On randomly colouring locally sparse graphs

Alan Frieze and Juan Vera[†]

Department of Mathematics
Carnegie Mellon University
Pittsburgh PA15213, USA

alan@random.math.cmu.edu
jvera@andrew.cmu.edu

received Aug 30, 2004, *accepted Apr 13, 2006*.

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 for all $v \in V$ the average degree of the subgraph H_v induced by the neighbours of $v \in V$ is $\ll \Delta$ where Δ is the maximum degree and $\Delta > c_1 \ln n$ then for sufficiently large c_1, this chain mixes rapidly provided $q/\Delta > \alpha$, where $\alpha \approx 1.763$ is the root of $\alpha = e^{1/\alpha}$. For this class of graphs, which includes planar graphs, triangle free graphs and random graphs $G_{n,p}$ with $p \ll 1$, this beats the $11\Delta/6$ bound of Vigoda [20] for general graphs.

Keywords: Counting Colourings, Sampling, Markov Chains

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 [3], [15], [16] and estimating the permanent of a non-negative matrix [12]. It was used by Jerrum [10] 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. We cannot expect the chain to mix for $q \leq \Delta + 1$ and at present it is unknown as to whether or not it mixes rapidly for say $q = \Delta + 2$. Vigoda [20] improved Jerrum's result by reducing the lower bound on q to $11\Delta/6$. This is still the best lower bound on q for general graphs.

The lack of complete success on the general problem has led to the analysis of restricted classes of graphs. Suppose that we consider *Glauber dynamics* on the set $\mathcal{Q}$. 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

[†]Research supported by NSF grant CCR-0200945.

at random. 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 $X_t \in \mathcal{Q}$.

Dyer and Frieze [2] considered this process restricted to the class of graphs $\mathcal{G}(c_1, c_2)$: the set of graphs with n vertices, maximum degree $\Delta \geq c_1 \log n$ and girth $g \geq c_2 \log \Delta$. They showed using the idea of "burn-in" that for c_1, c_2 sufficiently large, Glauber Dynamics mixed in $O(n \log n)$ time, provided $q > \alpha\Delta$ where $\alpha \approx 1.763$ is the root of $\alpha = e^{1/\alpha}$. Molloy [17] improved this result by reducing the lower bound on q to being more than $\beta\Delta$ where $\beta \approx 1.489$ is the root of $(1 - e^{-1/\beta})^2 + \beta e^{-1/\beta} = 1$. The girth asumptions were then relaxed by Hayes [7] to $g \geq 5$ for $k/\Delta > \alpha$ and $g \geq 6$ for $k/\Delta > \beta$. Subsequently, Hayes and Vigoda [8] made considerable progress, using a non-Markovian coupling, and 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. This was replaced by $\Delta \geq \Delta_0$ in Dyer, Frieze, Hayes and Vigoda [5], with the same restrictions on girth as in [7]. Dyer, Flaxman, Frieze and Vigoda [4] show that for sparse random graphs, the number of colours required for rapid mixing is of order the average rather than maximum degree **whp**. Goldberg, Martin and Paterson [6] prove results on the related notion of strong spatial mixing.

In this paper we avoid girth restrictions and consider locally sparse graphs instead. We say that a graph $G = (V, E)$ is γ-*locally sparse* if for all $v \in V$, the average degree of the graph induced by the neighbourhood $N(v)$ is at most γ. Thus planar graphs are always 6-locally-sparse and triangle free graphs are 0-locally-sparse.

Theorem 1.1 *Suppose that $q \geq (\alpha + \epsilon)\Delta$ where ϵ is a small positive constant. Let G be an n-vertex γ-locally sparse graph with $\gamma \leq \epsilon^2\Delta/10$ and $\Delta \geq c_1 \log n$. If $c_1 = c_1(\epsilon)$ is sufficiently large then the Glauber dynamics converges to within variation distance e^{-1} from uniform over $\mathcal{Q}$ in at most $O(n \ln n)$.*

Notice that if $G = G_{n,p}$ and $\frac{c_1 \log n}{n} \leq p \leq \epsilon^2/11$ then **whp** G satisfies the conditions of the theorem. Note also that the chromatic number of a triangle-free graph is $O(\Delta/\log \Delta)$ – see Johansson [14] or Molloy and Reed [18] or Alon, Krivelevich and Sudakov [1] or Vu [21].

Our proof uses coupling and relies on a recent idea from Hayes and Vigoda [9] that utilises the fact that we can couple against the steady state distribution of the chain. Note that the theorem generalises Theorem 4 of [9].

