Randomly coloring sparse random graphs with fewer colors than the maximum degree

Martin Dyer [*] Abie Flaxman [†] Alan Frieze [‡] Eric Vigoda [§]

July 4, 2004

Abstract

We analyze Markov chains for generating a random k-coloring of a random graph $G_{n,d/n}$. When the average degree d is constant, a random graph has maximum degree $\log n/\log\log n$, with high probability. We efficiently generate a random k-coloring when $k = \Omega(\log\log n/\log\log\log n)$, i.e., with many fewer colors than the maximum degree. Previous results hold for a more general class of graphs, but always require more colors than the maximum degree.For $k \geq (\ln n)^{\alpha}$ we can also show that Glauber Dynamics mixes in polynomial time.

1 Introduction

We study Markov Chain Monte Carlo algorithms for generating a random (vertex) k-coloring of an input graph $G = (V, E)$. When $k > \Delta$ where Δ is the maximum degree of G, it is trivial to construct a proper k-coloring. Therefore, we can hope to generate a k-coloring (almost) uniformly at random from the collection of proper k-colorings.

Simple Markov chains, such as the *Glauber dynamics*, appear extremely effective. In the Glauber dynamics, at each step we randomly recolor a random vertex. More precisely, from a k-coloring X_t at time t, the transition $X_t \rightarrow X_{t+1}$ is defined as follows. First, a random vertex v_t is chosen. We then set $X_{t+1}(v_t)$ to a color chosen uniformly at random from those colors not appearing in the neighborhood of v_t in X_t. For all $w \neq v_t$, we set $X_{t+1}(w) = X_t(w)$. The stationary distribution of the Glauber dynamics (and the other Markov chains we will study in this paper) is uniformly distributed over k-colorings. We are interested in the *mixing time* of such Markov chains, which is the number of steps until the chain is within variation distance 1/4 of the stationary distribution, regardless of our initial k-coloring X_0 (see Jerrum [13] for relevant background on finite Markov chains).

Jerrum [13] proved that whenever $k > 2\Delta$ the mixing time of the Glauber dynamics is $O(n \log n)$. Vigoda [19] improved Jerrum's result, by analyzing an alternative chain, reducing the lower bound

[*]School of Computing, University of Leeds, Leeds LS2 9JT, UK. Supported by EPSRC grant GR/S76151/01.

[†]Department of Mathematics, Carnegie Mellon University, Pittsburgh PA 15213, USA.

[‡]Department of Mathematics, Carnegie Mellon University, Pittsburgh PA15213, USA. Supported by NSF grant CCR-0200945.

[§]Department of Computer Science,University of Chicago, Chicago, IL 60637,USA. Supported by NSF grant CCR-0237834.

on k to $11\Delta/6$. This is still the best lower bound on k for general graphs.

Subsequent work, beginning with Dyer and Frieze [6], used the so-called burn-in method, and analyzed the Glauber dynamics on restricted classes of graphs. Building upon [6, 16, 10], Hayes and Vigoda [11] proved the Glauber dynamics has $O(n \log n)$ mixing time when $k > (1+\epsilon)\Delta$ for all $\epsilon > 0$, assuming G has girth > 9 and $\Delta = \Omega(\log n)$. Dyer, Frieze, Hayes and Vigoda [7] reduced the lower bound on Δ to a sufficiently large constant, assuming $k > 1.489\Delta$ and girth $g > 5$. Further improvements were recently proved for amenable graphs (without any lower bound on Δ) by Goldberg, Martin and Paterson [9], and for "locally sparse" graphs (assuming $\Delta = \Omega(\log n)$) by Frieze and Vera [8, 12].

In many classes of graphs, e.g., random graphs and planar graphs, the chromatic number is intimately related to the average degree, as opposed to the maximum degree. This paper focuses on randomly coloring sparse random graphs. These graphs have constant average degree d and much larger maximum degree Δ. We randomly color such graphs with many fewer than Δ colors.

We work with $G = G_{n,p}$ where $p = d/n$, $d > 1$ is the random graph with vertex set $[n] = \{1, 2, \ldots, n\}$ and where each possible edge is independently included with probability p. Such graphs have relatively few vertices of large degree. Thus, it would seem that we might be able to randomly colour such a graph with many fewer than Δ colours. We will prove this in our main theorem below.

The main difficulty in dealing with high degree vertices in the analysis is that in many colorings, these vertices have few color choices, i.e., almost all of the colors might appear in their neighborhood. Thus, the color choice of the neighbor of a high degree vertex v can have a large influence on the color choice of v when v's color is updated. To avoid this, we cluster the high degree vertices into sets of nearby vertices. We then pad these sets with a radius r of low degree vertices. The radius r is chosen sufficiently large so that these padded sets are not overly influenced by the color choices of their neighbors. We analyze a Markov chain tailored to our clustering of high degree vertices.

We need some notation before formally defining the Markov chain we analyze. For $b \geq 1$, let $L_b = \{v : \deg(v) \geq b\}$ denote those vertices with minimum degree b. For $r \geq 1$, let N_b denote those vertices within distance $\leq r$ of some vertex in L_b. Finally, let H_b be the subgraph of G induced by $V_b = L_b \cup N_b$.

