The Glauber dynamics for colourings of bounded degree trees

B. Lucier[*]   M. Molloy[†]

October 22, 2008

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 $n^{\Theta(1+\Delta/(k \log \Delta))}$. For $k \geq 4$ we show that the mixing time on *every* tree is at most $n^{O(1+\Delta/(k \log \Delta))}$. Our proof uses a weighted canonical paths analysis and a variation of the block dynamics in which we exploit the differing relaxation times of blocks.

---

[*] Dept of Computer Science, University of Toronto, blucier@cs.toronto.edu.

[†] Dept of Computer Science, University of Toronto, molloy@cs.toronto.edu.

1 Introduction

The Glauber dynamics is a Markov chain for sampling k-colourings of an input graph. Such chains have generated a great deal of interest for a variety of reasons. For one thing, 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. For another, k-colourings are equivalent to the antiferromagnetic Potts model from statistical physics, and there is a large body of research into Markov chains for this and similar models.

The Glauber dynamics has received a very large part of this interest (see eg. [12]). It is particularly appealing because it is a natural and simple algorithm and it underlies more substantial procedures such as block dynamics and systematic scan (see [12, 5]). It is also commonly used in practice, eg. in simulations of physical systems. Furthermore, it is very closely related to other important areas such as infinite-volume Gibbs distributions[2, 10, 13]. It is generally conjectured that the Glauber dynamics mixes in polynomial time on every graph of maximum degree Δ so long as $k \geq \Delta + 2$. Vigoda[18] has shown polynomial mixing time for $k \geq \frac{11}{6}\Delta$.

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, and there are much easier ways to do so. Nevertheless, people have studied the the Glauber dynamics on trees as a means of understanding how the chain performs on more general graphs, and because the performance on trees is particularly relevant to related areas such as Gibbs distributions. Berger et al. [1] 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[7] showed that it mixes in polytime 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 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, is the mixing time superpolynomial for the complete $(\Delta - 1)$-ary tree with $k = 3$ and $\Delta = O(1)$? Proposition 2.5 of Berger et al. [1] shows that the mixing time is in fact 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). In our first theorem, we determine the degree of the polynomial, up to a constant multiple.

Theorem 1.1. *For $k \geq 3$, the mixing time of the Glauber dynamics on k-colourings of the complete Δ-regular tree is $n^{\Theta(1+\Delta/k\log\Delta)}$.*

To be clear: in all of our theorems, the implicit constants are independent of Δ, k and (in Theorem 1.4) b.

Recently and independently, Goldberg, Jerrum and Karpinski[6] obtained the same lower bound as that in Theorem 1.1 and an upper bound of $n^{O(\Delta/\log\Delta)}$ for the case where $k, \Delta = O(1)$. They use different techniques than those used herein.

For $k \geq 4$ we can show that the same upper bound on the mixing time holds for *every* tree with $\Delta = O(1)$:

Theorem 1.2. *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 constant value of $\Delta = O(1)$, we have polytime mixing on every tree and for every $k \geq 4$. But if Δ grows with n, no matter how slowly, then on some trees we require the $\Omega(\Delta/\log\Delta)$ colours for polytime mixing that Hayes, Vera and Vigoda noted were required at $\Delta = n^{\epsilon}$.

Let us give some intuition into the difficulties that occur when $k = o(\Delta/\log\Delta)$. A bound of $k \geq \Delta+2$ is natural since it ensures that no vertex will ever be *frozen*; i.e. there will always be at least one colour that it can switch to. (It is also a natural bound because it corresponds to the threshold for unique infinite-volume Gibbs distributions[10].) Much of the difficulty in showing rapid mixing for smaller values of k is in dealing with the frozen variables. From this perspective, $k \geq C\Delta/\log\Delta$ for $C > 1$ becomes another natural threshold: if the neighbours of a vertex are assigned independently random colours (e.g. in the steady state distribution) then we expect that the vertex will not be frozen. But if

$k < (1-\epsilon)\Delta/\log\Delta$, then even in the steady state distribution the vast majority of the degree Δ vertices on a tree will be frozen. This creates most of our difficulties.

If the children of a vertex u change colours enough times, then eventually u will become unfrozen and change colours. Roughly speaking, this allows vertices to unfreeze, level by level, much like in the level dynamics of [7]. Of course, this is a slow process: the number of times that the children of u have to change before u is unfrozen is (roughly) exponential in Δ/k. However, this value is manageable for $\Delta = O(1)$. It takes a long time for this process to reach the top level, but the running time works out to be 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, this proof doesn't work and we have to take a more complicated approach.

For the smallest value of k for which we are interested, i.e. $k = 3$, we can only prove that the same upper bound from Theorem 1.2 holds for *balanced* trees. We say that a rooted tree T is ϵ-balanced if for every v, the subtree T_v rooted at v has the following property: For every child u of v, $|T_u| \leq (1-\epsilon)|T_v|$.

Theorem 1.3. *For any constant $\epsilon > 0$, the Glauber dynamics on 3-colourings of any ϵ-balanced tree with maximum degree Δ mixes in time $n^{O(1+\Delta/\log\Delta)}$.*

Clearly the complete $(\Delta-1)$-ary tree is $(1-(\Delta-1)^{-1})$-balanced, and so Theorem 1.1 shows that Theorem 1.3 is tight up to the constant factor in the exponent.

Our work raises the possibility that for trees with maximum degree $\Delta = O(1)$, the $(\Delta-1)$-ary tree is *not* the most problematic tree for the Glauber dynamics on 3-colourings, contrary to the implicit guess in the aforementioned open question from [7]. The structures that appear particularly difficult to deal with are trees containing long path-like pieces. Note that on a path, it is possible for a 3-colouring to freeze every vertex other than the two endpoints, in contrast to the complete tree in which $\Theta(n)$ of the vertices are always unfrozen. The value $k = 3$ is already known to be at least a bit more problematic than other values for the Glauber dynamics on paths: Dyer, Goldberg and Jerrum[5] showed that the mixing time is $\Theta(n^3 \log n)$ whereas it is only $O(n\log n)$ for $k \geq 4$ (as implied by eg. [8, 14]).

The proofs of our main theorems make use of a variation of the well-known *block dynamics* which takes account of differing mixing times amongst the blocks. To the best of our knowledge, this is the first time that such a variation has been used.

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 a tree. Markov chains on trees with the colours of leaves fixed are well-studied. In the case where every leaf is fixed, Martinelli, Sinclair and Weitz[13] proved rapid mixing for $k \geq \Delta + 2$; at $k \leq \Delta + 1$ the chain might not be ergodic. In our setting, k may be much smaller and so we have to bound the number of fixed leaves. Our proof of Theorem 1.2 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. In Section 4, we prove the upper bound for Theorem 1.1. These techniques are then used in Section 5 to give the proof of Theorem 1.2. In Section 6 we explain how to modify our arguments further to yield Theorems 1.3 and 1.4. The lower bound for Theorem 1.1 is given in Section 7.

2 Preliminaries

2.1 Graph Colourings

Let $G = (V, E)$ be a finite graph, and let $A = \{0, 1, \ldots, k-1\}$ 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 $\Omega \subset A^V$ 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. [3, 16], show that our theorems also hold for the heat-bath version.) Informally, the behaviour of this process is as follows: each vertex has an associated (rate 1) poisson clock. When the clock for vertex v rings, a colour a is chosen uniformly from A. 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 all trees (see eg. [1]).

In many applications we are interested in the discrete analog of the Glauber dynamics. This Markov chain is given by transition matrix $P = I + \frac{1}{n}\mathcal{L}$. We can therefore think of $K[\sigma \to \eta]$ as being the probability of moving from colouring σ to colouring η in the discrete Markov chain, scaled by a factor of n. The mixing time for the discrete chain is precisely n times the mixing time for the corresponding continuous process (see eg. [1]). Therefore the bounds on mixing time given in this paper apply to the discrete setting as well as the continuous 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

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

Now suppose $\mathcal{L}$ is the generator for an ergodic markov process over Ω. Then there is a unique measure π on Ω that satisfies $\pi\mathcal{L} = \pi$. We say that π is the *stationary distribution* for $\mathcal{L}$. For example, if $\mathcal{L}$ is the generator for the Glauber dynamics, it is well-known that π is the uniform distribution over Ω (since the Glauber dynamics is reversible).

Suppose $\mathcal{L}$ is ergodic with stationary distribution π. 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

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

Now recall that $\mathcal{L}$ is a $|\Omega| \times |\Omega|$ matrix. 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. [16]):

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

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 to mean the set of vertices that forms the exterior boundary of T in G. That is, $\partial T = \{x \in V \backslash T : N(x) \cap T \neq \emptyset\}$. Note that for each $x \in \partial T$ there is a unique $y \in T$ adjacent to x.

We will analyze the Glauber dynamics over a subtree T of G with boundary configuration σ.

Claim 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 $\sigma \in \Omega$, with uniform stationary distribution.*

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, we can move from η to ω by changing one vertex of T at a time, so that at each step we have a proper colouring.

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 $\eta = \omega_r^{\eta(r)}$. We conclude $(\eta, \omega) \in \Gamma$ and we are done.

Now suppose the tree 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 $\beta \in \Omega_{T'}^\eta$ 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.

Similarly, we can change the colours of all children of r so that none are $\omega(r)$. There is therefore a colouring $\alpha \in \Omega_T^\sigma$ in which $\omega(r)$ does not appear in the neighbourhood of r, and there is a path from η to α in Γ. But this implies $(\alpha, \alpha_r^{\omega(r)}) \in \Gamma$. Let $\gamma = \alpha_r^{\omega(r)}$. Finally, 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', again by the induction hypothesis. We have thus found a path from η to ω in Γ and we are done. $\square$

3 Weighted Block Dynamics

In this section we present a generalization of the well-known block dynamics for local spin systems. We prove the result for the Glauber dynamics acting on a finite graph $G = (V, E)$. Our statement of the block dynamics actually applies 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.

