Randomly Colouring Graphs with Girth Five and Large Maximum Degree

Lap Chi Lau and Michael Molloy

Department of Computer Science,
University of Toronto,
Toronto, Ontario, Canada
{chi, molloy}@cs.toronto.edu

Abstract. We prove that the Glauber dynamics on the k-colourings of a graph G on n vertices with girth 5 and maximum degree $\Delta \geq 1000 \log^3 n$ mixes rapidly if $k = q\Delta$ and $q > \beta$ where $\beta = 1.645...$ is the root of $2 - (1 - e^{-1/\beta})^2 - 2\beta e^{-1/\beta} = 0$.

1 Introduction

The Glauber dynamics is a Markov chain on the proper colourings of a graph that has been widely studied in both computer science and statistical physics. For a given graph G and integer k which is at least the chromatic number of G, the Markov chain is described as follows: We start with an arbitrary k-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.

This chain is of great interest for a number of 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 also of interest in the statistical physics community, in part because of its relation to the Potts model.

The main question in this area is: For what values of k 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 $k \leq \Delta + 1$. It is conjectured that for every graph, the chain mixes in polytime for $k \geq \Delta + 2$, or at least for $k \geq \Delta + o(\Delta)$, but this appears to be a very difficult conjecture. Jerrum[11], and independently Salas and Sokal[14], showed that for all graphs the chain mixes in polytime for $k \geq 2\Delta$. Vigoda[15] showed that for all graphs, a different chain mixes in optimal time for $k \geq \frac{11}{6}\Delta$ and this implies that for the same values of k, the Glauber dynamics mixes in polytime. This is the best progress to date for general graphs.

A recent trend has been to study the performance of the Glauber dynamics on graphs with restrictions on the girth and maximum degree. At first, these restrictions were rather severe, and the number of colours remained far from Δ: Dyer and Frieze[4] showed that if Δ is at least $O(\log n)$ and the girth is at least

J.R. Correa, A. Hevia, and M. Kiwi (Eds.): LATIN 2006, LNCS 3887, pp. 665–676, 2006.

$O(\log \Delta)$ then we have rapid mixing for roughly $k = 1.763\Delta$ colours (note that $1.763 < 11/6$). Since then, several improvements[12, 7, 8, 9, 10, 5, 6] have reduced these restrictions substantially, and this line of research is producing surprisingly strong results and shedding much insight on the general conjecture. Some notable results are that we obtain rapid mixing for $\Delta = O(\log n)$, girth at least 9, and $k \geq (1 + \epsilon)\Delta$[8] and for Δ at least a particular large constant, girth at least 6 and k roughly 1.489Δ[5].

Recently, Hayes and Vigoda[10] introduced "coupling from the stationary distribution" (described below) with which they managed to improve the girth requirement from five to four in one of these results (from Hayes[7]). (An improvement of 1 may not seem like much at first glance, but when the numbers are this small, each such improvement can be a huge gain.) They showed that we have rapid mixing with $\Delta = O(\log n)$, girth at least 4 and with k roughly 1.763Δ. This value of k is one that is often obtained by using a particular property that we call the *first local uniformity condition* (defined below). Hayes[7] had also proved rapid mixing $\Delta = O(\log n)$, girth at least 6 and with k roughly 1.489Δ, a value that is often obtained by using the *second local uniformity condition.* The main result of this paper, is to incorporate that second local uniformity condition into a coupling from the stationary distribution argument, and reduce the girth requirement from the latter result to 5. In doing so, difficulties cost us in two ways: (i) we must increase the restriction on Δ somewhat, and (ii) we obtain a number larger than the usual 1.489....

Define $\beta = 1.645...$ to be the solution to

$$2 - (1 - e^{-1/\beta})^2 - 2\beta e^{-1/\beta} = 0.$$

Theorem 1. *The Glauber dynamics mixes in $O(n \log n)$ time on all graphs on n vertices with maximum degree $\Delta \geq 1000 \log^3 n$, when the number of colours is $k \geq (\beta + \epsilon)\Delta$ for any constant $\epsilon > 0$.*

Remark. We made no attempt to optimize the exponent "3" in the lower bound on Δ. It is not hard to reduce it somewhat.

1.1 Outline

The proof of our main result uses the framework of "coupling with the stationary distribution" developed by Hayes and Vigoda[10] to prove their aforementioned result. Here is the basic idea: To analyse the mixing time via a coupling argument, we can assume that one Markov chain X is distributed according to the uniform distribution. Given a graph of girth at least 4 and maximum degree $\Delta = \Omega(\log n)$, one can show that with high probability, X_t has the first local uniformity condition. Hayes and Vigoda then show that, given an *arbitrary* colouring Y_t, the Hamming distance between X_t and Y_t decreases in expectation for k roughly 1.763Δ so long as X_t has the first local uniformity. So, with high probability, X_t and Y_t tend to drift together, and their theorem follows.