In addition to the Glauber dynamics defined earlier, we consider the following Markov chain, which we refer to as the *modified Glauber dynamics*. Let $\lambda(v) = 1$ if $v \notin V_b$ and $\lambda(v) = 1/|C|$ if $v \in V_b$ and C is the component of H_b containing v. Let $\Lambda = \sum_{v \in V} \lambda(v)$. The transitions of the modified Glauber dynamics are defined as follows. From a coloring X_t, we choose $v_t \in V$ with probability $\lambda(v_t)/\Lambda$. If $v_t \in C$, then we randomly re-color the component C, otherwise we randomly re-color v.

We can now state our main theorems.

Theorem 1. *For all $d \geq 1$, with probability $1 - o(1)$ the random graph $g = G_{n,d/n}$ is such that*

(a) *if*

$$r = \ln \ln n, \quad b = \frac{(2 + \ln d) r}{\ln r} \quad \text{and} \quad k \geq 12b \tag{1}$$

then

- *Modified Glauber dynamics has mixing time $O(n \log n)$.*
- *A step of the modified Glauber dynamics can be implemented in time polynomial in $\log n$.*

(b) *if* $0 < \alpha < 1$ *and*

$$r = 2(1 + 1/\alpha) \text{ and } b = (\ln n)^\alpha \text{ and } k \geq 12b \tag{2}$$

then

- *Modified Glauber dynamics has mixing time* $O(n \log n)$*.*
- *A step of the modified Glauber dynamics can be implemented in time polynomial in* $\log n$*.*
- *Glauber dynamics has polynomial mixing time.*

Note that if $d < 1$ then **whp**[1] G consists of trees and unicyclic components of size $O(\log n)$ and then it is trivial to randomly color G. Also, we can allow d to grow with n, but once $d = \Omega(\log n)$ the result is subsumed by Jerrum's result.

It is well known that the maximum degree of G is $\sim \frac{\ln n}{\ln \ln n}$ **whp**. Thus the number of colors required for rapid mixing is $o(\log \Delta)$ in our first case, and $o(\Delta)$ in the second.

2 Ergodicity

We first show that Glauber dynamics (and hence modified Glauber dynamics) is ergodic for a random graph $G_{n,d/n}$ when $k \geq d + 2$.

For a graph $G = (V, E)$, the κ-core is the unique maximal set $S \subseteq V$ such that the induced subgraph on S has minimum degree at least κ. It follows from work of Pittel, Spencer and Wormald [17] that **whp** G has no κ-core for $\kappa \geq d$. A graph without a κ-core is k*-degenerate* i.e. its vertices can be ordered as $v_1, v_2, \ldots, v_n$ so that v_i has fewer than κ neighbors in $\{v_1, v_2, \ldots, v_{i-1}\}$. To see this, let v_n be a vertex of minimum degree and then apply induction.

Lemma 2. *If* $G = (V, E)$ *has no* κ*-core, then, for all* $k \geq \kappa + 2$*, the Glauber dynamics for* k*-colorings is ergodic.*

Proof. Let $v_1, \ldots, v_n$ denote an ordering of V such that v_i has degree $< \kappa$ in G_i, defined as the induced subgraph on $\{v_1, v_2, \ldots, v_i\}$. For $1 \leq i \leq n$, let Ω_i denote the k-colorings of G_i.

We need to show that the set Ω_n is connected with respect to transitions of the Glauber dynamics. We will prove the claim by induction. The claim is trivial for $n = 1$. Assume the set Ω_j, for all $j < i$, is connected. Consider a pair of colorings $X, Y \in \Omega_i$. Let X', and Y' respectively, denote the projection of these colorings on G_{i-1}.

We inductively know there exists a path of Glauber transitions (for G_{i-1}) connecting X' to Y'. Consider any such path, say it has length ℓ. Let (w_j, c_j) denote the $(vertex, color)$ update at step j of this path. We construct a path (of length $\leq 2\ell$) from X to Y along Glauber transitions for G_i.

For $j = 1, 2, \ldots, \ell$, we will re-color w_j to color c_j, if such a transition is valid (i.e., no neighbor of w_j has color c_j). If it is not valid, then v_i must be the only neighbor of w_j that is colored c_j. Since v_i has degree $< \kappa$ in G_i, there exists a new color for v_i which does not appear in its neighborhood. Thus, we first re-color v_i to any new (valid) color, and then we re-color w_j to c_j. Hence, the length of the path at most doubles. ∎

[1] Throughout this paper, we use the term with high probability, denoted **whp**, to refer to events which occur with probability $1 - o(1)$.

3 Structure results

In this section we will define some useful graph properties and show that G has these properties **whp**.

3.1 Case (a)

The graph properties of interest are the following:

P1 The maximum component size in H_b is at most $C_{\max} = (\ln n)^3 (2d)^r = (\ln n)^{O(1)}$.

P2 If $v \notin H_b$ and C is a component of H_b then v has at most 2 neighbors in C.

P3 Each component C of H_b has at most $|C|$ edges.

Theorem 3. *Under the hypotheses of Theorem 1(a), with probability* $1 - o(1)$ *properties* **P1-P3** *hold.*

3.2 Case (b)

We modify our claims about the structure of G under the hypotheses of Theorem 1(b):

P1 The maximum component size in H_b is at most $C_{\max} = (10d)^r \ln n$.

P2 If $v \notin H_b$ and C is a component of H_b then v has at most 2 neighbors in C.