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 $\mathcal{L}_{V_i}^\sigma$ be the generator for the Glauber dynamics restricted to V_i with boundary configuration σ. In other words, $\mathcal{L}_{V_i}^\sigma$ is a variant of the Glauber dynamics on V in which colours can change only for nodes in V_i. Recall that although we can think of $\mathcal{L}_{V_i}^\sigma$ as being equivalent to the behaviour of the Glauber dynamics acting on V_i with boundary configuration σ, the states of this process are formally considered to be colourings of the entire graph.

Suppose that $\mathcal{L}_{V_i}^\sigma$ is ergodic for each i and σ. Let $\pi_{V_i}^\sigma$ denote the stationary distribution of $\mathcal{L}_{V_i}^\sigma$. For each i, define g_i by

$$g_i = \inf_{\sigma \in \Omega} \operatorname{Gap}(\mathcal{L}_{V_i}^\sigma),$$

the minimum spectral gap for $\mathcal{L}_{V_i}^\sigma$ 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. $\tau(\mathcal{L}_{\mathcal{V}}) \le \tau(\mathcal{L}_D) \times (\max_{1\le i\le r} \tau_{V_i}) \times \sup_{x\in V} |\{i : x \in V_i\}|$.

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. Informally, we can still think of the weighted block dynamics as having poisson clocks that signal uniform recolouring of the blocks, but now the clock for block V_i will have rate g_i as opposed to 1. The main result for this section is the following variant of Proposition 3.1:

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

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 $n^{O(1+\Delta/k)}$ for general trees, while the modified block dynamics given here tightens the mixing time to $n^{O(1+\Delta/k \log \Delta)}$ (see Remark 5.9).

The proof of Proposition 3.2 is a simple modification to the proof of Proposition 3.1 [12]. We include it 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.

$$\begin{aligned}
\xi_D^*(f,f) &= \frac{1}{2}\sum_{\sigma,\eta\in\Omega} \pi[\sigma] K_D^*[\sigma\to\eta](f(\sigma)-f(\eta))^2 \\
&= \frac{1}{2}\sum_{\sigma\in\Omega} \pi[\sigma]\sum_{i=1}^{r} g_i \sum_{\eta\in\Omega_{V_i}^{\sigma}} \pi_{V_i}^{\sigma}[\eta](f(\sigma)-f(\eta))^2 \\
&= \frac{1}{2}\sum_{\sigma\in\Omega} \pi[\sigma]\sum_{i=1}^{r} g_i \mathrm{Var}_{V_i}^{\sigma}(f) \\
&\le \frac{1}{2}\sum_{\sigma\in\Omega} \pi[\sigma]\sum_{i=1}^{r} \xi_{V_i}^{\sigma}(f,f) \\
&= \frac{1}{2}\sum_{\sigma\in\Omega} \pi[\sigma]\sum_{i=1}^{r}\sum_{\eta\in\Omega_{V_i}^{\sigma}} \pi_{V_i}^{\sigma}[\eta]\sum_{x\in V_i}\sum_{a\in A} K[\eta\to\eta_x^a](f(\eta)-f(\eta_x^a))^2 \\
&\le \frac{1}{2}\sum_{\eta\in\Omega} \pi[\eta]\sum_{x\in V} N_x \sum_{a\in A} K[\eta\to\eta_x^a](f(\eta)-f(\eta_x^a))^2 \\
&\le N\xi_V(f,f)
\end{aligned}$$

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

□

4 An Upper Bound for Complete Trees

In this section we prove the upper bound from Theorem 1.1. That is, we show that the mixing time for the Glauber dynamics on a complete trees of degree Δ with $k \ge 3$ colours is $n^{O(1+\Delta/k\log\Delta)}$.

Let T be a complete rooted tree of degree Δ, possibly with a single external boundary node adjacent to its root. Let $n = |T|$, let v be the root of T, and let σ 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 k colours on T with boundary configuration σ. By Claim 2.1, this is ergodic. Note that, up to a possible relabelling of the boundary colour, the behaviour is completely determined by the height of T, say h. We therefore write $\tau(h) := \tau_T^{\sigma}$ to be the relaxation time of this process.

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

$$\tau(h) \le c\Delta k\left(\frac{k-1}{k-2}\right)^{\Delta}\tau(h-1). \tag{2}$$

Before proving Lemma 4.1, let us show how it implies the upper bound from 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 $\tau(h) \le \left(c\Delta k \left(\frac{k-1}{k-2} \right)^{\Delta} \right)^{h}$. 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

$$\begin{aligned}
\mathcal{M} &\le (n \log k)\tau(\lfloor \log_\Delta n \rfloor) \\
&\le (n \log k) \left(c\Delta k \left(1 + \frac{1}{k-2} \right)^{\Delta} \right)^{\log n / \log \Delta} \\
&< (n \log k)(c\Delta k)^{\log n / \log \Delta} \left(e^{\Delta/(k-2)} \right)^{\log n / \log \Delta} \\
&= (n \log k) n^{O(1)} n^{O(\Delta / k \log \Delta)} \\
&= n^{O(1 + \Delta / k \log \Delta)}
\end{aligned}$$

which is the upper bound from Theorem 1.1. It remains to prove Lemma 4.1, to which we devote the rest of this section.

Proof of Lemma 4.1. We begin by deriving (3) below, which is a special case of Lemma 2.8 of [1]. Our proof of (3) uses the same techniques as in [1]. We include the argument for completeness, and because we extend it in Section 5.1.

Suppose $h > 0$ and let $u_1, \dots, u_\Delta$ be the children of v in T. Let V_i be the subtree of T rooted at u_i, for each $1 \le i \le \Delta$; note that each V_i is a complete tree of height $h - 1$. Let $D = \{\{v\}, V_1, \dots, V_\Delta\}$. 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

$$\tau(h) \le \tau_D \max_{\sigma \in \Omega} \{\tau^{\sigma}_{\{v\}}, \tau^{\sigma}_{V_1}, \dots, \tau^{\sigma}_{V_\Delta}\} = \tau_D \tau(h-1) \tag{3}$$

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

$$\tau_D \le c\Delta k \left(\frac{k-1}{k-2} \right)^{\Delta}. \tag{4}$$

The block dynamics $\mathcal{L}_D$ resembles the action of the Glauber dynamics on a tree of height 1 rooted at v. The blocks $V_1, \dots, V_\Delta$ act as the leaves of this tree, and the interactions between colourings of the blocks are captured entirely by the colours of their roots. With this in mind, define G to be the induced subgraph of T containing v and $u_1, \dots, u_\Delta$, and with the same (if any) boundary condition on v. Then the following claim is a special case of Lemma 2 in [15].

Claim 4.2. $\tau_D \le \max\{1, \tau_G\}$

Thus, to prove (4), it is sufficient to bound τ_G. First let us give some intuition into the bound in (4) as it applies to τ_G. 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 $k - 1$ 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 $(\frac{k-2}{k-1})^{\Delta}$. Thus we expect it to take roughly $\frac{1}{k-1}(\frac{k-1}{k-2})^{\Delta}$ time before the colour of the root can change. The bound in (4) states that the mixing time is not much more than this.

We now proceed to bound τ_G by the method of canonical paths. We note that the construction we use is not the simplest that obtains the desired bound, but it will introduce techniques that will be useful when proving Theorem 1.2.

Choose two colourings $\alpha, \eta \in \Omega_G$. Our goal is to define a sequence of steps of the Glauber dynamics that begins in state α and ends in state η. If $\alpha(v) = \eta(v)$ this sequence is simple: the colours of nodes $u_1, \dots, u_\Delta$ are changed from α to η one at a time. If $\alpha(v) \ne \eta(v)$, our strategy is to first change the colours of $u_1, \dots, u_\Delta$ 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, but this method generates too much congestion (defined below) for our desired bound. We therefore introduce a slightly more complex path that uses three passes. To describe this path formally, we will need to define some colours.

If $\alpha(v) \neq \eta(v)$ then for each $1 \leq i \leq \Delta$ we will define two colours, a_i and c_i, that depend on α and η. The first colour, a_i, is easy to define:

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

That is, the colours $(a_1, \dots, a_\Delta)$ are simply the colours of the children of v in α, except that any occurrences of $\eta(v)$ are replaced with $\alpha(v)$.

The definition of colour c_i is more involved. These will be the colours to which we set the leaf nodes, in order to allow v to change from $\alpha(v)$ to $\eta(v)$. Rather than only modify leaves that have colour $\eta(v)$, we will potentially change the colours of all leaves to better distribute congestion. In particular, we will apply a function f that will map the colours $(\alpha(u_1), \dots, \alpha(u_\Delta))$ to a vector of colours $(c_1, \dots, c_\Delta)$ such that for all i, $c_i \notin \{\alpha(v), \eta(v)\}$. We want this function f to satisfy the following balance property: for all $1 \leq i \leq t$,

$$\#\{(x_1, \dots, x_t) : (x_j = \alpha(u_j)\ \forall j > i) \wedge (f(x_1, \dots, x_t)_j = c_j\ \forall j \leq i)\} \leq \left\lceil \left(\frac{k-1}{k-2}\right)^i \right\rceil. \tag{5}$$

That is, for any $1 \leq i \leq \Delta$, if we are given the first i entries of the image of f and the last $\Delta - i$ entries of the preimage of f, there are at most $\left\lceil \left(\frac{k-1}{k-2}\right)^i \right\rceil$ possibilities for the preimage of f. We defer the construction of f to Lemma 4.3 in Section 4.1.

This completes the definition of colours a_i and c_i. We note some properties of these colours. First, $a_i \neq \eta(v)$ and $c_i \notin \{\alpha(v), \eta(v)\}$. Second, suppose that α and η are fixed, but unknown. Suppose that we are given $\alpha(v)$, $\eta(v)$, and a_i. Then $\alpha(u_i)$ is determined: it is exactly a_i, unless $a_i = \alpha(v)$ in which case $\alpha(u_i) = \eta(v)$. Furthermore, if we are given $c_1, \dots, c_i$ and $\alpha(u_{i+1}), \dots, \alpha(u_\Delta)$, then there are at most $\left\lceil (\frac{k-1}{k-2})^i \right\rceil$ possibilities for the vector of colours $(\alpha(u_1), \dots, \alpha(u_\Delta))$.

