The Glauber dynamics for colourings of bounded degree trees

B. Lucier[*]   M. Molloy[†]   Y. Peres[‡]

April 12, 2009

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 *every* tree is at most $n^{O(1+\Delta/(k \log \Delta))}$. This bound is tight up to the constant factor in the exponent, as evidenced by the complete tree. 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.
[‡]Microsoft Research, peres@microsoft.com.

1 Introduction

The Glauber dynamics is a Markov chain over configurations of spin systems on graphs, of which k-colourings is a special case. Such chains have generated a great deal of interest 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, 14]. 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[19] 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. [14] 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). Independently, Goldberg, Jerrum and Karpinski [6] and Lucier and Molloy [13] showed a lower bound of $n^{\Omega(1+\Delta/k\log\Delta)}$ on the mixing time for the case of the complete tree. Goldberg, Jerrum and Karpinski also showed an upper bound of $n^{O(1+\Delta/\log\Delta)}$ for complete trees and constant Δ.

Our main result is an upper bound for *every* tree. Our upper bound is asymptotically tight, in that it matches the lower bound for complete trees up to a constant factor in the degree.

Theorem 1.1. *For $k \geq 3$, the Glauber dynamics on k-colourings of any tree with maximum degree Δ mixes in time at most $n^{O(1+\Delta/k\log\Delta)}$.*

Thus, for every $k \geq 3$ and $\Delta = O(1)$, we have polytime mixing on every tree. But if Δ grows with n, no matter how slowly, then on some trees (eg. complete trees) we require the $\Omega(\Delta/\log\Delta)$ colours for polytime mixing that Hayes, Vera and Vigoda noted were required at $\Delta = n^\epsilon$.

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.

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 [14] 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 1.1 extends to show:

Theorem 1.2. *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. We prove Theorem 1.1 in Section 4. in Appendix A we sketch a proof of Theorem 1.2.

Remark 1.3. Our arguments can be extended to prove analogous upper bounds for a wide range of other instances of the Glauber dynamics, including the Ising model. Details will appear in a full version of the paper.

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, 17], 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.

We will additionally be interested in a variant of the Glauber dynamics, $\mathcal{L}_2$, that can also recolour pairs of adjacent vertices. That is, on each step of $\mathcal{L}_2$, a connected subgraph $S \subseteq T$ of size at most 2 is chosen uniformly at random. If the initial configuration is η, then the subgraph S is recoloured according to the uniform distribution on Ω_S^η.

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. [17]):

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

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, and $\overline{\partial}T$ for the interior boundary of T. That is, $\partial T = \{x \in V \backslash T : N(x) \cap T \neq \emptyset\}$ and $\overline{\partial}T = \{x \in T : N(x) \cap \partial T \neq \emptyset\}$. Note that for each $x \in \partial T$ there is a unique $y \in \overline{\partial}T$ adjacent to x.

We will analyze the Glauber dynamics over a subtree T of G with boundary configuration σ. The following simple claims analyze the ergodicity of the Glauber dynamics and 2-path Glauber dynamics on trees; their proofs are deferred to the appendix.

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.*

Claim 2.2. *Suppose $k = 3$ and let T be a subtree of G with $|\partial T| \leq 2$. Then the 2-path Glauber dynamics is ergodic over Ω_T^σ for all $\sigma \in \Omega$, with uniform stationary distribution.*

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, including the 2-path Glauber dynamics defined above. 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 (or 2-path 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

3

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

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

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 $\pi_{V_i}^\sigma$, where σ is the previous colouring.

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

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

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

$$K_D^*[\sigma \to \eta] = \begin{cases} g_i \pi_{V_i}^\sigma[\eta] & \text{for all } \eta, i \text{ such that } \eta \in \Omega_{V_i}^\sigma \\ 0 & \text{otherwise.} \end{cases}$$

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

The proof of Proposition 3.2 is a simple modification to the proof of Proposition 3.1 [12]. The details are given in the appendix. 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 4.8).

We next show that the weighted block dynamics is equivalent to a related process on a subgraph of T. Informally, we wish to "collapse" each block to its set of internal boundary nodes. We will assign colours to these boundary nodes according to the probability such a boundary configuration would occur in the block dynamics. More formally, suppose $D = \{V_1, \ldots, V_m\}$ is a set of blocks of vertices of T. Let $B = \cup_{i=1}^m \overline{\partial} V_i$. That is, B contains all internal boundary nodes for the blocks in D.

We now define a Markov process $\mathcal{L}_B$ on Ω_B, which simulates the behaviour of $\mathcal{L}_D$ restricted to the nodes in B. Given distribution π over Ω_T, $S \subseteq T$, and $\eta \in \Omega_S$, we will write $\pi_T[\eta' : \eta'(S) = \eta(S)]$ to denote $\sum_{\eta' : \eta'(S) = \eta(S)} \pi_T[\eta']$, the total probability that the configuration of S agrees with η. Then process $\mathcal{L}_B$ is defined by

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

In other words, a configuration η is chosen according to the probability that η appears as the colours on B after a step of the block dynamics on V_i. Our claim is that the relaxation times of $\mathcal{L}_D^*$ and $\mathcal{L}_B$ are the same. This Claim is similar to Claim 2.9 due to Berger et al [1]; the proof is in the Appendix.

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

4

4 An Upper Bound for General Trees

We now begin our proof of Theorem 1.1. Our approach is to decompose a tree into smaller subtrees, apply the block dynamics to the resulting subgraphs, and then use induction to bound the mixing time of the entire tree. Implicitly, this yields an iterative decomposition of the tree into smaller and smaller subtrees. How should we decompose a tree? A first idea is to root the tree, then take each subtree of the root as a block (with each subtree being rooted at a child of the root). A nice property of this decomposition is that each subtree has at most one boundary node, which must be adjacent to its root. In this case there will be h levels of recursion in the induction, where h is the height of tree T, and we will obtain a bound of the form c^h, where $c = c(\Delta, k)$ is the mixing time for an instance of the block dynamics. This method works for complete trees (and indeed was used in Berger et al. [1] in their analysis of complete trees), since they have logarithmic height. However, the height of a general tree could be much greater; in fact, it could be n. This leads to a super-polynomial bound on the mixing time.

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

There are some difficulties with the above approach that must be overcome. First, since we no longer split trees at their roots, a subtree T may have multiple boundary nodes, which complicates the behaviour of the block dynamics. We will limit this complication by making our choice of x carefully, so that subtree boundaries are of size at most 2. Second, for non-complete trees we might have blocks of vastly differing sizes, which makes a tight analysis of the block dynamics more difficult. We therefore use the weighted version of the block dynamics to perform our analysis.

In this section we describe our choice of blocks for the block dynamics. We then show that the upper bound of Theorem 1.1 holds, given a bound on the relaxation time of the block dynamics. The details of analyzing the block dynamics are encapsulated in Lemma 4.2, which is proved in Section 4.1.

