathcal{L}_B$ as a version of $\mathcal{L}_D^*$ wherein each block is treated like a single vertex. That is, B is a star with internal node v; we will refer to $u_1, \ldots, u_t$ as the leaf nodes of B. When such a leaf node, say u_i, is chosen by the dynamics, its colour updates with probability corresponding to the probability of seeing that colour as the root of V_i in $\mathcal{L}_D^*$. By Proposition 3.3, $\tau(\mathcal{L}_D^*) = \tau(\mathcal{L}_B^\mu)$. It is therefore sufficient to bound $\tau(\mathcal{L}_B^\mu)$. The following simple Lemma bounds the transition probabilities of $\mathcal{L}_B^\mu$.

Lemma 5.4. *Choose $S \subseteq T$ with $|\partial S| \leq 2$ and boundary configuration ξ, and suppose $x \in \overline{\partial} S$. Choose $c \in A$ and suppose there exists some $\eta \in \Omega_S^\xi$ with $\eta(x) = c$. Then $\pi_S^\xi[\omega : \omega(x) = c] \geq 1/k$.*

Proof. Think of x as the root of S and consider choosing a configuration for S by choosing a colour for each node, top-down. Since $x \in \overline{\partial} S$, there are at most $k-1$ choices for the colour of x. Given any such choice, there are $k-1$ choices for each subsequent node, except possibly for any node in $\overline{\partial} S$ for which there will be either $k-1$ or $k-2$ choices. Since $|\partial S| \leq 2$, there can be at most one node beside x in $\overline{\partial} S$. Thus, for any $c \in A$, $(k-1)^{n-2}(k-2) \leq |\{\eta \in \Omega_S^\xi : \eta(x) = c\}| \leq (k-1)^{n-1}$, as long as there is at least one colouring in which x has colour c. Thus

$$\pi_S^\xi[\eta : \eta(x) = c] = \frac{|\{\eta \in \Omega_S^\xi : \eta(x) = c\}|}{|\Omega_S^\xi|} \geq \frac{(k-1)^{n-2}(k-2)}{(k-1)^{n-1}(k-2) + (k-1)^{n-2}(k-2)} = \frac{1}{k}$$

as required. □

The following corollary is now immediate from Lemma 5.4 and the definition of $\mathcal{L}_B$.

Corollary 5.5. *Suppose $\alpha, \beta \in \Omega_B^\mu$, $K_B[\alpha \to \beta] > 0$, and $\alpha(u_i) \neq \beta(u_i)$. Then $K_B[\alpha \to \beta] \geq (k\tau_{V_i}^\mu)^{-1}$.*

5.1.2 Definition of Intermediate Configurations

Choose two colourings $\sigma, \eta \in \Omega_B$. Our goal is to define a sequence of steps of $\mathcal{L}_B$ that begins in state σ and ends in state η. If $\sigma(v) = \eta(v)$ this sequence is simple: the colours of nodes $u_1, \ldots, u_t$ are changed from σ to η one at a time. If $\sigma(v) \neq \eta(v)$, our strategy is to first change the colours of $u_1, \ldots, u_t$ so that none have colour $\eta(v)$, then change the colour of v to $\eta(v)$, and finally set the colours of the u_i nodes to match η. The obvious way to do this requires two "passes" of changes over the leaf nodes (as in our analysis for the complete tree in Section 4) but this method generates too much congestion in the canonical paths analysis (see Section 3.2). We therefore introduce a more complex path that uses three passes. We now define the colours used in the intermediate configurations of this path.

If $\sigma(v) \neq \eta(v)$ then for each $1 \leq i \leq t$ we will define three colours, a_i, b_i, and c_i, that depend on σ and η. The first two colours are easy to define:

$$a_i = \begin{cases} \sigma(u_i) & \text{if } \sigma(u_i) \neq \eta(v) \\ \sigma(v) & \text{otherwise} \end{cases} \qquad b_i = \begin{cases} \eta(u_i) & \text{if } \eta(u_i) \neq \sigma(v) \\ \eta(v) & \text{otherwise} \end{cases}$$

That is, $(a_1, \ldots, a_t)$ are the colours of the children of v in σ, with any occurrences of $\eta(v)$ replaced with $\sigma(v)$. Note that our assumption that u_i is not adjacent to the external boundary of T ensures that there exists a configuration in which u_i has colour a_i. We define b_i in the same way, but with the roles of σ and η reversed.

The definition of colour c_i is more involved. These will be the colours to which we set the leaf nodes to allow v to change from $\sigma(v)$ to $\eta(v)$. We will apply a function f that will map the colours $(\sigma(u_1), \ldots, \sigma(u_t))$ to a vector of colours $(c_1, \ldots, c_t)$ such that for all i, $c_i \notin \{\sigma(v), \eta(v)\}$. We want f to satisfy the following balance property: for all $1 \leq i \leq t$, writing $\mathbf{x}$ for $(x_1, \ldots, x_t)$,

$$\#\{\mathbf{x} : \ (x_j = \sigma(u_j) \ \forall j > i) \wedge (f(\mathbf{x})_j = c_j \ \forall j \leq i)\} \leq \left\lceil \left(\frac{k-1}{k-2}\right)^i \right\rceil. \tag{17}$$

That is, for any $1 \leq i \leq t$, if we are given $c_1, \ldots, c_i$ and $\sigma(u_{i+1}), \ldots, \sigma(u_t)$, there are at most $\left\lceil \left(\frac{k-1}{k-2}\right)^i \right\rceil$ possibilities for $\sigma(u_1), \ldots, \sigma(u_t)$. Such an f is guaranteed to exist; an explicit construction is given in Section 5.2.

