# Randomly colouring random graphs

Martin Dyer[*] Alan Frieze[†]

December 18, 2008

**Abstract**

We consider the problem of generating a colouring of the random graph $\mathbb{G}_{n,p}$ uniformly at random using a natural Markov chain algorithm: the Glauber dynamics. We assume that there are $\beta\Delta$ colours available, where $\Delta$ is the maximum degree of the graph, and we wish to determine the least $\beta = \beta(p)$ such that the distribution is close to uniform in $O(n \log n)$ steps of the chain. This problem has been previously studied for $\mathbb{G}_{n,p}$ in cases where $np$ is relatively small. Here we consider the "dense" cases, where $np \in [\omega \ln n, n]$ and $\omega = \omega(n) \to \infty$. Our methods are closely tailored to the random graph setting, but we obtain considerably better bounds on $\beta(p)$ than can be achieved using more general techniques.

## 1 Introduction

This paper is concerned with the problem of randomly colouring a graph with a given number of colours. The colouring should be proper in the sense that adjacent vertices get different colours and the distribution of the colouring should be close to uniform over the set of possible colourings. This problem has been the subject of intense study by researchers into Markov Chain Monte Carlo Algorithms (MCMC). In spite of this, the problem is not well solved.

Jerrum [10] showed that if $\Delta$ denotes the maximum degree of $G$ and the number of colours $q \geq 2\Delta$, then a simple algorithm can be used to generate a near random colouring in polynomial time. Vigoda [14] used a more complicated Markov Chain and improved this to $q \geq 11\Delta/6$. In spite of almost 10 years of effort, this is still the best bound known for the general case. Dyer and Frieze [4] were able to improve this bound by restricting attention to graphs of large enough maximum degree and large enough girth. There have been several improvements to this latter result and Frieze and Vigoda [5] surveys most of the known results.

The paper of Dyer, Flaxman, Frieze and Vigoda [3] considered this question in relation to random graphs, trying to bound the number of colours needed by something other than the maximum degree. They considered the random graph $G_{n,p}$ with edge probability $p = d/n$, $d$ constant and gave an upper bound on the number of colours needed that was $o(\Delta)$ **whp**. Mossel and Sly [13] have improved this result and shown that **whp** the number of colours needed satisfies $q \leq q_d$ where $q_d$ depends only on $d$. The random graph $G_{n,p}$ is the subject of this paper, but we consider graphs

[*]School of Computing, University of Leeds, Leeds LS2 9JT, UK. Research supported in part by EPSRC grant EP/E062172/1.

[†]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA 15213. Research supported in part by NSF grant CCF-0502793.

which are denser than those discussed in [3] and [13]. We will discuss the range $np \in [\omega \ln n, n]$ where $\omega = \omega(n) \to \infty$.

# 2 Glauber dynamics

We consider *colourings* $X : V \to C$ of a graph $G$, with colour set $C = [q]$. If $V' \subseteq V$, we use the notation $X(V') = \{X(v) : v \in V'\}$. A colouring is *proper* if $X(v) \neq X(w)$ $(\forall \{v,w\} \in E)$, but here colourings should not be assumed proper unless this is stated. The set of all proper colourings of $G$ will be denoted by $\Omega_G$, with $N_G = |\Omega_G|$.

We consider the *Metropolis Glauber dynamics* on the set $C^n$ of all colourings of $V$. For any colouring $X$ of a graph $G$, the *Metropolis update* is defined as follows.

1: **function** UPDATE($\mathsf{X}$)<br>
2:     $\mathsf{X}' \leftarrow \mathsf{X}$<br>
3:     choose $u \in V$ and $c \in C$ uar<br>
4:     **if** $\forall w \in \mathcal{N}_u$: $\mathsf{X}'(w) \neq c$ **then**<br>
5:         $\mathsf{X}'(u) \leftarrow c$<br>
6:     **return** $\mathsf{X}'$

We will refer to Steps 4–5 as an update of vertex $u$. An update will be called *successful* if Step 4 is executed, otherwise unsuccessful.

We analyse the following *Markov chain* $\mathsf{X}_t$ based on these dynamics, with initial colouring $\mathsf{X}$:

1: **function** GLAUBER($\mathsf{X}, \mathsf{t}$)<br>
2:     $\mathsf{X}_0 \leftarrow \mathsf{X}$<br>
3:     **for** $t \leftarrow 1$ **to** $\mathsf{t}$ **do**<br>
4:         $\mathsf{X}_t \leftarrow$ UPDATE($\mathsf{X}_{t-1}$)<br>
5:     **return** $\mathsf{X}_t$

We say the $t^{th}$ update occurs at *time* $t$, and the vertex and colour used are denoted by $u_t$, $c_t$.

For any initial colouring $\mathsf{X}$, the *equilibrium distribution* $\boldsymbol{\pi}_G$ of GLAUBER($\mathsf{X}, \infty$) is uniform on the set $\Omega_G$. Thus $\pi_G(X) = 1/N_G$ $(\forall X \in \Omega_G)$, and $\boldsymbol{\pi}_G$ is invariant under an update of any vertex.

Let $\tau(\varepsilon) = \min_t \{\mathrm{d_{TV}}(\mathsf{X}_t, \boldsymbol{\pi}_G) \leq \varepsilon\}$. Here $\mathrm{d_{TV}}$ is the *total variation distance* between distributions (see [11, p. 26]), and $\tau(\varepsilon)$ is called the *mixing time*. Our chain will be said to be *rapidly mixing* if for any initial colouring $\mathsf{X}$ and any fixed $\varepsilon$, $\tau(\varepsilon)$ is bounded by a polynomial in $n$. Our aim is to find sufficient conditions for the chain to be rapidly mixing.

## 2.1 Random graphs and colourings

Let $V^{(2)}$ denote the set of all *pairs* $\{v,w\}$ $(v, w \in V, v \neq w)$. Then $\mathbb{G}_{n,p}$ will denote the Erdős-Rényi probability space on $\mathcal{G}_n$. A random graph $\mathsf{G} = (V, \mathsf{E})$ from this distribution is defined by

$$\Pr(\{v,w\} \in \mathsf{E}) \;=\; p, \quad \text{independently for all pairs } \{v,w\} \in V^{(2)}.$$

See, for example, [9] for further information.

We consider $p(n) \geq \omega(n) \ln n/n$, where $\omega(n) \to \infty$ slowly, so that (say) $\omega = o(\log \log n)$. We also assume $1 - p(n) = \Omega(1)$ as $n \to \infty$.

We say that an event $\mathcal{E}$ occurs *with high probability* (**whp**) if $\Pr(\mathcal{E}) \geq 1 - n^{-k}$ for all $k \geq 0$. If events $\mathcal{E}_j$ $(j \in J)$ occur **whp**, it follows that the event $\bigcup_{j\in J} \mathcal{E}_j$ also occurs **whp**, whenever $|J|$ is bounded by any polynomial in $n$. We will call this the *union bound*.

We now define some constants that are used in our main theorem. First let $\beta_0 = 1.48909\ldots$ be the solution to

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

Now define $\beta_1 \approx 1.63457$ to be the solution to

$$\kappa_1(\beta) + \beta e^{-1/\beta} = 1$$

where

$$\kappa_1(\beta) = \frac{2}{\beta} \int_0^1 (1 - e^{-x/\beta}) \exp\Big( -\frac{x}{\beta} - \int_0^x \frac{\mathrm{d}y}{e^{(1-y)/\beta} - y} \Big) \mathrm{d}x. \tag{1}$$

In the next definition $\eta$ represents an arbitrarily small positive constant. Let

$$\beta(p) = (1+\eta) \times \begin{cases} \beta_0 & \omega \ln n/n \le p = o(1/\ln n) \\ \beta_1 & \Omega(1/\ln n) \le p = o(1) \\ \beta_2(p) & p = \Omega(1) \end{cases}.$$

The definition of $\beta_2(p)$ is complex and Figure 3 below provides a numerical plot.

The number of colours available will be parameterised as $\beta\Delta \approx \beta np$[1] , since **whp** every vertex of $\mathbb{G}_{n,p}$ will have vertex degree $\approx np$.

**Theorem 1.** *If $G = \mathbb{G}_{n,p}$ then Glauber Dynamics mixes in $O(n \ln n)$ time* **whp** *provided $\beta \ge \beta(p)$.* □

Some comments are in order. The value $\beta_0$ appears in Molloy [12] where $q > \beta_0\Delta$ is shown to be sufficient for rapid mixing when $\Delta = \Omega(\ln n)$ and the girth is $\Omega(\log \Delta)$. The important point to note is that for our range of values for $p$ and for almost all choices of $\mathbb{G}_{n,p}$, nothing better than Vigoda's bound of $11np/6$ can be inferred by previous results. It is perhaps of interest to point out here that **whp** $\chi(G_{n,p}) \approx \frac{np}{2 \log np}$.

# 3 Proof of Theorem 1

## 3.1 Proof Strategy

We make use of a particular choice of random initial distribution.

```
1: function RANDOM
2:     for v ∈ V do
3:         choose X(v) ∈ C uar
4:     return X
```

This initial colouring is unlikely to be a proper colouring of $\mathsf{G}$, but it is independent of $\mathsf{E}$. We will analyse $\textsc{Glauber}(\textsc{Random}, \mathsf{t}_0)$, $\mathsf{t}_0 = \lceil \lambda n \ln n \rceil$ (the value of $\lambda$ is given in Section 3.3). We will show in Lemmas 1, 2 below that this chain has nice properties. We couple the chains $\textsc{Glauber}(\mathsf{Y}, \mathsf{t}_0)$, where $\mathsf{Y} \sim \boldsymbol{\pi}_G$[2], and $\textsc{Glauber}(\textsc{Random}, \mathsf{t}_0)$. Using the aforementioned properties we show that these two chains will tend to coincide rapidly. We deduce that the equilibrium distribution also has the same nice properties. We can then use these properties to show convergence of $\textsc{Glauber}(X_0, \mathsf{t}_0)$

[1] $A_n \approx B_n$ iff $A_n = (1 + o(1))B_n$ as $n \to \infty$.

[2] We use $\sim$ to denote *has the same distribution as*

for an *arbitrary* initial colouring $X_0$. This is achieved by coupling the chains $\text{Glauber}(\mathsf{Y}, \mathfrak{t}_0)$ and $\text{Glauber}(X_0, \mathfrak{t}_0)$.

### 3.2 Coupling Argument

Let $\mathcal{G}_n = \{G = (V, E) : V = [n]\}$ be the set of $n$-vertex graphs. If $G \in \mathcal{G}_n$, the set of *neighbours* of a vertex $v \in V$ will be denoted by $\mathcal{N}_v = \{w \in V : \{v, w\} \in E\}$, and its *degree* $d_v = |\mathcal{N}_v|$. The *maximum degree* of $G$ is $\Delta(G) = \max_{v \in V} d_v$.

If $X$ is a colouring of $G$ and $v \in V$, let

$$\mathcal{A}_v(X) = \{c \in C : X(w) \neq c, \forall w \in \mathcal{N}_v\}$$

be the set of colours which are available for a successful update of $v$. Also, let $A_v(X) = |\mathcal{A}_v(X)|$. Clearly we always have $A_v(X) \geq q - \Delta = (\beta - 1)\Delta$.

We prove convergence of Glauber using a *coupling* due to Jerrum [10] and an idea of Hayes and Vigoda [7]. In [7] an arbitrary chain is coupled with one started in an equilibrium distribution whose properties are known. We use the following coupling, which is essentially that used by Jerrum [10]. Let $u_X, c_X$ and $u_Y, c_Y$ be the choices in Step 3 of $\text{Update}(X_t)$ and $\text{Update}(Y_t)$ respectively.

