Electronic Colloquium on Computational Complexity, Report No. 9 (2004)

# Randomly coloring constant degree graphs

Martin Dyer * Alan Frieze† Thomas P. Hayes‡ Eric Vigoda‡

January 22, 2004

**Abstract**

We study a simple Markov chain, known as the Glauber dynamics, for generating a random $k$-coloring of a $n$-vertex graph with maximum degree $\Delta$. We prove that the dynamics converges to a random coloring after $O(n \log n)$ steps assuming $k \geq k_0$ for some absolute constant $k_0$, and either: (i) $k/\Delta > \alpha^* \approx 1.763$ and the girth $g \geq 5$, or (ii) $k/\Delta > \beta^* \approx 1.489$ and the girth $g \geq 6$. Previous results on this problem applied when $k = \Omega(\log n)$, or when $k > 11\Delta/6$ for general graphs.

## 1 Introduction

Markov Chain Monte Carlo (MCMC) is an important tool in sampling from complex distributions. It has been successfully applied in several areas of Computer Science, most notably computing the volume of a convex body [6], [13], [14] and estimating the permanent of a non-negative matrix [11].

Generating a (nearly) random $k$-coloring of a $n$-vertex graph $G = (V, E)$ with maximum degree $\Delta$ is a well-studied problem in Combinatorics [2] and Statistical Physics [16]. Jerrum

*School of Computer Studies, University of Leeds, Leeds LS2 9JT, UK. Supported by EC Working Group RAND-APX.

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA 15213, USA. Supported in part by NSF grant ccr0200945 .

‡Toyota Technological Institute at Chicago and Department of Computer Science, University of Chicago, Chicago, IL 60637, USA. Supported in part by NSF grant CCR-0237834.



ISSN 1433-8092

[10] proved that a simple, popular Markov chain, known as the *Glauber dynamics*, converges to a random $k$-coloring after $O(n \log n)$ steps, provided $k/\Delta > 2$. This led to the challenging problem of determining the smallest value of $k/\Delta$ for which a random $k$-coloring can be generated in time polynomial in $n$.

Vigoda [17] gave the first significant improvement over Jerrum's result, reducing the lower bound on $k/\Delta$ to 11/6 by analyzing a different Markov chain. There has been no success in extending Vigoda's approach to smaller values of $k/\Delta$, and it remains the best bound for general graphs.

Dyer and Frieze [5] introduced a promising approach, known as the *burn-in method*, which improved the lower bound on $k/\Delta$ for the class of graphs with large maximum degree and large girth. They reduced the bound to $k/\Delta \geq \alpha$ for any $\alpha > \alpha^*$ where

$$\alpha^* \approx 1.763$$

is the root of

$$\alpha = e^{1/\alpha}.$$

They required lower bounds on the maximum degree $\Delta = \Omega(\log n)$ and on the girth $g = \Omega(\log \Delta)$. With the same restrictions on the maximum degree and girth, Molloy [15] improved the lower bound to $k/\Delta \geq \beta$ for any $\beta > \beta^*$ where

$$\beta^* \approx 1.489$$

is the root of

$$(1 - e^{-1/\beta})^2 + \beta e^{-1/\beta} = 1.$$

However, the girth and maximum degree requirements were still significant obstacles.

The girth assumptions were the first to be (nearly) removed. Hayes [7] reduced the girth requirements to $g \geq 5$ for $k/\Delta > \alpha^*$ and $g \geq 6$ for $k/\Delta > \beta^*$. Subsequently, Hayes and Vigoda [8] (using a non-Markovian coupling) reduced the lower bound on $k/\Delta$ to $(1 + \epsilon)$ for all $\epsilon > 0$, which is nearly optimal. Their result requires girth $g \geq 9$. However, the large maximum degree restriction remained a serious bottleneck for extending the burn-in approach to general graphs. The assumption $\Delta = \Omega(\log n)$ is required in all of the improvements relying on the burn-in approach.

We dramatically improve the maximum degree assumption, only requiring $\Delta$ to be a sufficiently large constant, independent of $n$.

Before formally stating our theorem we will define the Glauber dynamics. All of the aforementioned results (except Vigoda's [17]) analyze the Glauber dynamics, which is a simple and popular Markov chain for generating a random $k$-coloring.

Let $\mathcal{K}$ denote the set of proper $k$-colorings of $G$. For technical purposes, the state space of the Glauber dynamics is $\Omega = [k]^V \supseteq \mathcal{K}$ where $[k] = \{1, 2, \ldots, k\}$. From a coloring $Z_t \in \Omega$, the evolution $Z_t \rightarrow Z_{t+1}$ is defined as follows:

**(a)** Choose $v = v(t)$ uniformly at random from $V$.

**(b)** Choose color $c = c(t)$ uniformly at random from the set of colors $[k] \setminus Z_t(N(v))$ available to $v$. The set $N(v)$ denotes the neighbors of vertex $v$.

