[redacted]

Random Sampling of 3-colourings in $\mathbb{Z}^2$ *

Leslie Ann Goldberg, Russell Martin, Mike Paterson
Department of Computer Science
University of Warwick
Coventry CV4 7AL
United Kingdom

ABSTRACT

We consider the problem of uniformly sampling proper 3-colourings of an $m \times n$ rectangular region of $\mathbb{Z}^2$. We show that the single-site "Glauber-dynamics" Markov chain is rapidly mixing. Our result complements an earlier result of Luby, Randall and Sinclair, which demonstrates rapid mixing when there is a fixed boundary (whose colour cannot be changed).

1. INTRODUCTION

A proper k-colouring of a graph G is a labeling of the vertices with elements from $\{1, \ldots, k\}$ such that neighbouring vertices are assigned different colours. Often, Markov chain simulation is used to sample proper colourings. A natural chain is the *Glauber-dynamics* Markov chain, whose state space is the set of all proper colourings. To move from one colouring to another, the chain selects a vertex x and a colour c uniformly at random. The vertex x is re-coloured with c if and only if this results in a proper colouring. Provided that $k \geq \Delta + 2$, where Δ is the maximum degree of the graph, the set of proper colourings is connected and this Markov chain converges to the uniform distribution over all proper colourings of G.

Much effort has gone into determining how many steps of the chain are necessary before it is "close to" the uniform distribution. Using a Markov chain for sampling

* This work was partially supported by the EPSRC grant GR/R44560/01 "Analysing Markov-chain based random sampling algorithms" and by the IST Programme of the EU under contract numbers IST-1999-14186 (ALCOM-FT) and IST-1999-14036 (RAND-APX). Part of the work was done during a visit to the Isaac Newton Institute.

is generally only useful if it is *rapidly mixing*, i.e., when the expected number of steps needed to get within ε of the stationary distribution is bounded above by a polynomial in n and $\log \varepsilon^{-1}$, where n is a measure of the size of each configuration.

Jerrum [9], and, independently, Salas and Sokal [16] proved Glauber dynamics mixes rapidly on any graph provided the number of available colours is at least 2Δ. Vigoda [17] demonstrated rapid mixing for an arbitrary graph provided there are at least $\frac{11}{6}\Delta$ colours. This argument was somewhat indirect as a chain that allowed small clusters to be recoloured was examined first, then the comparison method of Diaconis and Saloff-Coste [5] was used to infer rapid mixing of the Glauber chain.

For more specialized cases, Bubley, Dyer and Greenhill [4] gave a proof of rapid mixing in the case of 5-colouring graphs with maximum degree 3. Their method relied on the method of path coupling, using a computer to solve a large number of linear programming problems (in particular, transportation problems) to establish the necessary couplings. The method was also extended to 7-colourings of triangle-free 4-regular graphs. To date, there is no known "proof by hand" of these results (and none seems forthcoming, due to the number of linear constraints necessary to find the required couplings).

Graph colourings have arisen as configurations in several models from statistical physics. Proper graph colourings, for example, correspond to configurations in the so-called anti-ferromagnetic Potts model at zero temperature. The Potts model is a generalization of the Ising model, in which (proper and non-proper) 2-colourings of graphs are used to model and examine properties of magnetic materials. There is a correspondence between proper 3-colourings of regions of $\mathbb{Z}^2$ and configurations in the "6-vertex ice model," one of the many models of the structure of water when it freezes [2]. More generally, there is a bijection between proper 3-colourings of certain 4-regular graphs and Eulerian orientations of such graphs, where an *Eulerian orientation* is an assignment of directions to the edges so that there are the same number of incoming and outgoing edges at each vertex.

Mihail and Winkler [13] gave a polynomial-time algorithm for sampling Eulerian orientations. Their algorithm is based on a reduction to sampling perfect matchings. It allows one to sample 3-colourings of any 4-regular graph. The method can also be extended to graphs such as rectangular regions of $\mathbb{Z}^2$.

In this paper, we are concerned with a particular algorithm, namely *Glauber dynamics,* which is often used to sample colourings. Luby, Randall, and Sinclair [10] showed that Glauber dynamics is rapidly mixing when the graph is a rectangular region in $\mathbb{Z}^2$ with the restriction that the colours of the vertices on the boundary of the region are fixed and cannot be altered. Their proof uses the bijection between 3-colourings and Eulerian orientations. The boundary has the effect of fixing the direction of certain edges in the orientation.

They first considered a different Markov chain for sampling proper colourings of their graph. The Markov chain is similar to the single-site Glauber-dynamics chain, but it contains extra moves which recolour a linear sequence of vertices in one step (see Section 5 for details). They used coupling to show that their chain is rapidly mixing. Randall and Tetali [15] later applied the comparison method of Diaconis and Saloff-Coste [5] to prove that Glauber dynamics is also rapidly mixing.