1: **procedure** $\text{Couple}(X, Y)$
2: Choose $u \in V$ uar
3: $u_X \leftarrow u$, $u_Y \leftarrow u$
4: $\mathcal{A}_X \leftarrow \mathcal{A}_u(X)$, $\mathcal{A}_Y \leftarrow \mathcal{A}_u(Y)$
5: $A_X \leftarrow |\mathcal{A}_X|$, $A_Y \leftarrow |\mathcal{A}_Y|$
6: Construct a bijection $f : C \to C$ such that
7: $f(c) \leftarrow c \quad (\forall c \in \mathcal{A}_X \cap \mathcal{A}_Y)$
8: **if** $A_X \leq A_Y$ **then** $f(\mathcal{A}_X) \subseteq \mathcal{A}_Y$ **else** $f(\mathcal{A}_Y) \subseteq \mathcal{A}_X$
9: Choose $c \in C$ uar
10: $c_X \leftarrow c$, $c_Y \leftarrow f(c)$

Note that Couple ensures that $c_Y$ will be sampled uar from $C$, as required. For brevity, we will make use of the notation defined in Couple below.

Let $X', Y'$ be the colourings after updating $X$ and $Y$ with $u$, $c_X = c$, $c_Y = f(c)$ chosen by Couple. Let $\mathcal{W} = \{w \in V : X(w) \neq Y(w)\}$ and $W = |\mathcal{W}| = \mathrm{d}(X, Y)$, where $\mathrm{d}(\cdot, \cdot)$ is the *Hamming distance*. For any $v \in V$, let $\mathcal{W}_v = \mathcal{W} \cap \mathcal{N}_v$ and $W_v = |\mathcal{W}_v|$. Note that

$$\textstyle\sum_{v \in V} W_v \;\leq\; \sum_{w \in \mathcal{W}} d_w \;\leq\; \Delta W.$$

Note that $c \in \mathcal{A}_X \setminus \mathcal{A}_Y$ implies there is a $w \in \mathcal{N}_u$ such that $X(w) \neq c$ and $Y(w) = c$. Hence $|\mathcal{A}_X \setminus \mathcal{A}_Y| \leq W_u$, and similarly $|\mathcal{A}_Y \setminus \mathcal{A}_X| \leq W_u$. Let

$$A_u = \max\{A_X, A_Y\}$$

so that we have

$$A_u - |\mathcal{A}_X \cap \mathcal{A}_Y| \;=\; \max\{|\mathcal{A}_X \setminus \mathcal{A}_Y|, |\mathcal{A}_Y \setminus \mathcal{A}_X|\} \;\leq\; W_u. \tag{2}$$

Also, from Step 8 of Couple, there are $q - A_u$ pairs $c, f(c)$ such that $c \notin \mathcal{A}_X$, $f(c) \notin \mathcal{A}_Y$. Hence there are at most $A_u - |\mathcal{A}_X \cap \mathcal{A}_Y| \leq W_u$ pairs which result in $X'(u) \neq Y'(u)$. Letting $W' = \mathrm{d}(X', Y')$, and assuming that we have a bound

$$A_u \geq \alpha\Delta \tag{3}$$

for some $\alpha > 1$ that holds **whp** for all $t \le n^2$ say, then we have

$$\mathbb{E}[W'-W] \;\le\; -\sum_{u\in\mathcal{W}} \frac{A_u - W_u}{nq} + \sum_{u\notin\mathcal{W}} \frac{W_u}{nq} \le -\frac{\alpha\Delta W}{nq} + \sum_{u\in V} \frac{W_u}{nq} \le -\frac{\alpha\Delta W}{nq} + \frac{\Delta W}{nq} = -\frac{(\alpha-1)W}{n\beta}, \tag{4}$$

since $q = \beta\Delta$.

Therefore,

$$\mathbb{E}[W'] \le \Big(1 - \frac{\alpha-1}{n\beta}\Big)W. \tag{5}$$

Using the Coupling Lemma (see [11, Lemma 4.7]),

$$\mathrm{d_{TV}}(\mathsf{X}_t, \mathsf{Y}_t) \;\le\; \Pr(\mathsf{X}_t \ne \mathsf{Y}_t) \;\le\; \mathbb{E}[\mathrm{d}(\mathsf{X}_t, \mathsf{Y}_t)] \;\le\; n\Big(1 - \frac{\alpha-1}{n\beta}\Big)^t \;\le\; ne^{-(\alpha-1)t/n\beta} \;\le\; \varepsilon,$$

for $t \ge \beta n \ln(n/\varepsilon)/(\alpha-1)$.

Thus one part of the analysis involves finding a lower bound as in (3).

A further improvement is possible using a modification of a method of Molloy [12]. Now, for any $v \in V$, let us define

$$\mathcal{W}^*_v = \{w \in \mathcal{N}_v : X(w) \notin Y(\mathcal{N}_v) \vee Y(w) \notin X(\mathcal{N}_v)\}, \text{ and } W^*_v = |\mathcal{W}^*_v|.$$

Note that $\mathcal{W}^*_v \subseteq \mathcal{W}_v$, so $W^*_v \le W_v$, since $X(w) = Y(w)$ clearly implies $X(w) \in Y(\mathcal{N}_v)$ and $Y(w) \in X(\mathcal{N}_v)$. Also, we will define

$$\mathcal{N}^*_v = \{w \in \mathcal{N}_v : X(v) \notin Y(\mathcal{N}_w) \vee Y(v) \notin X(\mathcal{N}_w)\} \text{ and } d^*_v = |\mathcal{N}^*_v|. \tag{6}$$

We obviously have $\mathcal{N}^*_v \subseteq \mathcal{N}_v$, and hence $d^*_v \le d_v$. Note also that $w \in \mathcal{N}^*_v$ iff $v \in \mathcal{W}^*_w$. Furthermore, $\mathcal{N}^*_v = \emptyset$ if $X(v) = Y(v)$, since this implies $X(v) \in Y(\mathcal{N}_w)$ and $Y(v) \in X(\mathcal{N}_w)$ for all $w \in \mathcal{N}_v$. Thus $d^*_v = 0$ for all $v \notin \mathcal{W}$. Hence,

$$\textstyle\sum_{v\in V} W^*_v \;=\; \sum_{w\in\mathcal{W}} d^*_w.$$

Note that $c \in \mathcal{A}_X \setminus \mathcal{A}_Y$ implies there is a $w \in \mathcal{N}_u$ such that $Y(w) = c$ and $c \notin X(\mathcal{N}_u)$. Hence $|\mathcal{A}_X \setminus \mathcal{A}_Y| \le W^*_u$, and similarly $|\mathcal{A}_Y \setminus \mathcal{A}_X| \le W^*_u$. Hence

$$A_u - |\mathcal{A}_X \cap \mathcal{A}_Y| \;=\; \max\{|\mathcal{A}_X \setminus \mathcal{A}_Y|, |\mathcal{A}_Y \setminus \mathcal{A}_X|\} \;\le\; W^*_u.$$

which is a strengthening of (2).

So,

$$W' - W \le -\sum_{u\in\mathcal{W}} \frac{A_u - W^*_u}{nq} + \sum_{u\notin\mathcal{W}} \frac{W^*_u}{nq} \le -\frac{\alpha W}{nq} + \sum_{u\in V} \frac{W^*_u}{nq} = -\frac{\alpha W}{nq} + \sum_{v\in\mathcal{W}} \frac{d^*_v}{nq}.$$

We will prove an upper bound

$$\mathbb{E}[d^*_v \mid v \in \mathcal{W}] \le \gamma\Delta$$

that holds for $t \le n^2$, say and then we will have a strengthened (5) to

$$\mathbb{E}[W'] \le \Big(1 - \frac{\alpha-\gamma}{n\beta}\Big)W. \tag{7}$$

and convergence will occur in time $O(n\log n)$ provided $\alpha > \gamma$ as opposed to $\alpha > 1$.

The rest of the paper provides the following bounds for $\alpha, \gamma$: The quantities in the next two lemmas refer to the chain Glauber(Random, $t$). The lemmas describe the nice properties referred to in Section 3.1.

**Lemma 1.** $A_v \gtrsim \alpha_p\Delta$ **whp** *for* $v \in V$, $0 \le t \le n^2$ *where*

$$\alpha_p = \alpha_p(\beta) = \begin{cases} \beta - 1, & \text{if } q \ge n, \\ \beta - \mu + \mu(1-p)^{1/\mu p}(1-p)^{1/\beta p}, & \\ \qquad\qquad \text{where } \mu = \min\left\{\beta(1-e^{-1/\alpha p}), 1/p\right\} & \text{if } q < n,\ p = \Omega(1), \\ \beta e^{-1/\beta}, & \text{if } q < n,\ p = o(1). \end{cases}$$

**Lemma 2.** $\mathbb{E}[d_v^* \mid v \in \mathcal{W}] \lesssim \gamma_p\Delta$ **whp** *for* $v \in V$, $0 \le t \le n^2$ *where*

$$\gamma_p = \gamma_p(\beta) = \begin{cases} 1 - (1-e^{-1/\beta})^2 & \omega \ln n/n \le p = o(1/\ln n) \\ 1 - \kappa_1(\beta) & \Omega(1/\ln n) \le p = o(1) \\ 1 - \kappa_2(\beta) & p = \Omega(1) \end{cases}$$

*Here* $\kappa_1(\beta)$ *is as defined in (1) and* $\kappa_2(\beta)$ *is defined in (41) and has only been computed numerically.* Putting them together with (7) yields Theorem 1.

### 3.3 Proof of Lemma 1

Our analysis will be split into three cases: where $1/p \ge \omega \ln n$, where $1/p \le \omega \ln n$ but $p = o(1)$, and where $p = \Theta(1)$. Slightly different techniques are required in each case. For notational convenience, we write $\lambda(n) = \sqrt[3]{\omega}$ and $\epsilon(n) = 1/\lambda$, so $\lambda \to \infty$ and $\epsilon \to 0$ as $n \to \infty$. We will also use $\theta(n)$ to denote any function such that $\theta = \Omega(\lambda)$ as $n \to \infty$. We will use $\epsilon_0, \varepsilon_0$ to indicate small constants, where $\varepsilon_0$ will always be a small probability.

For any colouring $X$ of $G$ and $c \in C$, the *colour class* of $c$ is $S_c(X) = \{v \in V : X(v) = c\}$, with $s_c(X) = |S_c(X)|$.

