The Glauber dynamics on colourings of a graph with high girth and maximum degree

Michael Molloy*
Department of Computer Science
University of Toronto
Toronto, Canada

October 30, 2001

Abstract

We prove that the Glauber dynamics on the C-colourings of a graph G on n vertices with girth g and maximum degree Δ mixes rapidly if (i) $C = q\Delta$ and $q > q^*$ where $q^* = 1.4890...$ is the root of $(1 - e^{-1/q})^2 + qe^{-1/q} = 1$; and (ii) $\Delta \geq D \log n$ and $g \geq D \log \Delta$ for some constant $D = D(q)$. This improves the corresponding result with $q \geq 1.763\Delta$ obtained by Dyer and Frieze [FOCS01] *for the same class of graphs. It also improves $q \geq 11\Delta/6 \approx 1.833\Delta$ obtained by Vigoda* [FOCS99] *for general graphs.*

For a given graph G and integer C which is at least the chromatic number of G, the Glauber dynamics on the C-colourings of G is the Markov chain described as follows. We start with an arbitrary C-colouring, and at each step we choose a uniformly random vertex v, and a uniformly random colour c from $L(v)$, the list of colours which do not appear on any neighbours of v. Then we change the colour of v to c. (If $L(v)$ is empty then we do not choose a colour; note that $L(v)$ is never empty if $C > \Delta$.)

Unless specified otherwise, we consider all colourings to be proper, i.e. no two adjacent vertices can have the same colour. In section 3, it will be important to note that we can apply this step even to a non-proper colouring of G. But note that if a colouring is proper, then applying a step of the chain cannot produce an improper colouring.

*This work is supported by an NSERC Research Grant and a Sloan Research Fellowship.

Different versions of the Glauber dynamics on the C-colourings of G appear in the literature. For example, we might choose c uniformly at random from $\{1, ..., C\}$ and then only change the colour of v if $c \in L(v)$. But as far as the asymptotics of the mixing rate goes, they are all equivalent.

This chain is of great interest for a few reasons. For one, it is the most natural chain on the colourings of a graph, and so is an obvious attempt at a procedure to approximately count the colourings of a graph and to generate such a colouring nearly uniformly at random. (It is well known that if the Markov chain mixes rapidly, then it can be used to accomplish both of these tasks). It is also of interest in the statistical physics community because of its relation to the Potts model.

The main question in this area is: For what values of C does this Markov chain mix in polytime? Usually this is studied in terms of Δ, the maximum degree of G. It is well known that for some graphs, the chain does not mix for $C \leq \Delta + 1$. Jerrum [6] showed that for all graphs, the chain mixes in polytime for $C \geq 2\Delta$ and in optimal time, i.e. $O(n \log n)$ time, for $C \geq 2\Delta + 1$. Salas and Sokal [9] independently obtained the latter result. Vigoda [11] showed that for all graphs, a different chain mixes in optimal time for $C \geq \frac{11}{6}\Delta$ and this implies that for the same values of C, the Glauber dynamics mixes in polytime. Dyer, Greenhill and Molloy [5] showed that the Glauber dynamics mixes in optimal time for $C \geq (2-\epsilon)\Delta$ where ϵ is a small positive constant (see also [8]). Some work has been done on special classes of graphs. Dyer et al[3] showed that the Glauber dy-

namics mixes in optimal time on triangle-free graphs when $C \geq (2-x)\Delta$ for a different small positive constant x. More recently, Dyer and Frieze[2] showed that if the girth and maximum degree of G are both at least $D \log n$ for some sufficiently large constant D, then it mixes in optimal time for $C \geq q\Delta$ for any constant $q > \beta$ where $\beta = 1.763... < 11/6$ is the root of $\beta e^{-1/\beta} = 1$. Here we improve on this latter result:

We define $q^* = 1.4890...$ to be the root of $(1 - e^{-1/q})^2 + qe^{-1/q} = 1$.

Theorem 1 *For any $q > q^*$ and integer ω, there exists D for which: Suppose G has n vertices, maximum degree $\Delta \geq D \log n$ and no vertex lies in more than ω cycles of length less than $D \log \Delta$. Then the Glauber dynamics mixes in time $O(n \log n)$ for $C = q\Delta$.*

Of course, this covers all graphs with girth at least $D \log \Delta$. It also covers natural models of random graphs such as $G_{n,p=c/n}$ for $c \approx D \log n$ and random Δ-regular graphs for $\Delta = D \log n$. Dyer and Frieze[2] showed that their theorem also extends to such graphs.

Throughout the paper, we assume that n is large enough for various asymptotic bounds to hold. We can also assume that q is sufficiently close to q^*. We use $N(v)$ to denote the neighbourhood of v, i.e. the set of vertices which are adjacent to v. We define $d(v) = |N(v)|$ to be the degree of v. A *short cycle* is a cycle of length less than $D \log \Delta$.

1 Some intuition

The proofs of all results mentioned above, except for that of Vigoda [11], all come from the following idea[1]. Consider two colourings X, W which differ only at one vertex v. We will carry out one step of the process on each colouring, where we couple these two random steps maximally. Specifically, we first choose a uniform vertex u for both colourings. if $L_X(u)$ and $L_W(u)$ are the sets of allowable colours for u in X, W respectively, then we take two mappings $f_X : [0,1] \to L_X(u), f_W : [0,1] \to L_W(u)$, such that

- for each $c \in L_X(u)$, $|f_X^{-1}(c)| = 1/|L_X(u)|$ and similarly for W, and
- $\{x : f_X(x) \neq f_W(x)\}$ is as small as possible.

Then we take a uniform random real $x \in [0,1]$ and choose $u, f_X(x)$ for X and $u, f_W(x)$ for W. Note that since $L_X(u), L_W(u)$ differ in at most one colour per list, we will have $|\{x : f_X(x) \neq f_W(x)\}| \leq \max\{|L_X(u)|^{-1}, |L_W(u)|^{-1}\}$. Note further that $L_X(u) = L_W(u)$ unless u is a neighbour of v.

Using the path-coupling technique of Bubley and Dyer [1], it usually suffices to show that the probability of X, W converging after one step is greater than the probability of them differing in a second vertex after one step (we elaborate on this in Section 3). They converge iff we choose $u = v$, which occurs with probability $1/n$. They differ on a second vertex iff we choose some $u \in N(v)$ and we choose $x \in \{x : f_X(x) \neq f_W(x)\}$. Since no list can ever be smaller than $C - \Delta$, this occurs with probability at most $\frac{\Delta}{n} \times \frac{1}{C-\Delta}$ which is less than $\frac{1}{n}$ so long as $C > 2\Delta$.

The bound was improved slightly in [3] by showing that for triangle-free graphs, after a burn-in period of $O(n \log n)$ steps, most vertices will tend to have many repeated colours in their neighbourhoods. Thus, their lists of available colours will tend to be somewhat greater than $C - \Delta$, and this leads to a gain in the above calculations. The same idea played a key role in [5].

In [2], Dyer and Frieze showed that for graphs with large girth and maximum degree, after $O(n \log n)$ steps, with high probability every vertex will have a list of size at least roughly $qe^{-1/q}\Delta$. If X, W are such that all vertices has lists of this size, then this yields that the probability of X, W differing on a second colour is at most $\frac{\Delta}{n} \times \frac{1}{qe^{-1/q}\Delta}$. For $q \geq 1.763..$ we have $qe^{-1/q} > 1$, and so this probability is less than $\frac{1}{n}$.