5.1.3 The Path Definition

Let Γ be the transition graph over Ω_B with $(\alpha, \beta) \in \Gamma$ if and only if $K_B[\alpha \to \beta] > 0$. We are now ready to define the path $\gamma(\sigma, \eta)$ for each $\sigma, \eta \in \Omega_B$. If $\sigma(v) = \eta(v)$, our path simply changes the colour of each u_i from $\sigma(u_i)$ to $\eta(u_i)$, one at a time. If $\sigma(v) \neq \eta(v)$, we use the following path:

1. For each u_i in increasing order: recolour from $\sigma(u_i)$ to b_i, then to c_i.

2. Recolour v from $\sigma(v)$ to $\eta(v)$.

3. For each u_i in decreasing order: recolour from c_i to $\eta(u_i)$, then to a_i.

4. For each u_i in increasing order: recolour from a_i to $\eta(u_i)$.

The reader is encouraged to verify that all steps are valid transitions according to $\mathcal{L}_B^\mu$. At first glance, the number of changes to the colour of each u_i seems excessive. In particular, step 4, the second half of step 3, and the first half of step 1 may seem redundant. However, we define our path in this way to maintain two important properties. First, each change to any vertex is from a colour derived from α to a colour derived from η, or vice-versa. Second, whenever the colours $c_1, \ldots, c_t$ appear on some of the leaves, they appear on the low-index nodes; and each remaining leaf (except possibly one) will have colour $\sigma(u_i)$ or a_i. These properties will be important in our analysis of the congestion generated by this set of paths.

5.1.4 Analysis of Weighted Path Congestion

We wish to apply the canonical paths bound to our choice of paths. However, the standard bound, Proposition 3.4, will not give us a tight result when the sizes of the subtrees in our decomposition are highly skewed. Let us briefly provide some intuition. The bound in Proposition 3.4 depends highly on the length of the paths we have chosen. In our case, our paths always have length linear in t. However, in our decomposition of tree T into t subtrees, it may be that only some small number t' of these subtrees trees are large and the remainder are very small. In this case our tree is "close to" a tree with only t' subtrees, so we would like for the congestion of our paths to depend on t' rather than t. Thus, when calculating the length of our paths, we would rather use a weighted length that takes into account the differing sizes of the blocks being reconfigured.

We will address this issue by applying Proposition 3.5, the weighted canonical paths bound, to our choice of paths. We will choose our weights such that changes to nodes corresponding to larger subtrees have greater weight, and furthermore the total weight of a path that changes the colour of each node once will be bounded by a constant. Specifically, we will set $w(\alpha, \beta) = 1$ if α and β differ on the colour of v, and set $w(\alpha, \beta) = i^{-2}$ if α and β differ on the colour of vertex u_i. Recall that the weight of path $\gamma(\sigma, \eta)$ is given by $|\gamma(\sigma, \eta)|_w = \sum_{(\alpha,\beta)\in\gamma(\sigma,\eta)} w(\alpha, \beta)$. Then note that for all σ and η, recalling the definition of $\gamma(\sigma, \eta)$ from Section 5.1.3, $|\gamma(\sigma, \eta)|_w \leq 1 + 5\sum_{i=1}^t i^{-2} < 1 + 5\left(\frac{\pi^2}{6}\right) < 10$. For each edge $(\alpha, \beta) \in \Gamma$, the weighted congestion of that edge, $\rho_w(\alpha, \beta)$, is

$$\rho_w(\alpha, \beta) := \frac{1}{w(\alpha, \beta)} \left(\sum_{\substack{\sigma,\eta: \\ \gamma(\sigma,\eta)\ni(\omega,\beta)}} \frac{\pi[\sigma]\pi[\eta]}{\pi[\alpha]K_B[\alpha \to \beta]} \right).$$

The weighted congestion for our set of paths is $\rho_w := \sup_{\omega,\beta} \rho_w(\omega, \beta)$. Proposition 3.5 is then

$$\tau_B^\mu \leq \max_{\sigma,\eta} |\gamma(\sigma, \eta)|_w \times \rho_w \leq 10\rho_w. \tag{18}$$

It remains to bound $\rho_w(\alpha, \beta)$ for each $(\alpha, \beta) \in \Gamma$. Uniformity of π implies

$$\rho_w(\alpha, \beta) \leq \frac{1}{w(\alpha, \beta)} \times |\{\gamma(\sigma, \eta) \ni (\alpha, \beta)\}| \times \frac{1}{(k-1)^{t+1} K_B[\alpha \to \beta]}. \tag{19}$$

We now consider cases depending on the nature of the transition (α, β).

Case 1: α and β differ on the colour of v. Note that $w(\alpha, \beta) = 1$. Also, from the definition of $\mathcal{L}_B$, we have $K_B[\alpha \to \beta] = \inf_{\mu\in\Omega} \text{gap}(\mathcal{L}_{\{v\}}^\mu)\pi_{\{v\}}^\alpha[\phi : \phi(v) = \beta(v)]$. But note that $\text{gap}(\mathcal{L}_{\{v\}}^\mu) = 1$ for all boundary conditions, and $\pi_{\{v\}}^\alpha$ is the uniform distribution over a set of at most $k-1$ colours.We conclude

$$K_B[\alpha \to \beta] \geq \frac{1}{k-1}. \tag{20}$$

Consider the number of (σ, η) such that $(\alpha, \beta) \in \gamma(\sigma, \eta)$. This occurs precisely when $\sigma(v) = \alpha(v)$, $\eta(v) = \beta(v)$, and moreover σ and η are such that $c_i = \alpha(u_i) = \beta(u_i)$ for all u_i.

