Electronic Journal of Probability

Vol. 12 (2007), Paper no. 16, pages 481–497.

Journal URL
http://www.math.washington.edu/~ejpecp/

Sampling 3-colourings of regular bipartite graphs

David Galvin
Department of Mathematics, University of Pennsylvania
209 South 33rd Street, Philadelphia PA 19104
dgalvin@math.upenn.edu.*

Abstract

We show that if $\Sigma = (V, E)$ is a regular bipartite graph for which the expansion of subsets of a single parity of V is reasonably good and which satisfies a certain local condition (that the union of the neighbourhoods of adjacent vertices does not contain too many pairwise non-adjacent vertices), and if $\mathcal{M}$ is a Markov chain on the set of proper 3-colourings of Σ which updates the colour of at most $\rho|V|$ vertices at each step and whose stationary distribution is uniform, then for $\rho \approx .22$ and d sufficiently large the convergence to stationarity of $\mathcal{M}$ is (essentially) exponential in $|V|$. In particular, if Σ is the d-dimensional hypercube Q_d (the graph on vertex set $\{0,1\}^d$ in which two strings are adjacent if they differ on exactly one coordinate) then the convergence to stationarity of the well-known Glauber (single-site update) dynamics is exponentially slow in $2^d/(\sqrt{d}\log d)$. A combinatorial corollary of our main result is that in a uniform 3-colouring of Q_d there is an exponentially small probability (in 2^d) that there is a colour i such the proportion of vertices of the even subcube coloured i differs from the proportion of the odd subcube coloured i by at most .22. Our proof combines a conductance argument with combinatorial enumeration methods.

Key words: Mixing time, 3-colouring, Potts model, conductance, Glauber dynamics, discrete hypercube.

AMS 2000 Subject Classification: Primary 05C15, 82B20.

Submitted to EJP on August 8 2006, final version accepted April 5 2007.

---

*This work was begun while the author was a member of the Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540 and was supported in part by NSF grant DMS-0111298.

1 Introduction and statement of the result

Markov chain Monte Carlo algorithms (MCMC's) occur frequently in computer science in algorithms designed to sample from or estimate the size of large combinatorially defined structures; they are also used in statistical physics and the study of networks to help understand the behavior of models of physical systems and networks in equilibrium. In this paper we study a class of natural MCMC's that sample from proper 3-colourings of a regular bipartite graph.

Let $\Sigma = (V, E)$ be a simple, loopless, finite graph on vertex set V and edge set E. (For graph theory basics, see *e.g.* (4), (9).) For a positive integer q write $\mathcal{C}_q = \mathcal{C}_q(\Sigma)$ for the set of proper q-colourings of Σ; that is,

$$\mathcal{C}_q = \{\chi : V(\Sigma) \to \{0, 1, \ldots, q-1\} : xy \in E(\Sigma) \Rightarrow \chi(x) \neq \chi(y)\}.$$

Let $\pi_q = \pi_q(\Sigma)$ be the uniform probability distribution on $\mathcal{C}_q$.

The notion of q-colouring is fundamental in graph theory; see *e.g.* (3, Chapter 5) for a survey. The notion also occurs in statistical physics; the pair $(\mathcal{C}_q, \pi_q)$ is the zero-temperature limit of the q-state antiferromagnetic Potts model (see *e.g.* (27; 28)).

Glauber dynamics for proper q-colourings is the single-site update Markov chain $\mathcal{M}_q = \mathcal{M}_q(\Sigma)$ on state space $\mathcal{C}_q$ with transition probabilities $P_q(\chi_1, \chi_2)$, $\chi_1, \chi_2 \in \mathcal{C}_q$, given by

$$P_q(\chi_1, \chi_2) = \begin{cases} 0 & \text{if } |\{v \in V : \chi_1(v) \neq \chi_2(v)\}| > 1 \\ \frac{1}{|V|}\frac{1}{q} & \text{if } |\{v \in V : \chi_1(v) \neq \chi_2(v)\}| = 1 \\ 1 - \sum_{\chi_1 \neq \chi_2' \in \mathcal{C}_q} P_q(\chi_1, \chi_2') & \text{if } \chi_1 = \chi_2. \end{cases}$$

We may think of $\mathcal{M}_q$ dynamically as follows. From a q colouring χ, choose a vertex v uniformly from V and a colour j uniformly from $\{0, \ldots, q-1\}$. Then define a function $\chi' : V \to \{0, \ldots, q-1\}$ by

$$\chi'(w) = \begin{cases} \chi(w) & \text{if } w \neq v \\ j & \text{if } w = v. \end{cases}$$

Finally, move to χ' if χ' is a proper q-colouring, and stay at χ otherwise. (A variant of Glauber dynamics chooses j uniformly from $\{0, \ldots, q-1\} \setminus \{\chi(w) : w \sim v\}$, ensuring that χ' is always a proper colouring. This changes the transition probabilities, but does not significantly change the qualitative behavior of the chain.)

For all Σ the chain $\mathcal{M}_q$ is aperiodic, but it is not in general irreducible (consider, for example, $\Sigma = K_q$, the complete graph on q vertices), and so not ergodic. In the case when $\mathcal{M}_q$ is ergodic (*e.g.*, when Σ has maximum degree Δ and $q \geq \Delta + 2$; see (15)) it is readily checked that it has (unique) stationary distribution π_q. (One only has to check that $\mathcal{M}_q$ is reversible with respect to π_q; that is, that it satisfies the detailed balance equations $\pi_q(\chi_1)P_q(\chi_1, \chi_2) = \pi_q(\chi_2)P_q(\chi_2, \chi_1)$ for all $\chi_1, \chi_2 \in \mathcal{C}_q$.) A natural and important question to ask about $\mathcal{M}_q$ in this case is how quickly it converges to its stationary distribution. We define the *mixing time* $\tau_{\mathcal{M}_q}$ of $\mathcal{M}_q$ by

$$\tau_{\mathcal{M}_q} = \min\left\{t \ : \ d_{TV}(P_q^t, \pi_q) \leq \frac{1}{e}\right\}$$

where $P_q^t(\chi, \chi')$ is the probability of moving from χ to χ' in t steps and

$$d_{TV}(P_q^t, \pi_q) = \max_{\chi_1 \in \mathcal{C}_q} \frac{1}{2} \sum_{\chi_2 \in \mathcal{C}_q} |P_q^t(\chi_1, \chi_2) - \pi_q(\chi_2)|$$

is total variation distance. The mixing time of $\mathcal{M}_q$ captures the speed at which the chain converges to its stationary distribution: for every $\varepsilon > 0$, in order to get a sample from $\mathcal{C}_q$ which is within ε of π_q (in total variation distance), it is necessary and sufficient to run the chain from some arbitrarily chosen distribution for some multiple (depending on ε) of the mixing time. For surveys of issues related to the mixing time of a Markov chain, see *e.g.* (1; 21; 22).

Jerrum (15) and Salas and Sokal (24) independently showed that if Σ has maximum degree Δ and $q > 2\Delta$ then there is rapid mixing of the Glauber dynamics; *i.e.*, $\tau_{\mathcal{M}_q}(\Sigma)$ is polynomial in $|V|$. In fact, they showed that the mixing time is optimal ($O(|V| \log |V|)$). Bubley and Dyer (7) showed that there is rapid mixing for $q = 2\Delta$ and Molloy (20) improved this to optimal mixing. In a breakthrough result Vigoda (29) showed rapid mixing for $q \geq (11/6)\Delta$. More recently Dyer, Greenhill and Molloy (11) exhibited optimal mixing for $q \geq (2 - \varepsilon)\Delta$ for a small positive constant ε.

