Very rapid mixing of the Glauber dynamics for proper colourings on bounded-degree graphs

Martin Dyer* Catherine Greenhill† Mike Molloy‡

23 August 2000

Abstract

Recent results have shown that the Glauber dynamics for graph colourings has optimal mixing time when (i) the graph is triangle-free and D-regular and the number of colours k is a small constant fraction smaller than $2D$, or (ii) the graph has maximum degree D and $k = 2D$. We extend both these results to prove that the Glauber dynamics has optimal mixing time when the graph has maximum degree D and the number of colours is a small constant fraction smaller than $2D$.

1 Introduction

Let $G = (V, E)$ be a graph and let C be a set of k colours. A proper k-colouring of G is a map $X : V \to C$ such that $X(v) \neq X(w)$ whenever v and w are neighbours (i.e. whenever $\{v, w\} \in E$). Consider the problem of almost uniformly sampling from the set of all proper k-colourings of a graph of maximum degree D. This is a fundamental combinatorial problem, and also has applications in statistical physics (see [8, 7]). A standard approach to the problem of approximate sampling is the Markov chain Monte Carlo approach. Here a Markov chain with uniform stationary distribution is simulated for sufficiently many steps, and the final state is returned as a sample from the set of proper colourings. For this approach to be efficient, the number of steps required for the chain to approach its stationary distribution must be small. (Note that efficient approximately uniform sampling can be used to perform efficient approximate counting of proper colourings; see [8].) We say that the chain is rapidly mixing if the mixing time is bounded above by some polynomial in $|V| = n$ (this will be made more precise in Section 2 below).

*School of Computing, University of Leeds, Leeds LS2 9JT, U.K. Supported by ESPRIT Working Group RAND-APX.

†Department of Mathematics and Statistics, University of Melbourne, Parkville VIC 3010 AUSTRALIA. Supported by an Australian Research Council Postdoctoral Fellowship.

‡Department of Computer Science, University of Toronto, Toronto, CANADA. Supported by an NSERC Research Grant.

The Glauber dynamics is the simplest possible Markov chain for graph colourings. Here a vertex v and a colour c are chosen uniformly at random at each step, and v is recoloured with c if c is not equal to the colour of any neighbours of v. The name 'Glauber dynamics' comes from the statistical physics literature. Jerrum [8] and independently, Salas and Sokal [11] showed that the Glauber dynamics is rapidly mixing when $k \geq 2D$. In fact, Jerrum proved that the Glauber dynamics mixes in $O(n \log n)$ steps when $k > 2D$. Bubley and Dyer [2] noted that the Glauber dynamics is also rapidly mixing when $k = 2D$. They could show an $O(n^3)$ bound on the mixing time.

Next, effort was expended to find a Markov chain for graph colourings which was rapidly mixing when fewer than k colours were used. Bubley, Dyer and Greenhill [3] described a Markov chain which was rapidly mixing when $D = 3$ and $k = 5$, or when $k = 7$ and the graph was 4-regular and triangle-free. The proof was computer-assisted and infeasible for large degrees. Vigoda [12] made a breakthrough when he showed that his chain, which changes the colours of up to six vertices per step, is rapidly mixing when $k \geq 11D/6$. Using the comparison technique of Diaconis and Saloff-Coste [5] he was able to show that the Glauber dynamics is also rapidly mixing on this range. In fact, an $O(n \log n)$ upper bound on the mixing time of Vigoda's chain gave an $O(n^2 \log n)$ upper bound on the mixing time of the Glauber dynamics, for $k \geq 11D/6$.

However, it is intuitively clear that at least $\Omega(n \log n)$ steps are required for rapid mixing of the Glauber dynamics, since this is the number of steps required before each vertex has been visited. A rigorous argument to support this intuition can be found in [6], in the case of the graph with no edges. Therefore the challenge became proving optimal mixing of the Glauber dynamics for some values of k in the range $k \leq 2D$.

So far, there have been two results in this direction. Dyer et al. [6] proved that the Glauber dynamics has optimal mixing time for triangle-free D-regular graphs, when k satisfies $(2-\xi)D \leq k \leq 2D$ for some small positive constant ξ (specifically, $\xi = 8 \times 10^{-9}$). Molloy [9] showed that the Glauber dynamics has optimal mixing time for graphs with maximum degree D when $k = 2D$. Both results involved applying path coupling over multiple steps, and analysing the coupling at one or more specially-defined stopping times. More will be said about these approaches below.

The contribution of this paper is to extend both these results as follows:

Theorem 1 *For any graph G with n vertices and maximum degree Δ, the Glauber dynamics on the k-colourings of G mixes in $O(n \log n)$ time so long as $k \geq (2 - \eta)D$, where $\eta > 0$ is a small positive constant.*

We prove Theorem 1 for $\eta = 10^{-12}$. However no attempt is made to optimize the constants used in this paper, and it is easy to improve η significantly.

The remainder of the paper is organised as follows. Definitions relating to the mixing time are given in Section 2, together with results relating to path coupling, including the use of stopping times. Some preliminary results involving path coupling for graph colourings are given in Section 3. Various cases are identified and analysed in Section 4. The results are combined in Section 5.

2 Notation, and path coupling with stopping times

First we give some standard notation relating to the mixing time of Markov chains. Let Ω be a finite set and let $\mathcal{M}$ be a Markov chain with state space Ω, transition matrix P and unique stationary distribution π. If the initial state of the Markov chain is x then the distribution of the chain at time t is given by $P_x^t(y) = P^t(x,y)$. The *total variation distance* of the Markov chain from π at time t, with initial state x, is defined by

$$\mathrm{d_{TV}}(P_x^t,\, \pi) = \frac{1}{2}\sum_{y\in\Omega} |P^t(x,y) - \pi(y)|.$$

Following Aldous [1], let $\tau_x(\varepsilon)$ denote the least value T such that $\mathrm{d_{TV}}(P_x^t,\, \pi) \le \varepsilon$ for all $t \ge T$. The *mixing time* of $\mathcal{M}$, denoted by $\tau(\varepsilon)$, is defined by $\tau(\varepsilon) = \max\{\tau_x(\varepsilon) : x \in \Omega\}$. A Markov chain is be said to be *rapidly mixing* if the mixing time is bounded above by some polynomial in n and $\log(\varepsilon^{-1})$, where n is a measure of the size of the elements of Ω. Throughout this paper all logarithms are to base e.

