

Local Uniformity Properties for Glauber Dynamics on Graph Colorings

Thomas P. Hayes*

March 5, 2009

Abstract

We investigate some local properties which hold with high probability for randomly selected colorings of a fixed graph with no short cycles. We show that, when the number of colors is sufficiently large relative to the maximum degree of the graph, the aforementioned properties can be used to establish improved bounds on the mixing time for the single-site dynamics, a Markov chain for sampling colorings uniformly at random. For a large class of graphs, this approach yields the most efficient known algorithms for sampling random colorings.

1 Introduction

This paper studies random k -colorings of an input graph $G = (V, E)$ with maximum degree Δ . A (proper) k -coloring is a function $f : V \rightarrow [k] = \{1, \dots, k\}$ which satisfies, for every $\{v, w\} \in E$, $f(v) \neq f(w)$. We restrict our attention to the setting $k > \Delta$. With this many colors, a trivial greedy algorithm can construct such a coloring in linear time.

We have two broad foci in this paper. Our first goal is to establish a number of “local” properties of graph colorings which hold with very high probability for uniformly random colorings, ideally for triangle-free graphs. Thus, we refer to these properties as *local uniformity* properties. We believe these results are interesting in their own right, they also serve as a starting point for establishing our second set of results.

*University of New Mexico Department of Computer Science. e-mail: hayes@cs.unm.edu. Partially supported by an NSF Postdoctoral Fellowship, the Toyota Technological Institute at Chicago, and the University of Chicago.

Our second focus is to analyze how quickly a coloring generated by the Glauber dynamics acquires the local uniformity properties. The Glauber dynamics (formally defined in Section 2.2) is a simple Markov chain which recolors a randomly chosen vertex in each step and whose stationary distribution is a uniformly random coloring. It is well-studied in statistical physics as a model of how physical systems reach equilibrium [12], and in computer science as a tool for approximately counting and randomly sampling colorings [8].

The prototypical example of a local uniformity property, studied earlier by M. Dyer and A. Frieze [3], is the number of distinct colors assigned to the neighbors $N(v)$ of a vertex, v . We define the set of *available colors* at a vertex v , under coloring f , as

$$A(f, v) := [k] \setminus f(N(v)).$$

If the colors of the neighbors of v are independently colored with uniformly distributed color choices, then the expected number of *available colors* for v is $\approx k(1 - e^{-d(v)/k})$ where $d(v)$ is the degree of v .

Our first result is an easy (relative to the remainder of the paper) proof which (roughly speaking) says that for triangle-free graphs with high probability a random coloring has at least $\approx k(1 - e^{-d(v)/k})$ colors available. See [7] for a similar proof and its implications for proving rapid mixing of the Glauber dynamics.

Theorem 1. *Let $G = (V, E)$ be triangle-free, let $v \in V$, let $\varepsilon > 0$ and let $k \geq \Delta + 2$. Let X be a uniformly random proper k -coloring of G . Then*

$$\Pr(|A(X, v)| < k(e^{-d(v)/k} - \varepsilon)) < (d(v) + 1)e^{-\varepsilon^2 k/50}.$$

Under a few additional conditions, perhaps most significantly, when the graph has girth at least six and $k/\Delta > 1 + \delta$ for constant $\delta > 0$, we establish concentration of the number of available colors by proving that the color choices on the neighborhood are roughly independent.

Theorem 2. *Let $1 > \delta, \varepsilon > 0$ and let $\Delta_0 = \Delta_0(\delta, \varepsilon)$. Let $G = (V, E)$ have girth ≥ 6 and $\Delta > \Delta_0$. Let $v \in V$ and let $k \geq (1 + \delta)\Delta$. Let X be a uniformly random proper k -coloring of G . Then with probability at least $1 - \exp(-\varepsilon^4 \delta^2 \Delta 10^{-8})$*

$$||A(X, v)| - ke^{-d(v)/k}| \leq \varepsilon k.$$

We then show how to establish analogous results for colorings generated by $O(n)$ or $O(n \log \Delta)$ steps of the Glauber dynamics. Such local uniformity properties can be very useful for proving upper bounds on the mixing time for the Glauber dynamics. (See

Section 2.3 for a formal definition of mixing time.) For general graphs, the best bound is Vigoda's result [13] establishing polynomial mixing time for the Glauber dynamics whenever $k > 11\Delta/6$. Dyer and Frieze [3] improved this for graphs with $\Delta = \Omega(\log n)$ and girth $g = \Omega(\log \Delta)$ by establishing $O(n \log n)$ mixing time when $k > 1.763... \Delta$. The key to their proof is showing an analog of Theorem 1 for the Glauber dynamics under the same conditions on Δ and g . This was improved by Molloy [10] to $k > 1.489... \Delta$ by establishing stronger local uniformity properties, an upper bound on available colors as considered in Theorem 2 is an example. In [5] we extended Molloy's approach to girth $g \geq 6$. However, that approach does not extend to $k/\Delta > 1 + \delta$ for all $\delta > 0$.

In this paper we establish a strong set of local uniformity properties. The following theorem is an example of the very general statement we obtain for any coloring obtained after $O(n \log \Delta)$ steps of the Glauber dynamics. For "nice" initial colorings, only $O(n)$ steps of the dynamics are necessary where nice is quantified using the following notion.

Definition 3. Let $G = (V, E)$ be a graph of maximum degree Δ , and let $C > 0$. Let $f : V \rightarrow [k]$ be a coloring, let $c \in [k]$, and let $v \in V$. We say f is C -heavy for color c at v if $|f^{-1}(c) \cap B_2(v)| \geq C\Delta$ or $|f^{-1}(c) \cap N(v)| \geq C\Delta/\log \Delta$.

Fix a vertex v , and a subset $S \subset N(v)$. For every color c , $i \geq 0$, and coloring X , let $S_{c,i}(X)$ denote the set of $w \in S$ such that exactly i neighbors z of w , excluding v , satisfy $X(z) = c$. We call this the subset of S which is " i times blocked for c ."

Theorem 4. Let $0 < \delta, \varepsilon < 1/20$, let $\Delta_0 = \Delta_0(\varepsilon, \delta)$, let $C = C(\varepsilon, \delta)$, and let $k \geq (1 + \delta)\Delta$. Let $I = [t_0, t_1]$ be a time interval with $t_0 \geq Cn \log \Delta$. Let $G = (V, E)$ have girth ≥ 7 and $\Delta > \Delta_0$. Let $(X_t)_{t \geq 0}$ be the continuous-time (or discrete-time) Glauber dynamics on G with arbitrary X_0 . Let $v \in V$ and $c \in [k]$.

$$\Pr \left((\exists t \in I) \left| |A(X_t, v)| - ke^{-d(v)/k} \right| > \varepsilon k \right) \leq \left(1 + \frac{t_1 - t_0}{n} \right) e^{-\Delta/C}. \quad (1)$$

Moreover, for every $S \subset N(v)$, $c_1 \neq c_2 \in [k]$ and non-negative integers i_1, i_2 ,

$$\begin{aligned} \Pr \left((\exists t \in I) \left| |S_{c_1, i_1}(X_t) \cap S_{c_2, i_2}(X_t)| - \frac{1}{i_1! i_2!} \sum_{w \in S} e^{-2d(w)/k} \left(\frac{d(w)}{k} \right)^{i_1 + i_2} \right| > \varepsilon \Delta \right) \\ \leq \left(1 + \frac{t_1 - t_0}{n} \right) e^{-\Delta/C}. \quad (2) \end{aligned}$$

$$\begin{aligned}
\text{and } \Pr \left((\exists t \in I) \left| |S_{c_1, i_1}(X_t)| - \frac{1}{i_1!} \sum_{w \in S} e^{-d(w)/k} \left(\frac{d(w)}{k} \right)^{i_1} \right| > \varepsilon \Delta \right) \\
\leq \left(1 + \frac{t_1 - t_0}{n} \right) e^{-\Delta/C}. \quad (3)
\end{aligned}$$

Moreover, if X_0 is not $(4/\delta)$ -heavy for any color c and vertex $w \in B_R(v)$, where $R = R(\varepsilon, \delta)$, then the same conclusions hold for $t_0 \geq CRn$.

The above result is used in the proofs of $O(n \log n)$ mixing time of the Glauber dynamics when: (i) $k > 1.489 \dots \Delta$, girth $g \geq 7$ and Δ is a sufficiently large constant [4]; (ii) for all $\varepsilon > 0$, $k > (1 + \varepsilon)\Delta$, girth $g > 10$ and $\Delta = \Omega(\log n)$ [6].

Our basic approach is the method of conditional independence: conditioned on the right information, a set of random variables of interest become fully independent. This is also the basic approach taken in [5]. However, when proving our local uniformity properties for the Glauber dynamics, the picture is a bit more complicated. In this setting, in order to get the desired conditional independence, we are forced to modify the dynamics to completely prevent the flow of color information between vertices of interest.

Essentially, the modification is to work with a graph G^* , which is G modified so that all the edges within a certain distance of a vertex v are oriented along the shortest path towards v ; if the dynamics only consider the colors on *in-neighbors* of the vertex being updated, then no color information can be propagated between neighbors of v , because no directed paths exist from one to another. A major component of our proof, then, is an argument comparing this dynamics to the original Glauber dynamics. In Section 4, we are able to show that, over periods of $O(n)$ steps, started from the same initial conditions, the two dynamics do not diverge too much.

In [5], I had stated the concentration bound (1) for girth ≥ 6 , and given a proof sketch under the additional assumption that $k/\Delta \geq \sqrt{2}$. However, that proof technique, based on a recurrence of Molloy [10], breaks down irreparably when $k/\Delta < 1.05 \dots$, because the fixed point of the recurrence in question undergoes a bifurcation, and there is no longer convergence to the desired fixed point. In order to get a proof which works for all values of $k/\Delta \geq 1 + \varepsilon$, the current paper makes use of a different recurrence relation, which was introduced by Jonasson [9] for his analysis of colorings of trees.

We now give a brief outline of the remainder of the present paper. Section 2 introduces some essential notation and concepts, some of which may be familiar. Our uniformity properties are defined and motivated in more detail in Section 2.5. In Section 3, we establish our local uniformity properties for random colorings. This is quite a bit easier than for colorings obtained by the Glauber dynamics, which helps to illustrate

most of the main steps in our argument. Section 4 introduces the modified graphs G^* , and proves the main comparison theorems relating the dynamics on G to the dynamics on G^* . In Section 5, we establish the local relations which are the main technical tool in establishing the desired uniformity properties. The remaining two sections of the paper are devoted to deducing the uniformity properties as a consequence.

2 Preliminaries

In this section, we introduce just enough concepts to formally state our main results.

2.1 Notation

For a graph $G = (V, E)$, positive integer r and $v \in V$, we define the *sphere and ball of radius r centered at v* as follows:

$$\begin{aligned} S_r(v) &= \{w \in V : \text{dist}(v, w) = r\} \\ B_r(v) &= \{w \in V : \text{dist}(v, w) \leq r\}, \end{aligned}$$

2.2 Glauber dynamics

Let $G = (V, E)$ be a fixed graph, and let $k > 0$ be an integer. By a *k -coloring of G* , we mean a function $f : V \rightarrow [k]$ which satisfies, for every $\{v, w\} \in E$, $f(v) \neq f(w)$. That is, adjacent vertices are assigned distinct colors.

By *Glauber dynamics*, we will mean the following Markov chain on the space of k -colorings of G , which is more explicitly the heat-bath version of the Glauber dynamics. **Glauber dynamics.** Given a k -coloring X_t , the coloring X_{t+1} is defined by the following procedure:

1. Choose $v \in V$ uniformly at random.
2. Choose $c \in [k] \setminus X_t(N(v))$ uniformly at random, where $N(v)$ denotes the set of neighbors of v .
3. Define

$$X_{t+1}(w) = \begin{cases} c & \text{if } w = v \\ X_t(w) & \text{otherwise.} \end{cases}$$

2.3 Mixing Time

We will use $\|\cdot\|$ to denote the total-variation norm: for distributions μ, ν , $\|\mu - \nu\| = \frac{1}{2} \sum_x |\mu(x) - \nu(x)|$. We define mixing time, τ_{mix} as follows:

$$\tau_{\text{mix}} = \max_{X_0} \min \left\{ t : \|X_t - \pi\| \leq \frac{1}{2e} \right\}.$$

The constant $\frac{1}{2e}$ is chosen purely for algebraic convenience; this choice ensures (see [1]) that $\|X_t - \pi\| \leq \varepsilon$ for every $t \geq \lceil \ln \varepsilon^{-1} \rceil \tau_{\text{mix}}$.

2.4 Continuous time

We're going to state our main results for the continuous-time version of the Glauber dynamics, in which each vertex is updated according to an independent Poisson process, with rate $1/n$. This will simplify many of our proofs. The following observation can be used to generically convert high-probability results proved in continuous time into high-probability results in discrete time.

Fact 5. *Let (X_t) be any discrete Markov chain on state space Ω , and let $(Y_t), t \geq 0$ be the corresponding continuous-time chain. Then, for any property $P \subset \Omega$ and positive integer t ,*

$$\Pr(X_t \notin P) \leq e\sqrt{t} \Pr(Y_t \notin P).$$

Proof. Since the number of updates of the chain Y in $[0, t]$ follows the Poisson distribution, it follows that

$$\begin{aligned} \Pr(Y_t \notin P) &= \sum_{i \geq 0} e^{-t} \frac{t^i}{i!} \Pr(X_i \notin P) \\ &\geq e^{-t} \frac{t^t}{t!} \Pr(X_t \notin P) \\ &\geq \frac{1}{e\sqrt{t}} \Pr(X_t \notin P). \end{aligned}$$

where the last step follows by Stirling's formula. □

The above general fact suffices for our purposes when $\Delta = \Omega(\log n)$, but not for Glauber dynamics on graphs of constant degree. For the latter case, instead of focusing on specific times t , we will establish guarantees for larger intervals of times, and apply the following result.

Fact 6. Let (X_t) be any discrete Markov chain on state space Ω , and let $(Y_t), t \geq 0$ be the corresponding continuous-time chain. Let $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{Z}$ be the “time correspondence” function, so that for $t \geq 0$, we have $Y_t = X_{f(t)}$. Then, for any property $P \subset \Omega$ and time intervals I and J ,

$$\Pr((\exists t \in I) X_t \notin P) \leq \Pr((\exists t \in J) Y_t \notin P) + \Pr(I \not\subset f(J)).$$

In particular, when $I = [t_0, t_1]$ and $J = [t_0 + C\sqrt{t_0}, t_1 - C\sqrt{t_1}]$, then

$$\Pr((\exists t \in I) X_t \notin P) \leq \Pr((\exists t \in J) Y_t \notin P) + \exp(-C^2/2). \quad (4)$$

We omit the (easy) proof.

2.5 Uniformity Properties

Let $G = (V, E)$ be a graph, and let $f : V \rightarrow [k]$ be a k -coloring of G . Recall our definition of the set of *available colors* at a vertex v , under f , as

$$A(f, v) := [k] \setminus f(N(v)).$$

Obviously, these sets play a role of particular interest in both versions of the Glauber dynamics, since the new color for the chosen vertex is always chosen from its available colors.