P3 Each component C of H_b has at most $|C|$ edges.

Theorem 4. *Under the hypotheses of Theorem 1(b), with probability* $1 - o(1)$ *properties* **P1-P3** *hold.*

We prove Theorem 3 in Section 6.1 and Theorem 4 in Section 6.3.

3.3 Implementing modified Glauber dynamics

Implementing a transition of the modified Glauber dynamics is equivalent to generating a random list coloring of the updated component C. In the list coloring problem every vertex $v \in C$ has a set $L(v)$ of valid colors, where $|L(v)| \subseteq \{1, 2, \ldots, k\}$, and v can only receive a color in $L(v)$. In our case, $L(v)$ are those colors not appearing in $N(v) \cap \bar{S}$.

For a tree on ℓ vertices, using dynamic programming we can exactly compute the number of list colorings in time ℓk. Therefore, we can also generate a random list coloring of a tree. By property P3, our components are trees or unicyclic. For a unicyclic component, we can simply consider all $\leq k^2$ colorings for the endpoints of the extra edge, and then recurse on the remaining tree. By property P1, this implies that the modified Glauber dynamics can be efficiently implemented.

4 Coupling Analysis: Proof of Theorem 1(a)

In this section, we prove Theorem 1(a), using the structure results from Theorem 3.

We use path coupling [4]. For a pair of colorings X, Y, our metric $d(X,Y)$ is Hamming distance:

$$d(X,Y) = \sum_{v \in V} 1_{X(v) \neq Y(v)},$$

We are therefore obliged to extend the state space to include improper colorings as transient states. Hence, for all (X_t, Y_t) where $d(X_t, Y_t) = 1$, we need to define a coupling $(X_t, Y_t) \to (X_{t+1}, Y_{t+1})$ such that

$$\mathbf{E}(d(X_{t+1}, Y_{t+1}) \mid X_t, Y_t) < (1 - 1/2n) d(X_t, Y_t).$$

This implies mixing time $O(n \log n)$ by a standard application of the path coupling method.

Each chain chooses the same random vertex w, and both chains re-color w, if $w \notin V_b$, or re-color the component $C_w \ni w$, if $w \in V_b$. The choices will be coupled as described below. We divide the coupling analysis into two cases, depending on whether X_t and Y_t differ at a (unique) vertex $v \in V_b$, or at a vertex $v \notin V_b$. Recall that $\Lambda = n - o(n)$, from **Q1**.

Case 1: $X_t(v) \neq Y_t(v)$ and $v \in C_v \in \mathcal{C}_b$.

If we re-color component C_v, then both chains can choose the same coloring and $X_{t+1} = Y_{t+1}$. Consider $w \in N(v)$. If $w \in V_b$, then $w \in C_v$ and $X_{t+1}(w) = Y_{t+1}(w)$. If $w \notin V_b$ then $\deg(w) < b$, and there are at least $k - b$ colors not appearing in $X_t(N(w))$, and similarly for $Y_t(N(w))$. Using the maximal coupling, there is at most one choice for $X_{t+1}(w)$ which results in $X_{t+1}(w) \neq Y_{t+1}(w)$, i.e. $X_{t+1}(w) = Y_t(v)$. It follows that

$$\mathbf{Pr}\left(X_{t+1}(w) \neq Y_{t+1}(w) \mid \xi_t = w\right) \leq \frac{1}{k-b}$$

where ξ_t is the random vertex chosen at step t.

We can now bound the expected change in distance after a coupled transition,

$$\begin{aligned}
\mathbf{E}(d(X_{t+1}, Y_{t+1}) - d(X_t, Y_t)) &= -\mathbf{Pr}\left(\xi_t \in C_v\right) + \sum_{w \in N(v) \setminus C_v} \mathbf{Pr}\left(\xi_t = w \wedge X_{t+1}(w) \neq Y_{t+1}(w)\right) \\
&\leq -\frac{1}{\Lambda} + \frac{b}{(k-b)\Lambda} \\
&\leq -\frac{1}{2n} \quad \text{for } k \geq 4b \text{ and } n \text{ large.}
\end{aligned}$$

Case 2: $X_t(v) \neq Y_t(v)$ and $v \notin V_b$.

For $w \in N(v) \setminus V_b$, using the maximal coupling, the probability w receives a different color in the two chains is bounded by $1/(k-b)\Lambda$ as above. Otherwise we will couple the colorings of C_w in X and Y, as described below, so as to have few disagreements. Let $\Phi(w)$ be the expected number of disagreements between X_{t+1} and Y_{t+1} in C_w, i.e.

$$\mathbf{E}\big(d(X_{t+1}(C_w), Y_{t+1}(C_w)) \mid \xi_t = w\big) = \Phi(w),$$

and $\Phi = \max_w \Phi(w)$. Then, bounding the expected change in distance,

$$\begin{aligned}
\mathbf{E}(d(X_{t+1}, Y_{t+1}) - d(X_t, Y_t)) &< -\mathbf{Pr}(\xi_t = v) + \frac{|N(v) \setminus V_b|}{(k-b)\Lambda} + \sum_{w \in N(v) \cap V_b} \mathbf{Pr}(\xi_t \in C_w)\Phi(w) \\
&\le -\frac{1}{\Lambda} + \frac{b}{(k-b)\Lambda} + \frac{1}{\Lambda} \sum_{w \in N(v) \cap V_b} \Phi(w) \\
&\le -\frac{1}{\Lambda} + \frac{b}{(k-b)\Lambda} + \frac{b\Phi}{\Lambda} \\
&\le -\frac{1}{2n} \quad \text{for } k \ge 6b, \;\; 4b\Phi \le 1, \text{ and } n \text{ large.}
\end{aligned}$$