In this paper our aim is to explore the limitations of Glauber dynamics as a sampling tool by exhibiting a class of graphs for which the mixing time is essentially as far from optimal as possible. In this direction, Łuczak and Vigoda (19) have exhibited families of planar graphs for which $\mathcal{M}_q$ is not rapidly mixing for each fixed $q \geq 3$ and families of bipartite graphs with maximum degree Δ for which $\mathcal{M}_q$ is not rapidly mixing for any $3 \leq q \leq O(\Delta/\log \Delta)$. A drawback of these negative results is that the families exhibited consist of random graphs. Here, we attempt to remedy this by constructing explicit families of graphs for which Glauber dynamics is inefficient. We focus exclusively on the case $q = 3$ (we cannot see at the moment how to apply our techniques to any $q > 3$) and Σ regular bipartite. Specifically, we establish certain local and expansion conditions in a regular bipartite graph Σ that force $\tau_{\mathcal{M}_3(\Sigma)}$ to be (almost) exponential in $|V|$. The discrete hypercube is among the families of graphs which satisfy our conditions.

Our techniques actually apply to the class of ρ-local chains (considered in (6) and also in (10), where the terminology $\rho|V|$-*cautious* is employed) for suitably small ρ. A Markov chain $\mathcal{M}$ on state space $\mathcal{C}_q$ is ρ-*local* if in each step of the chain at most $\rho|V|$ vertices have their colour changed; that is, if

$$P_{\mathcal{M}}(\chi_1, \chi_2) \neq 0 \Rightarrow |\{v \in V : \chi_1(v) \neq \chi_2(v)\}| \leq \rho|V|.$$

Before stating our main result, we establish some notation. From now on, $\Sigma = (V, E)$ will be a d-regular bipartite graph with partition classes $\mathcal{E}$ and $\mathcal{O}$. For $u, v \in V$ we write $u \sim v$ if there is an edge in Σ joining u and v. Set $N(u) = \{w \in V : w \sim u\}$ ($N(u)$ is the *neighbourhood* of u) and for $A \subseteq V$ set $N(A) = \cup_{w \in A} N(w)$. For $A \subseteq \mathcal{E}$ (or $A \subseteq \mathcal{O}$) set

$$[A] = \{x \in V : N(x) \subseteq N(A)\}$$

(we think of $[A]$ as an external closure of A) and say that such an A is *small* if $|[A]| \leq |V|/4$. Note that $N(A)$ determines $[A]$ but not A itself.

Define the *bipartite expansion* of Σ by

$$\delta(\Sigma) = \min \left\{ \frac{|N(A)| - |[A]|}{|N(A)|} : A \subseteq \mathcal{E} \text{ (or } \mathcal{O}) \text{ small}, \ A \neq \emptyset \right\};$$

note that $0 \leq \delta < 1$. The second inequality is clear. To see the first, note that since Σ is regular and bipartite it has a perfect matching, and so satisfies

$$|X| \leq |N(X)| \text{ for all } X \subseteq \mathcal{E} \text{ or } \mathcal{O}. \tag{1}$$

That $0 \leq \delta$ now follows from $|[A]| \leq |N([A])| = |N(A)|$. The bipartite expansion constant is a measure of the proportion by which the neighbourhood size of a small set exceeds the size of the set itself, in the worst case.

Finally, define the *locality* $\ell(\Sigma)$ of Σ to be the largest $\ell \geq 0$ such that for all $x \sim y \in V$ and for all independent sets I (sets of vertices spanning no edges) in the subgraph of Σ induced by $N(x) \cup N(y)$ we have $|I| \leq 2d - \ell$. (So, for example, if Σ is the d-regular tree then $\ell(\Sigma) = 2$ since the subgraph induced by the neighbourhoods of adjacent vertices contains an independent set of size $2d - 2$; whereas if Σ is the complete d-regular bipartite graph then $\ell(\Sigma) = d$.)

Our main result is the following. Recall that $H(x) = -x \log x - (1-x)\log(1-x)$ is the usual binary entropy function.

Theorem 1.1. *Fix $\rho > 0$ satisfying $H(\rho) + \rho < 1$. There are constants $d_0, C_1, C_1', C_2 > 0$ all depending on ρ such that if Σ is a d-regular bipartite graph on N vertices with bipartite expansion δ and locality $\ell > 0$ satisfying*

$$\delta \geq \max\left\{\frac{C_1 \log^3 d}{d}, \frac{C_1' \log d}{\ell}\right\} \tag{2}$$

and with $d \geq d_0$ and if $\mathcal{M}(\Sigma)$ is an ergodic ρ-local Markov chain on state space $\mathcal{C}_3(\Sigma)$ with stationary distribution $\pi_3(\Sigma)$ then

$$\tau_{\mathcal{M}(\Sigma)} \geq \exp_2\left\{\frac{C_2 N \delta}{\log d}\right\}.$$

Note that for all $\rho \leq .22$ we have $H(\rho) + \rho < 1$. Here and throughout we use "log" for $\log_2$ and write $\exp_2 x$ for 2^x.

Remark 1.2. *The second inequality in (2) implies $\ell \geq \Omega(\log d/\delta)$. This condition appears in the derivation of (10), where it is only used in the weaker form $\ell = \omega(1)$ (which follows since $\delta \leq 1$). It is used in a more essential way in the derivation of (16) where it serves to limit, somewhat artificially, the number of 3-colourings of a bipartite graph with a given pre-image of 0. We expect that Theorem 1.1 should remain true with the second inequality in (2) removed.*

We now return to Glauber dynamics. This changes the colour of at most one vertex at each step, and so (as long as the underlying graph has at least five vertices) it is a ρ-local chain for $\rho = .2$. Fixing ρ to this value, all of the constants in Theorem 1.1 become absolute, and we have the following corollary.

Corollary 1.3. *There are constants $d_0, C_1, C_1', C_2 > 0$ such that if Σ satisfies the conditions of Theorem 1.1 and if the Glauber dynamics chain $\mathcal{M}_3(\Sigma)$ is ergodic then*

$$\tau_{\mathcal{M}_3(\Sigma)} \geq \exp_2\left\{\frac{C_2 N \delta}{\log d}\right\}.$$

Let us apply Theorem 1.1 to the case $\Sigma = Q_d$, the d-dimensional Hamming cube. This is the graph on vertex set $\{0,1\}^d$ with $x \sim y$ iff x and y differ on exactly one coordinate. For $d \geq 2$ we have $\ell(Q_d) = d$ (the graph induced by the union of the neighbourhoods of adjacent vertices is a perfect matching, so all independent sets are of size at most d), and $\delta(Q_d) = \Omega(1/\sqrt{d})$ (see *e.g.* (18, Lemma 1.3)). So the following is an immediate corollary of Theorem 1.1.

Corollary 1.4. *Fix $\rho > 0$ satisfying $H(\rho) + \rho < 1$. There is a constant $C = C(\rho) > 0$ such that for all $d \geq 2$, if $\mathcal{M}(Q_d)$ is an ergodic ρ-local Markov chain on state space $\mathcal{C}_3(Q_d)$ with stationary distribution $\pi_3(Q_d)$ then*

$$\tau_{\mathcal{M}} \geq \exp_2\left\{\frac{C2^d}{\sqrt{d}\log d}\right\}.$$

In particular this result applies to the Glauber dynamics chain, although in this case it is not necessary to hypothesize ergodicity.

Corollary 1.5. *There is a constant $C > 0$ such that for all $d \geq 2$,*

$$\tau_{\mathcal{M}_3(Q_d)} \geq \exp_2\left\{\frac{C2^d}{\sqrt{d}\log d}\right\}.$$

Proof: In the presence of Corollary 1.4, it suffices to show that the chain $\mathcal{M}_3(Q_d)$ is ergodic. We will show that if χ_1 is a 3-colouring of Q_d with $\chi_1(v_0) = 0$ for some $v_0 \in \mathcal{E}$ then there is a sequence of steps in the Glauber dynamics chain that takes χ_1 to a 2-colouring χ_2 of Q_d with $\chi_2(v) = 0$ for all $v \in \mathcal{E}$. This suffices, since it is clear that any one of the six 2-colourings of Q_d can be reached from any other via steps in the chain.