**(c)** Define $Z_{t+1}$ by

$$Z_{t+1}(u) = \begin{cases} Z_t(u) & u \neq v \\ c & u = v \end{cases}$$

It is straightforward to verify that the stationary distribution is uniformly distributed over the set $\mathcal{K}$. For $\delta > 0$, the *mixing time* $\tau_{\text{mix}}(\delta)$ is the number of transitions until the dynamics is within variation distance at most $\delta$ of the stationary distribution, assuming the worst initial coloring $Z_0$.

We prove the following theorem.

**Theorem 1.** *Let $\alpha^* \approx 1.763$ and $\beta^* \approx 1.489$ be the constants defined earlier. For all $\epsilon > 0$, there exists $C > 0$, such that for every graph $G$ on $n$ vertices with maximum degree $\Delta$ and girth $g$, if either:*

**(a)** $k \geq \max\{(1+\epsilon)\alpha^*\Delta, C\}$ *and* $g \geq 5$, *or*

**(b)** $k \geq \max\{(1+\epsilon)\beta^*\Delta, C\}$ *and* $g \geq 6$,

*then for all $\delta > 0$, the mixing time of the Glauber dynamics on $k$-colorings of $G$ satisfies*

$$\tau_{\text{mix}}(\delta) \leq Cn \log(n/\delta).$$

Our proof analyzes a simple coupling over $T = \Theta(n)$ steps for an arbitrary pair of colorings which initially differ at a single vertex $v_0$. We prove that the expected Hamming distance after $T$ steps is at most $3/4$. We do this by breaking the analysis into two scenarios. In the advantageous scenario, during the entire $T$ steps, the Hamming distance stays small and all disagreements are close to $v_0$. If both of these events occur, after an initial burn-in period of $T_b < T$ steps, every updated vertex near $v_0$ will have certain local uniformity properties (the same properties used by [5, 15, 7]). It will then be straightforward to prove that the Hamming distance decreases in expectation over the final $T - T_b$ steps. In the disadvantageous scenario where one of the events fails, we use a crude upper bound on the Hamming distance.

# 2 Preliminaries

For $X_t, Y_t \in \Omega$, denote their difference by

$$D_t = \{v : \ X_t(v) \neq Y_t(v)\},$$

and their cumulative difference by

$$D_t^* = \bigcup_{t' \leq t} D_t.$$

Denote their Hamming distance by $H_t = |D_t|$, and let $H_t^* = |D_t^*|$. For $x, y \in V$, let $d(x, y)$ denote the length of the shortest path from $x$ to $y$ in the graph $G$. Finally, for vertex $v$, let $d(v)$ denote its degree and $N(v)$ denote its neighborhood.

We will prove convergence using path coupling [3] for $T$ steps of Glauber dynamics, where $T$ will be defined shortly. Therefore, to prove Theorem 1, for all $X_0, Y_0 \in \Omega$ where $H_0 = 1$, we need to define a coupling such that

$$\mathbf{E}(H(X_T, Y_T)) \leq \frac{3}{4} \tag{1}$$

Applying the path coupling approach of Bubley and Dyer [3], it is clear this implies the mixing time satisfies

$$\tau_{\text{mix}}(\delta) = O(T \log(n/\delta)).$$

We use Jerrum's optimal one-step coupling [10]. At every time $t$ we choose a random vertex $v = v(t)$, and update $v$ in both chains $X_t$ and $Y_t$. We maximally couple the available colors for $v$ to define $X_{t+1}(v)$ and $Y_{t+1}(v)$.

It will be useful to consider the notion of the propagation of disagreements. If for some $t, v = v(t)$ we have $X_t(v) = Y_t(v)$ and $X_{t+1}(v) \neq Y_{t+1}(v)$ then there exists a neighbor $w$ of $v$ which propagating its disagreement to $v$ in the following sense: in chain $X$ we chose color $c(t+1) = Y_t(w)$ or in chain $Y$ we chose $c(t+1) = X_t(w)$. In this way, a new disagreement at time $t$ can be traced back via a *path of disagreements* to $v_0$, the initial vertex of disagreement.

## 3 Proof of Theorem 1(a)

Fix a small positive constant $\delta > 0$ and assume that $k = (1+\delta)\alpha^*\Delta$. (For convenience we can assume that $\delta < 1/4$; larger $\delta$ are covered by Jerrum's result [10]).

Our proof makes use of various constants which we list here for convenience.

$$
\begin{array}{lll}
C_b = 200 \ln(1/\delta) & T_b = C_b n & C = 10C_b/\delta \\
T = Cn & D_{\max} = e^{20C} & R = \ln(D_{\max}) \\
\Theta_0 = (1-\delta/2)k\exp(-\Delta/k) & &
\end{array}
$$

Consider a pair $X_0, Y_0 \in \Omega$ where $D_0 = \{v_0\}$, and we will prove (1) holds. We begin with the definitions of the "bad" events of interest during our coupling period of $T$ steps. If none of these events occur, we will prove that the Hamming distance contracts in expectation over the remaining $T - T_b$ steps. If any of these events occur, we will use a crude upper bound on the Hamming distance.