Let T be any tree with maximum degree Δ. Suppose $|T| = n$ and $|\partial T| \leq 2$ (that is, T has at most two external boundary nodes). Let σ be a boundary configuration for T. If $k \geq 4$, then let $\mathcal{L}$ denote the Glauber dynamics on T with k colours and boundary configuration σ. If $k = 3$, then take $\mathcal{L}$ to be the 2-path Glauber dynamics on T with boundary configuration σ. Either way, since $|\partial T| \leq 2$, Claims 2.1 and 2.2 imply that $\mathcal{L}$ is ergodic. Let τ_T^σ denote the relaxation time for $\mathcal{L}$. 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 1.1 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 $2 \leq i \leq \Delta$,

$$\tau_2(n) \leq ci^2 \left(\frac{k-1}{k-2}\right)^{i+1} \tau_2\left(\lfloor n/i \rfloor\right). \tag{3}$$

First let us show how (3) implies Theorem 1.1 when $k \geq 4$. By induction on n, (3) implies that $\tau_2(n) \leq n^{d(1+\Delta/k \log \Delta)}$ for some constant d (since we can assume $k \leq 2\Delta$, as otherwise the result is known [7]). Then, by (1), the mixing time of the Glauber dynamics satisfies $\mathcal{M}(\mathcal{L}) \leq (n \log k)\tau_G \leq (n \log k)\tau_2(n) = n^{O(1+\Delta/k \log \Delta)}$ as required. For $k = 3$, (3) implies that the 2-path Glauber dynamics mixes in time $n^{O(1+\Delta/k \log \Delta)}$. The result of Theorem 1.1 for $k = 3$ then follows from Lemma 4.1 below.

Lemma 4.1. *Let $\mathcal{L}_1$ denote the Glauber dynamics with $k = 3$ colours, and $\mathcal{L}_2$ denote the 2-path Glauber dynamics again with $k = 3$ colours. For any T with $|\partial T| \leq 1$ and boundary configuration ξ, $\tau(\mathcal{L}_1^\xi) \leq n^{O(\Delta/ \log \Delta)}\tau(\mathcal{L}_2^\xi)$.*

Proof (sketch). We wish to apply the comparison method of Diaconis and Saloff-Coste [3]. We note that this application is not immediate, since a step of $\mathcal{L}_2$ cannot always be simulated by a small number of steps of $\mathcal{L}_1$. We therefore consider an intermediate process, which performs a cyclic shift of all colours of

a subtree of T in one step. It is easy to see that such a process can be used to simulate a step of $\mathcal{L}_2$. To compare with $\mathcal{L}_1$, we simulate a rotation step by changing the colours of nodes in a bottom-up fashion. If these changes are ordered carefully, one can simulate a rotation of colours in $O(n)$ steps of $\mathcal{L}_1$, where each step has a congestion of $n^{O(\Delta/\log \Delta)}$. The term $n^{O(\Delta/\log \Delta)}$ derives from a bound on the number of siblings of ancestors of a given node. Additional proof details can be found in the Appendix. $\qquad\square$

We now turn to proving (3). The following Lemma will be the workhorse for our proof.

Lemma 4.2. *Suppose $k \geq 3$ and let T be a subtree of a tree G with $|\partial T| \leq 2$ and let $\sigma \in \Omega$ be a boundary condition for T. Choose $v \in T$ and let $D_v = \{\{v\}, V_1, \ldots, V_t\}$ be a partition of T into disjoint connected subtrees, where $1 \leq t \leq \Delta$. Suppose $|\partial V_i| \leq 2$ for each V_i. Then there exists constant c such that*

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

We will prove Lemma 4.2 in Section 4.1. First, let us show how it implies (3). We will first consider trees with boundaries of size one, then trees with boundaries of size two.

Claim 4.3. *For some $2 \leq i \leq \Delta$, we have $\tau_1(n) \leq ci^2 \left(\frac{k-1}{k-2}\right)^i \tau_2\left(\lfloor n/i \rfloor\right)$.*

Proof. Let T be a subtree of G with $|\partial T| \leq 1$. 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 $|V_i| \leq \lfloor n/2 \rfloor$ for all $1 \leq i \leq t$ (see eg. [11]). We will choose our indices so that $|V_1| \geq |V_2| \geq \ldots \geq |V_t|$. Note that since $|\partial T| \leq 1$, we will have $|\partial V_i| \leq 2$ for all i. By Lemma 4.2, $\tau_T \leq ci^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}$ for some $1 \leq i \leq t$. 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 $\tau_T \leq ci^2 \left(\frac{k-1}{k-2}\right)^i \tau_2\left(\lfloor n/i \rfloor\right)$ 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 $\tau_T \leq c \left(\frac{k-1}{k-2}\right) \tau_{V_1} < c(2)^2 \left(\frac{k-1}{k-2}\right)^2 \tau_2\left(\lfloor n/2 \rfloor\right)$ as required. $\qquad\square$

Claim 4.4. *For some $2 \leq i \leq \Delta$, $\tau_2(n) \leq c^2 i^2 \left(\frac{k-1}{k-2}\right)^{i+1} \tau_2(\lfloor n/i \rfloor)$.*

Proof. Let T be a subtree with $|T| = n$ and $|\partial T| = 2$, say $\partial T = \{z_1, z_2\}$. Choose x as in Claim 4.3, with 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. Let $D_x = \{\{x\}, V_1, \ldots, V_t\}$ be a partition into disjoint connected subtrees, indexed so that $|V_1| \geq \ldots \geq |V_t|$; note that $|\partial V_i| \leq 2$ for all i. We then apply Lemma 4.2 as in Claim 4.3 and obtain the desired result.

Now suppose that x is not on the boundary path for T. Consider T to be rooted at some $r \in \overline{\partial} T$. Let y be the least ancestor of x that lies on the boundary path. Consider $D_y = \{\{y\}, V_1, \ldots, V_t\}$. Since x separates T into subtrees of size at most $\lfloor n/2 \rfloor$, in particular the subtree containing y must have size at most $\lfloor n/2 \rfloor$. This implies that the subtree separated by y that contains x must contain at least $\lfloor n/2 \rfloor$ nodes, and is therefore V_1, the largest subtree separated by y. Also, all the subtrees separated by y have boundary of size at most 2 (since y is on the boundary path for T). Lemma 4.2 applied to D_y yields:

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

for some i. If $i > 1$ then we obtain the desired result since $|V_i| \leq \lfloor n/i \rfloor$. If $i = 1$, then 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 4.3 yields:

$$\tau_T \leq c \left(\frac{k-1}{k-2}\right) \tau_1(|V_1|) \leq c \left(\frac{k-1}{k-2}\right) \tau_1(n)$$

$$\leq c^2 i^2 \left(\frac{k-1}{k-2}\right)^{i+1} \tau_2\left(\lfloor n/i \rfloor\right) \quad \text{for some } 2 \leq i \leq \Delta.$$

$\qquad\square$

We have now derived (3), completing the proof of Theorem 1.1.