Consider the possibilities for η. Configuration β determines $\eta(v)$, and there are $(k-1)^t$ choices for η given $\eta(v)$ (consider choosing the colours for $u_1, \ldots, u_t$, which cannot be $\eta(v)$). Now consider σ: the colour $\sigma(v)$ is determined by α, as are $(c_1, \ldots, c_t)$. Thus by (17) there are at most $\lceil (\frac{k-1}{k-2})^t \rceil$ possibilities for $(\sigma(u_1), \ldots, \sigma(u_t))$, which determines σ. Putting this together, the total number of colourings σ and η that satisfy $(\alpha, \beta) \in \gamma(\sigma, \eta)$ is at most $(k-1)^t \left\lceil \left(\frac{k-1}{k-2}\right)^t \right\rceil$. Substituting this and (20) into (19), we conclude

$$\rho_w(\alpha, \beta) \leq 10(1)(k-1)^t \left\lceil \left(\frac{k-1}{k-2}\right)^t \right\rceil \frac{k-1}{(k-1)^{t+1}} \leq 20 \left(\frac{k-1}{k-2}\right)^t.$$

Case 2: α and β differ on the colour of u_i for some i. In this case, $w(\gamma(\sigma, \eta)) = i^{-2}$. Also, since there exists a colouring of V_i in which u_i has colour $\beta(u_i)$ (recalling our assumption that $u_i \notin \overline{\partial} T$), Corollary 5.5 implies

$$K_B[\alpha \to \beta] \geq (k\tau_{V_i})^{-1}. \tag{21}$$

How many paths in $\gamma(\sigma, \eta)$ use the transition (α, β)? We consider subcases for σ and η.

Case 2.1: $\sigma(v) = \eta(v)$. Recall that in this case a special, simple canonical path is used. We know $\sigma(v) = \eta(v) = \alpha(v) = \beta(v)$. Also, we know $\sigma(u_j) = \alpha(u_j)$ for all $j \geq i$, and $\eta(u_j) = \beta(u_j)$ for all $j \leq i$. So for each $j < i$ there are $(k-1)$ possibilities for $\sigma(u_j)$, and for each $j > i$ there are $(k-1)$ possibilities for $\eta(u_j)$. The total number of possibilities for σ and η is therefore at most $(k-1)^{t-1}$.

Case 2.2: $\sigma(v) \neq \eta(v)$ and (α, β) is the first change to u_i in $\gamma(\sigma, \eta)$. That is, (α, β) is the first change in Step 1 of the canonical path description in Section 5.1.3. In this case we know $\sigma(v) = \alpha(v)$, $\sigma(u_j) = \alpha(u_j)$ for all $j \geq i$, $b_i = \beta(u_i)$, and $c_j = \beta(u_j)$ for all $j < i$. We wish to count the number of colourings σ and η that satisfy these conditions.

First consider η. There are at most $k-1$ possibilities for $\eta(v)$, since $\eta(v) \neq \sigma(v) = \alpha(v)$. Given $\eta(v)$, there are $k-1$ possibilities for $\eta(u_j)$ for each $j \neq i$. Note that β determines b_i, from which $\eta(v)$ determines $\eta(u_i)$. Thus the total number of possibilities for η is $(k-1)^t$.

Next consider σ. Note that α determines $\sigma(v)$ and also $\sigma(u_j)$ for all $j \geq i$. Also, β determines c_j for all $j < i$. Then (17) implies that the number of possibilities for $\sigma(u_1), \ldots, \sigma(u_t)$ is at most $\left\lceil \left(\frac{k-1}{k-2}\right)^{i-1} \right\rceil$. We conclude that for this subcase the total number of possibilities for σ and η is at most

$$\left\lceil \left(\frac{k-1}{k-2}\right)^{i-1} \right\rceil (k-1)^t.$$

Case 2.3: $\sigma(v) \neq \eta(v)$ and (α, β) is the second change to u_i in $\gamma(\sigma, \eta)$. This is the second change in Step 1 of the canonical paths description in Section 5.1.3. This case is nearly identical to Case 2.2; the only difference is that for node u_i we know $b_i = \alpha(u_i)$ and $c_i = \beta(u_i)$.

The only effect that this has on the analysis is that now c_i is determined instead of $\sigma(u_i)$. Given c_i (instead of $\sigma(u_i)$), the factor due to (17) becomes $\lceil (\frac{k-1}{k-2})^i \rceil$. We conclude that the number of possibilities for σ and η is at most

$$\left\lceil \left(\frac{k-1}{k-2}\right)^{i} \right\rceil (k-1)^t.$$

Case 2.4: $\sigma(v) \neq \eta(v)$ and (α, β) is the third change to u_i in $\gamma(\sigma, \eta)$. This is the first change in Step 3 of the canonical paths description in Section 5.1.3. In this case, $\eta(v) = \alpha(v)$, $c_j = \alpha(u_j)$ for all $j < i$, and $a_j = \beta(u_j)$ for all $j > i$. Further, $c_i = \alpha(u_i)$ and $\eta(u_i) = \beta(u_i)$.

Note first that α determines $c_1, \ldots, c_i$ and β determines $a_{i+1}, \ldots, a_t$. Consider possibilities for η: β determines $\eta(v)$ and $\eta(u_i)$. For each $j \neq i$, there are $(k-1)$ possibilities for $\eta(u_j)$. The number of possibilities for η is thus at most $(k-1)^{t-1}$.