We make use of a correspondence between proper 3-colourings of Q_d and homomorphisms from Q_d to $\mathbb{Z}$ that send v_0 to 0. Formally, set

$$\mathcal{F}^{v_0} = \{f : V \to \mathbb{Z} : f(v_0) = 0 \text{ and } x \sim y \Rightarrow |f(x) - f(y)| = 1\}.$$

(This set was introduced in (2) and further studied in (12; 17).) Then, as observed by Randall (23), there is a bijection from $\mathcal{F}^{v_0}$ to $\mathcal{C}_3^{v_0} := \{\chi \in \mathcal{C}_3 : \chi(v_0) = 0\}$ given by $f \longrightarrow \Phi(f)$ where $\Phi(f)(v) = i$ iff $f(v) \equiv i \pmod 3$. Before verifying that this is indeed a bijection, we use the correspondence to establish the corollary.

For $f \in \mathcal{F}^{v_0}$ set $R(f) = \{f(v) : v \in V\}$. Now consider $\chi_1 \in \mathcal{C}_3^{v_0}$. If $|R(\Phi^{-1}(\chi_1))| = 2$, then we may take $\chi_2 = \chi_1$ and we are done. If $|R(\Phi^{-1}(\chi_1))| = k > 2$, then it suffices to exhibit a sequence of steps in the chain that takes χ_1 to some $\chi_3 \in \mathcal{C}_3^{v_0}$ with $|R(\Phi^{-1}(\chi_3))| = k - 1$.

Without loss of generality we may assume that $\Phi^{-1}(\chi_1)$ takes on some strictly positive values. Let ℓ be the largest such value, and let $v \in V$ be any vertex satisfying $\Phi^{-1}(\chi_1)(v) = \ell$. Note that $\ell - 2 \in R(\Phi^{-1}(\chi_1))$. Let $f : V \to \mathbb{Z}$ be the function that agrees with $\Phi^{-1}(\chi_1)$ off v and satisfies $f(v) = \ell - 2$. Since $\Phi^{-1}(\chi_1)(y) = \ell - 1$ for all $y \in N(v)$ and $v \neq v_0$ we have that $f \in \mathcal{F}^{v_0}$ and $\Phi(f) \in \mathcal{C}_3^{v_0}$ and that the Glauber dynamics chain permits a move from χ_1 to $\Phi(f)$. But we also have that $R(f) \subseteq R(\Phi^{-1}(\chi_1))$ and $|\{v \in V : f(v) = \ell\}| < |\{v \in V : \Phi^{-1}(\chi_1)(v) = \ell\}|$, so that by repeating the above described procedure m more times (where $m = |\{v \in V : f(v) = \ell\}|$) we arrive at the desired χ_3.

It remains to verify that Φ is a bijection. That it is injective is clear. To see that it is surjective, consider $\chi' \in \mathcal{C}_3^{v_0}$. We shall construct from χ' an $f \in \mathcal{F}^{v_0}$ with $\Phi(f) = \chi'$ by setting $f(v_0) = 0$

and then extending f level by level, where the k^{th} level of Q_d ($k = 0, \ldots, d$) is $\mathcal{L}_k := \{v \in V : dist(v, v_0) = k\}$ (here we are using $dist(\cdot, \cdot)$ for the usual graph distance). Note that for $v \in \mathcal{L}_k$, $N(v) \subseteq \mathcal{L}_{k-1} \cup \mathcal{L}_{k+1}$ and that for $f \in \mathcal{F}^{v_0}$ the values that f takes on $\mathcal{L}_k$ must all have the same parity.

So suppose we have specified f up to $\mathcal{L}_k$ for some $0 \leq k \leq d-1$. Consider $v \in \mathcal{L}_{k+1}$. If f is constant on $N(v) \cap \mathcal{L}_k$ then (since the construction of f has succeeded up to $\mathcal{L}_k$) we also have that χ' is constant on $N(v) \cap \mathcal{L}_k$ with $\chi'(y) \equiv f(y) \pmod 3$ for all $y \in N(v) \cap \mathcal{L}_k$. In this case we choose $f(v)$ such that $|f(v) - f(y)| = 1$ for all $y \in N(v) \cap \mathcal{L}_k$ and $\chi'(v) \equiv f(v) \pmod 3$.

If f is not constant on $N(v) \cap \mathcal{L}_k$, then we claim that there is some $\ell \in \mathbb{Z}$ such that f takes on only the values ℓ and $\ell + 2$ on $N(v) \cap \mathcal{L}_k$. For if not, then we have $y_1, y_2 \in N(v) \cap \mathcal{L}_k$ with $|f(y_1) - f(y_2)| \geq 4$. But by the structure of Q_d there must be $v' \in \mathcal{L}_{k-1}$ with $v' \sim y_1$ and $v' \sim y_2$, which forces $|f(y_1) - f(y_2)| \leq 2$. This contradiction establishes the two-value claim. We now set $f(v) = \ell + 1$, allowing the construction to continue. Since $\mathcal{L}_{k+1}$ is an independent set in Q_d, we may repeat the above-described procedure on each vertex of $\mathcal{L}_{k+1}$ independently, thus extending the construction of f to all of $\mathcal{L}_{k+1}$. $\square$

Remark 1.6. *Glauber dynamics for q-colourings of Q_d is not in general ergodic for $3 < q < \Delta(Q_d) + 1$. Indeed, it is straightforward to construct a 4-colouring χ of Q_3 which is frozen in the sense that $P_4(\chi, \chi') = 0$ for all $\chi' \neq \chi$; one simply assigns the colours 0, 1, 2 and 3 to a particular vertex and its three neighbours and then extend to a colouring of the whole of Q_3 according to the rule that on each face (4-cycle) of Q_3 all of the colours 0, 1, 2 and 3 must appear.*

Remark 1.7. *While this paper was under review, Galvin and Randall (13) used methods different to those of the present work to extend Corollary 1.5 to the discrete torus $T_{L,d}$, the graph on vertex set $\{0, \ldots, L-1\}^d$ in which two strings are adjacent if they differ on exactly one coordinate and differ by 1 (mod L) on that coordinate. The main result of (13) is that for $L \geq 4$ even and d large, the Glauber dynamics chain $\mathcal{M}_3$ on $\mathcal{C}_3(T_{L,d})$ satisfies $\tau_{\mathcal{M}_3} \geq \exp\{L^{d-1}/(d^4 \log^2 L)\}$.*

We prove Theorem 1.1 via a well-known conductance argument (introduced in (16)). A particularly useful form of the argument was given by Dyer, Frieze and Jerrum (10). Let $\mathcal{M}$ be an ergodic Markov chain on state space Ω with transition probabilities P and stationary distribution π. Let $A \subseteq \Omega$ and $M \subseteq \Omega \setminus A$ satisfy $\pi(A) \leq 1/2$ and $\omega_1 \in A, \omega_2 \in \Omega \setminus (A \cup M) \Rightarrow P(\omega_1, \omega_2) = 0$. Then from (10) we have

$$\tau_{\mathcal{M}} \geq \frac{\pi(A)}{8\pi(M)}.$$

We may think of M as a bottleneck set through which any run of the chain must pass in order to mix; if the bottleneck has small measure, then the mixing time is high.

Now let us return to the setup of Theorem 1.1. Set

$$\mathcal{C}_3^{b,\rho,0} = \mathcal{C}_3^{b,\rho,0}(\Sigma) = \{\chi \in \mathcal{C}_3 : \big||\chi^{-1}(0) \cap \mathcal{E}| - |\chi^{-1}(0) \cap \mathcal{O}|\big| \leq \rho N/2\}$$

($\mathcal{C}_3^{b,\rho,0}$ is the set of 3-colourings that are balanced with respect to 0) and

$$\mathcal{C}_3^{\mathcal{E},\rho,0} = \mathcal{C}_3^{\mathcal{E},\rho}(\Sigma) = \{\chi \in \mathcal{C}_3 : |\chi^{-1}(0) \cap \mathcal{E}| > |\chi^{-1}(0) \cap \mathcal{O}| + \rho N/2\}.$$

We may assume without loss of generality that $\pi_3(\mathcal{C}_3^{\mathcal{E},\rho,0}) \le 1/2$. Notice that since $\mathcal{M}$ changes the colour of at most ρN vertices in each step, we have that if $\chi_1 \in \mathcal{C}_3^{\mathcal{E},\rho,0}$ and $\chi_2 \in \mathcal{C}_3 \setminus (\mathcal{C}_3^{\mathcal{E},\rho,0} \cup \mathcal{C}_3^{b,\rho,0})$ then $P_{\mathcal{M}}(\chi_1, \chi_2) = 0$. We therefore have

$$\tau_{\mathcal{M}} \ge \frac{\pi_3(\mathcal{C}_3^{\mathcal{E},\rho,0})}{8\pi_3(\mathcal{C}_3^{b,\rho})} \ge \frac{2^{N/2}}{8|\mathcal{C}_3^{b,\rho,0}|},$$

the second inequality coming from the trivial lower bound $|\mathcal{C}_3^{\mathcal{E},\rho,0}| \ge 2^{N/2}$ (consider those χ with $\chi(v) = 0$ for all $v \in \mathcal{E}$). Theorem 1.1 thus follows from the following theorem, whose proof will be the main business of this paper.