Our first class of uniformity properties are concentration inequalities on the values of $|A(X_t, v)|$, where $(X_t)_{t \geq 0}$ is an evolution of the Glauber dynamics from an arbitrary initial k -coloring, X_0 .

Suppose that f is a k -coloring, v is a vertex, and c is a color. Our second class of uniformity properties are concentration inequalities on the following parameter,

$$P(f, v, c) = \sum_{w \sim v} \frac{\mathbf{1}\{c \in A_v^*(f, w)\}}{|A_v^*(f, w)|},$$

where

$$A_v^*(f, w) := [k] \setminus f(N(w) \setminus \{v\}),$$

denotes the number of available colors for w under f , ignoring which color is assigned to v . Assuming G is triangle-free, $P(f, v, c)$ equals the expected number of occurrences of a color c in the neighborhood of v , if $N(v)$ were recolored uniformly at random, with each new color being chosen from its available colors under f , and ignoring the color of v . Note that, for every coloring f and vertex v ,

$$\sum_c P(f, v, c) = d(v),$$

where $d(v) = |N(v)|$ is the degree of v in G . Our main goal in this regard will be to prove that for colorings X_t obtained via the Glauber dynamics, with high probability,

$$P(f, v, c) \approx d(v)/k$$

for every $c \in [k]$. (The reason the color of v is ignored in the definition of $P(f, v, c)$, is that otherwise, the above approximate equality would only hold for every $c \in [k] \setminus \{f(v)\}$.)

3 Random colorings

In this section, we show how to derive the uniformity properties we are interested in, for uniformly random colorings. We will be able to re-use most of the ideas when proving our uniformity properties for colorings obtained from the Glauber dynamics, in Sections 5 and 7.

3.1 Lower bound on available colors

In this section, we will prove Theorem 1, our lower bound on $|A|$ for uniformly random colorings. The ideas introduced here will be built upon in Sections 5, 6 and 7, to prove analogous results for colorings found by the Glauber dynamics.

We first prove the following technical lemma, which establishes a “local relation” satisfied by $|A|$ with high probability.

Lemma 7. *Let $G = (V, E)$ be triangle-free, let k be at least the chromatic number of G , and let $v \in V$ be fixed. Let X be a uniformly random proper k -coloring of G . Then*

$$\Pr \left(|A(X, v)| \leq k \prod_{w \sim v} \left(1 - \frac{1}{|A(X, w)|} \right)^{|A(X, w)|/k} - a \right) \leq e^{-a^2/2k}. \quad (5)$$

In addition, condition on the restriction of X to the complement of $N(v)$; let $\mathcal{F}$ denote this conditional information. We have:

$$\Pr(|A(X, v)| \leq \mathbf{E}(|A(X, v)| \mid \mathcal{F}) - a) \leq e^{-a^2/2k} \quad (6)$$

$$\Pr(|A(X, v)| \geq \mathbf{E}(|A(X, v)| \mid \mathcal{F}) + a) \leq e^{-a^2/2k}. \quad (7)$$

Proof. Conditioned on $\mathcal{F}$, the colors $X(w)$ for $w \in N(v)$ become fully independent random variables, since G is triangle-free. Hence,

$$\mathbf{E}(|A(X, v)| \mid \mathcal{F}) = \sum_{c \in [k]} \Pr(c \in A(X, v) \mid \mathcal{F}) = \sum_{c \in [k]} \prod_{w \sim v} \Pr(X(w) \neq c \mid \mathcal{F}).$$

Now, since each $X(w)$ is uniformly distributed over $A(X, w)$, we have

$$\mathbf{E}(|A(X, v)| \mid \mathcal{F}) = \sum_{c \in [k]} \prod_{\substack{w \sim v: \\ c \in A(X, w)}} \left(1 - \frac{1}{|A(X, w)|}\right).$$

By the arithmetic-geometric mean inequality, this implies

$$\begin{aligned} \mathbf{E}(|A(X, v)| \mid \mathcal{F}) &\geq k \prod_{c \in [k]} \prod_{\substack{w \sim v: \\ c \in A(X, w)}} \left(1 - \frac{1}{|A(X, w)|}\right)^{1/k} \\ &= k \prod_{w \sim v} \prod_{c \in A(X, w)} \left(1 - \frac{1}{|A(X, w)|}\right)^{1/k} \\ &= k \prod_{w \sim v} \left(1 - \frac{1}{|A(X, w)|}\right)^{|A(X, w)|/k}. \end{aligned} \quad (8)$$

Finally, since $|A(X, v)| = \sum_{c \in [k]} \mathbf{1}\{c \in A(X, v)\}$, and these indicator variables are negatively associated, conditioned on $\mathcal{F}$, Chernoff's bound (see Dubhashi and Ranjan [2]) implies

$$\Pr(|A(X, v)| \leq \mathbf{E}(|A(X, v)| \mid \mathcal{F}) - a) \leq e^{-a^2/2k} \quad (9)$$

$$\Pr(|A(X, v)| \geq \mathbf{E}(|A(X, v)| \mid \mathcal{F}) + a) \leq e^{-a^2/2k}. \quad (10)$$

This proves (6) and (7), and by combining (6) with (8) we have (5). $\square$

The local relation stated in Lemma 7 is the main ingredient in the high-probability lower bound on $|A|$ stated in Theorem 1, which we now prove.

Proof of Theorem 1. Assume $d(v) \geq 1$, and $\varepsilon < d(v)/k < 1$; if either condition fails, the event in question has probability zero. We may further assume $\varepsilon k > 25$, since otherwise the upper bound on the probability exceeds 1.

Apply Lemma 7 to v and to each of its $d(v)$ neighbors, with $a = \varepsilon k/5$. By a union bound, the local relation holds for all of these with the right probability. Now, since $k \geq \Delta + 2$, all $z \in V$ satisfy $|A(X, z)| \geq 2$. For each $w \in N(v)$, the local relation implies

$$\begin{aligned} |A(X, w)| &\geq k \prod_{z \sim w} \left(1 - \frac{1}{|A(X, z)|}\right)^{|A(X, z)|/k} - \frac{\varepsilon k}{5} \\ &\geq k(1/4)^{d(w)/k} - k/5 \\ &\geq k/4 - k/5 \\ &= k/20. \end{aligned}$$

The local relation for v implies

$$\begin{aligned}
|A(X, v)| &\geq k \prod_{w \sim v} \left(1 - \frac{1}{|A(X, w)|}\right)^{|A(X, w)|/k} - \varepsilon k/5 \\
&\geq k \left(1 - \frac{20}{k}\right)^{d(v)/20} - \varepsilon k/5 \\
&\geq k \left(\frac{1 - 20/k}{e}\right)^{d(v)/k} - \varepsilon k/5 \\
&\geq k e^{-d(v)/k} (1 - 20/k) - \varepsilon k/5 \\
&\geq k e^{-d(v)/k} - 20 - \varepsilon k/5 \\
&\geq k e^{-d(v)/k} - \varepsilon k,
\end{aligned}$$

which completes the proof. Note the way the local relation allows the very weak lower bound $|A| \geq 2$ to be first strengthened to $|A| \geq k/20$, and then, bootstrapping, to the final bound $|A| \geq k(e^{-d(v)/k} - \varepsilon)$. $\square$

We note that Theorem 2 shows that, at least for graphs of girth ≥ 6 , the lower bound of Theorem 1 is asymptotically tight. Subsequently, Molloy and Lau [11] extended this tightness result to graphs of girth ≥ 5 , although their proof is probably only valid for $k > 1.45\Delta$ and for $\Delta = \Omega(\log n)$. On the other hand, it is not hard to see that, for the complete bipartite graph $K_{\Delta, \Delta}$ most k -colorings have $|A| \approx \beta k$, where β is substantially larger than $e^{-\Delta/k}$.

3.2 Further uniformity properties

We now prove several additional uniformity properties for random colorings under slightly stronger assumptions on the girth of the graph.

In Section 2.5 we introduced the following measure of “color bias,”

$$P(f, v, c) = \sum_{w \sim v} \frac{\mathbf{1}\{c \in A_v^*(f, w)\}}{|A_v^*(f, w)|}.$$

(Recall that by definition, $A_v^*(f, w) := [k] \setminus \{f(z) \mid z \in N(w) \setminus \{v\}\}$ is the set of available colors for w under f , ignoring the color of v .) We now show that, for random colorings, almost surely no vertex has much bias toward any color.

We will need the following observation.

Observation 8. Let X, Y be two non-negative random variables, and let $\min Y > 0$. Let $0 \leq \theta < \min Y/2$, and suppose $p \geq \Pr(|Y - \mathbf{E}(Y)| \geq \theta)$. Then with probability at least $1 - p$,

$$\mathbf{E}\left(\frac{X}{Y}\right) \in \frac{\mathbf{E}(X)}{Y \pm 2\theta} \pm \frac{p \max X}{\min Y}.$$

Proof. We will prove that

$$\mathbf{E}\left(\frac{X}{Y}\right) \in \frac{\mathbf{E}(X)}{\mathbf{E}(Y) \pm \theta} \pm \frac{p \max X}{\min Y},$$

from which the claimed probabilistic result follows immediately.

We first consider the lower bound on $\mathbf{E}\left(\frac{X}{Y}\right)$. Note that if $\mathbf{E}(X) \leq p \max X$ then the lower bound is trivial. When $\mathbf{E}(X) > p \max X$, we have

$$\begin{aligned} \mathbf{E}\left(\frac{X}{Y}\right) &\geq \frac{p \max X}{\max Y} + \frac{\mathbf{E}(X) - p \max X}{\mathbf{E}(Y) + \theta} \\ &\geq \frac{\mathbf{E}(X)}{\mathbf{E}(Y) + \theta} - \frac{p \max X}{\mathbf{E}(Y) + \theta} \end{aligned}$$

Now we consider the upper bound. If $\mathbf{E}(X) < p \max X$ we have:

$$\mathbf{E}\left(\frac{X}{Y}\right) \leq \frac{\mathbf{E}(X)}{\min Y} \leq \frac{p \max X}{\min Y} \leq \frac{\mathbf{E}(X)}{\mathbf{E}(Y) - \theta} + \frac{p \max X}{\min Y}$$

In the remaining case we have:

$$\mathbf{E}\left(\frac{X}{Y}\right) \leq \frac{p \max X}{\min Y} + \frac{\mathbf{E}(X) - p \max X}{\mathbf{E}(Y) - \theta} \leq \frac{\mathbf{E}(X)}{\mathbf{E}(Y) - \theta} + \frac{p \max X}{\min Y}$$

□

Lemma 9. Let $1 > \delta, \varepsilon > 0$ and let $\Delta_0 = \Delta_0(\delta, \varepsilon)$. Let $G = (V, E)$ have girth ≥ 6 and $\Delta > \Delta_0$. Let $v \in V$ and let $k \geq (1 + \delta)\Delta$. Let X be a uniformly random proper k -coloring of G . Then with probability at least $1 - \exp(-\varepsilon^2 \delta^2 \Delta 10^{-6})$,

$$\left| P(X, v, c) - \sum_{w \in N(v)} \frac{\exp(-P(X, w, c))}{|A(X, w)|} \right| \leq \varepsilon.$$

Proof. Let $S_2(v)$ denote the set of vertices at distance exactly 2 from v . Let $\mathcal{G}$ denote the restriction of X to the complement of $S_2(v)$. Note that, since the girth is ≥ 6 , there is no edge joining any pair of vertices in $S_2(v)$, and hence, conditioned on $\mathcal{G}$, the colors assigned by X to $S_2(v)$ are fully independent.

By linearity of expectation,

$$\begin{aligned} \mathbf{E}(P(X, v, c) \mid \mathcal{G}) &= \sum_{w \in N(v)} \mathbf{E} \left(\frac{\mathbf{1}\{c \in A_v^*(X, w)\}}{|A_v^*(X, w)|} \mid \mathcal{G} \right) \\ &\in \sum_{w \in N(v)} \mathbf{E} \left(\frac{\mathbf{1}\{c \in A_v^*(X, w)\}}{|A(X, w)| \pm 1} \mid \mathcal{G} \right) \end{aligned}$$

where $N_w^* = N(w) \setminus \{v\}$.

Note, $\prod_{z \in N_w^*} \mathbf{1}\{X(z) \neq c\} \in \{0, 1\}$ and $|A(X, w)| \geq \delta\Delta$. By (6) and (7) in Lemma 7 we have that except with probability $\leq e^{-\varepsilon^2\Delta/500000}$:

$$|A(X, v)| \in \mathbf{E}(|A(X, v)| \mid \mathcal{F}) \pm k\varepsilon/100$$

Hence, by Observation 8, except with probability $\leq \Delta^3 e^{-\varepsilon^2\Delta/500000}$:

$$\begin{aligned} \mathbf{E}(P(X, v, c) \mid \mathcal{G}) &\in \sum_{w \in N(v)} \left(\frac{\mathbf{Pr}(c \in A_v^*(X, w) \mid \mathcal{G})}{|A(X, w)| \pm (\varepsilon k/50 + 1)} \pm \frac{\Delta^2 e^{-\varepsilon^2\Delta/500000}}{\delta k} \right) \\ &\subset (1 \pm \varepsilon/15) \sum_{w \in N(v)} \frac{\mathbf{Pr}(c \in A_v^*(X, w) \mid \mathcal{G})}{|A(X, w)|} \end{aligned} \tag{11}$$

where the last line uses some kind of lower bound on $|A(X, w)|$ (high probability or absolute), and a similar lower bound on $\mathbf{Pr}(c \in A_v^*(X, w) \mid \mathcal{G})$ based on the lower bound on $|A(X)|$.