Now consider σ. There are at most $k-1$ possibilities for $\sigma(v)$. Recall that colours $a_{i+1}, \ldots, a_t$ and colours $c_1, \ldots, c_i$ are determined. But then $\sigma(u_{i+1}), \ldots, \sigma(u_t)$ can be recovered (using $\sigma(v)$) and by (17) there are at most $\lceil (\frac{k-1}{k-2})^i \rceil$ possibilities for $(\sigma(u_1), \ldots, \sigma(u_t))$. The number of possibilities for σ is therefore at most $(k-1)\left\lceil \left(\frac{k-1}{k-2}\right)^i \right\rceil$. We conclude that for this subcase the total number of possibilities for σ and η is at most

$$\left\lceil \left(\frac{k-1}{k-2}\right)^i \right\rceil (k-1)^t.$$

Case 2.5: $\sigma(v) \neq \eta(v)$ and (α, β) is the fourth change to u_i in $\gamma(\sigma, \eta)$. This is the second change in Step 3 of the canonical paths description in Section 5.1.3. This case is nearly identical to Case 2.4; the only difference is that for node u_i we know $\eta(u_i) = \alpha(u_i)$ and $a_i = \beta(u_i)$.

The only effect that this has on the analysis is that now a_i is determined instead of c_i. This causes the factor due to (17) to become $\lceil (\frac{k-1}{k-2})^{i-1} \rceil$. We conclude that the number of possibilities for σ and η is at most

$$\left\lceil \left(\frac{k-1}{k-2}\right)^{i-1} \right\rceil (k-1)^t.$$

Case 2.6: $\sigma(v) \neq \eta(v)$ and (α, β) is the fifth change to u_i in $\gamma(\sigma, \eta)$. This is the change in Step 4 of the canonical paths description in Section 5.1.3. In this case we know $\eta(v) = \alpha(v)$, $a_j = \alpha(u_j)$ for all $j > i$, and $\eta(u_j) = \beta(u_j)$ for all $j < i$. For u_i, we know $a_i = \alpha(u_i)$ and $\eta(u_i) = \beta(u_i)$.

In this case there are at most $(k-1)$ choices for $\sigma(v)$. The colours $a_i, \ldots, a_t$ plus $\eta(v)$ are determined by α. From these colours (plus $\sigma(v)$) the colours $\sigma(u_i), \ldots, \sigma(u_t)$ are determined. Furthermore, $\eta(u_1), \ldots, \eta(u_i)$ are determined from β.

From this point onward the analysis is identical to that of Case 2.1. Taking into account the $k-1$ possibilities for $\sigma(v)$, we conclude that the number of possible options for σ and η is at most $(k-1)^t$.

This concludes our subcase analysis. Summing over all subcases, the total number of possibilities for σ and η, given that (α, β) is a change in the colouring of u_i, is at most $12\left(\frac{k-1}{k-2}\right)^i (k-1)^t$. Substituting this and (21) into (19), we have

$$\rho_w(\omega, \beta) \leq 120 i^2 \left(\frac{k-1}{k-2}\right)^i (k-1)^t \left(\frac{\tau_{V_i} k}{(k-1)^{t+1}}\right) \leq 180 i^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}.$$

This concludes our case analysis. Cases 1 and 2 (and the fact that $\tau_{V_t} \geq 1$) imply

$$\rho_w \leq \max_{1 \leq i \leq t} 180 i^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}.$$

Applying the canonical paths bound and Proposition 3.2 we conclude

$$\tau_T^\sigma \leq \tau_D^\sigma \leq 180 \max_{1 \leq i \leq t} i^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i} \tag{22}$$

as required.

Remark 5.6. We note the effect of using the weighted canonical paths bound. If we had used the standard canonical paths bound, then we would replace the factor of i^2 in Lemma 5.1 by the maximum length of a path, which is $5\Delta + 1$. This would lead to an extra factor of Δ^h on our mixing time bound, where h is the "height" of the tree under our given decomposition method. However, since each recursive application of our decomposition guarantees only a reduction of the tree size by half, this extra factor is can be as large as $\Delta^{\log n} = n^{\log \Delta}$. This leads to a bound of $n^{O(\log \Delta + \Delta/k \log \Delta)}$ on the mixing time of the Glauber dynamics, which is weaker than $n^{O(1+\Delta/k \log \Delta)}$.

Remark 5.7. We also note the effect of using the weighted block dynamics. If we had applied Proposition 3.1 instead of Proposition 3.2, the bound in (21) would become $K_B[\omega \to \beta] \geq (k\tau)^{-1}$, where $\tau = \max_i \tau_{V_i}$. This would lead to a bound of $\tau_T^\sigma \leq ct^2 \left(\frac{k-1}{k-2}\right)^t \max_{1 \leq i \leq t} \tau_{V_i}$ for Lemma 5.1. With this modified Lemma, the bound in (16) would become $\tau_2(n) \leq ct^2 \left(\frac{k-1}{k-2}\right)^t \tau_2(\lceil n/2 \rceil)$, leading to a mixing time bound of $n^{O(1+\Delta/k)}$, which is weaker than $n^{O(1+\Delta/k \log \Delta)}$.

5.1.5 Handling the case $u_i \in \overline{\partial} T$

We now modify the proof of Lemma 5.1 to handle the case that there exist i such that $u_i \in \overline{\partial} T$. Note first that if $u_i \in \overline{\partial} T$, then u_i can be adjacent to only one node in ∂T, since $|\partial V_i| \leq 2$ and $v \in \partial V_i$.

We used the assumption $u_i \notin \overline{\partial} T$ when defining our canonical paths with colours a_i, b_i, c_i: this allowed us to assume that there existed colourings of V_i in which the colour of u_i was a_i (or b_i, or c_i). If $u_i \in \overline{\partial} T$, it's possible that one or more of these colours will conflict with the boundary configuration, so it may not be possible to use these colours in the construction of our canonical path. Our approach will be to construct alternative "fixed" versions of these colours, then show that it is easy to reconstruct the original colours from their fixed counterparts.