It remains to show that $4b\Phi \le 1$. We use the "disagreement percolation" coupling construction of van den Berg and Maes [1, Theorem 1]. We wish to couple $X(C_w)$ and $Y(C_w)$ as closely as possible, but the identity coupling is precluded by the disagreement at v. The technique of [1] assembles the coupling in a stepwise fashion working away from w. In our case, it may be viewed as follows. From **P3** we know that C_w is a tree with at most one additional edge. Also, from the definition of H_b, it has degree at most b except for a central "kernel" of higher-degree vertices at distance r from its boundary. The disagreement at v propagates into C_w along paths from w. A disagreement at vertex $x \in C_w$ at (edge) distance ℓ from w propagates to a neighbor z at distance $\ell + 1$ if $X(z) \neq Y(z)$. The distributions of $X(z)$, $Y(z)$ are invariant under Glauber heat-bath dynamics. Thus, if z is not in the kernel, we may couple $X(z), Y(z)$, using the maximal coupling, to have $\mathbf{Pr}(X(z) \neq Y(z)) \le 2/(k-b) = \zeta$ (say), since

(i) z can have at most two neighbors which disagree, since C_w is almost a tree,

(ii) each such neighbor of z will have at most one disagreement,

(iii) there are at least $k - b$ colors available at z.

The disagreement percolation is dominated by an independent process. Thus a disagreement propagates to a vertex at distance $\ell < r$ from w with probability at most $\zeta^{\ell+1}$. Moreover there are at most b^ℓ such vertices. It propagates to a vertex in the kernel with probability at most $\zeta^{r+1}(\ln n)^2$ for large n, using **Q3**. If this happens, we couple arbitrarily with the remaining probability, and concede $|C_w|$ disagreements. If $k \ge 12b$, it follows that

$$\begin{aligned}
\Phi(w) &\le \sum_{\ell=0}^{r-1} b^\ell \zeta^{\ell+1} + |C_w|\zeta^{r+1}(\ln n)^2 \\
&\le \zeta \sum_{\ell=0}^{\infty} \left(\frac{2b}{k-b}\right)^\ell + o(\zeta), \\
&\le \frac{1}{5b}\big(1 + o(1)\big), \\
&\le \frac{1}{4b} \quad \text{for } n \text{ large,}
\end{aligned}$$

using $|C_w| = (\log n)^{O(1)}$ from **P1**, and $r = \Omega(\log \log n)$ so $\zeta^{-r} = (\log n)^{\Omega(\log\log\log n)}$. □

5 Proof of Theorem 1: Part (b)

We first analyze the mixing time of the modified Glauber dynamics in Section 5.1. Then, in Section 5.2, we use the comparison method of Diaconis and Saloff-Coste [5] to bound the mixing time of the Glauber dynamics.

5.1 Coupling Analysis

In this section we bound the mixing time of the modified Glauber dynamics as claimed in Theorem 1(b).

We follow the argument of Section 4. The only place we might run into trouble is showing that $|C_w|\zeta^{r+1}\ln n = o(\zeta)$, noting that the kernel now has size $O(\log n)$. The remaining parts of the argument are unchanged. But, since $\zeta = \Omega\big((\log n)^{-\alpha}\big)$, for large n we have

$$|C_w|\zeta^{r+1}\ln n \leq A_5(\ln n)^2 \left(\frac{20d}{(A_4-1)(\ln n)^\alpha}\right)^{2(1+1/\alpha)} = o(\zeta).$$

5.2 Comparison: Part(b)

We now bound the mixing time of the Glauber dynamics. Let τ_G denote the mixing time of the Glauber dynamics, and τ_M denote the mixing time of the modified Glauber dynamics. Let P_G and P_M denote their corresponding transition matrices. Let Ω denote the k-colorings of the graph of interest.

Lemma 5. *Under the hypotheses of Theorem 1(b),*

$$\tau_G \leq d^{O(\log n)}\tau_M \log|\Omega|,$$

for the dynamics on $G_{n,d/n}$ **whp***.*

We will use the comparison technique of Diaconis and Saloff-Coste [5]. For all $I, F \in \Omega$ where $P_M(I,F) > 0$, we will define a path $\gamma_{IF} = (Z_0 = I, Z_1, \ldots, Z_\ell = F)$ such that $P_G(Z_i, Z_{i+1}) > 0$, for all $1 \leq i < \ell$. For $t = (Z, Z') \in \Omega^2$ where $P_G(Z, Z') > 0$, let

$$cp(t) = \{(I,F) \in \Omega^2 : t \in \gamma_{IF}\},$$

denote the set of canonical paths which contain t. We are interested in its congestion:

$$\begin{aligned}
\rho(t) &= \frac{1}{\pi(Z)P_G(Z,Z')} \sum_{(I,F)\in cp(t)} |\gamma_{IF}|\pi(I)P_M(I,F) \\
&\leq nk \sum_{(I,F)\in cp(t)} |\gamma_{IF}|P_M(I,F) && (3) \\
&\leq nk|cp(t)|\gamma_{\max}, && (4)
\end{aligned}$$

where $\gamma_{\max} = \max_{(I,F)\in\Omega^2} |\gamma_{IF}|$. Let

$$\rho = \max_t \rho(t).$$

Then, e.g., see [18],

$$\tau_G \leq \rho\tau_M \log|\Omega|.$$

Now we're ready to prove Lemma 5.

