Randomly coloring constant degree graphs

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

June 8, 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 Δ. 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 Δ is a well-studied problem in Combinatorics [2] and Statistical Physics [16]. Jerrum [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/Δ for which a random k-coloring can be generated in time polynomial in n.

*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.

Vigoda [17] gave the first significant improvement over Jerrum's result, reducing the lower bound on k/Δ to 11/6 by analyzing a different Markov chain. There has been no success in extending Vigoda's approach to smaller values of k/Δ, 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/Δ for the class of graphs with large maximum degree and large girth. *Coupling* is a standard approach for estimating the mixing time of a Markov chain. We take two copies (X_t, Y_t) of a Markov chain $\mathcal{M}$ and then bound the variation distance d_t of the t-step distribution from the steady state via the *coupling inequality*

$$d_t \leq \Pr(X_t \neq Y_t). \tag{1}$$

We are free to choose our coupling and we endeavour to minimise the RHS of (1). Often we define a distance function *dist* between states such that $X_t \neq Y_t$ implies $dist(x_t, Y_t) \geq 1$ and then try to prove that our coupling satisfies

$$\mathbf{E}(dist(X_{t+1}, Y_{t+1}) \mid X_t, Y_t) \leq \alpha dist(X_t, Y_t) \tag{2}$$

for some $\alpha < 1$.

One must consider all possible X_t, Y_t and so it would seem that we have to take a worst-case pair here. (We should point out that *path-coupling* [3] does ameliorate this). In the burn-in method, we allow the chains to run uncoupled for a sufficient amount of time (the burn-in period) so that only *average* pairs of states can be considered. Using this idea Dyer and Frieze 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 $g = \Omega(\log \Delta)$. Under these assumptions, Dyer and Frieze proved that after the burn-in period, the colorings X_t and Y_t satisfy certain properties in the local neighborhood of *every* vertex, so called local uniformity properties. Assuming these local uniformity properties avoided the worst case pair in (2).

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/Δ 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 Δ to be a sufficiently large constant, independent of n. When Δ is constant, in a typical coloring a constant fraction of the vertices do not satisfy the desired local uniformity properties. This is the main obstacle our proof overcomes.

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 $Z_t \to Z_{t+1}$ 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 δ 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 Δ 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.

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 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

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

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

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

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 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.

For a colouring X_t and vertex v, let

$$A(X_t, v) = |[k] \setminus X_t(N(v))|,$$

denote the number of available colors for v in X_t.

For colours $c_1 \neq c_2$, $w \in V$, $v \in N(v)$, 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 $N(w) \setminus \{v\}$. We use the following lemma from Hayes [7].

Lemma 2 (Hayes [7]). *For every $\epsilon > 0$ there exists $C^* > 0$ such that for every graph $G = (V, E)$ with maximum degree $\Delta > C^*$ and for $k > 1.45\Delta$, all $X_0 \in \Omega$, for every $t > C^* n \log \Delta$,*

1. *If the girth of G is ≥ 5, then for all $v \in V$,*

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

2. *If the girth of G is ≥ 6, then for all $v \in V$, for every pair of colours $c_1, c_2 \in [k]$,*

$$\mathbf{Pr}\left(\sum_{w \in N(v)} \frac{\mathbf{1}(U(X_t, w, v, c_1, c_2))}{A(X_t, w)} \geq (1+\epsilon) \frac{\Delta(1 - e^{-\Delta/k})^2}{k \exp(-\Delta/k)}\right) \leq \exp(-\epsilon^2 \Delta/100).$$

3 Proof of Theorem

Fix a small positive constant $\delta > 0$. For part (a) assume that $k = (1+\delta)\alpha^*\Delta$, where $\delta < .3$ (otherwise, Jerrum's result [10] for $k > 2\Delta$ applies). For part (b), assume that $k = (1+\delta)\beta^*\Delta$, where $\delta < .3$ (otherwise, part (a) applies).

3.1 Definitions

Our proof makes use of various constants which we list here for convenience. The constant C^* in the definition of C_b is the same C^* postulated in Lemma 2.

$C_b = C^* \ln(1/\delta) \log \Delta$	$T_b = C_b n$
$C = 40 C_b/\delta$	$T = Cn$
$D_{\max} = e^{20C}$	$R = \ln(D_{\max})$
$\Theta_0 = (1-\delta/2) k \exp(-\Delta/k)$	$\Psi_0 = (1+\delta/2)\frac{\Delta(1-e^{-\Delta/k})^2}{k\exp(-\Delta/k)}$