Theorem 1.8. *Fix $\rho > 0$ satisfying $H(\rho) + \rho < 1$. There are constants $d_0, C_1, C_1', C_2 > 0$ all depending on ρ such that if Σ is a d-regular bipartite graph on N vertices with bipartite expansion δ and locality ℓ satisfying (2) and with $d \ge d_0$ then*

$$|\mathcal{C}_3^{b,\rho,0}| \le \exp_2\left\{\frac{N}{2}\left(1 - \frac{C_2\delta}{\log d}\right)\right\}.$$

Theorem 1.8 says more about the structure of $\mathcal{C}_3$ than just that the dynamics mixes slowly. From it, we can infer that for Σ and ρ satisfying the conditions of the theorem, $\mathcal{C}_3$ breaks naturally into six sets in such a way that once a ρ-local chain enters one of these dominant sets, it tends to remain there for an exponential time. These sets are characterized by a predominance of one (of three) colours on one (of two) partition classes. Indeed, defining $\mathcal{C}_3^{b,\rho,1}$ and $\mathcal{C}_3^{b,\rho,2}$ by analogy with $\mathcal{C}_3^{b,\rho,0}$ and setting $\mathcal{R}_3 = \mathcal{C}_3 \setminus \cup_{i=0}^2 \mathcal{C}_3^{b,\rho,i}$, we may partition $\mathcal{R}_3$ into six pieces by

$$\mathcal{R}_3 = \cup_{(x,y,z)\in\{\mathcal{E},\mathcal{O}\}^3\setminus\{(\mathcal{E},\mathcal{E},\mathcal{E}),(\mathcal{O},\mathcal{O},\mathcal{O})\}} \mathcal{R}_3^{(x,y,z)}$$

where $\mathcal{R}_3^{(x,y,z)} = \{\chi \in \mathcal{R}_3 : \chi \in \mathcal{C}_3^{x,\rho,0} \cap \mathcal{C}_3^{y,\rho,1} \cap \mathcal{C}_3^{z,\rho,2}\}$. If a ρ-local chain leaves $\mathcal{R}_3^{(x,y,z)}$ (for any (x, y, z)) it must enter $\cup_{i=0}^2 \mathcal{C}_3^{b,\rho,i}$ which, by Theorem 1.1 and a union bound, has exponentially small measure.

Before turning to the proof of Theorem 1.8 we pause to give a pleasing combinatorial corollary in the special case $\Sigma = Q_d$.

Corollary 1.9. *Fix ρ satisfying $H(\rho) + \rho < 1$. There is a constant $C = C(\rho) > 0$ such that for all $d \ge 2$, if χ is a uniformly chosen 3-colouring of Q_d then*

$$\mathbb{P}\left(\exists i : \left|\frac{|\chi^{-1}(i) \cap \mathcal{E}|}{|\mathcal{E}|} - \frac{|\chi^{-1}(i) \cap \mathcal{O}|}{|\mathcal{O}|}\right| \le \rho\right) \le \exp_2\left\{-\frac{C2^d}{\sqrt{d}\log d}\right\}.$$

In other words, the typical 3-colouring of Q_d exhibits strong $\mathcal{E}/\mathcal{O}$ imbalance on all colours.

Proof of Corollary 1.9: As previously observed, $\ell(Q_d) = d$ and $\delta(Q_d) \le \Omega(1/\sqrt{d})$, so (2) is satisfied for large enough d. It follows that there is a $C'(\rho)$ such that for large enough d and for each $i = 0, 1, 2$,

$$\left|\{\chi \in \mathcal{C}_3(Q_d) : \left||\chi^{-1}(i) \cap \mathcal{E}| - |\chi^{-1}(i) \cap \mathcal{O}|\right| \le \rho 2^{d-1}\}\right| \le \exp_2\left\{2^{d-1} - \frac{C'2^d}{\sqrt{d}\log d}\right\}.$$

Using $2^{2^{d-1}}$ as a lower bound on $|\mathcal{C}_3(Q_d)|$ (consider those colourings for which $\chi^{-1}(0) = \mathcal{E}$) we obtain

$$\mathbb{P}\left(\left||\chi^{-1}(i) \cap \mathcal{E}| - |\chi^{-1}(i) \cap \mathcal{O}|\right| \leq \rho 2^{d-1}\right) \leq \exp_2\left\{-\frac{C' 2^d}{\sqrt{d}\log d}\right\}$$

for χ chosen uniformly from $\mathcal{C}_3(Q_d)$. The stated bound follows for large d (with a constant C'' slightly larger than C') via a union bound and the fact that $|\mathcal{E}| = |\mathcal{O}| = 2^{d-1}$; we may obtain the bound for all d by appropriately modifying the constant C''. $\square$

2 Proof of Theorem 1.8

2.1 Overview of the proof

In this section we give an informal overview of the proof of Theorem 1.8.

We bound the number of balanced 3-colourings by bounding, for each pair $E \subseteq \mathcal{E}$, $O \subseteq \mathcal{O}$ with $||E| - |O|| \leq \rho N/2$ and $E \not\sim O$ (that is, with no edge in Σ joining E and O), the number of 3-colourings of Σ in which $E \cup O$ is the pre-image of 0. We then sum over all choices of E and O.

How many ways are there to 3-colour Σ given that $E \cup O$ is the pre-image of 0? Write $I(E)$ for the set of vertices in $N(E)$ all of whose neighbours are in E, and $I(O)$ for the neighbours of O all of whose neighbours are in O. There are two choices for each vertex in $I(E)$ and two for each vertex in $I(O)$, as well as two choices for each component in the graph obtained from Σ by removing E, O, $I(E)$ and $I(O)$ (each such component is a connected bipartite graph), all choices independent. The first step in the proof is an easy graph theory lemma that shows that the contribution from components in $\Sigma - (E \cup O \cup I(E) \cup I(O))$ is negligible. (This step uses the locality of Σ in an essential way.) This reduces the problem of bounding the number of balanced 3-colorings to the problem of estimating a sum of the form

$$\sum_{E,O:E\not\sim O} 2^{|I(E)|+|I(O)|}. \tag{3}$$

When $|E|$ and $|O|$ are both small (less than cN for a suitably small constant c) a naive count suffices to give an appropriate bound. For larger E and O, we must work harder. We partition the set of pairs (E, O) according to the parameters $a = |[E]|$, $g = |N(E)|$, $b = |I(E)|$, $h = |N(I(E))|$, $b' = |I(O)|$ and $h' = |N(I(O))|$. Within each class, each pair gives the same contribution ($2^{b+b'}$) to the sum in (3). The main point of the proof is an estimate on the size of $\mathcal{H} = \{(E, O) : (E, O)$ has parameters a, g, b, h, b' and $h'\}$ of the form

$$|\mathcal{H}| \leq \exp_2\left\{\frac{N}{2} - b - b' - \frac{cN\delta}{\log d}\right\} \tag{4}$$

for sufficiently large $d = d(\rho)$ and suitable $c = c(\rho)$. The proof is completed by invoking (4) and summing over all choices of a, g, *et cetera.*

The proof of (4) involves the idea of approximation. To bound $|\mathcal{H}|$, we produce a small set $\mathcal{U}$ with the properties that each $(E, O) \in \mathcal{H}$ is approximated (in an appropriate sense) by some $U \in \mathcal{U}$, and for each $U \in \mathcal{U}$, the number of $(E, O) \in \mathcal{H}$ that could possibly be approximated by U is small. (Each $U \in \mathcal{U}$ will consist of six parts; one each approximating E, $N(E)$, $I(E)$, $N(I(E))$, $I(O)$ and $N(I(O))$.) The product of the bound on $|\mathcal{U}|$ and the bound on the number of those $(E,O) \in \mathcal{H}$ that may be approximated by any U is then a bound on $|\mathcal{H}|$.