Proof. Consider a Glauber transition t which recolors a vertex v. We only need to consider the case $v \in H_b$. Say v is in a component S of H_b. Fix an arbitrary coloring σ of $\bar{S} = V \setminus S$. Let $\Omega(S)$ denote the set of colorings of S consistent with σ.

We'll begin with an easy bound on $\rho(t)$, which suffices when $k = O(1)$. Clearly,

$$cp(t) \subseteq \Omega(S)^2.$$

Since $|S| = O(\log n)$, we trivially have

$$|cp(t)| \leq |\Omega|^2 \leq k^{2|S|} = k^{O(\log n)}.$$

Using the canonical paths implied by the ergodicity proof implies $\gamma_{\max} \leq 2^{|S|}$. Hence, from (4), we have

$$\rho \leq \exp(O(\log n \log k)).$$

And, for constant number of colors k, we have a polynomial bound on the mixing time of the Glauber dynamics.

We'd like to get a polynomial bound when np is constant, and $k = \Omega(1)$. So we'll fine-tune the above argument, and use (3).

Recall, with probability $\geq 1 - o(1)$, our input graph has no d-core. Fix an ordering $v_1, \dots, v_\ell, \ell = |S|$ such that v_i has degree $< d$ in the induced subgraph on $S_i = \{1, \dots, i\}$. Let G_i denote the induced subgraph on $S_i \cup \bar{S}$. Note, vertex v_i has degree $< \delta := d + bd$ in G_i. Hence, in any coloring of G_i, vertex v_i has at least two valid color choices. Let Ω_j denote the colorings of S_j in G_j ($\bar{S}$ has the fixed coloring σ).

Consider a pair of colorings $I, F \in \Omega_i$. We'll inductively define the canonical path $\gamma_i(I, F)$ along Glauber transitions for G_i. Let I', F' denote the projections of I, F onto G_{i-1}. We inductively have a path $\gamma_{i-1}(I', F')$ connecting I', F'. Let (v_j, c_j) denote the j-th transition on $\gamma_{i-1}(I', F')$. We will attempt the same transitions, in order, with a possible recoloring of v_i before and/or after each transition of $\gamma_{i-1}(I', F')$, in order to: (i) free up v_i's color for a neighbor of v_i (as in the ergodicity proof), and (ii) keep v_i colored with $I(v_i)$ unless a neighbor of v_i has color $I(v_i)$.

More precisely, consider the j-th transition, updating v_j to color c_j, and let Z denote the current coloring. Before the update (v_j, c_j), if $v_i \in N(v_j)$ and $c_j = Z(v_i)$, then choose an arbitrary new valid color for v_i. This ensures the recoloring of v_j to c_j is valid. After the update (v_j, c_j), if $v_j \in N(v_i)$ and $I(v_i) \notin Z(N(v_i))$, we recolor v_i to $I(v_i)$. This, of course, may be redundant if v_i already has color $I(v_i)$. We are trying to "remember" the initial coloring. Finally, after all of the transitions of the path $\gamma_{i-1}(I', F')$, we recolor v_ℓ to $F(v_\ell)$.

Note, the length of these paths are at most $3^{|S|}$. We bound the congestion with a similar inductive construction. For a Glauber transition t_i in G_i, let $cp_i(t_i)$ denote the set of canonical paths crossing t_i. We inductively assume, for all $j < i$, all t_j,

$$|cp_j(t_j)| \leq |\Omega_j|(1 + \delta)^{2j}. \tag{5}$$

Moreover, consider an injective map, or "encoding",

$$\eta_{t_j} : cp_j(t_j) \rightarrow \Omega_j \times \{0, \dots, \delta\}^j \times \{0, \dots, \delta\}^j.$$

Consider a transition $t_i = Z \rightarrow Z'$ in G_i. Suppose t_i recolors a vertex $v_j \neq v_i$. Then, let t_{i-1} denote the corresponding transition in G_{i-1}.

For $(I, F) \in cp_i(t_i)$, we define $\eta_{t_i}(I, F)$ by a simple modification of $\eta_{t_{i-1}}(I', F')$. Let

$$\eta_{t_{i-1}}(I', F') = (C', \{\alpha_1, \ldots, \alpha_{i-1}\}, \{\beta_1, \ldots, \beta_{i-1}\}),$$

where $C' \in \Omega_{i-1}$, and for all $1 \leq j < i$, $\alpha_j, \beta_j \in \{0, \ldots, \delta\}$. Now we'll define a coloring $C \in \Omega_i$ and α_i, β_i, which will define η_{t_i}. Let $w_1, \ldots, w_{d'}, d' \leq \delta$, denote the neighbors of v_i.