For each $1 \leq i \leq t$ we define colours a'_i, b'_i, and c'_i. If there exists a colouring of V_i in which u_i has colour a_i, then set $a'_i = a_i$. Otherwise it must be that $u_i \in \overline{\partial} T$, and in this case we set a'_i to be any other colour not in $\{a_i, \alpha(v), \eta(v)\}$. Such a colour must exist since $k \geq 4$. Also, since u_i is adjacent to at most one vertex in ∂T and $a'_i \neq a_i$, there must exist some colouring of V_i in which the colour of u_i is a'_i. We define b'_i and c'_i in a similar way, corresponding to b_i and c_i. We then use a'_i, b'_i, and c'_i in the definition of our canonical paths instead of colours a_i, b_i, and c_i.

How does this affect our analysis? When we compute the number of (α, η) such that $(\omega, \beta) \in \gamma(\alpha, \eta)$, there may be one or more vertices u_j for which we have to reconstruct a_j, b_j, and/or c_j from their fixed versions a'_j, b'_j, and/or c'_j. For any such j, if $u_j \notin \overline{\partial} T$ (which is determined by the structure of T and does not depend on the configuration) then the fixed and original colours are the same. When $u_j \in \overline{\partial} T$, there will be at most two possibilities for each of these colours (ie. either $a_j = a'_j$ or a_j is the colour that conflicts with the boundary configuration, and similarly for b_j and c_j). Since there are at most 2 nodes u_j in $\overline{\partial} T$, this adds a factor of at most 4 to our analysis.

We conclude that the analysis for Lemma 5.1 leads to the same result, with an extra factor of 4. This gives Lemma 6.1 as required.

5.2 A balanced mapping function

We now present the mapping function f used in the proof of Lemma 5.1, which satisfies (17). We will actually prove the existence of f in the following, equivalent, arena.

Lemma 5.8. *For all $k \geq 3$ and $1 \leq t \leq \Delta$, there is a function $f\colon [k-1]^t \to [k-2]^t$ such that for all*

$y \in [k-2]^t$, $z \in [k-1]^t$, *and* $1 \le i \le t$,

$$|\{x : (x_j = z_j \ \forall j > i) \wedge (f(x)_j = y_j \ \forall j \le i)\}| \le \left\lceil \left(\frac{k-1}{k-2}\right)^i \right\rceil .$$

Proof. Given $x \in [k-1]^t$, interpret x as the representation of an integer d in base $k-1$. Let y be the representation of $d \mod (k-2)^t$ in base $k-2$. Then we define f to be the function mapping x to y.

To see that f satisfies the required property, fix some $1 \le i \le t$. Consider the image of f on all $z \in [k-1]^t$ such that $x_j = z_j$ for all $j > i$, in lexicographic order. This image is simply a sequence of $(k-1)^i$ consecutive integers, modulo $(k-2)^t$, in base $k-2$. In particular, each pattern of i least significant digits occurs once every $(k-2)^i$ values, and hence occurs at most $\left\lceil \left(\frac{k-1}{k-2}\right)^i \right\rceil$ times over the sequence of integers. This is therefore a bound on the size of the preimage of f restricted to this set, as required. □

6 Extending to general boundary conditions

In the proofs of Theorem 1.1 and Lemma 5.1, we were careful to consider only subtrees with at most 2 boundary vertices. This was enough to prove Theorem 1.1 and simplified our arguments. However, this restriction can be relaxed when $k > 4$. Indeed, all that is required by our technique is that $|\partial T| \le k-2$.

Theorem 1.4 is a variant of Theorem 1.1 that uses this relaxation. We will prove it by making use of the following variant of Lemma 5.1.

Lemma 6.1. *Suppose $k \ge 3$ and let T be a subtree of a tree G with $b := |\partial T| \le k-2$ and let $\sigma \in \Omega$ be a boundary condition for T. Choose $x \in T$ and consider $D_x = \{\{x\}, V_1, \dots, V_t\}$, where $1 \le t \le \Delta$. Suppose $|\partial V_i| \le b$ for each V_i. Then, for some constant c,*

$$\tau_T^\sigma \le cb2^b \max_{1 \le i \le t} i^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}.$$

Before proving Lemma 6.1, we will discuss how it implies Theorem 1.4. Indeed, this implication follows the deduction of Theorem 1.1 from Lemma 5.1 almost exactly. Define $\tau_b(n)$ as in Section 5.1. Then the argument from Lemma 5.2 yields

$$\tau_{b-1}(n) \le cb2^b \max_{1 \le i \le t} i^2 \left(\frac{k-1}{k-2}\right)^i \tau_b(\lfloor n/i \rfloor). \tag{23}$$

To bound $\tau_b(n)$, we proceed as in Lemma 5.3. Define the *boundary tree* of T to be the union of the paths between vertices of ∂T in T. Note that this is, indeed, a subtree of T. Choose a vertex $x \in T$ that separates T into subtrees with at most $n/2$ vertices, then let y be the vertex that is the least ancestor of x that is in the boundary tree. Just as in Lemma 5.3 we can apply Lemma 6.1 (using the partition $D_y = \{\{y\}, V_1, \dots, V_t\}$), then use (23), to obtain the bound

$$\tau_b(n) \le c^2 b^2 2^{2b} \max_{1 \le i \le t} i^2 \left(\frac{k-1}{k-2}\right)^{i+1} \tau_b(n/i).$$

Induction then implies that $\tau_b(n) \le n^{d(1+b+\Delta/k \log \Delta)}$ for some sufficiently large constant d. This follows in precisely the same way that (16) implies Theorem 1.1 in Section 5.

It remains to give the proof of Lemma 6.1, which mirrors the proof of Lemma 5.1 with two changes. First, we require the following more general version of Lemma 5.4.