We are now ready to formally begin our canonical paths argument. Let Γ be the transition graph over Ω_G with $(\omega, \beta) \in \Gamma$ if and only if $K[\omega \to \beta] > 0$. That is, if and only if colourings ω and β differ on exactly one vertex of G. For each $\alpha, \eta \in \Omega_G$ we will define a path in Γ, denoted $\gamma(\alpha, \eta)$.

If $\alpha(v) = \eta(v)$, our path changes the colour of each u_i from $\alpha(u_i)$ to $\eta(u_i)$, one at a time. If $\alpha(v) \neq \eta(v)$, the most natural path would change each u_i to some colour other than $\eta(v)$, then set v to $\eta(v)$, and finally set each u_i to $\eta(u_i)$. However, in order to tighten our bounds, we use a more indirect path:

1. For each u_i in increasing order: recolour from $\alpha(u_i)$ to c_i.
2. Recolour v from $\alpha(v)$ to $\eta(v)$.
3. For each u_i in decreasing order: recolour from c_i to a_i.
4. For each u_i in increasing order: recolour from a_i to $\eta(u_i)$.

Let L be the maximum length of any such path; then $L \leq 3\Delta + 1$.

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

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

The congestion for our set of paths is

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

Then by the method of canonical paths, introduced in [9] (see also e.g. [1, 12, 17] for other applications) we have that

$$\tau_D \leq L\rho \leq (3\Delta + 1)\rho. \tag{6}$$

Our goal now is to bound ρ. First we recall from the definition of the Glauber dynamics that

$$K[\omega \to \beta] = \frac{1}{k}. \tag{7}$$

We also note that

$$|\Omega_G| \geq (k-1)^{\Delta+1}, \tag{8}$$

which follows by choosing a colouring for G 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 G has an external boundary).

Using (8) and the fact that π is the uniform distribution, we have

$$\begin{aligned} \rho(\omega, \beta) &= \left(|\{\alpha, \eta : \gamma(\alpha, \eta) \ni (\omega, \beta)\}| \times \frac{1}{|\Omega_G| K[\omega \to \beta]} \right) \\ &\leq \left(|\{\alpha, \eta : \gamma(\alpha, \eta) \ni (\omega, \beta)\}| \times \frac{k}{(k-1)^{\Delta+1}} \right). \end{aligned} \tag{9}$$

To bound ρ, it remains to compute the number of paths $\gamma(\alpha, \eta)$ that include (ω, β) for each $(\omega, \beta) \in \Gamma$. We consider cases depending on the nature of the transition (ω, β).

Case 1: ω and β differ on the colour of v. Note that in this case (ω, β) would be the change from Step 2 of the canonical path description. Then α and η must satisfy $\alpha(v) = \omega(v), \eta(v) = \beta(v)$, and $c_i = \omega(u_i)$ for all $1 \leq i \leq \Delta$.

Consider the possibilities for colouring η. Colouring β determines $\eta(v)$, and there are $(k-1)^\Delta$ choices for η given $\eta(v)$ (consider choosing the colours for $u_1, \ldots, u_\Delta$, which cannot be $\eta(v)$). Now consider α: the colour $\alpha(v)$ is determined by ω, as are $(c_1, \ldots, c_\Delta)$. Thus by (5) there are at most $\lceil (\frac{k-1}{k-2})^\Delta \rceil$ possibilities for $(\alpha(u_1), \ldots, \alpha(u_\Delta))$, which determines α.

Putting this together, we conclude that the total number of colourings α and η that satisfy $(\omega, \beta) \in \gamma(\alpha, \eta)$ is at most

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

Substituting this into (9), we have

$$\begin{aligned} \rho(\omega, \beta) &\leq (k-1)^\Delta \left\lceil \left(\frac{k-1}{k-2} \right)^\Delta \right\rceil \frac{k}{(k-1)^{\Delta+1}} \\ &\leq 3 \left(\frac{k-1}{k-2} \right)^\Delta \end{aligned}$$

Case 2: ω and β differ on the colour of u_i for some i.

Now as in Case 1, we must count the number of α, η such that $(\omega, \beta) \in \gamma(\alpha, \eta)$. We do this by considering four different possibilities for the position of (ω, β) in a path. We will count the number of α, η in each subcase separately.

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

$$(k-1)^{\Delta-1}.$$

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

First consider η. There are at most $k-1$ possibilities for $\eta(v)$, since $\eta(v) \neq \alpha(v)$. Given $\eta(v)$, there are $k-1$ possibilities for $\eta(u_j)$ for each $1 \leq j \leq \Delta$. Thus the number of possibilities for η is at most $(k-1)^{\Delta+1}$.

Next consider α. Note that ω determines $\alpha(v)$ and also $\alpha(u_j)$ for all $j \geq i$. Furthermore, β determines c_j for all $1 \leq j < i$. So by (5) there are at most $\lceil (\frac{k-1}{k-2})^{i-1} \rceil$ possibilities for $(\alpha(u_1), \ldots, \alpha(u_\Delta))$, which determines α. So the number of possibilities for α 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 α and η is

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

Case 2.3: $\alpha(v) \neq \eta(v)$ and (ω, β) is the second change to u_i in $\gamma(\alpha, \eta)$. This is the change in Step 3 of the canonical paths description. We know that $\eta(v) = \omega(v)$, $c_j = \omega(u_j)$ for all $j \leq i$, and $a_j = \beta(u_j)$ for all $j \geq i$.

First consider η. Note that β determines $\eta(v)$. For each $1 \leq j \leq \Delta$, there are $k-1$ possibilities for $\eta(u_j)$ (since $\eta(u_j) \neq \eta(v)$). Thus the number of possibilities for η is at most $(k-1)^{\Delta}$.

Now consider α. There are at most $k-1$ possibilities for $\alpha(v)$. Colours $a_i, \ldots, a_\Delta$ are determined by ω and colours $c_1, \ldots, c_{i-1}$ are determined by β. But then $\alpha(u_i), \ldots, \alpha(u_\Delta)$ can be recovered and by (5) there are at most $\lceil (\frac{k-1}{k-2})^{i-1} \rceil$ possibilities for $(\alpha(u_1), \ldots, \alpha(u_\Delta))$. This determines α, so the number of possibilities for α is at most $(k-1) \left\lceil \left(\frac{k-1}{k-2}\right)^{i-1} \right\rceil$. We conclude that for this subcase the total number of possibilities is at most

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

Case 2.4: $\alpha(v) \neq \eta(v)$ and (ω, β) is the third change to V_i in $\gamma(\alpha, \eta)$. This is the change in Step 4 of the canonical paths description. In this case we know $\eta(v) = \omega(v)$, $a_j = \omega(u_j)$ for all $j \geq i$, and $\eta(u_j) = \beta(u_j)$ for all $j \leq i$.

In this case there are at most $(k-1)$ choices for $\alpha(v)$. The colours $a_i, \ldots, a_\Delta$ plus $\eta(v)$ are determined by ω, and from these (plus $\alpha(v)$) the colours $\alpha(u_i), \ldots, \alpha(u_\Delta)$ are determined. On the other hand, $\eta(u_1), \ldots, \eta(u_i)$ 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 $\alpha(v)$, we conclude that the number of possible options for α and η is at most

$$(k-1)^{\Delta}.$$

This concludes our subcase analysis. Note that in each case, the number of possibilities for α and η was at most

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

Summing up over all cases, we get that the total number of possibilities for α and η, given that (ω, β) is a change in the colouring of V_i for some i, is at most

$$4 \left\lceil \left(\frac{k-1}{k-2}\right)^{i-1} \right\rceil (k-1)^{\Delta+1} \leq 8 \left(\frac{k-1}{k-2}\right)^{\Delta-1} (k-1)^{\Delta+1}.$$

Substituting this into (9), we have

$$\begin{aligned} \rho(\omega, \beta) &\leq 8 \left(\frac{k-1}{k-2}\right)^{\Delta-1} (k-1)^{\Delta+1} \left(\frac{k}{(k-1)^{\Delta+1}}\right) \\ &\leq 8k \left(\frac{k-1}{k-2}\right)^{\Delta}. \end{aligned}$$

This completes our analysis of Case 2. Considering all cases, we conclude

$$\rho \leq 8k \left(\frac{k-1}{k-2}\right)^{\Delta} .$$

Applying (6), we conclude that

$$\tau_D \leq L\rho \leq (3\Delta+1)(8k)\left(\frac{k-1}{k-2}\right)^{\Delta} \leq 32\Delta k \left(\frac{k-1}{k-2}\right)^{\Delta}$$

which is (4) with $c = 32$. This completes the proof of Lemma 4.1.

$\square$

4.1 A balanced mapping function

We now present the mapping function f used in the proof of Lemma 4.1, which satisfies (5). We will actually prove the existence of f in the following, equivalent, arena.

Lemma 4.3. *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 $1 \leq i \leq t$,*

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

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 $1 \leq i \leq t$. 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. $\square$

5 An Upper Bound for General Trees with $k \geq 4$

In this section we will prove Theorem 1.2. Up to this point we have only considered complete trees; let us briefly discuss the complications that arise when we consider general trees. First, the height of a general tree will be greater than $\log_\Delta n$ in general, so (2) would be insufficient to bound the relaxation time. We get around this issue by applying the block dynamics to a different choice of subtrees as blocks. This yields two problems: First, we deal with blocks of vastly differing sizes. Second, an arbitrary subtree T may have $|\partial T| > 1$, which introduces problems with ergodicity and complicates the behaviour of τ_D. We get around the first issue by using the weighted version of the block dynamics. For the second, we must choose blocks carefully to limit boundary size, and handle extra complications that arise.