The key new idea used in this paper is to show that, after $O(n \log n)$ steps, many neighbours u of

[1] Jerrum's and Salas and Sokal's original proofs predated this idea, but the idea yields a simpler proof.

v will satisfy $L_X(u) = L_Y(u)$. In fact, with high probability, there will be roughly $(1 - \mathrm{e}^{-1/q})^2\Delta$ such neighbours. This improves our bound on the probability of X, W differing on a second colour to $\frac{(1-(1-\mathrm{e}^{-1/q})^2)\Delta}{n} \times \frac{1}{q\mathrm{e}^{-1/q}\Delta}$, which is less than $\frac{1}{n}$ for $q > q^*$.

As we discuss in section 4, this is the limit to which this proof technique can be extended, even if we restrict our attention to graphs with high girth and maximum degree. It would be interesting to determine whether there is some small $x > 0$ such that the Glauber dynamics mix rapidly on such graphs for $C > (q^* - x)\Delta$, since to prove the existence of such an x would require a different approach (for example, perhaps by analyzing the more complicated Markov chain in [11]). To prove that no such x exists would be very interesting since it would provide a negative answer to the question of whether for every constant $x > 0$ the Glauber dynamics mixes rapidly on every graph of maximum degree Δ so long as $C \geq (1+x)\Delta$.

1.1 Uniform-like colour sets

Suppose that $d(v) = \Delta$ and that every vertex in $N(v)$ is assigned an independent uniform colour from $\{1, ..., C\}$. Then the probability that some colour c does not appear on any neighbour of v is $(1 - 1/C)^\Delta = \mathrm{e}^{-1/q} + o(1)$. Thus we would expect that each list would have size roughly $C \times \mathrm{e}^{-1/q} = q\mathrm{e}^{-1/q}\Delta$. This explains, at least intuitively, the lower bound that Dyer and Frieze obtain.

Now suppose that every vertex of distance 2 from v is also assigned an independent uniform colour from $\{1, ..., C\}$. Suppose further that we change the colour of v in our colouring X to obtain another colouring W. For each $u \in N(v)$, $L_X(u) = L_W(u)$ iff the colours $X(v)$ and $W(v)$ both appear on $N(u) - v$. The probability that this occurs is $(1-\mathrm{e}^{-1/q})^2+o(1)$, and so this explains the result in this paper.

Of course, the colours appearing on the neighbours of v are far from independent. But intuitively, since there are few short cycles near v, after $O(n \log n)$ steps the colours on vertices close to v are "close enough" to being independent. Much of the work in this paper can be viewed as proving this statement.

Our situation is somewhat more complicated than that in [2]. To illustrate this, suppose that $N(v) = w_1, ..., w_\Delta$ and consider any assignment of colours to the vertices in $N(w_1), ..., N(w_\Delta)$. Now, assign to each w_i a uniform colour from amongst those not appearing on $N(w_i)$, independently of the choice for any other w_j. It turns out, that for any assignment of colours to $N(w_1), ..., N(w_\Delta)$, the expected number of colours not appearing on $N(v)$ is at least $\mathrm{e}^{-1/q} + o(1)$. Thus, Dyer and Frieze did not have to ensure that the colours appearing on $N(w_1), ..., N(w_\Delta)$ are nearly uniform. Unfortunately, when analyzing our other parameter, we do not have this advantage, and we need to show that the colours appearing at distance 2 and 3 from v will be nearly uniform. To do this, we require a complicated iterative analysis.

2 The main lemmas

Our two main lemmas are:

Lemma 2 *For every $\epsilon, \rho > 0$, there exist constants D, τ such that with probability at least $1 - O(n^{-6})$ for every vertex v with $d(v) > \rho\Delta$ and time $\tau n \log n \leq t \leq n^2$, we have $C\mathrm{e}^{-d(v)/C} - \epsilon\Delta < |L(v)| < C\mathrm{e}^{-d(v)/C} + \epsilon\Delta$.*

Lemma 3 *For every $\epsilon, \rho > 0$, there exist constants D, τ such that with probability at least $1 - O(n^{-6})$, for every vertex v, colours c_1, c_2 and time $\tau n \log n \leq t \leq n^2$, v has fewer than $(1 - (1 - \mathrm{e}^{-1/q})^2)d(v) - \epsilon\Delta$ neighbours w with $d(w) \geq \rho\Delta$ and at least one of c_1, c_2 in $L(w)$.*

We start with the following recursive definition, for each $0 \leq d \leq \Delta$:

- $\alpha_1 = 0, \beta_1 = 1/(q-1), \lambda_1^{(d)} = q$;
- $\alpha_{k+1} = \mathrm{e}^{-\beta_k}/\lambda_k^{(\Delta)}$;
- $\beta_{k+1} = \mathrm{e}^{-\alpha_k}/(q\mathrm{e}^{-1/q})$;
- $\lambda_{k+1}^{(d)} = \frac{q\beta_k - 1}{\beta_k - \alpha_k}\mathrm{e}^{-\alpha_k d/\Delta} + \frac{1 - q\alpha_k}{\beta_k - \alpha_k}\mathrm{e}^{-b_k d/\Delta}$.

Lemma 4 $\lim_{k\to\infty} \alpha_k = \lim_{k\to\infty} \beta_k = 1/q$, $\lim_{k\to\infty} \lambda_k^{(d)} = q\mathrm{e}^{-d/q\Delta}$.

We postpone the proof of Lemma 4 to the appendix. But for now, we require the following simple observation:

Lemma 5 *For all $k \geq 1$ and $0 \leq d \leq \Delta$:*

(a) $\alpha_k < 1/q < \beta_k$, *and*

(b) $qe^{-\beta_k d/\Delta} \leq \lambda_{k+1}^{(d)} \leq qe^{-\alpha_k d/\Delta}$.

Proof (b) follows since $\lambda_{k+1}^{(d)}$ is a linear combination of $qe^{-\beta_k d/\Delta}$ and $qe^{-\alpha_k d/\Delta}$. (a) follows from a simple induction. □

Now, fix some small δ to be named later, and choose k^* such that $\alpha_{k^*} > \frac{1}{q} - \delta/2$, $\beta_{k^*} < \frac{1}{q} + \delta/2$ and $\lambda_{k^*} < qe^{-1/q} + \delta/2$ (such an k^* exists by Lemma 4). Choose a sufficiently small constant ζ and constants $a_1, ..., a_{k^*}, b_1, ..., b_{k^*}, \ell_1^{(0)}, ..., \ell_{k^*}^{(\Delta)}$ such that

(i) $a_1 = \alpha_1, b_1 = \beta_1, \ell_1^{(d)} = \lambda_1^{(d)}$;

(ii)
- $a_{k+1} < e^{-b_k}/\ell_k - \zeta$;
- $b_{k+1} > e^{-a_k}/(qe^{-(1+\zeta)/q}) + \zeta$;
- $\ell_{k+1}^{(d)} = \frac{qb_k - 1}{b_k - a_k} e^{-(1-\zeta)a_k d/\Delta} + \frac{1 - qa_k}{b_k - a_k} e^{-(1-\zeta)b_k d/\Delta}$;

(iv) $a_{k^*} > \frac{1}{q} - \delta, b_{k^*} < \frac{1}{q} + \delta$.

The existence of such constants follows easily from the continuity of the relevant functions over the region $a \neq b$.