The main advantage of using the coupling with the stationary distribution is that one only needs to prove that a uniformly random colouring has local uniformity properties rather than a colouring generated by the Markov chain. This allows one to skip the analysis of the burn-in period, which is the most technical part of many previous papers. In addition, short cycles are a bit less harmful in uniform colourings than in "burn-in" colourings; this allowed the girth requirement to be reduced by one in [10]. One substantial drawback to this technique is that it does not accommodate path-coupling, a very useful technique introduced in [2]. This means that one needs to analyse the expected change in Hamming distance between two colourings with *arbitrary* Hamming distance, rather than just analyzing the much simpler case where the Hamming distance is one. Carrying out that analysis turned out to be manageable in [10] where they were able to adopt the original coupling argument from Jerrum[11], which predated path-coupling.

The main thrust of this paper is to incorporate the second local uniformity into the framework of "coupling with the uniform condition". In this case, it is much more difficult to extend the path coupling analysis to the case where two colourings have arbitrary distance. In fact, we are unable to do so without some loss, and this is why our result requires 1.645Δ colours rather than 1.489Δ. This portion of our analysis makes use of a novel "charging" argument (Lemma 1). That argument does not make use of any special structure of G (such as its girth or maximum degree) and so it might be useful in other settings. This argument appears in Section 2.

A second difficulty that arises in this paper is in proving that the second uniformity condition holds for a uniformly random colouring when the girth requirement is reduced from 6 (in [7]) to 5. The main problem is that the second local uniformity condition is defined in terms of vertices of distance two from a specific vertex v. Every previous paper that established a uniformity condition made crucial use of the fact that the vertices which defined the condition were very close to being an independent set. This is true in our setting for girth 6 graphs, but girth 5 graphs can have many edges between those vertices. The difficulties caused by these edges are what require us to increase the bound on Δ from $O(\log n)$ to $O(\log^3 n)$. We present this part of the proof in Section 3.

Remark. Our main theorem applies to graphs with maximum degree Δ. However, for brevity and ease of presentation, we only present the proof for the case where the graph is Δ-regular. For the most part, it is straightforward to extend the proof to non-regular graphs. The material in Section 3 is not as straightforward to extend, but the arguments used in [12] will suffice.

1.2 Definitions

In a graph G, we define $N(v)$ to be the set of neighbours of vertex v.

For a colouring X of G, we define $X(v)$ to be the colour at vertex v. We denote by $L_X(v)$ the list of available colours at v in X; i.e. the colours that do not appear on any neighbours of v. We denote by L_X the minimum of $|L_X(v)|$ over all possible v. Given two colourings X, Y, $P_v(X,Y) := L_Y(v) - L_X(v)$ and $P_v(Y,X) := L_X(v) - L_Y(v)$. In other words, $P_v(X,Y)$ is the set of colours appearing in the neighbourhood of v in X but not appearing in the neighbourhood of v in Y. Given a colouring X, suppose we recolour v by colour c and denote the resulting colouring by X'; then $R_X(v,c)$ is defined to be the set of neighbours w of v such that $L_X(w) = L_{X'}(w)$. In other words, $R_X(v,c)$ is the set of vertices $w \in N(v)$ such that $X(v)$ and c both appear in $N_G(w) - v$. We further define R_X to be the minimum of $|R_X(v,c)|$ over all possible v and c.

For the purposes of this paper, the local uniformity conditions are defined as follows. Set $q = k/\Delta$.

First Local Uniformity Condition[4]
For every $\zeta > 0$, $(qe^{-1/q} - \zeta)\Delta < L_X < (qe^{-1/q} + \zeta)\Delta$.

Second Local Uniformity Condition[12]
For every $\zeta > 0$, $((1 - e^{-1/q})^2 - \zeta)\Delta < R_X < ((1 - e^{-1/q})^2 + \zeta)\Delta$.

Given a particular value of k, we define Ω to be the set of k-colourings of G.

1.3 A Concentration Tool

We will make use of the following inequality, which is particularly useful in this paper because it can be applied to random trials that are not independent. The version that we use is from [13] and is a distillation of Azuma's original statement[1].

Azuma's Inequality. *Let X be a random variable determined by n trials $T_1, ..., T_n$, such that for each i, and any two possible initial sequences of outcomes $t_1, ..., t_i$ and $t_1, ..., t_{i-1}, t_i'$ that differ only on the ith outcome:*

$$|\exp(X|T_1 = t_1, ..., T_i = t_i) - \exp(X|T_1 = t_1, ..., T_i = t_i')| \le \gamma_i$$

then

$$\mathbf{Pr}(|X - \exp(X)| > \tau) \le 2e^{-\tau^2/(2\sum_{i=1}^{n}\gamma_i^2)},$$

for every $\tau > 0$.

2 Distance Decreasing with the Local Uniformities

Consider two colourings X, Y of G. We use $d(X,Y)$ to denote the Hamming distance of X, Y; i.e. the number of vertices on which they differ. The key lemma in this paper is the following:

Lemma 1. *For any two colourings X, Y of a Δ-regular graph,*

$$\sum_{w \in V} \max\{P_w(X,Y), P_w(Y,X)\} \le (1 - \frac{R_X}{2})\Delta d(X,Y).$$

We defer its proof until the end of the section.