**Lemma 3.** *If* $\mathsf{X}_0 \sim$ Random *then* **whp**, *for all* $c \in C$, $0 \le t \le \mathsf{t}_0$,

$$s_c(\mathsf{X}_t) \le \begin{cases} (1+\epsilon)/(\beta-1)p < \epsilon^2 n/\ln n = o(n/\log n), & \text{if } 1/p \ge \omega \ln n. \\ \lambda\sqrt{\ln n}/(\beta-1)p < \epsilon^4 \ln^2 n \ \ = o(\log^2 n), & \text{if } 1/p \le \omega \ln n. \end{cases}$$

*Proof.* Fix a value of $t$. For all $v \in V$ and $c \in C$, we have $\Pr(\mathsf{X}(v) = c) \le 1/(q-\Delta)$. This is true initially since $\mathsf{X}_0 =$ Random, and hence $\Pr(\mathsf{X}_0(v) = c) = 1/q$. It remains true thereafter because, conditional on a successful update of $v$, Update independently selects a colour uar from $\mathcal{A}_v(\mathsf{X})$, which has size at least $(q-\Delta)$. Thus $1_{\mathsf{X}(v)=c} \le \mathsf{B}_v$, where $\mathsf{B}_v \in \{0,1\}$ are iid[3] with $\Pr(\mathsf{B}_v = 1) = 1/(q-\Delta)$. Therefore $s_c(\mathsf{X}_t)$ is dominated by $\mathsf{B} = \sum_{v\in V} \mathsf{B}_v \sim \mathrm{Bin}(n, 1/(q-\Delta))$. We have $\mathbb{E}[\mathsf{B}] = n/(q-\Delta) = 1/(\beta-1)p$. If $1/p \ge \omega \ln n$, let $k = (1+\epsilon)/(\beta-1)p \le \epsilon^2 n/\ln n$, for large enough $n$. The Chernoff bound gives

$$\Pr(s_c(\mathsf{X}_t) \ge k) \le \exp\Big(-\frac{\epsilon^2}{3(\beta-1)p}\Big) \le n^{-\theta}.$$

If $1/p \le \omega \ln n$, let $k = \epsilon^4 \ln^2 n$. Now

$$k \ge 7\mathbb{E}[s_c] = 7/(\beta-1)p = O(\lambda^3 \ln n),$$

since $\lambda^7 = o(\ln n)$, so a variant of the Chernoff bound [9, (2.11)] gives

$$\Pr(s_c(\mathsf{X}_t) \ge k) \le \exp(-\epsilon^4 \ln^2 n) \ < \ n^{-\theta}.$$

The conclusion now follows, using the union bound. □

[3] *independently and identically distributed*

**Lemma 4.** *If* $\mathsf{Y}_0 \sim \boldsymbol{\pi}$, $s_c(\mathsf{Y}_t) < \epsilon^2 n/\ln n = o(n/\log n)$ $(\forall c \in C,\ 0 \le t \le \mathfrak{t}_0)$ **whp**.

*Proof.* Since $\mathsf{Y}_t$ is invariant under a successful update of any vertex, $\Pr(\mathsf{Y}_t(v) = c) \le 1/(q-\Delta)$, $(\forall v \in V,\ c \in C)$ and the proof of Lemma 3 applies. □

The key fact underlying our analysis is that, for $p > \omega \ln n/n$ and $q = \beta\Delta$ for large enough $\beta$, Glauber(Random, $\mathfrak{t}_0$) does not inspect many edges of $\mathsf{G}$ before it terminates. This allows us to perform most of our calculations in the random graph model $\mathbb{G}_{n,p}$. To show that this is valid, we will analyse Update with respect to a fixed vertex $v$. To emphasise its dependence on $v$, we will rewrite Update as follows. Note, however, that the algorithm is effectively unchanged.

1: **function** Update($\mathsf{X}, v$)
2: $\quad \mathsf{X}' \leftarrow \mathsf{X}$
3: $\quad$ choose $u \in V$ and $c \in C$ uar
4: $\quad$ **if** $u = v$ **then**
5: $\quad\quad$ **if** $\forall w \in S_c(\mathsf{X}')$: $\{v, w\} \notin E$ **then**
6: $\quad\quad\quad \mathsf{X}'(v) \leftarrow c$
7: $\quad$ **else**
8: $\quad\quad$ **if** $c \ne \mathsf{X}'(v)$ **then**
9: $\quad\quad\quad$ **if** $\forall w \in S_c(\mathsf{X}')$: $\{u, w\} \notin E$ **then**
10: $\quad\quad\quad\quad \mathsf{X}'(u) \leftarrow c$
11: $\quad\quad$ **else**
12: $\quad\quad\quad$ **if** $\{u, v\} \notin E$ **then**
13: $\quad\quad\quad\quad$ **if** $\forall w \in S_c(\mathsf{X}') \setminus \{v\}$: $\{u, w\} \notin E$ **then**
14: $\quad\quad\quad\quad\quad \mathsf{X}'(u) \leftarrow c$
15: $\quad$ **return** $\mathsf{X}'$

We will say that Glauber *exposes* the pair $\{v, w\} \in V^{(2)}$ if it is necessary to determine whether or not $\{v, w\} \in \mathsf{E}$ during some update. We say Glauber *exposes an edge* $e = \{v, w\}$ if the exposed pair $\{v, w\} \in \mathsf{E}$. Otherwise it exposes a *non-edge* $\{v, w\}$. Let $\mathcal{D}_v$ be the set of pairs $\{v, w\}$ exposed by Glauber(Random, $\mathfrak{t}_0$), and let $D_v = |\mathcal{D}_v|$. Note that no pairs are exposed when $t = 0$.

**Lemma 5.** $D_v \le \lambda^2 \ln n/p \le \epsilon n = o(n)$ $(\forall v \in V)$ **whp**.

*Proof.* Update($X, v$) exposes pairs $\{v, w\}$ $(w \in V)$ only in Steps 5 and 12. Step 5 is executed, independently with probability $1/n$, when $u = v$ in Update. Let $\mathsf{K}_1$ be the number of times Step 5 of Update is executed in Glauber(Random, $\mathfrak{t}_0$). Then the Chernoff bound gives

$$\Pr(\mathsf{K}_1 > 2\lambda \ln n) \le \exp(-\tfrac{1}{3}\lambda \ln n) = n^{-\lambda/3} \;=\; n^{-\theta}.$$

Assume first that $1/p \ge \omega \ln n$. When Step 5 is executed, all pairs $\{v, w\}$ for $w \in S_c(X_t)$ are exposed, for some $c$. By Lemma 3, this is at most $2/(\beta-1)p$ **whp**. Thus **whp** at most $2\lambda \ln n \times 2/(\beta-1)p = 4\lambda \ln n/(\beta-1)p$ pairs are exposed by occurrences of Step 5 of Update in Glauber(Random, $\mathfrak{t}_0$). If $1/p \le \omega \ln n$ then by Lemma 3 the number of pairs exposed is at most $2\lambda \ln n \times \lambda\sqrt{\ln n}/(\beta-1)p = 2\lambda^2 \ln^{3/2} n/(\beta-1)p$.

Step 12 is executed, independently with probability less than $1/q$, if $u \ne v$ and $c = X_t(v)$ in Update. Let $\mathsf{K}_2$ be the number of times Step 12 of Update is executed in Glauber(Random, $\mathfrak{t}_0$). The Chernoff bound gives

$$\Pr(\mathsf{K}_2 > 2\lambda \ln n/\beta p) \le \exp(-\tfrac{1}{3}\lambda \ln n/\beta p) < n^{-\lambda/3} \;=\; n^{-\theta}.$$

Each execution of Step 12 exposes one pair. Thus **whp** at most $2\lambda \ln n/\beta p$ pairs are exposed by occurrences of Step 12. Thus in total, we expose at most $\epsilon n$ pairs **whp**. The union bound completes the proof. □

We now show that we have not exposed many edges $\{v,w\} \in \mathsf{E}$.

**Lemma 6.** $|\{w \in \mathcal{D}_v : \{v,w\} \in \mathsf{E}\}| \le 2\lambda^2 \ln n \le 2\epsilon np = o(\Delta)$ $(v \in V)$ **whp**.

*Proof.* Let $\mathsf{K} = |\{w \in \mathcal{D}_v : \{v,w\} \in \mathsf{E}\}|$. When any pair $\{v,w\}$ is exposed, $\Pr(\{v,w\} \in \mathsf{E}) = p$ independently. Since $D_v < \lambda^2 \ln n/p$ by Lemma 5, the Chernoff bound gives

$$\Pr(\mathsf{K} > 2\lambda^2 \ln n) \le \exp(-\tfrac{1}{3}\lambda^2 \ln n) = n^{-\lambda^2/3} < n^{-\theta}.$$ □

We will now bound $A_v(\mathsf{X}_t)$ below, for all $\mathsf{X}_t = \text{GLAUBER}(\text{RANDOM}, t)$ and all $v \in V$.

**Lemma 7.** *For all* $0 \le t \le \mathsf{t}_0$, **whp**,

$$A_{\min} \gtrsim^{4} \Delta \times \begin{cases} \beta - 1, & \text{if } q \ge n, \\ \beta(1-p)^{1/\beta p}, & \text{if } q < n,\ p = \Omega(1), \\ \beta e^{-1/\beta}, & \text{if } p = o(1). \end{cases}$$

*Proof.* Fix $v \in V$ and $t \le \mathsf{t}_0$. By Lemma 5, $\text{GLAUBER}(\text{RANDOM}, \mathsf{t}_0)$ exposes at most $\epsilon n$ pairs and at most $2\epsilon np$ edges adjacent to $v$. Thus $n' \ge (1-\epsilon)n$ pairs $\{v,w\}$ are not exposed. Thus we can write $\Pr(c \in \mathcal{A}_v) = p_c$ independently for all $c \in C$. Here $p_c = 0$ for $c \in C' = \{\mathsf{X}_t(w) : \{v,w\} \textit{ exposed},\ \{v,w\} \in \mathsf{E}\}$ and $p_c \ge (1-p)^{s_c}$ otherwise. Let $M = \{c \in C : s_c > 0\}$ be the set of $m = |M|$ colours appearing in $\mathsf{X}_t$. Note that $\sum_{c \in M} s_c = n$. Let $\mathcal{U}_t = (C \setminus C') \setminus \{X_t(w) : \{v,w\} \text{ unexposed}, \{v,w\} \in \mathsf{E}\}$, and $U_t = |\mathcal{U}_t| \le A_v(\mathsf{X}_t)$. Then, using the arithmetic-geometric mean inequality (see, for example, [6]), and $|C'| \le 2\epsilon np$ **whp**,

$$\mathbb{E}[U_t] \ge \sum_{c \in C \setminus C'} (1-p)^{s_c} \ge q - |C'| - m + \sum_{c \in M \setminus C'} (1-p)^{s_c} \gtrsim q - m + m(1-p)^{n/m}. \tag{8}$$

Thus $\mathbb{E}[U_t] \gtrsim \min\{q - m + m(1-p)^{n/m} : 0 \le m \le \min\{q,n\}\}$. We also always have the bound $U_t \ge (\beta-1)\Delta$. Thus, since $\Delta \approx np$ and $q = \beta\Delta$,

$$\mathbb{E}[U_t] \gtrsim \hat{A} = \begin{cases} q - \Delta = (\beta-1)\Delta, & \text{if } q \ge n, \\ q(1-p)^{n/q} \approx \beta(1-p)^{1/\beta p}\Delta, & \text{if } q < n,\ p = \Omega(1), \\ qe^{-\Delta/q} = \beta e^{-1/\beta}\Delta, & \text{if } p = o(1). \end{cases}$$

Now $U_t$ is a sum of $q$ independent binary random variables, since $\mathsf{G} \sim \mathbb{G}_{n,p}$. Therefore, since $\mathbb{E}[U_t] \ge (\beta-1)\Delta = (\beta-1)\lambda^3 \ln n$, Hoeffding's inequality [9, Thm. 2.8] gives

$$\Pr(U_t \le (1-\epsilon)\mathbb{E}[U_t]) \le \exp(-\tfrac{1}{3}\epsilon^2\mathbb{E}[U_t]) \le \exp(-\tfrac{1}{3}\epsilon^2\lambda^3 \ln n) = n^{-\lambda/3} = n^{-\theta}.$$

Finally, we use the union bound over all $v \in V$, and $0 \le t \le \mathsf{t}_0$. □

This proves Lemma 1 except for the case $p = \Omega(1)$. When $p = \Omega(1)$, we can marginally improve Lemma 7 by observing that some colours will not appear. Let $M$ and $m$ be as defined in the proof of Lemma 7. Let $\bar{A}$ be a uniform lower bound on the right hand side of (8) which holds **whp**. Clearly we may take $\bar{A} \ge \hat{A}$, so $\bar{A} = \Theta(n)$. Let $\alpha = \bar{A}/\Delta \approx \bar{A}/np$ and $\mu = m/\Delta \approx m/np$.