Note that we can assume that ρ is as small as we wish. In particular, we will assume that

- $1/(q - \rho) < e^{(1+\zeta-a_{k^*})/q}/q$; and
- $e^{-\rho/(q-1)} > e^{1/q - b_{k^*}}$.

We define $L_t(v)$ to be the set of colours available for v at time t; we sometimes omit the subscript t when it is not necessary. For any vertex u and time t, we define $\overline{t}(u,t)$ to be the last time before t that u is selected. To be precise, we define $\overline{t}(u,t)$ to be zero in the event that v is not selected before time t.

$$T_t(v,c) = \sum_{w \in N(v), c \in L_{\overline{t}(w,t)}(w)} \frac{1}{|L_{\overline{t}(w,t)}(w)|}.$$

Note that

$$\sum_{c=1}^{C} T_t(v,c) = \sum_{w \in N(v)} \sum_{c \in L_{\overline{t}(w,t)}(w)} \frac{1}{|L_{\overline{t}(w,t)}(w)|} = d(v).$$

We will prove inductively that

Lemma 6 *For each $1 \leq k \leq k^*$, with probability at least $1 - n^{-6}$, for every v with $d(v) > \rho\Delta$, colour c and $30kn \log n \leq t \leq n^2$, we have:*

(a) $qe^{-(1+\zeta)d(v)/C}\Delta \leq |L_t(v)| \leq \ell_k^{(d(v))}\Delta$;

(b) $a_k d(v)/\Delta \leq T_t(v,c) \leq b_k d(v)/\Delta$.

The lower bound in Part (a) is essentially Lemma 4.1 and (7) of Dyer and Frieze[2], and our proof is similar to theirs.

Proof The proof is by induction on k. The base case $k = 1$ holds trivially. So assume that it holds for some k and consider $k + 1$.

Let the neighbours of v be $w_1, ..., w_{d(v)}$ and for each i let $t_i = \overline{t}(w_i, t)$ be the last time before t that w_i is selected. For each i, let $N_{w_i} - v = \{w_{i,1}, ..., w_{i,d(w_i)-1}\}$ and let $t_{i,j} = \overline{t}(w_{i,j}, t_i)$ be the last time before t_i that $w_{i,j}$ is selected. Similarly, for each i, j let $N_{w_{i,j}} - w_i = \{w_{i,j,1}, ..., w_{i,j,d(w_{i,j})-1}\}$ and let $t_{i,j,r} = \overline{t}(w_{i,j,r}, t_{i,j})$ be the last time before $t_{i,j}$ that $w_{i,j,r}$ is selected. (If v lies in some cycles of length at most 6, then some vertices will receive more than one label, but this does not create a problem.)

The idea of focussing on the colours assigned at these times was used by Dyer and Frieze [2].

With high probability, each $t_i \geq t - 10n \log n$. Indeed, the probability that this is not true is at most $\Delta(1 - \frac{1}{n})^{10n \log n} < n^{-9}$. The same calculations give the same bound on the probability that some $t_{i,j}$ is less than $t - 20n \log n$ or that some $t_{i,j,r}$ is less than $t - 30n \log n$.

We start by proving part (a). Expose the values of $t_1, ..., t_{d(v)}, t_{1,1}, ..., t_{d(v),d(w_{d(v)})-1}$.

It would be nice if we could say that the colours assigned to $w_1, ..., w_{d(v)}$ were independent, as that would greatly simplify our calculations. However, this is clearly not the case, since the colour assigned to w_i has an effect on the next colour assigned to the

neighbour of w_i, and this effect can propagate along a path which eventually leads to some w_j. This can be a very short path, if it goes through v; we will deal with such paths later. Otherwise, unless w_i is one of the at most $O(1)$ vertices lying on a short cycle through v, the path must have length at least $D\log\Delta - 2$. Our first concern will be such long paths.

Consider the following procedure: GLAUB(i). It follows the usual Glauber dynamics, but after step t_i, all neighbours of w_i ignore w_i. More specifically, for each $u \in N(w_i)$, $L(u)$ is the set of colours which do not appear on $N(u) - w_i$, and whenever u is selected, u is assigned a uniformly random colour from $L(u)$. Thus, for example, w_i might have the same colour as some of its neighbours.

We use GLAUB to denote the usual Glauber dynamics, and we consider the two procedures to be coupled in that each has the same initial state and chooses the same vertex at each step. The color choice at each step is coupled maximally.

A *long path of disagreement* from w_i is a path P of length at least $\ell = (D/3)\log\Delta$ beginning at w_i and not going through v, such that the colour of each vertex in the path differs between GLAUB and GLAUB(i). (Such paths were used in a slightly different way by Dyer and Frieze [2] and earlier, in a different setting, by van den Berg and Steif[10].) We say that w_i is *influential* if there is a long path of disagreement from w_i or if by changing the colour assigned to w_i at time t_i, and not changing any future colour/vertex choices in GLAUB(i), it is possible to create a long path of disagreement from w_i.

We define B to be the set of neighbours w_i such that either w_i lies in a cycle through v of length less than $D\log\Delta$ or w_i is influential.

Lemma 7 *With probability at least* $1 - n^{-10}$, $|B| < \Delta/\log\Delta$.

Proof The set of neighbours which lie on a short cycle through v is at most 2ω. Therefore, we only have to count the influential vertices which do not lie on any short cycle through v. Consider such a vertex w_i; we will bound the probability that it is influential.

We will bound this probability by the probability that either w_i is influential, or $t_i < t - 20n\log\Delta$. The probability of the latter is $(1-1/n)^{20n\log\Delta} < \frac{1}{2}\Delta^{-11}$.

If w_i is influential and $t_i \geq t - 20n\log\Delta$, then there is a path $P : w_j = p_0, p_1, ..., p_\ell$, a colour c to assign to w_i at time t_i, and time steps $t - 20n\log\Delta < s_1 < ... < s_\ell \leq t$ such that at step s_r, p_r is selected and assigned different colours in the two procedures.

Suppose that P is the lexicographically first such path formed. Thus, at each step s_r, if the colours of some neighbour u of p_r disagree in the two procedures, then there is a path of disagreement from u to w_j which does not go through p_r. Since that path has length less than ℓ, then p_r and u lie in a cycle of length less than $D\log\Delta$. Since p_r lies in at most ω such cycles, there can be at most 2ω such neighbours. Thus, since $|L(p_r)|$ is always at least $C - \Delta = (q-1)\Delta$, the probability that p_r chooses a different colour for each procedure, under the maximal coupling, is less than $2\omega/((q-1)\Delta)$. Therefore, the probability that such a path is formed is at most:

$$\Delta \times (\Delta-1)^\ell \times \binom{20n\log\Delta}{\ell} \times \left(\frac{2\omega}{(q-1)\Delta}\right)^\ell < \Delta \times \left(\frac{100\mathrm{e}\omega}{D}\right)^\ell < \frac{1}{2}\Delta^{-11},$$

for D sufficiently large and q near q^*. Thus, after adding the probability that $t_i < t - 20n\log\Delta$, the probability that w_i is influential is at most Δ^{-11}.