We now approximate $\mathbf{Pr}(c \in A_v^*(X, w) \mid \mathcal{G})$ from the expression above.

$$\begin{aligned}
\Pr(c \in A_v^*(X, w) \mid \mathcal{G}) &= \mathbf{E} \left(\prod_{z \in N_w^*} (1 - \mathbf{1}\{X(z) = c\}) \mid \mathcal{G} \right) \\
&= \prod_{z \in N_w^*} \mathbf{E}((1 - \mathbf{1}\{X(z) = c\}) \mid \mathcal{G}) \text{ by conditional independence} \\
&\in (1 \pm O(1/k^2))^\Delta \prod_{z \in N_w^*} (\exp(-\Pr(X(z) = c \mid \mathcal{G}))) \\
&= (1 \pm O(1/k^2))^\Delta \exp \left(- \sum_{z \in N_w^*} \Pr(X(z) = c \mid \mathcal{G}) \right)
\end{aligned}$$

Plugging in (11), we have, except with probability $\leq \Delta^5 e^{-\varepsilon^2 \Delta / 5000}$:

$$\begin{aligned}
&\mathbf{E}(P(X, v, c) \mid \mathcal{G}) \\
&\in (1 \pm \varepsilon/14) \sum_{w \in N(v)} \frac{\exp(-\sum_{z \in N_w^*} \Pr(X(z) = c \mid \mathcal{G}))}{|A(X, w)|} \\
&= (1 \pm \varepsilon/14) \sum_{w \in N(v)} \frac{\exp\left(-\sum_{z \in N_w^*} \frac{\mathbf{1}\{c \in A(X, z)\}}{|A(X, z)|}\right)}{|A(X, w)|} \\
&\subseteq (1 \pm \varepsilon/13) \sum_{w \in N(v)} \frac{\exp\left(-\sum_{z \in N_w^*} \frac{\mathbf{1}\{c \in A(X, z)\}}{|A_w^*(X, z)|}\right)}{|A(X, w)|} \\
&= (1 \pm \varepsilon/13) \left(\sum_{w \in N(v): X(w)=c} \frac{1}{|A(X, w)|} + \sum_{w \in N(v): X(w) \neq c} \frac{\exp\left(-\sum_{z \in N_w^*} \frac{\mathbf{1}\{c \in A_w^*(X, z)\}}{|A_w^*(X, z)|}\right)}{|A(X, w)|} \right) \\
&\subseteq (1 \pm \varepsilon/12) \left(\sum_{w \in N(v): X(w)=c} \frac{1}{|A(X, w)|} + \sum_{w \in N(v): X(w) \neq c} \frac{\exp(-P(X, w, c))}{|A(X, w)|} \right).
\end{aligned}$$

Now, since there are always at least $\delta\Delta$ available colors, $|X^{-1}(c) \cap N(v)| < \varepsilon\delta\Delta/2$ with probability $\geq 1 - \exp(-\varepsilon^2\delta^2\Delta/10)$ for sufficiently large Δ , which leaves us with the

approximation

$$\begin{aligned}
& \mathbf{E}(P(X, v, c) \mid \mathcal{G}) \\
& \in (1 \pm \varepsilon/12) \left(\sum_{w \in N(v)} \frac{\exp(-P(X, w, c))}{|A(X, w)|} \pm \frac{\varepsilon}{2} \right) \\
& \subseteq \left(\sum_{w \in N(v)} \frac{\exp(-P(X, w, c))}{|A(X, w)|} \right) \pm (\varepsilon/2 + \varepsilon^2/24 + \frac{\varepsilon}{12} \frac{\Delta}{A_{\min}}) \\
& \subseteq \left(\sum_{w \in N(v)} \frac{\exp(-P(X, w, c))}{|A(X, w)|} \right) \pm 7\varepsilon/8 \quad \text{when } A_{\min} \geq \Delta/4 \quad (12)
\end{aligned}$$

which holds with probability $\geq 1 - \Delta^5 \exp(-\varepsilon^2 \delta^2 \Delta / 5000)$ over $\mathcal{G}$.

Finally, we observe that $P(X, v, c)$ is concentrated near its expectation. Since the random variables $\frac{\mathbf{1}_{\{c \in A_v^*(X, w)\}}}{|A_v^*(X, w)|}$ are conditionally independent given $\mathcal{G}$, and take values in $[0, 1/\delta\Delta]$, Chernoff's bound implies that

$$P(X, v, c) \in \mathbf{E}(P(X, v, c) \mid \mathcal{G}) \pm \varepsilon/8$$

with probability at least $1 - 2 \exp(-\varepsilon^2 \delta^2 \Delta / 128)$. Plugging into (12) completes the proof. $\square$

Clearly, the assignment $P(X, v, c) = d(v)/k$ satisfies the approximate local relation from the previous lemma. To prove our next theorem, we show that this solution is unique, ignoring small perturbations. The argument is an adaptation of the proof of a closely related result due to Jonasson [9, Theorem 2.1].

We will rely on the following technical lemma, which is very closely related to the calculations in the last page or so of the proof of Theorem 2.1 of [9].

Lemma 10. *For $R > 1$ and $y \geq 0$,*

$$y - y^R < \frac{\ln R}{e}.$$

Proof. Define $f(y, R) := e(y - y^R) - \ln R$. Differentiating with respect to R , we have

$$\frac{\partial f(y, R)}{\partial R} = -e \ln(y) y^R - \frac{1}{R}.$$

Differentiating with respect to y shows that $\frac{\partial f(y, R)}{\partial R}$ has a maximum of 0, attained at $y = \exp(-1/R)$, for every $R > 1$. Hence, for every $y \geq 0$, $f(y, R)$ is strictly decreasing in R for $R \geq 1$. Since $f(y, 1) = 0$, the desired conclusion follows. $\square$

Next we show how to convert our local relation (Lemma 9) for P into a concentration result.

Theorem 11. *Let $0 < \varepsilon < 1/10$, let $k \geq e^{10\varepsilon}\Delta$, and let $R \geq \frac{2}{\varepsilon} \ln\left(\frac{1}{\varepsilon}\right)$ be a positive integer. Suppose X is a coloring which satisfies the following relations for all vertices z at distance $\leq R$ from v .*

$$A(X, z) \geq k/e$$

$$P(X, z, c) \in [e^{-\varepsilon^2}, e^{\varepsilon^2}] \sum_{u \in N(z)} \frac{\exp(-P(X, u, c))}{|A(X, u)|}.$$

Then

$$\left| P(X, v, c) - \frac{d(v)}{k} \right| \leq \varepsilon.$$

Proof. Fix two colors $c, c' \in [k]$. For each $0 \leq i \leq R$,

$$\alpha_i := \max_z |P(X, z, c) - P(X, z, c')|$$

over all vertices z at distance $\leq i$ from v .

By hypothesis, we know, for all z at distance $i \leq R$ from v ,

$$\begin{aligned} P(X, z, c') &\leq e^{\varepsilon^2} \sum_{u \in N(z)} \frac{\exp(-P(X, u, c'))}{|A(X, u)|} && \text{by hypothesis} \\ &\leq e^{\varepsilon^2} \sum_{u \in N(z)} \frac{\exp(\alpha_{i+1} - P(X, u, c))}{|A(X, u)|} && \text{by def. of } \alpha_i \\ &\leq \exp(\alpha_{i+1} + 2\varepsilon^2) P(X, z, c) && \text{by hypothesis.} \end{aligned}$$

Applying our hypothesised relation again, this implies, for all w at distance $i \leq R-1$

from v , that

$$\begin{aligned}
& P(X, w, c') - P(X, w, c) \\
& \leq \sum_{z \sim w} \frac{e^{\varepsilon^2} \exp(-P(X, z, c)) - e^{-\varepsilon^2} \exp(-P(X, z, c'))}{|A(X, z)|} \\
& \leq \frac{e}{k} \sum_{z \sim w} e^{\varepsilon^2} \exp(-P(X, z, c)) - e^{-\varepsilon^2} \exp(-P(X, z, c')) \\
& \leq \frac{e}{k} \sum_{z \sim w} e^{\varepsilon^2} \exp(-P(X, z, c)) - e^{-\varepsilon^2} \exp(-\exp(\alpha_{i+2} + 2\varepsilon^2)P(X, z, c)) \\
& \leq \frac{e}{k} \sum_{z \sim w} \max_{0 \leq y \leq 1} e^{\varepsilon^2} y - e^{-\varepsilon^2} y^{\exp(\alpha_{i+2} + 2\varepsilon^2)} \\
& \leq \frac{e}{k} \sum_{z \sim w} e^{\varepsilon^2} - e^{-\varepsilon^2} \left(1 - \frac{\alpha_{i+2} + 2\varepsilon^2}{e}\right) \quad \text{by Lemma 10} \\
& = \frac{d(w)}{k} e^{-\varepsilon^2} \left(\alpha_{i+2} + 2\varepsilon^2 + e^{1+2\varepsilon^2} - e\right) \\
& \leq e^{-10\varepsilon - \varepsilon^2} (\alpha_{i+2} + 10\varepsilon^2) \quad \text{since } 0 < \varepsilon < 1/10 \text{ and } k \geq e^{10\varepsilon} \Delta \\
& \leq \max\{e^{-\varepsilon} \alpha_{i+2}, \varepsilon\} \quad \text{by a case analysis.}
\end{aligned}$$

Since $\alpha_R \leq 1$ and $R \geq \frac{2}{\varepsilon} \ln\left(\frac{1}{\varepsilon}\right)$, it follows by induction that $\alpha_0 \leq \varepsilon$. Since $\sum_{c \in [k]} P(X, v, c) = d(v)$, it follows that for every $c \in [k]$, $P(X, v, c) \in [d(v)/k \pm \varepsilon]$. $\square$

Corollary 12. *Let $1/10 > \delta, \varepsilon > 0$ and let $\Delta_0 = \Delta_0(\delta, \varepsilon)$. Let $G = (V, E)$ have girth ≥ 6 and $\Delta > \Delta_0$. Let $k \geq (1 + \delta)\Delta$. Let X be a uniformly random proper k -coloring of G . Let $v \in V$ and $c \in [k]$. Then with probability at least $1 - \exp(-\varepsilon^2 \delta^2 \Delta 10^{-6})$ the following holds:*

$$\left| P(X, v, c) - \frac{d(v)}{k} \right| \leq \varepsilon.$$

Proof. Let $\eta > 0$ be a parameter to be determined. Theorem 1 gives us the first condition for Theorem 11. Now, the high-probability event from the conclusion of Lemma 9 is of the form

$$\left| P(X, v, c) - \sum_{w \in N(v)} \frac{\exp(-P(X, w, c))}{|A(X, w)|} \right| \leq \eta.$$

To convert this additive approximation into a multiplicative approximation, observe that because of our lower bound $A_{\min} \geq k/e$, it follows immediately from the definitions that

for each time t , vertex v , and color c ,

$$\frac{k}{e} \leq A(X_t, v) \leq k \text{ and } 0 \leq P(X_t, v, c) \leq e$$

Hence,

$$\sum_{w \in N(v)} \frac{\exp(-P(X, w, c))}{|A(X, w)|} \in d(w) \left[\frac{e^{-e}}{k}, \frac{e^0}{k/e} \right] \subseteq \frac{d(w)}{k} [e^{-e}, e]$$

into Theorem 11. □

In order to derive additional uniformity results from Theorem 11, we will apply the following lemma, which basically says that any sum of independent indicator variables approaches a Poisson distribution, as the maximum event probability tends to zero. We are not aware of a reference, so we include a proof for completeness.

Lemma 13. *Let $\varepsilon > 0$. Suppose $\eta_1, \eta_2, \dots$ are independent $\{0, 1\}$ -valued random variables, and for each j , denote $\mu_j := \mathbf{E}(\eta_j)$. Suppose further that $\mu = \sum_j \mu_j$, and $\hat{\mu} = \sum_j \frac{\mu_j}{1-\mu_j}$ are both finite. Let P denote the distribution of $\sum_j \eta_j$. Let Q denote the Poisson distribution with mean $\hat{\mu} = \sum_j \frac{\mu_j}{1-\mu_j}$. Then the total variation distance between P and Q satisfies*

$$\|P - Q\|_{\text{TV}} \leq e^{\hat{\mu} - \mu} - 1.$$

Proof. For convenience, let us denote $P_i := P(\{i\})$ for $i \geq 0$. Note that

$$P_0 = \prod_j (1 - \mu_j) \in [e^{-\hat{\mu}}, e^{-\mu}]$$

Next, note that, for every $i \geq 0$,

$$\hat{\mu}^i \geq i! \frac{P_i}{P_0},$$

which follows by counting the summands in the expansion of $\hat{\mu}^i$ which are the product of i distinct terms $\mu_j/(1 - \mu_j)$. Hence,

$$P_i \leq P_0 \hat{\mu}^i / i! = P_0 e^{\hat{\mu}} Q(\{i\}).$$

It follows that, for every event A ,

$$P(A) \leq P_0 e^{\hat{\mu}} Q(A).$$

From this, it follows generically that the total variation distance between P and Q is at most $P_0 e^{\hat{\mu}} - 1$. Since $P_0 \leq e^{-\mu}$, the stated result follows. □

Corollary 14. Let $G = (V, E)$ be a triangle-free graph, let $v \in V$, and let X be a uniformly random proper k -coloring of G . Let $\mathcal{F}$ denote the restriction of X to $V \setminus N(v)$. For each $i \geq 0$ and $c \in [k] \setminus \{X(v)\}$,

$$\left| \Pr(|X^{-1}(c) \cap N(v)| = i \mid \mathcal{F}) - \exp(-P(X, v, c)) \left(\frac{P(X, v, c)^i}{i!} \right) \right| \leq \frac{100P(X, v, c)}{A_{\min}},$$

where $A_{\min} = \min_{w \in N(v)} |A(X, w)|$.

Proof. Note that, conditioned on $\mathcal{F}$, the colors assigned by X to the neighbors of v are fully independent. Now apply Lemma 13 to the indicator variables for the events $X(w) = c$, $w \in N(v)$. Note that this gives us $\mu = P(X, v, c)$ and $\max \mu_j = 1/A_{\min}$. The conclusion is that, for all $i \geq 0$,

$$\left| \Pr(|X^{-1}(c) \cap N(v)| = i \mid \mathcal{F}) - \exp(-P(X, v, c)) \left(\frac{P(X, v, c)^i}{i!} \right) \right| \leq \frac{100P(X, v, c)}{A_{\min}}.$$

□

We are now ready to prove our upper bound on the number of available colors in a random coloring.

Proof of Theorem 2. Let $\mathcal{F}$ denote the restriction of X to $V \setminus N(v)$. Now, conditioned on $\mathcal{F}$, the expected number of available colors for v equals the sum of the conditional probabilities that v has exactly zero neighbors colored c , for $c \in [k]$. Hence

$$\begin{aligned} \mathbf{E}(|A(X, v)| \mid \mathcal{F}) &= 1 + \sum_{c \neq X(v)} \Pr(c \in A(X, v) \mid \mathcal{F}) \\ &= 1 + \sum_{c \neq X(v)} \Pr(|X^{-1}(c) \cap N(v)| = 0 \mid \mathcal{F}) \\ &\in 1 + \sum_{c \neq X(v)} \left(e^{-P(X, v, c)} \pm \frac{100P(X, v, c)}{A_{\min}} \right) && \text{by Corollary 14} \\ \text{w.h.p. } &\subset 1 + \left(\sum_{c \neq X(v)} \exp\left(-\frac{d(v)}{k} \pm \frac{\varepsilon}{3}\right) \right) \pm \frac{100d(v)}{A_{\min}} && \text{(see below)} \\ &\subseteq ke^{-d(v)/k} \pm \frac{\varepsilon k}{2} && \text{for } \Delta \geq \Delta_0. \end{aligned}$$

In the penultimate step, the “w.h.p.” means that the containment relation holds with probability at least $1 - \exp(-\varepsilon^4 \delta^2 \Delta 10^{-8})$ over the choice of $\mathcal{F}$. This follows by plugging in the conclusions of Theorem 1 and Lemma 9 into Theorem 11.

Recall (6) and (7) from Lemma 7, which imply that, with probability at least $1 - 2e^{-\varepsilon^2 k/8}$,

$$|A(X, v)| \in \mathbf{E}(|A(X, v)| \mid \mathcal{F}) \pm \frac{\varepsilon k}{2},$$

which concludes the proof. $\square$

4 Modifying the Graph

For the next part of our argument, it will be helpful to study the Glauber dynamics on a modified graph G^* . Fix a vertex v . We obtain G^* from G by replacing all the undirected edges in the ball of radius 2 or 3 around v with edges directed towards v . Assuming the girth of G is at least 5 or 7, this specifies a unique direction for each such edge. The remaining edges of G are considered to be bi-directed in G^* .