The coloring C is the same as C' for all $v_j \neq v_i$. If no neighbor of v_i has color $F(v_i)$, we set $C(v_i) = F(v_i)$ and set $\alpha_i = 0$. Otherwise, we color v_i to an arbitrary valid color, and set

$$\alpha_i = \min\{1 \leq j \leq d' : C'(w_j) = F(v_i)\},$$

to "remember" the color $F(v_i)$.

Similarly, we set $\beta_i = 0$ if $Z(v_i) = I(v_i)$. (Recall, the transition is $t_i = Z \to Z'$.) Otherwise, set

$$\beta_i = \min\{1 \leq j \leq d' : Z(w_j) = I(v_i)\},$$

to "remember" the color $I(v_i)$. Note, we defined our canonical paths so that, for all colorings W on the path, $W(v_i) = I(v_i)$ or a neighbor of v_i has color $I(v_i)$ in W.

From the encoding and the transition t_i we can uniquely recover $(I, F) \in cp_i(t_i)$. Hence, our new mapping is again injective. For a transition t_i which recolors v_i, define the encoding identically to the adjacent transition which recolors some $v_j \in N(v_i)$.

We can now bound the congestion via (3). Note, for all $Z, Z' \in \Omega$, $P_M(Z, Z') = 1/|\Omega|$. Hence, applying (5) with (3), we have

$$\rho \leq nk3^\ell(1 + \delta)^{2\ell}$$

This completes the proof of the lemma. □

6 Proof of Structure Results

6.1 Proof of Theorem 3

In this section we assume r, b, k are defined as in Theorem 1(a). To show that properties **P1-P3** hold **whp**, it is convenient to define the following additional properties and prove that they also hold **whp**.

Q1 $|H_b| = o(n)$, and thus $\Lambda = n - o(n)$.

Q2 For all $k \leq n/(2e^3d^2)$, there is no subgraph of G with k vertices which contains more than $2k$ edges.

Q3 There does not exist $S \subseteq L_b$ such that $|S| \geq k = 2r^{-1} \ln n$ and S induces a connected subgraph in G^r.

Q4 For $v \in V$ let $B(v, r)$ denote the set of vertices at distance at most r from vertex v. Then $|B(v, r)| \leq A_3(2d)^r \ln n$ for all $v \in V$.

Lemma 6. *Properties* **Q1-Q4** *hold* **whp**.

First we will show how Lemma 6 implies Theorem 3. Then we perform the calculations necessary to verify Lemma 6.

Proof of Theorem 3

P1: *The maximum component size in* H_b *is at most* $(\ln n)^2 b^r$ **whp**.

Let C be a component of H_b. Let $K = C \cap L_b$. Then from **Q4** we have $|C| \leq |K| \Delta A_3 (2d)^r \ln n$. But $\Delta = o(\ln n)$ **whp** and by Lemma 6, **Q3** holds **whp** which implies that $|K| \leq 2r^{-1} \ln n$ and so **P1** also holds **whp**. □

P2,P3: *If* $v \notin H_b$ *then* **whp** v *has at most 2 neighbors in the same component of* H_b.

Let $N_{i,b}$ denote the set of vertices within distance i of L_b. Thus, $N_b = \bigcup_{i=1}^r N_{i,b}$. To prove these properties, we fix a "typical" degree sequence $\mathbf{d} = d_1, d_2, \ldots, d_n$ for $G_{n,p}$ and generate a random graph with this degree sequence using the configuration model as described in Bollobás [2]. Let $m = (d_1 + \cdots + d_n)/2$. We construct a random pairing F of the points in $W = \bigcup_{i=1}^n W_i$, $|W_i| = d_i$ and interpret them as edges of a (multi-)graph on $[n]$. A typical degree sequence is such that the probability it is simple is bounded away from zero by a function of d only. We first expose all the pairs $\{x_1, x_2\}$ in F such that $\{x_1, x_2\} \cap \left(\bigcup_{i \in L_b} W_i\right) \neq \emptyset$. This will define $N_{1,b}$. Then we expose all pairs $\{x_1, x_2\} \in F$ such that $\{x_1, x_2\} \cap \left(\bigcup_{i \in L_b} N_{1,b}\right) \neq \emptyset$. This will define $N_{2,b}$. Continuing in this way define $N_{i,b}, i = 1, 2, \ldots, r$ and then the components of H_b will be determined. Conditioning on **P1,Q1** we see that the probability any component C gets $|C| + 1$ edges is at most

$$n((\Delta C_{\max})^2 (2m - o(n))^{-1})^2 = o(1).$$

We continue by exposing all remaining pairs $\{x_1, x_2\}$ for which both points x_j lie in $\bigcup_{i \in N_b} W_i$. The rest of F will be a random pairing of the points of W which are (i) not incident with $\bigcup_{i \in L_b} W_i$ and (ii) meet $X = \bigcup_{i \in N_b} W_i$ in at most one point. We may generate this by randomly pairing the unpaired points in X and then randomly pairing up the remaining points. We consider one component C of H_b and estimate the probability that 3 vertices have a common neighbor outside H_b. Now, since all the vertices with edges still unassigned are not in L_b, each has at most b edges left to assign. So, if **P1** and **Q1** hold and $m = |W| \geq dn/3$, then probability there exists a vertex $v \notin H_b$ with 3 neighbors in a component C of H_b is at most the sum over v, C of the expected number of triples of vertices in $C \cap N_b$ which are adjacent to v, which is at most

$$n^2 \binom{|C \cap N_b|}{3} \left(\frac{b^2}{2m - o(n)}\right)^3 = O(n^{-1+o(1)})$$

Let $\mathbf{R1} = (\neg\mathbf{P2} \cup \neg\mathbf{P3}) \cap \mathbf{P1} \cap \mathbf{Q1}$. Then, we have $\mathbf{Pr}[\mathbf{R1} \mid \mathbf{d}] = o(1)$.