4.1 Proof of Lemma 4.2

We now proceed with the proof of Lemma 4.2, whose goal is to bound the relaxation time on a tree with respect to the relaxation times for subtrees. Our approach is to use a canonical paths argument to bound the behaviour of the block dynamics. Indeed, there is a simple canonical path to move from some configuration σ to another configuration η: modify the configuration of each V_i to an intermediate state so that v is free to change colour to $\eta(v)$, make that change to $\eta(v)$, then set the configuration of each V_i to $\eta(V_i)$. The block dynamics paired with this path yields a bound on the relaxation time. However, that bound is not tight enough to imply the mixing rate we desire; indeed, it only implies a mixing time of $n^{O(\Delta)}$. We therefore apply the following sequence of improvements to the above approach.

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

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

3. We apply the weighted block dynamics, so that we can differentiate between large and small subtrees. In our paths, we always change configurations of blocks in order of subtree size. This improves our bound to $n^{O(\log \Delta + \Delta/k \log \Delta)}$. See Remark 4.8.

4. We apply weights to our canonical path to discount the congestion on smaller subtrees. The net effect is that the presence of many small subtrees does not influence the congestion of our paths. This improves our bound to $n^{O(1 + \Delta/k \log \Delta)}$. See Remark 4.7.

We now describe our full construction, elaborating on each of the above refinements.

4.1.1 THE BLOCK DYNAMICS

Recall now the conditions of Lemma 4.2. Suppose $k \geq 3$ 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 $D = \{\{v\}, V_1, \ldots, V_t\}$, where $1 \leq t \leq \Delta$. Suppose we choose v so that $|\partial V_i| \leq 2$ for each V_i. We will think of T as being rooted at v; then let u_i denote the root of V_i (ie. the neighbour of v in V_i).

To simplify our argument, we will make the assumption that $u_i \notin \overline{\partial} T$ for all i. At the conclusion of our proof we will discuss the (simple) extension to handle the case that $u_i \in \overline{\partial} T$ for some i.

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

Next recall the definition of graph B and dynamics $\mathcal{L}_B$ from Proposition 3.3. In this context, we can view $\mathcal{L}_B$ as a version of $\mathcal{L}_D$ wherein each block is treated like a single vertex. That is, B is a star with internal node v; we will refer to $u_1, \ldots, u_t$ as the leaf nodes of B. When such a leaf node, say u_i, is chosen by the dynamics, its colour updates with probability corresponding to the probability of seeing that colour as the root of V_i in $\mathcal{L}_D$. By Proposition 3.3, $\tau(\mathcal{L}_D^\sigma) = \tau(\mathcal{L}_B^\sigma)$. It is therefore sufficient to bound $\tau(\mathcal{L}_B^\sigma)$. Note that this is true even for the special case of $k = 3$, as $\mathcal{L}_B^\sigma$ depends only on the ergodicity of $\mathcal{L}$ (the 2-path Glauber dynamics) and its stationary distribution, which is uniform. The following simple Lemma will allow us to bound the transition probabilities of $\mathcal{L}_B^\sigma$; its proof can be found in the appendix.

Lemma 4.5. *Choose $S \subseteq T$ with $|\partial S| \leq 2$ and boundary configuration ξ, and suppose $x \in \overline{\partial} S$. Choose $c \in A$ and suppose there exists some $\eta \in \Omega_S^\xi$ with $\eta(x) = c$. Then $\pi_S^\xi[\omega : \omega(x) = c] \geq 1/k$.*

The following corollary is now immediate from Lemma 4.5 and the definition of $\mathcal{L}_B$.

Corollary 4.6. *Suppose $\alpha, \omega \in \Omega_B^\sigma$, $K_B[\alpha \to \omega] > 0$, and $\alpha(u_i) \neq \omega(u_i)$. Then $K_B[\alpha \to \omega] \geq (k\tau_{V_i}^\sigma)^{-1}$.*

4.1.2 DEFININITION OF INTERMEDIATE CONFIGURATIONS

Choose two colourings $\alpha, \eta \in \Omega_B$. Our goal is to define a sequence of steps of $\mathcal{L}_B$ that begins in state α and ends in state η. If $\alpha(v) = \eta(v)$ this sequence is simple: the colours of nodes $u_1, \ldots, u_t$ are changed from α to η one at a time. If $\alpha(v) \neq \eta(v)$, our strategy is to first change the colours of $u_1, \ldots, u_t$ so that none have colour $\eta(v)$, then change the colour of v to $\eta(v)$, and finally set the colours of the u_i nodes to match η. The obvious way to do this requires two "passes" of changes over the leaf nodes, 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 $1 \leq i \leq t$ we will define three colours, a_i, b_i, and c_i, that depend on α and η. The first two colours are easy to define:

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

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

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 $(\alpha(u_1), \ldots, \alpha(u_t))$ to a vector of colours $(c_1, \ldots, c_t)$ such that for all i, $c_i \notin \{\alpha(v), \eta(v)\}$. We want f to satisfy the following balance property: for all $1 \leq i \leq t$,

$$\#\{(x_1, \ldots, x_t) : \ (x_j = \alpha(u_j) \ \forall j > i) \wedge (f(x_1, \ldots, x_t)_j = c_j \ \forall j \leq i)\} \leq \left\lceil \left(\frac{k-1}{k-2} \right)^i \right\rceil. \tag{4}$$

That is, for any $1 \leq i \leq t$, if we are given $c_1, \ldots, c_i$ and $\alpha(u_{i+1}), \ldots, \alpha(u_t)$, there are at most $\left\lceil \left(\frac{k-1}{k-2} \right)^i \right\rceil$ possibilities for $\alpha(u_1), \ldots, \alpha(u_t)$. We defer the construction of f to Appendix B.4.

4.1.3 THE PATH DEFINITION

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

1. For each u_i in increasing order: recolour from $\alpha(u_i)$ to b_i, then 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 $\eta(u_i)$, then to a_i.
4. For each u_i in increasing order: recolour from a_i to $\eta(u_i)$.

The reader is encouraged to verify that all steps of this path are valid transitions according to $\mathcal{L}_B^\sigma$. The number of changes to the colour of each u_i seems excessive, but we define our path this way to maintain an important property: each change is from a colour derived from α to a colour derived from η, or vice-versa. This will be important in our analysis of the path congestion, defined below.