In what follows we will assume that n is sufficiently large and ϵ is sufficiently small to satisfy our inequalities.

2 Preliminaries

We will consider two copies of Glauber Dynamics, $(X_t,\, t \geq 0)$ and $(Y_t,\, t \geq 0)$. Here X_0 is an arbitrary colouring and Y_0 is chosen from the uniform (*stationary*) distribution over $\mathcal{Q}$. At time t, the Hamming distance between X_t, Y_t is defined by

$$H(X_t, Y_t) = \sum_{v \in V} 1_{X_t(v) \neq Y_t(v)}.$$

We will couple the two processes as in Jerrum [10]. Here v_t is the same in both processes and then the choice of colours is maximally coupled. For vertex w let

$$A(X_t, w) = \{c \in [q] :\ c \notin X_t(N(w))\}$$

be the set of colours available to colour w in X_t if $v_t = w$.

Let $a(X_t, w) = |A(X_t, w)|$ and define the terms $A(Y_t, w), a(Y_t, w)$ analogously.

It is shown in [9] that

$$\mathbf{E}(H(X_{t+1}, Y_{t+1}) - H(X_t, Y_t)) \le -\frac{1}{n} H(X_t, Y_t) + \frac{1}{n} \sum_{w \in V} \frac{|\{u \in N(w) : X_t(u) \neq Y_t(u)\}|}{\max\{a(X_t, w), a(Y_t, w)\}}. \quad (1)$$

We will show that for $w \in V$ and $\delta = \epsilon/10$,

$$\Pr(a(Y_t, w) \le \Delta/(1-\delta)) \le n^{-4}. \quad (2)$$

Assuming that $a(Y_t, w) \ge \Delta/(1-\delta)$ in (1) we get

$$\begin{aligned} \mathbf{E}(H(X_{t+1}, Y_{t+1}) - H(X_t, Y_t)) &\le -\frac{1}{n} H(X_t, Y_t) + \frac{1}{n} \frac{H(X_t, Y_t)\Delta}{\Delta/(1-\delta)} \\ &\le -\frac{\delta}{n} H(X_t, Y_t). \end{aligned}$$

So conditional on an event of probability $1 - O(n^{-3})$, we have

$$\mathbf{E}(H(X_{t+1}, Y_{t+1}) \mid X_t, Y_t) \le \left(1 - \frac{\delta}{n}\right) H(X_t, Y_t).$$

Thus if $T = n(1 + \ln n)\delta^{-1}$ then conditional on an event of probability $1 - O(n^{-2} \log n)$, we have

$$\mathbf{E}(H(X_T, Y_T)) \le e^{-1}$$

and so unconditionally

$$\mathbf{E}(H(X_T, Y_T)) \le e^{-1} + o(1).$$

Hence the mixing time of the Glauber Dynamics is $O(n \ln n)$ as claimed.

3 Bounding the number of available colours

Fix $v \in W$ and let H_v be the subgraph of G induced by $N(v)$. Let $B(v)$ be the vertices of $N(v)$ that have degree at least $\gamma\delta^{-1}$ in H_v. Note that $\gamma\delta^{-1} \le \epsilon\Delta$ and

$$|B(v)| \le \delta |N(v)|, \quad (3)$$

since G is γ-locally-sparse.

Let

$$N^*(v) = N(v) \setminus B(v) = \{w_1, w_2, \ldots, w_d\}.$$

Now let let us fix the colours $\kappa(v)$ used at

$$v \in W_v = V \setminus N^*(v).$$

Let us use the term *allowable* for colorings of $N^*(v)$ which respect this conditioning. Let Ω be the set of allowable colourings of $N^*(v)$.

Let $a^*(\sigma, v)$ be the number of colours not used on $N^*(v)$. Note that (3) implies

$$a(\sigma, v) \geq a^*(\sigma, v) - \delta|N(v)|. \tag{4}$$

Now consider the following process $\mathcal{P}_\sigma$ for producing an allowable colouring of H_v. Here $\sigma \in \Omega$. We let $\sigma_0 = \sigma$ and for $j = 1, 2, \ldots, d$ let σ_j be obtained from σ_{j-1} as follows: Keep $\sigma_j(w_k) = \sigma_{j-1}(w_k)$ for $k \neq j$ and choose $\sigma_j(w_j)$ randomly from what is available to it.

Let Z_σ be the number of colours not appearing on a vertex in $N^*(v)$ if we start with $\sigma_0 = \sigma$.