Now consider any collection $w_{i_1}, ..., w_{i_m}$ of $m = \Delta/\log\Delta - 2\omega$ neighbours which don't lie in short cycles through v. Using the fact that long paths of disagreement from any two such neighbours must be disjoint, similar calculations yield that the probability of all m neighbours being influential is at most $(\Delta^{-11})^m$. Therefore, the probability that at least m such neighbours of v are influential is at most

$$\binom{\Delta}{m} \times \Delta^{-11m} < \Delta^{-10m} < n^{-10},$$

for $D > 1$. □

Now, for each i, j, we expose the colour assigned to $u_{i,j}$ at time $t_{i,j}$. We also expose the set B. We denote this set of information, along with the values of $t_1, ...t_{d(v)}, t_{1,1}, ...t_{d(v),d(w_{d(v)})-1}$, by $\mathcal{H}$. We say that

$\mathcal{H}$ is *good* if every $t_i, t_{i,j} \geq t - 30n\log n$, if the sets of colour assignments satisfy the conditions of Lemma 6 for k and if $|B| \leq \Delta/\log\Delta$. We will show by induction that the probability of $\mathcal{H}$ not being good is at most $O(n^{-9})$.

Lemma 8 *For any good $\mathcal{H}$, the conditional probability that, at a particular time $t \geq 30(k+1)n\log n$ during MOD-GLAUB, $|L(v)| \leq q\mathrm{e}^{-(1+\zeta/2)d(v)/C}\Delta$ or $|L(v)| \geq \ell_{k+1}^{(d(v))}\Delta - \zeta/2$ is at most n^{-9}.*

Proof Rather than analyzing $|L(v)|$ directly, we will focus on $|L^*(v)|$, the set of colours which do not appear on any $w_i \in N(v) - B$. Note that $|L^*(v)| \geq |L(v)| \geq |L^*(v)| - |B| = |L^*(v)| - o(\Delta)$.

Rather than dealing with GLAUB directly, we consider the procedure MOD-GLAUB, whereby after time $t - 30n\log n$, neighbours of v ignore the colour on v.

The advantage of focussing on $|L^*(v)|$ and MOD-GLAUB is that the colour assignments to the vertices in $N(v) - B$ are independent. This is because no combination of assignments to these vertices can produce a long path of disagreement from one such vertex to another, and, since the vertices ignore the colour on v and no two lie on a short cycle through v, there are no short paths of disagreement between these vertices.

Furthermore, exposing the fact that $w_i \notin B$, only exposes that w_i is not influential. Since the definition of "influential" does not depend on the particular colour assigned to w_i, this does not expose anything about that choice of colour.

So each w_i receives a uniformly random colour from amongst those colours which $\mathcal{H}$ dictates to not appear on $N(w_i) - v$ at time t_i, and the choices for $w_1, ..., w_\Delta$ are independent. Thus, for any colour c the conditional probability that c belongs to $L^*(v)$ at time t is

$$\prod_{w\in N(v)-B, c\in L(w)} 1 - \frac{1}{L(w)} = \exp(-T(v,c)) + o(1),$$

since every $L(w)$ has size at least $C - \Delta = (q-1)\Delta$, and since $|B| = o(\Delta)$. Therefore, the expected size of $L^*(v)$ at time t is equal to $\sum_c \exp(-T(v,c)) + o(\Delta)$.

Recall that $\sum_c T(v,c) = d(v)$. Furthermore, since $\mathcal{H}$ is good, each $T(v,c)$ lies between $a_k d(v)/\Delta$ and $b_k d(v)/\Delta$. Subject to these constraints, $\sum_c \exp(-T(v,c))$ is easily seen to be minimized when every $T(v,c) = d(v)/C$, and maximized when every $T(v,c)$ is either $a_k d(v)/\Delta$ or $b_k d(v)/\Delta$. This yields

$$\begin{aligned} &q\mathrm{e}^{-d(v)/C}\Delta \leq |L(v)| \\ &\leq \Delta \times \left(\frac{qb_k - 1}{b_k - a_k}\mathrm{e}^{-a_k d(v)/\Delta} + \frac{1 - qa_k}{b_k - a_k}\mathrm{e}^{-b_k d(v)/\Delta}\right). \end{aligned}$$

Note that these calculations can be used to show that for all d, k,

$$\ell_k^{(d)} > q\mathrm{e}^{-d/q}. \tag{1}$$

Since the colour choice for w_i can affect $|L^*(v)|$ by at most 1, Azuma's Inequality implies that $|L^*(v)|$ is highly concentrated and, in particular, that the probability of it differing from its expected value by $\Theta(\Delta)$ is $\mathrm{e}^{-\Theta(\Delta)}$. Since $d(v) > \rho\Delta$, this proves our bound on $|L^*(v)|$ in MOD-GLAUB.

Now we extend this bound to GLAUB. If the colour of w_i differs in the two procedures, then at some step after $t - 30n\log n$, w_i is assigned a colour that appears on v and on no other neighbour of w_i. Since $|L(w_i)|$ is always greater than $C - \Delta = (q-1)\Delta$, this occurs with probability at most $((q-1)n\Delta)^{-1}$ at any one time step, so the probability that it occurs at least once is at most $30n\log n/((q-1)n\Delta) < 100/D$.

So the expected number of vertices in $N(v)$ which are affected in this way is at most $100\Delta/D = O(\log n)$. A simple application of the Chernoff Bounds shows that this number is highly concentrated, and so the probability that it is higher than $\Theta(\Delta)$ is at most $\mathrm{e}^{-\Theta(\Delta)} < n^{-10}$, for $\Delta \geq D\log n$ where D is sufficiently large in terms of ν. This proves the lemma. □

Adding the probability of $n^{-10} + O(n^{-9})$ that $\mathcal{H}$ is not good, and multiplying by the $n \times n^2$ choices for v, t establishes part (a).

Part (b) follows in the same manner. We define B^* in a similar way to B, with the exception that a neighbour w_i is in B^* if (i) $w_i \in B$, (ii) it is possible, by changing the colour assigned to a neighbour u of w_i, to form a long path of disagreement from u not

passing through w_i, or (iii) some neighbour u of w_i lies in a short cycle through w_i. The same analysis shows that $\mathbf{Pr}(|B^*| > \Delta/\log\Delta) < n^{-11}$. (An extra factor of Δ appears in the expected number calculation, and this is not enough to raise that expected number significantly.) We then restrict our attention to

$$T^*_t(v,c) = \sum_{w \in N(v) - B*, c \in L_{\overline{t}(w,t)}(w)} \frac{1}{|L_{\overline{t}(w,t)}(w)|}.$$

$\mathcal{H}$ exposes B^*, all times $t_i, t_{i,j}, t_{i,j,r}$ and the colours assigned to each $u_{i,j,r}$ at time $t_{i,j,r}$. Then we modify our procedure as follows: All vertices ignore the colours on $v, w_1, ..., w_\Delta$, and after time $t_{i,j}$, the neighbours of $w_{i,j}$ ignore the colour of $w_{i,j}$. If some $w_{i,j}$ is adjacent to $w_{i',j'}$, then they ignore each others colours. We refer to this modified procedure as MOD-GLAUB2.