4.1.4 ANALYSIS OF WEIGHTED PATH CONGESTION

We will now define the weighted congestion of our choice of paths. For each $(\omega, \beta) \in \Gamma$, we will define a weight $w(\omega, \beta) > 0$. 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 vertex u_i. We define the weight of a path by $w(\gamma(\alpha, \eta)) = \sum_{(\omega, \beta) \in \gamma(\alpha, \eta)} w(\omega, \beta)$. Then note that for all $\gamma(\alpha, \eta)$, $w(\gamma(\alpha, \eta)) \leq 1 + 5 \sum_{i=1}^{t} i^{-2} < 1 + 5 \left(\frac{\pi^2}{6} \right) < 10$. For each edge $(\omega, \beta) \in \Gamma$, define the weighted congestion of that edge, $\rho_w(\omega, \beta)$, as

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

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 $\tau_D^\sigma \leq \rho_w$. We note that this bound and its proof are implicit in [4] (see their Remark on page 38). Note that the standard use of canonical paths 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,

8

our approach obtains a bound of $n^{O(1+\Delta/k\log\Delta)}$, whereas the standard canonical paths bound would only give a bound of $n^{O(\log\Delta+\Delta/k\log\Delta)}$ (see Remark 4.7).

Our result now follows from bounding $\rho_w(\omega,\beta)$. Using the uniformity of π, note that

$$\rho_w(\omega,\beta) \leq 10\left(\frac{1}{w(\omega,\beta)} \times |\{\gamma(\alpha,\eta) \ni (\omega,\beta)\}| \times \frac{1}{(k-1)^{t+1}K_B[\omega\to\beta]}\right). \tag{5}$$

We now 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 $\mathcal{L}_B$, we have $K_B[\omega\to\beta] = \inf_{\sigma\in\Omega}\mathrm{gap}(\mathcal{L}^\sigma_{\{v\}})\pi^\omega_{\{v\}}[\phi:\phi(v)=\beta(v)]$. But note that $\mathrm{gap}(\mathcal{L}^\sigma_{\{v\}})=1$ for all boundary conditions, since this is the set of proper colourings on a single vertex which mixes in a single step. Also, $\pi^\omega_{\{v\}}$ is the uniform distribution over $\Omega_{\{v\}}$, which is the set of at most $k-1$ colours not appearing on $\partial\{v\}$ in ω. We conclude

$$K_B[\omega\to\beta] \geq \frac{1}{k-1}. \tag{6}$$

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 $\alpha(u_i)=\omega(u_i)$ for all u_i.

Consider the possibilities for colouring η. Configuration β determines $\eta(v)$, and there are $(k-1)^t$ choices for η given $\eta(v)$ (consider choosing the colours for $u_1,\ldots,u_t$, which cannot be $\eta(v)$). Now consider α: the colour $\alpha(v)$ is determined by ω, as are $(c_1,\ldots,c_t)$. Thus by (4) there are at most $\lceil(\frac{k-1}{k-2})^\Delta\rceil$ possibilities for $(\alpha(u_1),\ldots,\alpha(u_t))$, which determines α. Putting this together, the total number of colourings α and η that satisfy $(\omega,\beta)\in\gamma(\alpha,\eta)$ is at most $(k-1)^t\left\lceil\left(\frac{k-1}{k-2}\right)^t\right\rceil$. Substituting this and (6) into (5), we conclude

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

Case 2: ω and β differ on the colour of u_i for some i. In this case, $w(\gamma(\alpha,\eta))=i^{-2}$. Also, since there exists a colouring of V_i in which u_i has colour $\beta(u_i)$ (recalling our assumption that $u_i\notin\overline{\partial}T$), Corollary 4.6 implies

$$K_B[\omega\to\beta] \geq (k\tau_{V_i})^{-1}. \tag{7}$$

How many paths in $\gamma(\alpha,\eta)$ use the transition (ω,β)? We consider subcases for α and η. We give only one subcase here; the remaining 5 cases (which are very similar) are described in the Appendix.

Subcase: $\alpha(v)\neq\eta(v)$ and (ω,β) is the first change to u_i in $\gamma(\alpha,\eta)$. That is, (ω,β) is the first 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 $j\geq i$, $b_i=\beta(u_i)$, and $c_j=\beta(u_j)$ for all $j<i$. How many α,η satisfy these conditions?

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

Next consider α. Note that ω determines $\alpha(v)$ and also $\alpha(u_j)$ for all $j\geq i$. Also, β determines c_j for all $j<i$. By (4), the number of possibilities for $\alpha(u_1),\ldots,\alpha(u_t)$ is at most $\left\lceil\left(\frac{k-1}{k-2}\right)^{i-1}\right\rceil$. The total number of α and η is therefore at most $\left\lceil\left(\frac{k-1}{k-2}\right)^{i-1}\right\rceil(k-1)^t$. This completes the subcase.

In each subcase, the number of possibilities for α and η is at most $\left\lceil\left(\frac{k-1}{k-2}\right)^i\right\rceil(k-1)^t$. Summing up over all subcases, we get that the total number of possibilities for α and η, given that (ω,β) is a change in the colouring of u_i, is at most $12\left(\frac{k-1}{k-2}\right)^i(k-1)^t$. Substituting this and (7) into (5), we have

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

9

This concludes our case analysis. Cases 1 and 2 (and the fact that $\tau_{V_t} \geq 1$) yield that $\rho_w \leq \max_{1 \leq i \leq t} 180 i^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}$. Applying the canonical paths bound and Proposition 3.2 we conclude that

$$\tau_T^\sigma \leq \tau_D^\sigma \leq 180 \max_{1 \leq i \leq t} i^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}. \tag{8}$$

4.1.5 REMOVING RESTRICTIONS ON BOUNDARY NODES

Recall that in the analysis above we assumed that no u_i was in $\overline{\partial}T$. We now sketch the method for removing this assumption; additional details appear in Appendix B.6. We made use of the assumption $u_i \notin \overline{\partial}T$ in our use of the colours a_i, b_i, c_i: this assumption allowed us to assume that there existed colourings of V_i in which the colour of u_i was a_i (or b_i, or c_i). If $u_i \in \overline{\partial}T$, it's possible that one or more of these colours will conflict with the boundary configuration, and the block dynamics we described might not be ergodic (when $k = 3$).

We will modify our selection of blocks to guarantee that no leaf of B is adjacent to the boundary of T. Specifically, we replace block $\{v\}$ with a block $R \subseteq T$ that contains v and any neighbouring nodes in $\overline{\partial}T$. Note $|R| \leq 3$. Our new set of blocks D will contain R and all subtrees separated by R. The net effect is that B (from Proposition 3.3) will no longer be a star, but rather a tree with at most 3 internal nodes. We then bound the relaxation time of $\mathcal{L}_B$ as before, extending our set of canonical paths in the natural way. For path $\gamma(\alpha, \eta)$, modify the leaf colours to allow the nodes of R to change from α to η, make this change to R in one step, then modify the leaf colours again to agree with η. The congestion analysis for this new set of canonical paths is very similar to the original, and we obtain the same result up to a constant factor.