There are relatively few methods available to prove that a Markov chain is rapidly mixing. One such method is *coupling.* A coupling for $\mathcal{M}$ is a stochastic process $(X_t,\, Y_t)$ on $\Omega\times\Omega$ such that each of (X_t), (Y_t), considered marginally, is a faithful copy of $\mathcal{M}$. The moves of the coupling are correlated to encourage the two copies of the Markov chain to *couple*: i.e. to achieve $X_t = Y_t$. This gives a bound on the total variation distance using the *Coupling Lemma* (see for example, Aldous [1]), which states that

$$\mathrm{d_{TV}}(P_x^t,\, \pi) \le \mathrm{Prob}[X_t \ne Y_t]$$

where $X_0 = x$ and Y_0 is drawn from the stationary distribution π. The following standard result is used to obtain an upper bound on this probability and hence an upper bound for the mixing time (the proof is omitted).

Theorem 2 *Let $(X_t,\, Y_t)$ be a coupling for the Markov chain $\mathcal{M}$ and let ρ be any integer valued metric defined on $\Omega\times\Omega$. Suppose that there exists $\beta \le 1$ such that $\mathbf{E}[\rho(X_{t+1},\, Y_{t+1})] \le \beta\,\rho(X_t,\, Y_t)$ for all t, and all $(X_t,\, Y_t) \in \Omega\times\Omega$. Let D be the maximum value that ρ achieves on $\Omega\times\Omega$. If $\beta < 1$ then the mixing time $\tau(\varepsilon)$ of $\mathcal{M}$ satisfies $\tau(\varepsilon) \le \log(D\varepsilon^{-1})/(1-\beta)$. If $\beta = 1$ and there exists $\alpha > 0$ such that*

$$\mathrm{Prob}[\rho(X_{t+1},\, Y_{t+1}) \ne \rho(X_t,\, Y_t)] \ge \alpha$$

for all t, and all $(X_t,\, Y_t) \in \Omega\times\Omega$, then $\tau(\varepsilon) \le \lceil eD^2/\alpha\rceil \lceil \log(\varepsilon^{-1})\rceil$.

From now on, assume that all couplings are Markovian unless explicitly stated. The path coupling method, introduced in [2], is a variation of traditional coupling which allows us to restrict our attention to a certain subset S of $\Omega\times\Omega$. If we view S as a relation, the transitive closure of S must equal Ω. The rate of convergence of the chain is measured with respect to a (quasi)metric ρ on $\Omega\times\Omega$, which can be defined by lifting a proximity function on S to the whole of $\Omega\times\Omega$ (see [7] for details). The path coupling lemma says the following. Let $(X,Y) \mapsto (X',Y')$ be a coupling defined on all pairs in S. Suppose there exists a constant β such that $0 < \beta \leq 1$ and for all $(X,Y) \in S$ we have

$$\mathbf{E}\left[\rho(X',Y')\right] \leq \beta\, \rho(X,Y). \tag{1}$$

Then we can conclude that (1) holds for all $(X,Y) \in \Omega \times \Omega$, and apply Theorem 2.

Suppose however that the smallest value of β for which (1) holds for all $(X,Y) \in S$ satisfies $\beta > 1$. Then path coupling is not good enough to allow us to apply Theorem 2. In some cases, we may make an improvement by defining a multiple-step path coupling. The following lemma is the "delayed path coupling lemma" [4, Lemma 4.2] of Czumaj et al., which shows how the mixing time of a Markov chain may be related to the behaviour of a t-step path coupling (which may be non-Markovian). The proof is obtained by applying the path coupling lemma to the transition matrix P^t.

Lemma 1 *Let $S \subseteq \Omega \times \Omega$ be such that the transitive closure of S is the whole of $\Omega \times \Omega$. Let ρ be an integer-valued metric on $\Omega \times \Omega$ which takes values in $\{0,\ldots,D\}$. Given $(X_0,Y_0) \in S$, let (X_0,Y_0), $(X_1,Y_1),\ldots,(X_t,Y_t)$ be the t-step evolution of a (possibly non-Markovian) coupling starting from (X_0,Y_0). Suppose that there exists a constant γ such that $0 < \gamma < 1$ and*

$$\mathbf{E}\left[\rho(X_t,Y_t)\right] \leq \gamma \rho(X_0,Y_0) \tag{2}$$

for all $(X_0,Y_0) \in S$. Then the mixing time $\tau(\varepsilon)$ of $\mathcal{M}$ satisfies

$$\tau(\varepsilon) \leq \frac{\log(D\varepsilon^{-1})}{1-\gamma} \cdot t.$$

In this paper, we apply Lemma 1 taking ρ to be the Hamming distance (defined below) and S to be the set of all pairs at Hamming distance 1 apart.

3 Preliminary results

Let G be a graph on n vertices with maximum degree D. Let C be a fixed set of k colours, and let $\Omega_k(G)$ denote the set of proper k-colourings of G.

Fix $\eta = 10^{-12}$ and assume that $(2-\eta)D \leq k < 2D$, since by the main results of [8] and [9], Theorem 1 holds when $k \geq 2D$. Note that, since k is an integer, these bounds on k imply that $D > \frac{1}{\eta}$, and thus $n > \frac{1}{\eta}$.