Consider any $w_i \notin B^*$ and let $I(w_i,c)$ be the indicator variable that $c \in L(w_i)$ at time t_i. Suppose that $d(w_i) > \rho\Delta$. As in the proof of part (a), $\mathbf{Pr}(I(w_i,c)=1) = \exp(-T_{t_i}(w_i,c)) + o(1)$. Therefore, using the bounds on $T_{t_i}(w_i,c)$ and the list sizes from the fact that $\mathcal{H}$ is good, the expected value of $I(w_i,c)/|L_{t_i}(w_i)|$ is at most

$$\begin{aligned}(\mathrm{e}^{-a_k d(w_i)/\Delta} + o(1)) \times \frac{1}{q\mathrm{e}^{-(1+\zeta)d(w_i)/C}\Delta} \\ \leq \frac{1}{\Delta} \times \mathrm{e}^{-a_k/q}/(q\mathrm{e}^{-(1+\zeta)/q}) + o(1),\end{aligned}$$

since the fact that $a_k < 1/q$ implies that the LHS is maximized at $d(w_i) = \Delta$.

We must take more care in proving our lower bound. We use L to denote $|L(w_i)|$ at time t_i, we let R be the event that $L \leq \ell^{(d(v))}_{k+1}\Delta$, and we let I_R be the indicator variable for R. In part (a), we proved that $\mathbf{Pr}(I_R = 0) = O(n^{-9})$. Therefore,

$$\begin{aligned}\mathbf{Exp}(\frac{I(w_i,c)}{L}) &\geq \mathbf{Exp}(\frac{I(w_i,c)}{L} \times I_R) \\ &\geq \frac{\mathbf{Pr}(I(w_i,c)=1) - \mathbf{Pr}(I_r = 0)}{\ell^{(d(v))}_{k+1}\Delta} \\ &= \frac{\mathbf{Pr}(I(w_i,c))}{\ell^{(d(v))}_{k+1}\Delta} + o(1/\Delta).\end{aligned}$$

Furthermore, we have

$$\begin{aligned}\ell^d_{k+1} &= \frac{1}{b_k - a_k}\Big((qb_k - 1)\mathrm{e}^{-(1-\zeta)a_k d/\Delta} \\ &\qquad +(1-qa_k)\mathrm{e}^{-(1-\zeta)b_k d/\Delta}\Big) \\ &< \frac{1}{b_k - a_k}\Big((qb_k - 1)\mathrm{e}^{-(1-\zeta)a_k} \\ &\qquad +(1-qa_k)\mathrm{e}^{-(1-\zeta)b_k}\Big) \times \mathrm{e}^{(1-\zeta)b_k(1-d/D)} \\ &= \ell^{(\Delta)}_{k+1} \times \mathrm{e}^{(1-\zeta)b_k(1-d/D)}.\end{aligned}$$

Thus, the expected value of $I(w_i,c)/|L(w_i)|$ at time t is at least

$$\begin{aligned}&(\mathrm{e}^{-b_k d(w_i)/\Delta} + o(1)) \times \frac{1}{\ell^{(d(w))}_{k+1}\Delta} \\ \geq\ & \frac{1}{\Delta} \times \frac{\mathrm{e}^{-b_k}}{\ell^{(\Delta)}_{k+1}} \\ &\qquad \times \mathrm{e}^{(b_k - (1-\zeta)b_k)(1-d(w_i)/\Delta)} + o(1) \\ \geq\ & \frac{1}{\Delta} \times \frac{\mathrm{e}^{-b_k}}{\ell^{(\Delta)}_{k+1}} + o(1) \\ \geq\ & \frac{1}{\Delta} \times \frac{\mathrm{e}^{-b_k}}{\ell^{(\Delta)}_{k}} + o(1).\end{aligned}$$

If $d(w_i) \leq \rho\Delta$, then $|L(w_i)|$ is never less than $C - \rho\Delta = (q-\rho)\Delta$. Therefore, by our assumptions on how small ρ is, the expected value of $I(w_i,c)/|L(w_i)|$ is at most

$$\frac{1}{(q-\rho)\Delta} < \frac{1}{\Delta} \times \mathrm{e}^{-a_k/q}/(q\mathrm{e}^{-(1+\zeta)/q}).$$

Furthermore, since no list is ever smaller than $(q-1)\Delta$, $\mathbf{Pr}(I(w_i,c)) \geq (1 - \frac{1}{(q-1)\Delta})^{\rho\Delta}$ which, by (1) and our assumptions on the size of ρ, is greater than $\frac{C}{\Delta} \times \mathrm{e}^{-b_k}/\ell^{(\Delta)}_k$. Therefore, the expected value of $I(w_i,c)/|L(w_i)|$ is again at least $\frac{1}{\Delta} \times \mathrm{e}^{-b_k}/\ell^{(\Delta)}_k$.

This implies that the expected value of $T^*_t(v,c)$ is at most $(a_{k+1} - \zeta)d(v)/\Delta$ and at least $(b_{k+1} - \zeta)d(v)/\Delta$. By viewing the colour assignments to $w_{i,1}, ..., w_{i,\Delta-1}$ as one single random choice, we have Δ choices, each of which can affect $T^*_t(v,c)$ by at most $1/((q-1)\Delta)$, since no $L(u)$ can have size less than $(q-1)\Delta$. Therefore by Azuma's Inequality, the probability that, under the procedure MOD-GLAUB2, $T_{v,c}$ differs from its expected value by

more than $(\zeta/2)d(v)/\Delta$ is at most $\mathrm{e}^{-\Theta(\Delta)} < n^{-10}$, for D sufficiently large in terms of ζ. (Again, we use the fact that $d(v) \geq \rho\Delta$.)

Virtually the same argument as that used for Part (a) proves that with sufficiently high probability, $T_t^*(v,c)$ in GLAUB is within $(\zeta/4)d(v)/C$ of its value in MOD-GLAUB2. Furthermore, if $\mathcal{H}$ is good then $|T_t^*(v,c) - T_t(v,c)| < (\Delta/\log\Delta) \times (1/(q-1)\Delta) = o(1)$. This proves part (b).

Finally, we need to note that, since with probability at least $1 - n^{-9}$ we have each $t_i, t_{i,j} \geq t - 30n\log n \geq 30kn\log n$, then we have by induction that the probability $\mathcal{H}$ is not good is at most $2n^{-9}$. □

Lemma 2 obviously follows from Lemma 6. We now show that Lemma 3 also does.

Proof of Lemma 3 We prove Lemma 3 in the same way as the inductive step for Lemma 6(b). We consider the procedure MOD-GLAUB2. We expose $\mathcal{H}$ and, by taking $\tau \geq 30k^*$, we can assume it is (with probability at least $1 - O(n^{-6})$) such that for each w_i with $d(w_i) > \rho\Delta$, $(\frac{1}{q} - \delta)d(w_i)/\Delta < T(w_i,c_1), T(w_i,c_2) < (\frac{1}{q} + \delta)d(w_i)/\Delta$.