Let $B_R = \{y \in V \mid d(v_0, y) \leq R\}$ denote the ball of radius $R$ centered at $v_0$ where $D_0 = \{v_0\}$. For $t > T_b$, we define the following *bad* events:

- $\mathcal{D}(t)$ denotes the event $H_t^* \geq D_{\max}$.
- $\mathcal{B}_1(t)$ denotes the event $D_t^* \not\subseteq B_R$.
- $\mathcal{B}_2(t)$ denotes the event that there exists $C_b n \leq \tau \leq t$ such that at time $\tau$ there exists $v \in B_R$ such that
$$|U(X_\tau, v)| < \Theta_0,$$
where $U(X_\tau, v) = [k] \setminus X_\tau(N(v))$.

Then we let

$$\mathcal{B}(t) = \mathcal{B}_1(t) \cup \mathcal{B}_2(t).$$

and finally we define our good event to be

$$\mathcal{G}(t) = \overline{\mathcal{D}}(t) \cap \overline{\mathcal{B}}(t).$$

For all of these events when the time $t$ is dropped, we are referring to the event at time $T$. For an event $A$, we will use the notation $\mathbf{1}_A$ to refer to the $\{0,1\}$-valued indicator variable for the event $A$, i. e.,

$$\mathbf{1}_A = \begin{cases} 1 & \text{if } A \\ 0 & \text{if } \overline{A}. \end{cases}$$

We will bound the Hamming distance by conditioning on the above events in the following manner,

$$\begin{aligned} \mathbf{E}\left(H_T\right) &= \mathbf{E}\left(H_T \mathbf{1}_{\mathcal{D}}\right) + \mathbf{E}\left(H_T \mathbf{1}_{\overline{\mathcal{D}}} \mathbf{1}_{\mathcal{B}}\right) + \mathbf{E}\left(H_T \mathbf{1}_{\mathcal{G}}\right) \\ &\leq \mathbf{E}\left(H_T \mathbf{1}_{\mathcal{D}}\right) + D_{\max} \mathbf{Pr}\left(\mathcal{B}\right) + \mathbf{E}\left(H_T \mathbf{1}_{\mathcal{G}}\right). \end{aligned} \tag{2}$$

We will bound each of the terms in (2) by $1/4$, thus ensuring that $\mathbf{E}(H_T) \leq 3/4$. This will verify (1) and Theorem 1(a).

**Lemma 2.** $\mathbf{E}\left(H_T \mathbf{1}_{\mathcal{D}}\right) < 1/4$.

**Proof** We will prove that for every integer $1 \leq D \leq n$,

$$\mathbf{Pr}\left(H_T^* \geq D\right) \leq \exp(-De^{-2C}). \tag{3}$$

For $1 \leq i \leq D$, let $t_i$ be the time at which the $i$'th disagreement is generated (possibly counting the same vertex multiple times). Denote $t_0 = 0$. Let $\eta_i := t_i - t_{i-1}$ be the waiting time for the formation of the $i$'th disagreement. Conditioned on the evolution at all times in $[0, t_i]$, the distribution of $\eta_i$ stochastically dominates a geometric distribution with success probability $i\rho$, where $\rho = \Delta/kn$. This is because at all times prior to $t_i$ we have $H_t \leq i$ and thus the set $H_t^*$ increases with probability at most $i\Delta/kn$ at each step, regardless of the history. Hence $\eta_1 + \cdots + \eta_D$ stochastically dominates the sum of independent geometrically distributed random variables with success probabilities $\rho, 2\rho, \ldots, D\rho$. Now for any real $x > 0$,

$$\Pr(\eta_i \geq x) = (1 - i\rho)^{\lceil x \rceil - 1} \geq \exp\left\{-\frac{i\rho}{1 - i\rho} x\right\} \geq e^{-2i\rho x}$$

since $i\rho \le n\rho = \Delta/k \le 4/7$.

Thus $\eta_1 + \cdots + \eta_D$ stochastically dominates the sum of exponential random variables $\zeta_1, \zeta_2, \dots,$ $\zeta_D$ with parameters $2\rho, 4\rho, \dots, D\rho$.

Now consider the problem of collecting $D$ coupons, assuming each coupon is generated by a Poisson process with rate $2\rho$. The delay between collecting the $i$'th coupon and the $i+1$'st coupon is exponentially distributed with rate $2(D-i)\rho$. Hence the time to collect all $D$ coupons has the same distribution as $\zeta_1 + \cdots + \zeta_D$. But the event that the total delay is less than $T$ is nothing but the intersection of the (independent) events that each coupon is hit in $[0,T]$. The probability of this is at most

$$\left(1 - e^{-2T\rho}\right)^D < \exp\left(-De^{-2C}\right).$$

This completes the proof of (3).