Throughout, we use $N(v)$ to denote the set of neighbours of v, and we use $d(v) = |N(v)|$ to denote the degree of v.

We will use path coupling with respect to the Hamming distance, where

$$H(X,Y) = |\,\{v \in V \mid X(v) \neq Y(v)\}\,|$$

for all $(X,Y) \in \Omega_k(G) \times \Omega_k(G)$. The set S of pairs used in path coupling is the set of colourings at Hamming distance 1 apart. In order to use path coupling with respect to

S, we must expand the state space of the Glauber dynamics to include all of C^V; that is, non-proper colourings as well as proper colourings. (Otherwise, we are not always able to form a path from X to Y with length $H(X, Y)$ and consisting of pairs in S.) This approach is now standard, and causes no problems since the non-proper colourings are transient states, and the stationary distribution of the extended chain is uniform over all proper colourings, and zero elsewhere. Although the chain is no longer reversible, the path coupling lemma still applies. Moreover, the mixing time of the chain on the original space is bounded above by the mixing time of the chain on the extended space.

Let X_0 and Y_0 be two (not necessarily proper) k-colourings of G which differ just at a single vertex v of degree $d(v)$. The standard path coupling for the Glauber dynamics is now described. From initial state (X_0, Y_0) produce the next state (X_1, Y_1) as follows:

(i) Choose $(w, c) \in V \times C$ uniformly at random,

(ii) if $w \notin N(v)$ or $c \notin \{X_0(v), Y_0(v)\}$ then attempt to recolour w with c in both X_0 and Y_0,

(iii) otherwise let c' be the unique element of $\{X_0(v), Y_0(v)\} \setminus \{c\}$ and attempt to recolour w with c in X_0 and w with c' in Y_0.

This coupling decreases the Hamming distance by 1 whenever v itself is chosen with a colour $c \notin \{X_0(w) \mid w \in N(v)\}$. There are at least $k - d(v)$ such choices. Similarly, the Hamming distance may increase by 1 when $w \in N(v)$ is chosen with colour $Y_0(v)$. There are $d(v)$ such choices. No other choices can affect the Hamming distance. Thus

$$\mathbf{E}[H(X_1, Y_1) - 1] \leq \frac{1}{kn} \left(-(k - d(v)) + d(v)\right) \leq -\frac{k - 2D}{kn}. \tag{3}$$

This is nonnegative when $k \geq 2D$, and gives an $O(n \log n)$ bound when $k > 2D$, by (1) and Theorem 2. However, when $k \leq 2D$ we must do more work to show that the Hamming distance tends to decrease.

The approach taken by Dyer et al. in [6] is a stopping times approach. They define a stopping time, which we shall denote T_1, and show that $\mathbf{E}[H(X_{T_1}, Y_{T_1})]$ is bounded above by a constant which is sufficiently smaller than 1. Particularly important are stopping times which occur within the first $n/8$ steps (for convenience we assume that $n/8$ is an integer). The stopping time T_1 is defined to be the first time at which a vertex-colour pair from the set $Q(X_0, Y_0)$ is chosen, where

$$Q(X_0, Y_0) = \{(v, c) \mid c \in C\} \cup \{(w, Y_0(v)) \mid w \in N(v)\}.$$

Note that this set contains all vertex-colour pair choices which can affect the Hamming distance. It is easy to verify that the calculations for (3) yield:

$$\mathbf{E}[H(X_{T_1}, Y_{T_1}) - 1] \leq -\frac{k - 2D}{|Q(X_0, Y_0)|} < \eta, \tag{4}$$

and that this bound still holds when conditioning on $T_1 = t$ for any $t > 0$.

They improve this bound on $\mathbf{E}[H(X_{T_1}, Y_{T_1}) \mid T_1 \leq n/8]$ by considering the number of distinct colours present around v just before T_1. Formally, the random variable $\mathcal{C}$ is defined by

$$\mathcal{C} = d(v) - |\{X_{T_1-1}(w) : w \in N(v)\}|.$$

So $\mathcal{C}$ measures the amount by which the number of distinct colours present on $N(v)$ just before the stopping time differs from $d(v)$, which is the maximum possible. The following lemma is embedded in the calculations of [6]. We present a proof here, since we will be using this approach for some of our own calculations.

Lemma 2 *Let (X_0, Y_0) be a pair of colourings which differ just at $v \in V$. Let η be a positive constant and let $(2-\eta)D \leq k \leq 2D$. Suppose that*

$$\mathbf{E}[\mathcal{C} \mid T_1 \leq n/8] \geq 2000\eta\, D$$

where $\mathcal{C}$ is as defined above. Then

$$\mathbf{E}[H(X_{T_1}, Y_{T_1}) \mid T_1 \leq n/8] \leq 1 - 500\eta.$$

Proof. There are $k + d(v)$ choices for the vertex-colour pair chosen at step T_1. Of these, at most $d(v)$ cause the Hamming distance to increase by one, and exactly $k - (d(v) - \mathcal{C})$ choices cause the Hamming distance to decrease by one. Therefore

$$\begin{aligned}
\mathbf{E}[H(X_{T_1}, Y_{T_1}) \mid T_1 \leq n/8] &\leq 1 - \frac{k - (d(v) - \mathbf{E}[\mathcal{C} \mid T_1 \leq n/8])}{k + d(v)} + \frac{d(v)}{k + d(v)} \\
&\leq 1 - \frac{(2-\eta)D - 2D + 2000\eta D}{k + d(v)} \\
&= 1 - \frac{1999\eta\, D}{k + d(v)} \\
&< 1 - 500\eta
\end{aligned}$$

as claimed. □

For the calculations of [6], a lower bound on the conditional expectation of T_1 is needed. A more complicated and more general bound was proven in [6]. For completeness we include a short proof of a simpler bound, which will be sufficient for our purposes.