The main inspiration for our approximation scheme is the work of A. Sapozhenko, who, in (26), gave a relatively simple derivation for the asymptotics of the number of independent sets in Q_d, earlier derived in a more involved way in (18). We produce the set $\mathcal{U}$ by appealing to a lemma from (14) where a similar approximation scheme was used to show that the mixing time of Glauber dynamics for the hard-core model on Q_d with activity λ is (essentially) exponential in 2^d for large enough λ. The proof that each $U \in \mathcal{U}$ approximates only a small number of $(E,O) \in \mathcal{H}$ is a modification of a similar proof from (12) in which it is shown that a uniformly chosen homomorphism from Q_d to $\mathbb{Z}$ almost surely takes on at most 5 values, and also that the number of proper 3-colourings of Q_d is asymptotic to $2e2^{2^{d-1}}$ as d goes to infinity.

2.2 The proof

We begin by establishing some more notation. From now on, we write M for $N/2$. For $A \subseteq \mathcal{E}$ and $B \subseteq \mathcal{O}$ write $A \not\sim B$ if for all $x \in A$ and $y \in B$, $x \not\sim y$ (this is equivalent to both $N(A) \cap B = \emptyset$ and $N(B) \cap A = \emptyset$). For $S \subseteq V$ write $d_S(u)$ for $|N(u) \cap S|$ and $\mathrm{comp}(S)$ for the number of components of the subgraph induced by S. Finally for $T \subseteq \mathcal{E}$ (or $\mathcal{O}$) set

$$I(T) = \{x \in N(T) : N(x) \subseteq T\} \ (= \{x \in V : N(x) \subseteq T\}).$$

We think of $I(T)$ as an internal closure of $N(T)$. Note that for all $T \subseteq \mathcal{E}$ (or $\mathcal{O}$), $[I(T)] = I(T)$, $I(T) \subseteq N(T)$ and $N(I(T)) \subseteq T$.

For $\chi \in \mathcal{C}_3^{b,\rho,0}$ set

$$\begin{aligned} E &= \chi^{-1}(0) \cap \mathcal{E}, \\ O &= \chi^{-1}(0) \cap \mathcal{O}, \\ I &= I(E), \\ J &= I(O) \end{aligned}$$

and

$$R = V \setminus (E \cup O \cup I \cup J).$$

We assume the convention that whenever E and O have been specified, I, J and R will be used as shorthand for $I(E)$, $I(O)$ and $V \setminus (E \cup O \cup I \cup J)$.

For $E \subseteq \mathcal{E}$ and $O \subseteq \mathcal{O}$ set

$$\mathcal{C}_3(E,O) = \{\chi \in \mathcal{C}_3 : E(\chi) = E,\ O(\chi) = O\}.$$

Note that $\mathcal{C}_3(E,O) \neq \emptyset$ iff $E \not\sim O$. For $\mathcal{C}_3(E,O) \neq \emptyset$ we have

$$|\mathcal{C}_3(E,O)| = 2^{|I|+|J|+\mathrm{comp}(R)}.$$

To see this, note that once we have specified that the set of vertices coloured 0 is $E \cup O$, we have a free choice between 1 and 2 for the colour at $x \in I \cup J$, with each choice independent. This accounts for the factor $2^{|I|+|J|}$. The subgraph induced by R breaks into $\mathrm{comp}(R)$ components, each of which is bipartite and may be coloured in exactly two ways using the colours 1 and 2. This accounts for the factor $2^{\mathrm{comp}(R)}$. We therefore have

$$|\mathcal{C}_3^{b,\rho,0}| = \sum_{\substack{E\subseteq\mathcal{E},\ O\subseteq\mathcal{O}:\\ ||E|-|O||\leq\rho M,\ E\not\sim O}} 2^{|I|+|J|+\mathrm{comp}(R)}.$$

A key observation is the following.

Proposition 2.1. *For $E \not\sim O$, $\mathrm{comp}(R) \leq 2M/\ell$.*

Proof: Let C be a component of $V \setminus (E \cup O)$. If $C = \{v\}$ consists of a single vertex, then (depending on the parity of v) we have either $N(v) \subseteq E$ or $N(v) \subseteq O$ and so $v \in I(E) \cup I(O)$. Otherwise, let vw be an edge of C. We have

$$|(N(v) \cup N(w)) \cap (E \cup O)| \leq 2d - \ell$$

(recall that $E \cup O = \chi^{-1}(0)$ is an independent set), and so $|C| \geq \ell$. The result follows. $\square$

We now decompose $\mathcal{C}_3^{b,\rho,0}$ into four pieces. Set

$$\alpha = \sup\left\{\alpha' \in \left[0, \frac{1}{2} - \rho\right] : 2\alpha' + \rho + H(\alpha') + H(\rho + \alpha') \leq \frac{1}{2}\left(1 + \rho + H(\rho)\right)\right\}. \tag{5}$$

Since $H(\rho) + \rho < 1$, α is a strictly positive constant depending on ρ. Set

$$\mathcal{C}_3^{b,\rho,0}(triv, \mathcal{E}) = \{\chi \in \mathcal{C}_3^{b,\rho,0} : |E| \leq \alpha M, |E| \leq |O|\}$$

and define $\mathcal{C}_3^{b,\rho,0}(triv, \mathcal{O})$ analogously. Set

$$\mathcal{C}_3^{b,\rho,0}(nt, \mathcal{E}) = \{\chi \in \mathcal{C}_3^{b,\rho,0} \setminus (\mathcal{C}_3^{b,\rho,0}(triv, \mathcal{E}) \cup \mathcal{C}_3^{b,\rho,0}(triv, \mathcal{O})) : E \text{ small}\}$$

(recall that E is small if $|[E]| \leq M/2$) and define $\mathcal{C}_3^{b,\rho,0}(nt, \mathcal{O})$ similarly. Since Σ has a perfect matching, it is easy to see that for $\chi \in \mathcal{C}_3$ at least one of $|E| \leq M/2$, $|O| \leq M/2$ holds; moreover, it is straightforward to check that at least one of $|[E]| \leq M/2$, $|[O]| \leq M/2$ holds also; that is, that at least one of E, O is small, and so

$$\mathcal{C}_3^{b,\rho,0} = \mathcal{C}_3^{b,\rho}(triv, \mathcal{E}) \cup \mathcal{C}_3^{b,\rho,0}(triv, \mathcal{O}) \cup \mathcal{C}_3^{b,\rho}(nt, \mathcal{E}) \cup \mathcal{C}_3^{b,\rho,0}(nt, \mathcal{O}).$$

In what follows we make extensive use of a result concerning the sums of binomial coefficients which follows from the Chernoff bounds (8) (see also (5), p.11):

$$\sum_{i=0}^{[\beta M]} \binom{M}{i} \leq 2^{H(\beta)M} \qquad \text{for } \beta \leq \tfrac{1}{2}. \tag{6}$$

Also, since $H(x) \leq 2x \log 1/x$ for $x \leq e^{-1}$,

$$\sum_{i=0}^{[\beta M]} \binom{M}{i} \leq 2^{2\beta M \log(1/\beta)} \qquad \text{for } \beta \leq e^{-1}. \tag{7}$$