Lemma 6.2. *Choose $S \subseteq T$ with $b := |\partial S| \leq k-2$ and boundary configuration ξ, and suppose $x \in \overline{\partial} S$. Choose $c \in A$ and suppose there exists some $\eta \in \Omega_S^\xi$ with $\eta(x) = c$. Then $\pi_S^\xi[\omega : \omega(x) = c] \geq 1/(b+1)k$.*

Proof. Think of x as the root of S and consider choices of colours top-down. Then for any colour c, since $|\partial S| = b$, $(k-1)^{n-2}(k-b+1) \leq |\{\eta \in \Omega_S^\xi : \eta(x) = c\}| \leq (k-1)^{n-1}$, as long as there is at least one colouring in which x has colour c. Thus

$$\pi_S^\xi[\eta : \eta(x) = c] = \frac{|\{\eta \in \Omega_S^\xi : \eta(x) = c\}|}{|\Omega_S^\xi|} \geq \frac{(k-1)^{n-2}(k-b+1)}{(k-1)^{n-1}(k-2) + (k-1)^{n-2}(k-b+1)}$$
$$= \frac{1}{1 + \frac{(k-2)(k-1)}{k-b+1}} \geq \frac{1}{(b+1)k}$$

as required. □

This change to Lemma 5.4 affects (19), adding a factor of $(b+1)$ to our analysis.

Second, recall our discussion of the case that $u_i \in \overline{\partial} T$ at the end of the proof of Lemma 5.4. We will give a simple extension of this approach. For each $1 \leq i \leq t$ we will define colours a_i', b_i', and c_i'. If there exists a colouring of V_i in which u_i has colour a_i, then set $a_i' = a_i$. Otherwise it must be that $u_i \in \overline{\partial} T$, and in this case we set a_i' to be any other colour not in $\{a_i, \sigma(v), \eta(v)\}$. This colour a_i' must exist since $k \geq b+2$, and there will exist some colouring of V_i in which u_i has colour a_i'. We define b_i' and c_i' in a similar way. We then use a_i', b_i', and c_i' in the definition of our canonical paths instead of colours a_i, b_i, and c_i.

How does this affect our analysis? When we compute the number of (σ, η) such that $(\alpha, \beta) \in \gamma(\sigma, \eta)$, we must reconstruct a_j, b_j, and/or c_j from their fixed versions. If $u_j \notin \overline{\partial} T$ then the fixed and original colours are the same. When $u_j \in \overline{\partial} T$, there will be at most two possibilities for each of these colours (ie. either $a_j = a_j'$ or a_j is the colour that conflicts with the boundary configuration, and similarly for b_j and c_j). Since there are at most b nodes u_j in $\overline{\partial} T$, this adds a factor of at most 2^b to our analysis.

We conclude that the analysis for Lemma 5.1 leads to the same result, with an extra factor of $(b+1)2^b$. This gives Lemma 6.1 as required.

7 A Lower Bound

We will now prove Theorem 1.3: that the mixing time for the Glauber dynamics on the complete tree of degree Δ is $n^{\Omega(\Delta/k \log \Delta)}$. Recall that we are working in the continuous-time setting.

Let T be the complete tree with degree Δ and height h, where a singleton is said to have height 0. Let $n = |T|$ and note $h \geq \log_\Delta n - 1$. For each $0 \leq i \leq h$, define T_i by

$$T_i = \frac{1}{10(k-1)} \left(\frac{1}{20(k-1)^2} \left(1 - \frac{9}{10(k-1)}\right)^{-\Delta} \right)^i.$$

Note that, in particular,

$$\begin{aligned}
T_h &= \frac{1}{10(k-1)}\left(\frac{1}{20(k-1)^2}\left(1-\frac{9}{10(k-1)}\right)^{-\Delta}\right)^h \\
&\geq \frac{1}{10(k-1)}\left(\frac{1}{20(k-1)^2}\left(1-\frac{9}{10(k-1)}\right)^{-\Delta}\right)^{\log_\Delta n-1} \\
&> \frac{1}{10(k-1)}(2^{\log n})^{-\log(20(k-1)^2)/\log\Delta}(2^{\log n})^{\frac{9}{10(k-1)}\Delta/\log\Delta} \\
&= n^{\Omega(\Delta/k\log\Delta-\log k/\log\Delta)} \\
&= n^{\Omega(\Delta/k\log\Delta)}.
\end{aligned}$$

Choose a vertex v, say at height i. We will prove the following:

Lemma 7.1. *The probability that v changes colour before time T_i during a run of the Glauber dynamics, starting from a uniformly chosen colouring, is at most $\frac{1}{10(k-1)}$.*

Before proving the Lemma, we show how to use it to get our lower bound. Lemma 7.1 implies that the probability that the root r changes colour in T_h steps is at most $\frac{1}{10(k-1)}$. There must therefore be a particular initial colouring η_0 such that $Pr[\eta_{T_h}(r) = \eta_0(r)] \geq 1 - \frac{1}{10(k-1)}$.

We now have

$$||\eta_{T_h} - \pi||_{TV} \geq 1 - \frac{1}{10(k-1)} - \frac{1}{k} > \frac{1}{3}$$

and therefore the mixing time must be greater than $T_h \geq n^{\Omega(\Delta/k\log\Delta)}$, giving us the desired bound.

Proof of Lemma 7.1. Let $\sigma = \sigma_0$ denote an initial uniformly chosen colouring. Denote by σ_t the colouring at time t. Note that σ_t is uniformly distributed for every t.