**Lemma 8.** *For all* $v \in V$ *and* $0 \le t \le \mathsf{t}_0$, $A_v(\mathsf{X}_t) \gtrsim \alpha\Delta$ **whp**, *where* $\alpha = \beta - \mu + \mu(1-p)^{1/\mu p}$ *and* $\mu = \min\{\beta(1 - e^{-1/\alpha p}), 1/p\}$.

[4] $A_n \gtrsim B_n$ if $A_n \ge 1 - o(1))B_n$ as $n \to \infty$.

*Proof.* Let $\bar{m} = q - m$ denote the number of colours which *do not* appear anywhere in the graph. Sort $V$ in increasing order of the epochs at which vertices were last successfully recoloured before $t$. Vertices that have not been successfully recoloured are put before those that have, in arbitrary order. Then relabel $V$ so that this order is $1, 2, \ldots, n$. Let $\bar{m}_k$ be the number of colours not present in $\{X_t(1), X_t(2), \ldots, X_t(k)\}$, so $\bar{m}_0 = q$. Now, since each $X_t(i)$ ($i \in [n]$) is chosen uar from a set of size at least $\bar{A}$, we have

$$\mathbb{E}[\bar{m}_{k+1}] \;\geq\; \bar{m}_k - \bar{m}_k/\bar{A} \;=\; (1 - 1/\bar{A})\bar{m}_k,$$

and hence $\mathbb{E}[\bar{m}_k] \;\geq\; q(1 - 1/\bar{A})^k$. In particular,

$$\mathbb{E}[\bar{m}] = \mathbb{E}[\bar{m}_n] \geq q(1 - 1/\bar{A})^n \approx qe^{-n/\bar{A}} \approx qe^{-1/\alpha p} = \Theta(n),$$

since $q, \bar{A} = \Omega(n)$. Let us define $Z_k = (1 - 1/\bar{A})^{n-k}\bar{m}_k$, so $Z_n = \bar{m}_n$. Then

$$\mathbb{E}[Z_{k+1}] = (1 - 1/\bar{A})^{n-k-1}\mathbb{E}[\bar{m}_{k+1}] \;\geq\; (1 - 1/\bar{A})^{n-k}\bar{m}_k \;=\; Z_k,$$

and hence $Z_k$ is a submartingale. Also, $|\bar{m}_{k+1} - \bar{m}_k| \leq 1$, $\bar{m}_k \leq q$ and $\bar{A} \geq q - \Delta$ and so

$$\begin{aligned} |Z_{k+1} - Z_k| \;&=\; (1 - 1/\bar{A})^{n-k-1}|\bar{m}_{k+1} - \bar{m}_k + \bar{m}_k/\bar{A}| \\ &\leq\; 1 + q/(q - \Delta) \\ &=\; (2q - \Delta)/(q - \Delta) \\ &=\; (2\beta - 1)/(\beta - 1), \end{aligned}$$

and hence $Z_k$ is a bounded-difference submartingale. Thus, by the martingale inequality [9, p. 37],

$$\Pr(\bar{m} < (1 - \epsilon)qe^{-1/\alpha p}) \;\approx\; \Pr(Z_n < (1 - \epsilon)\mathbb{E}[Z_n]) \;\leq\; \exp(-\epsilon^2\Theta(n)) \;\leq\; \exp(-\Theta(n^{1/3})). \qquad (9)$$

Now, from (8) we have $\alpha = \beta - \mu + \mu(1 - p)^{1/\mu p}$, and from (9) we have $\mu \lesssim \beta(1 - e^{-1/\alpha p})$. The bound $\mu \leq 1/p$, which is equivalent to $m \leq n$, is trivial. ☐

The proof of Lemma 1 is now complete. The resulting values of $\beta(p)$ for $p = \Theta(1)$ are plotted in Figure 1. ☐

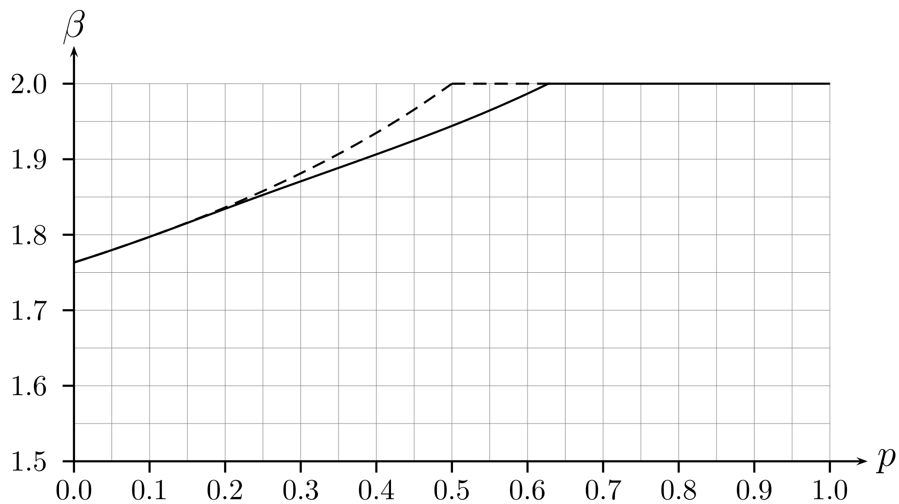


Figure 1: Bound on $\beta$ for convergence of Glauber, using Lemmas 7 and 8.

## 3.4 Proof of Lemma 2

We split the proof of Lemma 2 into the three ranges that we have indicated.

### 3.4.1 The case $p = o(1/\log n)$

We will first analyse in more detail the number of occurrences of each colour after sufficiently many GLAUBER updates. We show that the size of all colour sets is highly concentrated after sufficiently many rounds of GLAUBER.

**Lemma 9.** *Let $0 < \epsilon_0 \le 1$ be a constant and $\delta_0 = n^{-K}$ for some constant $K > 0$. Then, with probability at least $1-\delta_0$, we have $s_c \ge (1-\epsilon_0)n/q$ ($\forall c \in C$) for all $\mathsf{t}_1 \le t \le \mathsf{t}_0$, where*

$$\mathsf{t}_1 \;=\; e^{1/\beta}\ln(\beta/\epsilon_0)n\ln\big(\ln(\beta/\epsilon_0)n/\delta_0\big).$$

*Proof.* Let $\nu = \lfloor n/q \rfloor \approx 1/\beta p$. If $s_c(X_t) \le \nu$, the probability that GLAUBER updates a known neighbour $v$ of $S_c(X_t)$ is at most $3\epsilon nps_c/n \lesssim 3\epsilon/\beta$, from Lemma 6. Therefore

$$\Pr\big(X_{t+1}(u_{t+1})=c\big) \;\gtrsim\; (1-p)^{s_c}/q \;\approx\; e^{-ps_c}/q \;\gtrsim\; e^{-1/\beta}/q,$$

since the pairs $\{v,w\}$ ($w \in S_c$) are exposed. Thus

$$\Pr\big(s_c(\mathsf{X}_{t+1}) = s+1 \mid s_c(\mathsf{X}_t)=s\big) \;\gtrsim\; e^{-1/\beta}/q \qquad (s=0,1,2,\ldots,\nu-1). \tag{10}$$

If $s_c(X_t) > \nu$, we will assume only that $\Pr\big(s_c(\mathsf{X}_{t+1}) = s+1 \mid s_c(\mathsf{X}_t)=s\big) \ge 0$. We may ignore $(1 \pm o(1))$ factors in the probability estimates since we will state our lower bounds on $\beta$ for rapid mixing only as strict inequalities. We do not seek to analyse the equality cases, which are rather more problematic. See, for example, [1].

Suppose $\varphi$ is such that **whp** $s_c(X_t) \ge (1-\varphi)\nu$ ($\forall c \in C$, $0 \le t \le \mathsf{t}_0$). Note that $0 \le \varphi \le 1$. It follows, as in the proof of Lemma 7 (in particular (8)) that **whp**

$$A_v(X_t) \;\lesssim\; \sum_{c\in C}(1-p)^{(1-\varphi)\nu} \;\approx\; qe^{-(1-\varphi)/\beta}. \tag{11}$$

Then we have

$$\Pr\big(s_c(\mathsf{X}_{t+1}) = s-1 \mid s_c(X_t)=s\big) \;\lesssim\; \frac{s}{n}\,\frac{qe^{-(1-\varphi)/\beta}}{q} \;=\; \frac{se^{-(1-\varphi)/\beta}}{n} \qquad (s=1,2,\ldots,\nu) \tag{12}$$

Putting $\boldsymbol{\sigma}_{t+1} = s_c(\mathsf{X}_t)$ and $\rho = ne^{-\varphi/\beta^*}/q \approx \nu e^{-\varphi/\beta^*} \le \nu$ where $\beta^* = (1-o(1))\beta$ for a suitable choice of $o(1)$ we see that $s_c(\mathsf{X}_t)$ dominates the process $\boldsymbol{\sigma}_t$ with state space $\{0,1,\ldots,\nu\}$ such that $\boldsymbol{\sigma}_0 = 0$ and

$$\begin{aligned}
\Pr\big(\boldsymbol{\sigma}_{t+1} = \sigma+1 \mid \boldsymbol{\sigma}_t=\sigma\big) \;&=\; \frac{e^{-1/\beta^*}}{q}, && (\sigma = 0,1,\ldots,\nu-1)\\
\Pr\big(\boldsymbol{\sigma}_{t+1} = \sigma-1 \mid \boldsymbol{\sigma}_t=\sigma\big) \;&=\; \frac{\sigma e^{-(1-\varphi)/\beta^*}}{n}, && (\sigma = 0,1,\ldots,\nu)\\
\Pr\big(\boldsymbol{\sigma}_{t+1} = \sigma \mid \boldsymbol{\sigma}_t=\sigma\big) \;&=\; 1-\frac{e^{-1/\beta^*}}{q}-\frac{\sigma e^{-(1-\varphi)/\beta^*}}{n}, && (\sigma = 0,1,\ldots,\nu-1)\\
\Pr\big(\boldsymbol{\sigma}_{t+1} = \nu \mid \boldsymbol{\sigma}_t=\nu\big) \;&=\; 1-\frac{\nu e^{-(1-\varphi)/\beta^*}}{n}.
\end{aligned}$$

This is a reversible Markov chain with equilibrium distribution,

$$\Pr\big(\boldsymbol{\sigma}_\infty = \sigma\big) \;=\; \frac{\rho^\sigma e^{-\rho}}{\psi\,\sigma!} \quad (\sigma=0,1,\ldots,\nu), \qquad \psi \;=\; \sum_{\sigma=0}^{\nu}\frac{\rho^\sigma e^{-\rho}}{\sigma!}.$$

as may be verified using the *detailed balance* equations

$$\frac{e^{-1/\beta^*}}{q}\,\frac{\rho^{\sigma-1}\,e^{-\rho}}{\psi\,(\sigma-1)!} \;=\; \frac{\sigma e^{-(1-\varphi)/\beta^*}}{n}\,\frac{\rho^\sigma e^{-\rho}}{\psi\,\sigma!} \qquad (\sigma=1,2,\ldots,\nu).$$

The equilibrium distribution is clearly Poiss($\rho$) conditional on $\sigma \le \nu$. If $x \sim$ Poiss($\rho$) and $\varphi = \Omega(\epsilon)$, the Chernoff bound gives

$$\Pr\big(x < (1-\epsilon)\rho \vee x > (1+\epsilon)\rho\big) \;\le\; 2\exp(-\tfrac{1}{3}\epsilon^2\rho) \;=\; \exp(-\Omega(\epsilon^2\omega\log n)) \;\le\; n^{-\theta}. \tag{13}$$