Now consider any $w_i \in N(v) - B^*$ with $d(w_i) > \rho\Delta$. Let E_1, E_2 be the events that $c_1 \notin L(w_i)$, $c_2 \notin L(w_i)$ respectively. We will estimate $\mathbf{Pr}(E_1 \cap E_2) = 1 - \mathbf{Pr}(\overline{E_1}) - \mathbf{Pr}(\overline{E_2}) + \mathbf{Pr}(\overline{E_1} \cap \overline{E_2})$. By our assumption on $\mathcal{H}$, each $\mathbf{Pr}(\overline{E_r})$ is at most

$$\exp(-T(w_i,c_r)) \leq \exp(-(\frac{1}{q} - \delta)d(w_i)/\Delta).$$

Also, $\mathbf{Pr}(\overline{E_1} \cap \overline{E_2})$ is at least

$$\begin{aligned}
&\prod_{u \in N(w_i); c_1 \in L(u); c_2 \notin L(u)} \left(1 - \frac{1}{|L(u)|}\right) \\
&\quad \times \prod_{u \in N(w_i); c_2 \in L(u); c_2 \notin L(u)} \left(1 - \frac{1}{|L(u)|}\right) \\
&\quad \times \prod_{u \in N(w_i); c_1, c_2 \in L(u)} \left(1 - \frac{2}{|L(u)|}\right) \\
= \;& \exp(-T(w_i,c_1) - T(w_i,c_2)) + o(1) \\
\geq \;& \exp(-2(\frac{1}{q} + \delta)d(w_i)/\Delta) + o(1).
\end{aligned}$$

For δ sufficiently small in terms of ϵ, this yields $\mathbf{Pr}(E_1 \cap E_2) > (1 - \mathrm{e}^{-1/q})^2 - \epsilon/2$, and so the expected number of w_i with $c_1, c_2 \notin L(w_i)$ is at least $(1 - \mathrm{e}^{-1/q})^2 d(v) - (\epsilon/2)\Delta$. As before, this number is highly concentrated and so the probability that it is less than $(1 - \mathrm{e}^{-1/q})^2 d(v) - (3\epsilon/4)\Delta$ is at most n^{-10} for D sufficiently large.

As before, the probability that it differs by more than $(\epsilon/4)\Delta$ from GLAUB to MOD-GLAUB is at most n^{-10} for D sufficiently large. Also, if the number of $w_i \in N(v) - B^*$ with $c_1 \notin L(w_i)$, $c_2 \notin L(w_i)$ is large, then of course, the number of such w_i in $N(v)$ is large. This proves Lemma 3. □

3 Path Coupling

Here, we prove Theorem 1. We consider a fixed small ϵ, ρ to be named later. We begin with a burn-in period of $\tau n \log n$ steps where $\tau > 30$ is as in Lemmas 2, 3. All of our analysis will assume that $\tau n \log n < t < n^2$. For now, we assume that for every vertex v, and pair of colours c_1, c_2, we have (i) $|L(v)| \geq C\mathrm{e}^{-d(v)/C} - \epsilon\Delta$ and (ii) v has fewer than $\gamma d(v) - \epsilon\Delta$ neighbours w with $d(w) \geq \rho\Delta$ and $c_1, c_2 \notin L(w)$. Later we will account for the $O(n^{-5})$ probability that this is not the case.

We use the path coupling technique of Bubley and Dyer[1]. To do so, we couple two chains $X_0, X_1, \ldots$ and $W_0, W_1, \ldots$ with arbitrary initial colourings X_0, W_0, and show that, with high probability, they coincide within $O(n\log n)$ steps. To prove this fact using path coupling, at any time t, we consider a "path" of possibly improper colourings $X_t = Z_0, Z_1, \ldots, Z_h = W_t$. We define this path as follows. Consider an arbitrary ordering of the vertices $v_1, \ldots, v_n$. To form Z_1 from Z_0, we change the colour of the first vertex on which X_t and W_t differ from its colour in X_t to its colour in W_t. To form Z_2 from Z_1, we change the colour of the second vertex on which X_t, W_t differ, and so on. Thus, h is the Hamming distance between X_t, W_t, i.e. the number of vertices on which they differ. (If $X_t = Y_t$ then $h = 0$ and $X_t = Z_0 = Y_t$.)

We couple the chains as follows: We carry out a step of $Z_0 = X_t$ thus obtaining $Z_0' = X_{t+1}$. Then we maximally couple a random choice for Z_1 to the choice for Z_0, thus obtaining Z_1'. (Recall that, even though Z_r may not be a proper colouring, we can apply a step of our process to it.) Repeatedly, we

maximally couple Z_i to Z_{i-1} obtaining $Z_i^{'}$, finally yielding $W_{t+1} = Z_h^{'}$.

For any two (not necessarily proper) colourings X, W, we define their *weighted Hamming distance* $H^{'}(X,W)$ to be the number of vertices v with $d(v) > \rho\Delta$ and $X(v) \neq W(v)$ plus 2ρ times the number of vertices v with $d(v) \leq \rho\Delta$ and $X(v) \neq W(v)$. Note that $H^{'}(X_t, W_t) = \sum_{i=0}^{h-1} H^{'}(Z_i, Z_{i+1})$. Note further that $H^{'}(X_{t+1}, W_{t+1}) \leq \sum_{i=0}^{h-1} H^{'}(Z_i^{'}, Z_{i+1}^{'})$, since if $X(v) \neq W(v)$ then $Z_i^{'}(v) \neq Z_{i+1}^{'}(v)$ for at least one i.

We will prove that, after an $O(n \log n)$ burn-in period, the expected value of the change of the weighted Hamming distance between any pair Z_r, Z_{r+1} is at most $-\psi/n$ for some constant $\psi > 0$. Thus, the expected value of the change of the weighted Hamming distance between X, W is at most $-h \times \psi/n < -\psi/n$.

To prove this, we need to know that, with high probability, the bounds in Lemmas 2, 8 apply to each Z_r. So for each $0 \leq s \leq n$ and step t, we define M_t^s to be the (possibly improper) colouring in which vertices $v_1, ..., v_s$ have their colour from W_t and $v_{s+1}, ..., v_n$ have their colours from X_t. Note that, at time t, each Z_r is equal to M_t^s for at least one value of s. At time t, for each vertex v and for each $0 \leq s \leq n$, we define $L^s(v)$ to be the set of colours which do not appear in M_t^s on the neighbourhood of v. $L^X(v) = L^0(v)$ is the set of colours which do not appear in X on $N(v)$, and $L^W(v) = L^n(v)$ is the set of colours which do not appear in W on $N(v)$.

Lemma 9 *For every $\epsilon, \rho > 0$, there exist constants D, τ such that with probability at least $1 - n^{-5}$ for every vertex v with $d(v) > \rho n$, $0 \leq s \leq n$ and time $\tau n \log n \leq t \leq n^2$, we have $C\mathrm{e}^{-\overline{d(v)/C}} - \epsilon\Delta < |L^s(v)| < C\mathrm{e}^{-\overline{d(v)/C}} + \epsilon\Delta$.*

Lemma 10 *For every $\epsilon, \rho > 0$, there exist constants D, τ such that with probability at least $1 - n^{-5}$, for every vertex v, colours c_1, c_2, $0 \leq s \leq n$ and time $\tau n \log n \leq t \leq n^2$, v has fewer than $(1 - (1 - \mathrm{e}^{-1/q})^2)d(v) - \epsilon\Delta$ neighbours w with $d(w) \geq \rho\Delta$ and at least one of c_1, c_2 in $L^s(w)$.*

To prove Lemma 9 we modify the proof of Lemma 2 as follows. For each $w_i = v_j$, $\mathcal{H}$ exposes the colours that $w_{i,1}, ..., w_{i,\Delta-1}$ receive in the chain W if $j \leq s$, and exposes the colours they recieve in X otherwise. The rest of the proof is the same. The exponent of n in the probability bound changes from -6 to -5 because of the extra n choices for s.

