Randomly colouring graphs with lower bounds on girth and maximum degree

Martin Dyer*

School of Computing
University of Leeds
Leeds LS2 9JT, UK

dyer@comp.leeds.ac.uk

Alan Frieze†

Department of Mathematics
Carnegie Mellon University
Pittsburgh PA15213, USA

alan@random.math.cmu.edu

19 November 2002

Abstract

We consider the problem of generating a random q-colouring of a graph $G = (V, E)$. We consider the simple Glauber Dynamics chain. We show that if the maximum degree $\Delta > c_1 \ln n$ and the girth $g > c_2 \ln \Delta$ $(n = |V|)$ for sufficiently large c_1, c_2, then this chain mixes rapidly provided $q/\Delta > \beta$, where $\beta \approx 1.763$ is the root of $\beta = e^{1/\beta}$. For this class of graphs, this beats the $11\Delta/6$ bound of Vigoda [14] for general graphs. We extend the result to random 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 volume computation [5], [12] and estimating the permanent of a non-negative matrix [10]. It was used by Jerrum [8] to generate a random q-colouring of a graph G, provided $q > 2\Delta$. This has led to the challenging problem of determining the smallest value of q for which it is possible to generate a (near)-uniform sample from the set $\mathcal{Q}$ of proper q-colourings of G in polynomial time. Vigoda [14] recently improved Jerrum's result by reducing the lower bound on q to $11\Delta/6$. In this paper we obtain an improvement in this bound for a restricted class of graphs.

For any constants $c_1, c_2 > 0$, let $\mathcal{G}(c_1, c_2)$ be the class of graphs $G = (V, E)$ such that, when $n = |V|$, G has maximum degree $\Delta \geq c_1 \ln n$ and girth $g \geq c_2 \ln \Delta$. We consider *Glauber dynamics* on the

*Research supported in part by EPSRC grant "Sharper Analysis of Randomized Algorithms", and by EC Working Group RAND-APX.

†Research supported by NSF grant CCR-9818411. The work was carried out in part during a visit by the first author to CMU, financed under grant CCR-9818411.

set $\mathcal{Q}$, for q an integer such that $\beta < q/\Delta \leq 2$, where $\beta \approx 1.763$ is the root of $\beta = e^{1/\beta}$. Let the colour set $Q = [q]$, where throughout we use $[m]$ to denote the set $\{1, 2, \ldots, m\}$. Specifically we will consider the *heat bath* dynamics, which may be described as follows. We start from an arbitrary proper q-colouring $X_0 \in \mathcal{Q}$. At step $t > 0$ of the process, in state $X_{t-1} \in \mathcal{Q}$, we choose a vertex $v_t \in V$ uniformly at random. (We will say v_t is *visited* at time t.) Then we choose j_t uniformly at random from the colours with which v_t may be properly coloured, given $X_{t-1}(V \setminus v_t)$. We recolour v_t with j_t to give $X_t \in \mathcal{Q}$. We show that this process converges to a distribution close to uniform on $\mathcal{Q}$ when $t = O(n \ln n)$ if c_1, c_2 are large enough. Note that $\beta < 11/6 \approx 1.833$, the best lower bound known on q/Δ for rapid mixing of any dynamics on degree-bounded graphs, a result due to Vigoda[14]. We prove

Theorem 1.1 *Let $\epsilon > 0$ be a constant. There exist constants $\lambda(\epsilon), \xi_1(\epsilon), \xi_2(\epsilon)$ such that the Glauber dynamics with $q \geq (\beta + \epsilon)\Delta$ colours converges to within variation distance e^{-1} from uniform in at most $\lambda n \ln n$ steps for all $G \in \mathcal{G}(c_1, c_2)$, whenever $c_1 \geq \xi_1$, $c_2 \geq \xi_2$.*

Our result easily extends to graphs in which no vertex v lies on more than c circuits, for constant c. We will show that this class includes the class of random graphs $G(n, p)$ [7] with $p = O((\ln n)^r/n)$ for constant r. The class of graphs we consider is clearly rather special, but we believe that our results may be the first step on the way to proving a corresponding theorem for more general graphs.

Counting Colourings: Generating random colourings is closely related to approximately counting them, as is well known in a more general context (Jerrum, Valiant and Vazirani [11]). For any edge e of G we can estimate the proportion ρ of proper colourings of $H = G \setminus e$ which are also proper colourings of G. This is done by (near) random sampling and relies on the fact that $\rho \geq \frac{(\beta-1)\Delta}{(\beta-1)\Delta+1}$. See Jerrum [8] for the argument. Applying this in a standard way [8], we obtain the following result.

Theorem 1.2 *Suppose $q > \max\{\beta\Delta, c_1 \ln n\}$ and the girth of G is at least $c_2 \ln \Delta$, then there is a fully polynomial randomized approximation scheme for the number of proper q-colourings of G.*

2 Basic Model

Consider the following model for the selection of colours in the neighbourhood $\mathcal{N}(v) = \{w : \{v, w\} \in E\}$ of any vertex $v \in V$.

