The Glauber Dynamics for Colourings of Bounded Degree Trees

B. Lucier[1], M. Molloy[2], and Y. Peres[3]

[1] Dept of Computer Science,University of Toronto, blucier@cs.toronto.edu.
[2] Dept of Computer Science,University of Toronto, molloy@cs.toronto.edu.
[3] Microsoft Research, peres@microsoft.com.

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 that exploits the differing relaxation times of blocks.

1 Introduction

The Glauber dynamics is a Markov chain over configurations of spin systems on graphs, of which k-colourings is a special case. Such chains have generated a great deal of interest 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 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, and is 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. Nevertheless, people have studied the Glauber dynamics on trees as a means of understanding its performance 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 $k \geq \Delta + 2$.

* This extended abstract presents two pieces of work. The first [13] proves the case $k \geq 4$ (amongst other things); it has been submitted to a journal. The second covers the case $k = 3$; a full version is in progress.

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 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 give 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 bound is asymptotically tight, matching the lower bound up to a constant factor in the degree.

Theorem 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 describe the difficulties that occur when $k = o(\Delta/\log\Delta)$. If $k \geq \Delta+2$ then no vertex will ever be *frozen*; i.e. there will always be at least one colour that it can switch to. (It also 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 frozen variables. From this perspective, $k \geq C\Delta/\log\Delta$ for $C > 1$ is another natural threshold: if the neighbours of a vertex are assigned independently random colours then we expect that the vertex will not be frozen. But if $k < (1-\epsilon)\Delta/\log\Delta$, then even in the steady state distribution most degree Δ vertices on a tree will be frozen.

If the children of a vertex u change colours enough times, then eventually u will become unfrozen and change colours. This allows vertices to unfreeze, level by level, much like in the level dynamics of [7]. 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 $\Delta = O(1)$: the running time is a high degree polynomial rather than superpolynomial. For balanced trees, it is very helpful that there are only $O(\log n)$ levels. For taller trees, a more complicated approach is necessary.

The proofs of our main theorems use 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 this 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 fixed leaves are well-studied. When every leaf is fixed, Martinelli, Sinclair and Weitz [14] prove 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 must bound the number of fixed leaves. Theorem 1 extends to show:

Theorem 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 $n^{O(1+b+\Delta/k \log \Delta)}$.*

Due to space constraints, some proofs are omitted from this extended abstract and may be found in the full versions of the papers.

Remark 1. Our arguments can be extended to other instances of the Glauber dynamics, e.g. 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 a (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 trees (see eg. [1]).

In many applications we are interested in the discrete analog of the Glauber dynamics. We then think of $K[\sigma \to \eta]$ as the probability of moving from colouring σ to colouring η, scaled by a factor of n. The mixing time for the discrete chain is precisely n times the mixing time for the corresponding continuous process (see eg. [1]), so our bounds on mixing time apply to the discrete setting.

We will additionally be interested in a variant of the Glauber dynamics, the 2-*path 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

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

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

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

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

We define the *spectral gap* of $\mathcal{L}$, $\mathrm{Gap}(\mathcal{L})$, to be the second-largest eigenvalue of $-\mathcal{L}$. The *relaxation time* of $\mathcal{L}$, denoted $\tau(\mathcal{L})$, is defined as the inverse of the spectral gap. We will use the following standard bound (see eg. [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 and $\overline{\partial} T$ to mean the exterior and interior boundaries 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.

The following simple Lemma analyzes the ergodicity of the Glauber dynamics and 2-path Glauber dynamics on trees.

Lemma 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 $\sigma \in \Omega$. If additionally $k = 3$ and $|\partial T| \leq 2$, the 2-path Glauber dynamics is also ergodic over Ω_T^σ for all $\sigma \in \Omega$.*

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 $1 \le i \le r$ and $\sigma \in \Omega$, let $\mathcal{L}^{\sigma}_{V_i}$ be the generator for the Glauber dynamics (or 2-path Glauber dynamics) restricted to V_i with boundary configuration σ. In other words, colours can change only for nodes in V_i.

Suppose that $\mathcal{L}^{\sigma}_{V_i}$ is ergodic for each i and σ. Let $\pi^{\sigma}_{V_i}$ denote the stationary distribution of $\mathcal{L}^{\sigma}_{V_i}$. For each i, define $g_i := \inf_{\sigma \in \Omega} \mathrm{Gap}(\mathcal{L}^{\sigma}_{V_i})$, the minimum spectral gap for $\mathcal{L}^{\sigma}_{V_i}$ over all choices of boundary configurations. The *block dynamics* is a continuous-time Markov process with generator $\mathcal{L}_D$ defined by

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

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

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

Proposition 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. The main result for this section is the following variant of Proposition 1:

Proposition 2. $\tau(\mathcal{L}_{\mathcal{V}}) \le \tau(\mathcal{L}^*_D) \times \sup_{x \in V} |\{i : x \in V_i\}|$.