To prove Lemma 10, we make a similar change. We define

$$T_t^s(v,c) = \sum_{v_j \in N(v); j \leq s; c \in L^W_{\overline{t}(v_j,t)}(v_j)} \frac{1}{|L^W_{\overline{t}(v_j,t)}(v_j)|} + \sum_{v_j \in N(v); j > s; c \in L^X_{\overline{t}(v_j,t)}(v_j)} \frac{1}{|L^X_{\overline{t}(v_j,t)}(v_j)|},$$

and we modify Lemma 6 to:

Lemma 11 *For each $1 \leq k \leq k^*$, with probability at least $1 - n^{-6}$, for every v with $d(v) > \rho\Delta$, colour c, $0 \leq s \leq n$ and $30kn \log n \leq t \leq n^2$, we have at time t:*

(a) $q\mathrm{e}^{-(1+\zeta)d(v)/C}\Delta \leq |L^X(v)|, |L^W(v)| \leq \ell_k^{(d(v))}\Delta$;

(b) $a_k d(v)/\Delta \leq T_t^s(v,c) \leq b_k d(v)/\Delta$.

Part (a) is essentially the same as in Lemma 6. To prove part (b), for each $w_i = v_j$, $\mathcal{H}$ exposes the colours that $w_{i,1,1}, ..., w_{i,\Delta-1,\Delta-1}$ receive in the chain W if $j \leq s$, and exposes the colours they receive in X otherwise. The rest of the proof is the same. Now we prove Lemma 10 from Lemma 11 by defining $\mathcal{H}$ in the same way. □

For now we assume that for every vertex v, $0 \leq s \leq n$ and pair of colours c_1, c_2, we have (i) $|L^s(v)| \geq C\mathrm{e}^{-d(v)/C} - \epsilon\Delta$ and (ii) v has fewer than $(1 - (1 - \mathrm{e}^{-1/q})^2)d(v) + \epsilon\Delta$ neighbours w with $d(w) > \rho\Delta$ and at least one of c_1, c_2 in $L^s(w)$. Later we will account for the $O(n^{-5})$ probability that this is not the case.

Now, consider any Z_r and Z_{r+1}. They differ on exactly one vertex, say v which has colour c_1 in Z_r and c_2 in Z_{r+1}. We apply one step of our process to Z_r and to Z_{r+1}, coupled as described in the introduction.

Suppose that $d(v) > \rho\Delta$. Then the weighted Hamming distance between Z_r and Z_{r+1} decreases by 1 iff we select v. This has probability $1/n$ of

occurring. The Hamming distance increases iff we choose a neighbour u of v, and assign it colour c_1 in Z_r and/or assign it colour c_2 in Z_{r+1}. If $d(u) > \rho\Delta$ then it increases by 1, otherwise it increases by 2ρ. There are at most $(1-(1-\mathrm{e}^{-1/q})^2+\epsilon)\Delta$ neighbours of v with degree greater than $\rho\Delta$ and for which it is possible to make at least one of those two assignments. Of course, there are at most Δ neighbours of v with degree at most $\rho\Delta$ and for which it is possible to make at least one of those two assignments. By Lemma 9, each neighbour has lists of size at least $(q\mathrm{e}^{-1/q}-\epsilon)\Delta$ in both Z_r and Z_{r+1}, since $C\mathrm{e}^{-d(v)/C} \geq q\mathrm{e}^{-1/q}\Delta$ for $d(v) \leq \Delta$. Therefore, the probability that the Hamming distance increases by 1 is at most $(1-(1-\mathrm{e}^{-1/q})^2+\epsilon)\Delta/(n(q\mathrm{e}^{-1/q}-\epsilon)\Delta)$, and the probability that it increases by 2ρ is at most $\Delta/(n(q\mathrm{e}^{-1/q}-\epsilon)\Delta)$. Therefore, the expected change in the Hamming distance is at most

$$\frac{1}{n} \times \left(-1 + \frac{1-(1-\mathrm{e}^{-1/q})^2+\epsilon+2\rho}{q\mathrm{e}^{-1/q}-\epsilon} \right)$$

which is negative if we choose ϵ and ρ to be sufficiently small in terms of q, since $(1-\mathrm{e}^{-1/q})^2 + q\mathrm{e}^{-1/q} > 1$.

Suppose that $d(v) \leq \rho\Delta$. Then the weighted Hamming distance decreases by 2ρ with probability $\frac{1}{n}$ and it increases with probability at most $\rho\Delta/n(q\mathrm{e}^{-1/q}-\epsilon)\Delta$. Since it never increases by more than 1, the expected change in the Hamming distance is at most

$$\frac{1}{n} \times \left(-2\rho + \frac{\rho}{q\mathrm{e}^{-1/q}-\epsilon} \right),$$

which is negative for ϵ sufficiently small since $q\mathrm{e}^{-1/q}$ is easily seen to be greater than $\frac{1}{2}$ for $q > q^*$.

Thus, in either case, the expected change in the weighted Hamming distance is less than $-\psi$ for some $\psi = \psi(q) > 0$. This implies that with sufficiently high probability, the weighted Hamming distance drops to 0 within $O(n \log n)$ steps.

Now we still have to account for the $O(n^{-5})$ chance that for some v, t, s, c_1, c_2, $L^s(v)$ is too small or $N(v)$ has too many high degree neighbours with either c_1 or c_2 in their lists. Let t^* be the first time at which this occurs. Consider the random variable $H^*(t)$ defined as follows. Until time t^*, $H^*(t) = H'(X_t, W_t)$. After time t^*, if $H^*(t-1) > 0$ then $H^*(t) = H^*(t-1) - 1$ with probability $\frac{1}{n}$ and $H^*(t) = H^*(t-1) + 1$ with probability $\frac{1-\psi}{n}$; if $H^*(t-1) = 0$ then $H^*(t) = 0$. $H^*(t)$ is a simple random walk with negative drift after time $t = \tau n \log n$, and it is straightforward to verify that with high probability, $H^*(t) = 0$ when $t = O(n \log n)$. Furthermore, with probability $1 - O(n^{-4})$, $H(X_t, W_t) = H^*(t)$ for each $1 \leq t \leq n^2$. Therefore, with high probability, $H(X_t, W_t) = 0$ when $t = O(n \log n)$. This is enough to prove Theorem 1. See, for example, [4] for the standard argument. □

4 Some final comments

In this section, we note that this coupling argument cannot be used for the case $C = q\Delta$ for any $q < q^*$. To see this, consider any polynomial n^x and any regular graph G with $\Delta \geq D \log n$ and girth at least $D \log \Delta$ for some sufficiently large D in terms of x. Suppose that $1.4890 \leq q < q^*$.

Note that Lemmas 2 to 11 hold for all $q > 1.4890$ (and in fact, if needed, we could show that they hold for even smaller q). The only place where we required $q > q^*$ was in the Proof of Theorem 1 in Section 3. Furthermore, the upper bound $t \leq n^2$ in their statements can be easily increased to n^x. Therefore, if the Glauber dynamics mixes rapidly, then a "typical" colouring will satisfy that for all v, c, $|L_v|$ is arbitrarily close to $C\mathrm{e}^{-1/q}$ and $T(v, c)$ is arbitrarily close to $1/q$. Thus, such a colouring must exist, call it Ψ.

Suppose that we choose Ψ as our initial colouring. Then for the first n^x steps, with high probability, for all v, c, $|L_v|$ is arbitrarily close to $C\mathrm{e}^{-1/q}$ and $T(v, c)$ is arbitrarily close to $1/q$. This implies that with high probability, for any v, c_1, c_2, the number of neighbours of v which have either c_1 or c_2 in their list is arbitrarily close to $(1-(1-\mathrm{e}^{-1/q})^2)\Delta$. So if we couple the colouring arising at any time less than n^x with another colouring which differs in exactly one vertex, then since $q < q^*$, the expected change in their Hamming distance will be positive.