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). For any $\sigma \in \Omega$, consider the Glauber dynamics on T with $k \geq 4$ colours and boundary configuration σ, say with generator $\mathcal{L}$. Since $|\partial T| \leq 2$, Claim 2.1 implies that the Glauber dynamics on T is ergodic when $k \geq 4$; this is the critical reason that we require $k \geq 4$. Again, we can assume $k \leq 2\Delta$. Let τ_T^σ denote the relaxation time for this instance of the Glauber dynamics. We wish to consider the maximum relaxation time over all boundary configurations and trees of a certain size. To this end, we define

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

We will prove Theorem 1.2 by showing the slightly stronger result that $\tau_2(n) = n^{O(1+\Delta/k\log\Delta)}$. What we will show is that, for some fixed constant c and some $2 \le i \le \Delta$,

$$\tau_2(n) \le ci^2 \left(\frac{k-1}{k-2}\right)^{i+1} \tau_2\left(\lfloor n/i \rfloor\right). \tag{10}$$

First let us show how (10) implies Theorem 1.2. By (1), the mixing time of the Glauber dynamics satisfies

$$\mathcal{M}(\mathcal{L}) \le (n\log k)\tau_G \le (n\log k)\tau_2(n).$$

It suffices to show that $\tau_2(n) \le n^{d(1+\Delta/k\log\Delta)}$ for some constant d, which we prove by induction on n. The result holds for the base case of small n by choosing d sufficiently large. In general, (10) implies that

$$\begin{aligned}
\tau_2(n) &\le ci^2\left(1+\frac{1}{k-2}\right)^{i+1}\tau_2\left(\lfloor n/i\rfloor\right) \\
&\le (i^{\log c/\log i})i^2(e^{1/(k-2)})^{i+1}(n/i)^{d(1+\Delta/k\log\Delta)} \\
&\le (i^{2+\log c/\log i})(i^{(i+1)\log e/(k-2)\log i})(i^{-d(1+\Delta/k\log\Delta)})\left[n^{d(1+\Delta/k\log\Delta)}\right] \\
&\le (i^{2+\log c/\log 2+6\Delta\log e/k\log\Delta})(i^{-d(1+\Delta/k\log\Delta)})\left[n^{d(1+\Delta/k\log\Delta)}\right]
\end{aligned}$$

where we used the fact that $\log i \ge \log 2$ and $i/\log i < \Delta/\log\Delta$ since $2 \le i \le \Delta$. Thus, if d is large enough that $d > \max\{2+\log c/\log 2, 6\log e\}$, we conclude that

$$\tau_2(n) \le n^{d(1+\Delta/k\log\Delta)}$$

as required.

We now turn to proving (10). We begin with a definition that will be helpful when we define blocks for the block dynamics.

Definition 5.1. Choose $v \in T$. Then the *separation of G by v* is the collection of disjoint subsets $D_v = \{\{v\}, V_1, \ldots, V_r\}$ where each V_r is a maximal connected component of T. We say that each V_i is a *subtree separated by v*.

The following generalization of (2) is the workhorse for our proof.

Lemma 5.2. *Suppose $k \ge 4$ and let T be a subtree of a tree G with $|\partial T| \le 2$ and let $\sigma \in \Omega$ be a boundary condition for T. Choose $v \in T$ and consider $D_v = \{\{v\}, V_1, \ldots, V_t\}$, where $1 \le t \le \Delta$. Suppose $|\partial V_i| \le 2$ for each V_i. Then*

$$\tau_T^\sigma \le 4000 \max_{1\le i\le t} i^2\left(\frac{k-1}{k-2}\right)^i \tau_{V_i}.$$

Furthermore, this bound holds for $k = 3$ if we additionally require that $|\partial V_i| \le 1$ for each V_i.

We will prove Lemma 5.2 in Section 5.1. First, let us show how it implies (10). We will first consider trees with boundaries of size one, then trees with boundaries of size two.

Claim 5.3. *For some $2 \le i \le \Delta$, we have $\tau_1(n) \le 4000i^2\left(\frac{k-1}{k-2}\right)^i \tau_2\left(\lfloor n/i\rfloor\right)$.*

Proof. Let T be a subtree of G with $|\partial T| \le 1$. Choose a root r for T. If $|\partial T| = 1$ then choose r to be adjacent to the boundary of T; otherwise choose r arbitrarily.

It is well-known that we can find a vertex $x \in T$ such that if $D_x = \{\{x\}, V_1, \ldots, V_t\}$, we will have $|V_i| \le \lfloor n/2\rfloor$ for all $1 \le i \le t$ (see eg. [11]). We will choose our indices so that $|V_1| \ge |V_2| \ge \ldots \ge |V_t|$. Note that since $|\partial T| \le 1$, we will have $|\partial V_i| \le 2$ for all i. By Lemma 5.2, $\tau_T \le 4000i^2\left(\frac{k-1}{k-2}\right)^i \tau_{V_i}$ for some $1 \le i \le t$. We now consider two cases for the value of i.

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

$$\tau_T \leq 4000 i^2 \left(\frac{k-1}{k-2}\right)^i \tau_2\left(\lfloor n/i \rfloor\right)$$

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

$$\tau_T \leq 4000 \left(\frac{k-1}{k-2}\right) \tau_{V_1} < 4000(2)^2 \left(\frac{k-1}{k-2}\right)^2 \tau_2\left(\lfloor n/2 \rceil \rfloor\right)$$

as required. ☐

Claim 5.4. *For some* $2 \leq i \leq \Delta$, $\tau_2(n) \leq 4000^2 i^2 \left(\frac{k-1}{k-2}\right)^{i+1} \tau_2(\lfloor n/i \rfloor)$.

Proof. Let T be a subtree with $|T| = n$ and $|\partial T| = 2$, say $\partial T = \{z_1, z_2\}$. Choose root r for T and vertex x as in Claim 5.3, with r adjacent to z_1 and x separating T into subtrees of size at most $\lfloor n/2 \rfloor$. We will call the unique path in G from z_1 to z_2 the *boundary path* for T. Suppose x is on the boundary path for T. Then we consider $D_x = \{\{x\}, V_1, \ldots, V_t\}$, where blocks are indexed such that $|V_1| \geq \ldots \geq |V_t|$, and note that $|\partial V_i| \leq 2$ for all i. We can then apply Lemma 5.2 as in Claim 5.3 and obtain the desired result.

Now suppose that x is not on the boundary path for 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, all the subtrees separated by y have boundary of size at most 2 (since y is on the boundary path for T). Lemma 5.2 applied to D_y yields:

$$\tau_T \leq 4000 i^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}$$

for some i. If $i > 1$ then we obtain the desired result since $|V_i| \leq \lfloor n/i \rfloor$. If $i = 1$, then note that $|V_1| < n$ and V_1 has boundary of size 1 (namely, $\partial V_1 = \{y\}$, by our choice of y). Therefore Claim 5.3 yields:

$$\begin{aligned} \tau_T &\leq 4000 \left(\frac{k-1}{k-2}\right) \tau_1(|V_1|) \leq 4000 \left(\frac{k-1}{k-2}\right) \tau_1(n) \\ &\leq 4000^2 i^2 \left(\frac{k-1}{k-2}\right)^{i+1} \tau_2\left(\lfloor n/i \rfloor\right) \quad \text{for some } 2 \leq i \leq \Delta. \end{aligned}$$

☐

We have now derived (10) (with $c = 4000^2$), completing the proof of Theorem 1.2.

5.1 Proof of Lemma 5.2

We now proceed with the proof of Lemma 5.2. Our approach is similar to the proof of Lemma 4.1 in Section 4. Recall that in that proof we gave a bound for the Glauber dynamics in terms of the block dynamics, then used a canonical paths argument to bound the behaviour of the block dynamics. We take the same approach, but with two major changes. First, we use weighted block dynamics and a weighted version of canonical paths to deal with differing block sizes. Second, the block dynamics is no longer equivalent to the Glauber dynamics on a star (i.e. the analogue of Claim 4.2 does not hold in this setting). Therefore in our analysis we will need to recolour entire subtrees at each step. Third, we must change the construction of our canonical paths to account for the effect of external boundary nodes on subtrees.

Remark 5.5. We note that while the analogue of Claim 4.2 does not hold in our setting, the techniques of [15] can be applied to obtain a result similar to Claim 4.2 for general trees, and to obtain a polynomial bound on the mixing time given that k is sufficiently large in terms of Δ. However, to the best of our understanding, this approach will not lead to the asymptotically tight bounds in Theorem 1.2.

Note that in the statement of Lemma 5.2 we require that $|\partial T| \leq 2$. This requirement is not crucial; it is made to simplify the argument (and is sufficient to prove Theorem 1.2). Indeed, in Section 6.2 we will discuss what happens when we relax this condition.

Recall the conditions of Lemma 5.2. Suppose $k \geq 4$ and let T be a tree with $|\partial T| \leq 2$ and let $\sigma \in \Omega$ be a boundary condition for T. Choose $v \in T$ and consider $D_v = \{\{v\}, V_1, \ldots, V_t\}$, where $1 \leq t \leq \Delta$. Suppose $|\partial V_i| \leq 2$ for each V_i.

Let $D := D_v$ for notational convenience. We will consider T to be rooted at v. Let $u_1, \ldots, u_t$ be the children of v in T, so that u_i is the root of V_i. Recall that given any colouring $\alpha \in \Omega_T^\sigma$, we will write $\alpha(V_i)$ to refer to the colours assigned to the nodes of V_i by α. Note that for all V_i we have $1 \leq |\partial V_i| \leq 2$, $v \in \partial V_i$, and no other member of ∂V_i is in T. Thus if $|\partial V_i| = 2$ then $|\partial V_i \cap \partial T| = 1$ and so $|\partial V_i| = 2$ for at most two blocks V_i.

As we did for complete trees in Lemma 4.1, we will define a method for moving from one colouring to another using steps of the block dynamics. However, unlike before, we will explicitly define steps that recolour entire blocks of vertices. Our overall strategy for each step will be to choose a desired colour a for the root u_i of a block V_i, then recolour the entire block V_i so that u_i has colour a. Note that the way in which we choose the colours for the block is important as it will affect the bound we obtain. The following definition describes the way in which we will colour blocks.