Recall that i is the height of v. We will first prove the claim for the case $i = 0$. When $i = 0$ we have that v is a leaf. Then $T_0 = \frac{1}{10(k-1)}$ and the time until v is selected by the Glauber dynamics is exponentially distributed with mean 1. Thus the probability that v changes colour before time T_0 is at most $1 - e^{-1/(10(k-1))} < \frac{1}{10(k-1)}$, as required.

We will now assume $i \geq 1$, so that v is not a leaf. Assume now that $T_i \leq T_{i-1}$. This implies that

$$\left(\frac{1}{20(k-1)^2}\left(1-\frac{9}{10(k-1)}\right)^{-\Delta}\right) \leq 1.$$

But then we have $T_i \leq T_{i-1} \leq T_{i-2} \leq \ldots \leq T_0$. By the same argument as for the case $i = 0$, the probability that v is selected by the Glauber dynamics before time T_i is at most $\frac{1}{10(k-1)}$. Thus $\frac{1}{10(k-1)}$ is a bound on the probability that v changes colour and we are done. We can therefore assume that $T_i > T_{i-1}$.

Let the children of v be $u_1, \ldots, u_\Delta$. We say that the children of v *avoid a colour* at time t if there is a colour $a \neq \sigma_t(v)$ such that $\sigma(u_j) \neq a$ for all $1 \leq j \leq \Delta$. Let A be the event that the children of v avoid a colour at some time before T_i.

We will now prove the stronger result that $Pr[A] \leq \frac{1}{10(k-1)}$. Note that this is truly a stronger result, since if v changes colour before time T_i then it must be that event A has occurred. In fact, an even stronger event must have occurred, since the same colour must not be on the parent of v as well; however, focusing on event A will suffice for our lower bound.

We proceed by induction on i. Consider first the base case $i = 1$, so that the children of v are leaf nodes. For any time $t \geq 0$, define the time interval $I_t := [t, t + T_0)$. Let $A(t)$ denote the event that there is at least one time in I_t at which the children of v avoid a colour. Given colour $c \neq \sigma_t(v)$, let $A(t, c)$ be the event that there is at least one time in I_t at which the children of v avoid colour c. Finally, for all $1 \leq j \leq \Delta$, let $A(t, j, c)$ denote the event that either $\sigma_t(u_j) \neq c$ or u_j is selected by the Glauber dynamics at some time in I_t.

The union bound implies that

$$\begin{aligned}
Pr[A(t,j,c)] &\leq Pr[\sigma_t(u_j) \neq c] + Pr[u_j \text{ is selected in } I_t] \\
&< 1 - \frac{1}{k-1} + 1 - e^{-1/(10(k-1))} \\
&\leq 1 - \frac{9}{10(k-1)}.
\end{aligned}$$

Furthermore, we note that the events $A(t, 1, c), \ldots, A(t, \Delta, c)$ are mutually independent (due to the uniformity of σ_t and the independence of vertex selection by the Glauber dynamics). This implies

$$\begin{aligned}
Pr[A(t)] \leq \sum_{c \neq \sigma_t(v)} Pr[A(t,c)] &\leq \sum_{c \neq \sigma_t(v)} \prod_{1 \leq j \leq \Delta} Pr[A(t,j,c)] \\
&\leq (k-1)\left(1 - \frac{9}{10(k-1)}\right)^{\Delta}.
\end{aligned}$$

Denote by X_t the random variable

$$X_t = \sum_{j=0}^{\lfloor t/T_0 \rfloor} \mathbf{1}_{A(jT_0)}.$$

Then X_t denotes the number of time intervals I_t of length T_0, starting at times between 0 and t that are multiples of T_0, in which the children of v avoid a colour at some point. Note that, for all $t > T_0$,

$$\begin{aligned}
E[X_t] &= \sum_{j=0}^{\lfloor t/T_0 \rfloor} Pr[A(jT_0)] \\
&\leq (\lfloor t/T_0 \rfloor + 1)(k-1)\left(1 - \frac{9}{10(k-1)}\right)^{\Delta}. \\
&\leq \frac{2t}{T_0}(k-1)\left(1 - \frac{9}{10(k-1)}\right)^{\Delta}
\end{aligned}$$

Then since $T_i > T_{i-1}$, it must be that

$$\begin{aligned}
E[X_{T_1}] &\leq \frac{2T_1}{T_0}(k-1)\left(1 - \frac{9}{10(k-1)}\right)^{\Delta} \\
&\leq 2\left(\frac{1}{20(k-1)^2}\right)\left(1 - \frac{9}{10(k-1)}\right)^{-\Delta}(k-1)\left(1 - \frac{9}{10(k-1)}\right)^{\Delta} \\
&= \frac{1}{10(k-1)}
\end{aligned}$$

so by Markov's Inequality

$$Pr[X_{T_1} \geq 1] \leq \frac{1}{10(k-1)}.$$

But now we note that $Pr[A] \leq Pr[X_{T_1} \geq 1]$, and hence $Pr[A] \leq \frac{1}{10(k-1)}$ as required. This concludes the base case.

Now consider general $i \geq 1$. Let the children of v be $u_1, \ldots, u_\Delta$. For each $t \geq 0$, let I_t be the interval of length T_{i-1} starting at time t. Let $A(t)$ denote the event that there is at least one time in I_t at which the children of v avoid a colour. Given colour $c \neq \sigma_t(v)$, let $A(t,c)$ be the event that there is at least one time in I_t at which the children of v avoid colour c. Let $A(t,j,c)$ be the event that either $\sigma_t(u_j) \neq c$ or the children of u_j avoid a colour at some time in I_t. Note that $A(t)$ and $A(t,c)$ are defined similarly as in the base case, while $A(t,j,c)$ is defined differently.