We obtain the result unconditionally by this summing over values of $\mathbf{d}$ for which $\frac{1}{2}\sum_{v=1}^n \mathbf{d}_v \geq dn/3$ and then

$$\begin{aligned}\mathbf{Pr}[\neg\mathbf{P2} \cup \neg\mathbf{P3}] &\leq \mathbf{Pr}[\neg\mathbf{P1}] + \mathbf{Pr}[\neg\mathbf{Q1}] + \mathbf{Pr}[m \leq dn/3] + \sum_{\mathbf{d}: m \geq dn/3} \mathbf{Pr}[\mathbf{R1} \mid \mathbf{d}]\mathbf{Pr}[\mathbf{d}] \\ &= o(1).\end{aligned}$$

□

6.2 Proof of Lemma 6

Q1: $|H_b| = o(n)$.

$$\mathbf{E}(|H_b|) \leq n \sum_{i=0}^{r} n^i p^i \mathbf{Pr}(\mathrm{Bin}(n,p) \geq b-1). \tag{6}$$

We verify **Q1** by showing that the RHS of (6) is $o(n)$ and using the Markov inequality. (The RHS of (6) bounds the expected number of vertices within distance r of a vertex in L_b). But, $\mathbf{Pr}(\mathrm{Bin}(n,p) \geq b-1) \leq (de/(b-1))^{b-1}$ and we are done here.

Q2: *For all $k \leq n/(2e^3d^2)$, there is no subgraph of G with k vertices which contains more than $2k$ edges.*

Let $\mu_d = n/(2e^3d^2)$ and define $\mathcal{B}$ to the event that there exists a set S with $|S| = k \leq \mu_d$ such that S contains at least $2k$ edges. Then

$$\begin{aligned}
\mathbf{Pr}(\mathcal{B}) &\leq \sum_{k=1}^{\mu_d} \binom{n}{k} \binom{\binom{k}{2}}{2k} p^{2k} \\
&\leq \sum_{k=1}^{\mu_d} \left(\frac{ne}{k}\right)^k \left(\frac{k^2 e}{4k}\right)^{2k} \left(\frac{d}{n}\right)^{2k} \\
&= \sum_{k=1}^{\mu_d} \left(\frac{e^3 d^2 k}{16n}\right)^k \\
&= \sum_{k=1}^{\log_2 n} \left(\frac{e^3 d^2 k}{16n}\right)^k + \sum_{k=\log_2 n}^{\mu_d} \left(\frac{e^3 d^2 k}{16n}\right)^k \\
&\leq (\log_2 n) \left(\frac{e^3 d^2 \log_2 n}{16n}\right) + \mu_d \left(\frac{1}{32}\right)^{\log_2 n} \\
&= o(1).
\end{aligned}$$

□

Q3: *There does not exist $S \subseteq L_b$ such that $|S| \geq k = 2r^{-1} \ln n$ and S induces a connected subgraph in G^r.*

If S exists then we can assume that $|S| = k$ and that there exists a tree T in G such that (i) $T \cap L_b = S$, (ii) $t = |T| \leq kr$ and (iii) the leaves of T are in S. We can also assume that S contains at most $2k$ edges, from Property **Q2**.

Suppose that T has leaves L and $|L| = \ell$. We use the identity

$$\ell = 2 + \sum_{v \in T \setminus L} (\deg_T(v) - 2). \tag{7}$$

Let $T_b = (T \setminus L) \cap L_b$ and $D = \sum_{v \in L_b} \deg_T(v)$. Then (7) implies $\ell \geq D - 2(k - \ell)$ and from this we deduce that $D \leq 2k$.

Then let M be the number of edges joining $L \cup T_b$ to $V \setminus T$. We need a bound on M.

$$M \geq \ell(b-1) + (k-\ell)b - D - 2k \geq (b-5)k.$$

(Subtract $2k$ for edges in S).

So,

$$\begin{aligned}
\mathbf{Pr}(\exists S) &\leq \sum_{t=k}^{kr} \binom{n}{t}\binom{t}{k} t^{t-2}p^{t-1}\binom{k(n-k-t)}{(b-5)k}p^{(b-5)k} \\
&\leq \sum_{t=k}^{kr} n(2ed)^t(3db^{-1})^{(b-5)k} \\
&\leq 2n((2ed)^r(3db^{-1})^{b-5})^k \qquad (8)\\
&\leq 2ne^{-kr} \\
&= o(1).
\end{aligned}$$

Q4: $|B(v,r)| \leq A_3(2d)^r \ln n$ *for all* $v \in V$.