The proof of Proposition 2 is a simple modification to the proof of Proposition 1 [12]. It is worth noting the difference between Proposition 2 and the original block dynamics, Proposition 1. In the original version, the block dynamics Markov process can be thought of as having a poisson clock of rate g for each block, where g is the minimum over all g_i. In other words, each block is chosen with the same rate, that being the worst case over all blocks. On the other hand, in the modified version each block is chosen with the rate corresponding to that block. The original version yields a simpler Markov process, but a looser bound on the gap of the original process. In particular, applying the original block dynamics to our main result yields a mixing time of $n^{O(1+\Delta/k)}$, while the modified block dynamics tightens the bound to $n^{O(1+\Delta/k \log \Delta)}$ (see Remark 4).

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

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

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

Proposition 3. $\tau(\mathcal{L}_D^*) = \tau(\mathcal{L}_B)$.

4 An Upper Bound for General Trees

We now begin our proof of Theorem 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 at a vertex v, then take each subtree rooted at a child of v as a block (and v itself as a block of size 1). A nice property of this decomposition is that each subtree has at most one boundary node, adjacent to its root. In this case there will be h levels of recursion in the induction, where h is the height of tree T, and we will obtain a bound of the form c^h, where $c = c(\Delta, k)$ is the mixing time for an instance of the block dynamics. This method works for complete trees (and indeed was used by Berger et al. [1]) since they have logarithmic height. However, the height of a general tree could be much greater, leading to a super-polynomial bound.

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

There are difficulties with the above approach that must be overcome. First, a subtree T may have multiple boundary nodes, which complicates the behaviour of the block dynamics. We therefore make our choice of x carefully, so that boundaries are of size at most 2. Second, for non-complete trees we might have blocks of vastly differing sizes, which makes a tight analysis of the block dynamics more difficult. We therefore use the weighted version of the block dynamics.

In this section we describe our choice of blocks for the block dynamics. We then show that the upper bound of Theorem 1 holds, given a bound on the relaxation time of the block dynamics. The details of analyzing the block dynamics are encapsulated in Lemma 3, 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 $|\partial T| \leq 2$, Lemma 1 implies 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

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

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

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

First let us show how (3) implies Theorem 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]). 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 $n^{O(1+\Delta/k \log \Delta)}$. Theorem 1 then follows from Lemma 2 below.

Lemma 2. *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. 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. Details are given in the full version.

We now turn to proving (3). The following Lemma will be our main tool.

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

Lemma 4. *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. Suppose $|\partial T| \leq 1$. 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 $|V_1| \geq |V_2| \geq \ldots \geq |V_t|$. Since $|\partial T| \leq 1$, we have $|\partial V_i| \leq 2$ for all i. By Lemma 3, $\tau_T \leq ci^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}$ for some $1 \leq i \leq t$.

If $i \geq 2$, we get $\tau_{V_i} \leq \tau_2(|V_i|) \leq \tau_2(\lfloor n/i \rfloor)$, since the V_i are given by increasing size. Thus $\tau_T \leq ci^2 \left(\frac{k-1}{k-2}\right)^i \tau_2\left(\lfloor n/i \rfloor\right)$ 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.

Proposition 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 Lemma 4, 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 3 as in Lemma 4 and obtain the desired result.

Now suppose that x is not on the boundary path for T. Consider T to be rooted at some $r \in \overline{\partial} T$. Let y be the least ancestor of x that lies on the boundary path. Consider $D_y = \{\{y\}, V_1, \ldots, V_t\}$. Since x separates T into subtrees of size at most $\lfloor n/2 \rfloor$, in particular the subtree containing y must have size at most $\lfloor n/2 \rfloor$. This implies that the subtree separated by y that contains x must contain at least $\lfloor n/2 \rfloor$ nodes, and is therefore V_1, the largest subtree separated by y. Also, $|\overline{\partial} V_i| \leq 2$ for all i, since y is on the boundary path for T. Lemma 3 implies

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

for some i. If $i > 1$ then we obtain the desired result since $|V_i| \leq \lfloor n/i \rfloor$. If $i = 1$, then since $|V_1| < n$ and $|\overline{\partial} V_1| = 1$ (by our choice of y), Lemma 4 implies

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

We have now derived (3), completing the proof of Theorem 1.

4.1 Proof of Lemma 3

We now proceed with the proof of Lemma 3, which bounds the relaxation time on a tree with respect to the relaxation times for subtrees. Our approach is to use a canonical paths argument to bound the behaviour of the block dynamics. Indeed, there is a simple canonical path to move between configurations σ and 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 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: it only implies a mixing time of $n^{O(\Delta)}$. We therefore apply the following sequence of improvements to the above approach.

1. We explicitly describe an intermediate configuration for the neighbours of v, in order to balance congestion over all start and end configurations. This improves the bound on the mixing time to $n^{O(\log \Delta + \log k + \Delta/k)}$.
2. Our path shifts between 3 different intermediate configurations to maximize the dependency on the start and end configurations at each step. This improves our bound to $n^{O(\log \Delta + \Delta/k)}$.
3. We apply the weighted block dynamics, to differentiate between large and small subtrees. We always change configurations of blocks in order of subtree size. This improves our bound to $n^{O(\log \Delta + \Delta/k \log \Delta)}$. See Remark 4.
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 3.