We begin by bounding $|\mathcal{C}_3^{b,\rho,0}(triv,\mathcal{E})|$. Noting that $|I| \le |E|$ and $|J| \le |O|$ always (this follows from (1)) we have

$$
\begin{aligned}
|\mathcal{C}_3^{b,\rho,0}(triv,\mathcal{E})| &= \sum_{\substack{E\subseteq\mathcal{E},\ O\subseteq\mathcal{O}:\\ |E|\le\alpha M,\ |O|\le(\alpha+\rho)M}} 2^{|E|+|O|+\mathrm{comp}(R)} \\
&\le \exp_2\left\{\frac{2M}{\ell}\right\} \sum_{\substack{E\subseteq\mathcal{E},\ O\subseteq\mathcal{O}:\\ |E|\le\alpha M,\ |O|\le(\alpha+\rho)M}} 2^{|E|+|O|} \qquad (8)\\
&\le \exp_2\left\{M\left(2\alpha+\rho+\frac{2}{\ell}\right)\right\}\sum_{i=0}^{[\alpha M]}\binom{M}{i}\sum_{i=0}^{[(\alpha+\rho)M]}\binom{M}{i} \\
&\le \exp_2\left\{M\left(2\alpha+\rho+H(\alpha)+H(\rho+\alpha)+\frac{2}{\ell}\right)\right\} \qquad (9)\\
&\le \exp_2\left\{\frac{N}{2}\left(1-\frac{\delta}{\log d}\right)\right\} \qquad (10)
\end{aligned}
$$

for sufficiently large $d = d(\rho)$. In (8) we have used Proposition 2.1. In (9) we use (6) while in (10) we use (5) to obtain

$$2\alpha+\rho+H(\alpha)+H(\rho+\alpha)+\frac{2}{\ell} \le 1-\varepsilon+\frac{2}{\ell}$$

for some $\varepsilon = \varepsilon(\alpha)$, and then use the second inequality in (2) (in the weak form that $\ell = \omega(1)$) to obtain

$$1-\varepsilon+\frac{2}{\ell} \le 1-\frac{\delta}{\log d}$$

(note that $\delta/\log d = o(1)$). Similarly, we have

$$|\mathcal{C}_3^{b,\rho,0}(triv,\mathcal{O})| \le \exp_2\left\{\frac{N}{2}\left(1-\frac{\delta}{\log d}\right)\right\} \qquad (11)$$

for suitable d.

Next we turn to $\mathcal{C}_3^{b,\rho,0}(nt,\mathcal{E})$ and $\mathcal{C}_3^{b,\rho,0}(nt,\mathcal{O})$. Without loss of generality we may assume $|\mathcal{C}_3^{b,\rho,0}(nt,\mathcal{E})| \le |\mathcal{C}_3^{b,\rho,0}(nt,\mathcal{O})|$. Bearing (10) and (11) in mind, Theorem 1.8 now follows from

$$|\mathcal{C}_3^{b,\rho,0}(nt,\mathcal{E})| \le \exp_2\left\{\frac{N}{2}\left(1-\frac{c\delta}{\log d}\right)\right\} \qquad (12)$$

for some constant $c = c(\rho)$.

For integers a, g, b, h, b' and h', set

$$\mathcal{H}(a,g,b,h,b',h') = \left\{(E,O) : \begin{array}{c} E\subseteq\mathcal{E},\ O\subseteq\mathcal{O},\ E\not\sim O,\ |[E]|=a,\ |N(E)|=g, \\ |I|=b,\ |N(I)|=h,\ |J|=b',\ |N(J)|=h' \end{array}\right\}.$$

Our main lemma is the following (*cf.* (14, Theorem 2.1)).

Lemma 2.2. *For each $\beta_1,\beta_2>0$, there are constants $d_0, c>0$ depending on both β_1 and β_2 such that the following holds. If G is a d-regular bipartite graph with bipartite expansion $\delta \ge d^{-\beta_1}$ and $d \ge d_0$ and if a satisfies $\beta_2 M \le a \le M/2$, then for any g, b, h, b' and h' we have*

$$|\mathcal{H}(a,g,b,h,b',h')| \le \exp_2\left\{M\left(1+\frac{15\log^2 d}{d}\right)-b-b'-\frac{c\delta g}{\log d}\right\}.$$

For $G = \Sigma$ we may take $\beta_1 = 2$ (say). Note that for each $(E, O) \in \mathcal{C}_3^{b,\rho,0}(nt, \mathcal{E})$ with $|I| = b$ and $|J| = b'$, $(E, O) \in \mathcal{H}(a, g, b, h, b', h')$ for some a, g, h and h' with $\alpha M \le a \le M/2$. With the steps justified below, we therefore have

$$\begin{aligned}
|\mathcal{C}_3^{b,\rho,0}(nt, \mathcal{E})| &\le \exp_2\left\{\frac{2M}{\ell}\right\} \sum_{\substack{a,g,b,h,b',h':\\ \alpha M \le a \le M/2}} |\mathcal{H}(a, g, b, h, b', h')| 2^{b+b'} \\
&\le \exp_2\left\{M\left(1 + \frac{2}{\ell} + \frac{15\log^2 d}{d}\right)\right\} \sum_{\substack{a,g,b,h,b',h':\\ \alpha M \le a \le M/2}} \exp_2\left\{-\frac{c\delta g}{\log d}\right\} && (13) \\
&\le \exp_2\{M\left(1 + \frac{2}{\ell} + \frac{21\log^2 d}{d}\right) \max_{\substack{\alpha M \le a \le M/2\\ a \le g}} \left\{\exp_2\left\{-\frac{c\delta g}{\log d}\right\}\right\} && (14) \\
&\le \exp_2\left\{M\left(1 + \frac{2}{\ell} + \frac{21\log^2 d}{d} - \frac{c\alpha\delta}{\log d}\right)\right\} && (15) \\
&\le \exp_2\left\{\frac{N}{2}\left(1 - \frac{c'\delta}{\log d}\right)\right\} && (16)
\end{aligned}$$

verifying (12) and completing the proof of Theorem 1.8. The main point, (13), is an application of Lemma 2.2. Here the constant c depends on α and therefore on ρ. In (14) we use that $M \le \exp\{M \log^2 d/d\}$ for all d. In (15) we have chosen $g = \alpha M$ to maximize the exponent. Finally in (16) we may (for example) take $C_1 = 43/(\alpha c)$ and $C_1' = 4/(\alpha c)$, and use both inequalities in (2). The final constant c' depends only on c and α and therefore only on ρ, as claimed.

To prove Lemma 2.2, we use a notion of approximation introduced in (25). An *approximation* for $A \subseteq \mathcal{E}$ is a pair $(F, S) \subseteq \mathcal{O} \times \mathcal{E}$ satisfying

$$F \subseteq N(A), \quad S \supseteq [A],$$
$$d_F(u) \ge d - \sqrt{d} \quad \forall u \in S$$

and

$$d_{\mathcal{E}\setminus S}(v) \ge d - \sqrt{d} \quad \forall v \in \mathcal{O} \setminus F.$$

For $A \subseteq \mathcal{O}$ we make the analogous definition.

The following lemma is from (14) (a combination of Lemmata 3.2 and 3.3). We use the shorthand $\binom{t}{\le k}$ for $\sum_{0 \le i \le k} \binom{t}{i}$.

Lemma 2.3. *Let G be a d-regular bipartite graph with $2M$ vertices. For each a and g set*

$$\mathcal{A}(a, g) = \{A \subseteq \mathcal{E} : |[A]| = a, |N(A)| = g\}.$$

There is a family $\mathcal{W} = \mathcal{W}(a, g) \subseteq 2^{\mathcal{O}} \times 2^{\mathcal{E}}$ with

$$|\mathcal{W}| \le \binom{M}{\le \frac{2g\log d}{d}} \binom{2g\log d}{\le \frac{2g}{d}} \binom{2d^3 g\log d}{\le \frac{2(g-a)}{\sqrt{d}}} \binom{2g\log d}{\le (g-a)\frac{\sqrt{d}}{(d-\sqrt{d})}}$$

such that every $A \in \mathcal{A}(a, g)$ has an approximation in $\mathcal{W}$. The analogous result holds with $\mathcal{O}$ replacing $\mathcal{E}$ in the definition of $\mathcal{A}(a, g)$.