Lemma 3 *With T_1 as defined above,*

$$\mathbf{E}[T_1 \mid T_1 \leq n/8] \geq \frac{n}{32}.$$

Proof. Let $p = (k + d(v))/kn$, which is the probability of choosing an element of $Q(X_0, Y_0)$ at any particular step. Then

$$\text{Prob}[T_1 = s] = p\,(1 - p)^{s-1}$$

which decreases as s increases. Note also that $\text{Prob}[T_1 \leq n/8] = 1 - (1 - p)^{n/8}$. Hence

$$\begin{aligned}
\mathbf{E}[T_1 \mid T_1 \leq n/8] &= \text{Prob}[T_1 \leq n/8]^{-1} \sum_{s=1}^{n/8} \text{Prob}[T_1 = s] \cdot s \\
&\geq \frac{\text{Prob}[T_1 = n/8]}{\text{Prob}[T_1 \leq n/8]} \sum_{s=1}^{n/8} s \\
&\geq \frac{p\,(1-p)^{n/8}}{(1 - (1-p)^{n/8})} \cdot \frac{n^2}{128} \\
&\geq \frac{p\left(1 - \frac{k+d(v)}{8k}\right)}{\frac{k+d(v)}{8k}} \cdot \frac{n^2}{128} \\
&= \left(7 - \frac{d(v)}{k}\right) \cdot \frac{n}{128} \\
&> \frac{n}{32},
\end{aligned}$$

as claimed. The fourth line follows since $x/(1 - x)$ decreases as x decreases, and then using the well-known inequality

$$(1 - x)^r \geq 1 - rx \tag{5}$$

which holds when r is a natural number and $0 < rx < 1$. □

The result of Molloy [9] also involves path coupling and stopping times. Rather than focus on the expected number of repeated colours, he considers the Hamming distance itself. He obtains bounds on the expected value of the Hamming distance by analysing up to three distinct stopping times (and conditioning on them all occuring within a certain number of steps). We will use a similar approach for some of our calculations.

4 Calculations

Let $(X_0, Y_0) \in C^V \times C^V$ be a pair of colourings of G which differ just at the vertex v. We say that vertex v has one of four possible types (depending also on X_0):

Type (i) $d(v) < 19D/20$,

Type (ii) there are fewer than $9D/10$ distinct colours in $\{X_0(w) \mid w \in N(v)\}$, and v does not have type (i),

Type (iii) there are fewer than $D^2/2000$ edges between the neighbours of v, and v does not have type (i) or (ii),

Type (iv) none of the above.

Each type is analysed separately.

Let T be the stopping time defined by T_1 if v has type (i), (ii) or (iii), and defined in Section 4.4 otherwise. In each case we aim to show that

$$\mathbf{E}[H(X_T, Y_T) \mid T \leq n/8] \leq 1 - 500\eta. \tag{6}$$

If v has type (i) then this can easily be shown directly. For types (ii) or (iii), we use the approach of [6]. If v has type (iv) then we analyse the expected value of the Hamming distance at a second stopping time $T = T_2 \geq T_1$, defined in Section 4.4. Here the arguments used are similar to those given in [9]. The calculations will be combined in Section 5 below to give an optimal bound for the mixing time of the Glauber dynamics on graphs with maximum degree D, where $(2-\eta)D \leq k \leq 2D$ and $\eta = 10^{-12}$.

4.1 Type (i): v has low degree

Suppose that v has type (i). By definition, $d(v) < 19D/20$. Here we can argue directly to show (6) holds, without any conditioning on T_1. There are exactly $k + d(v)$ vertex-colour pair choices for step T_1. Of these, at most $d(v)$ increase the Hamming distance by 1, and at least $k - d(v)$ decrease the Hamming distance by 1. Hence

$$\begin{aligned}
\mathbf{E}[H(X_{T_1}, Y_{T_1}) - 1] &\leq -\frac{k - 2d(v)}{k + d(v)} \\
&\leq -\frac{(2-\eta)D - 19D/10}{k + d(v)} \\
&< -\frac{D}{20(k + d(v))} \\
&< -500\eta
\end{aligned}$$

since $\eta = 10^{-12}$. This establishes (6) when v has type (i).

4.2 Type (ii): many repeated colours around v in X_0

Next suppose that v has type (ii). That is, there are fewer than $9D/10$ distinct colours in $\{X_0(w) : w \in N(v)\}$, and $d(v) \geq 19D/20$. We will establish (6) by applying Lemma 2. Here it is easy to show that you can form at least $D/40$ disjoint monochromatic pairs $\{u, w\} \subseteq N(v)$. To see this, suppose that there are n_i vertices in $N(v)$ coloured i. By assumption, there are at fewer than $9D/10$ colours i with $n_i > 0$. We can certainly form $\lfloor n_i/2 \rfloor$ disjoint pairs of vertices where both elements of each pair are coloured i. Hence the number of disjoint monochromatic pairs which we can form is at least

$$\sum_{i \in C} \left\lfloor \frac{n_i}{2} \right\rfloor \geq \frac{1}{2}\left(\sum_{i \in C} n_i - \frac{9D}{10}\right) = \frac{1}{2}\left(d(v) - \frac{9D}{10}\right) \geq \frac{D}{40}.$$

Now calculate the probability that neither u nor w are is chosen in the first $T_1 - 1$ steps, for such a pair $\{u, w\}$. There are $kn - (k + d(v)) \geq kn - (k + D)$ choices at each of the first $T_1 - 1$ steps, since we know that no vertex-colour pair in $Q(X_0, Y_0)$ has been chosen in these steps. We wish to avoid (u, c), (w, c) at each step, for all $c \in C \setminus \{Y_0(v)\}$, which means avoiding $2k - 2$ vertex-colour pairs. Hence the probability that neither u nor w is chosen in the first $T_1 - 1$ steps, for a fixed value of T_1, is at least

$$\left(1 - \frac{2k-2}{kn - (k+D)}\right)^{T_1 - 1} \geq \left(1 - \frac{4}{n}\right)^{T_1}.$$

Conditioning on $T_1 \leq n/8$ and using (5), it follows that

$$\mathbf{E}\left[\mathcal{C} \mid T_1 \leq n/8\right] \geq \left(1 - \frac{4}{n}\right)^{n/8} \cdot \frac{D}{40} \geq \frac{D}{80} > 2 \times 10^{-9} \cdot D = 2000\eta\, D.$$

Hence (6) holds, by applying Lemma 2.