The Block Dynamics Recall the conditions of Lemma 3. 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). Due to space limitations, we prove Lemma 3 under the assumption that $u_i \notin \overline{\partial} T$ for all i. The (simple) extension to remove this assumption is discussed at the conclusion of the section; see Remark 2.

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 of $\mathcal{L}_D^*$. Since no vertex lies in more than one block, Proposition 2 implies $\tau_T^\sigma \leq \tau_D^\sigma$.

Next recall the definition of graph B and dynamics $\mathcal{L}_B$ from Proposition 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, $\tau(\mathcal{L}_D^\sigma) = \tau(\mathcal{L}_B^\sigma)$. 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 bounds the transition probabilities of $\mathcal{L}_B^\sigma$.

Lemma 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 $\pi_S^\xi[\omega : \omega(x) = c] \geq 1/k$.*

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

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). We therefore introduce a more complex path that uses three passes. We now define the colours used in the intermediate configurations of this path.

If $\alpha(v) \neq \eta(v)$ 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, with occurrences of $\eta(v)$ 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 configuration in which u_i has colour a_i. We define b_i 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 to allow v to change from $\alpha(v)$ to $\eta(v)$. 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 $1 \leq i \leq t$, writing $\mathbf{x}$ for $(x_1, \ldots, x_t)$,

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

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 $\alpha(u_1), \ldots, \alpha(u_t)$. Such an f is guaranteed to exist; see Lucier and Molloy [13] for a construction.

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 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.

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 $\rho_w := \sup_{\omega,\beta} \rho_w(\omega, \beta)$. 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). The standard canonical path bound sets $w(\omega, \beta) = 1$ 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 (see Remark 3).

Our result follows by bounding $\rho_w(\omega, \beta)$. Uniformity of π implies

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

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 $K_B[\omega \to \beta] = \inf_{\sigma\in\Omega} \operatorname{gap}(\mathcal{L}^{\sigma}_{\{v\}})\pi^{\omega}_{\{v\}}[\phi : \phi(v) = \beta(v)]$. But note that $\operatorname{gap}(\mathcal{L}^{\sigma}_{\{v\}}) = 1$ for all boundary conditions, and $\pi^{\omega}_{\{v\}}$ is the uniform distribution over a set of at most $k-1$ colours.We conclude

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

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

Consider the possibilities for η. Configuration β determines $\eta(v)$, and there are $(k-1)^t$ choices for η given $\eta(v)$ (consider choosing the colours for $u_1, \ldots, u_t$, which cannot be $\eta(v)$). Now consider 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 1 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 omitted due to space constraints.

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?

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. 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.

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

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

This concludes our case analysis. Cases 1 and 2 (and the fact that $\tau_{V_t} \ge 1$) imply $\rho_w \le \max_{1\le i\le t} 180 i^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}$. Applying the canonical paths bound and Proposition 2 we conclude $\tau_T^\sigma \le \tau_D^\sigma \le 180 \max_{1\le i\le t} i^2 \left(\frac{k-1}{k-2}\right)^i \tau_{V_i}$ as required.

Remark 2. 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 the full version of this paper. We used the assumption to guarantee that no leaf of B was adjacent to the boundary of T. We modify our selection of blocks to maintain this property: we replace block $\{v\}$ with a block $R \subseteq T$ that contains v and any neighbouring nodes in $\overline{\partial} T$. Our new set of blocks D will contain R and all subtrees separated by R. Then B will no longer be a star, but rather a tree or forest with few internal nodes. We then bound the relaxation time of $\mathcal{L}_B$ as before, extending our set of canonical paths in the natural way. The congestion analysis for this set of paths is similar to the original, and we obtain the same result up to a constant factor.

Remark 3. We note the effect of using the weighted canonical paths bound. If we had used the standard canonical paths bound, then we would replace the factor of i^2 in Lemma 3 by the maximum length of a path, which is $5\Delta + 1$. 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 $n^{O(1+\Delta/k \log \Delta)}$.

Remark 4. We also note the effect of using the weighted block dynamics. If we had applied Proposition 1 instead of Proposition 2, the bound in (7) would become $K_B[\omega \to \beta] \ge (k\tau)^{-1}$, where $\tau = \max_i \tau_{V_i}$. This would lead to a bound of $\tau_T^\sigma \le ct^2 \left(\frac{k-1}{k-2}\right)^t \max_{1\le i\le t} \tau_{V_i}$ for Lemma 3. With this modified Lemma, the bound in (3) would become $\tau_2(n) \le ct^2 \left(\frac{k-1}{k-2}\right)^t \tau_2(\lceil n/2 \rceil)$, leading to a mixing time bound of $n^{O(1+\Delta/k)}$, which is weaker than $n^{O(1+\Delta/k \log \Delta)}$.

5 Open Problems

Our results raise questions about the Glauber dynamics on planar graphs of bounded degree. 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. The same is true for any Δ that grows with n [13]. But for $\Delta = O(1)$, Theorem 1 shows that no trees require $k > 3$. Is there a constant K such that for every $k \ge K$ 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 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 $2k-2$.

Acknowledgements

We thank Nayantara Bhatnagar, Jian Ding, 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 (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).