We can now bound the expected Hamming distance at time $T$ as follows:

$$\begin{aligned}
\mathbf{E}\left(H_T \mathbf{1}_{\mathcal{D}}\right) &\le \mathbf{E}\left(H_T^* \mathbf{1}_{\mathcal{D}}\right) \\
&= \sum_{D=D_{\max}}^{n} D\,\mathbf{Pr}\left(H_T^* = D\right) \\
&= D_{\max}\mathbf{Pr}\left(H_T^* \ge D_{\max}\right) + \sum_{D=D_{\max}+1}^{n} \mathbf{Pr}\left(H_T^* \ge D\right) \\
&< \sum_{D \ge D_{\max}} D_{\max}\mathbf{Pr}\left(H_T^* \ge D\right) \\
&< \sum_{D \ge D_{\max}} D_{\max}\exp(-De^{-2C}) \qquad \text{by (3)} \\
&= \frac{D_{\max}\exp(-D_{\max}e^{-2C})}{1 - \exp(-e^{-2C})} \\
&< D_{\max}\exp(3C - D_{\max}e^{-2C})
\end{aligned}$$

Since $D_{\max} = e^{20C}$, the above quantity is $e^{23C - e^{18C}} < 1/4$. This completes the proof of the Lemma. □

To prove Lemma 4 (below) we use the following lemma from Hayes [7].

**Lemma 3.** *For every $\epsilon > 0$, for all sufficiently large $\Delta \ge \Delta_0(\epsilon)$, for every graph $G = (V, E)$ having girth $\ge 5$ and maximum degree $\Delta$, for $k > (1+\epsilon)\Delta$, for every $t > 100\ln(1/\epsilon)$, all*

$w \in V$,

$$\mathbf{Pr}\left(U(X_t, w) < (1-\epsilon)k\exp(-\Delta/k)\right) \leq \exp(-\epsilon^2\Delta/100).$$

**Lemma 4.** $\mathbf{Pr}\left(\mathcal{B}\right) \leq 1/4D_{\max}$.

**Proof** Let $R = \ln(D_{\max})$. We can bound the probability of the event $\mathcal{B}_1$ by a standard paths of disagreement argument. Recall $D_{\max} = e^{20C}$.

$$\begin{aligned} \mathbf{Pr}\left(\mathcal{B}_1\right) &\leq \Delta^R \binom{T}{R} \frac{1}{(n(k-\Delta))^R} \\ &< (2Ce/R)^R \\ &< 1/8D_{\max}. \end{aligned} \tag{4}$$

To bound the probability of the event $\mathcal{B}_2$, we first bound the number of re-colorings of interest. Let

$$S = \{T_b < t \leq T : v(t) \in B_R\}.$$

For $\sigma = 100C\Delta^{R+1}$,

$$\mathbf{Pr}\left(|S| \geq \sigma\right) \leq \binom{T - T_b}{\sigma} (\Delta^{R+1}/n)^\sigma \leq (Ce\Delta^{R+1}/\sigma)^\sigma < 1/16D_{\max}. \tag{5}$$

At each time $t \in S$, by Lemma 3, with $\epsilon = \delta/2$, the desired bound on the number of available colors of $v(t)$ fails with probability at most $\exp(-\delta^2\Delta/400)$. Therefore,

$$\mathbf{Pr}\left(\mathcal{B}_2\right) \leq 1/16D_{\max} + \sigma\Delta^{R+1}\exp(-\delta^2\Delta/400) < 1/8D_{\max}. \tag{6}$$

□

**Lemma 5.** $\mathbf{E}\left(H_T \mathbf{1}_{\mathcal{G}}\right) < 1/4$.

**Proof** Let $A_t = V \setminus D_t$ be the set of vertices whose colors agree in $X$ and $Y$ at time $t$. Thus $H(X_t, Y_t) = |D_t|$.

Define $a_t(v) = |\{u \in N(v) : u \in A_t\}|$ if $v \in D_t$, and $b_t(v) = |\{u \in N(v) : u \in D_t\}|$ if $v \in A_t$.

Using Jerrum's coupling, the probability that different colors are chosen in $X$ and $Y$ for a vertex $v \in A_t$ is at most $b_t(v)/\Theta(v,t)$ where

$$\Theta(v,t) = \min\{|U(X_t,v)|, |U(Y_t,v)|\}$$

and $U(X_t,v) = [k] \setminus X_t(N(v))$ is the set of unused colors in the neighbourhood of $v$.