Remark 4.7. We note the effect of using the weighted version of the canonical paths bound in our proof of Lemma 4.2. If we had used the standard canonical paths bound, then we would replace the factor of i^2 in (8) 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 4.8. We also note the effect of using the weighted block dynamics. If we had applied Proposition 3.1 instead of Proposition 3.2, the bound in (7) would become $K_B[\omega \to \beta] \geq (k\tau)^{-1}$, 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 4.2. With this modified Lemma, the bound in (3) 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 $n^{O(1+\Delta/k)}$, which is weaker than $n^{O(1+\Delta/k \log \Delta)}$.

5 Open Problems

Our results raise some natural questions about the Glauber dynamics on planar graphs of bounded degree. As described in the introduction, Hayes, Vera and Vigoda [7] noted that when $\Delta \geq n^\eta$ for any $\eta > 0$ then certain trees require $k \geq c\Delta/\log \Delta$ for polytime mixing, where c is an absolute constant. The same is true for any Δ that grows with n [13]. But for $\Delta = O(1)$, Theorem 1.1 shows that there are no trees that require $k > 3$. Is there a constant K such that for every $k \geq K$ and constant Δ, the Glauber dynamics mixes in polytime on k-colourings of every planar graph with maximum degree Δ?

Another question is how far Theorem 1.2 can be extended. In other words, how many leaves can we fix and still guarantee polytime mixing? It is easy to fix the colours of $k-1$ neighbours of each of two adjacent vertices u, v so that the chain is not ergodic, so the answer lies between $k-2$ and $2k-2$.

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 (2008).

[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] B. Lucier, M. Molloy. *The Glauber dynamics for colourings of bounded degree trees.* Submitted to journal (2008).

[14] 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).

[15] 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).

[16] 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).

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

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

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

A   Extending to general boundary conditions

In the proofs of Theorem 1.1 and Lemma 4.2, we were careful to consider only subtrees with at most 2 boundary vertices. This was enough to prove Theorem 1.1 and simplified our arguments. However, this restriction can be relaxed when $k > 4$. Indeed, all that is required by our technique is that $|\partial T| \leq k - 2$.

Theorem 1.2 is a variant of Theorem 1.1 that uses this relaxation. We will prove it by making use of the following variant of Lemma 4.2.

Lemma A.1. *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 A.1, we will discuss how it implies Theorem 1.2. Indeed, this implication follows the deduction of Theorem 1.1 from Lemma 4.2 almost exactly. Define $\tau_b(n)$ as in Section 4.1.

Then the argument from Claim 4.3 yields

$$\tau_{b-1}(n) \leq cb2^b \max_{1 \leq i \leq t} i^2 \left(\frac{k-1}{k-2}\right)^i \tau_b(\lfloor n/i \rfloor). \tag{9}$$

To bound $\tau_b(n)$, we proceed as in Claim 4.4. 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 4.4 we can apply Lemma A.1 with $v = y$, then use (9), to obtain the bound

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

Induction then implies that $\tau_b(n) \leq n^{d(1+b+\Delta/k \log \Delta)}$ for some sufficiently large constant d. This follows in precisely the same way that (3) implies Theorem 1.1 in Section 4.

It remains to give the proof of Lemma A.1, which mirrors the proof of Lemma 4.2 with two changes. First, we require the following more general version of Lemma 4.5.

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

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

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

$$= \frac{1}{1 + \frac{(k-2)(k-1)}{k-b+1}} \geq \frac{1}{(b+1)k}$$

as required. $\square$

This change to Lemma 4.5 affects (5), adding a factor of $(b+1)$ to our analysis.

Second, recall our discussion of the case that $u_i \in \overline{\partial} T$ at the end of the proof of Lemma 4.5. We will handle this case in a different way. For each $1 \leq i \leq t$ we will define colours a_i', b_i', and c_i'. If there exists a colouring of V_i in which u_i has colour a_i, then set $a_i' = a_i$. Otherwise it must be that $u_i \in \overline{\partial} T$, and in this case we set a_i' to be any other colour not in $\{\alpha(v), \eta(v)\}$. We define b_i' and c_i' in a similar way. We think of a_i', b_i', and c_i' as "fixed" versions of the original colours. We then use a_i', b_i', and c_i' in the definition of our canonical paths instead of colours a_i, b_i, and c_i.

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

We conclude that the analysis for Lemma 4.2 leads to the same result, with an extra factor of $(b+1)2^b$. This gives Lemma A.1 as required.

B Missing Proofs

B.1 Proofs from Section 2

Proof of Claim 2.1. 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. $\square$

Proof of Claim 2.2. Write $\mathcal{L}_2^\sigma(T)$ for the 2-path Glauber dynamics on subtree T. We first show that $\mathcal{L}_2^\sigma(T)$ is ergodic whenever $|\partial T| \leq 1$. Choose any $\eta_1, \eta_2 \in \Omega_T^\sigma$. Choose any $x \in T$ such that $\overline{\partial} T \subseteq \{x\}$, and consider x to be the root of T. We proceed by induction on the height of this rooted tree. For any $v \in T$, write $T(v)$ for the subtree of T rooted at v. Then for each child y of x, $|\partial T(y)| \leq 1$, and y is the unique node of $T(y)$ such that $\overline{\partial} T(y) \subseteq \{y\}$. Then, by induction, $\mathcal{L}_2(T(y))$ is ergodic for any boundary condition. The configuration of $T(y)$ can therefore be changed to that the colour of y is not $\eta_2(x)$, while the colour of x does not change from $\eta_1(x)$. We repeat this process for all children of x, then change the colour of x to $\eta_2(x)$. Finally, for each child y of x, we can recolour $T(y)$ to agree with η_2, again by ergodicity of $T(y)$ by induction. We have then reached configuration η_2 on T, as required.

We next show that $\mathcal{L}_2^\sigma(T)$ is ergodic when $|\partial T| = 2$. Choose some $r \in \overline{\partial} T$ and consider r to be the root of T. Let P be the path in T connecting the vertices in $\overline{\partial} T$. Note that since $\overline{\partial} T \subseteq P$, $|P| \geq 2$. We now proceed by induction on $|P|$. Choose any $\eta_1, \eta_2 \in \Omega_T^\sigma$. We will construct a sequence of configurations that move from η_1 to η_2 using valid transitions of $\mathcal{L}_2^\sigma$.

Suppose first that $|P| = 2$. For each $x \in P$, let y be a child of x that is not in P. Then $|\partial T(y)| \leq 1$, so $\mathcal{L}_2(T(y))$ is ergodic. We can therefore recolour $T(y)$ so that y has a colour other than $\sigma(x)$ and $\eta(x)$ (note that such a colouring exists, since $y \notin P$ implies $y \notin \overline{\partial} T$). We can perform a similar recolouring operation for each child of x, for each $x \in P$. We then change the colour of both nodes in P to agree with η in a single move, as they form a 2-path. Finally, we recolour each subtree of each node in P to agree with η, again making use of ergodicity for boundary-size 1.