Experiment

Let $Q = [q]$, and let $S_i \subseteq Q$ $(i \in [\Delta])$ be such that $|S_i| \geq q - \Delta > 0$. Select $s_i \in S_i$ independently and uniformly at random, and let $C = \{s_i : i \in [\Delta]\}$.

Lemma 2.1 *Let $D = |Q \setminus C|$, then* $\mathbf{E}[D] \geq q\left(1 - \dfrac{1}{q - \Delta}\right)^{\Delta(q-\Delta)/q}$.

Proof: Let $m_i = |S_i|$, and let $a_{ij} = 1$ if $j \in S_i$, $a_{ij} = 0$ otherwise. Thus $m_i = \sum_{j=1}^{q} a_{ij}$, and

$$\mathbf{E}[D] = \sum_{j=1}^{q} \prod_{i=1}^{\Delta} \left(1 - \frac{1}{m_i}\right)^{a_{ij}}. \tag{1}$$

Using the inequality between the arithmetic and geometric mean, this implies

$$\begin{aligned}
\mathbf{E}[D] &\geq q \left(\prod_{j=1}^{q} \prod_{i=1}^{\Delta} \left(1 - \frac{1}{m_i}\right)^{a_{ij}} \right)^{1/q} \\
&= q \left(\prod_{i=1}^{\Delta} \left(1 - \frac{1}{m_i}\right)^{m_i} \right)^{1/q}.
\end{aligned}$$

But $(1 - 1/m_i)^{m_i} \geq (1 - 1/(q - \Delta))^{q-\Delta}$ for all $i \in [\Delta]$, since $(1 - 1/m)^m$ increases with m, and the conclusion follows. □

Corollary 2.2 *If $q - \Delta = \Theta(\Delta)$, then $\mathbf{E}[D] \geq qe^{-\Delta/q} - O(1)$.*

Proof: Since $q - \Delta = \Theta(\Delta)$, $(1 - 1/(q - \Delta))^{q-\Delta} = e^{-1} - O(1/\Delta)$. The conclusion follows using the binomial theorem and $\Delta/q = O(1)$, $q = O(\Delta)$. □

In the context of Glauber dynamics, for a given vertex $v \in V$, we interpret $Q \setminus S_i$ as the colours used in $\mathcal{N}(w_i)$, where $\mathcal{N}(v) = \{w_i : i \in [\Delta]\}$. Thus S_i is the set of colours available for recolouring w_i. If the experiment actually models the process for selecting colours in $\mathcal{N}(v)$, then we expect more unused colours than the $q - \Delta$ worst case. This will be true, for example, if G is triangle-free and all colours in $\mathcal{N}(v)$ are chosen simultaneously. But colour selections in the dynamics take place sequentially, and it is possible that earlier choices can influence later ones. In extreme situations, this will completely invalidate our model. Consider, for example, colouring $\mathcal{N}(v)$ for v a vertex of a $(\Delta + 1)$-clique. The model predicts more than the $(q - \Delta)$ unused colours that we know must always occur. Therefore, our aim will be to show that this model is a close approximation to the truth for the class of graphs to which we have restricted our attention. Let $D_{v,t}$ be the number of colours unused in $\mathcal{N}(v)$ after step t of the Glauber dynamics. We will first show, in section 3, that our experiment is at least a good approximation in expectation.

3 Lower bounding $\mathbf{E}[D_{v,t}]$

We will use "with high probability" to mean "with probability $(1 - O(1/n^6))$". We will generally assume that high probability events actually occur, and account for the resulting errors in section 5.

Furthermore, we assume throughout that n, Δ are always sufficiently large to justify any inequality.

The following bounds, Theorem A.1.15 of [1] will be used several times. Let X be a Poisson random variable with mean μ. Then for any $\epsilon > 0$,

$$\Pr(X \leq (1-\epsilon)\mu) \leq e^{-\epsilon^2\mu/2}, \qquad \Pr(X \geq (1+\epsilon)\mu) \leq (e^{\epsilon}(1+\epsilon)^{-(1+\epsilon)})^{\mu}. \tag{2}$$

It will be helpful to modify the dynamics slightly. First, we move to a continuous time model for vertex selection. The vertex sequence will be chosen by n independent Poisson processes, one for each vertex, each with rate $1/n$. All processes are run for fixed time $T_0 = An\ln n$, where A is a large enough constant. The observed total number of steps T is then a Poisson random variable, with mean T_0. Moreover, T will be sharply concentrated around T_0. From (2) we obtain

$$\Pr(T \notin T_0 \pm n^{3/4}\ln n) = n^{-\Omega(n^{1/2})}.$$

Thus there is little difference from the original vertex choice process, using a fixed number of steps T_0, but we have gained independence of choices. We have also shown that a bound of T_0 on the mixing time of this continuous-time version implies a bound of $T_0 + n^{3/4}\log n$ on the mixing time of the discrete-time version.