4.3 Type (iii): few edges in $N(v)$

In this section, we assume that v has type (iii). That is, there are fewer than $D^2/2000$ edges between the neighbours of v, there are at least $9D/10$ distinct colours appearing on the neighbours of v, and $d(v) \geq 19D/20$. Again, we establish (6) by applying Lemma 2. Define the set $\mathcal{X} \subseteq N(v)$ by

$$\mathcal{X} = \{u \in N(v) : |N(u) \cap N(v)| < D/20\}.$$

It is not difficult to see that $|\mathcal{X}| \geq 9D/10$. For otherwise, at least $d(v) - 9D/10 \geq D/20$ elements of $N(v)$ have at least $D/20$ neighbours in common with v, giving at least $D^2/800$ edges in $N(v)$. Similarly, at least $D^2/4$ pairs of vertices in $\mathcal{X}$ are non-adjacent, otherwise there are too many edges in $N(v)$.

Let $\{u, w\} \subseteq \mathcal{X}$ be a non-adjacent pair of vertices. Define $S(\{u, w\})$ by

$$S(\{u, w\}) = N(v) \setminus (N(u) \cup N(w) \cup \{u, w\}).$$

Thus $S(\{u, w\})$ is the set of common non-neighbours of u, w in $N(v) \setminus \{u, w\}$. Note that $|S(\{u, w\})| \geq 4D/5$. For otherwise there are at least $d(v) - 2 - 4D/5 > D/10$ vertices in $N(v) \setminus \{u, w\}$ which are joined to u or to w. (This holds since D is large.) Hence, by the pigeonhole principle, at least one of u or w has at least $D/20$ neighbours in $N(v)$. This contradicts the definition of $\mathcal{X}$. We also know that $S(\{u, w\})$ is coloured with at least $7D/10$ distinct colours, since there are at most $D/5$ vertices in $N(v) \setminus S(\{u, w\})$.

We say that $\{u, w\}$ is a *Type A* pair if there are at least $2D/5$ distinct colours appearing on $S(\{u, w\})$, which are acceptable to at least one of u, w. (Here *acceptable* means not appearing on the neighbours of the given vertex in the colouring X_0.) Otherwise say $\{u, w\}$ is a *Type B* pair. Recall that there are at least $D^2/4$ non-adjacent pairs in

$\mathcal{X}$. Thus, by the pigeonhole principle, there are at least $D^2/8$ pairs of the same type (either Type A or Type B). This leads to two possibilities.

Many Type A pairs.

Suppose first that there are at least $D^2/8$ Type A pairs. Let $\{u, w\}$ be a Type A pair. Then there is at least one member of the pair, u say, such that at least $D/5$ colours appearing on $S(\{u, w\})$ are acceptable to u, by the pigeonhole principle. Call such an element *good*. There are at least $D/8$ good vertices, since each Type A pair produces at least one good vertex, and any good vertex can belong to at most D Type A pairs.

Choose exactly $\lceil D/8 \rceil$ good vertices and place them together in the set $\mathcal{U}$. For each good vertex u there are at least $D/5$ distinct colours appearing on $N(v) \setminus \{u\}$ which are acceptable to u. Even if we delete all the chosen good vertices and count again, we find there are at least $D/5 - D/8 - 1 > D/20$ distinct colours which appear on a vertex $z \in N(v) \setminus \mathcal{U}$ and which are acceptable to u. (This follows since D is large.) Consider the probability that u changes its colour to $X_0(z)$, and that z is never chosen, within the first $T_1 - 1$ steps. We also insist that no neighbour of u is ever chosen with the colour $X_0(z)$, and that u is never revisted. Thus we must choose $(u, X_0(z))$ at some step, and avoid $2k + D - 3$ vertex-colour pairs during the other $T_1 - 2$ steps. This probability of this, for a fixed value of T_1, is at least

$$\frac{1}{kn} \cdot (T_1 - 1) \cdot \left(1 - \frac{2k + D - 3}{kn - (k + D)}\right)^{T_1 - 2}.$$

If $T_1 \leq n/8$ then, using (5) we have

$$\left(1 - \frac{2k + D - 3}{kn - (k + D)}\right)^{T_1 - 2} \geq \left(1 - \frac{2(2k + D)}{kn}\right)^{n/8} \geq 1 - \frac{2k + D}{4k} \geq \frac{1}{4}.$$

Recall that there are at least $D/20$ choices for z. Using Lemma 3, the probability that u ends up recoloured the same colour as some other, untouched, neighbour of v just before the stopping time T_1, conditioned on $T_1 \leq n/8$, is at least

$$\frac{D}{20} \cdot \frac{1}{kn} \cdot \left(\frac{n}{32} - 1\right) \cdot \frac{1}{4} \geq \frac{1}{10240}.$$

Multiplying this value by the number of chosen good vertices we find that

$$\mathbf{E}\left[\mathcal{C} \mid T_1 \leq n/8\right] \geq \frac{1}{81920} \cdot D > 2 \times 10^{-9} \cdot D = 2000\eta\, D.$$

Hence (6) holds here, by applying Lemma 2.

Many Type B pairs.