Throughout this section, we will adopt the convention that “neighbor” means “in-neighbor.” For a vertex w , we denote $N(w) := \{u \mid (u, w) \in G^*\}$. For $S \subset V$, we denote $\delta(S) = \{(u, w) \in E \mid w \in S\}$. We will denote $A(X, w) = [k] \setminus X(N(w))$. Note that, the way we have defined G^* above, for a vertex $w \in N(v)$, $A(X, w) = A_v^*(X, w)$ as studied earlier.

The heat-bath Glauber dynamics on G^* is defined by the following update rule. Recall that $\Omega = [k]^V$ includes all colorings, not just proper colorings. Given $X_t \in \Omega$, we define $X_{t+1} \in \Omega$ by the following procedure:

1. Choose $u \in V$ uniformly at random.
2. Choose $c \in [k] \setminus X_t(N(u))$ uniformly at random, where $N(u)$ denotes the set of in-neighbors of u .
3. Define

$$X_{t+1}(w) = \begin{cases} c & \text{if } w = u \\ X_t(w) & \text{otherwise.} \end{cases}$$

Note that, unlike on G , the Glauber dynamics on G^* is not reversible, and includes moves from proper colorings of G to improper colorings. In this regard, it is a rather unusual object of study. However, the Glauber dynamics on G^* has two important points in its favor. The first is that, because G^* has no directed paths between neighbors of v

(for radius 2) or neighbors of neighbors of v (for radius 3), the colors assigned to these vertices are conditionally independent given the evolution of the dynamics outside of the ball. We will exploit this property in various ways in the next section. The second desirable property is that, over short time periods (enough for $O(n)$ vertex updates), the Glauber dynamics on G^* and on G can be coupled with few disagreements.

More precisely, consider the natural coupling of the dynamics on G^* with the dynamics on G , in which the update times and vertices are the same for both chains, and the color choices are paired maximally at each update. Then we have the following main result, which, loosely speaking, says that for the rough behavior of “local properties,” such as the number of available colors at a vertex, there is almost no difference between the Glauber dynamics on G and on G^* , at least over relatively short time periods.

Notation 15. For two directed graphs G, G^* , we will denote by $G \oplus G^*$ the symmetric difference of their edgesets. By our usual choice of G^* , this will be a directed tree of depth 2 (or 3) rooted at v and oriented toward its leaves.

Notation 16. Let $X, X' : V \rightarrow [k]$ be two colorings. The *disagreement set*, $X \oplus X'$ denotes the set of vertices v for which $X(v) \neq X'(v)$.

We will use the following tail inequalities for a sum of independent exponentially distributed random variables.

Lemma 17. Let $X_1, \dots, X_s$ be independent random variables, where each X_i is exponentially distributed with mean μ_i . Let $X = \sum_{i=1}^s X_i$ and $\mu = \sum_{i=1}^s \mu_i$. Then, for every $0 < \varepsilon < 1$,

$$\Pr(X \leq (1 - \varepsilon)\mu) \leq \min_{\lambda > 0} e^{\lambda(1-\varepsilon)\mu} \prod_{i=1}^s \frac{1}{1 + \lambda\mu_i}.$$

The proof is a straightforward application of the exponential moment method. I will uncomment it later.

We will make frequent use of the following corollaries.

Corollary 18. Let $X_1, \dots, X_s$ be i.i.d. exponentially distributed random variables, each with mean μ_1 . Let $X = \sum_{i=1}^s X_i$ and $\mu = s\mu_1$. Then, for every $0 < \varepsilon < 1$,

$$\Pr(X \leq (1 - \varepsilon)\mu) \leq ((1 - \varepsilon)e^\varepsilon)^s \leq e^{-\varepsilon^2 s/2}.$$

Proof. The first inequality follows from Lemma 17 upon setting $\lambda = \varepsilon/((1 - \varepsilon)\mu_1)$. The second inequality follows from the Taylor expansion of $\ln(1 - \varepsilon)$, which, exponentiated, becomes

$$1 - \varepsilon = \exp\left(-\sum_{i \geq 1} \frac{\varepsilon^i}{i}\right).$$

□

Corollary 19. *Let $X_1, \dots, X_s$ be independent exponentially distributed random variables, where each X_i has mean μ_i . Let $X = \sum_{i=1}^s X_i$, $\mu = \sum_{i=1}^s \mu_i$ and $V = \sum_{i=1}^s \mu_i^2$. Then, for every $0 < \delta < 1$,*

$$\Pr\left(X \leq \mu - \sqrt{2V \ln(1/\delta)}\right) \leq \delta.$$

Proof. Follows from Lemma 17 upon setting $\varepsilon\mu = \sqrt{2V \ln(1/\delta)}$ and $\lambda = \varepsilon\mu/V$. The inequality

$$1 + x \geq \exp(x - x^2/2),$$

which holds for $x \geq 0$, is used to upper bound $1/(1 + \lambda\mu_i)$ in the first step. □

Theorem 20. *For every $\varepsilon > 0, \delta > 0, C_1 > 0$, there exists $\Delta_0 > 0, C > 0$ such that the following holds. Suppose $G = (V, E)$ has girth $g \geq 5$ and maximum degree $\Delta \geq \Delta_0$. Let $v \in V$, and let G^* be G with two layers of edges around v oriented toward v . Let $k \geq (1 + \delta)\Delta$. Let X_0 be an arbitrary coloring, and let (X_t, X_t^*) denote a maximal coupling of the Glauber dynamics on G with the Glauber dynamics on G^* , starting from $X_0 = X_0^*$. Then,*

$$\Pr(\forall t \leq C_1 n, |(X_t \oplus X_t^*) \cap N(v)| \leq \varepsilon\Delta) \geq 1 - \exp(-\Delta/C).$$

Remark 21. Our proof, below, gives a dependence of C on ε, δ, C_1 which looks something like

$$C = \frac{\exp(4C_1/\delta)}{\varepsilon^2}.$$

Proof of Theorem 20. We begin by defining several “bad” events, $\mathcal{B}_i$, and show that each has low probability. Let $R = \varepsilon'\Delta$ be a positive integer, where $\varepsilon' > 0$ is a parameter to be specified later. Let $D = \cup_{t \leq C_1 n} X_t \oplus X_t^*$ denote the total set of disagreeing vertices. Let $\mathcal{B}_1$ denote the event that $D \not\subseteq B_R(v)$. Let $\mathcal{B}_2$ denote the event that $|D| \geq C'\Delta$, where C' is a suitably large constant to be fixed later.

To bound the probability of $\mathcal{B}_1$, we use a standard paths of disagreement argument. Since any disagreement outside of $B_R(v)$ must have arisen via a path of disagreements starting within $B_2(v)$, there are at most Δ^R such paths, and each involves a particular sequence of $R - 2$ disagreements being formed (in order) within the first $C_1 n$ time units. Since the rate at which a disagreement percolates from a particular site to a particular

neighbor is bounded by $1/(nA_{\min})$, it follows that

$$\begin{aligned}\Pr(\mathcal{B}_1) &\leq \Delta^R e^{-C_1/A_{\min}} \sum_{j \geq R-2} \frac{(C_1/A_{\min})^j}{j!} \\ &\leq 2\Delta^R e^{-C_1/A_{\min}} \frac{(C_1/A_{\min})^{R-2}}{(R-2)!} \\ &\leq 2\Delta^2 \left(\frac{C_1 \Delta e}{A_{\min}(R-2)} \right)^{R-2}\end{aligned}$$

Note that, since $A_{\min} \geq \delta\Delta$ and $R = \varepsilon'\Delta$, this bound is at most $\exp(-\varepsilon''\Delta)$ for some constant ε'' , for all sufficiently large Δ .

To bound the probability of $\mathcal{B}_2$, we consider the waiting time τ_i for the i 'th disagreement. Note that, for all i , even conditioned on the history $\tau_1, \dots, \tau_i$, τ_{i+1} is exponentially distributed with mean $\frac{nA_{\min}}{|G \oplus G^*| + i\Delta} \geq \frac{n\delta}{\Delta + i}$. Also note that $\mathcal{B}_2$ equals the event that $\sum_{i=1}^{C'\Delta} \tau_i \leq C_1 n$. Applying Corollary 19 yields

$$\Pr(\mathcal{B}_2) \leq \exp(-(\mu - C_1 n)^2 / 2V)$$

where $\mu = \sum_{i=1}^{C'\Delta} \frac{n\delta}{\Delta + i} \approx n\delta \ln(C' + 1)$ and $V = \sum_{i=1}^{C'\Delta} \frac{n^2 \delta^2}{(\Delta + i)^2} \approx \frac{n\delta C'}{\Delta(C' + 1)}$, assuming $\mu \geq C_1 n$. (The approximations above come from replacing the sums by integrals.) We ensure that the condition $\mu \geq C_1 n$ is satisfied, by choosing $C' \geq \exp(C_1/\delta)$. Thus we conclude that

$$\Pr(\mathcal{B}_2) \leq \exp\left(-\frac{(\ln(C' + 1) - C_1/\delta)^2 \Delta}{2C'^2}\right) = O(\exp(-\Delta/C)).$$

Finally, assuming neither $\mathcal{B}_1$ and $\mathcal{B}_2$ occurs, we show that with high probability, no vertex in $B_R(v)$ is likely to have more than $\varepsilon\Delta$ disagreements in its neighborhood. Fix a vertex $w \in B_R(v)$. Consider the rate at which neighbors of w become disagreements. This can occur due to their neighbors being disagreements, which because $\mathcal{B}_1$ nor $\mathcal{B}_2$ occurs, never occurs with rate faster than $\frac{(C'+1)\Delta}{nA_{\min}}$. (The “+1” comes from the possibility that w itself is a disagreement—note that since the girth is ≥ 5 , w is the only vertex with more than one neighbor in $N(w)$.) Disagreements on $N(w)$ may also be caused by incident edges in $G \oplus G^*$, however, since this has maximum in-degree 1, these occur with rate at most $\Delta/nA_{\min}$. Thus the total rate of disagreement formation in $N(w)$ is never more than $(C' + 2)\Delta/nA_{\min}$. Let $C'' = C' + 2$.

Hence the probability that more than $\varepsilon\Delta$ neighbors of w become disagreements before

time $C_1 n$ is bounded by a Poisson tail:

$$\begin{aligned} \Pr(|D \cap N(w)| \geq \varepsilon \Delta \text{ and not } \mathcal{B}_1 \text{ and not } \mathcal{B}_2) &\leq \exp(-C_1 C''' \Delta / A_{\min}) \sum_{j \geq \varepsilon \Delta} \frac{(C_1 C''' \Delta / A_{\min})^j}{j!} \\ &\leq 2 \left(\frac{C_1 C''' \Delta e}{\varepsilon \Delta A_{\min}} \right)^{\varepsilon \Delta} \\ &\leq 2 \left(\frac{C'''}{\varepsilon \delta \Delta} \right)^{\varepsilon \Delta} \end{aligned}$$

since $A_{\min} \geq \delta \Delta$. The desired conclusion now follows by a union bound over the $\Delta^R = \Delta^{\varepsilon \Delta/2}$ vertices in $B_R(v)$. $\square$

In the case when girth ≥ 7 , and three layers of edges around v have been oriented towards v , the conclusion of Theorem 20 holds under the same hypotheses, except the hidden constants will be slightly worse, as we now prove.

Theorem 22. *For every $\varepsilon > 0, \delta > 0, C_1 > 0$, there exists $\Delta_0 > 0, C > 0$ such that the following holds. Suppose $G = (V, E)$ has girth $g \geq 7$ and maximum degree $\Delta \geq \Delta_0$. Let $v \in V$, and let G^* be G with three layers of edges around v oriented toward v . Let $k \geq (1 + \delta)\Delta$. Let X_0 be an arbitrary coloring, and let (X_t, X_t^*) denote a maximal coupling of the Glauber dynamics on G with the Glauber dynamics on G^* , starting from $X_0 = X_0^*$. Then,*

$$\Pr(\forall w \in B_1(v), \forall t \leq C_1 n, |(X_t \oplus X_t^*) \cap N(w)| \leq \varepsilon \Delta) \geq 1 - \exp(-\Delta/C).$$

Proof. The general proof outline is the same. The main difference is that now $|G \oplus G^*| = \Theta(\Delta^3)$ rather than Δ^2 . Consequently, the disagreement set will grow to size $\Theta(\Delta^2)$ rather than $\Theta(\Delta)$. In this case, more work must be done to argue that the disagreements intersect $S_2(w)$ in a small set.

Analogously to the proof of Theorem 20, we define “bad” sets, $\mathcal{B}_i$. As before, $\mathcal{B}_1$ denotes the event that the cumulative disagreement set D escapes a ball $B_R(v)$, where $R = \varepsilon' \Delta$. We define $\mathcal{B}_2$ to be the event that $|D| > C' \Delta^2$.

We omit the proof of the bound on $\Pr(\mathcal{B}_1)$, which is exactly analogous to that in Theorem 20, except with $R - 3$ in place of $R - 2$.

The proof of the bound on $\Pr(\mathcal{B}_2)$, is analogous to that in Theorem 20, except with $|G \oplus G^*| = \Delta^3$ rather than Δ^2 , and $C' \Delta^2$ in place of $C' \Delta$. In this case, the waiting time τ_i for the i 'th disagreement. satisfies

$$\mathbf{E}(\tau_{i+1} \mid \tau_1, \dots, \tau_i) = \frac{n A_{\min}}{|G \oplus G^*| + i \Delta} \geq \frac{n \delta}{\Delta^2 + i},$$

and $\mathcal{B}_2$ equals the event that $\sum_{i=1}^{C'\Delta^2} \tau_i \leq C_1 n$. In this case, Corollary 19 gives an even stronger bound on $\Pr(\mathcal{B}_2)$ than before, since there are more independent variables being summed.

From here on, let us assume that $\mathcal{B}_1$ and $\mathcal{B}_2$ do not occur. This time, we break our argument into two steps. Fix a vertex $w \in B_R(v)$. In the first step, we argue that $|D \cap S_2(w)| \leq C''\Delta$ with overwhelming probability. In the second step, we argue, exactly as in Theorem 20 that $|D \cap N(w)| \leq \varepsilon\Delta$ with overwhelming probability.

Consider the rate at which vertices in $S_2(w)$ become disagreements. This can occur due to their neighbors being disagreements, but since neither $\mathcal{B}_1$ nor $\mathcal{B}_2$ occurs, this contribution to the rate is never more than $\frac{(C'+1)\Delta^2}{nA_{\min}}$. (As before, this step uses the girth assumption, in this case $g \geq 7$, which implies that only the Δ neighbors in $N(w)$ have more than 1 neighbor in $S_2(w)$.) Disagreements can also arise due to the edges of $G \oplus G^*$ with $S_2(w)$, but since $G \oplus G^*$ has maximum in-degree 1, this contributes at most $\Delta^2/nA_{\min}$ to the rate, leading to a total rate of at most $\frac{(C'+2)\Delta^2}{nA_{\min}}$.