Suppose now that each vertex v independently generates a list of colours, chosen uniformly and independently from Q. At any visit to v, we attempt to recolour it with the first colour on its list. If this is impossible, we remove this colour and proceed to the next. Otherwise, we recolour v and remove the colour from the list. For any $v \in V$, let ζ_v denote the number of visits to v during time T_0. Then, if $r \geq c'\ln n$ and c' is large enough,

$$\Pr(\exists v : \zeta_v > r) \leq n \sum_{s>r} \frac{T_0^s e^{-T_0/n}}{n^s s!} \leq n^{1-A} \sum_{s>r} \left(\frac{eA\ln n}{s}\right)^s < \frac{1}{n^6}. \tag{3}$$

The probability that the first colour on the list is used at any visit is at least $(q-\Delta)/q = (\beta-1)/\beta$. Thus, with high probability, the maximum list length required by any vertex is $O(\ln n)$, since the probability there is a vertex requiring a list of more than $10c'\ln n$ colours is at most

$$n\binom{10c'\ln n}{c'\ln n}\beta^{-9c'\ln n} \leq n\left(\frac{10e}{\beta^9}\right)^{c'\ln n} < \frac{1}{n^6},$$

if c' is large enough. So we assume that each vertex initially generates a list of $O(\ln n)$ colours. Given the sequence of vertices chosen at each step, these lists completely determine the dynamics, but now all random choices are made *ab initio*. We will couple processes by giving them the same colour lists and sequence of vertex choices.

Consider a fixed vertex v and (continuous) time t. Let us define the *phase* $\mathcal{P}(v,t)$ for v ending at time t by $\mathcal{P}(v,t) = [t', t]$, where $t - t' = 4n\ln\Delta$. If $t < 4n\ln\Delta$, $\mathcal{P}(v,t)$ is undefined. We first bound the probability that some neighbour of v is not visited during $\mathcal{P}(v,t)$.

Lemma 3.1 *Let $\mathcal{W}$ be the event that there is a $w \in \mathcal{N}(v)$ not visited in $\mathcal{P}(v,t)$. Then* $\Pr(\mathcal{W}) < \Delta^{-3}$.

Proof: $\Pr(\mathcal{W}) \leq \Delta e^{-4n\ln\Delta/n} = \Delta^{-3}$. □

Suppose v has degree $d \leq \Delta$ and label the vertices $w_i \in \mathcal{N}(v)$ ($i \in [d]$) in the order they are last visited during $\mathcal{P}$. (We simply write $\mathcal{P}$ for $\mathcal{P}(v,t)$ when there is no ambiguity.) Note that w_i $(i > 1)$ may be recoloured many times during $\mathcal{P}$, but the last recolouring is the one we will identify with the choice s_i in our experiment, so $s_i = X_t(w_i)$.

If $\mathrm{dist}(v,w)$ denotes the edge distance from v to w in G, let $\mathcal{N}_k(v)$ be the set of vertices v' with $\mathrm{dist}(v,v') \le k$. Let $H_v = (W_v, F_v)$ be the subgraph of G induced by $\mathcal{N}_\ell(v)$, where $\ell = \lfloor 30 \ln \Delta \rfloor$. Observe that H_v is a *tree* in view of our assumption on the girth of G. Let N be the total number of vertices visited in G during $\mathcal{P}$.

Lemma 3.2 *If* $n > 100$, $\Pr(N \ge 5n \ln \Delta) < e^{-n}$

Proof: Using (2) with $\epsilon = 1/4$, $\mu = 4n \ln \Delta$,

$$\Pr(N \ge 5n \ln \Delta) \le \left(e^{1/4}(5/4)^{-5/4}\right)^{4n \ln \Delta} < e^{-n}.$$

□

Next we bound the number of visits to vertices close to v in H_v.

Lemma 3.3 *Let* η_w *be the number of visits to* w *during* $\mathcal{P}$. *Then*

$$\Pr(\exists w \in \mathcal{N}_2(v) : \eta_w \ge 16 \ln \Delta) < \Delta^{-7}.$$

Proof: Using (2) with $\epsilon = 3$, $\mu = 4 \ln \Delta$,

$$\Pr(\eta_w \ge 16 \ln \Delta) \le \left(e^3 4^{-4}\right)^{4 \ln \Delta} < \Delta^{-10}.$$

Since there are at most $\Delta^2 + 1$ such w, the Lemma follows. □

We will assume henceforward that neither of the events whose probability we have bounded in Lemmas 3.1 and 3.3 occurs, and consider the effect of these errors at a later point in the argument.

We now consider how the colour choices s_i $(i \in [d])$ may be correlated. We will do so by coupling the process X_t with a process Y_t on the graph G', where G' is *G with all edges incident to v deleted.* We will have $Y_{t'} = X_{t'}$ and, during $\mathcal{P}$, the vertex and colour choices will be coupled as described above. Note, however, that a colour will be removed from a list if it is considered in either X or Y. Thus the remaining colour lists are always identical in X and Y at every visit to a vertex. This modification does not require us to significantly lengthen the lists. The effect of this coupling propagates through G as X_t and Y_t disagree on choice of colour. This occurs through a *path of disagreement*, as in the "disagreement percolation" technique of van den Berg and Steif [13]. First we show that, with high probability, no path of disagreement leaves H_v during $\mathcal{P}$.