Now suppose that there are at least $D^2/8$ Type B pairs. Let $\{u, w\}$ be a Type B pair. Then there are at least $D/5$ colours which are acceptable at both u and w. To see this, note that there are at least $3D/10$ colours appearing on $S(\{u, w\})$ which are acceptable at neither u nor w (since $\{u, w\}$ is not a Type A pair). Even if all remaining vertices around u and w are coloured using disjoint sets of distinct colours, there are at most $17D/10$ colours which are acceptable to at most one of u and w. Since there are at least $19D/10$ colours available, there are at least $D/5$ colours acceptable to both u and w, as claimed.

Let N be the number of Type B pairs and let Z be the number of (pair, colour) tuples consisting of a Type B pair and a colour which is acceptable to both members of the pair. We have just shown that $Z \geq DN/5$. We now prove that there are at least $D/10$ colours which are acceptable to both members of at least $N/20$ Type B pairs. For otherwise we can bound Z from above as follows:

$$Z \leq \left(\frac{D}{10} - 1\right) \cdot N + \left(\frac{19D}{10} + 1\right) \cdot \frac{N}{20} = \frac{39DN}{200} - \frac{19N}{20} < \frac{DN}{5}.$$

This is a contradiction. Hence there are at least $D/10$ *popular* colours, where a colour is called popular if it is available to both members of at least $N/20$ Type B pairs.

Suppose that i is a popular colour. The probability that i is repeated is bounded below by the probability that *exactly one* Type B pair is recoloured with the colour i. This probability, for a fixed value of T_1, is at least

$$\frac{D^2}{160} \cdot \frac{1}{k^2n^2} \cdot \binom{T_1 - 1}{2} \cdot \left(1 - \frac{3D + 2k - 6}{kn - (k + D)}\right)^{T_1 - 3}.$$

To see this, note that there are at least $N/20 \geq D^2/160$ choices for the Type B pair $\{u, w\}$ to be recoloured. The probability that (u, i) and (w, i) are both chosen at two different time steps is at least $1/(k^2n^2) \cdot (T_1 - 1)(T_1 - 2)/2$. Then we will never choose u or w again, and we never choose any other element of a Type B pair with colour i (there are at most $D - 2$ other vertices in Type B pairs), and we never choose a neighbour of u or w with colour i. (This is very similar to the calculation performed in [6]; however for completeness we derive rough bounds here.) Note that, conditional on $T_1 \leq n/8$ we have

$$\left(1 - \frac{3D + 2k - 6}{kn - (k + D)}\right)^{T_1 - 3} \geq \left(1 - \frac{2(3D + 2k)}{kn}\right)^{n/8} \geq 1 - \frac{3D + 2k}{4k} > \frac{2}{19}$$

using (5) and the fact that $k > 19D/10$. Applying Jansen's inequality and using Lemma 3, we have

$$\mathbf{E}\left[\binom{T_1 - 1}{2} \mid T_1 \leq n/8\right] \geq \binom{\mathbf{E}[T_1 \mid T_1 \leq n/8] - 1}{2} \geq \frac{n^2}{8192}.$$

Hence

$$\text{Prob}\left[i \text{ repeated} \mid T_1 \leq n/8\right] \geq \frac{D^2}{160} \cdot \frac{1}{k^2n^2} \cdot \frac{n^2}{8192} \cdot \frac{2}{19} \geq \frac{1}{49807360} > 2 \times 10^{-8}.$$

Since there are at least $D/10$ popular colours we find that

$$\mathbf{E}\left[\mathcal{C} \mid T_1 \leq n/8\right] \geq 2 \times 10^{-9} \cdot D = 2000\eta\, D.$$

Applying Lemma 2, we conclude that (6) holds.

4.4 Type (iv): several edges in $N(v)$

In this section we assume that v has type (iv). In particular, there are at least $D^2/2000$ edges between the neighbours of v, and $19D/20 \leq d(v) \leq D$. We will show that (6) holds. To do so, we will not focus on $\mathcal{C}$; rather we consider the situation after step T_1. To that end, we define T_2 to be the first time-step greater than T_1 in which we either (i) choose a vertex on which X_{T_1}, Y_{T_1} differ, or (ii) choose a neighbour, w, of such a vertex, v, and choose for w the colour $Y_{T_1}(v)$. If $H(X_{T_1}, Y_{T_1}) = 0$ then we set $T_2 = T_1$.

We begin by considering the various possibilities for (X_{T_1}, Y_{T_1}). The colourings differ on at most 2 vertices. The case of 0 vertices is trivial - in this case $H(X_{T_2}, Y_{T_2}) = 0$. In the case of 1 vertex, v, equation (4) applies (with time 0 replaced by T_1 and time T_1 replaced by T_2), giving

$$\mathbf{E}[H(X_{T_2}, Y_{T_2})] \leq 1 + \eta.$$

So we turn our attention to the case where they differ on exactly 2 vertices, u, v. Note that u, v must be adjacent. Define $r = |N(u) \cap N(v)|$. There are only two possible cases for the colour assignments to u, v:

Case 1: $X_{T_1}(u) = R, Y_{T_1}(u) = B; X_{T_1}(v) = B, Y_{T_1}(v) = R.$

This case arises when the choice for X in step T_1 is (u, R), and u has no neighbour other than v coloured either R or B.

From time $T_1 + 1$ to time T_2, we will couple our two chains as follows: If, for X we choose any vertex $w \in N(u) \cup N(v) - \{u, v\}$ and we choose R (resp. B) for w, then in Y we choose w and we choose B (resp. R) for w. Otherwise, we make the same choice in Y as we do in X.

There is no change in $H(X, Y)$ before time T_2. At that point, $H(X, Y)$ decreases by one if we choose either u or v, and we choose some colour, other than R, B which does not appear on the neighbourhood of that vertex. There are at least $2k - d(u) - d(v) - 2$ such possibilities.

For $H(X, Y)$ to increase by one, we must choose a vertex $w \neq u, v$ which is adjacent to exactly one of u, v and choose B if $w \in N(u)$ or R if $w \in N(v)$. There are at most $d(u) + d(v) - 2r - 2$ such possibilities.