Thus **whp** we have

$$\boldsymbol{\sigma}_\infty \;\ge\; (1-\epsilon)\rho \approx \frac{ne^{-\varphi/\beta^*}}{q} \;\ge\; \Big(1-\frac{\varphi}{\beta^*}\Big)\frac{n}{q}. \tag{14}$$

We may bound the mixing time of the process $\boldsymbol{\sigma}_t$ using *path coupling* [2]. We consider two copies $x_t$, $y_t$ of the process $\boldsymbol{\sigma}_t$, where $y_0$ is sampled from the equilibrium distribution. We will use the metric $\mathrm{d}(x_t, y_t) = |x_t - y_t|$. We define the coupling on adjacent states $x_t = \sigma - 1$, $y_t = \sigma$ $(1 \le \sigma \le \nu)$, so $\mathrm{d}(x_t, y_t) = 1$. Conditionally on $x_t = \sigma - 1$, $y_t = \sigma$, for $0 < \sigma \le \nu$, we couple the processes as follows

$$\Pr(x_{t+1} = x,\, y_{t+1} = y) \;=\; \begin{cases} \dfrac{(\sigma-1)e^{-(1-\varphi)/\beta^*}}{n}, & \text{if} \quad x = \sigma-2, \quad y = \sigma-1, \\ 1 - \dfrac{\sigma e^{-(1-\varphi)/\beta^*}}{n} - \dfrac{e^{-1/\beta^*}}{q}, & \text{if} \quad x = \sigma-1, \quad y = \sigma, \\ \dfrac{e^{-1/\beta^*}}{q}, & \text{if} \quad x = \sigma, \quad y = \min\{\sigma+1, \nu\}, \\ \dfrac{e^{-(1-\varphi)/\beta^*}}{n}, & \text{if} \quad x = \sigma-1, \quad y = \sigma-1. \end{cases}$$

It is now easy to see that

$$\mathbb{E}[\mathrm{d}(x_{t+1}, y_{t+1}) \mid \mathrm{d}(x_t, y_t) = 1] \;\le\; 1 - \frac{e^{-(1-\varphi)/\beta^*}}{n},$$

and hence

$$\Pr(x_t \ne y_t) \;\le\; \mathbb{E}[\mathrm{d}(x_t, y_t)] \;\le\; \mathrm{d}(x_0, y_0)\Big(1 - \frac{e^{-(1-\varphi)/\beta^*}}{n}\Big)^t \;\le\; \frac{n}{q}\Big(1 - \frac{e^{-1/\beta^*}}{n}\Big)^t,$$

from which it follows, using the Coupling Lemma, that for any $\delta_0$ we will have

$$\mathrm{d}_{\mathrm{TV}}(\boldsymbol{\sigma}_t, \boldsymbol{\sigma}_\infty) \;\le\; \frac{n}{q}\Big(1 - \frac{e^{-1/\beta^*}}{n}\Big)^t \;\le\; \frac{n}{q}\exp\Big(-\frac{te^{-1/\beta^*}}{n}\Big) \;\le\; \frac{\delta_0}{q},$$

when $t \ge e^{1/\beta^*} n\ln(n/\delta_0)$. Thus, when $t \ge e^{1/\beta^*} n\ln(n/\delta_0)$, we will have $s_c \ge (1 - \varphi/\beta^*)n/q$ $(\forall c \in C)$ with probability at least $1 - \delta_0$.

We will run this process in $k$ stages. We take $\varphi_0 = 1$, assuming only $s_c \ge 0$ $(\forall c \in C)$. For $t \ge e^{1/\beta^*} n\ln(kn/\delta_0)$ we will have $s_c \ge (1 - \varphi_0/\beta^*)n/q$ with probability $1 - \delta_0/k$. Thus we can take $\varphi_1 = \varphi_0/\beta^*$ and repeat the process. After time $t \ge ke^{1/\beta^*} n\ln(kn/\delta_0)$ we will have

$$s_c \;\ge\; \Big(1 - \frac{\varphi_0}{(\beta^*)^k}\Big)\frac{n}{q} \;=\; \Big(1 - \frac{1}{(\beta^*)^k}\Big)\frac{n}{q} \qquad (\forall c \in C)$$

with probability $1 - k(\delta_0/k) = 1 - \delta_0$. So, in time $t \ge e^{1/\beta^*}\ln(\beta^*/\epsilon_0)n\ln\big(\ln(\beta^*/\epsilon_0)n/\delta_0\big) \approx e^{1/\beta}\ln(\beta/\epsilon_0)n\ln\big(\ln(\beta/\epsilon_0)n/\delta_0\big)$ we will have $s_c \ge (1-\epsilon_0)n/q$ $(\forall c \in C)$ with probability at least $1 - \delta_0$. ☐

**Corollary 1.** *For all* $\mathfrak{t}_1 \le t \le \mathfrak{t}_0$ *and* $v \in V$*, we have*

$$\beta e^{-1/\beta}\Delta \;\lesssim\; A_v(X_t) \;\lesssim\; \beta e^{-(1-\epsilon_0)/\beta}\Delta, \tag{15}$$

*with probability at least* $1 - \delta_0$*.*

*Proof.* The left hand inequality follows from Lemma 7. The right hand inequality follows from the proof of Lemma 9, in particular (11). ☐

We note here that the better bound on the number of available colours $A_v(X_t)$ from Corollary 1 can be used to improve the bound on $\beta(p)$ for $p = \Theta(1)$ using (3)–(5). The results are plotted in Figure 2.

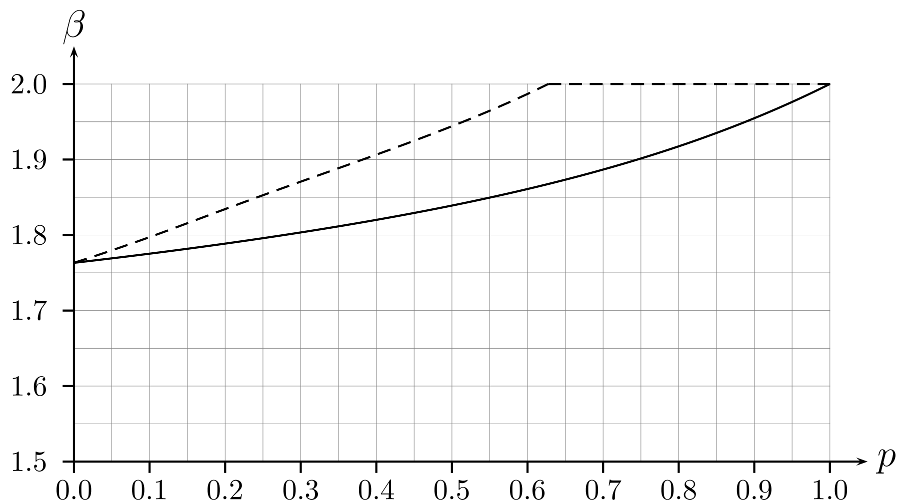


Figure 2: Bound on $\beta$ for convergence of Glauber using Corollary 1, also showing the bound from Lemma 8 for comparison.

We can now verify the upper bound on $d_v^*$ in this lemma (see (6) for the definition).

Let $v \in V$, $c_X = X_t(v)$ and $c_Y = Y_t(v)$. Also, let $S_X = S_{c_Y}(X_t)$, $S_Y = S_{c_X}(Y_t)$ with $s_X = |S_X|$, $s_Y = |S_Y|$. From Lemma 3 and Lemma 9 with $\delta_0 = n^{-10}$, we know that **whp** $(1-\epsilon_0)n/q \le s_X, s_Y \le \epsilon^2 n/\ln n$ when $t \ge \mathfrak{t}_1$. ¿From Lemma 5, we have exposed at most $\epsilon n s_X$ pairs adjacent to vertices in $S_X$, $\epsilon n s_Y$ pairs adjacent to vertices in $S_Y$, and $\epsilon n$ pairs adjacent to $v$. Let $V' = V \setminus (S_X \cup \{v\})$. Then, by simple counting, there can be at most $\epsilon n + \sqrt{\epsilon} n + \sqrt{\epsilon} n = o(n)$ vertices $w \in V'$ such that either the pair $\{v, w\}$ is exposed, or there are more than $\sqrt{\epsilon} s_X = o(n/q)$ exposed pairs $\{u, w\}$ $(u \in S_X)$ or more than $\sqrt{\epsilon} s_Y = o(n/q)$ exposed pairs $\{u, w\}$ $(u \in S_Y)$. Thus there is a set $V'' \subseteq V'$, of size $\approx n$, such that for $w \in V''$, $\{v, w\}$ is unexposed and there are $\approx s_X$ unexposed pairs $\{u, w\}$ $(u \in S_X)$ and $\approx s_Y$ unexposed pairs $\{u, w\}$ $(u \in S_Y)$. Hence, for $w \in V''$,

$$\begin{aligned} \Pr(w \in \mathcal{N}_v^*) &\lesssim p\big((1-p)^{s_X} + (1-p)^{s_Y} - (1-p)^{s_X + s_Y - |S_X \cap S_Y|}\big) && (16) \\ &\le p\big((1-p)^{s_X} + (1-p)^{s_Y} - (1-p)^{s_X + s_Y}\big) \\ &= p\big(1 - (1-(1-p)^{s_X})(1-(1-p)^{s_Y})\big) && (17) \\ &\le p\big(1 - (1-(1-p)^{(1-\epsilon_0)n/q})^2\big) \\ &\lesssim p\big(1 - (1-e^{-(1-\epsilon_0)/\beta})^2\big). \end{aligned}$$

Moreover, these bounding events are independent for all $w \in V''$. Let $\gamma = 1 - (1 - e^{-(1-\epsilon_0)/\beta})^2$. Then, using the Chernoff bound,

$$\Pr\left(d_v^* > (1+\epsilon)\gamma\Delta\right) \le \exp(-\Omega(\epsilon^2 \Delta)) = n^{-\theta}. \qquad (18)$$

This completes the proof of Lemma 2 and Theorem 1 for this case.

### 3.4.2 The case $p = o(1)$, $p = \Omega(1/\log n)$

**Lemma 10.** *Let $0 < \epsilon_0, \delta_0 \le 1$ be constants. Then, with probability at least $1 - \delta_0$, we have that for all $\mathsf{t}_1 \le t \le \mathsf{t}_0$, $s_c \ge (1 - \epsilon_0)n/q$ for almost all $c \in C$, where*

$$\mathsf{t}_1 \;=\; e^{1/\beta} \ln(\beta/\epsilon_0) n \ln\big(\ln(\beta/\epsilon_0)n/\delta_0\big).$$

*Proof.* When $p = \Omega(1/\log n)$, the analysis of Section 3.4.1 fails, even for $p = o(1)$. The equilibrium distributions of individual $s_c(\mathsf{X}_t)$ are not sufficiently concentrated even to prove Corollary 1 directly. However, conditional on the $s_c$ $(c \in C)$, we may still prove concentration of $\mathcal{A}_v(X_t)$ as in Lemma 7. We may then prove Corollary 1 as follows.