Similarly, if $v \notin A_t$ the probability that the same color is chosen in $X$ and $Y$ is at least $1 - (\Delta - a_t(v))/\Theta(v,t)$. Thus, given $X_t, Y_t$,

$$\mathbf{E}(H(X_{t+1},Y_{t+1})) - H(X_t,Y_t) \leq \frac{1}{n}\left(\sum_{v \in A_t} \frac{b_t(v)}{\Theta(v,t)} - \sum_{v \in D_t}\left(1 - \frac{\Delta - a_t(v)}{\Theta(v,t)}\right)\right). \qquad (7)$$

Now $\Theta(v,t) \geq k - \Delta \geq 3\Delta/4$ and

$$\sum_{v \in A_t} b_t(v) = \sum_{v \in D_t} a_t(v). \qquad (8)$$

Therefore, given $X_t, Y_t$,

$$\begin{aligned}
\mathbf{E}(H(X_{t+1},Y_{t+1}) - H(X_t,Y_t)) &\leq \frac{1}{n}\left(\sum_{v \in A_t} \frac{d(v)}{3\Delta/4} - \sum_{v \in D_t}\left(1 - \frac{\Delta - a(v)}{3\Delta/4}\right)\right) \\
&= \frac{1}{n}\sum_{v \in D_t} \frac{1}{3} \\
&= \frac{1}{3n} H(X_t,Y_t). \qquad (9)
\end{aligned}$$

This bound will only be used for the *burn-in* phase of $T_b$ steps. We will need to do significantly better for the remaining $T - T_b$ steps of an *epoch*.

Let us now re-write (7) as

$$\mathbf{E}(H(X_{t+1},Y_{t+1}) - H(X_t,Y_t)) \leq$$
$$\mathbf{E}\left(\sum_{v \in A_t} \frac{b_t(v)}{\Theta(v,t)} \mathbf{1}_{v=v(t)} - \sum_{v \in D_t}\left(1 - \frac{\Delta - a_t(v)}{\Theta(v,t)}\right)\mathbf{1}_{v=v(t)}\right). \qquad (10)$$

Note that assuming the bad event $\mathcal{B}_2(t)$ does not occur, we can replace the terms $\Theta(v,t)$ in (10) by the bound $\Theta_0$, $\mathbf{E}(\mathbf{1}_{v=v(t)})$ by $1/n$ and then use (8) to obtain

$$\begin{aligned}
\mathbf{E}(H(X_{t+1},Y_{t+1}) - H(X_t,Y_t)) &\leq \frac{1}{n}\sum_{v \in D_t}\left(-1 + \frac{\Delta}{\Theta_0}\right) \\
&\leq -\frac{\delta}{4n} H(X_t,Y_t). \qquad (11)
\end{aligned}$$

Let $t \in [T_b, T-1]$. Then

$$\begin{aligned}
\mathbf{E}\left(H_{t+1}\mathbf{1}_{\mathcal{G}(t)}\right) &= \mathbf{E}\left(\mathbf{E}\left(H_{t+1}\mathbf{1}_{\mathcal{G}(t)} \mid X_0, Y_0, \ldots, X_t, Y_t\right)\right) \\
&= \mathbf{E}\left(\mathbf{E}\left(H_{t+1} \mid X_0, Y_0, \ldots, X_t, Y_t\right)\mathbf{1}_{\mathcal{G}(t)}\right) \\
&\leq (1-\delta/4n)\mathbf{E}\left(H_t\mathbf{1}_{\mathcal{G}(t)}\right) \\
&\leq (1-\delta/4n)\mathbf{E}\left(H_t\mathbf{1}_{\mathcal{G}(t-1)}\right)
\end{aligned}$$

The above derivation deserves some words of explanation. In brief, the first equality is Fubini's Theorem, the second is because $\mathcal{G}(t)$ is determined by $X_0, Y_0, \ldots, X_t, Y_t$. The first inequality uses (11) and the definition of $\mathcal{G}(t)$, and the second inequality uses $\mathcal{G}(t) \subset \mathcal{G}(t-1)$.

By induction, it follows that

$$\mathbf{E}\left(H_T\mathbf{1}_{\mathcal{G}}\right) \leq (1-\delta/4n)^{T-T_b}\,\mathbf{E}\left(H_{T_b}\mathbf{1}_{\mathcal{G}(T_b)}\right).$$

And by (9) and the exact same argument for $t \in [0, T_b - 1]$,

$$\mathbf{E}\left(H_T\mathbf{1}_{\mathcal{G}}\right) \leq (1-\delta/4n)^{T-T_b}\,(1+1/3n)^{T_b}\,H_0. \tag{12}$$

The result follows from our choice of constants (recall that we assumed $H_0 = 1$.) □

# 4 Proof of Theorem 1(b)

Fix a small positive constant $\delta > 0$ and assume that $k = (1+\delta)\alpha^*\Delta$. (For convenience we can assume that $\delta < .3$; larger $\delta$ are covered by part (a)). Let $\rho = \delta/16$.