3.2 Path coupling

Consider a pair $X_0, Y_0 \in \Omega$ where $D_0 = \{v_0\}$, and we will prove (3) 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 $B_R = \{y \in V \mid d(v_0, y) \le R\}$ 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 $H_t^* \ge D_{\max}$.
- $\mathcal{B}_1(t)$ denotes the event $D_t^* \not\subseteq B_R$.
- For part (a) of Theorem 1, let $\mathcal{B}_2(t)$ denote the event that there exists $C_b n \le \tau \le t$ such that at time τ, $v(\tau) \in B_R$ and there exists there exists $z \in B_R$ such that

$$A(X_\tau, z) < \Theta_0.$$

  For part (b) of Theorem 1, let $\mathcal{B}_2(t)$ denote the event that there exists $C_b n \le \tau \le t$ such that at time τ, $v(\tau) \in B_R$ and there exists $z \in B_R$ and colours c_1, c_2 such that

$$\sum_{w \in N(z)} \frac{\mathbf{1}(U(X_t, w, z, c_1, c_2))}{A(X_\tau, w)} \ge \Psi_0.$$

- $\mathcal{B}_3(t)$ denotes the event that there exists $C_b n \le \tau \le t$ such that at time τ, there exists $z \in B_R$ such that
$$|\{w \in N(z):\ X_\tau(w) \neq Y_\tau(w)\}| \geq \Delta^{1/3}.$$
- $\mathcal{B}_4(t)$ denotes the event that there exists $C_b n \le \tau \le t$ such that at time τ, there exists $z \in B_R$ such that
$$|\{w \in N^2(z):\ X_\tau(w) \neq Y_\tau(w)\}| \geq \Delta^{2/3}.$$

Then we let
$$\mathcal{B}(t) = \mathcal{B}_1(t) \cup \mathcal{B}_2(t) \cup \mathcal{B}_3(t) \cup \mathcal{B}_4(t),$$
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.

We will bound the Hamming distance by conditioning on the above events in the following manner,
$$\begin{aligned} \mathbf{E}\left(H_T\right) &= \mathbf{E}\left(H_T \mathbf{1}(\mathcal{D})\right) + \mathbf{E}\left(H_T \mathbf{1}(\overline{\mathcal{D}})\mathbf{1}(\mathcal{B})\right) + \mathbf{E}\left(H_T \mathbf{1}(\mathcal{G})\right) \\ &\leq \mathbf{E}\left(H_T \mathbf{1}(\mathcal{D})\right) + D_{\max}\mathbf{Pr}\left(\mathcal{B}\right) + \mathbf{E}\left(H_T \mathbf{1}(\mathcal{G})\right). \end{aligned} \tag{4}$$
We will bound each of the terms in (4) by 1/4, thus ensuring that $\mathbf{E}(H_T) \leq 3/4$. This will verify (3) and Theorem 1.

Lemma 3. $\mathbf{E}\left(H_T \mathbf{1}(\mathcal{D})\right) < 1/4$.

Proof We will prove that for every integer $1 \le D \le n$,
$$\mathbf{Pr}\left(H^*_T \geq D\right) \leq \exp(-De^{-2C}). \tag{5}$$

For $1 \le i \le D$, 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 η_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 $H_t \le i$ 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 $i\rho \leq n\rho = \Delta/k \leq 4/7$.

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

Now consider the problem of collecting D coupons, assuming each coupon is generated by a Poisson process with rate 2ρ. The delay between collecting the i'th coupon and the $i+1$'st coupon is exponentially distributed with rate $2(D-i+1)\rho$. 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 all coupons are generated in $[0, T]$. The probability of this is

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

This completes the proof of (5).

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

$$\begin{aligned}
\mathbf{E}\left(H_T \mathbf{1}(\mathcal{D})\right) &\leq \mathbf{E}\left(H_T^* \mathbf{1}(\mathcal{D})\right) \\
&= \sum_{D=D_{\max}}^{n} D \,\mathbf{Pr}\left(H_T^* = D\right) \\
&= D_{\max}\mathbf{Pr}\left(H_T^* \geq D_{\max}\right) + \sum_{D=D_{\max}+1}^{n} \mathbf{Pr}\left(H_T^* \geq D\right) \\
&< \sum_{D \geq D_{\max}} D_{\max}\mathbf{Pr}\left(H_T^* \geq D\right) \\
&< \sum_{D \geq D_{\max}} D_{\max} \exp(-De^{-2C}) \qquad \text{by (5)} \\
&= \frac{D_{\max} \exp(-D_{\max}e^{-2C})}{1 - \exp(-e^{-2C})} \\
&< D_{\max} \exp(3C - D_{\max}e^{-2C})
\end{aligned}$$

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

We now bound the probability of one of the bad events occuring.

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

Proof We can bound the probability of the event $\mathcal{B}_1$ by a standard paths of disagreement argument.

$$\begin{aligned}\mathbf{Pr}\left(\mathcal{B}_1\right) &\leq \Delta^R \binom{T}{R} \frac{1}{(n(k-\Delta))^R} \\ &< (2Ce/R)^R \\ &< 1/20D_{\max}. \end{aligned} \tag{6}$$

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

$$\mathbf{Pr}\left(|S| \geq \sigma\right) \leq \binom{T - T_b}{\sigma} (\Delta^{R+1}/n)^\sigma \leq (Ce\Delta^{R+1}/\sigma)^\sigma < 1/40D_{\max}. \tag{7}$$