Lemma 3.4 *Suppose* $c_2 \ge 100$, *and let* $\mathcal{H}$ *be the event*

$$\{\exists \tau \in \mathcal{P}, w \notin H_v : X_\tau(w) \ne Y_\tau(w)\}.$$

Then $\Pr(\mathcal{H}) < \Delta^{-20}$ *for all large enough* n.

Proof: We overestimate the probability of a disagreement at any vertex in $\mathcal{N}_1(v)$ by assuming that this occurs with probability 1. Then a path of disagreement leaving H_v starts from some $v' \in \mathcal{N}(v)$, and has length at least $\ell - 1$. Consider one of the (at most) $\Delta(\Delta-1)^{\ell-1}$ such paths from $\mathcal{N}(v)$ to the boundary of H_v along which a disagreement is propagated. Then

(a) the vertices of the path must be visited in outward sequence during $\mathcal{P}$;

(b) if w is the vertex of disagreement furthest from v at any step in $\mathcal{P}$, the colours of $\mathcal{N}(w') \setminus w$ agree for each $w' \in \mathcal{N}(w) \setminus x$ where x is the neighbour of w on the path to v, (because H_v is a tree);

(c) there are at least $q - \Delta$ available colours at w', so the probability that the coupling disagrees is at most $1/(q-\Delta)$;

(d) from (a) and (c), an event of probability at most $1/n(q-\Delta)$ occurs each time the path of disagreement is extended by one more vertex.

Thus,

$$\Pr(\mathcal{H}) < \Delta(\Delta-1)^{\ell-1}\binom{N}{\ell-1}\left(\frac{1}{n(q-\Delta)}\right)^{\ell-1} < \Delta\left(\frac{5e\ln\Delta}{(\ell-1)(\beta-1)}\right)^{\ell-1} < \Delta^{-20}, \tag{4}$$

provided $\ell \geq 30\ln\Delta$ and n is large enough. □

Henceforward we will restrict attention to the tree H_v and correct for this small error later. If $s_i^Y = Y_t(w_i)$, observe that the s_i^Y are *independent*, since they cannot "communicate" through v. Thus $D_{t,v}(Y)$ accords with the experiment of section 2, [1] and hence $\mathbf{E}[D_{t,v}(Y)] \geq qe^{-\Delta/q} - O(1)$. We will show that $\mathbf{E}[D_{t,v}(X)]$ does not differ too much from $\mathbf{E}[D_{t,v}(Y)]$. To see this, consider the colour disagreements in $\mathcal{N}(w_i)$. Vertex $u \in \mathcal{N}(w_i)$ will disagree only if a path of disagreements reaches w_i. Suppose this happens. The probability that the path is then extended to u is at most $1/(q-\Delta)$. Let ρ be the maximum over i of the total number of disagreements in $\mathcal{N}(w_i)$. Then, from Lemma 3.3,for $r \geq 80\ln\Delta$,

$$\Pr(\rho \geq r) \leq \Delta\binom{\Delta}{r}\left(\frac{16\ln\Delta}{q-\Delta}\right)^r \leq \Delta\left(\frac{16e\ln\Delta}{(\beta-1)r}\right)^r < \frac{1}{\Delta^{20}}. \tag{5}$$

Again we will assume this is actually true and correct later.

Finally, we bound the difference between $\mathbf{E}[D_{t,v}(X)]$ and $\mathbf{E}[D_{t,v}(Y)]$. Note that $s_i^X \neq s_i^Y$ only if, at its last visit, w_i chooses the colour

(a) $X_t(v)$;

(b) $X_t(u)$ or $Y_t(u)$, where $X_t(u) \neq Y_t(u)$, for some $u \in \mathcal{N}(w_i)$.

But, by (5), there are at most $1 + 160\ln\Delta$ such colours for each w_i. Thus, since there are at least $q - \Delta$ possible colour choices for w_i,

$$\Pr(X_t(w_i) \neq Y_t(w_i)) \leq \frac{1+160\ln\Delta}{q-\Delta} = O\left(\frac{\ln\Delta}{\Delta}\right). \tag{6}$$

[1]Strictly speaking we should replace the factor $\left(1-\frac{1}{m_i}\right)^{a_{ij}}$ by $\left(1-\frac{1}{m_i}+O(\Delta^{-20})\right)^{a_{ij}}$ to account for the conditioning on $\overline{\mathcal{H}}$, but this will not affect our lower estimate for $\mathbf{E}[D_{t,v}(Y)]$