Remark 2.4. *If a and g satisfy $g - a \geq d^{-\beta} g$ for some constant β then using (7) the bound on $|\mathcal{W}|$ from Lemma 2.3 may be rewritten as*

$$|\mathcal{W}| \leq \exp_2 \left\{ \frac{5M \log^2 d}{d} + \frac{(6\beta + 17)(g-a)\log d}{\sqrt{d}} \right\}$$

as long as d is sufficiently large (as a function of β).

Say that a sextuple $(F, S, P, Q, P', Q') \subseteq \mathcal{O} \times \mathcal{E} \times \mathcal{E} \times \mathcal{O} \times \mathcal{O} \times \mathcal{E}$ is an *approximation* for $(E, O) \in \mathcal{H}(a, g, b, h, b', h')$ if (F, S) is an approximation for E, (P, Q) is an approximation for I and (P', Q') is an approximation for J.

Lemma 2.5. *Let G, a, g, b, h, b' and h' be as in Lemma 2.2. There are constants $c_1 = c_1(\beta_1) > 0$ and $c_2 = c_2(\beta_1, \beta_2) > 0$ and a family $\mathcal{X} = \mathcal{X}(a, g, b, h, b', h') \subseteq 2^{\mathcal{O}} \times 2^{\mathcal{E}} \times 2^{\mathcal{E}} \times 2^{\mathcal{O}} \times 2^{\mathcal{O}} \times 2^{\mathcal{E}}$ with*

$$|\mathcal{X}| \leq \exp_2 \left\{ \frac{15M \log^2 d}{d} + \frac{c_1(g-a)\log d}{\sqrt{d}} + \frac{c_1(h-b)\log d}{\sqrt{d}} + \frac{c_2(h'-b')\log d}{\sqrt{d}} \right\}$$

such that every $(E, O) \in \mathcal{H}(a, g, b, h, b', h')$ has an approximation in $\mathcal{X}$.

Proof: We apply Lemma 2.3 (in the form given in Remark 2.4) to each of E, I and J independently. Note that $g - a \geq d^{-\beta_1} g$ and $h - b \geq d^{-\beta_1} h$ follow from the assumptions on δ in Lemma 2.2 (recall $|[E]| = a \leq M/2$ and $|[I]| = |I| \leq |E| \leq M/2$), justifying the first two applications of Lemma 2.3 and the dependence of c_1 on β_1 alone.
For the third application, note that if $b' \leq M/2$ we have $h' - b' \geq d^{-\beta_1} h'$ (recall $|[J]| = |J|$). If $b' > M/2$, then $|\mathcal{O} \setminus N(J)| \leq M/2$. Since $[J] = J$ we also have $[\mathcal{O} \setminus N(J)] = \mathcal{O} \setminus N(J)$ and $N(\mathcal{O} \setminus N(J)) = \mathcal{E} \setminus J$ and so (by the bound on δ) $(M - b') - (M - h') \geq d^{-\beta_1}(M - b')$. Using $h' \geq b'$ it follows that $h' - b' \geq d^{-\beta_1}(M - h')$. But since $N(J) \cap N(E) = \emptyset$ we have $h' \leq M - g \leq M - a \leq (1 - \beta_2)M$ and so $h' - b' \geq d^{-\beta_1}(1/(1-\beta_2) - 1)h' \geq d^{-c} h'$, where the constant c depends on both β_1 and β_2. □

Before going on to the final step in the proof of Lemma 2.2, we need the following simple inequalities ((14, Lemma 3.1) in the case $\psi = \sqrt{d}$). If (F, S, P, Q, P', Q') is an approximation for $(E, O) \in \mathcal{H}(a, g, b, h, b', h')$ then for suitably large d

$$|S| \leq |F| + \frac{3(g-a)}{\sqrt{d}}, \qquad |Q| \leq |P| + \frac{3(h-b)}{\sqrt{d}} \qquad \text{and} \qquad |Q'| \leq |P'| + \frac{3(h'-b')}{\sqrt{d}}. \tag{17}$$

Bearing this and the fact that $g - a \geq \delta g$ in mind, Lemma 2.2 is implied by Lemma 2.5 and the following reconstruction lemma.

Lemma 2.6. *Let G, a, g, b, h, b' and h' be as in Lemma 2.2. There are constants $c_1 = c_1(\beta_1) > 0$ and $c_2 = c_2(\beta_1, \beta_2) > 0$ such that for each $(F, S, P, Q, P', Q') \subseteq \mathcal{O} \times \mathcal{E} \times \mathcal{E} \times \mathcal{O} \times \mathcal{O} \times \mathcal{E}$ satisfying (17) there are at most*

$$\exp_2 \left\{ M - b - b' - \frac{c_1(g-a)}{\log d} - \frac{c_1(h-b)}{\log d} - \frac{c_2(h'-b')}{\log d} \right\}$$

pairs $(E, O) \subseteq \mathcal{E} \times \mathcal{O}$ satisfying

$$F \subseteq N(E), \quad S \supseteq [E], \quad P \subseteq N(I), \quad Q \supseteq I, \quad P' \subseteq N(J) \quad \text{and} \quad Q' \supseteq J. \tag{18}$$

Proof: For notational convenience, write t for $g-a$, s for $h-b$ and s' for $h'-b'$. Say that S is *tight* if $|S|<g-c_1't/\log d$ and *slack* otherwise, that Q is *tight* if $|Q|<b+c_1's/\log d$, and *slack* otherwise, and that Q' is *tight* if $|Q'|<b'+c_2's'/\log d$, and *slack* otherwise, where $c_1'=c_1'(\beta_1)>0$ and $c_2'=c_2'(\beta_1,\beta_2)>0$ are constants that will be specified presently.

We now describe a procedure which, for input (F,S,P,Q,P',Q') satisfying (17), produces an output (E,O) which satisfies (18). The procedure involves a sequence of choices, the nature of the choices depending on whether S, Q and Q' are tight or slack.

We begin by identifying a subset D of E which can be specified relatively cheaply: if Q is tight, we pick $I\subseteq Q$ with $|I|=b$ and take $D=N(I)$; if Q is slack, we simply take $D=P$ (recalling that $P\subseteq N(I)\subseteq E$).

If S is tight, we complete the specification of E by choosing $E\setminus D\subseteq S\setminus D$. If S is slack, we first complete the specification of $N(E)$ by choosing $N(E)\setminus F\subseteq N(S)\setminus F$. We then complete the specification of E by choosing $E\setminus D\subseteq[E]\setminus D$ (noting that we do know $[E]\setminus D$ at this point).

Next we turn to the specification of O. As with E, we begin by identifying a subset D' of O: if Q' is tight, we pick $J\subseteq Q'$ with $|J|=b'$ and take $D'=N(J)$; if Q' is slack, we simply take $D'=P'$. From here, we complete the specification of O by choosing $O\setminus D'\subseteq\mathcal{O}\setminus(N(E)\cup D')$ (recall that $E\not\sim O$).

This procedure produces all pairs (E,O) satisfying (18). Before bounding the number of outputs, we gather together some useful observations.

First note that as established in the proof of Lemma 2.5 we have

$$g-a\geq d^{-\beta_1}g,\qquad h-b\geq d^{-\beta_1}h\qquad\text{and}\qquad h'-b'\geq d^{-c}h' \tag{19}$$

where the constant $c>0$ depends on both β_1 and β_2, while from (17) we have

$$|S|\leq 2g,\qquad |Q|\leq 2h\qquad\text{and}\qquad |Q'|\leq 2h' \tag{20}$$

for suitably large d.

If Q is tight then there are at most

$$\sum_{i\leq c_1's/\log d}\binom{|Q|}{|Q|-i}\leq\sum_{i\leq c_1's/\log d}\binom{2h}{i}\leq 2^{s/2}$$

possibilities for D (for sufficiently small choice of the constant c_1', depending on β_1), and in this case $|D|=h$. Here we are using (7) and (19). If Q is slack there is just one possibility for D, and in this case (using (17))

$$|D|>b+c_1's/\log d-3s/\sqrt{d}\geq b+c_1's/2\log d \tag{21}$$

for suitably large d.

Similarly if Q' is tight then there are at most $2^{s'/2}$ possibilities for D' (for suitably small c_2' depending on both β_1 and β_2; here we use (19)), and in this case $|D'|=h'$; while if Q' is slack there is just one possibility for D', and in this case $|D'|>b'+c_2's'/2\log d$ for suitably large d.