Hence the probability that more than $C''\Delta$ vertices in $S_2(w)$ become disagreements can be bounded by a Poisson tail:

$$\begin{aligned} & \Pr(|D \cap S_2(w)| \geq C''\Delta \text{ and not } \mathcal{B}_1 \text{ and not } \mathcal{B}_2) \\ & \leq \exp(-C_1(C'+1)\Delta^2/A_{\min}) \sum_{j \geq C''\Delta} \frac{(C_1(C'+1)\Delta^2/A_{\min})^j}{j!} \\ & \leq 2 \left(\frac{C_1(C'+1)\Delta^2 e}{A_{\min} C''\Delta} \right)^{C''\Delta} \\ & = 2 \left(\frac{C_1(C'+1)e}{\delta C''} \right)^{C''\Delta} \end{aligned}$$

which is $O(e^{-\Delta})$ for sufficiently large C'' as a function of δ and C_1 .

The final step, in which we show that, assuming $|D \cap S_2(w)| \leq C''\Delta$, then with high probability also $|D \cap N(w)| \leq \varepsilon\Delta$, is exactly as in Theorem 20. $\square$

Theorem 22 ensures that the available colors satisfy $A(X_t, z) \approx A(X_t^*, z)$, as stated more precisely in Corollary 25. Corollary 25 also gives a precise statement that $P(X_t, w, c) \approx P(X_t^*, w, c)$ with high probability. Our next result, which controls the distribution of disagreements involving a particular color, will be the key ingredient in that result, and will also be applied directly in Theorem 4.

Theorem 23. *For every $1/4 > \varepsilon, \delta > 0$, for every $C_1 > 0$, there exists $\Delta_0 > 0, C > 0$ such that the following holds. Suppose $G = (V, E)$ has girth $g \geq 7$ and maximum degree*

$\Delta \geq \Delta_0$. Let $v \in V$, and let G^* be G with three layers of edges around v oriented toward v . Let $k \geq (1 + \delta)\Delta$. Let X_0 be a coloring which is not $2/\varepsilon$ -heavy for a color c at vertex v . For every $w \in V$, let $S_w \subseteq N(w)$ be fixed. Let (X_t, X_t^*) denote a maximal coupling of the Glauber dynamics on G with the Glauber dynamics on G^* , starting from $X_0 = X_0^*$. Then

$$\Pr(\forall t \leq C_1 n, \forall w \in V, \forall c \in [k] \forall t \leq C_1 n \mid \{z \in N(S_w) : z \in X_t \oplus X_t^*, (X_t(z) = c \text{ or } X_t^*(z) = c)\} \mid \leq \varepsilon \Delta) \geq 1 - \exp(-\Delta/C).$$

Proof. As before, let $D = \cup_{t \leq C_1 n} X_t \oplus X_t^*$ denote the cumulative set of disagreements. We will also need a way to keep track of those disagreements which involve the color c . To this end, denote $D_c = \bigcup_{t \leq C_1 n} \{w \in B_R(v) : X_t(w) = c \neq X_t^*(w) \text{ or } X_t(w) \neq c = X_t^*(w)\}$. Thus, D_c is the cumulative set of disagreements involving color c .

We now summarize the rates at which these two types of disagreements spread. A disagreement involving c may propagate along any incident edge to form a new disagreement involving c , with rate $1/nA_{\min}$. An ordinary disagreement (not involving c) may propagate along an incident edge to form a new disagreement involving c , whenever exactly one of the two colors involved in the disagreement is blocked, but this occurs with rate at most $1/(nA_{\min}(A_{\min} + 1))$, since the extra color is equally likely to be coupled with each of the other colors. An ordinary disagreement can also propagate to form a new ordinary disagreement, with rate at most $1/nA_{\min}$. Furthermore, disagreements can also form at the heads of directed edges in $G \oplus G^*$, even when there is no disagreement involved. If the color of the vertex at the tail of this edge is c , then a disagreement involving c can form with rate $1/nA_{\min}$. If the color at the tail is not c , then a disagreement involving c can form with rate at most $1/nA_{\min}^2$, while an ordinary disagreement can form with rate at most $1/nA_{\min}$.

For this proof, we will focus on the following “bad” events $\mathcal{B}_i$. As before, let $\mathcal{B}_1$ denote the event that $D \not\subseteq B_R(v)$. Let $\mathcal{B}_2$ denote the event that $|D| \geq C'\Delta^2$. Let $\mathcal{B}_3$ denote the event that $|D_c| \geq C'\Delta$.

The probabilities of the events $\mathcal{B}_1$ and $\mathcal{B}_2$ are bounded analogously to the previous two theorems, and we omit the details. To bound the probability of $\mathcal{B}_3$, we observe that, so long as $|D| \leq C'\Delta^2$, the rate of formation of the $(i + 1)$ ’st disagreement involving c can be bounded by

$$\frac{|G \oplus G^*| + xA_{\min} + C'\Delta^3 + i\Delta A_{\min}}{nA_{\min}^2},$$

where x denotes the number of edges in $G \oplus G^*$ where the tail is colored c in X_t or X_t^* . Now, since the initial coloring X_0 is not $2/\varepsilon$ -heavy, we know that, at least initially, $x \leq 2\Delta^2/\delta$. Moreover, it is straightforward to show that, with high probability, x never

increases to more than twice this value over the course of $C_1 n$ steps. Analogously with the earlier bounds on $\Pr(\mathcal{B}_2)$, an application of Corollary 19 implies that with high probability, $|D_c| \leq C''' \Delta$.

Now, fix a vertex $w \in B_{R-3}(v)$. Assuming the bad events $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ don't occur, there are never more than $O(\Delta^2)$ disagreements in $S_3(w)$, of which at most $O(\Delta)$ involve color c . Under this assumption, consider the rate at which disagreements involving c can appear in $N(S_w)$.

Because the girth is at least 7, each vertex outside is adjacent to at most one vertex in $N(S_w)$. which means the rate is at most

$$\frac{\alpha + \beta A_{\min} + O(\Delta^2) + \Delta^2 + O(\Delta) A_{\min} + \gamma \Delta A_{\min}}{n A_{\min}^2}$$

where

- α is the number of edges in $G \oplus G^*$ having their heads in $N(S_w)$,
- β is the number of edges in $G \oplus G^*$ having their head in $N(S_w)$ and their tails colored c .
- γ is the number of disagreements in S_w colored c .

Note that $\alpha \leq \Delta^2$, due to our definition of G^* and because we assumed girth ≥ 7 (recall that $G \oplus G^*$ is a tree oriented towards its leaves, and hence every vertex has in-degree at most 1). β is $O(\Delta^2 / \log \Delta)$ with high probability because v is initially not C -heavy, and with high probability is never more than $2C$ -heavy. (This worst case is only achievable when $w = v$.) Finally, γ can be shown to be $O(\Delta / \log \Delta)$ with high probability, just by considering the number of neighbors of w which receive color c at some point after the initial coloring. [Note that the expected number of times elements of S_w are recolored is $C_1 |S_w|$, and each time the probability of choosing color c in one of the chains is at most $1/A_{\min}$; hence the expected number of recolorings of elements of S_w to color c in either chain is $O(1)$. An application of Chernoff's bound shows that the actual number is $O(\Delta / \log \Delta)$ with probability $1 - e^{-\Delta/C}$.]

This shows that, with high probability, at all times, the rate of disagreements involving c appearing in $N(S_w)$ is $O(\Delta / (n \log \Delta))$. Our familiar Poisson argument, also used in the proofs of Theorems 20 and 22 to prove a tail bound on the number of disagreements in a set, implies that with high probability the total number of disagreements is with high probability at most $\varepsilon \Delta$ after $C_1 n$ time units (for any fixed $\varepsilon > 0$ and Δ sufficiently large). $\square$

Definition 24. Fix a vertex v , and a subset $S \subset N(v)$. For every color c , $i \geq 0$, and coloring X , let $S_{c,i}(X)$ denote the set of $w \in S$ such that exactly i neighbors z of w , excluding v , satisfy $X(z) = c$. We call this the subset of S which is “ i times blocked for c .”

Corollary 25. For every $\varepsilon > 0, \delta > 0, C_1 > 0$, there exist $\Delta_0 > 0, C > 0$ such that the following holds. Suppose $G = (V, E)$ has girth $g \geq 7$ and maximum degree $\Delta \geq \Delta_0$. Let $v \in V$, $c \in [k]$ and let G^* be G with three layers of edges around v oriented toward v . Let $k \geq (1 + \delta)\Delta$. Let X_0 be a coloring which is not $2/\varepsilon$ -heavy for color c at v . Let (X_t, X_t^*) denote a maximal coupling of the Glauber dynamics on G with the Glauber dynamics on G^* , starting from $X_0 = X_0^*$. Then with probability at least $1 - \exp(-\Delta/C)$ the following hold for all $w \in V$ and all $t \leq C_1 n$:

$$||A(X_t, w)| - |A(X_t^*, w)|| \leq \varepsilon \Delta. \quad (13)$$

and

$$|P(X_t, w, c) - P(X_t^*, w, c)| \leq \varepsilon \quad (14)$$

Moreover, for any $w \in V$ and set $S \subseteq N(w)$, color c and non-negative integer i , with probability at least $1 - \exp(-\Delta/C)$,

$$|S_{c,i}(X_t) \oplus S_{c,i}(X_t^*)| \leq \varepsilon \Delta, \quad (15)$$

where $\oplus$ denotes the symmetric difference of sets.

Proof. Suppose the high-probability events of Theorems 22 and 23 both hold—in the latter case, taking $S_w = N(w)$ for all $w \in V$. That is, no vertex has more than $\varepsilon \Delta$ disagreements in its neighborhood, and no vertex has more than $\varepsilon \Delta$ disagreements involving color c in its sphere of radius 2. From the first fact, (13) follows immediately. From the second fact, (15) follows immediately.

To establish (14), which compares $P(X_t)$ and $P(X_t^*)$, we apply the definition of P , and then enumerate the sources of differences: the contributions of edges where G and G^* differ, and the contributions of the disagreement sets D and D_c . Thus,

$$\begin{aligned} & |P(X_t, w, c) - P(X_t^*, w, c)| \\ &= \sum_{y \in N(w)} \left| \frac{\mathbf{1}\{c \in A(X_t, w)\}}{|A(X_t, w)|} - \frac{\mathbf{1}\{c \in A(X_t^*, w)\}}{|A(X_t^*, w)|} \right| \end{aligned}$$

We split this into two sums, according to how many of the terms are nonzero (1 or 2), noting that the denominators are both at least $A_{\min}$ and cannot differ by more than 1:

$$\begin{aligned} &\leq \sum_{y \in N(w)} \frac{\mathbf{1}\{(y, w) \notin G^* \text{ or } c \in A(X_t, y) \oplus A(X_t^*, y)\}}{A_{\min}} \\ &\quad + \sum_{y \in N(w)} \sum_{w \neq z \in N(y)} \frac{\mathbf{1}\{(z, y) \notin G^* \text{ or } z \in D\}}{A_{\min}(A_{\min} + 1)} \end{aligned}$$

We further note that $c \in A(X_t, y) \oplus A(X_t^*, y)$ implies that either {some neighbor of y is a disagreement involving c }, or that {a directed edge (z, y) was deleted from G to form G^* and z received color c }:

$$\begin{aligned} &\leq \sum_{y \in N(w)} \frac{\mathbf{1}\{(y, w) \notin G^*\}}{A_{\min}} \\ &\quad + \sum_{y \in N(w)} \sum_{w \neq z \in N(y)} \frac{\mathbf{1}\{z \in D_c\}}{A_{\min}} + \frac{\mathbf{1}\{(z, y) \notin G^* \text{ and } X_t(z) = c\}}{A_{\min}} + \frac{\mathbf{1}\{(z, y) \notin G^* \text{ or } z \in D\}}{A_{\min}(A_{\min} + 1)} \end{aligned}$$

Let us bound these four terms one by one.

1. Since $G \oplus G^*$ has maximum in-degree 1, the first term is nonzero for at most one value of y , and therefore contributes a total of at most $1/A_{\min} \leq \varepsilon/4$.
2. By Theorem 23 applied with $\eta = \varepsilon\delta/4$, the second term contributes at most $\eta\Delta/A_{\min} \leq \eta/\delta = \varepsilon/4$.
3. For the third term, note that, because $G \oplus G^*$ has maximum in-degree 1, there are at most Δ pairs (z, y) which could possibly contribute, and these are fixed. By the definition of C -heavy, we know that under X_0 , $O(\Delta/\log \Delta)$ of these were colored c . Of the ones which are recolored, with high probability $O(\Delta/\log \Delta)$ received color c at their final recoloring. So, overall, with high probability $O(\Delta/\log \Delta)$ of these end with color c in X_t . Hence the third term contributes a total of at most $\varepsilon/4$.
4. For the final term, recall that we already observed that at most Δ pairs $(z, y) \in G \oplus G^*$. Moreover, by Theorem 22 applied with error parameter $\varepsilon\delta^2/4$, $|D \cap S_2(w)| \leq \varepsilon\delta^2\Delta^2/4$, and hence the total contribution is $\leq \varepsilon/4$.

This completes the proof of Corollary 25. $\square$

5 Local Relations for Glauber dynamics

In this section, we will prove analogues of Lemmas 7 and 9 in the setting of the continuous-time Glauber dynamics. As in the case of uniformly random colorings, these local relations are the key to all our uniformity properties.

5.1 Local Relation for $|A|$

We will need the following result, due to Dyer and Frieze [3, Lemma 2.1]. It says that, in a sequence of independent color selections for which no one color is very likely in any stage, the number of *missed* colors will not be much less than if each color were chosen independently and uniformly from all of $[k]$. Although originally stated with the stronger hypothesis that each color is selected *uniformly* from a subset of $[k]$, the original proof suffices for the current version as well. We include a proof for completeness.

Lemma 26 (Dyer and Frieze). *Let k, s be positive integers, and let $c_1, \dots, c_s$ be independent (but not identically distributed) random variables taking values in $[k]$. Let p be the maximum over $1 \leq i \leq s$ of the probability of the most likely value for c_i . Let $A = [k] \setminus \{c_1, \dots, c_s\}$ be the set of missed colors. Then*

$$\mathbf{E}(|A|) \geq k(1-p)^{s/kp} \geq ke^{-s/k}(1-p)^{s/k},$$

and for every $a > 0$, $\Pr(|A| \leq \mathbf{E}(|A|) - a) \leq e^{-a^2/2k}$.