At each time $t \in S$, by Lemma 2, with $\epsilon = \delta/2$, the desired bound on the local uniformity property of a vertex z fails with probability at most $\exp(-\delta^2\Delta/400)$. Therefore,

$$\mathbf{Pr}\left(\mathcal{B}_2\right) \leq 1/40D_{\max} + \sigma\Delta^{R+1}\exp(-\delta^2\Delta/400) < 1/20D_{\max}. \tag{8}$$

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

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

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

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

□

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

Proof We will bound the expected change in $H(X_t, Y_t)$ 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 w_i from $X_t(w_i)$ to $Y_t(w_i)$. We maximally couple W_i and W_{i+1} in one step of the Glauber Dynamics to obtain W_i', W_{i+1}'. More precisely, both chains recolor the same vertex, and maximize the probability of choosing the same new color for the chosen vertex.

Consider a pair W_i, W_{i+1}. With probability $1/n$ both chains recolor w_i to the same color, and the distance decreases by one. Consider $z \in N(w_i)$, and let $c_1 = W_i(w_i)$ and $c_2 = W_{i+1}(w_i)$. Note, color c_1 is not valid for z in W_i, however, it is valid in W_{i+1} if $c_1 \notin W_{i+1}(N(z) \setminus \{w_i\})$. Similarly, color c_2 is valid in W_{i+1}, but it is valid in W_i if $c_2 \notin W_i(N(z) \setminus \{w_i\})$. If at least one of these two cases hold, with probability at most $1/n \min\{A(W_i, z), A(W_{i+1}, z)\}$, vertex z is recolored to different colors in the two chains. Therefore, given W_i, W_{i+1},

$$\mathbf{E}(H(W_i', W_{i+1}')) - H(W_i, W_{i+1}) \leq -\frac{1}{n} + \frac{1}{n} \sum_{z \in N(w_i)} \frac{\mathbf{1}(U(W_i, z, w_i, c_1, c_2))}{\min\{A(W_i, z), A(W_{i+1}, z)\}} \tag{9}$$

In any coloring every vertex has at least $k - \Delta$ available colors. Since $k - \Delta \geq \Delta/3$, we have the following trivial bound. Given W_i, W_{i+1},

$$\mathbf{E}(H(W_i', W_{i+1}')) - H(W_i, W_{i+1}) \leq -\frac{1}{n} + \frac{\Delta}{n} \frac{3}{\Delta} = \frac{2}{n}. \tag{10}$$

Therefore, given X_t, Y_t,

$$\mathbf{E}(H(X_{t+1}, Y_{t+1})) - H(X_t, Y_t) \leq \frac{2}{n} H(X_t, Y_t). \tag{11}$$

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*.

Assume that $\mathcal{G}(t)$ holds. We will bound the distance in (9) separately for part (a) and part (b) of Theorem 1.

Suppose G has girth ≥ 5 and $k = (1 + \delta)\alpha^* \Delta, \delta < .3$. For all $0 \leq i \leq h$, $z \in B_R$, all $t \in [T_b, T - 1]$,

$$A(W_i, z) \geq A(X_t, z) - \Delta^{1/3} \geq \Theta_0 - \Delta^{1/3}$$

Hence, for $t \in [T_b, T]$, given W_i, W_{i+1},

$$\mathbf{E}(H(W_{i+1}', W_i') - H(W_{i+1}, W_i)) \leq -\frac{1}{n} + \frac{\Delta}{(\Theta_0 - \Delta^{1/3})n} \leq -\frac{\delta}{4n} \tag{12}$$

Suppose G has girth ≥ 6 and $k = (1+\delta)\beta^*\Delta, \delta < .3$. For all $0 \leq i \leq h$, $z \in B_R, c_1, c_2 \in [k]$, all $t \in [T_b, T-1]$,

$$\sum_{y \in N(z)} \frac{\mathbf{1}(U(W_i, y, z, c_1, c_2))}{\min\{A(W_i, y), A(W_{i+1}, y)\}} \leq \sum_{y \in N(z)} \frac{\mathbf{1}(U(X_t, y, z, c_1, c_2))}{A(X_t, y) - \Delta^{1/3}} + \frac{\Delta^{2/3}}{k - \Delta} \leq \Psi_0 + 5\Delta^{2/3}, \tag{13}$$

since $A(X_t, y) \geq k - \Delta > \Delta/3$. Plugging (13) into (9) proves (12) for part (b) of the Theorem.

Therefore, for parts (a) and (b) of the Theorem, for $t \in [T_b, T-1]$, given X_t, Y_t,

$$\mathbf{E}(H(X_{t+1}, Y_{t+1}) - H(X_t, Y_t)) \leq -\frac{\delta}{4n} H(X_t, Y_t). \tag{14}$$

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 (14) 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 (11) and the exact same argument for $t \in [0, T_b - 1]$,

$$\mathbf{E}\left(H_T \mathbf{1}(\mathcal{G})\right) \leq (1 - \delta/4n)^{T - T_b} (1 + 1/3n)^{T_b} H_0. \tag{15}$$

The result follows from the choice of constants (note, $H_0 = 1$). □

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.