We combine the $q = \Omega(n/\log n)$ colours into $\ell \approx q/r$ groups, $C_1, C_2, \ldots, C_\ell$, each of size $r = \lceil \omega \ln n \rceil$ or $r + 1$. Consider a particular group $C_j$. For $c \in C_j$, we may bound $s_c$ by a random process $\boldsymbol{\sigma}_t$, as in Section 3.4.1. Moreover, the processes are close to being independent for all $c \in C_j$. This follows from the facts that $q = \Omega(n/\log n)$, and $s_c \le \epsilon^4 \ln^2 n$, from Lemma 3, so $\sum_{c \in C_j} s_c = o(\log^3 n)$ **whp**. Using this, we can bound the transition probabilities of each $s_c$ $(c \in C_j)$ conditionally on the rest, and show independence to a sufficiently close approximation. More precisely we can replace (10) and (12) by

$$\Pr\big(s_c(\mathsf{X}_{t+1}) = s + 1 \mid s_c(\mathsf{X}_t) = s, s_{c'}, c' \in C_j \setminus \{c\}\big) \;\gtrsim\; e^{-1/\beta}/q \qquad (0 \le s < \nu). \tag{19}$$

$$\Pr\big(s_c(\mathsf{X}_{t+1}) = s - 1 \mid s_c(X_t) = s, s_{c'}, c' \in C_j \setminus \{c\}\big) \;\lesssim\; \frac{s e^{-(1-\varphi)/\beta}}{n} \qquad (1 \le s \le \nu) \tag{20}$$

Then the argument for a given $s_c$ follows the proof of Lemma 9, using the process defined there, so we will simply indicate the differences.

The value of $\rho$ is not large enough to get the RHS of (13). Instead, all we can say is that for $x \sim \text{Poiss}(\rho)$,

$$\Pr\big(x < (1 - \varphi)\rho\big) \;\le\; \epsilon_\varphi = \exp(-\tfrac{1}{2}\varphi^2\rho) \;= o(1).$$

So we weaken the hypothesis $s_c(\mathsf{X}_t) \ge (1 - \varphi)\nu$ to

$$\Pr\big(s_c(\mathsf{X}_t) < (1 - \varphi)\nu\big) < \epsilon_\varphi \;(\forall c \in C,\; 0 \le t \le \mathsf{t}_0).$$

(This is enough to prove the lemma. The value of $\varphi$ tends to zero as we iterate.)

Note that then

$$\mathbb{E}\big[(1 - p)^{s_c(\mathsf{X}_t)}\big] \;\le\; (1 - p)^{(1-\varphi)\nu} + \epsilon_\varphi \;\le\; e^{-(1-\varphi)/\beta} + \epsilon_\varphi \;\approx\; e^{-(1-\varphi)/\beta},$$

and hence

$$\mathbb{E}[\mathcal{A}_v(X_t)] \lesssim q e^{-(1-\varphi)/\beta} \;\approx\; \ell r e^{-(1-\varphi)/\beta}.$$

However, to show that (11) remains true, we must prove concentration. Now $0 \le (1 - p)^{s_c} \le 1$, and $\mathbb{E}[(1 - p)^{s_c}] \lesssim e^{-(1-\varphi)/\beta}$, so we may use Hoeffding's inequality [8] to show that

$$\Pr\Big(\textstyle\sum_{c \in C_j}(1 - p)^{s_c} \ge (1 + \epsilon) r e^{-(1-\varphi)/\beta}\Big) \;\lesssim\; \exp\big(-\Omega(\epsilon^2 \omega \log n)\big) \;=\; n^{-\theta},$$

and hence

$$\Pr\Big(\textstyle\sum_{c \in C}(1 - p)^{s_c} \ge (1 + \epsilon) q e^{-(1-\varphi)/\beta}\Big) \;\le\; n^{-\theta}.$$

This shows that the required upper bound (11) for $\mathcal{A}_v(X_t)$ holds **whp**. The remainder of the proof now follows that of Lemma 9, except that (14) can be proved to hold only with probability $1-\epsilon_\varphi$. However, this is the inductive hypothesis, so we may conclude that Corollary 1 holds for all $p = o(1)$. □

When $p = \Omega(1/\log n)$, the bound (18) is not valid. However, observe that we only require the bound on $d_v^*$ to hold in expectation. We must deal with conditioning on $d_v^*$, but only by the event $v \in \mathcal{W}$, not by the whole of $\mathcal{W}$. Therefore, we need to estimate $\mathbb{E}[d_v^* \mid v \in \mathcal{W}]$ at a given time $t^*$. To do this, we must bound the random variable $s_c(\mathsf{X}_{t^*})$ conditional on $\mathsf{X}_{t^*}(v) = c$ and $v \in \mathcal{W}$. Since we cannot ignore the conditioning here, the approach must be different from that of Section 3.4.1.

If $v \in \mathcal{W}$, then at time $t_v < t^*$, the last successful update of $v$ occurred either in $X_{t_v}$ or $Y_{t_v}$. Thus $u = v$, $c_X = X_{t_v}(v)$, $c_Y = Y_{t_v}(v)$ and $c_X \neq c_Y$. We have $X_t(v) = c_X$, $Y_t(v) = c_Y$ for all $t_v < t \leq t^*$. We wish to bound $s_X = s_{c_Y}(\mathsf{X})$ and $s_Y = s_{c_X}(\mathsf{Y})$ from below (see (16)). But the evolution of $s_X$ during $(t_v, t^*)$ is independent of $X_t(v)$, since $X_t(v) = c_X$ throughout. Similarly the evolution of $s_Y$ during $(t_v, t)$ is independent of $Y_t(v)$.

We assume that $t \geq \mathfrak{t}_1$, so the process has been running long enough that the inequality in Corollary 1 is true, with the claimed probability. Consequently $A_v(\mathsf{X}_t)$ and $A_v(\mathsf{Y}_t)$ are both close to $\beta e^{-1/\beta}\Delta$ for all $v$ and $t \leq \mathfrak{t}_0$. The construction of Couple implies that the probability that an update of $v$ is successful in $\mathsf{X}$ but unsuccessful in $\mathsf{Y}$ is at most $3\epsilon_0$ say. For a given $v \in \mathcal{W}$, let $t_v$ be the last epoch before $t^*$ at which a successful update of $v$ occurred in $\mathsf{X}$, and let $T = t^* - t_v$. Let $A_L, A_R$ denote the quantities on LHS and RHS of (15). Let $\rho_L = n/A_L$ and $\rho_R = n/A_R$. Then the (conditional) probability of a successful update at $v$ is at least $\frac{A_L}{nq} = \frac{1}{\rho_L q}$ and so the distribution of $T$ satisfies

$$\Pr(T \geq t) \;\leq\; \left(1 - \frac{1}{\rho_L q}\right)^t \quad (t = 0, 1, \ldots, t^*) \tag{21}$$

Let $b = \beta/(\beta-1)$ and assume $t^* \geq \mathfrak{t}_2 = \mathfrak{t}_1 + 4bn\ln n$. Now $A_L \geq (q - \Delta)$ and so $1/\rho_L q \geq 1/bn$. So $\Pr(T > 4bn\ln n) \leq (1 - 1/bn)^{4bn\ln n} < n^{-4}$. Thus we can take $t_v \geq \mathfrak{t}_1$ for all $v \in V$ and $\mathfrak{t}_2 \leq T \leq \mathfrak{t}_0$ with probability $1 - o(1/n)$.

**Analysis of a bounding process**

We now consider two copies of a bounding process that will be dominated by the processes $s_X, s_Y$. The processes can be assumed independent, as discussed in (19), (20).

We consider the process

$$\Pr(\boldsymbol{\sigma}_{t+1} = \sigma + 1 \mid \boldsymbol{\sigma}_t = \sigma) \;=\; \frac{(1-p)^\sigma}{q} \qquad (\sigma = 0, 1, 2, \ldots) \tag{22}$$

$$\Pr(\boldsymbol{\sigma}_{t+1} = \sigma - 1 \mid \boldsymbol{\sigma}_t = \sigma) \;=\; \frac{\sigma}{n}\frac{\tilde{A}}{q} = \frac{\sigma}{q\rho} \qquad (\sigma = 0, 1, 2, \ldots), \tag{23}$$

where $\tilde{A}$ is a uniform upper bound on $A_v(\mathsf{X}_t)$ $(v \in V, \mathfrak{t}_1 \leq t \leq \mathfrak{t}_0)$, and $\rho = n/\tilde{A}$. From Lemma 7, we know that $\tilde{A} \gtrsim \beta np(1-p)^{1/\beta p}$, so $\rho = O(1)$ when $p = \Theta(1)$.

Thus, if we assume $s_X = s_Y = 0$ at time $t_v$, we can bound $s_X$ and $s_Y$ at time $t^*$ by two independent processes of the form (22)–(23) during $(t_v, t^*)$. In particular we wish to examine the transient behaviour of $\boldsymbol{\sigma}_t$, with this distribution and $\rho = ne^{1/\beta}/q$.

(Of course, we can only assume a $\rho$ which fluctuates close to this value, but simple bounding arguments justify the use of this particular value). In particular, we wish to determine certain properties of $\boldsymbol{\sigma}_{\mathsf{T}}$ at a random time $\mathsf{T}$ such that

$$\Pr(\mathsf{T} \geq t) \;=\; \Big(1 - \frac{1}{\rho q}\Big)^t \qquad (t = 0, 1, 2, \ldots).$$

Now

$$\mathbb{E}[\boldsymbol{\sigma}_{t+1}] \;=\; \sigma_t + \frac{(1-p)^{\sigma_t}}{q} - \frac{\sigma_t}{q\rho}.$$

Putting $\psi_t = p\sigma_t$ we may re-write this as

$$\frac{\mathbb{E}[\psi_{t+1} - \psi_t]}{p} = \frac{e^{-\gamma\psi_t}}{\beta np} - \frac{\psi_t e^{-1/\beta}}{\beta np}.$$

Here $e^{-\gamma} \approx (1-p)^{1/p} \approx 1/e$ and the error in putting $q = np$ will be absorbed into $\gamma$ which is close to 1.

Let $x = \beta\psi_t$ and $z = e^{-1/\beta}t/n$ so that

$$\frac{\mathbb{E}[x(z + e^{-1/\beta}/n) - x(z)]}{\beta p} \;=\; \frac{e^{-\gamma x(z)/\beta}}{\beta np} - \frac{x(z)e^{-1/\beta}}{\beta np}. \tag{24}$$

Using the proof idea of Theorem 5.1 of Wormald [15] we see that (24) is closely approximated by the differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}z} \;=\; e^{(1-x)/\beta} - x, \quad \text{so} \quad z(x) = \int_0^x \frac{\mathrm{d}y}{e^{(1-y)/\beta} - y} \quad (0 \leq x < 1).$$

The upper bound $x < 1$ is due to the fact that $\mathrm{d}x/\mathrm{d}z = 0$ when $x = 1$, the lower bound $x \geq 0$ to the assumption $\sigma_0 = 0$.

Let $\boldsymbol{\sigma}_t$, $\boldsymbol{\sigma}'_t$ be iid processes with the distribution (22)–(23). We wish to determine (see (17))

$$\mathbb{E}\big[\big(1 - (1-p)^{\boldsymbol{\sigma}_t}\big)\big(1 - (1-p)^{\boldsymbol{\sigma}'_t}\big)\big] \approx \mathbb{E}\big[1 - e^{-p\boldsymbol{\sigma}_t}\big]\mathbb{E}\big[1 - e^{-p\boldsymbol{\sigma}'_t}\big] = \mathbb{E}\big[1 - e^{-p\boldsymbol{\sigma}_t}\big]^2 \approx (1 - e^{-x(z)/\beta})^2$$

at the random epoch $t = T$, where $T$ has the distribution $\Pr(T \geq t) \approx (1 - e^{-1/\beta}/n)^t \approx e^{-z}$. Hence, using parts integration,

$$\begin{aligned}
\mathbb{E}\big[(1 - (1-p)^{\boldsymbol{\sigma}_{\mathsf{T}}})^2\big] \;&=\; \int_0^\infty (1 - e^{-x(z)/\beta})^2 e^{-z}\,\mathrm{d}z \\
&= \int_0^1 (1 - e^{-x/\beta})^2 e^{-z(x)} \frac{\mathrm{d}z}{\mathrm{d}x}\mathrm{d}x \\
&= \Big[-(1 - e^{-x/\beta})^2 e^{-z(x)}\Big]_0^1 + \frac{2}{\beta}\int_0^1 (1 - e^{-x/\beta})e^{-z(x)}\,\mathrm{d}x \\
&= \frac{2}{\beta}\int_0^1 (1 - e^{-x/\beta})\exp\Big(-\frac{x}{\beta} - \int_0^x \frac{\mathrm{d}y}{e^{(1-y)/\beta} - y}\Big)\mathrm{d}x.
\end{aligned} \tag{25}$$

Thus we can write

$$\Pr(w \in \mathcal{N}_v^* \mid v \in \mathcal{W}) \leq \kappa_1(\beta) + \delta_0 + 3\epsilon_0.$$

This completes our upper estimate $\mathbb{E}[d_v^* \mid v \in \mathcal{W}]$ required for Lemma 2 and so completes the proof of Theorem 1 for this case.