Hence we conclude that

$$|\mathbf{E}[D_{t,v}(X)] - \mathbf{E}[D_{t,v}(Y)]| = O(\ln \Delta). \tag{7}$$

We must now account for the errors associated with Lemmas 3.1 to 3.4 and equation (5). The probability that any of these "bad events" occurs is $O(\Delta^{-3})$. Now $D_{v,t} \geq q - \Delta$ deterministically, so such events can only decrease $\mathbf{E}[D_{v,t}]$ by $O(\Delta^{-2})$. Thus (7) remains valid. Hence we have

$$\mathbf{E}[D_{t,v}(X)] \geq qe^{-\Delta/q} - O(\ln \Delta). \tag{8}$$

4 Concentration

We must show that the lower bound implied by our experiment holds approximately with high probability. This requires showing that some of the error probabilities from section 3 are much smaller than we needed to bound the expectation. We will show that they can all be bounded as $O(1/n^6)$, though this will involve some weakening of the conclusions.

First, we will revisit Lemma 3.1. Recall $\Delta \geq c_1 \ln n$ and $T_0 = An \ln n$.

Lemma 4.1 *Let $t \in [0, T_0]$ and $v \in V$. Let $U_{v,t}$ be the number of vertices in $\mathcal{N}(v)$ not visited during $\mathcal{P}(v,t)$. Then, if c_1 is large enough,*

$$\Pr\left(\max_{v,t} U_{v,t} > \Delta/\ln \Delta\right) < 1/n^6,$$

Proof: With high probability, $O(n \ln n)$ vertex choices occur during $[0, T_0]$. Thus for $k > 2\Delta/\ln \Delta$,

$$\begin{aligned}
\Pr\left(\max_{v,t} U_{v,t} \geq k\right) &\leq O(n^2 \ln n) \binom{\Delta}{k} e^{-4k \ln \Delta}, \\
&\leq O(n^2 \ln n) \left(\frac{e}{k\Delta^3}\right)^k, \\
&< \frac{1}{n^6}.
\end{aligned}$$

□

Henceforward we assume the conclusion of Lemma 4.1 is satisfied. We now strengthen Lemma 3.3.

Lemma 4.2 *If, for any $v \in V$ and $t \in [0, T_0]$, η_w is the number of visits to w during $\mathcal{P}(v,t)$ then, provided c_1 is large enough,*

$$\Pr(\exists v, t, w : \eta_w \geq \Delta) < \frac{1}{n^6}.$$

Proof: Using (2) with $\epsilon = \frac{\Delta}{4 \ln \Delta} - 1$, $\mu = 4 \ln \Delta$,

$$\Pr(\exists v, t, w : \eta_w \geq \Delta) \leq O(n^3 \ln n) \left(\frac{4e \ln \Delta}{\Delta}\right)^{\Delta^2/(4 \ln \Delta)} < \frac{1}{n^6}.$$

□

We now reconsider Lemma 3.4. Let $H_v^k = H_v \setminus \mathcal{N}_{k-1}(v)$. Thus $H_v^0 = H_v$. We consider a process Y_t on H_v^1 coupled with X_t inside H_v^1 as before. However, we will require that the boundary colourings of X_t and Y_t are the same, i.e. $X_t(v') = Y_t(v')$ for every leaf v' of H_v^1.

Let asequence $(x_i, c_i, t_i), i = 1, \ldots, k$ be *bad* if

- x_i is visited at time t_i where $t_1 < t_2 < \cdots < t_k \leq t$.
- $x_1, x_2, \ldots, x_k$ is a path in H_v^2.
- For $i \geq 2$, $c_i \in \{c_{i-1}, \hat{c}_{i-1}\}$ where $\hat{c}_{i-1}$ is the colour of x_{i-1} at time $t_{i-1} - 1$.

Then let $w \in \mathcal{N}(v)$ be *bad* if there is a bad sequence of length $\frac{1}{3}c_2 \ln \Delta$ for which $x_1 \in N(w), x_1 \neq v$. Let $\mathcal{B} \subseteq \mathcal{N}(v)$ be the set of bad vertices. if $w \notin \mathcal{B}$ then a disagreement at w cannot reach the boundary of H_v during $\mathcal{P}$. We now argue that $|\mathcal{B}| \leq \Delta/\ln \Delta$ with high probability.

Arguing similarly to (4) with $\ell = \frac{1}{3}c_2 \ln \Delta$ and $N = 5n \ln \Delta$ (see Lemma 3.2), we see that for $w \in N(v)$,

$$\Pr(w \in \mathcal{B}) < \Delta \binom{q}{2} (\Delta - 1)^{\ell-1} \binom{N}{\ell - 1} \left(\frac{2}{n(q - \Delta)} \right)^{\ell-1} < \Delta^{-20}$$

since $|\mathcal{N}(w)| \leq \Delta$and there are $\binom{q}{2} < 2\Delta^2$ possible colour pairs for the disagreement at u.