Now note that

$$\begin{aligned}
Pr[A_{t,j,c}] &\leq Pr[\sigma_t(u_j) \neq c] + Pr[\text{children of } u_j \text{ avoid a colour at some time in } I_t] \\
&\leq 1 - \frac{1}{k-1} + \frac{1}{10(k-1)} \\
&= 1 - \frac{9}{10(k-1)}
\end{aligned}$$

by induction.

We now claim that the events $A(t,1,c), \ldots, A(t,\Delta,c)$ are mutually independent. For each i, $1 \leq i \leq \Delta$, let T_i be the subtree rooted at u_i. Consider $\sigma_t(T_i)$, the colouring of T_i at time t, plus the sequence of attempted assignments of colours to vertices in T_i in the time period I_t. We claim that event $A(t,i,c)$ is determined entirely by $\sigma_t(T_i)$ plus these attempted assignments. This follows since the colouring of T_i can be affected by other events only if the colour of u_i changes, but event $A(t,i,c)$ is a prerequisite for a change in the colour of u_i.

But now note that the trees $T_1, \ldots, T_\Delta$ are disjoint, so the attempted colouring events over different trees are independent by the definition of the Glauber dynamics. Also, the colourings of these trees at time t are also independent, since σ_t is uniformly chosen from the set of colourings. We conclude that the events $A(t,1,c), \ldots, A(t,\Delta,c)$ must be independent as well.

Using mutual independence, we have that

$$\begin{aligned}
Pr[A_t] &\leq \sum_c Pr[A_{t,c}] \\
&\leq \sum_c \prod_j Pr[A_{t,j,c}] \\
&\leq (k-1)\left(1 - \frac{9}{10(k-1)}\right)^\Delta.
\end{aligned}$$

Just as in the base case, we define random variable X_t as

$$X_t = \sum_{j=0}^{\lfloor t/T_{i-1} \rfloor} \mathbf{1}_{A(jT_{i-1})}.$$

Then, by the same argument used in the base case, we find

$$Pr[A] \leq Pr[X_{T_i} \geq 1] \leq \frac{1}{10(k-1)}$$

as required.

□

8 Open Problems

Our results raise questions about the Glauber dynamics on planar graphs of bounded degree. Hayes, Vera and Vigoda [9] noted that when $\Delta \geq n^\eta$ for any $\eta > 0$ then certain trees require $k \geq c\Delta/\log\Delta$ for polytime mixing, where c is an absolute constant. Our results imply that the same is true for any Δ that grows with n. But for $\Delta = O(1)$, Theorem 1.1 shows that no trees require $k > 4$. Is there a constant K such that for every $k \geq K$ and constant Δ, the Glauber dynamics mixes in polytime on k-colourings of every planar graph with maximum degree Δ?

Another question is how far Theorem 1.4 can be extended. In other words, how many leaves can we fix and still guarantee polytime mixing? It is easy to fix the colours of $k-1$ neighbours of each of two adjacent vertices u, v so that the chain is not ergodic, so the answer lies between $k-2$ and $2k-2$.

Acknowledgements

We thank Nayantara Bhatnagar, Jian Ding, Thomas Hayes, Juan Vera and Eric Vigoda for some insightful discussions. We also thank two anonymous referees for their many helpful comments.

References

[1] D. Achlioptas and A. Coja-Oghlan. Algorithmic barriers from phase transitions. In *Proc. 49th IEEE Symp. on Foundations of Computer Science*, pages 793–802, 2008.

[2] 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.

[3] N. Bhatnagar, J. Vera, E. Vigoda, and D. Weitz. Reconstruction for colorings on trees. To appear. Available at arXiv:0711.3664v2, 2008.

[4] G. Brightwell and P. Winkler. Random colorings of a cayley tree. In Bollobs and Bela, editors, *Contemporary Combinatorics*, pages 247–276. 2002.

[5] P. Diaconis and L. Saloff-Coste. Comparison theorems for reversible markov chains. *Annals of Applied Probability*, 3(3):696–730, 1993.

[6] P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of markov chains. *Annals of Applied Probability*, 1(1):36–61, 1991.

[7] M. Dyer, L. Goldberg, and M. Jerrum. Systematic scan for sampling colorings. *Annals of Applied Probability*, 16(1):185–230, 2006.

[8] L. Goldberg, M. Jerrum, and M. Karpinski. The mixing time of glauber dynamics for coloring regular trees. *Random Structures and Algorithms*, 36(4):464–476, 2010.

[9] T. Hayes, J. Vera, and E. Vigoda. Randomly coloring planar graphs with fewer colors than the maximum degree. In *Proc. 38th ACM Symp. on Theory of Computing*, pages 450–458, 2007.

[10] M. Jerrum and A. Sinclair. Approximating the permanent. *SIAM Journal on Computing*, 18(6):1149–1178, 1989.

[11] J. Jonasson. Uniqueness of uniform random colourings of regular trees. *Statistics and Probability Letters*, 57(3):243–248, 2002.

[12] F. Martinelli. Lectures on glauber dynamics for discrete spin models. *Lecture Notes in Mathematics*, 1717:93–191, 2000.

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

[14] D. Randall and P. Tetali. Analyzing glauber dynamics by comparison of markov chains. In *Proc. 3rd Latin American Symposium on Theoretical Informatics*, pages 292–304, 1998.

[15] A. Sinclair. Improved bounds for mixing rates of markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1:351–370, 1992.

[16] A. Sly. Reconstruction of random colorings. *Communications in Mathematical Physics*, 288(1):943–961, 2009.

[17] P. Tetali, J. Vera, E. Vigoda, and L. Yang. Phase transition for the mixing time of the glauber dynamics for coloring regular trees. In *Proc. 21st ACM Symp. on Discrete Algorithms*, pages 1646–1656, 2010.