If S is slack then (17) implies $|N(E)\setminus F|<2c_1't/\log d$ and since $|N(S)\setminus F|\leq d|S|\leq 2dg$ (see (20)) the number of possibilities for $N(E)\setminus F$ is at most

$$\sum_{i<2c_1't/\log d}\binom{2gd}{i}\leq 2^{t/2} \tag{22}$$

for suitable small c'_1 depending on β_1 (here again we use (19)).

Now we assume that d, c'_1 and c'_2 are suitably chosen so that all the previously made observations hold. We bound the number of outputs of the procedure, considering first the four cases determined by whether S and Q are slack or tight, and then considering the two cases of whether Q' is slack or tight. If S and Q are both tight then the number of possibilities for E is at most

$$\exp_2\{(s/2) + (g - c'_1 t/\log d - h)\} \leq \exp_2\{g - c'_1 t/\log d - b - s/2\}. \tag{23}$$

(The first term in the exponent on the left-hand side corresponds to the choice of D (using (21)), and the second to the choice of $E \setminus D \subseteq S \setminus D$ (note that since S and Q are both tight, $|S \setminus D| \leq g - c'_1 t/\log d - h$).

If S is tight and Q is slack then the total is at most

$$\exp_2\{g - c'_1 t/\log d - b - c'_1 s/2\log d\}. \tag{24}$$

(Here there is no choice for D, and the exponent corresponds to the choice of $E \setminus D \subseteq S \setminus D$ (using (21)).)

If Q is tight then $|[E] \setminus D| = a - h$, so that if S is slack (and Q tight) then the number of possibilities for E is at most

$$\exp_2\{(s/2) + (t/2) + (a - h)\} \leq \exp_2\{g - t/2 - b - s/2\}. \tag{25}$$

(The first term in the exponent on the left-hand side corresponds to the choice of D (using (21)), the second to the choice of $N(E) \setminus F$ (using (22)) and the third to the choice of $E \setminus D$.)

If Q is slack then $|[E] \setminus D| \leq a - b - c'_1 s/2\log d$ (see (21)), so that if S and Q are both slack the number of possibilities for E is at most

$$\exp_2\{(t/2) + (a - b - c'_1 s/2\log d)\} \leq \exp_2\{g - t/2 - b - c'_1 s/\log d\}. \tag{26}$$

(The first term in the exponent on the left-hand side corresponds to the choice of $N(E) \setminus F$ and the second to the choice of $E \setminus D$.)

Now we consider the number of choices for O, given our choice of E. Note that $O \subseteq (\mathcal{O} \setminus N(E))$, a set of size $M - g$. If Q' is tight then the number of possibilities for O is at most

$$\exp_2\{(s'/2) + (M - g - h')\} \leq \exp_2\{M - g - b' - s'/2\}. \tag{27}$$

(The first term in the exponent on the left-hand side corresponds to the choice of D', and the second to the choice of $O \setminus D' \subseteq \mathcal{O} \setminus (N(E) \cup D')$.)

Finally if Q' is slack then the number of possibilities for O is at most

$$\exp_2\{M - g - b' - c'_2 s'/2\log d\}. \tag{28}$$

(Here there is no choice for D', and the exponent corresponds to the choice of $O \setminus D' \subseteq \mathcal{O} \setminus (N(E) \cup D')$.)

Combining (23), (24), (25), (26), (27) and (28) we obtain the lemma. $\square$

References

[1] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, monograph in preparation, available at `http://stat-www.berkeley.edu/users/aldous/RWG/book.html`.

[2] I. Benjamini, O. Häggström and E. Mossel, On random graph homomorphisms into $\mathbb{Z}$, *J. Combinatorial Th. (B)* **78** no. 1 (2000), 86–114. MR1737627

[3] B. Bollobás, *Extremal Graph Theory*, Academic Press, New York, 1978. MR0506522

[4] B. Bollobás, *Modern Graph Theory*, Springer, New York, 1998. MR1633290

[5] B. Bollobás, *Random Graphs*, Cambridge University Press, Cambridge, 2001. MR1864966

[6] C. Borgs, J. Chayes, A. Frieze, J. Kim, P. Tetali, E. Vigoda, V. Vu, Torpid Mixing of some Monte Carlo Markov Chain algorithms in Statistical Physics, *Proc. IEEE FOCS '99*, 218–229. MR1917562

[7] R. Bubley and M. Dyer, Path coupling: a technique for proving rapid mixing in Markov chains, *Proc. IEEE FOCS '97*, 223–231.

[8] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Ann. Math. Statistics* **23** (1952), 493–507. MR0057518

[9] R. Diestel, *Graph Theory*, Springer, New York, 2005. MR1448665

[10] M. Dyer, A. Frieze and M. Jerrum, On counting independent sets in sparse graphs, *SIAM J. Comp.* **31** (2002), 1527–1541. MR1936657

[11] M. Dyer, C. Greenhill, M. Molloy, Very rapid mixing of the Glauber dynamics for proper colorings on bounded degree graphs, *Random Struc. & Alg.* **20** (2002), 98–114. MR1871953

[12] D. Galvin, On homomorphisms from the Hamming cube to $\mathbb{Z}$, *Isr. J. Math.* **138** (2003), 189–213. MR2031957

[13] D. Galvin and D. Randall, Torpid Mixing of Local Markov Chains on 3-Colorings of the Discrete Torus, *Proc. ACM–SIAM SODA '07*, 376–384.

[14] D. Galvin and P. Tetali, Slow mixing of Glauber dynamics for the hard-core model on the Hamming cube, *Random Structures & Alg.* **28** (2006) 427-443. MR2225701

[15] M. Jerrum, A very simple algorithm for estimating the number of k-colourings of a low-degree graph, *Random Struc. & Alg.* **7** (1995), 157–165. MR1369061

[16] M. Jerrum and A. Sinclair, Conductance and the rapid mixing property for Markov chains: the approximation of the permanent resolved, *Proc. ACM STOC '88*, 235–243.

[17] J. Kahn, Range of cube-indexed random walk, *Isreal J. Math.* **124** (2001) 189–201. MR1856513

[18] A. Korshunov and A. Sapozhenko, The number of binary codes with distance 2, *Problemy Kibernet.* **40** (1983), 111–130. (Russian) MR0717358

[19] T. Łuczak and E. Vigoda, Torpid mixing of the Wang-Swendsen-Kotecký algorithm for sampling colorings, *J. Discrete Alg.* **3** no. 1 (2005), 92–100. MR2167765

[20] M. Molloy, Very rapidly mixing Markov chains for 2Δ-colourings and for independent sets in a 4-regular graph, *Random Struc. & Alg.* **18** (2001), 101–115. MR1809717

[21] R. Montenegro and P. Tetali, Mathematical aspects of mixing times in Markov chains, Foundations and Trends in Theoretical Computer Science **1** no. 3 (2006), 237–354.

[22] D. Randall, Mixing, *Proc. IEEE FOCS '03*, 4–15.

[23] D. Randall, personal communication.

[24] J. Salas and A. Sokal, Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem, *J. Stat. Phys.* **86** (1997), 551–579. MR1438965

[25] A. Sapozhenko, On the number of connected subsets with given cardinality of the boundary in bipartite graphs, *Metody Diskret. Analiz.* **45** (1987), 42–70. (Russian) MR0946363

[26] A. A. Sapozhenko, The number of antichains in ranked partially ordered sets, *Diskret. Mat.* **1** (1989), 74–93. (Russian; translation in Discrete Math. Appl. 1 no. 1 (1991), 35–58) MR1072642

[27] A. Sokal, Chromatic Polynomials, Potts Models and All That, *Physica* **A**279 (2000), 324–332.

[28] A. Sokal, A Personal List of Unsolved Problems Concerning Lattice Gases and Antiferromagnetic Potts Models, *Markov Process. Related Fields* **7** (2001), 21–38. MR1835744

[29] E. Vigoda, Improved bounds for sampling colorings, *J. Math. Phys.* **41** (2000), 1555–1569. MR1757969