Lemma 3.1 *If σ is chosen uniformly from Ω then for any $c > 0$,*

$$\Pr(a^*(\sigma, v) \geq c) = \Pr(Z_\sigma \geq c).$$

Proof We first prove that

$$\text{If } \sigma_0 \text{ is chosen uniformly from } \Omega \text{ then } \sigma_d \text{ is also uniform over } \Omega. \tag{5}$$

We do this by induction on j, with base case $j = 0$.

$$\begin{aligned}
\Pr(\sigma_j = \sigma) &= \sum_{\sigma' \in \Omega} \Pr(\sigma_j = \sigma \mid \sigma_{j-1} = \sigma') \Pr(\sigma_{j-1} = \sigma') \\
&= \frac{1}{|\Omega|} \sum_{\sigma' \sim \sigma} \Pr(\sigma_j = \sigma \mid \sigma_{j-1} = \sigma')
\end{aligned}$$

Here $\sigma' \sim \sigma$ if σ, σ' differ only at w_j.

$$\begin{aligned}
&= \frac{1}{|\Omega|} \sum_{\sigma' \sim \sigma} \frac{1}{|\{\sigma' : \sigma' \sim \sigma\}|} \\
&= \frac{1}{|\Omega|}.
\end{aligned}$$

Now $a^*(\sigma_d, v) = Z_{\sigma_0}$ and so

$$\Pr(a^*(\sigma_d, v) \geq c) = \Pr(Z_{\sigma_0} \geq c)$$

and the lemma follows from (5) □

For $w \in N^*(v)$ let

$$L(w) = [q] \setminus \{\kappa(u) : u \in N(w) \setminus N^*(v)\}$$

be the colours not specifically barred from w by the current conditioning. Then let

$$L^*(w_j) = [q] \setminus \{\sigma_{j-1}(u) : u \neq w_j\} \qquad \text{for } j = 1, 2, \ldots, d$$

be the colours available to w_j when it is re-coloured by σ_j.

We will first estimate the (conditional) expectation of Z_σ for arbitrary σ. Suppose that $x \in [q]$. Let $\theta_{x,j} = 1_{x \in L(w_j)}$ and let $\theta^*_{x,j} = 1_{x \in L^*(w_j)}$. Then we have

$$\begin{aligned}
\Pr(x \notin \sigma_d(N^*(v))) &= \prod_{j=1}^{d} \Pr(\sigma_d(w_j) \neq x \mid \sigma_d(w_i) \neq x,\, 1 \leq i < j) \\
&= \prod_{j=1}^{d} \mathbf{E}\left(\left(1 - \frac{1}{|L^*(w_j)|}\right)^{\theta^*_{x,j}}\right) \\
&\geq \prod_{j=1}^{d} \left(1 - \frac{1}{|L(w_j)| - \gamma\delta^{-1}}\right)^{\theta_{x,j}}
\end{aligned}$$

since $|L^*(w_j)| \geq |L(w_j)| - \gamma\delta^{-1}$ and $L^*(w_j) \subseteq L(w_j)$ implying $\theta^*_{x,j} \leq \theta_{x,j}$.

Then, following [2],

$$\begin{aligned}
\mathbf{E}(Z_\sigma) &\geq \sum_{x \in [q]} \prod_{j=1}^{d} \left(1 - \frac{1}{|L(w_j)| - \gamma\delta^{-1}}\right)^{\theta_{x,j}} \\
&\geq q \left(\prod_{x \in [q]} \prod_{j=1}^{d} \left(1 - \frac{1}{|L(w_j)| - \gamma\delta^{-1}}\right)^{\theta_{x,j}}\right)^{1/q} \\
&= q \left(\prod_{j=1}^{d} \left(1 - \frac{1}{|L(w_j)| - \gamma\delta^{-1}}\right)^{|L(w_j)|}\right)^{1/q} \\
&\geq q \exp\left\{-\frac{1}{q} \sum_{j=1}^{d} \frac{|L(w_j)|}{|L(w_j)| - 1 - \gamma\delta^{-1}}\right\}, \qquad \text{using } 1 - x \geq e^{-x/(1-x)} \text{ for } 0 < x < 1, \\
&\geq q \exp\left\{-\frac{\Delta}{q} \cdot \frac{q - \Delta}{q - \Delta - 1 - \gamma\delta^{-1}}\right\} \\
&\geq \left(1 + \frac{\epsilon}{2}\right) \Delta. \qquad (6)
\end{aligned}$$

(If $f(x) = xe^{-1/x}$ then $f(\alpha) = 1$ and $f'(\alpha) \sim .891$.)