Definition 5.6. Choose some V_i with root u_i. Choose $\sigma \in \Omega$ and $a \in A$. Then the **cyclic recolouring** of V_i that sets u_i to a , denoted $\sigma_{V_i}^a$, is the (possibly not proper) colouring defined by

$$\sigma_{V_i}^a(y) = \begin{cases} \sigma(y) + a - \sigma(u_i) \pmod{k} & \text{if } y \in V_i \\ \sigma(y) & \text{otherwise.} \end{cases}$$

That is, we perform a cyclic permutation of the colours on V_i so that the colour of u_i is set to a.

We note the following useful property of cyclic recolourings. Choose $\alpha \in \Omega$, subtree V_i, and a colour a. Then if $\omega = \alpha_{V_i}^a$, it is the case that $\alpha = \omega_{V_i}^{\alpha(u_i)}$.

Choose two colourings $\alpha, \eta \in \Omega_T^\sigma$. If $\alpha(v) \neq \eta(v)$ then for each $1 \leq i \leq t$ we define colours a_i and c_i precisely as in the proof of Lemma 4.1. We also define colour b_i to be analagous to a_i, but with the roles of α and η reversed. That is,

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

The analogue of changing the colour of u_i in Lemma 4.1, say from $\alpha(u_i)$ to c_i, would be to cyclically recolour block V_i, say from $\alpha(V_i)$ to $\alpha_{V_i}^{c_i}(V_i)$. We might be tempted to try to use the canonical paths from Lemma 4.1 with this change. However, if $|\partial V_i| = 2$, the colouring $\alpha_{V_i}^{c_i}(V_i)$ might not be proper: it could conflict with the colour of the fixed external boundary node adjacent to V_i. We get around this problem by defining three more colours: a_i', b_i', and c_i'. The colours a_i', b_i', and c_i' can be thought of as versions of a_i, b_i, and c_i that are "repaired" so that the cyclic recolourings we will use in our canonical paths construction will lead to proper colourings.

We begin with the definition of a_i'. If $|\partial V_i| = 1$ then set $a_i' = a_i$. Otherwise $|\partial V_i| = 2$, so $\partial V_i = \{v, z\}$ where $z \in \partial T$. Let y be the vertex in V_i adjacent to z. Consider the colour assigned to vertex y in $\alpha_{V_i}^{a_i}(V_i)$. If this colour is not $\sigma(z)$, then we again set $a_i' = a_i$. Otherwise, we note that $\alpha_{V_i}^{a_i}(V_i)$ conflicts with the fixed colour of the external boundary node z. In this case, we choose a_i' to be *any* colour that is not in $\{\alpha(v), a_i\}$. Such a colour must exist since $k \geq 4$. Then by the definition of a cyclic recolouring, $\alpha_{V_i}^{a_i'}(y) \neq \alpha_{V_i}^{a_i}(y) = \sigma(z)$, so $\alpha_{V_i}^{a_i'}(V_i)$ does not conflict with the boundary condition for T.

We define b_i' and c_i' in a similar way. We set $b_i' = b_i$ if $|\partial V_i| = 1$ or $\eta_{V_i}^{b_i}(V_i)$ is a proper colouring with respect to the boundary condition σ. Otherwise we choose b_i' to be any colour not in $\{\eta(v), b_i\}$.

Likewise, $c'_i = c_i$ if $|\partial V_i| = 1$ or $\alpha^{c_i}_{V_i}(V_i)$ is a proper colouring with respect to σ. Otherwise, choose c'_i to be any colour not in $\{\eta(v), \alpha(v), c_i\}$ (which exists since $k \geq 4$).

Our next claim shows that the preimage of this repairing process is not too large.

Claim 5.7. *Suppose that we are given $\alpha(v)$ and $\eta(v)$, and the remainder of α and η is fixed but unknown. Consider a block V_i. Then:*

1. *Given colouring $\alpha^{a'_i}_{V_i}(V_i)$, there are at most $|\partial V_i|$ possibilities for a_i.*
2. *Given colouring $\eta^{b'_i}_{V_i}(V_i)$, there are at most $|\partial V_i|$ possibilities for b_i.*
3. *Given colouring $\alpha^{c'_i}_{V_i}(V_i)$, there are at most $|\partial V_i|$ possibilities for c_i.*

Proof. Consider the case where we are given $\alpha^{a'_i}_{V_i}(V_i)$; the other two cases are similar. Write $\omega = \alpha^{a'_i}_{V_i}(V_i)$. Note that a'_i is determined by ω; indeed $a'_i = \omega(u_i)$. If $|\partial V_i| = 1$ then $a_i = a'_i$, so a_i is determined as required.

Suppose $|\partial V_i| = 2$; then $\partial V_i = \{v, z\}$ for some $z \in \partial T$. Let y be the vertex in V_i adjacent to z. From the definition of a'_i we know that either $a_i = a'_i$ or it must be that $\alpha^{a_i}_{V_i}(V_i)$ is not a proper colouring with respect to boundary condition σ. But since ω is a cyclic recolouring of α, it must be that a_i is the unique colour c such that $\omega^{c}_{V_i}(V_i)$ assigns colour $\sigma(z)$ to y. There are therefore two possibilities for a_i, namely a'_i or c, as required. □

We are now ready to proceed with the proof of Lemma 5.2. Let $\mathcal{L}^*_D$ be the generator for the weighted block dynamics corresponding to D and boundary configuration σ, with entries denoted K^*_D. Let τ^σ_D denote the relaxation time for this instance of the weighted block dynamics. Since no vertex lies in more than one block, by Proposition 3.2,

$$\tau^\sigma_T \leq \tau^\sigma_D. \tag{11}$$

Therefore to prove Lemma 5.2 it is sufficient to bound τ^σ_D, which we will do via a weighted canonical paths argument. Let Γ be the transition graph over Ω^σ_T with $(\omega, \beta) \in \Gamma$ if and only if $K^*_D[\omega \to \beta] > 0$. That is, if and only if the colours of ω and β differ on only one block. For each $\alpha, \eta \in \Omega^\sigma_T$ we will define a path $\gamma(\alpha, \eta)$ in Γ from α to η. If $\alpha(v) = \eta(v)$, our path simply changes the colouring of each block from $\alpha(V_i)$ to $\eta(V_i)$ in order.

If $\alpha(x) \neq \eta(x)$ then our path proceeds by the following steps.

1. For each block V_i in increasing order: change the colouring from $\alpha(V_i)$ to $\eta^{b'_i}_{V_i}(V_i)$, then to $\alpha^{c'_i}_{V_i}(V_i)$.
2. Recolour v from $\alpha(v)$ to $\eta(v)$.
3. For each block V_i in decreasing order: change the colouring from $\alpha^{c'_i}_{V_i}(V_i)$ to $\eta(V_i)$, then to $\alpha^{a'_i}_{V_i}(V_i)$.
4. For each block V_i in increasing order: change the colouring from $\alpha^{a'_i}_{V_i}(V_i)$ to $\eta(V_i)$.

The reader is encouraged to verify that all steps of this path are valid transitions according to $\mathcal{L}^*_D$. Note that this path is similar to the path used in Lemma 4.1, except that we are recolouring subtrees rather than single vertices, and instances of colours a_i and c_i are replaced by a'_i and c'_i. Also, the transitions in steps 1 and 3 are now done in two phases. This is so that each recolouring of a block moves from a cyclic recolouring of α to a cyclic recolouring of η or vice-versa, to limit congestion.

We will now define the weighted congestion of our choice of paths. For each $(\omega, \beta) \in \Gamma$, we will define a weight $w(\omega, \beta) > 0$. The precise weights used will depend on the choice of ω and β, and are described below. We define the weight of a path by

$$w(\gamma(\alpha, \eta)) = \sum_{(\omega,\beta)\in\gamma(\alpha,\eta)} w(\omega, \beta).$$

For each edge $(\omega, \beta) \in \Gamma$, define the weighted congestion of that edge, $\rho_w(\omega, \beta)$, as

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

The weighted congestion for our set of paths is $\rho_w := \sup_{\omega,\beta} \rho_w(\omega,\beta)$. Then the weighted canonical paths bound is

$$\tau_D^\sigma \le \rho_w. \tag{12}$$

We note that (12) and its proof are implicit in [4] (see their Remark on page 38). Note that (6), the standard statement of the canonical path bound, comes from taking $w(\omega,\beta) = 1$ for all $(\omega,\beta) \in \Gamma$. Our choice of a different weight function will allow us to tighten the bound we obtain on τ_D^σ. In particular, our approach obtains a bound of $n^{O(1+\Delta/k\log\Delta)}$, whereas the standard canonical paths bound (6) would only give a bound of $n^{O(\log\Delta+\Delta/k\log\Delta)}$ (see Remark 5.8).

We will use the following weight function. Set $w(\omega,\beta) = 1$ if ω and β differ on the colour of v, and set $w(\omega,\beta) = i^{-2}$ if ω and β differ on the colour of block V_i. Then note that for all $\gamma(\alpha,\eta)$,

$$w(\gamma(\alpha,\eta)) \le 1 + 5\sum_{i=1}^{t} i^{-2} < 1 + 5\left(\frac{\pi^2}{6}\right) < 10.$$

Let us also consider the size of the state space $|\Omega_T^\sigma|$. We can think of T as being rooted at some vertex r, which we choose to be adjacent to a boundary vertex if $|\partial T| > 0$. We then choose the colours for the vertices of $|\Omega_T^\sigma|$ in a top-down manner. There will be at least $k-1$ options for each vertex, except perhaps for the one vertex other than r adjacent to a boundary node if $|\partial T| = 2$, for which there would be at least $k-2$ options. We conclude that

$$|\Omega_T^\sigma| \ge (k-1)^{n-1}(k-2). \tag{13}$$

Using the fact that π is the uniform distribution, we now have

$$\begin{aligned}
\rho_w(\omega,\beta) &< \left(\frac{1}{w(\omega,\beta)} \sum_{\gamma(\alpha,\eta)\ni(\omega,\beta)} \frac{10\pi[\alpha]\pi[\eta]}{\pi[\omega]K_D^*[\omega\to\beta]}\right) \\
&\le 10\left(\frac{1}{w(\omega,\beta)} \times |\{\gamma(\alpha,\eta)\ni(\omega,\beta)\}| \times \frac{1}{|\Omega_T^\sigma| K_D^*[\omega\to\beta]}\right) \\
&\le 10\left(\frac{1}{w(\omega,\beta)} \times |\{\gamma(\alpha,\eta)\ni(\omega,\beta)\}| \times \frac{1}{(k-1)^{n-1}(k-2)K_D^*[\omega\to\beta]}\right).
\end{aligned} \tag{14}$$

It remains to evaluate (14) for different colourings ω and β. We consider cases depending on the nature of the transition (ω,β).