Since all vertex and colour choices in H_v^3 are independent, it follows that

$$\Pr(|\mathcal{B}| > \Delta/\ln \Delta) \leq \binom{\Delta}{\Delta/\ln \Delta} \left(\frac{1}{\Delta^{15}} \right)^{\Delta/\ln \Delta} \leq \left(\frac{e \ln \Delta}{\Delta^{15}} \right)^{\Delta/\ln \Delta} < \frac{1}{n^6},$$

provided c_1 is large enough.

Now condition on the colour lists for H_v^3 and the times when these vertices are visited. This defines the bad set $\mathcal{B}$. Now note that if vertex v is deleted, the colours assigned to the sets $N(w), w \notin \mathcal{B}$ are (conditionally) independent. Thus, without having conditioned the vertex or colour choices for $\mathcal{N}_2(v)$, we will partition the vertex and colour choices for $V \setminus \mathcal{N}_2(v)$ on the observed set $\mathcal{B}$. In each element of this partition we will have a specific set $\mathcal{B}$, and we may assume $|\mathcal{B}| \leq \Delta/\ln \Delta$.

We will use the following construction. For each $w \notin \mathcal{B}$, let $\mathcal{T}_w$ be the sub-tree of H_v^1 rooted at w. We will couple the process X_t on G with a process Y_t^w on a copy $\mathcal{T}'_w$ of $\mathcal{T}_w$, using the same coupling as before. Each $\mathcal{T}'_w$ $(w \notin \mathcal{B})$ will be connected to G at its outer boundary by identifying the boundary vertices with those of $\mathcal{T}_w$, the corresponding tree in G. The processes X_t, Y_t^w disagree only if $Y_t^w(w) = X_t(v)$ for any $t \in \mathcal{P}$. Notice, therefore, that the colours at the boundary of $\mathcal{T}'_w$ will always agree with those at the boundary of $\mathcal{T}_w$, since $w \notin \mathcal{B}$. We have not conditioned the vertex or colour choices of w and its neighbours in $\mathcal{T}'_w$, and these are independent for different $w \notin \mathcal{B}$.

By identical calculations to those leading to (5) and (6), we see that

$$\Pr(Y_t^w(w) \neq X_t(w)) = O\left(\frac{\ln \Delta}{\Delta} \right).$$

Here, in fact, we first condition on the vertex and choices outside $\mathcal{N}_2(v)$, but the argument leading to (5) and (6) remains valid. The random variables $Y_t^w(w)$ and events $\{Y_t^w(w) \neq X_t(w)\}$ are independent for different $w \notin \mathcal{B}$. Therefore, if D is the number of unused colours among the $X_t(w)$ ($w \in \mathcal{N}(v)$) and D' the number of unused colours among the $Y_t^w(w)$ ($w \notin \mathcal{B}$),

Lemma 4.3 $\Pr(|D' - D| > 3\Delta/\ln\Delta)) = O(1/n^6)$, *provided c_1 is large enough.*

Proof: Since there are at most $\Delta/\ln\Delta$ unvisited vertices in $\mathcal{N}(v)$ during $\mathcal{P}$, and $|\mathcal{B}| \leq \Delta/\ln\Delta$, we need only show that $|\{w \notin \mathcal{B} : Y_t^w(w) \neq X_t(w)\}| \leq \Delta/\ln\Delta$ with high probability. But

$$\begin{aligned}
\Pr(|\{w \notin \mathcal{B} : Y_t^w(w) \neq X_t(w)\}| > \Delta/\ln\Delta) &\leq \binom{\Delta}{\Delta/\ln\Delta}\left(\frac{O(\ln\Delta)}{\Delta}\right)^{\Delta/\ln\Delta} \\
&\leq \left(\frac{O(\ln^2\Delta)}{\Delta^2}\right)^{\Delta/\ln\Delta} \\
&\leq e^{-\Delta} \\
&\leq n^{-6}.
\end{aligned}$$

□

As each of the $Y_t^w(w)$ ($w \notin \mathcal{B}$) is exposed, D' changes by at most one. So we may use Hoeffding's martingale inequality [6] to bound the probability that D' is far below its expectation.

Lemma 4.4 $\Pr(D' < (1-\varepsilon)\mathbf{E}[D']) \leq 1/n^6$, *provided* $c_1 \geq 4\varepsilon^{-2}(\beta-1)^{-2}$.

Proof: Let $\mu = \mathbf{E}[D']$. For any $\theta > 0$, Hoeffding's inequality [7, p. 39] gives

$$\Pr(D' - \mu < -\theta) < \exp(-2\theta^2/\Delta).$$

Let $\theta = \varepsilon\mu$. Using $\mu \geq (\beta-1)\Delta$, which is true deterministically, we obtain

$$\Pr(D' < (1-\varepsilon)\mu) < \exp(-2\varepsilon^2(\beta-1)^2\Delta) \leq 1/n^6.$$

□

We may now account for the errors introduced by Lemmas 4.1–4.3 and give our desired concentration bound.