Luby et al. left open the question of whether Glauber dynamics is rapidly mixing in the absence of the fixed boundary. The free-boundary problem was considered by Madras and Randall [11] as a possible application of their new *decomposition* *method*, but the application contained an error [12]. In [12] they stated that the free-boundary problem was still open.

In this paper, we show that the Glauber-dynamics Markov chain is rapidly mixing in the free boundary case. Like Luby et al., we first analyse a variant of the Glauber-dynamics chain with additional moves. We use the path-coupling method of Bubley and Dyer [3] to show that our variant is mixing. Finally, like Randall and Tetali, we use comparison to show that the original chain (that uses only transitions of Glauber dynamics) is rapidly mixing.

The paper is outlined as follows: In Section 2 we define the notation that we use for Markov chains and 3-colourings. Section 3 contains a description of the Glauber-dynamics Markov chain $\mathcal{M}_{\mathrm{Gl}}$ that is the main interest of this paper. In Section 4, we review the path-coupling method, define a metric on the set of proper 3-colourings of the region, and describe a geometric method of interpreting this metric which is useful in our analysis. We feel this section can be of independent interest, and we found it to be a necessary part of our analysis to demonstrate the path-coupling result.

Section 5 describes the variant chain that we prove is rapidly mixing, and Section 6 contains the details of that path-coupling argument. In particular, Subsection 6.A defines a coupling under which neighbouring pairs of colourings are likely to move together. The proof of this fact is relatively straightforward, once the transition probabilities of the variant chain have been established. The past difficulty of using path coupling in the free-boundary case may have been due, in large part, to not being able to discover a good candidate Markov chain. The more interesting part of the path-coupling argument uses the geometric features of the metric given in Subsection 4.B. These features are relevant in Subsection 6.B when we prove that the coupled (variant) Markov chain "moves around" sufficiently easily in the set $\Omega \times \Omega$, where Ω is the set of all proper 3-colourings of the region in $\mathbb{Z}^2$. Finally, Section 7 reviews the comparison method of Diaconis and Saloff-Coste and applies the method to give a bound on the mixing time of the Glauber-dynamics chain $\mathcal{M}_{\mathrm{Gl}}$.

2. DEFINITIONS

Let $\mathcal{M}$ be an ergodic Markov chain with a finite state space Ω. Let P be the transition matrix of $\mathcal{M}$, and let π be its stationary distribution. The *mixing time* of $\mathcal{M}$ is the time that it takes to converge to π. This is measured in terms of the distance between the distribution at time t and the stationary distribution. For $x, y \in \Omega$ and $t \geq 1$, let $P^t(x,y)$ denote the t-step probability of going from x to y. The *total variation distance* of $\mathcal{M}$ at time t is

$$\|P^t, \pi\|_{tv} = \max_{x\in\Omega} \frac{1}{2} \sum_{y\in\Omega} |P^t(x,y) - \pi(y)|.$$

For $\varepsilon > 0$, the *mixing time* $\tau(\varepsilon)$ of $\mathcal{M}$ is given by

$$\tau(\varepsilon) = \min\{t : \|P^{t'}, \pi\|_{tv} \leq \varepsilon, \forall t' \geq t\}.$$

$\mathcal{M}$ is said to be *rapidly mixing* if its mixing time is bounded above by a polynomial in n and $\log \varepsilon^{-1}$ where n is a measure of the size of each configuration in Ω.

In this work, we focus on induced subgraphs of $\mathbb{Z}^2$ that are rectangular in shape. In particular, let G be the graph corresponding to an $m \times n$ grid. Dyer, Goldberg, and Jerrum [7] have recently examined the special case when $m = 1$ (where the graph G is a path on n vertices). In this case, they show that the Glauber-dynamics Markov chain (as defined in the next section) mixes in time $\Theta(n^3 \log n \log(\varepsilon^{-1}))$ by providing upper and lower bounds on the mixing time that differ by a constant factor.

With the above result in mind, we will assume without loss of generality that $2 \leq m \leq n$ (if this does not hold, then it can be achieved by rotating the figure by 90°).

The vertex set of G is $V = \{v_{i,j} \mid i \in [1, m], j \in [1, n]\}$. The edge set of G is

$$E = \bigcup_{i \in [1,m], j \in [1,n-1]} \{(v_{i,j}, v_{i,j+1})\} \cup \bigcup_{i \in [1,m-1], j \in [1,n]} \{(v_{i,j}, v_{i+1,j})\}.$$

We will consider proper 3-colourings of G using colours $0, 1$ and 2. That is, we will consider 3-colourings in which adjacent vertices do not receive the same colour. Let Ω be the set of all proper 3-colourings of G. It is clear that Ω is non-empty, as G is a bipartite graph so we can in fact colour it using just two colours.