There are exactly $2k + d(u) + d(v) - 2$ possibilities for step T_2. Therefore,

$$
\begin{aligned}
\mathbf{E}[H(X_{T_2}, Y_{T_2}) - 2] &\leq \frac{1}{2k + d(u) + d(v) - 2} \\
&\qquad \times ((d(u) + d(v) - 2r - 2) - (2k - d(u) - d(v) - 2)) \\
&= \frac{2d(u) + 2d(v) - 2r - 2k}{2k + d(u) + d(v) - 2} \\
&\leq \frac{4D - 2r - 2(2 - \eta)D}{2k + d(u) + d(v) - 2} \\
&\leq \frac{2\eta D}{2k + d(u) + d(v) - 2} - \frac{2r}{2k + d(u) + d(v) - 2} \\
&\leq \frac{\eta D}{k} - \frac{r}{3D} \\
&< \eta - \frac{r}{3D}.
\end{aligned}
$$

Case 2: $X_{T_1}(u) = R, Y_{T_1}(u) = G; X_{T_1}(v) = B, Y_{T_1}(v) = R.$

This case arises when the choice for X in step T_1 is (u, R), u has no neighbour other than v coloured R, but u has a neighbour, other than v, coloured B. Note that $X_{T_1-1}(u) = G$.

In this case, from time $T_1 + 1$ to time T_2, we will couple our chains as follows: (i) If, for X, we choose any vertex $w \in N(u) \cap N(v)$ and we choose G (resp. B) for w, then in Y we choose w and we choose B (resp. G) for w. (ii) If we choose any vertex $w \in N(u) - N(v) - v$ and we choose R (resp. G) for w, then in Y we choose w and we choose G (resp. R) for w. (iii) If we choose any vertex $w \in N(v) - N(u) - u$ and we choose R (resp. B) for w, then in Y we choose w and we choose B (resp. R) for w. (iv) Otherwise, we make the same choice for Y as for X.

Again, there are at least $2k - d(u) - d(v) - 2$ possible choices which will cause $H(X, Y)$ to decrease by one. For $H(X, Y)$ to increase by one, we must either (i) choose a vertex $w \neq u, v$ which is adjacent to exactly one of u, v and choose G if $w \in N(u)$ or R if $w \in N(v)$; or (ii) choose a vertex $w \in N(u) \cap N(v)$ and choose G. There are at most $d(u) + d(v) - r - 2$ such possibilities. The same calculations as in Case 1 yield:

$$\mathbf{E}[H(X_{T_2}, Y_{T_2})] \leq 2 + \eta - \frac{r}{6D}.$$

All other cases are isomorphic to one of these two.

Having analysed our cases, we turn our attention to the ways in which these cases can arise for various values of r. Recall that X_0, Y_0 differ only at v. For each neighbour $u \in N(v)$, we use r_u to denote $|N(u) \cap N(v)|$.

If, at time T_1, we choose $u \in N(v)$ and increase $H(X, Y)$, then we are in either Case 1 or Case 2 above, and so the conditional expectation of $H(X_{T_2}, Y_{T_2})$ is at most $2 + \eta - \frac{r_u}{6D}$. If we choose u and we do not increase $H(X, Y)$ then the conditional expectation of $H(X_{T_2}, Y_{T_2})$ is at most $1 + \eta < 2 + \eta - \frac{r_u}{6D}$.

If, at time T_1, we choose v and do not decrease $H(X,Y)$ then the conditional expectation of $H(X_{T_2},Y_{T_2})$ is at most $1+\eta$. There are at most $d(v)$ choices which can do this.

If, at time T_1, we choose v and decrease $H(X,Y)$, then the conditional expectation of $H(X_{T_2},Y_{T_2})$ is trivially equal to 0.

Therefore, since there are exactly $k+d(v)$ possibilities for time T_1, we have

$$\begin{aligned}
\mathbf{E}[H(X_{T_2},Y_{T_2})] &\le \frac{1}{k+d(v)} \times \left((1+\eta)d(v) + \sum_{u\in N(v)} 2+\eta-\frac{r_u}{6D} \right) \\
&\le \frac{1}{(3-\eta)d(v)} \times \left((1+\eta)d(v) + (2+\eta)d(v) - \frac{1}{6D}\sum_{u\in N(v)} r_u \right) \\
&= \frac{3+2\eta}{3-\eta} - \frac{1}{6(3-\eta)Dd(v)} \sum_{u\in N(v)} r_u \\
&< 1+2\eta - \frac{1}{18D^2}\sum_{u\in N(v)} r_u \\
&< 1+2\eta - \frac{1}{18000} \\
&< 1-500\eta,
\end{aligned}$$

as required. Here the second last line uses the fact that the number of edges in $N(v)$ is at least $D^2/2000$, and the final line follows since $\eta = 10^{-12}$.

Finally, observe that all of our bounds still hold even upon conditioning on $T_1 = i, T_2 = j$ for any i, j such that $1 \le i < j$. In particular, (6) holds.