## 4.1 Path coupling

We use a weighted Hamming distance as was done by Molloy [15]. For $v \in V$ let its *weight* $\zeta(v)$ be defined by

$$\zeta(v) = \begin{cases} 4\rho & \text{if } d(v) < \rho\Delta \\ 1 & \text{otherwise} \end{cases}$$

Let $D_t$ be as before. For a pair $X_t, Y_t \in \Omega$ define their weighted Hamming distance as

$$\widehat{H}_t = \sum_{v \in D_t} \zeta(v)$$

We will define a coupling for all $X_0, Y_0 \in \Omega$ with $|D_0| = 1$ such that

$$\mathbf{E}\left(\widehat{H}_T\right) < \frac{3}{4}\widehat{H}_0 \tag{13}$$

The length $T$ is a function of the pair $X_0, Y_0$. If $D_0 = \{v_0\}$ then we have $T = 1$ if $d(v_0) < \rho\Delta$, and $T = cn$ if $d(v_0) \geq \rho\Delta$. Since the coupling length $T$ is not fixed, we can not immediately apply the path coupling approach. Instead we use the variable length coupling approach of Hayes and Vigoda [9] to justify this choice of $T$. While it is intuitively clear that (13) is sufficient for out purpose, we refer the reader to [9] (Theorem 3) for a proper justification.

The analysis is straightforward for pairs $X_0, Y_0 \in \Omega$ with $\widehat{H}_0 = 4\rho$. As such $X_0, Y_0$ differ only at a single vertex $v$ with $d(v) < \rho\Delta$. With probability $1/n$, the chains update $v$, and the chains are identical after the coupled transition. If the chains update a neighbor $w$ of $v$, then with probability at most $1/(k-\Delta)$, vertex $w$ receives a different color in the two chains. Thus, for $k > 4\Delta/3$,

$$\mathbf{E}\left(\widehat{H}_1 - \widehat{H}_0\right) \leq -\frac{4\rho}{n} + \frac{\rho\Delta}{(k-\Delta)n} \leq -\frac{\rho}{n} = -\frac{\delta}{16n}$$

It remains to bound $\mathbf{E}\left(\widehat{H}_T\right)$ for pairs $X_0, Y_0 \in \Omega$ with $D_0 = \{v_0\}$ with $d(v_0) > \rho\delta$. Therefore, in the remainder of the proof we consider such a pair $X_0, Y_0$. The analysis will be similar to the proof of Theorem 1(a).

## 4.2 Definitions

We now redefine the constants of the previous section. There is an additional factor of $\log\Delta$ in the definition of $C_b$. The constant $A$ in the definition of $C_b$ is the same $A$ postulated in Lemma 7 below.

| | |
|---|---|
| $C_b = A\ln(1/\delta)\log\Delta$ | $T_b = C_b n$ |
| $C = 40C_b/\delta$ | $T = Cn$ |
| $D_{\max} = e^{3C}$ | $R = \ln(D_{\max})$ |
| $\Theta_0 = (1-\delta/2)k\exp(-\Delta/k)$ | $\Psi_0 = (1-\delta/2)\Delta(1-\exp(-\Delta/k))^2$ |

## 4.3 Events

We will follow the same basic proof strategy as in part (a). We first re-define $\mathcal{B}(t)$ by adding two more bad events $\mathcal{B}_3(t), \mathcal{B}_4(t)$ defined as follows:

- $\mathcal{B}_3(t)$ denotes the event that there exists $C_b n \leq \tau \leq t$ such that at time $\tau$, there exists $v \in B_R$ such that
$$|\{w \in N(v) : X_\tau(w) \neq Y_\tau(w)\}| \geq \Delta^{1/3}.$$
- $\mathcal{B}_4(t)$ denotes the event that there exists $C_b n \leq \tau \leq t$ such that at time $\tau$, there exists $v \in B_R$ such that
$$|\{w \in N^2(v) : X_\tau(w) \neq Y_\tau(w)\}| \geq \Delta^{2/3}.$$
- For colours $c_1 \neq c_2$, $v \in V$, time $t$ and $1 \geq \rho > 0$ let
$$L(X_t, v, c_1, c_2, \rho) = |\{w \in N(v) : \{c_1, c_2\} \not\subseteq X_t(N(w) \setminus \{v\}) \text{ and } d(w) > \rho\Delta\}|$$
be the number of neighbours $w$ of $v$ which have large degree and $c_1$ and/or $c_2$ do not appear on $N(w) \setminus \{v\}$.

  $\mathcal{B}_5(t)$ denotes the event that there exists $C_b n \leq \tau \leq t$ such that at time $\tau$, $v(\tau) \in B_R$ and there exists $v \in B_R$ and colours $c_1, c_2$ such that
$$L(X_\tau, v, c_1, c_2, \rho) \geq \Psi_0.$$

Then as before we let

$$\mathcal{B}(t) = \mathcal{B}_1(t) \cup \mathcal{B}_2(t) \cup \mathcal{B}_3(t) \cup \mathcal{B}_4(t) \cup \mathcal{B}_5(t).$$

and finally we define our good event to be

$$\mathcal{G}(t) = \overline{\mathcal{D}}(t) \cap \overline{\mathcal{B}}(t).$$

## 4.4 Proof of Lemmas

Since the weighted Hamming distance is bounded by the Hamming distance we can use an analog of inequality (2). More precisely, we prove

$$\mathbf{E}\left(\widehat{H}_T\right) \leq \mathbf{E}\left(H_T \mathbf{1}_{\mathcal{D}}\right) + D_{\max}\mathbf{Pr}\left(\mathcal{B}\right) + \mathbf{E}\left(\widehat{H}_T \mathbf{1}_{\mathcal{G}}\right). \tag{14}$$

We will again bound each of the terms by $1/4$.