We will now prove that for all $\sigma \in \Omega$, Z_σ is concentrated around its mean via the use of the Azuma-Hoeffding martingale inequality. To this end, let $x_1, x_2, \ldots, x_d$ be the colours assigned to $w_1, w_2, \ldots, w_d$. Thus we can write $Z_\sigma = Z_\sigma(x_1, x_2, \ldots, x_d)$. Now let

$$Z_{\sigma,i} = Z_{\sigma,i}(x_1, x_2, \ldots, x_i) = \mathbf{E}(Z \mid x_1, x_2, \ldots, x_i).$$

We will show next that for all feasible colours $x_1, x_2, \ldots, x_i, x'_i$ that

$$|Z_{\sigma,i}(x_1, x_2, \ldots, x_{i-1}, x_i) - Z_{\sigma,i}(x_1, x_2, \ldots, x_{i-1}, x^*_i)| \leq 2. \qquad (7)$$

The aforementioned inequality will then imply that for any $t \geq 0$,

$$\Pr(Z_\sigma - \mathbf{E}(Z_\sigma) \leq -t) \leq e^{-t^2/(2d)}$$

and then taking $t = \epsilon\Delta/4$ and using (6) we get

$$\Pr\left(Z_\sigma \leq \left(1 + \frac{\epsilon}{4}\right)\Delta\right) \leq e^{-\epsilon^2\Delta/32}.$$

This together with Lemma 3.1 and (4) implies (2).

To prove (7), fix $i, x_1, x_2, \ldots, x_i, x_i^*$. In one instance of $\mathcal{P}_\sigma$ we start by colouring $w_1, w_2, \ldots, w_i$ with $x_1, x_2, \ldots, x_i$ to produce colouring τ. In another instance we start by colouring $w_1, w_2, \ldots, w_i$ with $x_1, x_2, \ldots, x_i^*$ to produce colouring τ^*.

We couple these two constructions in order to minimise the expected difference in the number of vertices U with a different colour. A *paths of disagreement* argument gives that

$$\mathbf{E}(U) \leq 1 + \sum_{j=i+1}^{d} \left(\frac{\gamma\delta^{-1}}{|L(w_j)| - \gamma\delta^{-1}}\right)^{j-i} \leq 2 \tag{8}$$

and (7) follows. □

Explanation of (8)**:** We claim that if c_j, c_j^* is the colour of v_j in σ_d, σ_d^* respectively, then

$$\Pr(c_j \neq c_j^*) \leq \left(\frac{\gamma\delta^{-1}}{|L(w_j)| - \gamma\delta^{-1}}\right)^{j-i}.$$

This is because if $c_j \neq c_j^*$ then there is a path of disagreements $v_{i_1}, v_{i_2}, \ldots, v_{i_s}$ where $i = i_1 < i_2 < \cdots < i_s = j$ such that $c_{i_r} \neq c_{i_r}^*$ for $1 \leq r \leq s$. There are at most $(\lambda\delta^{-1})^{j-i}$ such paths and each has probability at most $(|L(w_j)| - \gamma\delta^{-1})^{i-j}$ of all vertices being coloured differently. □

References

[1] N. Alon, M. krivelevich and B. Sudakov, Coloring graphs with sparse neighborhoods, *Journal of Combinatorial Theory, Ser. B* 77 (1999) 73-82.

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

[3] 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.

[4] M.E. Dyer, A. Flaxman, A.M. Frieze and E. Vigoda, Randomly colouring sparse random graphs.

[5] M.E. Dyer, A.M. Frieze, T. P. Hayes and E. Vigoda, Randomly coloring constant degree graphs.

[6] L. Goldberg, R. Martin and M. Paterson, Strong spatial mixing for graphs with fewer colours.

[7] T. P. Hayes, Randomly coloring graphs of girth at least five, *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, (2003) 269–278

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

[9] T. P. Hayes and E. Vigoda, Coupling with the stationarity distribution and improved sampling fror colorings and independent sets.

[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 (1995), 157–165.

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

[12] 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.

[13] 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.

[14] A. Johansson, Asymptotic choice number for triangle-free graphs, DIMACS technical report 1996.

[15] 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.

[16] L. Lovász and S. Vempala, Simulated annealing in convex bodies and an $O^*(n^4)$ volume algorithm, to appear in JCSS.

[17] M. Molloy, The Glauber dynamics on colorings of a graph with high girth and maximum degree. To appear in *SIAM Journal on Computing*.

[18] M. Molloy and B. Reed, Graph Colouring and the Probabilistic Method, Springer, 2002.

[19] 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.

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

[21] V. Vu, An upper bound on the list chromatic number of locally sparse graphs, *Combinatorics, Probability and Computing* 11 (2002), 103-111