Let X', Y' denote random colourings generated by applying one step of the Glauber dynamics to X, Y respectively. Following the notation in [10], we say that X, Y are δ-*distance decreasing* if there exists a coupling of X', Y' under which the expected value of $d(X', Y')$ is at most $(1-\delta)d(X, Y)$.

Recall that $\beta = 1.645...$ is defined in the introduction. Using Lemma 1, it is fairly straightforward to prove the following:

Lemma 2. *Suppose that $k \geq (\beta+\epsilon)\Delta$ for some $\epsilon > 0$. Then there exists $\zeta, \delta > 0$ such that if $X \in \Omega$ satisfies the first and the second local uniformity conditions for $\zeta > 0$, then for every $Y \in \Omega$, (X, Y) is δ-distance-decreasing.*

Proof. We need to prove, for every $Y \in \Omega$,

$$\mathbf{E}(d(X', Y')) \leq (1 - \frac{\delta}{n})d(X, Y)$$

for some $\delta > 0$. Let v be the vertex selected for recolouring at the first time step. First we bound the probability that the chains recolour v to different colours. For a colour c available to both chains, c will be chosen in X with probability $\frac{1}{|L_X(v)|}$ and in Y with probability $\frac{1}{|L_Y(v)|}$, and hence c will be chosen in both chains with probability $\frac{1}{\max\{|L_X(v)|,|L_Y(v)|\}}$ if we use, as usual, Jerrum's coupling. Therefore, the probability that v will be coloured differently is:

$$\begin{aligned}\mathbf{Pr}(X'(v) \neq Y'(v) \mid v) &= 1 - \frac{|L_X(v) \cap L_Y(v)|}{\max\{|L_X(v)|, |L_Y(v)|\}} \\ &= \frac{\max\{|L_X(v)|, |L_Y(v)|\} - |L_X(v) \cap L_Y(v)|}{\max\{|L_X(v)|, |L_Y(v)|\}} \\ &= \frac{\max\{|P_v(X, Y)|, |P_v(Y, X)|\}}{\max\{L_X(v), L_Y(v)\}}, \qquad (1)\end{aligned}$$

recall that $P_v(X, Y) := L_Y(v) - L_X(v)$ and $P_v(Y, X) := L_X(v) - L_Y(v)$. Now we bound the expected distance after one step.

$$\begin{aligned}&\mathbf{E}(d(X', Y')) \\ &= \sum_{w \in V} \mathbf{Pr}(X'(w) \neq Y'(w)) \\ &= \sum_{w \in V} \mathbf{Pr}(v \neq w \wedge X(w) \neq Y(w)) + \sum_{w \in V} \mathbf{Pr}(v = w \wedge X'(w) \neq Y'(w)) \\ &= \frac{n-1}{n} d(X, Y) + \frac{1}{n} \sum_{w \in V} \mathbf{Pr}(X'(w) \neq Y'(w) \mid v = w) \\ &= \frac{n-1}{n} d(X, Y) + \frac{1}{n} \sum_{w \in V} \frac{\max\{|P_w(X, Y)|, |P_w(Y, X)|\}}{\max\{|L_Y(w)|, |L_X(w)|\}} \quad \text{(by (1))} \\ &\leq \frac{n-1}{n} d(X, Y) + \frac{1}{nL_X} \sum_{w \in V} \max\{P_w(X, Y), P_w(Y, X)\} \\ &\leq \frac{n-1}{n} d(X, Y) + \frac{1}{nL_X} (1 - \frac{R_X}{2}) \Delta d(X, Y) \quad \text{(by Lemma 1)}\end{aligned}$$

$$\leq \frac{n-1}{n} d(X,Y) + \frac{1}{n} \frac{(2-(1-e^{-\frac{\Delta}{k}})^2+\zeta)\Delta}{2ke^{-\frac{\Delta}{k}}-\zeta} d(X,Y)$$
$$\leq \frac{n-1}{n} d(X,Y) + \frac{1}{n}(1-\delta) d(X,Y) = (1-\frac{\delta}{n}) d(X,Y) \quad \text{(since } k \geq (\beta+\epsilon)\Delta\text{)}$$

for some $\delta > 0$, if we take ζ to be sufficiently small in terms of ϵ. The second last inequality follows from the local uniformity properties. ∎

The following theorem about couplings which "usually" decrease distances is by Hayes and Vigoda[10].

Theorem 2. *Let $X_0, \ldots, X_T$, $Y_0, \ldots, Y_T$ be coupled Markov chains such that, for every $0 \leq t \leq T-1$,*

$$\mathbf{Pr}((X_t, Y_t) \text{ is not } \delta \text{ distance decreasing}) \leq \epsilon.$$

Then

$$\mathbf{Pr}(X_t \neq Y_t) \leq ((1-\delta)^T + \epsilon/\delta)\, diam(\Omega).$$

In Section 3, we will prove (Lemma 3) that if G is a graph of girth 5 and maximum degree $\Delta \geq 1000 \log^3 n$, and if X is a uniformly random k-colouring of G where $k \geq (1+\epsilon)\Delta$ for some $\epsilon > 0$ then X satisfies the first and second local uniformity properties. That will allow us to apply the preceding lemmas to such graphs, and thus prove the main result of this paper, which we do now.