For ease of visualizing elements of Ω, we will draw each colouring as an $m \times n$ rectangle divided into squares of unit area. The vertex $v_{i,j}$ corresponds to the square in row i and column j, and this is labelled with one of the three colours from the set $\{0, 1, 2\}$. Row 1 is the top row and column 1 is the leftmost column. See Figure 1.

2	0	1	0	2	1
0	1	0	1	0	2
1	0	2	0	2	1
2	1	0	2	1	2
1	0	2	1	2	0
2	1	0	2	0	1

Fig. 1. A proper 3–colouring of a 6×6 region of $\mathbb{Z}^2$.

3. GLAUBER DYNAMICS

Let Ω be the set of all proper 3-colourings of G, as defined in Section 2. The *Glauber-dynamics* Markov chain, which we will denote $\mathcal{M}_{\mathrm{Gl}}$, has state space Ω. To move from one colouring to another, this chain selects a vertex x and a colour c uniformly at random. The vertex x is re-coloured with c if and only if this results in a proper colouring.

In what follows, we use notation like "$t \in_u T$" to mean that t is chosen *uniformly at random* from the set T. Formally, the Glauber chain iterates these steps:

One step of the Markov chain $\mathcal{M}_{\mathrm{Gl}}$:

1. Pick $x \in_u V(G)$ and $c \in_u \{0, 1, 2\}$.
2. If x has no neighbours coloured c, then obtain σ' from σ by recolouring x with c.
3. Otherwise (if x has at least one neighbour coloured c), set $\sigma' = \sigma$.

Let P_{Gl} denote the transition matrix of $\mathcal{M}_{\mathrm{Gl}}$. From the description of $\mathcal{M}_{\mathrm{Gl}}$, we deduce that

$$P_{\mathrm{Gl}}(\sigma_1, \sigma_2) = \begin{cases} \frac{1}{3mn} & \text{if } \sigma_1, \sigma_2 \text{ differ on a single vertex,} \\ 0 & \text{if } \sigma_1, \sigma_2 \text{ differ on two or more vertices, and} \\ 1 - \sum_{z \neq \sigma_1} P_{\mathrm{Gl}}(\sigma_1, z) & \text{if } \sigma_1 = \sigma_2. \end{cases}$$

Note that for every σ we have $P_{\mathrm{Gl}}(\sigma, \sigma) \geq 1/3$. This, together with Lemma 4.4 below (which shows that $\mathcal{M}_{\mathrm{Gl}}$ is irreducible), implies that $\mathcal{M}_{\mathrm{Gl}}$ is ergodic. Since P_{Gl} is symmetric, the chain $\mathcal{M}_{\mathrm{Gl}}$ is reversible, and the stationary distribution π is uniform. We can now state our main theorem:

Theorem 3.1. *The Markov chain $\mathcal{M}_{Gl}$ is rapidly mixing. In particular, the mixing time $\tau_{\mathcal{M}_{Gl}}(\varepsilon)$ of $\mathcal{M}_{Gl}$ satisfies*

$$\tau_{\mathcal{M}_{Gl}}(\varepsilon) \in \mathrm{O}\big(m^4 n^9 \log(\varepsilon^{-1})\big) .$$

More details about the constants involved in our bound for $\tau_{\mathcal{M}_{\mathrm{Gl}}}(\varepsilon)$ are given in the proof.

4. PATH COUPLING AND A METRIC ON COLOURINGS

We will use the path-coupling method of Bubley and Dyer [3] to analyse a variant of the Markov chain $\mathcal{M}_{\mathrm{Gl}}$. We start with a brief introduction to this method.

A. Path coupling

Coupling is a popular method for analysing mixing times of Markov chains. A (Markovian) coupling for a Markov chain $\mathcal{M}$ with state space Ω is a stochastic process (X_t, Y_t) on $\Omega \times \Omega$ such that each of (X_t) and (Y_t), considered marginally, is a faithful copy of $\mathcal{M}$. The coupling lemma (see, for example, Aldous [1]) states that the total variation distance of $\mathcal{M}$ at time t is bounded above by $\Pr(X_t \neq Y_t)$. The *path-coupling* method, introduced in [3], is a powerful method for finding couplings. The idea is that one can find a coupling on a subset S of $\Omega \times \Omega$ and extend this to a coupling on $\Omega \times \Omega$. The following theorem, adapted from [8], summarizes the path-coupling method.