We first remark that the proof of Lemma 2 is unaffected by the increase in $C_b$. Restating the lemma we have

**Lemma 6.** $\mathbf{E}\left(H_T \mathbf{1}_{\mathcal{D}}\right) < 1/4$.

To prove the equivalent of Lemma 4 we need another lemma from Hayes [7]:

**Lemma 7.** *For every $\epsilon, \rho' > 0$, there exists $A > 0$ such that for every graph $G = (V, E)$ with maximum degree $\Delta > A$ and girth $\geq 6$, for $k > (1+\epsilon)\Delta$, for every $t > An\log(1/\epsilon)\log\Delta$, all $v \in V$, for every pair of colours $c_1, c_2 \in [k]$,*

$$\Pr(L(X_t, v, c_1, c_2, \rho') \geq (1+\epsilon)\Delta(1 - e^{-\Delta/k})^2) \leq \exp(-\epsilon^2\Delta/100).$$

With this we can prove

**Lemma 8.** $\Pr(\mathcal{B}) \leq 1/4D_{\max}$.

**Proof** The proofs of (4) and (6) are still valid, although now we will replace their right-hand sides by $1/20D_{\max}$.

We can bound the probability of the event $\mathcal{B}_3(t)$ by a standard paths of disagreement argument.

$$\begin{aligned}
\mathbf{Pr}\left(\mathcal{B}_3(t)\right) &\leq \Delta^R \binom{\Delta}{\Delta^{1/3}} \binom{T}{\Delta^{1/3}} \frac{1}{(n(k-\Delta))^{\Delta^{1/3}}} \\
&< \Delta^R \left( \frac{\Delta e^2 C n}{\Delta^{2/3} n (k-\Delta)} \right)^{\Delta^{1/3}} \\
&< 1/20D_{\max}.
\end{aligned}$$

We can bound the probability of the event $\mathcal{B}_4(t)$ in a similar way.

$$\begin{aligned}
\mathbf{Pr}\left(\mathcal{B}_4(t)\right) &\leq \Delta^R \binom{\Delta^2}{\Delta^{2/3}} \binom{T}{\Delta^{2/3}} \frac{1}{(n(k-\Delta))^{\Delta^{2/3}}} \\
&< \Delta^R \left( \frac{\Delta^2 e^2 C n}{\Delta^{4/3} n (k-\Delta)} \right)^{\Delta^{2/3}} \\
&< 1/20D_{\max}.
\end{aligned}$$

To bound the probability of the event $\mathcal{B}_5(t)$ we first observe that (5) continues to hold, even with the right-hand side reduced to $1/32D_{\max}$. At each time such that $v(t) \in B_R$, by Lemma 7, with $\epsilon = \delta/2, \rho' = \rho$, the desired bound on $L(X_t, v, c_1, c_2)$ fails with probability at most $\exp(-\delta^2\Delta/400)$. Therefore,

$$\mathbf{Pr}\left(\mathcal{B}_5(t)\right) \leq 1/32D_{\max} + k^2\sigma\Delta^{R+1}\exp(-\delta^2\Delta/400) < 1/16D_{\max}. \tag{15}$$

□

We now prove the equivalent of Lemma 5:

**Lemma 9.** $\mathbf{E}\left(\widehat{H}_T \mathbf{1}_{\mathcal{G}}\right) < 1/4$.

**Proof** Equation (9) continues to hold and so we must concentrate on the process after the burn-in period $t > C_b n$.

We will bound the expected change in $\widehat{H}(X_t, Y_t)$ using path coupling. Thus, let $W_0 = X_t, W_1, W_2, \ldots, W_h = Y_t$ be a sequence of colourings where $h = H(X_t, Y_t)$ and $W_{i+1}$ is obtained from $W_i$ by changing the color of one vertex $v$ from $X_t(v)$ to $Y_t(v)$. Assume that $\mathcal{G}(t)$ holds. Then for $0 \leq i \leq h$, $v \in B_R$, $c_1, c_2 \in [k]$, we have

$$\begin{aligned}
U(W_i, v) &\geq \Theta_0' = \Theta_0 - \Delta^{1/3} \\
L(W_i, v, c_1, c_2, \rho) &\leq \Psi_0' = \Psi_0 + \Delta^{2/3}.
\end{aligned}$$

We maximally couple $W_i$ and $W_{i+1}$ in one step of the Glauber Dynamics to obtain $W_i', W_{i+1}'$.