Furthermore, since the graph has large girth, we cannot apply the technique from [5] and [8] where

by analyzing the expected total change over a few steps, we were able to get some gain from edges in $N(v)$.

This raises the following question:

Question: *Is there any $q < q^*$ and $D > 0$ such that the Glauber dynamics for $q\Delta$-colourings mixes in polytime on graphs with girth at least D and maximum degree Δ at least $D \log n$?*

A negative answer to this question would be very interesting since it would establish that for some (in fact many) graphs, the Glauber dynamics does not mix rapidly for values of C which are much higher than Δ.

Acknowledgements: The author is grateful to Martin Dyer and Alan Frieze for some helpful discussions, for providing him with an early copy of their paper [2], and for pointing out an error in an early version.

References

[1] R. Bubley and M. Dyer. *Path coupling: a technique for proving rapid mixing in Markov chains.* Proceedings of the 28th Annual Symposium on Foundations of Computer Science (1997), 223 - 231.

[2] M. Dyer and A. Frieze. *Randomly colouring graphs with bounds on girth and maximum degree.* Proceedings of the 32nd Annual Symposium on Foundations of Computer Science (2001).

[3] M. Dyer, L. Goldberg, C. Greenhill, M. Jerrum and M. Mitzenmacher, *An extension of path coupling and its application to the Glauber dynamics for path colourings,* Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms (2000), 355 - 363.

[4] M.Dyer and C. Greenhill. *Random walks on combinatorial objects.* Surveys in Combinatorics, 1999, eds. J.D. Lamb and D.A. Preece. Cambridge University Press, 1999, 101 - 136.

[5] M. Dyer, C. Greenhill and M. Molloy. *Very rapid mixing of the Glauber dynamics for proper colourings on bounded-degree graphs.* Random Structures and Algorithms (to appear).

[6] M. 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.

[7] M. Jerrum, *Mathematical foundations of the Markov chain Monte Carlo method.* Probabilistic Methods for Algorithmic Discrete Mathematics, eds. M. Habib, C. McDiarmid, J. Ramirez-Alfosin and B. Reed. Springer, 1998, 116 - 165.

[8] M. Molloy. *Very rapidly mixing Markov Chains for 2Δ-colourings and for independent sets in a 4-regular graph.* Random Struc. & Alg. **18**, 101 - 115 (2001).

[9] J. Salas and A. Sokal. *Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem.* J. Stat. Physics **86** (1997), 551 - 579.

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

[11] E. Vigoda. *Improved bounds for sampling colourings.* J. Mathematical Physics **41** (2000), 1555 - 1569.

5 Appendix: Proof of Lemma 4

Recall that we can assume that $q > q^*$ is sufficiently small, and so we will take $q < 1.49$.

It is straightforward to show that α_k is strictly increasing and β_k is strictly decreasing, and so by Lemma 5, $\alpha = \lim_{k\to\infty} \alpha_k, \beta = \lim_{k\to\infty} \beta_k, \lambda = \lim_{k\to\infty} \lambda_k^{(\Delta)}$ exist. They must satisfy

$$\begin{aligned} \alpha &= \mathrm{e}^{-\beta}/\lambda \\ \beta &= \mathrm{e}^{-\alpha}/(q\mathrm{e}^{-1/q}) \\ \lambda &= \frac{q\beta - 1}{\beta - \alpha}\mathrm{e}^{-\alpha} + \frac{1 - q\alpha}{\beta - \alpha}\mathrm{e}^{-b} \end{aligned}$$

We will prove that this system has no roots for $0 \leq \alpha < 1/q$. This, along with Lemma 5, implies Lemma 4.

Rearranging the first equation of our system, we get $f(\alpha) = \alpha\lambda - \mathrm{e}^{-\beta} = 0$. We will bound the derivative of f with respect to α. We start by bounding

$$g(\alpha) = 1 - \beta q + \beta - \alpha\beta q.$$

Clearly $g(\alpha) \to 0$ as $\alpha \to 1/q$. Also, noting that the derivative of β with respect to α is $-\beta$, we have $g'(\alpha) = q\beta - \beta - q\beta + q\alpha\beta = q\alpha\beta - \beta < 0$ for $\alpha < 1/q$. Therefore, $g(\alpha) > 0$ for $\alpha < 1/q$.

Now, using the fact that by (1) $q\mathrm{e}^{-1/q} \le \lambda \le q\mathrm{e}^{-\alpha}$, we have:

$$\begin{aligned}
f'(\alpha) &= \lambda - \beta\mathrm{e}^{-\beta} + \frac{\alpha}{\beta-\alpha}\Big((1-2\beta q)\mathrm{e}^{-\alpha} \\
&\qquad +(\beta - q - \alpha\beta q)\mathrm{e}^{-\beta} + (\beta+1)\lambda\Big) \\
&\ge \lambda - \beta\mathrm{e}^{-\beta} \\
&\quad +\frac{\alpha\mathrm{e}^{-\beta}}{\beta-\alpha}(1 - 2\beta q + \beta - q - \alpha\beta q + q\beta + q) \\
&\quad +\frac{\mathrm{e}^{-\alpha} - \mathrm{e}^{-\beta}}{\beta-\alpha}(\alpha - 2\alpha\beta q) \\
&\ge \lambda - \beta\mathrm{e}^{-\beta} + \frac{\mathrm{e}^{-\alpha} - \mathrm{e}^{-\beta}}{\beta-\alpha}(\alpha - 2\beta) \\
&\ge q\mathrm{e}^{-1/q} - \beta\mathrm{e}^{-\beta} - (\mathrm{e}^{-\alpha} - \mathrm{e}^{-\beta}) - \beta\frac{\mathrm{e}^{-\alpha} - \mathrm{e}^{-\beta}}{\beta-\alpha} \\
&\ge q\mathrm{e}^{-\beta} - \beta\mathrm{e}^{-\beta} - (\mathrm{e}^{-\alpha} - \mathrm{e}^{-\beta}) - \beta\mathrm{e}^{-\beta}\frac{\mathrm{e}^{-(\alpha-\beta)} - 1}{\beta-\alpha} \\
&\ge q\mathrm{e}^{-\beta} - \beta\mathrm{e}^{-\beta} - (\mathrm{e}^{-\alpha} - \mathrm{e}^{-\beta}) - \beta\mathrm{e}^{-\beta}(1 + \frac{1}{2}(\beta-\alpha)) \\
&= (q - 2\beta)\mathrm{e}^{-\beta} - (\mathrm{e}^{-\alpha} - \mathrm{e}^{-\beta}) - (\beta-\alpha)\beta\mathrm{e}^{-\beta}.
\end{aligned}$$

As $\alpha \to 1/q$, we get $\alpha \to \beta$ and so the latter two terms tend to 0. So this allows us to bound $f'(\alpha)$ away from 0 when α is close to the discontinuity at $\alpha = 1/q$. In particular, for $1.489 < q < 1.49$ we have $f'(\alpha) > .01$ when $.64 \le a < 1/q$ and so $f(\alpha)$ has no roots in that range. Having dealt with this discontinuity, it is straightforward to check that for the same range of q, $f(\alpha)$ has no roots in $0 \le \alpha \le .58$, thus proving the lemma. □