### 3.4.3 The case $p = \Theta(1)$

By detailed balance, the equilibrium distribution $\boldsymbol{\sigma}_\infty$ of (22)–(23) is given by

$$\Pr(\boldsymbol{\sigma}_\infty = \sigma) = \frac{\rho^\sigma (1-p)^{\sigma(\sigma-1)/2}}{G(\rho)\sigma!} \qquad (\sigma = 0, 1, 2, \ldots), \tag{26}$$

where $G(\rho) = \sum_{\sigma=0}^{\infty} \rho^\sigma (1-p)^{\sigma(\sigma-1)/2}/\sigma!$, since

$$\frac{(1-p)^{\sigma-1}}{q}\,\frac{\rho^{\sigma-1}(1-p)^{(\sigma-1)(\sigma-2)/2}}{(\sigma-1)!} \;=\; \frac{\sigma}{\rho q}\,\frac{\rho^\sigma(1-p)^{\sigma(\sigma-1)/2}}{\sigma!} \qquad (\sigma = 1,2,\ldots).$$

We will denote a random variable with the distribution given in (26) by $\boldsymbol{\sigma}_\infty(\rho)$.

We can bound the mixing time of the process $\boldsymbol{\sigma}_t$ using path coupling. We define the coupling on adjacent states $x_t = \sigma - 1$, $y_t = \sigma$ $(\sigma > 0)$ as follows

$$\Pr(x_{t+1} = x,\, y_{t+1} = y) \;=\; \begin{cases} \dfrac{\sigma-1}{\rho q}, & \text{if} \quad x = \sigma-2, \quad y = \sigma-1, \\ \dfrac{(1-p)^\sigma}{q}, & \text{if} \quad x = \sigma, \qquad y = \sigma+1, \\ \dfrac{1}{\rho q}, & \text{if} \quad x = \sigma-1, \quad y = \sigma-1, \\ \dfrac{p(1-p)^{\sigma-1}}{q}, & \text{if} \quad x = \sigma, \qquad y = \sigma, \\ 1 - \dfrac{\sigma}{\rho q} - \dfrac{(1-p)^{\sigma-1}}{q} & \text{if} \quad x = \sigma-1, \quad y = \sigma. \end{cases}$$

It is now easy to see that

$$\mathbb{E}[\mathrm{d}(x_{t+1}, y_{t+1}) \mid \mathrm{d}(x_t, y_t) = 1] \;\le\; 1 - \frac{1}{\rho q} - \frac{p(1-p)^{x_t}}{q} \;<\; 1 - \frac{1}{\rho q}. \tag{27}$$

Since $\rho = O(1)$ and $q = O(n)$ in (27), convergence occurs in $O(n \log n)$ time.

We may now use the proof method from Section 3.4.2 to show that, for all $\mathfrak{t}_1 \le t \le \mathfrak{t}_0$ and $v \in V$, **whp** we have

$$A_v(\mathsf{X}_t) \;\lesssim\; \sum_{c \in C} \mathbb{E}\big[(1-p)^{s_c(\mathsf{X}_t)}\big] \;\approx\; q\mathbb{E}\big[(1-p)^{s_c(\mathsf{X}_t)}\big],$$

by symmetry. Hence we may take $\tilde{A}_0 = n/\rho_0$, $\rho_0 = n/q$, corresponding to $A_v(\mathsf{X}_t) \le q$, and then we will obtain **whp**

$$A_v(\mathsf{X}_t) \lesssim \tilde{A}_1 \;\approx\; q\mathbb{E}\big[(1-p)^{\boldsymbol{\sigma}_\infty(\rho_0)}\big].$$

As in Section 3.4.1, we may run this process iteratively. The iteration is analysed below in Section **??** and is shown to converge in $O(1)$ steps so that **whp**

$$A_v(\mathsf{X}_t) \lesssim A^* \;=\; q\mathbb{E}\big[(1-p)^{\boldsymbol{\sigma}_\infty(\rho^*)}\big] \;=\; \frac{q\mathbb{E}[\boldsymbol{\sigma}_\infty(\rho^*)]}{\rho^*} \;=\; \frac{n}{\rho^*} \;\approx\; \frac{\Delta}{p\rho^*},$$

where $\rho^*$ is the solution to

$$\mathbb{E}[\boldsymbol{\sigma}_\infty(\rho^*)] = \frac{n}{q} \approx \frac{1}{\beta p} \; \mathit{or} \; \beta \approx \frac{1}{p\mathbb{E}[\boldsymbol{\sigma}_\infty(\rho^*)}]. \tag{28}$$

The remaining analysis to get an upper bound on $\mathbb{E}[d_v^* \mid v \in \mathcal{W}]$ is similar to Section 3.4.2, except that we cannot use concentration of measure to approximate as we did there. We let

$$\kappa_2 \;=\; \mathbb{E}\big[(1 - (1-p)^{\boldsymbol{\sigma}_T})^2\big], \tag{29}$$

where $T$ has the geometric distribution with mean $\rho q$ (see (21)), and the process $\boldsymbol{\sigma}_t$ is governed by (22)–(23), with $\rho$ and the initial condition $\sigma_0 = 0$. This distribution depends on $p$ and $\rho$. We look for a solution to

$$(1-\kappa_2)\Delta \;\approx\; (1-\kappa_2)np \;\approx\; \hat{A} \approx \frac{n}{\rho}, \qquad \text{i.e. } 1-\kappa_2 \;=\; \frac{1}{\rho p}. \tag{30}$$

To determine the expectation in (29), we set $\tau = t/n$. Then, for large $n$, (21) becomes

$$\Pr(\tau \geq z) \;=\; e^{-\tau/\beta p\rho},$$

an exponential distribution with mean $\beta p\rho$. Thus $\kappa_2$ is a function of $\beta$, $p$ and $\rho$. However, we may use (28) to substitute for $\beta$ in (29), giving $\kappa_2$ as a function of $\rho$ and $p$. We may then substitute for $\kappa_2$ in (30), giving $\rho$ only as a function of $p$. Hence we can determine $\beta$ as a function of $p$. The analysis is given in Section 3.4.4 and the results are shown in Figure 3.

Mysteriously, the exact value obtained for $p = 1$ (i.e. the complete graph $K_n$) is $\beta = 11/6$, the same value obtained in Vigoda's analysis [14]. We do not know if this is mere coincidence, but we note that for $K_n$ a better analysis is is possible using path coupling. We may show $O(n \log n)$ convergence if $\beta = 1+\varepsilon$ for any $\varepsilon > 0$. We will sketch the proof here.

We start, as usual in path coupling, with colourings $X, Y$ which disagree at one vertex, which is red in $X$ and blue in $Y$, say. If coupling does not occur, we may easily create a second disagreement, which will be coloured blue in $X$ and red in $Y$. But then all other vertices have both red and blue in their neighbourhood in both $X$ and $Y$, so at subsequent steps no further disagreements are created. With probability $\Omega(1/n)$ we will destroy a disagreement, since $\beta = 1+\varepsilon$. It is then easy to show that the probability that coupling has not occurred in $O(n \log n)$ time is $O(1/n^2)$, say. Since there were at most $n$ disagreements at the outset, the result follows.

**Analysis of a bounding process**

Let $G(\rho) = \sum_{\sigma=0}^{\infty} \rho^{\sigma}(1-p)^{\sigma(\sigma-1)/2}/\sigma!$. Note that $1+\rho \leq G(\rho) \leq e^{\rho}$. We consider a random variable $\boldsymbol{\sigma} = \boldsymbol{\sigma}_\infty$ such that

$$\Pr(\boldsymbol{\sigma} = \sigma) = \frac{\rho^{\sigma}(1-p)^{\sigma(\sigma-1)/2}}{G(\rho)\sigma!} \qquad (\sigma = 0, 1, 2, \ldots).$$

We consider $\rho$ to be a parameter, so we will write, for example, $\mathbb{E}[\boldsymbol{\sigma}(\rho)]$ if we wish to specify the parameter. If the parameter is not specified, its value is $\rho$. Thus $\mathbb{E}[\boldsymbol{\sigma}] = \mathbb{E}[\boldsymbol{\sigma}(\rho)]$. We will also write $\rho = e^z$ and $g(z) = G(e^z)$. Let us write $\bar{p} = (1-p)$. Then easy calculations show that

$$G'(\rho) \;=\; G(\rho)\mathbb{E}[\boldsymbol{\sigma}]/\rho = G((1-p)\rho) \;=\; G(\rho)\mathbb{E}[(1-p)^{\boldsymbol{\sigma}}], \tag{31}$$

$$G''(\rho) \;=\; G(\rho)\mathbb{E}[\boldsymbol{\sigma}(\boldsymbol{\sigma}-1)]/\rho^2 \;= (1-p)G'\big((1-p)\rho\big), \tag{32}$$

$$g'(z) \;=\; g(z)\mathbb{E}[\boldsymbol{\sigma}], \qquad g''(z) \;=\; g(z)\mathbb{E}[\boldsymbol{\sigma}^2]. \tag{33}$$

We first show that $\mathbb{E}[\boldsymbol{\sigma}]$ is a strictly increasing function of $z$.

$$\frac{\mathrm{d}\mathbb{E}[\boldsymbol{\sigma}]}{\mathrm{d}z} \;=\; \frac{\mathrm{d}}{\mathrm{d}z}\Big(\frac{g'(z)}{g(z)}\Big) \;=\; \frac{g(z)g''(z) - g'(z)^2}{g(z)^2} \;= \mathbb{E}[\boldsymbol{\sigma}^2] - \mathbb{E}[\boldsymbol{\sigma}]^2 \;=\; \mathbb{V}\mathrm{ar}[\boldsymbol{\sigma}] \;>\; 0. \tag{34}$$

We wish to analyse the iteration

$$\rho_0 \;=\; \frac{n}{q}, \qquad \rho_{i+1} \;=\; \frac{n}{q\mathbb{E}[(1-p)^{\boldsymbol{\sigma}(\rho_i)}]} \;=\; \frac{n\rho_i}{q\mathbb{E}[\boldsymbol{\sigma}(\rho_i)]} \;=\; \frac{ne^{z_i}g(z_i)}{qg'(z_i)} \quad (i \geq 0),$$

using (31). Taking logarithms, and using (33), we consider instead the equivalent iteration,

$$z_0 \;=\; \ln(n/q), \qquad z_{i+1} \;=\; f(z_i) \;=\; z_i + \ln(n/q) + \ln g(z_i) - \ln g'(z_i) \quad (i \geq 0). \tag{35}$$

Now, using (31)–(34), we have

$$f'(z) \;=\; 1+\frac{g'(z)}{g(z)}-\frac{g''(z)}{g'(z)} \;=\; 1+\mathbb{E}[\boldsymbol{\sigma}]-\frac{\mathbb{E}[\boldsymbol{\sigma}^2]}{\mathbb{E}[\boldsymbol{\sigma}]} \;=\; 1-\frac{\mathbb{V}\mathrm{ar}[\boldsymbol{\sigma}]}{\mathbb{E}[\boldsymbol{\sigma}]} \tag{36}$$

$$= \mathbb{E}[\boldsymbol{\sigma}]-\frac{\mathbb{E}[\boldsymbol{\sigma}(\boldsymbol{\sigma}-1)]}{\mathbb{E}[\boldsymbol{\sigma}]} \;=\; \rho\frac{G'(\rho)}{G(\rho)}-(1-p)\rho\frac{G'\big((1-p)\rho\big)}{G\big((1-p)\rho\big)}$$

$$= \mathbb{E}\big[\boldsymbol{\sigma}(\rho)\big]-\mathbb{E}\big[\boldsymbol{\sigma}\big((1-p)\rho\big)\big]. \tag{37}$$

From (36) and (37) we see that $0<f'(z)<1$ for all $z\in\mathbb{R}$.