Lemma 4.5 *If* $t = O(n\ln n)$ *and* $c_1 = \Omega(\varepsilon^{-2})$, *then*

$$\Pr\left(\exists v, t : D_{v,t} < (1-\varepsilon)\mathbf{E}[D_{v,t}]\right) = O(1/n^3),$$

Proof: Follows easily from Lemmas 4.1–4.4 and the bound on the probability of a union of events. □

5 Convergence of the dynamics

We will use the method of *coupling*, due to Doeblin [3], to show that the dynamics approaches the uniform distribution on $\mathcal{Q}$ in $O(n\ln n)$ steps. Two copies X_t, W_t of the chain are coupled, where W_0

has the stationary distribution π, but X_0 is arbitrary. Convergence is monitored using the *coupling inequality*

$$\mathrm{d_{TV}}(\mathcal{L}(X_t), \pi)) \leq \Pr(X_t \neq W_t), \tag{9}$$

where $\mathrm{d_{TV}}$ is the *variation distance*. It is also possible to use the simpler variant of *path coupling* [2] here, but it requires some additional argument, so we will not do so.

Given states X_t, W_t of the two chains, we measure the difference between X_t, W_t by the *Hamming distance* $H(X_t, W_t)$, i.e. the number of vertices at which they disagree in colour. Suppose the number of colours available for (properly) colouring any vertex in G is always at least Θ, for some $\Theta > \Delta$. Using the coupling which selects the same random vertex in both X and W, and uses the maximal coupling on available colour choices, we will show below that

$$\mathbf{E}(H(X_{t+1}, W_{t+1})) \leq (1 - \alpha/n) H(X_t, W_t),$$

for $\alpha = (\Theta - \Delta)/\Theta > 0$. Then, since $H(X_0, W_0) \leq n$, it follows that

$$\mathrm{d_{TV}}(\mathcal{L}(X_t), \pi)) \leq \Pr(X_t \neq W_t) \leq \mathbf{E}(H(X_t, W_t)) \leq (1 - \alpha/n)^t H(X_0, W_0) \leq n e^{-\alpha t/n}.$$

Thus the Glauber dynamics will converge to variation distance e^{-2} in at most $\lceil \alpha^{-1} n(\ln n + 2) \rceil$ steps. Jerrum [8] used the deterministic bound $\Theta = q - \Delta$ to show that there is rapid mixing for the Metropolis Glauber dynamics if $q > 2\Delta$. A small improvement (to $(2 - \varepsilon)\Delta$ for some constant $\varepsilon > 0$) was achieved by Dyer, Goldberg, Greenhill, Jerrum and Mitzenmacher [4], using a probabilistic bound for Θ in the path coupling setting.

To achieve our result, we adapt an argument of Jerrum ([9], proof of Proposition 4.5) for proving rapid mixing of the heat-bath Glauber dynamics. Jerrum again uses $\Theta = q - \Delta$, and we merely need to use a better estimate for Θ.

Let $\mathcal{A}_t$ be the set of vertices whose colours agree in X and W at time t. Thus $H(X_t, W_t) = |V \setminus \mathcal{A}_t|$. Define $a(v) = |\{u \in \mathcal{N}(v) : u \in \mathcal{A}_t\}|$ if $v \notin \mathcal{A}_t$, and $d(v) = |\{u \in \mathcal{N}(v) : u \notin \mathcal{A}_t\}|$ if $v \in \mathcal{A}_t$. Then, by simple counting,

$$\sum_{v \in \mathcal{A}_t} d(v) = \sum_{v \notin \mathcal{A}_t} a(v).$$

Using the coupling described above, the probability that different colours are chosen in X and W for a vertex $v \in \mathcal{A}_t$ is clearly at most $d(v)/\Theta$. Similarly, if $v \notin \mathcal{A}_t$ the probability that the same colour is chosen in X and W is at least $1 - (\Delta - a(v))/\Theta = \alpha + a(v)/\Theta$. Thus

$$\mathbf{E}[H(X_{t+1}, W_{t+1})] - H(X_t, W_t) \leq \sum_{v \in \mathcal{A}_t} \frac{d(v)}{n\Theta} - \sum_{v \notin \mathcal{A}_t} \left(\frac{\alpha}{n} + \frac{a(v)}{n\Theta} \right) = -\frac{\alpha}{n} H(X_t, W_t).$$

Our result now follows. Using Lemma 4.5 and (8), for any $\varepsilon > 0$ we will have $\Theta \geq (1 - \varepsilon) q e^{-\Delta/q}$ for $c_1 > 8\varepsilon^{-2}(\beta - 1)^{-2}$ and n large enough. Let $\gamma = q/\Delta$. Then

$$\alpha = 1 - \frac{\Delta}{(1 - \varepsilon) q e^{-\Delta/q}} = 1 - \frac{e^{1/\gamma}}{(1 - \varepsilon)\gamma} > 0,$$

if $\varepsilon < (1 - e^{1/\gamma}/\gamma)$. Now $e^{1/\gamma}/\gamma$ decreases with γ. Thus, if $\gamma > \beta$ (the root of $e^{1/\beta} = \beta$), this will be satisfied for small enough $\varepsilon > 0$.