Case 1: ω and β differ on the colour of v. Note that $w(\omega,\beta) = 1$. Also, from the definition of the weighted block dynamics $\mathcal{L}_D^*$, we have

$$K_D^*[\omega\to\beta] = \inf_{\sigma\in\Omega} \operatorname{gap}(\mathcal{L}_{\{v\}}^\sigma)\pi_{\{v\}}^\omega[\beta].$$

But note that $\operatorname{gap}(\mathcal{L}_{\{v\}}^\sigma) = 1$ for all boundary conditions, since this is the set of proper colourings on a single vertex which mixes in a single step. Also, the stationary distribution $\pi_{\{v\}}^\omega$ is the uniform distribution over the set $\Omega_{\{v\}}^\omega$, which is the set of at most $k-1$ colours not appearing on $\partial\{v\}$ in ω. We conclude

$$K_D^*[\omega\to\beta] \le \frac{1}{k-1}. \tag{15}$$

We now consider the number of paths $\gamma(\alpha,\eta)$ that use (ω,β). This transition is used once for each α,η such that $\alpha(v) = \omega(v)$ and $\eta(v) = \beta(v)$, and in which $\alpha_{V_i}^{c_i'}(V_i) = \omega(V_i)$ for each V_i.

Consider the possibilities for colouring η. Colouring β determines $\eta(v)$, and there are $(k-1)^{n-1}$ choices for η given $\eta(v)$ (consider choosing the colours in η for each vertex, top-down, as in (13)). Now consider α: the colour $\alpha(v)$ is determined by ω, as are $(c_1',\dots,c_t')$. By Claim 5.7, there are at most $|\partial V_j|$ possibilities for each c_j, and hence there are at most $\prod_{j=1}^t |\partial V_j| \le 4$ possibilities for the colours $(c_1,\dots,c_t)$. Then by (5) there are at most $\lceil(\frac{k-1}{k-2})^t\rceil$ possibilities for $(\alpha(u_1),\dots,\alpha(u_t))$. But now these colours and ω determine α, since for each block V_i we have $\alpha(V_i) = \omega_{V_i}^{\alpha(u_i)}(V_i)$.

Putting this together, we conclude that the total number of colourings α and η that satisfy $(\omega, \beta) \in \gamma(\alpha, \eta)$ is at most

$$4(k-1)^{n-1} \left\lceil \left(\frac{k-1}{k-2}\right)^t \right\rceil.$$

Substituting this and (15) into (14), we have

$$\begin{aligned}
\rho_w(\omega, \beta) &\le 4(1)(k-1)^{n-1} \left\lceil \left(\frac{k-1}{k-2}\right)^t \right\rceil \frac{k-1}{(k-1)^{n-1}(k-2)} \\
&\le 8 \left(\frac{k-1}{k-2}\right)^{t+1} \\
&\le 16 \left(\frac{k-1}{k-2}\right)^t
\end{aligned}$$

Case 2: ω and β differ on the colouring of V_i for some i. In this case, $w(\gamma(\alpha, \eta)) = i^{-2}$. Also, by definition,

$$K_D^*[\omega \to \beta] = \inf_{\sigma \in \Omega} \operatorname{gap}(\mathcal{L}_{V_i}^{\sigma}) \pi_{V_i}^{\omega}[\beta].$$

But note that

$$\inf_{\sigma \in \Omega} \operatorname{gap}(\mathcal{L}_{V_i}^{\sigma}) = \left(\max_{\sigma \in \Omega} \tau_{V_i}^{\sigma}\right)^{-1} = (\tau_{V_i})^{-1}.$$

Also, since π is the uniform distribution, $\pi_{V_i}^{\omega}[\beta] = |\Omega_{V_i}^{\omega}|^{-1}$. We conclude

$$K_D^*[\omega \to \beta] = (\tau_{V_i} |\Omega_{V_i}^{\omega}|)^{-1}. \tag{16}$$

How many paths in $\gamma(\alpha, \eta)$ use the transition (ω, β)? As in the proof of Lemma 4.1, we can determine this by considering subcases for α and η. We will count the number of α, η in each subcase separately.

Case 2.1: $\alpha(v) = \eta(v)$. Recall that in this case a special, simple canonical path is used. We know $\alpha(v) = \eta(v) = \omega(v)$. Also, we know $\alpha(V_j) = \omega(V_j)$ for all $j \ge i$, and $\eta(V_j) = \beta(V_j)$ for all $j \le i$. So for all $j < i$ there are $|\Omega_{V_j}^{\omega}|$ possibilities for $\alpha(V_j)$, and for all $j > i$ there are $|\Omega_{V_j}^{\omega}|$ possibilities for $\eta(V_j)$. The total number of possibilities for α and η is therefore at most

$$\prod_{j \ne i} |\Omega_{V_j}^{\omega}|.$$

Now note $\prod_{1 \le j \le t} |\Omega_{V_j}^{\omega}| \le (k-1)^{n-1}$, as the colours for the $n-1$ vertices in the blocks $V_1, \dots, V_t$ can be chosen in a top-down manner. This implies that the number of possibilities for α and η is at most

$$\prod_{j \ne i} |\Omega_{V_j}^{\omega}| = \frac{1}{|\Omega_{V_i}^{\omega}|} \prod_{1 \le j \le t} |\Omega_{V_j}^{\omega}| \le \frac{(k-1)^{n-1}}{|\Omega_{V_i}^{\omega}|}.$$

Case 2.2: $\alpha(v) \ne \eta(v)$ and (ω, β) is the first change to V_i in $\gamma(\alpha, \eta)$. That is, (ω, β) is the first change in Step 1 of the canonical path description. In this case we know $\alpha(v) = \omega(v)$, $\alpha(V_j) = \omega(V_j)$ for all $j > i$, and $\alpha_{V_j}^{c_j'}(V_j) = \omega(V_j)$ for all $j < i$. For block V_i, we know $\alpha(V_i) = \omega(V_i)$ and $\eta_{V_i}^{b_i'}(V_i) = \beta(V_i)$. We wish to count the number of colourings α and η that satisfy these conditions.

Note first that ω determines $c_1', \dots, c_{i-1}'$ and β determines b_i'. Then, by Claim 5.7, the total number of possibilities for $c_1, \dots, c_{i-1}$ and b_i is at most $\prod_{j=1}^{i} |\partial V_j| \le 4$.

Now consider η. There are at most $k-1$ possibilities for $\eta(v)$, since $\eta(v) \ne \alpha(v)$. Then b_i and $\eta(v)$ determine $\eta(u_i)$. But now $\eta(V_i) = \beta_{V_i}^{\eta(u_i)}(V_i)$, so this is determined as well. Consider the number of possibilities for $\eta(V_j)$ for each $j \ne i$; note that this number of possibilities may vary depending on the choice made for $\eta(v)$. With this in mind, we define Ω_{V_j} to be the space of proper colourings of V_j with the colour of v chosen to maximize the number of proper colourings. That is, $|\Omega_{V_j}| \ge |\Omega_{V_j}^{\eta}|$ for all choices of $\eta(v)$. Then for each $j \neq i$ there are at most $|\Omega_{V_j}|$ possibilities for $\eta(V_j)$. Then the number of possibilities for η is at most

$$(k-1)\prod_{j\neq i}|\Omega_{V_j}| \le \frac{(k-1)}{|\Omega^{\omega}_{V_i}|}\prod_{1\le j\le t}|\Omega_{V_j}| \le \frac{(k-1)^n}{|\Omega^{\omega}_{V_i}|}. \tag{17}$$

Next consider α. Note that ω determines $\alpha(v)$ and also $\alpha(u_j)$ for all $j \ge i$. Furthermore, c_j is determined for all $1 \le j < i$. So by (5) there are at most $\lceil(\frac{k-1}{k-2})^{i-1}\rceil$ possibilities for $(\alpha(u_1), \ldots, \alpha(u_t))$. From each possibility all of α is determined, since $\alpha(V_j) = \omega(V_j)$ for all $j \ge i$ and $\alpha(V_j) = \beta^{\alpha(u_j)}_{V_j}(V_j)$ for all $j < i$. So the number of possibilities for α 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 is

$$4\left\lceil \left(\frac{k-1}{k-2}\right)^{i-1} \right\rceil \frac{(k-1)^n}{|\Omega^{\omega}_{V_i}|}.$$

Case 2.3: $\alpha(v) \neq \eta(v)$ and (ω, β) is the second change to V_i in $\gamma(\alpha, \eta)$. This is the second change in Step 1 of the canonical paths description. This case is nearly identical to Case 2.2; the only difference is that for block V_i we know $\eta^{b'_i}_{V_i}(V_i) = \omega(V_i)$ and $\alpha^{c'_i}_{V_i}(V_i) = \beta(V_i)$.

The only effect that this has on the analysis is that now c'_i is determined instead of $\alpha(u_i)$. By Claim 5.7 there are 2 possibilities for c_i given c'_i. Given c_i (instead of $\alpha(u_i)$), the factor due to (5) becomes $\lceil(\frac{k-1}{k-2})^{i}\rceil$. We conclude that the number of possibilities for α and η is at most

$$8\left\lceil \left(\frac{k-1}{k-2}\right)^{i} \right\rceil \frac{(k-1)^n}{|\Omega^{\omega}_{V_i}|}.$$

Case 2.4: $\alpha(v) \neq \eta(v)$ and (ω, β) is the third change to V_i in $\gamma(\alpha, \eta)$. This is the first change in Step 3 of the canonical paths description. We know that $\eta(v) = \omega(v)$, $\alpha^{c'_j}_{V_j}(V_j) = \omega(V_j)$ for all $j < i$, and $\alpha^{a'_j}_{V_j}(V_j) = \beta(V_j)$ for all $j > i$. Further, $\alpha^{c'_i}_{V_i}(V_i) = \omega(V_i)$ and $\eta(V_i) = \beta(V_i)$.