Proof (**Proof of Theorem 1).** The proof is along the same line as in Hayes and Vigoda[10]. Here we just give a quick sketch. For ease of exposition, we assume that G is Δ-regular.

Let X_0 be distributed according to π (the uniform distribution) and Y_0 be arbitrary. Generate $X_1, \ldots, X_T$, $Y_1, \ldots, Y_T$ using Jerrum's coupling with initial states X_0, Y_0. For every $t \geq 0$, X_t is distributed according to π. By Lemma 3, X_t has the first and the second local uniformity properties with sufficiently high probability. Hence, by Lemma 2, X_t and Y_t are δ-distance decreasing with high probability. Now, applying Theorem 2 gives the theorem. ∎

Finally, we close this section by proving the key lemma.

Proof (**Proof of Lemma 1).** Let $d := d(X,Y)$ be the Hamming distance between X and Y, and $v_1, \ldots, v_d$ be the d vertices with different colours in X and Y. Let Z_i be a colouring equal to X except $Z_j(v_j) = Y(v_j)$ for $1 \leq j \leq i$. Let $P_w(i) := \max\{P_w(X, Z_i), P_w(Z_i, X)\}$. So, $Z_d = Y$ and $\sum_{w \in V} P_w(d)$ is the value we would like to bound. To bound $\sum_{w \in V} P_w(d)$, we consider $\sum_{w \in V} P_w(i)$ for $1 \leq i \leq d$. Intuitively, we consider the colour changes one-at-a-time.

Note that $P_w(i) > P_w(i-1)$ only when w is a neighbour of v_i, and note also that $P_w(i) \leq P_w(i-1) + 1$ by definition. Since the maximum degree in G is Δ, it follows that $\sum_{w \in V} P_w(d) \leq d\Delta$. With the second local uniformity, however, we can give a better bound. For example, when the colour of v_1 is changed from $X(v_1)$ to $Y(v_1)$, for each vertex w in $R_X(v_1, Y(v_1))$, both colours $X(v_1)$

and $Y(v_1)$ appear in $N_G(w) - v$ and thus $L_X(w)$ and $L_{Z_1}(w)$ are the same. Hence, $P_w(1) = P_w(0)$ for $w \in R_X(v_1, Y(v_1))$. Since $|R_X(v_1, Y(v_1))| \geq R_X \Delta$ by definition, we have $\sum_{w \in V} P_w(1) \leq (1 - R_X)\Delta$. Notice that the above argument does not hold in general at time i for $i > 1$, since the colours have been changed at $v_1, \ldots, v_{i-1}$. But one may still hope that $P_w(i) = P_w(i-1)$ for "many" vertices in $R_X(v_i, Y(v_i))$. In light of this, we say that a *good* event happens at w at time i if $P_w(i) = P_w(i-1)$ when the colour of v_i is changed from $X(v_i)$ to $Y(v_i)$; otherwise a *bad* event if $P_w(i) = P_w(i-1) + 1$ at w at time i when the colour of v_i is changed from $X(v_i)$ to $Y(v_i)$. In the following, we focus on a bad event at w at time i where $w \in R_X(v_i, Y(v_i))$.

Let $a := X(v_i)$ and $b := Y(v_i)$. Consider a vertex w in $R_X(v_i, b)$ when the colour of v_i is changed from a to b. Suppose that a bad event happens at w (i.e. $P_w(i) = P_w(i-1) + 1$). Since w is in $R_X(v_i, b)$, by definition, there are two vertices $u_a, u_b \in N_G(w) - v_i$ so that $X(u_a) = a$ and $X(u_b) = b$. Recall that $P_w(i) := \max\{P_w(X, Z_i), P_w(Z_i, W)\}$. Since both colours a and b appear in the neighbourhood of w in X, we have $P_w(Z_i, X) = P_w(Z_{i-1}, X)$. Since we assume $P_w(i) = P_w(i-1) + 1$, it must be the case that $P_w(X, Z_i) = P_w(X, Z_{i-1}) + 1$. This can only happen when the colour a disappears in the neighbourhood of w at time i (i.e. $a \notin L_{Z_{i-1}}(w)$ and $a \in L_{Z_i}(w)$). In particular, this implies that the colour of u_a had been changed from a to some other colour in some j-th step where $j < i$. Consider that colour change of $u_a = v_j$ at the j-th step. At the j-th step, $Z_j(v_i)$ is still of colour a. Therefore, by changing the colour of u_a from a to some other colour, we have $P_w(X, Z_j) = P_w(X, Z_{j-1})$. Notice that $P_w(X, Y) \leq \sum_{i=1}^{d} |P_w(X, Z_i) - P_w(X, Z_{i-1})|$ and similarly $P_w(Y, X) \leq \sum_{i=1}^{d} |P_w(Z_i, X) - P_w(Z_{i-1}, X)|$. From the above argument, if a bad event happens at w at the i-th step, we have $|P_w(X, Z_i) - P_w(X, Z_{i-1})| = 0$ and $|P_w(Z_j, X) - P_w(Z_{j-1}, X)| = 0$ for some $j < i$. And thus the bad event at w at time i and the event at w at time j combine to contribute at most 1 to $P_w(d)$, in particular this implies that $P_w(d)$ is at most $d-1$. Formally, we map a bad event at w at time i to another event at w at time j where $j < i$ and $X(v_i) = X(v_j) = a$, so that they combine to contribute at most 1 to $P_w(d)$. We call the events in this mapping a *couple*. We can do the mapping for each bad event at w at time j for each $w \in R_X(v_j, Y(v_j))$ by the above argument. If there are T_1 disjoint couples and T_2 distinct good events such that no event appears therein more than once, then $\sum_{w \in V} P_w(d) \leq d\Delta - T_1 - T_2$. Suppose for now that each bad event at w at time j for $w \in R_X(v_j, Y(v_j))$ maps to a distinct good event (we will prove this claim in the next paragraph). Since $\sum_{j=1}^{d} |R_X(v_j, Y(v_j))| \geq R_X \Delta d$ and each event therein is either good or is in a distinct couple, we have $\sum_{w \in V} P_w(d) \leq d\Delta - (R_X \Delta d)/2 = ((1 - R_X/2)\Delta)d$, as desired. (The worst case is that there are $(R_X \Delta d)/2$ disjoint couples where each couple contains two distinct events therein).