In order that $\mathcal{P}(v,t)$ is defined for all v, we require a (continuous time) "burn in" period of length $4n \ln \Delta$ before the coupling can be started. Using (2) this will be at most $5n \ln \Delta$ steps, with high probability. Then, after at most $\lceil \alpha^{-1} n(\ln n + 2)\rceil + \lceil 2n \ln n \rceil$ further steps, the dynamics will reach variation distance $e^{-2} + O(\ln(n)/n)$ from the uniform distribution. This is at most e^{-1} for large enough n, and completes the proof of Theorem 1.1.

6 Graphs with few circuits

Suppose every vertex $v \in V$ lies on at most c circuits of length at most $c \ln \ln n$, for some constant $c > 0$. Our arguments generalize to this case with minor modification. First H_v may no longer be a tree. For each $w \in \mathcal{N}(v)$, let C_w be the connected component of $H_v \setminus v$ containing w. Write $w \sim w'$ if $C_w = C'_w$, and observe that $\sim$ is an equivalence relation with at least $(\Delta - c)$ equivalence classes. Choose one representative from each class. These now comprise the set $\{w_i\}$ whose colours are chosen in our experiment, and the consequent change to $\mathbf{E}[D_{v,t}]$ is $O(1)$. The following modifications are all that are needed.

(a) In the proof of Lemma 3.4, the vertex w' in claim (b) may now have up to $2c$ neighbours disagreeing in X and Y, since there are up to $2c$ distinct paths from any previous disagreement to w'. This gives an extra factor $2c$ in the in the bound on $\Pr(\mathcal{H})$. This means we now need $(\ell - 1) \geq 200c \ln \Delta$ and hence $c_2 \geq 800c$, say.

(b) In the proof of (5), there could be $2c$ paths from w_i to $u \in \mathcal{N}(w_i) \setminus \{v\}$. Thus the numerator of the bracketed term again requires an additional factor $2c$, so that (5) requires $r \geq 160c \ln \Delta$.

The argument then proceeds exactly as before, and Theorem 1.1 follows for this slightly wider class of graphs.

7 Random graphs

The generalization of section 6 allows us to prove rapid mixing with more than $\beta\Delta$ colours for random graphs with average degree $(\ln n)^{\Theta(1)}$. In such graphs it is easy to show that the maximum degree Δ is also $(\ln n)^{\Theta(1)}$. We will complete the proof by showing that, in random graphs of this density, there is only a small probability that any vertex lies on *two* circuits of size $O(\ln \ln n) = O(\ln \Delta)$.

Lemma 7.1 *Let $p = (\ln n)^{O(1)}/n$, and $\mathcal{D}_v$ be the event that $v \in G(n,p)$ lies on two distinct circuits of length $O(\ln \ln n)$. If $\mathcal{D} = \bigcup_{v \in V} \mathcal{D}_v$, then $\Pr(\mathcal{D}) \to 0$ as $n \to \infty$.*

Proof: If $\mathcal{D}_v$ occurs, let C_1, C_2 be any two such circuits. Since $\{v\} \subseteq C_1 \cap C_2$, $C = C_1 \cup C_2$ is a connected component of G with at least two distinct circuits and size $O(\ln \ln n)$. Therefore the number of edges in C must exceed its number of vertices by at least one. We will bound the probability of $\mathcal{D}$ by the probability that any such C exists in $G(n,p)$. Thus, letting $k \leq r = O(\ln \ln n)$

be the number of vertices in C,

$$\Pr(\mathcal{D}) \leq \sum_{k=1}^{r} \binom{n}{k} \binom{\binom{k}{2}}{k+1} p^{k+1} = \frac{(\ln n)^{O(\ln \ln n)}}{n} = o(1).$$

□

References

[1] N. Alon, J.H. Spencer and P. Erdös, *The probabilistic method*, John Wiley & Sons, New York, 1992.

[2] 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* (1997), 223–231.

[3] W. Doeblin, Exposé de la théorie des chaines simples constantes de Markov á un nombre fini d'états, *Revue Mathématique de l'Union Interbalkanique* 2 (1933), 77–105.

[4] M.E. Dyer, L. Goldberg, C. Greenhill, M.R. Jerrum and M. Mitzenmacher, An extension of path coupling and its application to the Glauber dynamics for graph colourings, *Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms* (2000), 616–624.

[5] 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 (1991), 1–17.

[6] W. Hoeffding, Probability inequalities for sums of bounded random variables, *Journal of the American Statistical Association* 58 (1963), 13–30.

[7] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, John Wiley & Sons, New York, 2000.

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

[9] M.R. Jerrum, *Counting, sampling and integrating: algorithms and complexity*, Lectures in Mathematics–ETH Zürich, Birkhäuser, to appear January 2003.

[10] M.R. Jerrum, A. Sinclair and E. Vigoda, A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries, *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms* (2001), 712–721.

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

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

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

[14] E. Vigoda, Improved bounds for sampling colorings, *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science* (1999), 51–59.