Fix $v \in V$ and let $B_i = B(v,i)$. We first observe that since $|B_{i+1}|$ is stochastically dominated by $\mathrm{Bin}(n|B_i|,p)$ we have

$$\begin{aligned}
\mathbf{Pr}(|B_{i+1}| \geq 2d|B_i| \mid |B_i| \geq 6\ln n) &\leq e^{-2d\ln n} \leq n^{-2}. \\
\mathbf{Pr}(|B_{i+1}| \geq 12\ln n \mid |B_i| \leq 6d^{-1}\ln n) &\leq e^{-2d\ln n} \leq n^{-2}.
\end{aligned}$$

Now **whp** $|B_1| = o(\ln n)$ and then **whp** either $|B(v,r)| \leq 12\ln n$ or there exists i_0 such that $|B_{i_0}| \in [6d^{-1}\ln n, 12\ln n]$. In the both cases we see that **whp** $|B(v,r| \leq 6(2d)^r \ln n$ as required.

6.3 Proof of Theorem 4

In this section we assume r, b, k are defined as in Theorem 1(b).

Proof Properties **P2,P3** are proved as for Theorem 3, the components are smaller now and so in some sense they follow from this part.

We have to work to prove **P1**. If this fails then there exists $S \subseteq V_b, T_0, T_1, \ldots T_r$ such that (i) S is connected in G^r and $T = S \cup T_0$ defines a tree in G with leaves $L \subseteq S$ (just as in **Q3** of Part (a)) and then (ii) T_1 is the neighbor set of T and T_{i+1} is the neighbour set of T_i for $0 \leq i < r$ and (iii) $|S| + |T_0| + \cdots + |T_r| \geq A_5 \ln n$.

The argument for **Q4** is still valid and this shows that **whp** $|T| \leq 2\log n$ for any $S \subseteq V_b$ which is connected in G^r We now prove that **whp** there does not exist a set Y such that (i) Y induces a connected subset of G and (ii) $|Y| \leq (\ln n)^2$ and (iii) $|N(Y)| \geq 9d|Y| + \ln n$. This will complete the verification of **P1**, since we have already shown that **whp** $|T| \leq 2\ln n$.

$$\begin{aligned}
\mathbf{Pr}(\exists Y) &\leq \sum_{t=1}^{(\ln n)^2} \binom{n}{t} t^{t-2}p^{t-1}\binom{t(n-t)}{10dt+\ln n}p^{9dt+\ln n} \\
&\leq n\sum_{t=1}^{(\ln n)^2} (de)^t(e/9)^{9dt+\ln n} \\
&= o(1).
\end{aligned}$$

□

7 Open questions

There are two natural questions that we would like to resolve:

1. Can we prove that the modified Glauber mixes rapidly if $k = O(d)$?
2. What can we say about the mixing time of the Glauber dynamics under the hypotheses of Theorem 1(a)?

References

[1] J. van den Berg and C. Maes, Disagreement percolation in the study of Markov fields, *Annals of Probability* 22 (1994), 749–763.

[2] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European Journal on Combinatorics* 1 (1980) 311–316.

[3] B. Bollobás, Martingales, isoperimetric inequalities and random graphs, *Colloq. Math. Soc. János Bolyai* 52 (1997) 113–139.

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

[5] P. Diaconis and L. Saloff-Coste, Comparison theorems for reversible Markov chains, *Annals of Applied Probabity* 3 (1993), 696–730.

[6] M. E. Dyer and A. M. Frieze, Randomly colouring graphs with lower bounds on girth and maximum degree, *Random Structures and Algorithms* 23 (2003) 167–179.

[7] M. E. Dyer, A. M. Frieze, T. P. Hayes and E. Vigoda, Randomly coloring constant degree graphs, to appear in *Proc. 45th Annual IEEE Symposium on Foundations of Computer Science*, 2004.

[8] A. M. Frieze and J. Vera, On randomly colouring locally sparse graphs, preprint, 2004.

[9] L. A. Goldberg, R. Martin and M. S. Paterson, Strong spatial mixing for lattice graphs with fewer colours, to appear in *Proc. 45th Annual IEEE Symposium on Foundations of Computer Science*, 2004.

[10] T. P. Hayes, Randomly coloring graphs of girth at least five, in *Proc. 35th Annual ACM Symposium on Theory of Computing*, pp. 269–278, 2003.

[11] T. P. Hayes and E. Vigoda, A non-Markovian coupling for randomly sampling colorings, in *Proc. 44th Annual IEEE Symposium on Foundations of Computer Science*, pp. 618-627, 2003.

[12] T. P. Hayes and E. Vigoda, Coupling with the stationary distribution and improved sampling for colorings and independent sets, preprint, 2004.

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

[14] M. R. Jerrum, *Counting, sampling and integrating: algorithms and complexity*, Birkhäuser, Basel, 2003.

[15] C. McDiarmid, Concentration, in *Probabilistic methods for algorithmic discrete mathematics* (M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, B. Reed eds.), Springer, 1998.

[16] M. Molloy, The Glauber dynamics on colorings of a graph with high girth and maximum degree, *SIAM Journal on Computing* 33 (2004), 721–737.

[17] B. G. Pittel, J. Spencer and N. C. Wormald, Sudden emergence of a giant k-core in a random graph, *Journal of Combinatorial Theory B* 67 (1996) 111-151.

[18] D. Randall and P. Tetali, Analyzing Glauber dynamics by comparison of Markov chains, *Journal of Mathematical Physics* 41 (2000), 1598–1615.

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