To finish the proof, it remains to show that each bad event at w at time j for $w \in R_X(v_j, Y(v_j))$ maps to a distinct good event. To see this, we need to review the mapping process. As argued previously, a bad event at w at time i for $w \in R_X(v_i, Y(v_i))$ happens only if the colour $X(v_i)$ disappears in the neighbourhood of w at time i. Then we map this bad event to another event at w at time j where $j < i$ and $X(v_i) = X(v_j)$. This implies that $Z_j(v_i) = X(v_i)$ and thus the colour $X(v_j)$ does not disappear in the neighbourhood of w at time j, and hence the event at w at time j is not a bad event. So, a bad event does not map to another bad event. Also, two bad events cannot map to the same event since a colour can disappear at most once in the neighbourhood of a vertex, as each vertex is recoloured at most once. This proves the claim and completes the proof. ∎

3 Uniform Colourings of Graphs with Girth 5

In this section, we establish that the uniform random colourings we consider satisfy the first and second uniformity properties. Recall that our setting is: G is a graph on n vertices with maximum degree Δ and with girth at least 5. We consider k-colourings of G where $k = q\Delta$ for some $q > 1$. For ease of exposition, we assume that G is Δ-regular.

Lemma 3. *Consider a uniform random k-colouring X of G and consider any $\zeta > 0$. With probability at least $1 - n^{-3}$ we have:*

(a) $(qe^{-1/q} - \zeta)\Delta < L_X < (qe^{-1/q} + \zeta)\Delta$;
(b) $((1 - e^{-1/q})^2 - \zeta)\Delta < R_X < ((1 - e^{-1/q})^2 + \zeta)\Delta$.

The lower bound in part (a) was proven in [10] for graphs of girth 4. The rest of Lemma 3 was (essentially) proven in [7] to hold for graphs of girth 6. Roughly speaking, having girth 6 is very helpful as follows: Define $N_2(v)$ to be the vertices of distance 2 from v. Note that $R_X(v, c)$ is a function of the colours appearing on $v \cup N_2(v)$ which, if G has girth at least 6, is an independent set. If we pretend that the colours on those vertices are independent uniformly random colours, then Lemma 3(b) follows easily. Of course they aren't independent: some dependency is induced by the edges joining the independent set to the rest of G; this is a bit difficult to deal with, but the techniques from [12] will suffice. When we reduce the girth requirement to 5, $N_2(v)$ can now have many edges, and these edges bring dependencies that are not straightforward to deal with. Fortunately, there are some restrictions - for example, no two vertices of $N_2(v)$ with a common "parent" in $N(v)$ can be joined. This allows us to overcome the dependency. The way we do so comprises the new ideas in the proof; the rest just follows techniques from [12].

In order to better control the edges within $v \cup N(v) \cup N_2(v)$, we partition it into smaller subgraphs. In particular, for each vertex v, we partition $N(v)$ into sets $U_1(v), ..., U_{\Delta^{2/3}}(v)$ each of size $\Delta^{1/3}$. (For convenience, we treat $\Delta^{1/3}$ as an integer; it is trivial to extend the argument to the non-integer case.) Instead of analyzing the colour distribution of colours on all of $N_2(v)$, we will sometimes consider, for each i, the distribution on the subset of $N_2(v)$ that is adjacent to $U_i(v)$.

Given a colouring of G, for any vertex v and colours c, c' we define:

- $T_{v,c} = \sum \frac{1}{|L(u)|}$ summed over all $u \in N(v)$ with c not appearing on $N(u) - v$;
- $T^j_{v,c} = \sum \frac{1}{|L(u)|}$ summed over all $u \in U_j(v)$ with c not appearing on $N(u) - v$.