More generally, if $|P| > 2$, then recall that $r \in \overline{\partial} T$ is the root of T. Let y be the neighbour of r in P and note $y \notin \overline{\partial} T$. Let c be some colour such that $c \notin \{\sigma(z), \eta(z)\}$. Note that there is a colouring of $T(y)$ in which y has c, since $y \notin \overline{\partial} T$. Also, $\overline{\partial} T(y) = 2$. By induction, we can recolour $T(y)$ so that y has colour c, without changing the colour of z. We can then recolour all other subtrees of z so that their roots have colour c, by ergodicity over single-boundary trees. At this point z can be recoloured to $\eta(z)$. Finally, $T(y)$ and the other subtrees of z can be recoloured to η again by induction and ergodicity for single-boundary trees. $\square$

B.2 Proofs from Section 3

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

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

13

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

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

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

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

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, $\text{Var}_D^*(f) = \text{Var}_V(f)$ 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.

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

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

$$\text{gap}(\mathcal{L}_V) = \inf_f \frac{\xi_V(f, f)}{\text{Var}_V(f)} \geq \inf_f \frac{\xi_D^*(f, f)}{\text{Var}_D^*(f)} N^{-1} = \text{gap}(\mathcal{L}_D^*)N^{-1}$$

as required.

$\square$

Proof of Proposition 3.3. Given dynamics $\mathcal{L}$ on configuration space Ω and function $f : \Omega_V \to \mathbb{R}$, we will write $T(\mathcal{L}, f) = \frac{Var(f)}{\xi(f, f)}$. We recall that

$$\tau(\mathcal{L}) = \sup \{T(\mathcal{L}, f) : \pi[f] = 0\} \tag{10}$$

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

$$\mathcal{L}_B(\sigma', \eta') = \sum_{\sigma:\sigma'=\sigma|_B} \sum_{\eta:\eta'=\eta|_B} \mathcal{L}_D^*(\sigma, \eta). \tag{11}$$

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

$$f(\sigma) = g(\eta) \text{ for all } \sigma, \eta \text{ such that } \eta|_B = \sigma. \tag{12}$$

Then we will have

$$
\begin{aligned}
T(\mathcal{L}_D^*, g) &= \frac{Var_D^*(g)}{\xi_D^*(g, g)} \\
&= \frac{\sum_{\sigma,\eta\in\Omega_T} \pi_D^*(\sigma)\pi_D^*(\eta)(g(\sigma) - g(\eta))^2}{\sum_{\sigma,\eta\in\Omega_T} \pi_D^*(\sigma)P_D^*(\sigma,\eta)(g(\sigma) - g(\eta))^2} \\
&= \frac{\sum_{\sigma',\eta'\in\Omega_B} \pi_B(\sigma')\pi_B(\eta')(f(\sigma') - f(\eta'))^2}{\sum_{\sigma',\eta'\in\Omega_B} \pi_B(\sigma')P_B(\sigma',\eta')(f(\sigma') - f(\eta'))^2} \\
&= \frac{Var_B(f)}{\xi_B(f, f)} \\
&= T(\mathcal{L}_B, f)
\end{aligned}
\tag{13}
$$

where we used (2) and (11) in the third equality.

Suppose the supremum in (10) for $\mathcal{L}_D$ occurs at a function g_1. That is, $g_1 : \Omega_V \to \mathbb{R}$ satisfies $\pi[g_1] = 0$ and $\tau(\mathcal{L}_D) = T(\mathcal{L}_D, g_1)$). Then g_1 must be an eigenfunction of $\mathcal{L}_D$, so $g_1 = \mathcal{L}_D(g_1)$. Choose $\sigma, \eta \in \Omega_V$ such that $\sigma(B) = \eta(B)$; then $(\mathcal{L}_D(g_1))(\sigma) = (\mathcal{L}_D(g_1))(\eta)$ from the definition of $\mathcal{L}_D$, and hence $g_1(\sigma) = g_1(\eta)$. We can therefore define function $f_1 : \Omega_B \to \mathbb{R}$ as follows: for each $\alpha \in \Omega_B$, $f_1(\alpha)$ will be the (unique) value of $g_1(\eta)$ for all η with $\eta|_B = \alpha$. Thus f_1 and g_1 satisfy (12), so (13) implies $\tau(\mathcal{L}_D) = T(\mathcal{L}_D, g_1) = T(\mathcal{L}_B, f_1) \le \tau(\mathcal{L}_B)$.

Next suppose that the supremum in (10) for $\mathcal{L}_B$ occurs at a function f_2. Then we can define function g_2 by $g_2(\sigma) = f_2(\sigma|_B)$, from which (13) and (10) imply $\tau(\mathcal{L}_B) = T(\mathcal{L}_B, f_2) = T(\mathcal{L}_D, g_2) \le \tau(\mathcal{L}_D)$. We therefore conclude $\tau(\mathcal{L}_B) = \tau(\mathcal{L}_D)$, as required. $\qquad\square$

B.3 Proofs from Section 4

Proof of Lemma 4.1. We describe an intermediate dynamics, $\mathcal{L}_R$, the rotate dynamics. Take T to be rooted, with any external boundary adjacent to the root. For $x \in T$ we will write $T(x)$ for the subtree rooted at x and $P(x)$ for the parent of x. Informally, $\mathcal{L}_R$ performs a cyclic shift of the colours on a subtree of T at each step. That is, on each step of $\mathcal{L}_R$, a node v and integer $0 \le t < k$ are chosen uniformly. We consider the configuration in which each node of $T(v)$ has its colour incremented by t, mod k. This will be the new configuration if it forms a proper colouring, otherwise the configuration will not change.

A simple application of the comparison method of Diaconis and Saloff-Coste [3] shows that $\tau(\mathcal{L}_2^\xi) \le c\Delta\tau(\mathcal{L}_R^\xi)$ for some constant c (as any step of $\mathcal{L}_2$ can be simulated by $O(\Delta)$ steps of $\mathcal{L}_R$). We now wish to apply the comparison method to show $\tau(\mathcal{L}_R^\xi) \le n^{O(\Delta/\log\Delta)}\tau(\mathcal{L}_1^\xi)$. This amounts to showing how to simulate a step of $\mathcal{L}_R^\xi$ using steps of $\mathcal{L}_1^\xi$, then bounding the congestion of such a simulation.