Note first that ω determines $c'_1, \ldots, c'_i$ and β determines $a'_{i+1}, \ldots, a'_t$. Then, by Claim 5.7, the total number of possibilities for $c_1, \ldots, c_i$ and $a_{i+1}, \ldots, a_t$ is at most $\prod_{j=1}^{t}|\partial V_j| \le 4$.

Consider η. Note that β determines $\eta(v)$ and $\eta(V_i)$. For each $j \neq i$, there are $|\Omega_{V_j}|$ possibilities for $\eta(V_j)$. Again following the methods used in (17), the number of possibilities for η is at most

$$\prod_{j\neq i}|\Omega_{V_j}| \le \frac{(k-1)^{n-1}}{|\Omega^{\omega}_{V_i}|}.$$

Now consider α. There are at most $k-1$ possibilities for $\alpha(v)$. Recall that colours $a_{i+1}, \ldots, a_t$ and colours $c_1, \ldots, c_i$ are determined. But then $\alpha(u_{i+1}), \ldots, \alpha(u_t)$ can be recovered and by (5) there are at most $\lceil(\frac{k-1}{k-2})^{i}\rceil$ possibilities for $(\alpha(u_1), \ldots, \alpha(u_t))$. These colours and ω determine the colouring α, since $\alpha(V_j) = \omega^{\alpha(u_j)}_{V_j}(V_j)$ for all $1 \le j \le t$. The number of possibilities for α 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 is at most

$$4\left\lceil \left(\frac{k-1}{k-2}\right)^{i} \right\rceil \frac{(k-1)^n}{|\Omega^{\omega}_{V_i}|}.$$

Case 2.5: $\alpha(v) \neq \eta(v)$ and (ω, β) is the fourth change to V_i in $\gamma(\alpha, \eta)$. This is the second change in Step 3 of the canonical paths description. This case is nearly identical to Case 2.4; the only difference is that for block V_i we know $\eta(V_i) = \omega(V_i)$ and $\alpha^{a'_i}_{V_i}(V_i) = \beta(V_i)$.

The only effect that this has on the analysis is that now a_i is determined instead of c_i. But a_i determines $\alpha(u_i)$, so the factor due to (5) becomes $\lceil(\frac{k-1}{k-2})^{i-1}\rceil$. We conclude that the number of possibilities for α and η is at most

$$4\left\lceil\left(\frac{k-1}{k-2}\right)^{i-1}\right\rceil\frac{(k-1)^n}{|\Omega_{V_i}^{\omega}|}.$$

Case 2.6: $\alpha(v) \neq \eta(v)$ and (ω, β) is the fifth change to V_i in $\gamma(\alpha, \eta)$. This is the change in Step 4 of the canonical paths description. In this case we know $\eta(v) = \omega(v)$, $\alpha_{V_j}^{a_j'}(V_j) = \omega(V_j)$ for all $j > i$, and $\eta(V_j) = \beta(V_j)$ for all $j < i$. For block V_i, we know $\alpha_{V_i}^{a_i'}(V_i) = \omega(V_i)$ and $\eta(V_i) = \beta(V_i)$.

In this case there are at most $(k-1)$ choices for $\alpha(v)$. The colours $a_i', \ldots, a_t'$ plus $\eta(v)$ are determined by ω. Claim 5.7 then implies that there are at most $\prod_{j=i}^{t}|\partial V_j| \leq 4$ possibilities for colours $a_i, \ldots, a_t$. From these colours (plus $\alpha(v)$) the colours $\alpha(u_i), \ldots, \alpha(u_t)$ are determined. Those colours then determine $\alpha(V_i), \ldots, \alpha(V_t)$ from ω. Furthermore, $\eta(V_1), \ldots, \eta(V_i)$ 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 $\alpha(v)$, we conclude that the number of possible options for α and η is at most

$$4\frac{(k-1)^n}{|\Omega_{V_i}^{\omega}|}.$$

This concludes our subcase analysis. Note that in each case, the number of possibilities for α and η was at most

$$8\left\lceil\left(\frac{k-1}{k-2}\right)^{i}\right\rceil\frac{(k-1)^n}{|\Omega_{V_i}^{\omega}|}.$$

Summing up over all cases, we get that the total number of possibilities for α and η, given that (ω, β) is a change in the colouring of V_i for some i, is at most

$$48\left\lceil\left(\frac{k-1}{k-2}\right)^{i}\right\rceil\frac{(k-1)^n}{|\Omega_{V_i}^{\omega}|} \leq 96\left(\frac{k-1}{k-2}\right)^{i}\frac{(k-1)^n}{|\Omega_{V_i}^{\omega}|}.$$

Substituting this and (16) into (14), we have

$$\begin{aligned}\rho_w(\omega, \beta) &\leq 960 i^2\left(\frac{k-1}{k-2}\right)^{i}\frac{(k-1)^n}{|\Omega_{V_i}^{\omega}|}\left(\frac{\tau_{V_i}|\Omega_{V_i}^{\omega}|}{(k-1)^{n-1}(k-2)}\right)\\ &\leq 2000 i^2\left(\frac{k-1}{k-2}\right)^{i}\tau_{V_i}.\end{aligned}$$

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

$$\rho_w \leq \max_{1\leq i\leq t} 2000 i^2\left(\frac{k-1}{k-2}\right)^{i}\tau_{V_i}.$$

Applying (12) and (11) we conclude that

$$\tau_T^{\sigma} \leq \tau_D^{\sigma} \leq 2000\max_{1\leq i\leq t} i^2\left(\frac{k-1}{k-2}\right)^{i}\tau_{V_i} \tag{18}$$

as required.

To see that this result holds when $k = 3$ given that $|\partial V_i| = 1$ for all V_i, note that the restriction $k \geq 4$ is used only to handle complications that occur when $|\partial V_i| = 2$; namely, to show that the Glauber dynamics is ergodic over such a block, and in the definition of colour c_i' when $|\partial V_i| > 1$. Thus the restriction $k \geq 4$ can be dropped if $|\partial V_i| < 2$ for all i. □

Remark 5.8. We note the effect of using the weighted version of the canonical paths bound, (12), in our proof of Lemma 5.2. If we had used the standard canonical paths bound, (6), then we would replace the factor of i^2 in (18) by the maximum length of a path, which is $5\Delta + 1$. Our final bound would then be of the form $\tau_T^\sigma \leq c\Delta \max_{1\leq i\leq t} \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}$. However, this would lead to a bound of $n^{O(\log\Delta + \Delta/k\log\Delta)}$ on the mixing time of the Glauber dynamics, which is weaker than $n^{O(1+\Delta/k\log\Delta)}$ when $k >> \Delta/\log^2\Delta$.

Remark 5.9. We also note the effect of making use of the weighted block dynamics. If we had applied Proposition 3.1 instead of Proposition 3.2, the bound in (16) would become $K_D[\omega \to \beta] = (\tau|\Omega^\omega_{V_i}|^{-1})$, where $\tau = \max_i \tau_{V_i}$. This would lead to a final 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 5.2. With this modified Lemma, the bound in (10) would become

$$\tau_2(n) \leq ct^2 \left(\frac{k-1}{k-2}\right)^t \tau_2\left(\lceil n/2\rceil\right).$$

This would yield an upper bound of $n^{O(1+\Delta/k)}$ on the mixing time of the Glauber dynamics, which is weaker than $n^{O(1+\Delta/k\log\Delta)}$.

6 Extensions

In this section we extend the techniques used to prove Theorem 1.2 to prove Theorems 1.3 and 1.4.

6.1 Upper Bound for Balanced Trees with $k = 3$

Let us now turn to a discussion of Theorem 1.3, which pertains to the case of three colours. The proof of Theorem 1.2 does not extend directly to the case $k = 3$. The critical issue is that Lemma 5.2 holds for $k = 3$ only if the subtrees under consideration have at most one boundary vertex, to ensure ergodicity of the Glauber dynamics. However, the decomposition techniques used in Claim 5.3 and Claim 5.4 generate subtrees with 2 boundary points.

On the other hand, the decomposition technique used in Theorem 1.1 (i.e. root the tree, and always choose the root of a subtree as the separator) does limit boundary sizes to at most 1. However, it does not limit the number of nodes in the generated subtrees. The class of balanced trees contains precisely those trees for which the sizes of subtrees are bounded under the decomposition method used in the proof of Theorem 1.1. We now give the details of the proof.

Proof of Theorem 1.3. Fix some $\epsilon > 0$. Define $\tau(n)$ to be the maximum relaxation time over any rooted ϵ-balanced tree T with $|T| \leq n$ and possibly with a fixed external boundary node adjacent to its root.

Claim 6.1. *There exists a constant c such that for all $n > 1$, either*

$$\tau(n) \leq c \max_{2\leq i\leq t} i^2 \left(\frac{k-1}{k-2}\right)^i \tau(\lfloor n/i\rfloor)$$

or

$$\tau(n) \leq c \left(\frac{k-1}{k-2}\right) \tau(\lfloor n(1-\epsilon)\rfloor).$$

Proof. Suppose T is a rooted ϵ-balanced tree with $|T| \leq n$ and root v. Let $u_1, \ldots, u_t$ be the children of v, with $1 \leq t \leq \Delta$. Then take $D = \{\{v\}, V_1, \ldots, V_t\}$, where $|V_1| \geq |V_2| \geq \ldots \geq |V_t|$ and V_i is the tree rooted at u_i. By Lemma 5.2, we know that

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

The claim follows by noting that $|V_i| \leq \lfloor n/i \rfloor$ for all $2 \leq i \leq t$ by our choice of indices, and $|V_1| \leq \lfloor n(1-\epsilon) \rfloor$ since T is ϵ-balanced. $\square$

It now follows by induction that $\tau(n) \leq n^{d(\frac{1}{\log(1/(1-\epsilon))}+\Delta/k\log\Delta)}$. Indeed, the base case of small n follows by judicious choice of d. For general n, we have one of two possibilities. First, it may be that $\tau(n) \leq i^2 \left(\frac{k-1}{k-2}\right)^i \tau(\lfloor n/i \rfloor)$ for some $2 \leq i \leq t$. But now the bound follows just as for Theorem 1.2, using the fact that $\frac{1}{k\log(1/(1-\epsilon))} < 1$.