Consider the equation $z=f(z)$. This is equivalent to $\mathbb{E}[\boldsymbol{\sigma}(\rho)]=n/q$ from (33) and (35), so it has at most one root because $\mathbb{E}[\boldsymbol{\sigma}]$ is strictly increasing. It is also easy to show that it has a root, since (31) and Jensen's inequality imply

$$\rho(1-p)^{\mathbb{E}[\boldsymbol{\sigma}]} \;\le\; \rho\mathbb{E}[(1-p)^{\boldsymbol{\sigma}}] \;=\; \rho\frac{G'(\rho)}{G(\rho)} \;=\; \mathbb{E}[\boldsymbol{\sigma}] \;\le\; \rho \qquad (\textit{using } G'(\rho)\le G(\rho) \;\; -- \;(31))$$

From the right inequality, we have $\mathbb{E}[\boldsymbol{\sigma}(\rho_0)] \le \rho_0 = n/q$. From the left, we see that $\mathbb{E}[\boldsymbol{\sigma}] > (1-p)\ln\rho$. Otherwise, since $x\ge -(1-x)\ln(1-x)$ $(0\le x\le 1)$ and $x>\ln x$ $(x>0)$, we would have

$$\mathbb{E}[\boldsymbol{\sigma}] \;\ge\; \rho(1-p)^{(1-p)\ln\rho} \;=\; \rho^{1+(1-p)\ln(1-p)} \;\ge\; \rho^{1-p} \;>\; \ln\rho^{1-p} \;=\; (1-p)\ln\rho, \tag{38}$$

a contradiction. Hence $\mathbb{E}[\boldsymbol{\sigma}(\rho)]>n/q$ for large enough $\rho$.

We may use (36) and (38) to give an explicit upper bound on $f'(z)$, as follows.

$$1-f'(z) \;=\; \frac{\mathbb{V}\mathrm{ar}[\boldsymbol{\sigma}]}{\mathbb{E}[\boldsymbol{\sigma}]} \;\ge\; \frac{\mathbb{E}[\boldsymbol{\sigma}]^2\Pr(\boldsymbol{\sigma}=0)}{\mathbb{E}[\boldsymbol{\sigma}]} \;=\; \frac{\mathbb{E}[\boldsymbol{\sigma}]}{G(\rho)} \;\ge\; (1-p)e^{-\rho}\ln\rho.$$

Let $z^*$ be the root of $z=f(z)$, and $\rho^*=e^{z^*}$. Now we have

$$z_0=\ln(n/q)=\ln\mathbb{E}[\boldsymbol{\sigma}(\rho^*)]\le\ln\rho^*=z^*$$

and

$$z^* \;=\; \ln\rho^* \;\le\; \frac{\mathbb{E}[\boldsymbol{\sigma}(\rho^*)]}{1-p} \;=\; \frac{n}{q(1-p)} \;\approx\; \frac{1}{\beta p(1-p)}$$

If $z_i<z^*$ then, for some $\hat{z}_i$ with $z_i\le\hat{z}_i\le z^*$,

$$z_{i+1} \;=\; f(z_i) \;=\; f\big(z^*+(z_i-z^*)\big) \;=\; f(z^*)+(z_i-z^*)f'(\hat{z}_i) \;=\; z^*-(z^*-z_i)f'(\hat{z}_i) \;<\; z^*,$$

Thus $\{z_i\}$ is a bounded increasing sequence, so convergent. More precisely, we have

$$f'(z_i) \;\lesssim\; 1-(1-p)e^{-\hat{\rho}}\ln\hat{\rho} \;=\; 1-\varphi,$$

where $\hat{\rho}=\exp\left(\frac{1}{\beta p(1-p)}\right)$ and $\varphi=(1-p)e^{-\hat{\rho}}\ln\hat{\rho}$. Therefore

$$z^*-z_{i+1} \;=\; (z^*-z_i)f'(\hat{z}_i) \;<\; (z^*-z_i)(1-\varphi) \;\le\; (z^*-z_0)(1-\varphi)^i \;\lesssim\; \frac{e^{-i\varphi}}{\beta p(1-p)}.$$

Thus we have $z^*-\epsilon_0<z_i<z^*$ when $i\ge i^*=\big\lceil\varphi\ln\big(\beta p(1-p)\epsilon_0\big)\big\rceil=O(1)$.

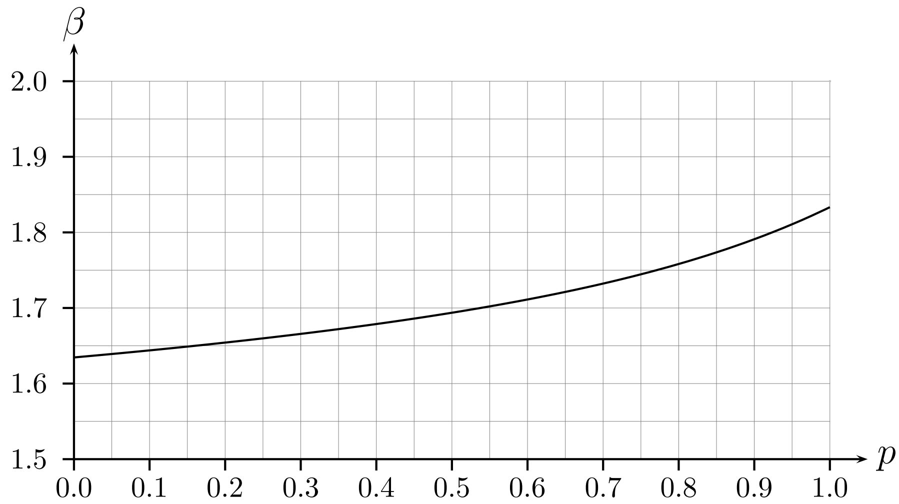


Figure 3: Bound on $\beta$ for convergence of GLAUBER from Section 3.4.4.

#### 3.4.4 Final Analysis:

Let $H_\sigma(t) = \Pr(\boldsymbol{\sigma}_t = \sigma)$. Note that $H_0(0) = \delta_\sigma$, where $\delta_0 = 1$ and $\delta_\sigma = 0$ for all $\sigma > 0$. Then, using (22)–(23),

$$H_\sigma(t+1) \;=\; \frac{(1-p)^{\sigma-1}}{q} H_{\sigma-1}(t) + \left(1 - \frac{(1-p)^\sigma}{q} - \frac{\sigma}{\rho q}\right) H_\sigma(t) + \frac{\sigma+1}{\rho q} H_{\sigma+1}(t),$$

where $\rho$ is the solution to $\mathbb{E}[\boldsymbol{\sigma}(\rho)] = 1/\beta p$, using the distribution of (26). Thus

$$\rho q\big(H_\sigma(t+1) - H_\sigma(t)\big) \;=\; \rho(1-p)^{\sigma-1} H_{\sigma-1}(t) - \big(\rho(1-p)^\sigma + \sigma\big) H_\sigma(t) + (\sigma+1) H_{\sigma+1}(t).$$

Let $\tau = t/n$. Then, as $n \to \infty$, this becomes

$$\beta p \rho H'_\sigma(\tau) \;=\; \rho(1-p)^{\sigma-1} H_{\sigma-1}(\tau) - \big(\rho(1-p)^\sigma + \sigma\big) H_\sigma(\tau) + (\sigma+1) H_{\sigma+1}(\tau), \tag{39}$$

with the initial conditions $H_0(0) = 1$, $H_0(\tau) = 0$ $(\tau > 0)$. Differentiating this gives

$$\beta p \rho H''_\sigma(\tau) \;=\; \rho(1-p)^{\sigma-1} H'_{\sigma-1}(\tau) - \big(\rho(1-p)^\sigma + \sigma\big) H'_\sigma(\tau) + (\sigma+1) H'_{\sigma+1}(\tau), \tag{40}$$

with the initial conditions $H'_0(0) = -\rho$, $H'_0(1) = \rho$, $H'_0(\tau) = 0$ $(\tau > 0)$. We used (39) and (40) as the basis of a second order method for approximating $H_\sigma(\tau)$ at sufficiently many values of $\tau$. Hence we could estimate $F(\tau) = \mathbb{E}[(1-p)^{\sigma_\tau}]$.

Now $\Pr(\mathsf{T} \geq n\tau) \approx (1 - 1/\rho q)^{n\tau} \approx e^{-\eta\tau}$, where $1/\eta = \beta\rho p$ so we could estimate

$$\kappa_2 = \int_0^\infty (1 - F(\tau))^2 e^{-\eta\tau} \eta d\tau, \tag{41}$$

and hence solve (30) for $\rho$. Then $\beta$ could be calculated from (28). This was used to obtain $\beta(p)$ to five decimal places for all values of $p$ from 0 to 1 in steps of 0.025. The results are plotted in Figure 3.

## References


[1] M. Bordewich and M. Dyer, Path coupling without contraction, *Journal of Discrete Algorithms* **5** (2007), 280–292.

[2] R. Bubley and M. Dyer, Path coupling: a technique for proving rapid mixing in Markov chains, in *Proc. 38$^{th}$ Annual IEEE Symposium on Foundations of Computer Science*, IEEE Computer Society, 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, *Random Structures and Algorithms* **29** (2006), 450–465.

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

[5] A. Frieze and E. Vigoda, A survey on the use of Markov chains to randomly sample colorings, in *Combinatorics, Complexity and Chance, A Tribute to Dominic Welsh*, (G. Grimmett, C. McDiarmid eds.), 2007, pp. 53–71.

[6] G. Hardy, J. Littlewood and G. Pólya, *Inequalities*, 2nd rev. edn., Cambridge University Press, 1988.

[7] T. Hayes and E. Vigoda, Coupling with the stationary distribution and improved sampling for colorings and independent sets, in *Proc. 16$^{th}$ Annual ACM-SIAM Symposium on Discrete Algorithms*, SIAM Press, 2005, pp. 971–979.

[8] W. Hoeffding, Probability inequalities for sums of bounded random variables, *Journal of the American Statistical Association* **58** (1963), 13–30.

[9] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley-Interscience, New York, 2000.

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

[11] M. Jerrum, *Counting, Sampling and Integrating: Algorithms and Complexity*, ETH Zürich Lectures in Mathematics, Birkhäuser, Basel, 2003.

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

[13] E. Mossell and A. Sly, Gibbs rapidly samples colorings of $G(n, d/n)$, to appear.

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

[15] N. Wormald, The differential equation method for random graph processes and greedy algorithms, in *Lectures on Approximation and Randomized Algorithms* (M. Karoński and H.J. Prömel, eds.), PWN, Warsaw, 1999, pp. 73–155.