Proof. For each $1 \leq i \leq s$, and $1 \leq j \leq k$, let $\eta_{i,j}$ denote the indicator variable for the event that $c_i = j$. Observe that

$$|A| = \sum_{j=1}^k \prod_{i=1}^s (1 - \eta_{i,j}).$$

By linearity of expectation and the independence of the color choices, and by the arithmetic-geometric mean inequality, this implies

$$\begin{aligned} \mathbf{E}(|A|) &= \sum_{j=1}^k \prod_{i=1}^s (1 - \mathbf{E}(\eta_{i,j})) \\ &\geq k \prod_{j=1}^k \prod_{i=1}^s (1 - \mathbf{E}(\eta_{i,j}))^{1/k}. \end{aligned}$$

Now, for each i , $\sum_{j=1}^k \mathbf{E}(\eta_{i,j}) = 1$, and moreover, for each i, j , $0 \leq \mathbf{E}(\eta_{i,j}) \leq p$. An easy shifting argument shows that, subject to these constraints, the minimum for the right hand side is achieved when as many as possible of the $\mathbf{E}(\eta_{i,j})$ equal p , and hence

$$\begin{aligned} \mathbf{E}(|A|) &\geq k(1-p)^{s\lfloor 1/p \rfloor/k} (p\lfloor 1/p \rfloor)^{s/k} \\ &\geq k(1-p)^{s/kp}. \end{aligned}$$

Note that $1-p \geq e^{-p/(1-p)}$, a handy variant of the standard inequality $1+x \leq e^x$. It follows that

$$(1-p)^{1/p} = (1-p)^{(1-p)/p} (1-p) \geq \frac{1-p}{e},$$

and hence $k(1-p)^{s/kp} \geq ke^{-s/k}(1-p)^{s/k}$.

Moreover, since every subset of the sk events $\{c_i = j\}$ are either fully independent (if all the i 's are distinct) or mutually exclusive (otherwise), it follows that the random variables $\eta_{i,j}$ are negatively associated (see Dubhashi and Ranjan [2] for a concise overview). Since decreasing functions of disjoint subsets of a family of negatively associated variables are also negatively associated, the k variables

$$\prod_{i=1}^s 1 - \eta_{i,j}$$

are negatively associated. Since Chernoff's bound applies to sums of bounded negatively associated variables (see [2][Proposition 1.5]), and since $|A| = \sum_{j=1}^k \prod_{i=1}^s 1 - \eta_{i,j}$, we conclude that for every $a > 0$, $\mathbf{Pr}(|A| \leq \mathbf{E}(|A|) - a) \leq e^{-a^2/2k}$. $\square$

Let $G^* = (V, E^*)$ be the graph G modified so that all the edges within the ball of radius 2 centered at v are directed inward. Next, we prove a local relation for $|A(X_t^*, v)|$, where $(X_t^*)_{t \geq 0}$ is the continuous-time heat bath Glauber dynamics on G^* , from an arbitrary initial coloring X_0^* of G . This result does not require the dynamics to be connected, so it applies whenever k is at least the chromatic number of G . An analogous relation for uniformly random colorings was proved as Lemma 7.

Lemma 27. *Let $0 < \varepsilon < 1/4$, let $\Delta_0 = \Delta_0(\varepsilon)$, and let $G = (V, E)$ have girth ≥ 5 and $\Delta > \Delta_0$. Let $v \in V$, and let G^* equal G modified so that 2 levels of edges around v are oriented towards v . Let k be at least the chromatic number of G . Let $(X_t^*)_{t \geq 0}$ be the continuous-time Glauber dynamics on G^* with arbitrary X_0^* . We will condition on the restriction of X_t^* to $V \setminus B_1(v)$, for all $t \in [0, T]$. Denote this by $\mathcal{F}$. Let $I = [T - \ell n, T] \subset [0, \infty)$ be a fixed non-empty time interval. Then*

$$\mathbf{Pr} \left(\frac{|A(X_T^*, v)|}{k} \leq \left(\frac{1 - e^{-\ell} - \frac{1}{A_{\min}}}{e} \right)^{d(v)/k} - \varepsilon \mid \mathcal{F} \right) \leq e^{-\varepsilon^2 k/2}.$$

where $A_{\min} = \min_{w,s} |A(X_s^*, w)|$, with the minimum taken over all $w \in N(v)$ and $s \in I$.

Proof. Let $\mathcal{F}$ denote the σ -algebra generated by $X_{T-\ell}^*$, together with the restrictions of X_s^* to the $V \setminus B_1(v)$, for all $s \in I$. The neighbor colors $X_T^*(w)$, for $w \in N(v)$, are conditionally fully independent, conditioned on $\mathcal{F}$. This is because (a) the chain is in continuous time, so the update times of the vertices are fully independent, (b) there are no paths in G from $B_1(v)$ to its complement, so information about the colors assigned to $N(v)$ do not affect the posterior probability of $\mathcal{F}$, and (c) within $B_1(v)$, there are no edges oriented into $N(v)$.

Next, we describe the conditional distribution of $X_T^*(w)$ given $\mathcal{F}$, where w is a neighbor of v . For each $c \in [k]$,

$$\Pr(X_T^*(w) = c \mid \mathcal{F}) = e^{-\ell} \mathbf{1}\{X_{T-\ell n}^*(w) = c\} + \int_{s=0}^{\ell} \frac{\mathbf{1}\{c \in A(X_{T-sn}^*, w)\}}{|A(X_{T-sn}^*, w)|} e^{-s} ds. \quad (16)$$

This expression can be derived by considering the amount of time from the last successful recoloring of w before T , until T ; this quantity is exponentially distributed, with a probability of $e^{-\ell}$ of being greater than ℓn , in which case $X_T^*(w) = X_{T-\ell n}^*(w)$.

Considering (16) and the definition of $A_{\min}$, it follows immediately that

$$\begin{aligned} \Pr(X_T^*(w) = c \mid \mathcal{F}) &\leq e^{-\ell} + \int_{s=0}^{\ell} \frac{1}{A_{\min}} e^{-s} ds \\ &\leq e^{-\ell} + \frac{1 - e^{-\ell}}{A_{\min}}. \end{aligned}$$

Applying Lemma 26 to the colors $X_T^*(w)$, with $p = e^{-\ell} + \frac{1}{A_{\min}}$ and $a = \varepsilon k$, we have

$$\Pr\left(\frac{|A(X_T^*, v)|}{k} \leq \left(\frac{1-p}{e}\right)^{d(v)/k} - \varepsilon \mid \mathcal{F}\right) \leq e^{-\varepsilon^2 k/2},$$

which completes the proof. $\square$

Next we apply our comparison techniques from the previous section to derive an analogous lower bound for the original undirected graph.

Lemma 28. *Let $0 < \delta, \varepsilon < 1/4$, let $\Delta_0 = \Delta_0(\delta, \varepsilon)$, $C = C(\delta, \varepsilon)$, and let $k \geq (1 + \delta)\Delta$. Let $G = (V, E)$ have girth ≥ 5 and $\Delta > \Delta_0$. Let $(X_t)_{t \geq 0}$ be the continuous-time Glauber dynamics on G with arbitrary X_0 . Let $v \in V$ and $T \geq n \ln(1/\varepsilon)$. Then*

$$\Pr\left(\frac{|A(X_T, v)|}{k} \leq e^{-d(v)/k} - 4\varepsilon\right) \leq 2e^{-k/C}.$$

Proof. Fix $X_{T-n\ln(1/\varepsilon)}$ and run a naive coupling (X_t, X_t^*) , for $T - n\ln(1/\varepsilon) \leq t \leq T$ (in continuous time) from initial state $X_{T-n\ln(1/\varepsilon)}^* = X_{T-n\ln(1/\varepsilon)}$, where (X_t^*) is the heat-bath Glauber dynamics on G^* , which is G with two layers of edges around v oriented toward v . By Lemma 27, we know that, with high probability (at most half the desired probability of error),

$$\begin{aligned} \frac{|A(X_T^*, v)|}{k} &\geq \left(\frac{1 - e^{-\ln(1/\varepsilon)} - \frac{1}{A_{\min}}}{e} \right)^{d(v)/k} - \varepsilon && \text{by Lemma 27} \\ &\geq e^{-d(v)/k} - 3\varepsilon, \end{aligned}$$

where the last line follows because $A_{\min} \geq k - \Delta \geq \delta\Delta \geq 1/\varepsilon$ for $\Delta \geq \Delta_0(\delta) \geq 1/\varepsilon\delta$.

Assuming the above high-probability event holds, it suffices to show that

$$|A(X_T, v)| \geq |A(X_T^*, v)| - \varepsilon k,$$

for which it is sufficient that

$$|(X_T \oplus X_T^*) \cap N(v)| \leq \varepsilon k.$$

Theorem 20 gives the desired upper bound on this probability, completing the proof. $\square$

Lemma 29. *Let $0 < \delta, \varepsilon < 1/4$, let $\Delta_0 = \Delta_0(\delta, \varepsilon)$, and let $k \geq (1 + \delta)\Delta$. Let $G = (V, E)$ have girth ≥ 7 and $\Delta > \Delta_0$, let $v \in V$, and let G^* equal G modified so that 3 levels of edges around v are oriented towards v . Let $(X_t^*)_{t \geq 0}$ be the continuous-time Glauber dynamics on G^* with arbitrary X_0^* . Let $v \in V$. We will condition on the restriction of X_t^* to $V \setminus B_2(v)$, for all $t \in [0, T]$. Denote this by $\mathcal{F}$. Then for all $a \geq 0$, for all $w \in N(v)$,*

$$\begin{aligned} \Pr(|A(X_T^*, w)| \leq \mathbf{E}(|A(X_T^*, w)| \mid \mathcal{F}) - a) &\leq e^{-a^2/2k} \\ \Pr(|A(X_T^*, w)| \geq \mathbf{E}(|A(X_T^*, w)| \mid \mathcal{F}) + a) &\leq e^{-a^2/2k}. \end{aligned}$$

Proof. The conditioning gives us full independence for the colors assigned to $N(v)$, so the result follows by Chernoff's bound. $\square$

5.2 Local Relation for P

Now we can state our local relation for P in the context of the Glauber dynamics. This may be viewed as a dynamical version of Lemma 9. As in Lemma 27, we work in continuous time and with a partially directed version of the original graph. We additionally assume that vertex v is not heavy for color c under the initial coloring.

Lemma 30. *Let $0 < \delta, \varepsilon < 1/20$, let $\Delta_0 = \Delta_0(\delta, \varepsilon)$, $C = C(\delta, \varepsilon)$ and let $k \geq (1 + \delta)\Delta$. Let $G = (V, E)$ have girth ≥ 7 and $\Delta > \Delta_0$. Let $(X_t^*)_{t \geq 0}$ be the continuous-time Glauber dynamics on G^* , which is G with three levels of edges around v oriented inwards. Let $v \in V$, $c \in [k]$, and suppose that X_0^* is not $4/\varepsilon$ -heavy with respect to color c at vertex v . Then for $T \geq 2n \log(1/(\varepsilon\delta))$,*

$$\Pr \left((\exists t \in [2n \log(1/\varepsilon), T]) \left| P(X_t^*, v, c) - \sum_{w \sim v} \frac{\exp(-\mathbf{E}_s P(X_s^*, w, c))}{|A(X_t^*, w)|} \right| > 50\varepsilon \right) < (10T/n) \exp(-\Delta/C).$$

Here, $s \in [0, t]$ is distributed as the last recoloring time prior to t , that is, for $0 \leq a \leq t$, $\Pr(s \leq a) = e^{-(t-a)/n}$.

Proof. Let $\mathcal{F}$ denote the entire history of X^* , up to time t , excluding the ball of radius 2 centered at v . More precisely, $\mathcal{F}$ is the σ -algebra generated by the restrictions $X_{t'}^*|_{V \setminus B_2(v)}$, for all $t' \in [0, t]$.

Recall that by definition, $P(X_t^*, v, c) = \sum_{w \in N(v)} Y_w$, where each Y_w is a function of the colors assigned to $N(w) \setminus \{v\}$ under X_t . Since, in G^* , there is no path from a vertex in $N(w_1) \setminus \{v\}$ to any vertex in $N(w_2) \setminus \{v\}$, where $w_1 \neq w_2 \in N(v)$, and since we are in continuous time, it follows that, conditioned on $\mathcal{F}$, the variables $Y_w, w \in N(v)$ are fully independent. Applying Chernoff's bound, we find that with high probability $P(X_t^*, v, c) \approx \mathbf{E}(P(X_t^*, v, c) | \mathcal{F})$. More precisely, since the variables Y_w take values in $[0, 1/(\delta\Delta)]$, Chernoff's bound yields

$$\Pr(|P(X_t^*, v, c) - \mathbf{E}(P(X_t^*, v, c) | \mathcal{F})| > \varepsilon | \mathcal{F}) \leq 2 \exp(-\delta^2 \varepsilon^2 \Delta/2). \quad (17)$$

Now, by the definition of P and linearity of expectation,

$$\begin{aligned} & \mathbf{E}(P(X_t^*, v, c) | \mathcal{F}) \\ &= \sum_{w \sim v} \mathbf{E} \left(\frac{\mathbf{1}\{c \in A(X_t^*, w)\}}{|A(X_t^*, w)|} \middle| \mathcal{F} \right) \\ &\in \sum_{w \sim v} \left(\frac{\Pr(c \in A(X_t^*, w) | \mathcal{F})}{|A(X_t^*, w)| \pm \varepsilon k/50} \pm \exp(-\delta^2 \varepsilon^2 \Delta 10^{-6}) \right) \quad \text{by Observation 8 and Lemma 29} \\ &\subseteq \left(1 \pm \frac{\varepsilon}{15} \right) \sum_{w \sim v} \frac{\Pr(c \in A(X_t^*, w) | \mathcal{F})}{|A(X_t^*, w)|}, \end{aligned} \quad (18)$$

where the last line uses that $A_{\min} \geq k/e$ with high probability by Lemma 28.