We will use the following recursive simulation of a move of $\mathcal{L}_R$. In tree T, consider sibling nodes to be ordered by subtree size, with larger subtrees on the left. Suppose (σ, η) is a step of $\mathcal{L}_R$, with node v and integer t. Let $u_1, \ldots, u_\ell$ be the children of v. We implement this step as follows:

1. For each u_i with $\sigma(u_i) = \eta(v)$, from left to right, rotate $T(u_i)$ by t.
2. Change the colour of v to $\eta(v)$.
3. For each u_i with $\sigma(u_i) \ne \eta(v)$, from right to left, rotate $T(u_i)$ by t.

We note that every step of this simulation is a proper colouring, since if $\sigma(u_i) = \eta(v)$ then $\eta(u_i) \ne \sigma(v)$, as σ and η are defined by a cyclic shift over an odd number of colours. Also, simple induction shows that

at most $O(n)$ steps of the Glauber dynamics are used in the above implementation. It remains to show that each possible step of the Glauber dynamics occurs in at most $n^{O(\Delta/\log \Delta)}$ such implementations; the desired result then follows from the comparison method.

Suppose (α, ω) is a step of the Glauber dynamics, with $\alpha(z) \neq \omega(z)$ for some $z \in T$. Let us count the number of rotate operations (σ, η) this step may appear in; we do so by attempting to reconstruct σ and η. There are at most n choices for the vertex v being rotated, and t is determined by $\alpha(z)$ and $\omega(z)$. What remains to be determined is which nodes in $T(v)$ have yet been rotated at the time of step (α, ω) within the simulation of (σ, η), and which have not. It is enough to determine this for the children, ancestors, and siblings-of-ancestors of z, as one of these nodes has been rotated if and only if all nodes in its subtree have been as well.

Consider the ancestors of z. There are two possibilities for v: either it has been rotated or it has not. If $v = z$ then we are done, otherwise we now claim that we can determine which of the remaining ancestors of z up to v have been rotated. First consider the child of v that is an ancestor of z, say y. We have two cases, depending on whether or not v has been rotated.

Case 1: v has not been rotated. Then $\sigma(v) = \alpha(v)$ and $\eta(v) = \sigma(v) + t \mod k$. In this case, since we are rotating $T(y)$ before changing the colour of v, it must be that $\sigma(y) = \eta(v)$ (from the definition of $\gamma(\sigma, \eta)$). Thus if $\alpha(y) = \eta(v)$ then y has not yet been rotated, otherwise it has.

Case 2: v has been rotated. Then $\eta(v) = \alpha(v)$ and $\sigma(v) = \eta(v) - t \mod k$. In this case, since we changed the colour of v before rotating $T(y)$, it must be that $\sigma(y) \neq \eta(v)$ (from the definition of $\gamma(\sigma, \eta)$). That is, $\sigma(y)$ is the one colour not in $\{\sigma(v), \eta(v)\}$, say c. Thus if $\alpha(y) = c$ then y has not yet been rotated, otherwise it has.

In either case, given that we know whether v has been rotated, we can determine whether y has been rotated. Repeating this argument top-down from v to z, we can determine precisely which ancestors of z have been rotated. We also note that, for any ancestor y of z (including z itself), the right siblings of y are either all rotated or all not rotated, and they are all rotated if and only if $P(y)$ has been rotated.

The only remaining unknowns are the children of z and the left siblings of the ancestors of z. How many such nodes can there be? There are at most Δ children of z. Let z_i be the ancestor of z at depth i, and let k_i be the number of left siblings of z_i. We wish to find an upper bound on $\sum_i k_i$. Recall that siblings are ordered from largest subtree to smallest, so $|T(z_i)| \leq |T(z_{i-1})|/(k_i + 1)$ for all i. This implies $\prod_i (k_i + 1) \leq |T(z_1)| = n$. Since $k_i \leq \Delta - 1$ for each i, convexity implies that $\sum_i k_i$ is maximized when $\log_{\Delta-1} n$ values of k_i are equal to $\Delta - 1$. That is, $\sum_i k_i \leq (\Delta - 1) \log_{\Delta-1} n$. We conclude that there are at most $O(\Delta \log_\Delta n)$ nodes x that are children of z or left siblings of ancestors of z. Hence the number of possibilities for their subtrees being rotated or not is at most $2^{O(\Delta \log_\Delta n)} = n^{O(\Delta/\log \Delta)}$.

We conclude that (α, ω) can appear in at most $(2n)(n^{O(\Delta/\log \Delta)}) = n^{O(\Delta/\log \Delta)}$ different rotate operations, as required. $\qquad \square$

Proof of Lemma 4.5. Think of x as the root of S and consider choosing a configuration for S by choosing a colour for each node, top-down. Since $x \in \overline{\partial} T$, there are at most $k - 1$ choices for the colour of x. Given any such choice, there are $k - 1$ choices for each subsequent node, except possibly for any node in $\overline{\partial} T$ for which there will be either $k - 1$ or $k - 2$ choices. Since $|\partial T| \leq 2$, there can be at most one node beside x in $\overline{\partial} T$. Thus, for any $c \in A$, $(k-1)^{n-2}(k-2) \leq |\{\eta \in \Omega_S^\xi : \eta(x) = c\}| \leq (k-1)^{n-1}$, as long as there is at least one colouring in which x has colour c. Thus

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

as required. $\qquad \square$

B.4   A balanced mapping function

We now present the mapping function f used in the proof of Lemma 4.2, which satisfies (4). We will actually prove the existence of f in the following, equivalent, arena.

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

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

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

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

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

B.5 Full Subcase Analysis from Lemma 4.2

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 $j \ge i$, and $\eta(u_j) = \beta(u_j)$ for all $j \le i$. 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 $(k-1)^{t-1}$.

Case 2.2: $\alpha(v) \ne \eta(v)$ **and** (ω, β) **is the first change to u_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 $\alpha(v) = \omega(v)$, $\alpha(u_j) = \omega(u_j)$ for all $j \ge i$, $b_j = \beta(u_j)$, and $c_j = \beta(u_j)$ for all $j < i$. 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 $\eta(v) \ne \alpha(v) = \omega(v)$. Given $\eta(v)$, there are $k-1$ possibilities for $\eta(u_j)$ for each $j \ne i$. Note that β determines b_i, from which $\eta(v)$ determines $\eta(u_i)$. Thus the total number of possibilities for η is $(k-1)^t$.

Next consider α. Note that ω determines $\alpha(v)$ and also $\alpha(u_j)$ for all $j \ge i$. Also, β determines c_j for all $j < i$. Then (4) implies that the number of possibilities for $\alpha(u_1), \ldots, \alpha(u_t)$ is at most $\left\lceil \left(\frac{k-1}{k-2}\right)^{i-1} \right\rceil$. We conclude that for this subcase the total number of possibilities for α and η is at most

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

Case 2.3: $\alpha(v) \ne \eta(v)$ **and** (ω, β) **is the second change to u_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 node u_i we know $b_i = \omega(u_i)$ and $c_i = \beta(u_i)$.