If $d(v_i) \geq \rho\Delta$ then

$$\begin{aligned}
\mathbf{E}(\widehat{H}(W_{i+1}', W_i') - \widehat{H}(W_{i+1}, W_i)) &\leq -\frac{1}{n} + \frac{\Psi_0' + 4\rho\Delta}{\Theta_0' n} \\
&\leq -\frac{\delta}{8n}.
\end{aligned}$$

If $d(v_i) < \rho\Delta$ then, for $k > 4\Delta/3$,

$$\begin{aligned}
\mathbf{E}(\widehat{H}(W_{i+1}', W_i') - \widehat{H}(W_{i+1}, W_i)) &\leq -\frac{4\rho}{n} + \frac{\rho\Delta}{(k-\Delta)n} \\
&\leq -\frac{\rho}{n} = -\frac{\delta}{16n}
\end{aligned}$$

Analogous to Inequality (11), we now have,

$$\mathbf{E}(\widehat{H}(X_{t+1}, Y_{t+1}) - \widehat{H}(X_t, Y_t)) \leq -\frac{\delta}{16n} H(X_t, Y_t). \tag{16}$$

Identical to the proof of Inequality (12), we now have,

$$\mathbf{E}\left(\widehat{H}_T \mathbf{1}_{\mathcal{G}}\right) \leq (1 - \delta/16n)^{T-T_b} (1 + 1/3n)^{T_b} H_0. \tag{17}$$

The result follows from the choice of constants (note, $H_0 = \widehat{H}_0 = 1$). □

# 5 Remarks

There seem to be several obstacles to extending the $k > (1+\epsilon)\Delta$ result of [8] to constant-degree graphs. In order to locally guarantee Molloy's burn-in property, it is necessary to run for $\Omega(n \log \Delta)$ steps. However, Hayes and Vigoda assume that $H_t = o(\Delta)$, an assumption which is almost sure to fail after $\Omega(n \log \Delta)$ steps without the full set of burn-in properties.

One approach to getting around this obstacle would be to directly define a non-Markovian coupling without using path coupling.

# References


[1] J. van den Berg and J.E. Steif, Percolation and the hard-core lattice gas model, *Stochastic Processes and their Applications* 49(2):179–197, 1994.

[2] G. R. Brightwell and P. Winkler, Random colorings of a Cayley tree. *Contemporary combinatorics*, 10:247–276, 2002.

[3] R. Bubley and M.E. Dyer, Path coupling: a technique for proving rapid mixing in Markov chains, *Proceedings of the 38th Annual IEEE Symposium on Foundations of Computer Science*, 223–231, 1997.

[4] W. Doeblin, Exposé de la théorie des chaînes simples constantes de Markov à un nombre fini d'états, *Revue Mathématique de l'Union Interbalkanique*, 2:77–105, 1938.

[5] M. Dyer, A. Frieze. Randomly colouring graphs with lower bounds on girth and maximum degree. *Random Structures and Algorithms*, 23 (2):167-179, 2003.

[6] M.E. Dyer, A.M. Frieze and R. Kannan, A random polynomial time algorithm for approximating the volume of convex bodies, *Journal of the Association for Computing Machinery* 38(1):1–17, 1991.

[7] T. P. Hayes. Randomly coloring graphs of girth at least five. Preprint. A preliminary version appears in *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, 269–278, 2003.

[8] T. P. Hayes and E. Vigoda. A Non-Markovian Coupling for Randomly Sampling Colorings. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, 618–627, 2003.

[9] T. P. Hayes and E. Vigoda. Variable length path coupling. In *Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, 96–103, 2004.

[10] M.R. Jerrum, A very simple algorithm for estimating the number of $k$-colourings of a low-degree graph, *Random Structures and Algorithms* 7(2):157–165, 1995.

[11] M.R. Jerrum, A. Sinclair and E. Vigoda, A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries. To appear in *Journal of the Association for Computing Machinery.* A preliminary version appears in *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, 712–721, 2001.

[12] M.R. Jerrum, L.G. Valiant and V.V. Vazirani, Random generation of combinatorial structures from a uniform distribution, *Theoretical Computer Science* 43(2-3):169–188, 1986.

[13] R. Kannan, L. Lovász and M. Simonovits, Random walks and an $O^*(n^5)$ volume algorithm for convex bodies, *Random Structures and Algorithms* 11(1):1–50, 1997.

[14] L. Lovász and S. Vempala, Simulated Annealing in Convex Bodies and an $O^*(n^4)$ Volume Algorithm. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, 650-659, 2003.

[15] M. Molloy, The Glauber dynamics on colorings of a graph with high girth and maximum degree. To appear in *SIAM Journal on Computing.* A preliminary version appears in *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, 91-98, 2002.

[16] 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.

[17] E. Vigoda, Improved bounds for sampling colorings, *Journal of Mathematical Physics* 41(3):1555-1569, 2000.