We approximate the numerators in the right hand side as follows:

$$\begin{aligned}
& \Pr(c \in A(X_t^*, w) \mid \mathcal{F}) \\
&= \prod_{z \in N(w)} \mathbf{E}(1 - \mathbf{1}\{X_t^*(z) = c\} \mid \mathcal{F}) \quad \text{by conditional independence} \\
&\in (1 \pm O(1/k^2))^\Delta \prod_{z \in N(w)} \exp(-\Pr(X_t^*(z) = c \mid \mathcal{F})) \\
&= (1 \pm O(1/k^2))^\Delta \exp\left(-\sum_{z \in N(w)} \Pr(X_t^*(z) = c \mid \mathcal{F})\right)
\end{aligned}$$

Plugging into (18) and using (17), we have, except with probability $\leq \Delta^5 e^{-\varepsilon^2 \Delta / 5000}$:

$$\begin{aligned}
& P(X_t^*, v, c) \\
&\in (1 \pm \varepsilon/14) \sum_{w \in N(v)} \frac{\exp(-\sum_{z \in N(w)} \Pr(X_t^*(z) = c \mid \mathcal{F}))}{|A(X_t^*, w)|} \\
&= (1 \pm \varepsilon/14) \sum_{w \in N(v)} \frac{\exp\left(-\sum_{z \in N(w)} \left(e^{-t/n} \mathbf{1}\{X_0^*(z) = c\} + \int_0^t \frac{\mathbf{1}\{c \in A(X_s^*, z)\}}{|A(X_s^*, z)|} e^{(s-t)/n} ds\right)\right)}{|A(X_t^*, w)|} \\
&= (1 \pm \varepsilon/14) \sum_{w \in N(v)} \frac{\exp\left(-e^{-t/n} |X_0^{*-1}(c) \cap N(w)| - \int_0^t P(X_s^*, w, c) e^{(s-t)/n} ds\right)}{|A(X_t^*, w)|}.
\end{aligned}$$

Now, since X_0^* is not $4/\varepsilon$ -heavy for c at v , it follows that except for up to $\varepsilon\Delta$ “bad” neighbors $w \in N(v)$, the number of blocking vertices is $\leq 4/\varepsilon\delta$. Since $e^{-t/n} \leq \varepsilon^2 \delta^2 \ll \varepsilon^2 \delta / 2$, we may conclude (with high probability)

$$P(X_t^*, v, c) \in (1 \pm \varepsilon/10) \left(\sum_{w \in N(v)} \frac{\exp\left(-\int_0^t P(X_s^*, w, c) e^{(s-t)/n} ds\right)}{|A(X_t^*, w)|} \right) \pm \varepsilon e.$$

Since $A_{\min} \geq k/e$ with high probability by Lemma 28, we have:

$$\begin{aligned} P(X_t^*, v, c) &\in \left(\sum_{w \in N(v)} \frac{\exp\left(-\int_0^t P(X_s^*, w, c) e^{(s-t)/n} ds\right)}{|A(X_t^*, w)|} \right) \pm 20\varepsilon \\ &\subseteq \left(\sum_{w \sim v} \frac{\exp(-\mathbf{E}_s P(X_s^*, w, c))}{|A(X_t^*, w)|} \right) \pm 40\varepsilon. \end{aligned}$$

Finally, a covering argument combined with a union bound extends the result from the single time t to the entire interval $[2n \log(1/\varepsilon), T]$, at the cost of increasing the error probability by a factor of $O(T/(n\varepsilon))$ and the accuracy by $O(\varepsilon)$. $\square$

Next we show that after $O(n \log \Delta)$ time units, the probability that X_t is heavy for a particular vertex is already very small, regardless of the original coloring X_0 .

Lemma 31. *Let $0 < \delta, \varepsilon < 1/20$, let $\Delta_0 = \Delta_0(\delta, \varepsilon)$, let $C = C(\delta, \varepsilon)$ and let $k \geq (1+\delta)\Delta$. Let $G = (V, E)$ have girth ≥ 5 and $\Delta > \Delta_0$. Let $(X_t)_{t \geq 0}$ be the continuous-time (or discrete-time) Glauber dynamics on G with arbitrary X_0 . Let $v \in V$, $c \in [k]$, and $t > 3n \log \Delta$. Then*

$$\Pr(X_t \text{ is } 2/\varepsilon\text{-heavy for } v \text{ at } c) < \exp(-\Delta/C).$$

Moreover, if X_0 is not $4/\varepsilon$ -heavy for v at c , then the same conclusion holds for $t \geq Cn$.

Proof. In continuous time, each vertex $w \in V$ has a probability of $e^{-t/n} < 1/\Delta^3$ to go unrecolored during the entire time interval. In discrete time, the probability is $(1 - 1/n)^t$, which is even smaller.

Assuming w is recolored at least once, it has probability at most $1/A_{\min} < 1/(\delta\Delta)$ to be assigned color c at the time of its last recoloring prior to time t . Since this is true regardless of what colors were assigned previously to the other vertices, it follows that, within any fixed vertex set S , the number of vertices colored c at time t is stochastically dominated by the number of heads in $|S|$ independent tosses of a biased coin which comes up heads with probability $(1/\Delta^3 + 1/\delta\Delta)$. Chernoff's bound implies that this number is, with high probability, at most $2|S|/\delta\Delta$.

Applying this to the sets $B_2(v)$, $N(v)$, and $N(w)$ for each $w \in B_6(v)$, implies that X_t is not $2/\varepsilon$ -heavy for c at v .

To see the proof for $t \geq Cn$ under the assumption that X_0 is not $4/\varepsilon$ -heavy for c at v , we note that the only unrecolored vertices which contribute to the count of color c are the at most $(4/\delta)\Delta$ which are colored c under X_0 . Otherwise, the proof is the same. $\square$

We now show that, once the property of not being $2/\varepsilon$ -heavy appears (after $O(n \log \Delta)$ steps), it persists with high probability for $\Omega(n \log n)$ steps.

Corollary 32. *Let $0 < \delta, \varepsilon < 1/20$, let $\Delta_0 = \Delta_0(\delta, \varepsilon)$, let $C = C(\delta, \varepsilon)$ and let $k \geq (1 + \delta)\Delta$. Let $G = (V, E)$ have girth ≥ 5 and $\Delta > \Delta_0$. Let $(X_t)_{t \geq 0}$ be the continuous-time (or discrete-time) Glauber dynamics on G with arbitrary X_0 . Let $v \in V, c \in [k]$. Then*

$$\Pr((\exists t \in [3n \log \Delta, T]) \text{ } X_t \text{ is } 2/\varepsilon\text{-heavy for } c \text{ at } v) < \frac{10T}{n} \exp(-\Delta/C).$$

Moreover, if X_0 is not $(4/\varepsilon)$ -heavy for c at v , then the same conclusion holds for $t \in [Cn, T]$.

Proof. As in Section 5.1, it suffices to check the property holds with high probability for short time intervals, then take a union bound to cover the long time interval $[Cn, T]$. For an interval of length $\varepsilon n / \Delta$, we expect at most $\varepsilon \Delta$ vertices in $B_2(v)$ to be recolored, and Chernoff's bound implies high probability that at most $2\varepsilon \Delta$ vertices are recolored. By Lemma 31 X_t is not $2/\varepsilon$ -heavy for a particular time, with high probability. Assuming this holds for each time in an εn -net for the entire time interval, X_t is never more than $4/\varepsilon$ -heavy in the time interval. Taking a union bound over the $T/\varepsilon n$ times in the covering completes the proof (we choose C and Δ_0 sufficiently large to absorb the $1/\varepsilon$ term). $\square$

Now we are ready to derive our local relation for P on the original undirected graph.

Lemma 33. *Let $0 < \delta, \varepsilon < 1/20$, let $\Delta_0 = \Delta_0(\delta, \varepsilon)$, let $C = C(\delta, \varepsilon)$ and let $k \geq (1 + \delta)\Delta$. Let $I = [t_0, t_1]$ be a time interval with $t_0 \geq Cn \log \Delta$. Let $G = (V, E)$ have girth ≥ 7 and $\Delta > \Delta_0$. Let $(X_t)_{t \geq 0}$ be the continuous-time (or discrete-time) Glauber dynamics on G with arbitrary X_0 . Let $v \in V$ and $c \in [k]$. Then*

$$\Pr\left((\exists t \in I) \left| P(X_t, v, c) - \sum_{w \sim v} \frac{\exp(-\mathbf{E}_s(P(X_s, w, c)))}{|A(X_t, w)|} \right| > 12\varepsilon \right) < 10 \left(1 + \frac{t_1 - t_0}{n} \right) \exp(-\Delta/C).$$

Here, $s \in [0, t]$ is distributed as the last recoloring time prior to t , that is, for $0 \leq a \leq t$, $\Pr(s \leq a) = e^{-(t-a)/n}$.

Moreover, if X_0 is not $(4/\varepsilon)$ -heavy for c at v , then the same conclusion holds for any $t_0 \geq Cn$.

Proof. By Corollary 32, we can discount the possibility that there exists $t \in [t_0/2, t_1]$, such that X_t is $(2/\delta)$ -heavy for c at v . Now let $t' = t - Cn/2 \geq t_0/2$. Condition on the coloring $X_{t'}$. Then we run a naive coupling (X_t, X_t^*) of the Glauber dynamics on G with the Glauber dynamics on G^* , where G^* is the modified graph of Lemma 30, starting at time t' with $X_{t'}^* = X_{t'}$. By Lemma 30, we have with high probability

$$\left| P(X_t^*, v, c) - \sum_{w \sim v} \frac{\exp(-\mathbf{E}_s(P(X_s^*, w, c)))}{|A(X_t^*, w)|} \right| \leq \varepsilon \quad (19)$$

Finally, Corollary 25 implies that with high probability, for every $s \in [t', t]$,

$$|P(X_s, v, c) - P(X_s^*, v, c)| \leq \varepsilon$$

and for every $s \in [t', t]$ and every $w \sim v$,

$$|A(X_s, w) - A(X_s^*, w)| \leq \varepsilon \Delta.$$

The triangle inequality together with some basic arithmetic lets us infer that

$$\begin{aligned} & \left| \sum_{w \sim v} \frac{\exp(-\mathbf{E}_s(P(X_s, w, c)))}{|A(X_t, w)|} - \sum_{w \sim v} \frac{\exp(-\mathbf{E}_s(P(X_s^*, w, c)))}{|A(X_t^*, w)|} \right| \\ & \leq \sum_{w \sim v} \left| \frac{\exp(-\mathbf{E}_s(P(X_s, w, c))) - \exp(-\mathbf{E}_s(P(X_s^*, w, c)))}{|A(X_t, w)|} \right| \\ & \quad + \sum_{w \sim v} \left| \frac{\exp(-\mathbf{E}_s(P(X_s^*, w, c)))}{|A(X_t, w)|} - \frac{\exp(-\mathbf{E}_s(P(X_s^*, w, c)))}{|A(X_t^*, w)|} \right| \\ & \leq \sum_{w \sim v} \frac{\varepsilon}{|A(X_t, w)|} + \sum_{w \sim v} \frac{\exp(-\mathbf{E}_s(P(X_s^*, w, c))) \varepsilon \Delta}{A_{\min}^2} \\ & \leq \frac{\varepsilon \Delta}{A_{\min}} + \frac{\varepsilon \Delta^2}{A_{\min}^2} \leq (e^2 + e) \varepsilon < 11 \varepsilon \end{aligned}$$

where the last line follows because with high probability, $A_{\min} \geq \Delta/e$. Plugging these bounds into (19) and applying the triangle inequality completes the proof. $\square$

6 Bias for Glauber dynamics

In this section, we apply the results of the previous section to derive an absolute (high probability) bound for P for the Glauber dynamics. The main result is an analog of Theorem 11.

Theorem 34. Let $0 < \delta, \varepsilon < 1/20$, let $\Delta_0 = \Delta_0(\varepsilon, \delta)$, let $C = C(\varepsilon, \delta)$, and let $k \geq (1 + \delta)\Delta$. Let $I = [t_0, t_1]$ be a time interval with $t_0 \geq Cn \log \Delta$. Let $G = (V, E)$ have girth ≥ 7 and $\Delta > \Delta_0$. Let $(X_t)_{t \geq 0}$ be the continuous-time Glauber dynamics on G with arbitrary X_0 . Let $v \in V$ and $c \in [k]$. Then with probability at least $1 - \frac{t_1 - t_0}{n} \exp(-\Delta/C)$ the following holds:

$$(\forall t \in I) \quad \left| P(X_t, v, c) - \frac{d(v)}{k} \right| \leq \varepsilon.$$

Moreover, if X_0 is not $(4/\varepsilon)$ -heavy for any $c' \in [k]$, $w \in B_R(v)$, where $R = R(\delta, \varepsilon)$, then the same conclusion holds for $t_0 \geq (R + 1)Cn$.

Proof. The basic approach of the proof is similar to the proof of Theorem 11.

Let R denote the positive integer $\lceil 5 \log(10/\varepsilon)/\delta \rceil$. Fix two colors c and c' .

We will restrict our attention to times $t > t_0 - RCn$, for which Lemma 33 tells us that with high probability, the local relation for P holds. (Note that our $C(\delta, \varepsilon)$ is therefore slightly bigger than the one in the statement of Lemma 33.)

For each time t and positive integer $i \leq R$, we define

$$\alpha_i := \max |P(X_s, z, c) - P(X_s, z, c')|,$$

where the maximum is taken over all times $s \in I_i := [t_0 - iCn, t_1]$ and over all vertices z at distance $\leq i$ from v .

We will assume henceforth that the high-probability event from Lemma 28 holds for every $w \in B_R(v)$, at all times $t \in [t_0 - CRn, t_1]$. In particular, if $A_{\min}$ denotes the minimum number of available colors at any vertex $w \in B_R(v)$, at any such time t , then $A_{\min} \geq k/e$. It follows that $\alpha_0 \leq \alpha_1 \leq \dots \leq \alpha_R \leq e^e$.

Now partition each time interval I_i into many short subintervals so that, with high probability, $o(\Delta)$ vertices from $B_i(v)$ are updated in any one subinterval. This allows us to focus on a single time t rather than an entire interval, and deduce the conclusion via a union bound.

Now, for a particular fixed vertex $z \in B_{R-2}(v)$ and fixed time $s \in I_i$, let us condition on X_{s-Cn} , and inspect X_s .

By Lemma 33, we know that

$$\begin{aligned} \Pr \left((\exists t \in I_i) \left| P(X_t, v, c) - \sum_{w \sim v} \frac{\exp(-\mathbf{E}_a(P(X_a, w, c)))}{|A(X_t, w)|} \right| > \varepsilon \delta / 200 \right) \\ < 10((t_1 + iCn - t_0)/n) \exp(-\Delta/C). \end{aligned}$$

where a is distributed as $t - \min\{Cn, \eta\}$, where η is an exponential random variable with mean n . Assuming this high-probability event holds for colors c and c' , it follows that

$$\begin{aligned}
P(X_s, z, c') &\leq \frac{\varepsilon\delta}{200} + \sum_{u \in N(z)} \frac{\exp(-\mathbf{E}_a(P(X_a, u, c')))}{|A(X_s, u)|} && \text{by Lemma 33} \\
&\leq \frac{\varepsilon\delta}{200} + \sum_{u \in N(z)} \frac{\exp(\alpha_{i+1} - \mathbf{E}_a(P(X_a, u, c)))}{|A(X_s, u)|} && \text{by def. of } \alpha_{i+1} \\
&\leq \frac{\varepsilon\delta}{200} + \exp(\alpha_{i+1}) \left(P(X_s, z, c) + \frac{\varepsilon\delta}{200} \right) && \text{by Lemma 33} \\
&= \exp(\alpha_{i+1}) P(X_s, z, c) + \frac{\varepsilon\delta}{200} (1 + \exp(\alpha_{i+1})) \\
&\leq \exp(\alpha_{i+1}) P(X_s, z, c) + \frac{\varepsilon\delta}{20} && \text{since } \alpha_{i+1} \leq e^e \quad (20)
\end{aligned}$$

Applying Lemma 33 again, this time for every $u \sim z$, colors c, c' , and times $t \in I_i$, we obtain, with high probability,

$$\begin{aligned}
&|P(X_s, z, c) - P(X_s, z, c')| \\
&\leq \frac{\varepsilon\delta}{100} + \sum_{u \sim z} \frac{\exp(-\mathbf{E}_a(P(X_a, u, c))) - \exp(-\mathbf{E}_a(P(X_a, u, c')))}{|A(X_s, z)|} && \text{by Lemma 33} \\
&\leq \frac{\varepsilon\delta}{100} + \sum_{u \sim z} \frac{\exp(-\mathbf{E}_a(P(X_a, u, c))) - \exp(-\mathbf{E}_a(e^{\alpha_{i+1}} P(X_a, u, c)) - \frac{\varepsilon\delta}{20})}{|A(X_s, z)|} && \text{by (20).} \\
&= \frac{\varepsilon\delta}{100} + \sum_{u \sim z} \frac{y_u - y_u^{\exp(\alpha_{i+1})} + (1 - e^{-\frac{\varepsilon\delta}{20}}) y_u^{\exp(\alpha_{i+1})}}{|A(X_s, z)|} && \text{(see below)} \\
&\leq \frac{\varepsilon\delta}{100} + \sum_{u \sim z} \frac{\alpha_{i+1}}{e|A(X_s, z)|} + \frac{(1 - e^{-\frac{\varepsilon\delta}{20}})}{|A(X_s, z)|} && \text{by Lemma 10 and since } y_u \leq 1 \\
&\leq \frac{\alpha_{i+1}\Delta}{eA_{\min}} + \frac{\frac{\varepsilon\delta}{10}\Delta}{A_{\min}} + \frac{\varepsilon\delta}{100} && \text{since } \frac{\varepsilon\delta}{20} < 1/2 \\
&\leq \frac{\alpha_{i+1}}{1+\delta} + \left(e\frac{\varepsilon\delta}{10} + \frac{\varepsilon\delta}{50} \right) && A_{\min} \geq k/e \text{ w.h.p.} \\
&\leq e^{-\delta/2} \alpha_{i+1} && \text{assuming } \alpha_{i+1} \geq \varepsilon.
\end{aligned}$$

In the above, we denoted $y_u := \exp(-\mathbf{E}_a(P(X_a, u, c)))$, and used the fact that $y_u \leq 1$.