The only effect that this has on the analysis is that now c_i is determined instead of $\alpha(u_i)$. Given c_i (instead of $\alpha(u_i)$), the factor due to (4) becomes $\lceil (\frac{k-1}{k-2})^i \rceil$. We conclude that the number of possibilities for α and η is at most

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

Case 2.4: $\alpha(v) \ne \eta(v)$ **and** (ω, β) **is the third change to u_i in** $\gamma(\alpha, \eta)$. This is the first change in Step 3 of the canonical paths description. We know that $\eta(v) = \omega(v)$, $c_j = \omega(u_j)$ for all $j < i$, and $a_j = \beta(u_j)$ for all $j > i$. Further, $c_i = \omega(u_i)$ and $\eta(u_i) = \beta(u_i)$.

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

Now consider 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 (using $\alpha(v)$) and by (4) there are at most $\lceil (\frac{k-1}{k-2})^i \rceil$ possibilities for $(\alpha(u_1), \ldots, \alpha(u_t))$. 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 for α and η is at most

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

Case 2.5: $\alpha(v) \neq \eta(v)$ **and** (ω, β) **is the fourth change to** u_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 u_i we know $\eta(u_i) = \omega(u_i)$ and $a_i = \beta(u_i)$.

The only effect that this has on the analysis is that now a_i is determined instead of c_i. This causes the factor due to (4) to become $\lceil (\frac{k-1}{k-2})^{i-1} \rceil$. We conclude that the number of possibilities for α and η is at most

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

Case 2.6: $\alpha(v) \neq \eta(v)$ **and** (ω, β) **is the fifth change to** u_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 $j > i$, and $\eta(u_j) = \beta(u_j)$ for all $j < i$. For u_i, we know $a_i = \omega(u_i)$ and $\eta(u_i) = \beta(u_i)$.

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. From these colours (plus $\alpha(v)$) the colours $\alpha(u_i), \ldots, \alpha(u_t)$ are determined. Furthermore, $\eta(u_1), \ldots, \eta(u_i)$ are determined from β.

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

B.6 A full treatment of the case $u_i \in \overline{\partial} T$

We now modify the proof of Lemma 4.2 to handle the case that there exist i such that $u_i \in \overline{\partial} T$. Note first that if $u_i \in \overline{\partial} T$, then u_i can be adjacent to only one node in ∂T, since $|\partial V_i| \leq 2$ and $v \in \partial V_i$.

We used the assumption $u_i \notin \overline{\partial} T$ when defining our canonical paths with colours a_i, b_i, c_i: this assmption allowed us to assume that there existed colourings of V_i in which the colour of u_i was a_i (or b_i, or c_i). If $u_i \in \overline{\partial} T$, it's possible that one or more of these colours will conflict with the boundary configuration, so it may not be possible to use these colours in the construction of our canonical path.

We proceed by altering our choice of blocks for the block dynamics. Let us define a small subgraph $R \subseteq T$. Informally, R is the smallest subgraph containing v whose external boundary doesn't intersect the internal boundary of T. More formally, begin by taking $R_0 = \{v\}$. Let $R_1 = R_0 \cup (\partial R_0 \cap \overline{\partial} T)$, and let $R_2 = R_1 \cup (\partial R_1 \cap \overline{\partial} T)$. Then we note that $|R_2| \leq 3$, and $\partial R_1 \cap \overline{\partial} T = \emptyset$ (since $|\partial T| \leq 2$). Set $R = R_2$.

Now think of T as being rooted at v, and consider the subtrees generated by removing R from T. In a slight abuse of notation, label these subtrees $V_1, \ldots, V_t$, with roots $u_1, \ldots, u_t$, where $t \leq 3\Delta$. Choose labels so that $|V_1| \geq \ldots \geq |V_t|$, and note that $|V_i| \leq n/2$ for all i (as the subtrees of v have this property). We will write $P(u_i)$ for the parent of u_i, which lies in R. Let D be the set of blocks $\{R, V_1, \ldots, V_t\}$. Define $\mathcal{L}_D^*$ and $\mathcal{L}_B$ as before. Note that B is either a tree with at most 3 internal nodes, or the union of two stars. For simplicity of presentation we will assume the former case, so that $B = R \cup \{u_1, \ldots, u_t\}$; the proof is very similar for the latter case.

We now construct a set of canonical paths for $\mathcal{L}_B$. Choose $\alpha, \eta \in \Omega_B$. For each u_i, define colours a_i, b_i, and c_i as before, replacing instances of v with $P(u_i)$. Our path $\gamma(\alpha, \eta)$ will now be as follows:

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

This path is precisely the same as the one used in Lemma 4.2, except that we now recolour all of R in a single step (rather than just a single node v).

We note that $\gamma(\alpha, \eta)$ uses only valid steps of $\mathcal{L}_B$. What is the congestion of this set of steps? We perform a weighted paths congestion analysis that is very similar to the one in Lemma 4.2; we only outline the differences here. We set the weight of any change to nodes in R to be 1, and the weight of a

change to node u_i to be i^{-2}. We then obtain the following variant of (5), with the term $t + |R|$ deriving from the size of B.

$$\rho_w(\omega, \beta) \le 10 \left(\frac{1}{w(\omega, \beta)} \times |\{\gamma(\alpha, \eta) \ni (\omega, \beta)\}| \times \frac{1}{(k-1)^{t+|R|} K_B[\omega \to \beta]} \right).$$

We must then count the number of choices for (α, η) given a move (ω, β) of a path. There are two main cases:

Case 1: (ω, β) **is a change to vertices in** R. We then know both $\alpha(R)$ or $\eta(R)$. We can then reconstruct $\alpha(u_1), \ldots, \alpha(u_t)$ from ω, β, and $\alpha(R)$. There are at most $(k-1)^t$ possibilities for $\eta(u_1), \ldots, \eta(u_t)$.

Case 2: (ω, β) **is a change to** u_i. We then know either $\alpha(R)$ or $\eta(R)$, and there are at most $(k-1)^{|R|}$ possibilities for the other. We then apply a subcase analysis as in the proof of Lemma 4.2 to show how to reconstruct α and η on the leaves. There will be at most $d \left(\frac{k-1}{k-2} \right)^i (k-1)^{t-1}$ ways to do this for some constant d (as before), for a total of $d \left(\frac{k-1}{k-2} \right)^i (k-1)^{t+|R|-1}$ possibilities.

We conclude that, in either case, there are at most $d \left(\frac{k-1}{k-2} \right)^i (k-1)^{t+|R|-1}$ possibilities for η and α for some i and some constant d. Plugging this bound into our weighted canonical paths equation, we end up with a final bound of

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

for some constant c, which matches (8) from the proof of Lemma 4.2 up to a constant factor. This implies the desired result of Lemma 4.2.