On the other hand, we may have

$$\tau(n) \leq c\left(\frac{k-1}{k-2}\right)\tau(n(1-\epsilon)).$$

For ease of notation define $a = (1-\epsilon)^{-1}$, and note $a > 1$. Then the induction hypothesis implies

$$\begin{aligned}
\tau(n) &\leq c\left(\frac{k-1}{k-2}\right) n^{d(\frac{1}{\log(a)}+\Delta/k\log\Delta)} a^{-d(\frac{1}{\log(a)}+\Delta/k\log\Delta)} \\
&\leq a^{\log c/\log a} a^{\log e/(k-2)\log a} a^{-d(\frac{1}{\log(a)})} n^{d(\frac{1}{\log(a)}+\Delta/k\log\Delta)} \\
&\leq n^{d(\frac{1}{\log(a)}+\Delta/k\log\Delta)}
\end{aligned}$$

as long as $d \geq \max\{\log c, \log e\}$.

So we conclude that $\tau(n) = n^{O(\frac{1}{\log(1/(1-\epsilon))}+\Delta/k\log\Delta)}$, which is $n^{O(1+\Delta/k\log\Delta)}$ for any constant $\epsilon > 0$. Then the bound (1) gives us that the mixing time for the Glauber dynamics with k colours on any ϵ-balanced tree T with n nodes is at most $n^{O(1+\Delta/k\log\Delta)}$, as required. $\square$

6.2 Extending to general boundary conditions

In the proofs of Theorem 1.2 and Lemma 5.2, we were careful to consider only subtrees with at most 2 boundary vertices. This was enough to prove Theorem 1.2 and simplified our arguments. However, this restriction can be relaxed when $k > 4$. Indeed, all that is required by our technique is that $|\partial T| \leq k-2$.

Theorem 1.4 is a variant of Theorem 1.2 that uses this relaxation. We will prove it by making use of the following variant of Lemma 5.2.

Lemma 6.2. *Suppose $k \geq 3$ and let T be a subtree of a tree G with $b := |\partial T| \leq k-2$ and let $\sigma \in \Omega$ be a boundary condition for T. Choose $x \in T$ and consider $D_x = \{\{x\}, V_1, \ldots, V_t\}$, where $1 \leq t \leq \Delta$. Suppose $|\partial V_i| \leq b$ for each V_i. Then, for some constant c,*

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

Before proving Lemma 6.2, we will discuss how it implies Theorem 1.4. Indeed, this implication follows the deduction of Theorem 1.2 from Lemma 5.2 almost exactly. Define $\tau_b(n)$ as in Section 5.1. Then the argument from Claim 5.3 yields

$$\tau_{b-1}(n) \leq cb2^b \max_{1\leq i\leq t} i^2 \left(\frac{k-1}{k-2}\right)^i \tau_b(\lfloor n/i \rfloor). \tag{19}$$

To bound $\tau_b(n)$, we proceed as in Claim 5.4. 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 Claim 5.4 we can apply Lemma 6.2 with $v = y$, then use (19), to obtain the bound

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

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 (10) implies Theorem 1.2 in Section 5.

It remains to give the proof of Lemma 6.2, which mirrors the proof of Lemma 5.2 with two changes. First, the bounds on $|\Omega_T^\sigma|$ in (13) must change, since there are now multiple boundary vertices that can affect the number of admissible colours. The minimum number of possible colours occurs when all boundary vertices are adjacent to a single vertex x in T, and they all have different colours in σ. In this case, there will be only $(k-b-1)$ possible colours for x. We conclude that

$$(k-1)^{n-1}(k-b-1) \le |\Omega_T^\sigma| \le k(k-1)^{n-1}.$$

This change affects (14), adding a factor of $\frac{k-1}{k-b-1}$ to our analysis. Since $b \le k-2$, this extra factor is at most $b+1$.

The second complication that arises is that when we invoke Claim 5.7 during the case analysis for the canonical paths, the number of possible values for a_i', b_i' or c_i' given a_i, b_i, or c_i is $|\partial V_i|$, which is now bounded by b rather than 2.

We modify the case analysis as follows. When counting the number of possibilities for colours α and η, it is necessary to determine colours a_j, b_j, or c_j from a_j', b_j' or c_j'. The number of possibilities is at most $\prod_{j=1}^t |\partial V_j|$. However, we can write $|\partial V_j| = 1 + e_j$, where $e_j = |\partial V_j \cap \partial T|$. Then we know $\sum_{j=1}^t e_j = b$. The product $\prod_{j=1}^t (1+e_j)$ is therefore maximized when b of the entries e_j are 1, and the rest are 0. Therefore the number of possibilities for colours a_j, b_j, or c_j given the colours a_j', b_j', or c_j' is at most 2^b.

We conclude that the analysis for Lemma 5.2 leads to the same result, with an extra factor of $(b+1)2^b$. This gives Lemma 6.2 as required.

7 A Lower Bound

We will now prove the lower bound from Theorem 1.1: that the mixing time for the Glauber dynamics on the complete tree of degree Δ is $n^{\Omega(\Delta/k\log\Delta)}$. Note that this plus the upper bound from Section 4 completes the proof of Theorem 1.1. 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 $h \ge \log_\Delta n - 1$. For each $0 \le i \le h$, 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,

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

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 $t > T_0$,

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

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

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

so by Markov's Inequality

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

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

Now consider general $i \ge 1$. Let the children of v be $u_1, \ldots, u_\Delta$. For each $t \ge 0$, 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 $c \ne \sigma_t(v)$, 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 $\sigma_t(u_j) \ne c$ 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

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

by induction.

We now claim that the events $A(t,1,c), \ldots, A(t,\Delta,c)$ are mutually independent. For each i, $1 \le i \le \Delta$, 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

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

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

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

as required.

□

8 Open Problems

The obvious remaining problem is whether Theorem 1.2 can be extended to the case $k = 3$.

Another 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 is somewhere between $k-2$ and $2k-2$.

Finally, this raises some natural questions about the mixing time of the Glauber dynamics on planar graphs of bounded degree. As described in the introduction, Hayes, Vera and Vigoda[7] noted that when $\Delta \ge n^{\eta}$ for any $\eta > 0$ then certain trees require $k \ge c\Delta/\log\Delta$ for polytime mixing, where c is an absolute constant. Theorem 1.1 shows that the same is true for any Δ that grows with n. But for $\Delta = O(1)$, Theorem 1.2 shows that there are no trees that require $k > 4$. Is there any constant K such that for every $k \ge K$ and constant Δ, the Glauber dynamics mixes in polytime on the k-colourings of every planar graph with maximum degree Δ?

Acknowledgements

We thank Nayantara Bhatnagar, Thomas Hayes, Juan Vera and Eric Vigoda for some helpful discussions.

References

[1] N. Berger, C. Kenyon, E. Mossel and Y. Peres. *Glauber dynamics on trees and hyperbolic graphs.* Prob. Th. Related Fields **131**, 311 - 340 (2005).

[2] G. Brightwell and P. Winkler. *Random colorings of a Cayley tree.* In *Contemporary Combinatorics.* Bolyai Society Mathematical Studies, vol 10. B. Bollobás (Ed .), 247- -276 (2002).

[3] P. Diaconis and L. Saloff-Coste. *Comparison theorems for reversible Markov chains.* Ann. Appl. Prob. **3**, 696 - 730 (1993).

[4] P. Diaconis and D. Stroock. *Geometric bounds for eigenvalues of Markov chains.* Ann. Appl. Prob. **1**, 36 - 61 (1991).

[5] M. Dyer, L. Goldberg and M. Jerrum. *Systematic scan for sampling colorings.* Ann. Appl. Prob. **18**, 185 - 230 (2006).

[6] L. Goldberg, M. Jerrum and M. Karpinski. *The mixing time of Glauber dynamics for colouring regular trees.* Available at arXiv:0806.0921v1.

[7] T. Hayes, J. Vera and E. Vigoda. *Randomly coloring planar graphs with fewer colors than the maximum degree.* Proceedings of STOC 2007.

[8] M. Jerrum. *A very simple algorithm for estimating the number of k-colourings of a low-degree graph.* Rand. Struct. Alg. **7**, 157 - 165 (1995).

[9] M. Jerrum and A. Sinclair. *Approximating the permanent.* Siam. Jour. Comput. **18**, 1149 - 1178 (1989).

[10] J. Jonasson. *Uniqueness of uniform random colourings of regular trees.* Stat. & Prob. Letters **57**, 243 - 248 (2002).

[11] C. Jordan. *Sur les assemblages de lignes.* J. Reine Angew. Math. 70, 185 - 190 (1869).

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

[13] F. Martinelli, A. Sinclair and D. Weitz. *Fast mixing for independent sets, colorings and other models on trees.* Rand. Struc. & Alg. **31**, 134 - 172 (2007).

[14] M. Molloy. *Very rapidly mixing Markov Chains for* (2Δ)*-colourings and for independent sets in a 4-regular graph.* Rand. Struc. & Alg. **18**, 101 - 115 (2001).

[15] E. Mossel and A. Sly. *Gibbs Rapidly Samples Colourings of* $G(n, d/n)$. Available at Arxiv http://front.math.ucdavis.edu/0707.3241 (2007).

[16] D. Randall. *Mixing.* Proceedings of FOCS 2003.

[17] A. Sinclair. *Improved bounds for mixing rates of Markov chains and multicommodity flow.* Combinatorics, Probability and Computing, 1:351 - 370 (1992).

[18] E. Vigoda. *Improved bounds for sampling colorings.* J. Math. Physics, **41**, 1555 - 1569 (2000).