To bound α_i , we apply the above argument to an $\varepsilon\delta/20\Delta$ -net covering I_i . By a Chernoff bound, with probability at least $1 - \exp(-\varepsilon^4\delta^4\Delta/200)$, at most $\varepsilon\delta\Delta/16$ vertices in $S_2(z)$ get updated between two consecutive times in our net. In this case, $P(X_s, z, c)$ can't change by more than $\varepsilon\delta/4$ in such a window. Thus $\alpha_i \leq \max\{\varepsilon\delta/4 + e^{-\delta/2}\alpha_{i+1}, \varepsilon\}$. As in the proof of Theorem 11, since α_R is suitably upper-bounded, and R is sufficiently large, it follows that $\alpha_0 \leq \varepsilon$.

The combined error probability in our analysis is based on the following high-probability events we assumed at various points in our proof:

- no big changes to P during $\text{poly}(\Delta)$ tiny time intervals.
- lower bound on available colors from Lemma 28, applied to $B_R(v)$ over time interval I_R .
- local relation for P from Lemma 33, applied to $B_R(v)$ over time interval I_R , for each pair of colors (c, c') .

The sum of the $(\text{poly}(\Delta))$ many error probabilities is clearly bounded by the expression given in the theorem statement. $\square$

7 Further Uniformity Properties: Proof of Theorem 4

In this section, we prove Theorem 4. We show how to parlay our upper bound on color bias into the other uniformity properties we need for our applications.

We are now ready to prove Part 2 of Theorem 4.

Proof of (2) from Theorem 4. Orient the edges in $B_3(v)$ towards v , and let X_t^* denote the Glauber dynamics on this modified graph, starting from $X_{t_0}^* = X_{t_0}$. Condition on X_{t_0} and on the restriction of X_t^* to $V \setminus B_2(v)$ for all $t \in [t_0, t_1]$. Denote this conditional information by $\mathcal{F}$. Note that, conditioned on $\mathcal{F}$, the colors $X_{t_1}^*(z)$ become fully independent, for $z \in S_2(v)$.

Let $w \sim v$, and for every neighbor $z \in N(w) \setminus \{v\}$, let η_z denote the event that $X_t^*(z) = c_1$, and let ν_z denote the event that $X_t^*(z) = c_2$.

Now we argue that $\sum_z \mathbf{E}(\eta_z \mid \mathcal{F}) \approx d(w)/k$.

$$\begin{aligned} \sum_{z \sim w} \mathbf{E}(\eta_z \mid \mathcal{F}) &= \sum_{z \sim w} \mathbf{E}_a \left(\frac{\mathbf{1}\{c_1 \in A(X_a^*, z)\}}{|A(X_a^*, z)|} \right) \\ &= \mathbf{E}_a(P(X_a, w, c_1)), \end{aligned}$$

where $a = t_1 - \min\{\gamma, Cn\}$ where γ is exponentially distributed with mean n . Applying Theorem 34, this is, with high probability, in the range $d(w)/k \pm \varepsilon$. The same bound holds for $\sum_{z \sim w} \mathbf{E}(\nu_z \mid \mathcal{F})$.

Now, applying Lemma 13 to the random variables η_z , we find that for sufficiently large k ,

$$\Pr(i \text{ neighbors of } w \text{ receive color } c_1 \mid \mathcal{F}) \in e^{-d(w)/k} \frac{(d(w)/k)^i}{i!} \pm 2\varepsilon.$$

Next we argue that the colors c_1 and c_2 can be treated as independent. If we redefine the random variables η_z, ν_z so that they are independent, but have the same marginals as before, an easy coupling argument shows that, in expectation, the Hamming distance between the original and modified vectors (η_z, ν_z) (for $z \in S_2(v)$) is $O(1)$. Indeed, since the colors assigned to distinct vertices are fully independent in both cases, a Chernoff bound can be applied to prove the Hamming distance is at most $\varepsilon\Delta$ with high probability. To see this near-independence, note that, whenever a neighbor $z \in N(w)$ was last recolored, there were at least $k - \Delta = \Omega(k)$ colors to choose from at that time. Hence the maximum conditional expectation of any of the indicator variables η_z or ν_z is $O(1/k)$. Hence the probability of a disagreement in our coupling at vertex z is $O(1/k^2)$, and the result follows by summing.

It follows that, for each $w \sim v$,

$$\Pr(w \in S_{c_1, i_1} \cap S_{c_2, i_2} \mid \mathcal{F}) \in e^{-2d(w)/k} \frac{(d(w)/k)^{i_1+i_2}}{i_1!i_2!} \pm 4\varepsilon.$$

Summing over w and applying Chernoff's bound gives the desired result on our modified graph.

Finally, noting that by Lemma 31, X_{t_0} is almost surely not $2/\varepsilon$ -heavy, we can apply (15) from Corollary 25 to get the desired result for the original graph. $\square$

The proof of (3) is just a simpler variant of that of (2).

Proof of (1). This proof is analogous to the proof of Theorem 2. To begin with, let us prove the result for a fixed value of t . As usual, it will suffice to work with the graph G^* obtained from G by orienting 3 layers of edges towards v . Let $\mathcal{F}$ denote the restriction of X_t^* to the complement of $B_2(v)$. Conditioned on $\mathcal{F}$, the colors assigned to $N(v)$ become fully independent. Now we estimate the probability that a color c is available for v . This equals the product over $w \in N(v)$ of the probability that c is not assigned to w . Now, assuming w has been recolored at least once, the probability that w receives color c equals the expectation over its last recoloring time t_w of the indicator variable that c

is available for w at time t_w , divided by the number of available colors for w at time t_w . On the other hand, since with high probability, the total number of unrecolor neighbors of v is close to $\exp(-T/n) < \varepsilon$, this assumption cannot affect $|A(X_t, v)|$ by more than $2\varepsilon\Delta$.

Since $\mathbf{Pr}(X_t(w) = c) = O(1/\Delta)$, it follows that the probability that w does not receive color c is roughly $\mathbf{E}_{t_w}(\exp(-\mathbf{1}\{c \in A(X_{t_w}, w)\}/|A(X_{t_w}, w)|))$. Since the distribution of t_w is the same for all w , when we sum this over w , we can bring the expectation outside the sum, obtaining

$$\mathbf{E}_s \left(\prod_w \exp(-\mathbf{1}\{c \in A(X_s^*, w)\}/|A(X_s, w)|) \right),$$

which equals $\mathbf{E}_s(\exp(-P(X_s^*, w, c)))$. Now, by Corollary 25 (14), we know that $P(X_s^*, w, c) \approx P(X_s, w, c)$. By Theorem 34, we know $P(X_s, w, c) \approx d(w)/k$, for all times $s \geq Cn/\varepsilon$.

Thus, with high probability, $A(X_t^*, v) \leq k \exp(-d(v)/k) + \varepsilon$. Corollary 25 (13) shows that essentially the same bound applies to $A(X_t, v)$. Collecting error terms completes the proof. $\square$

For colours $c_1 \neq c_2$, $w \in V$, $v \in N(w)$, coloring X_t , let

$$\mathbf{1}\{U(X_t, w, v, c_1, c_2)\} = \begin{cases} 1 & \text{if } \{c_1, c_2\} \not\subset X_t(N(w) \setminus \{v\}) \\ 0 & \text{otherwise} \end{cases}$$

be the indicator variable for the event that w is *unblocked* for c_1 or c_2 , i.e., at least one of c_1 and c_2 does not appear on $N(w) \setminus \{v\}$.

Corollary 35. *Let $0 < \delta, \varepsilon < 1/20$, let $\Delta_0 = \Delta_0(\varepsilon, \delta)$, let $C = C(\varepsilon, \delta)$, and let $k \geq (1 + \delta)\Delta$. Let $I = [t_0, t_1]$ be a time interval with $t_0 \geq Cn \log \Delta$. Let $G = (V, E)$ have girth ≥ 7 and $\Delta > \Delta_0$. Let $(X_t)_{t \geq 0}$ be the continuous-time (or discrete-time) Glauber dynamics on G with arbitrary X_0 . Let $v \in V$ and $c, c' \in [k]$.*

$$\begin{aligned} \mathbf{Pr} \left((\exists t \in I) \sum_{w \in N(v)} \frac{\mathbf{1}\{U(X_t, w, v, c_1, c_2)\}}{|A(X_t, w)|} \geq \frac{d(v)(1 - (1 - e^{-\Delta/k})^2)}{k \exp(-\Delta/k)} + \varepsilon \right) \\ \leq \left(1 + \frac{t_1 - t_0}{n} \right) \exp(-\Delta/C). \end{aligned} \quad (21)$$

Moreover, if X_0 is not $(4/\varepsilon)$ -heavy for any color c and vertex $w \in B_R(v)$, where $R = R(\varepsilon, \delta)$, then the same conclusions hold for $t_0 \geq CRn$.

Proof. We begin by partitioning $N(v)$ into $O(1/\varepsilon)$ sets, based on degree, so that each such set S , and $w, w' \in S$, $|d(w) - d(w')| \leq \varepsilon\Delta$. Now, for each such set S , we apply (2) from Theorem 4, with colors $c_1 = c, c_2 = c'$ and $i_1 = i_2 = 0$, and also (3) with color $c_1 = c, i_1 = 0$, and again with $c_1 = c', i_1 = 0$. Applying the principle of Inclusion and Exclusion to the elements of S which are unblocked with respect to c, c' gives us

$$\begin{aligned} \sum_{w \in S} \mathbf{1}\{U(X_t, w, v, c, c')\} &= \sum_{w \in S} (\mathbf{1}\{c \in A_v^*(X_t, w)\} + \mathbf{1}\{c' \in A_v^*(X_t, w)\} - \mathbf{1}\{\{c, c'\} \subset A_v^*(X_t, w)\}) \\ &= |S_{c,0}| + |S_{c',0}| - |S_{c,c',0,0}| \\ &\approx (2e^{-d/k} - e^{-2d/k})|S| \quad \text{by (3) and (2),} \end{aligned}$$

where d is the degree of any vertex in S . Informally, we have shown that the events of being unblocked for c and c' in S are roughly independent.

Since, by Lemma 28 we also know that for any $w \in S$, $|A(X_t, w)| \geq k(e^{-d/k} - \varepsilon)$, it follows that

$$\sum_{w \in S} \frac{\mathbf{1}\{U(X_t, w, v, c, c')\}}{|A(X_t, w)|} \leq \frac{(2 - e^{-d/k})|S|}{k} + \varepsilon = \frac{(1 - (1 - e^{-d/k})^2)|S|}{ke^{-d/k}} + \varepsilon.$$

Noting that the right-hand side is maximized when $d = \Delta$, and summing over all the sets S in our partition of $N(v)$, we obtain the desired conclusion,

$$\sum_{w \in N(v)} \frac{\mathbf{1}\{U(X_t, w, v, c, c')\}}{|A(X_t, w)|} \leq \sum_S \frac{(2 - e^{-\Delta/k})|S|}{k} + \varepsilon \leq \frac{(2 - e^{-\Delta/k})d(v)}{k} + \varepsilon$$

□

Acknowledgments

I am indebted to Eric Vigoda for his great help with this paper, through many conversations and suggestions. I would also like to thank Alan Frieze and Martin Dyer for their comments. Any flaws which remain are my own responsibility.

References

- [1] D. Aldous. Random walks on finite groups and rapidly mixing Markov chains. *Seminaire de Probabilités XVII 1981/82*, Springer, 243–297, 1983.

- [2] D. Dubhashi and D. Ranjan. Balls and bins: a study in negative dependence. *Random Structures and Algorithms*, 13(2):99–124, 1998.
- [3] M. Dyer, A. Frieze. Randomly colouring graphs with lower bounds on girth and maximum degree. *Random Structures and Algorithms* **23**(2): 167–179 (2003).
- [4] M. Dyer, A. Frieze, T. P. Hayes, E. Vigoda. Randomly coloring constant degree graphs. Preliminary version in: *Proceedings, 45th Annual IEEE Symposium on Foundations of Computer Science* (FOCS 2004), 582–589.
- [5] T. P. Hayes. Randomly coloring graphs of girth at least five. In *Proceedings, 35th Annual ACM Symposium on Theory of Computing* (2003) 269–278.
- [6] T. P. Hayes and E. Vigoda. A Non-Markovian Coupling for Randomly Sampling Colorings. In *Proceedings, 44th Annual IEEE Symposium on Foundations of Computer Science* (2003) 618–627.
- [7] T. P. Hayes, E. Vigoda. Coupling with the stationary distribution and improved sampling for colorings and independent sets. *Annals of Applied Probability*, **16**(3):1297–1318 (2006).
- [8] M.R. Jerrum, A very simple algorithm for estimating the number of k -colorings of a low-degree graph, *Random Structures and Algorithms*, 7(2):157–165, 1995.
- [9] J. Jonasson. Uniqueness of uniform random colorings of regular trees. *Statistics and Probability Letters* **57** (2002) 243–248.
- [10] M. Molloy. The Glauber dynamics on colorings of a graph with high girth and maximum degree. *SIAM J. Computing* **33**(3):721–737 (2004).
- [11] L. C. Lau, M. Molloy. Randomly Colouring Graphs with Girth Five and Large Maximum Degree. In: *Proceedings of LATIN 2006*, 665–676
- [12] J. Salas and A. Sokal, Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem, *Journal of Statistical Physics*, 86(3-4):551–579, 1997.
- [13] E. Vigoda, Improved bounds for sampling colorings, *Journal of Mathematical Physics*, 41(3):1555–1569, 2000.