We define $\alpha_0 = 0, \beta_0 = 1, \lambda_0 = 1$, and

- $\alpha_{i+1} = e^{-\beta_i}/\lambda_i$
- $\beta_{i+1} = e^{-\alpha_i}/qe^{-1/q}$
- $\lambda_{i+1} = \frac{q\beta_i - 1}{\beta_i - \alpha_i} e^{-\alpha_i} + \frac{1 - q\alpha_i}{\beta_i - \alpha_i} e^{-\beta_i}$

As shown in [12], $\lim_{i\to\infty} \alpha_i = \lim_{i\to\infty} \beta_i = 1/q$ and $\lim_{i\to\infty} \lambda_i = qe^{-1/q}$.

We will prove:

Lemma 4. *In a uniformly random k-colouring X of G, for every v, c, j, i we have: with probability at least $1 - (\Delta^3)^i \exp(-\Delta^{1/3})$:*

(a) $qe^{-1/q}\Delta - o(\Delta) < |L(v)| < \lambda_i \Delta + o(\Delta)$;
(b) $\alpha_i \Delta^{-2/3} - o(\Delta^{-2/3}) < |T^j_{v,c}| < \beta_i \Delta^{-2/3} + o(\Delta^{-2/3})$.

Proof. Many details are the same as those that have appeared already in several papers. So we gloss over those, focusing more on the details that are new.

The lower bound in (a) was proven in [10]. For the rest, we use induction on i. The base case $i = 0$ is trivial. So suppose it holds for i; we will prove that it holds for $i + 1$.

Expose the colours of every vertex except for those in $N(v)$. This yields a list $L(u)$ for each $u \in N(v)$.

By induction, and multiplying by Δ^2 vertices u in $N(v) \cup N_2(v)$ plus less than $2\Delta^{8/3}$ triples j, c, u with $u \in N(v)$, we see that with probability at least $1 - (\Delta^2 + 2\Delta^{8/3})(\Delta^3)^i \exp(-\Delta^{1/3})$, each $u \in N(v) \cup N_2(v)$ has $qe^{-1/q}\Delta - o(\Delta) < |L(u)| < \lambda_i \Delta + o(\Delta)$ and for every $u \in N(v)$ and c, $T_{u,c} = \sum_{j=1}^{\Delta^{2/3}} T^j_{u,c}$ is between $\alpha_i + o(1)$ and $\beta_i + o(1)$.

We will show that, if the exposed colours behave as described in the previous paragraph, then the probability that (a) is violated is at most $\exp(-\Delta^{1/3})$ and so is the probability that (b) is violated. This yields an overall bound of at most $(\Delta^2 + 2\Delta^{8/3})(\Delta^3)^i \exp(-\Delta^{1/3}) + 2\exp(-\Delta^{1/3}) < (\Delta^3)^{i+1} \exp(-\Delta^{1/3})$.

Part (a): Straightforward calculations, as in [12], show that $\mathbf{Exp}(|L(v)|) \leq \lambda_{i+1}\Delta + o(\Delta)$. Because G is triangle-free, we can regard the assignments of colours to $N(v)$ as independent uniform choices from the lists $L(u)$. Thus standard concentration bounds (such as Azuma's Inequality) easily yield that the probability of $|L(u)|$ differing from its mean by more than $\Delta^{2/3}$ is at most $\exp(-\theta(\Delta^{2/3})) < \exp(-\Delta^{1/3})$.

Part (b): Let $t^j_{v,c}$ denote the number of neighbours $u \in U_j(v)$ with $c \in L(u)$. We will show that, with sufficiently high probability, $\alpha_i \Delta^{1/3} + o(\Delta^{1/3}) < t^j_{v,c} < \alpha_i \Delta^{1/3} + o(\Delta^{1/3})$. This, along with the inductive bound on $|L(u)|$ will establish (b).

Let H be the subgraph induced by $\cup_{u \in U_j} N(u) - v$. Note that, since the girth of G is at least 5, no $w \in H$ can be adjacent to more than one neighbour of any $u \in U_i(v)$. Thus the maximum degree in H is at most $|U_j(v)| = \Delta^{1/3}$.

Expose the colours of every vertex except those in H. This yields a list $L(u)$ for each $u \in U_j(v)$. With probability at least $1-2ie^{-\Delta^{1/20}}$, each $w \in N(U_j(v))-v$ has $qe^{-1/q}\Delta - o(\Delta) < |L(w)| < \lambda_i \Delta + o(\Delta)$ and for every c and $u \in U_j(v)$, $T_{u,c}$ is between $\alpha_i + o(1)$ and $\beta_i + o(1)$.

Let Ω be the set of colourings of H in which each $w \in H$ has a colour from $L(w)$. Thus, the unexposed colours form a uniformly random member of Ω. Since H is not an independent set, we can't simply treat those colours as independent uniform choices from their lists, as we did in part (a). How we deal with this complication is the main new idea required for this section.