5 Synthesis

We can now combine the above analysis of types (i)–(iv), as defined in the introduction. As above, let (X_0,Y_0) be a pair of k-colourings which differ just at $v \in V$. Let T be the stopping time defined by

$$T = \begin{cases} T_1 & \text{if } v \text{ has type (i), (ii) or (iii),} \\ T_2 & \text{otherwise.} \end{cases}$$

Then the calculations of Sections 4.1–4.4 show that

$$\mathbf{E}[H(X_T,Y_T) \mid T \le n/8] \le 1-500\eta,$$

by definition of T. Of course, what we really want is $\mathbf{E}(H(X_{n/8},Y_{n/8}))$, so we must adjust this bound to account for two things: (I) the possibility that $T > n/8$, and (II) the possibility that H increases from time T to time $n/8$. We start by obtaining a lower bound on Prob$[T \leq n/8]$. Note that $T \leq T_2$, and that $T_2 \leq T^*$ where T^* is the second step in which we select v. Therefore,

$$\begin{aligned}
\text{Prob}[T \leq n/8] &\geq \text{Prob}[T^* \geq n/8] \\
&= 1 - \left(1 - \frac{1}{n}\right)^{n/8} - \frac{n}{8} \times \frac{1}{n} \times \left(1 - \frac{1}{n}\right)^{n/8} \\
&\geq 1 - \exp(-1/8) - \frac{1}{8}\exp(-1/8) \\
&\geq 1/200.
\end{aligned}$$

Next we note that if $T > n/8$ then either (a) $T_1 > n/8$, in which case $H(X_{n/8}, Y_{n/8}) = 1$, or (b) $T_1 \leq n/8 < T_2$ and so $H(X_{n/8}, Y_{n/8}) = H(X_{T_1}, Y_{T_1})$. By (4),

$$\mathbf{E}[H(X_{T_1}, Y_{T_1}) \mid T_1 \leq n/8] < 1 + \eta.$$

Therefore,

$$\mathbf{E}[H(X_{n/8}, Y_{n/8}) \mid T > n/8] < 1 + \eta.$$

We now bound the maximum amount by which the Hamming distance can increase from step T to step $n/8$. Let $\beta = 1 + \eta\, D/(kn)$. Then β is an upper bound on the increase of the Hamming distance per path coupling step (as in (3)). Thus (3) and a standard path-coupling argument (see, eg. the discussion in [7] or [6]) allows us to extend our coupling so that for any $t < n/8$,

$$\mathbf{E}[H(X_{n/8}, Y_{n/8}) \mid T = t] < \beta^{n/8-t} \times H(X_T, Y_T).$$

Thus

$$\mathbf{E}[H(X_{n/8}, Y_{n/8}) \mid T \leq n/8] \leq \beta^{n/8} \times \mathbf{E}[H(X_T, Y_T) \mid T \leq n/8].$$

Therefore

$$\begin{aligned}
\mathbf{E}[H(X_{n/8}, Y_{n/8})] &\leq \text{Prob}[T \leq n/8] \left(\beta^{n/8}\mathbf{E}[H(X_T, Y_T) \mid T \leq n/8]\right) \\
&\quad + \text{Prob}[T > n/8]\,(1 + \eta) \\
&\leq 1 + \text{Prob}[T \leq n/8] \left(\left(1 + \frac{\eta\, D}{kn}\right)^{n/8} (1 - 500\eta) - 1 \right) + \eta \\
&\leq 1 + \text{Prob}[T \leq n/8] \left(\exp\left(\frac{\eta}{8(2-\eta)} - 500\eta\right) - 1 \right) + \eta \\
&\leq 1 - \frac{1}{200} \cdot 400\eta + \eta \\
&= 1 - 10^{-12}.
\end{aligned}$$

Therefore, setting $\gamma = 1 - 10^{12}$, by Lemma 1 the mixing time of the Glauber dynamics satisfies

$$\tau(\varepsilon) \leq \frac{n}{8(1-\gamma)} \log(n\varepsilon^{-1}) \leq \frac{10^{12}}{8}\, n \log(n\varepsilon^{-1}) \leq 2 \times 10^{11}\, n \log(n\varepsilon^{-1}),$$

thus proving Theorem 1.

References

[1] D. J. Aldous, *Random walks on finite groups and rapidly mixing Markov chains*, in A. Dold and B. Eckmann, eds., *Séminaire de Probabilités XVII 1981/1982*, vol. 986 of *Springer–Verlag Lecture Notes in Mathematics*, Springer–Verlag, New York, 1983, pp. 243–297.

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

[3] R. Bubley, M. Dyer and C. Greenhill, *Beating the* 2Δ *bound for approximately counting colourings: a computer-assisted proof of rapid mixing*, in *9th Annual Symposium on Discrete Algorithms*, ACM–SIAM, New York–Philadelphia, 1998, pp. 355–363.

[4] A. Czumaj, P. Kanarek, M. Kutyłowski, and K. Loryś, *Delayed path coupling and generating random permutations via distributed stochastic processes*, in *10th Annual Symposium on Discrete Algorithms*, ACM–SIAM, New York–Philadelphia, 1999, pp. 271–280.

[5] P. Diaconis and L. Saloff-Coste, *Comparison theorems for reversible Markov chains*, The Annals of Applied Probability **6** (1996), pp. 695–750.

[6] M. Dyer, L. A. Goldberg, C. Greenhill, M. Jerrum and M. Mitzenmacher, *An extension of path coupling and its application to the Glauber dynamics for graph colourings*, in *11th Annual Symposium on Discrete Algorithms*, ACM–SIAM, New York–Philadelphia, 2000, pp. 616–624.

[7] M. Dyer and C. Greenhill, *Random walks on combinatorial objects*, in J. D. Lamb and D. A. Preece, eds., *Surveys in Combinatorics, 1999*, vol. 267 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge, 1999, pp. 101–136.

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

[9] M. Molloy, *Very rapidly mixing Markov chains for* 2Δ*-colourings and for independent sets in a* 4*-regular graph*, Preprint, 1999.

[10] D. Randall and P. Tetali, *Analysing Glauber dynamics by comparison of Markov chains*, in *3rd Latin American Symposium on Theoretical Informatics*, vol. 1380 of *Springer–Verlag Lecture Notes in Computer Science*, Springer–Verlag, New York, 1998.

[11] J. Salas and A. D. Sokal, *Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem*, Journal of Statistical Physics **86** (1997), pp. 551–579.

[12] E. Vigoda, *Improved bounds for sampling colorings*, Journal of Mathematical Physics **41** (2000), pp. 1555–1569.