Suppose that the vertices of H are $w_1, ..., w_t$; we will colour the vertices one-at-a-time, in order. After r vertices have been coloured, we use Ω_r to denote the set of completions of the partial colouring to a colouring of H taken from the lists $L(w)$; thus $\Omega_0 = \Omega$. When we colour w_r, we choose each colour c with probability $p_r(c)$ which is equal to the proportion of colourings in Ω_{r-1} in which w_r has c. Thus, the resultant colouring is a uniformly random member of Ω, as required.

Claim. For each r, c such that c is not assigned to a neighbour $w_{r'}$ of w_r with $r' < r$, we have: $p_r(c) = |L(w_r)|^{-1}(1 \pm O(\Delta^{-2/3}))$.

Proof. First note that since the maximum degree of H is at most $\Delta^{1/3}$, at any point in the colouring process, every w has at least $|L(w)| - \Delta^{1/3} = \Theta(\Delta)$ colours not appearing on any of its neighbours.

We prove the claim by induction on r. The base case is $r = t$; i.e. the last vertex coloured in H. Here, each colour in $L(w_t)$ that does not yet appear on a neighbour is equally likely, so $|L(w_r)|^{-1} \le p_t(c) \le (|L(w_r)| - \Delta^{1/3})^{-1} < |L(w_r)|^{-1}(1 + O(\Delta^{-2/3})$. Now assume that the claim holds for every $r' > r$; we will prove it for r.

Consider two colours $c, c' \in L(w_r)$ that don't appear on any neighbour of w_r. Let $\Omega(c), \Omega(c')$ denote the sets of colourings in Ω_{r-1} in which w_r gets c, c' respectively. Suppose that we set $w_r = c'$ and then continue our colouring process. By induction, the probability that at least one neighbour of w_r receives c is at most $\Delta^{1/3} \times O(\Delta^{-1}) = O(\Delta^{-2/3})$. Since this process yields a uniform member of $\Omega(c')$, this implies that at least $(1 - O(\Delta^{-2/3}))|\Omega(c')|$ of the colourings of $\Omega(c')$ can be mapped to a colouring in $\Omega(c)$ by switching the colour of w_r to c. Therefore, $|\Omega(c)| \ge |\Omega(c')|(1 - O(\Delta^{-2/3}))$. Since this is true for every pair c, c' the claim follows. ∎

Having proven the Claim, we now consider any $u \in U_j$; we will bound the probability that c is not assigned to any $w \in N(u) - v$. Since our colouring procedure yields a uniformly random member of Ω, this probability is not affected by the actual order in which we colour the vertices. So we can take $w_1, ...,$ the first vertices to be coloured, to be $N(u) - v$. Since $N(u) - v$ is an independent set, if $c \in L(w)$ for some $w \in N(u) - v$, c will still be eligible to be assigned to w when we come to choose the colour for w. Therefore by our claim, the desired probability is $\prod(1 - |L_w|^{-1} + O(\Delta^{-5/3}))$ over all $w \in N(u) - v$ with $c \in L(w)$, and this is between $\alpha_i + o(1)$ and $\beta_i + o(1)$ by the same calculations as in [12]. Thus, $\alpha_i \Delta^{1/3} + o(\Delta^{1/3}) < \mathbf{Exp}(t^j_{v,c}) < \beta_i \Delta^{1/3} + o(\Delta^{1/3})$.

Next we will use Azuma's Inequality to show that $t^j_{v,c}$ is concentrated.

Consider a particular w_r adjacent to $u_\ell \in U_j$. We want to measure how much the colour chosen for w_r can affect the expected value of $t^j_{v,c}$, where the expectation is over the remaining $t - r$ random colour assignments. The extreme case is when we choose to assign c to w_r (we omit the straightforward dispensation of the other cases). This will cause u_ℓ to not have $c \in L(u_\ell)$ and thus might reduce the conditional expectation of $t^j_{v,c}$ by 1. Possibly it will also have a further effect on the conditional expectation because it changes the probability that other members of U_j will have c in their lists; we bound that effect as follows: For each of the $\Delta^{1/3}$ neighbours w of w_r, the assignment of c to w_r drops the probability of w receiving c to zero; for every other w, by reasoning similar to that in our claim, this affects the probability of w receiving c by a negligible amount. Each $u \in N(v)$ has at most one neighbour adjacent to w_r and so the effect on the probability of c not appearing in $N(u)$ is at most a factor of $(1-1/\Theta(\Delta))$; since this probability is $\Theta(1)$, by the previous paragraph, the effect is an additive term of at most $O(1/\Delta)$. Thus, the overall affect on $\mathbf{Exp}(t_{v,c})$ is $1 + |U_i| \times O(1/\Delta) = 1 + O(\Delta^{-2/3}) < 2$.

Note that $t_{v,c}$ is determined by $\Delta^{4/3}$ trials - the colour choice for each neighbour of every member of U_i. If we try to apply Azuma's Inequality directly with each $\gamma_i = 2$ and with $\Delta^{4/3}$ trials, we fail. So we reduce the number of trials to $\Delta^{1/3}$ as follows: For each $u \in N(v)$, we treat the assignments to all of $N(u) - v$ as a single random choice. A simple concentration argument (details omitted) implies that with probability at least $1 - \exp(-\Theta(\Delta))$, no u has more than $\sqrt{\Delta}$ neighbours that receive c. Standard arguments (details omitted) allow us to assume that no u has more than $\Delta^{1/10}$ such neighbours, as far as the remainder of the argument is concerned. Thus, by the same calculations as in the previous paragraph, the maximum effect that any one of these random choices can have is $1+\sqrt{\Delta}\times O(\Delta^{-2/3}) < 2$. Thus we can apply Azuma's Inequality with $\Delta^{1/3}$ trials, each $\gamma_i = 2$ and with $\tau = O(\Delta^{9/30})$ to show that the probability of $t_{v,c}$ differing from its mean by more than $O(\Delta^{9/30}) = o(\Delta^{1/3})$ is at most $\exp(-\Theta(\Delta^{4/15}))$. This yields an overall probability of $t^j_{v,c}$ differing from its mean by more than $O(\Delta^{9/30})$ of less than $\exp(-\Theta(\Delta)) + \exp(-\Theta(\Delta^{-4/15})) < \exp(-\Delta^{1/3})$. ∎

Now we finish the proof of Lemma 3.

*Proof (***Proof of Lemma 3).** We will show that for every v, c, the probability that $L_X(v)$ violates part (a) or that $R_X(v, c)$ violates part (b) is at most $\exp(-\frac{1}{2}\Delta^{1/3})$. If $\Delta \geq 1000 \log^3 n$ then this is at most $1/n^5$. Thus, after multiplying by the fewer than n^2 choices for v, c, we obtain that conditions (a,b) hold for every n, c with probability at least $1 - n^{-3}$, as required.

The bound on the probability that $L_X(v)$ is in violation follows immediately from Lemma 4 by taking i to be a large enough constant that λ_i differs from its limit by at most $\zeta/2$ and noting that $(\Delta^3)^i \exp(-\Delta^{1/3}) < \exp(-\frac{1}{2}\Delta^{1/3})$.

For $R_X(v, c)$, we also apply Lemma 4 for a particular large value of i. Then we carry out an argument nearly identical to that in the proof of Lemma 4(b) to bound, for every v, c the number of neighbours $u \in N(v)$ with $X(v), c$ both in $N(u)-v$. Straightforward calculations, as in [12], show that the expected number is within $\zeta\Delta/2$ of $(1-e^{-1/q})^2\Delta$ so long as i is large enough that $\alpha_i, \beta_i, \lambda_i$ are sufficiently close to their limits. A concentration proof nearly identical to that in the proof of Lemma 4 shows that the probability that this number differs from its mean by at least $\zeta\Delta/2$ is less than $\exp(-\frac{1}{2}\Delta^{1/3})$; we omit the repetitive details. ∎

References

1. K. Azuma. *Weighted sums of certain dependent random variables.* Tokuku. Math. Journal, 19, pp. 357-367, 1967.
2. R. Bubley and M. Dyer. *Path coupling: a technique for proving rapid mixing in Markov chains.* Proceedings of FOCS 1997, pp. 223-231.
3. M. Dyer, A. Flaxman, A. Frieze and E. Vigoda. *Randomly coloring sparse random graphs with fewer colors than the maximum degree.* Manuscript, 2004.
4. M. Dyer and A. Frieze. *Randomly colouring graphs with lower bounds on girth and maximum degree.* Random Structures and Algorithms, 23, pp. 167-179, 2003.
5. M. Dyer, A. Frieze, T. Hayes, E. Vigoda. *Randomly coloring constant degree graphs.* Proceedings of FOCS 2004, pp. 582-589.
6. A. Frieze and J. Vera. *On randomly colouring locally sparse graphs.* Manuscript, 2004.
7. T. Hayes. *Randomly coloring graphs of girth at least five.* Proceedings of STOC 2003, pp. 269-278.
8. T. Hayes and E. Vigoda. *A non-Markovian coupling for randomly sampling colorings.* Proceedings of FOCS 2003, pp. 618-627.
9. T. Hayes and E. Vigoda. *Variable length path coupling.* Proceedings of SODA 2004, pp. 103-110.
10. T. Hayes and E. Vigoda. *Coupling with the stationary distribution and improved sampling for colorings and independent sets.* Proceedings of SODA 2005, pp. 971-979.
11. M. Jerrum. *A very simple algorithm for estimating the number of k-colourings of a low-degree graph.* Random Structures and Algorithms, 7, pp. 157-165, 1995.
12. M. Molloy. *The Glauber dynamics on colorings of a graph with high girth and maximum degree.* SIAM Journal on Computing, 33, pp. 721-737, 2004.
13. M. Molloy and B. Reed. *Graph colourings and the probabilistic method.* Springer, 2002.
14. J. Salas and A. Sokal. *Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem.* J. Statist. Phys., 86, pp. 551-579, 1997.
15. E. Vigoda. *Improved bounds for sampling colorings.* Journal of Mathematical Physics, 41, pp. 1555-1569, 2000.