Theorem 4.1. **[3, 8]** *Let S be a relation $S \subseteq \Omega^2$ such that S has transitive closure Ω^2. Let $\phi : S \to \mathbb{N}$ be a "proximity function" defined on pairs in S. We use*

ϕ to define a function Φ on Ω^2 as follows: For each pair $(\omega, \omega') \in \Omega^2$, let

$$\Phi(\omega, \omega') = \min_{\omega_0,\ldots,\omega_k} \sum_{i=0}^{k-1} \phi(\omega_i, \omega_{i+1}),$$

where the minimum is over all paths $\omega = \omega_0, \ldots, \omega_k = \omega'$ such that, for all $i \in [0, k-1]$, $(\omega_i, \omega_{i+1}) \in S$. Now suppose (X_t, Y_t) is a coupling for $\mathcal{M}$ defined over all pairs in S. Suppose that for all $(\sigma_1, \sigma_2) \in S$, we have

$$\mathbb{E}\left[\Phi(X_{t+1}, Y_{t+1}) \mid (X_t, Y_t) = (\sigma_1, \sigma_2)\right] \leq \beta \Phi(\sigma_1, \sigma_2),$$

where $\beta \leq 1$. Let D be the maximum value that Φ achieves on Ω^2 and suppose that there exists $\alpha > 0$ such that for all $(\sigma_1, \sigma_2) \in \Omega^2$,($\sigma_1 \neq \sigma_2$),

$$\Pr(\Phi(X_{t+1}, Y_{t+1}) \neq \Phi(\sigma_1, \sigma_2) \mid (X_t, Y_t) = (\sigma_1, \sigma_2)) \geq \alpha.$$

Then $\tau(\varepsilon) \leq \lceil eD^2/\alpha \rceil \lceil \log(\varepsilon^{-1}) \rceil$. (A better bound is also given for the case $\beta < 1$, but we will not use that result here.)

B. A metric on colourings

In our application of path coupling, the state space Ω will be the set of proper 3-colourings of G as defined in Section 2. S will be the set of edges in the Markov kernel of $\mathcal{M}_{\mathrm{Gl}}$. That is, $S = \{(\sigma_1, \sigma_2) \in \Omega^2 \mid \sigma_1 \neq \sigma_2 \text{ and } P_{\mathrm{Gl}}(\sigma_1, \sigma_2) > 0\}$. Equivalently, S is the set of pairs (σ_1, σ_2) which differ on the colour of a single vertex.

We show in Lemma 4.4 below that the transitive closure of S is Ω^2. We define the proximity function ϕ as follows: For all $(\sigma_1, \sigma_2) \in S$, $\phi(\sigma_1, \sigma_2) = 1$. Following Theorem 4.1, we define the function Φ on Ω^2 as follows. For each pair $(\sigma_1, \sigma_2) \in \Omega^2$, let

$$\Phi(\sigma_1, \sigma_2) = \min_{\omega_0,\ldots,\omega_k} \sum_{i=0}^{k-1} \phi(\omega_i, \omega_{i+1}) = \min_{\omega_0,\ldots,\omega_k} k,$$

where the minimum is over all paths $\sigma_1 = \omega_0, \ldots, \omega_k = \sigma_2$ such that, for all $i \in [0, k-1]$, $(\omega_i, \omega_{i+1}) \in S$. Thus $\Phi(\sigma_1, \sigma_2)$ is the minimum number of steps of $\mathcal{M}_{\mathrm{Gl}}$ needed to get from σ_1 to σ_2.

In Lemma 4.6 we show that the maximum value D that Φ achieves on Ω^2 is in $\mathrm{O}(mn^2)$. We can prove Lemmas 4.4 and 4.6 below without studying the geometric structure of Φ. Nevertheless, we need to establish some properties of Φ in order to determine the bound α when we apply Theorem 4.1. Thus, we introduce this structure now.

Recall that the $m \times n$ grid G is a bipartite graph. Let b be a function from the vertex set $V(G)$ to $\{0, 1\}$ which records parity: If the edge-distance from a vertex w to $v_{1,1}$ is even, then $b(w) = 0$. Otherwise, $b(w) = 1$. Every 3-colouring $\sigma \in \Omega$ is a function from the vertex set $V(G)$ to $\{0, 1, 2\}$. A *height function* [2, 10]

h corresponding to a proper colouring σ is a function $h : V(G) \to \mathbb{Z}$ satisfying the following properties[1]:

1. For every vertex $v \in V(G)$, $h(v) \equiv b(v) \pmod 2$.
2. For every vertex $v \in V(G)$, $h(v) \equiv \sigma(v) \pmod 3$.
3. For every pair of adjacent vertices v and w of G, $|h(v) - h(w)| = 1$.

2	0	1	0	2	1
0	1	0	1	0	2
1	0	2	0	2	1
2	1	0	2	1	2
1	0	2	1	2	0
2	1	0	2	0	1

2	3	4	3	2	1
3	4	3	4	3	2
4	3	2	3	2	1
5	4	3	2	1	2
4	3	2	1	2	3
5	4	3	2	3	4

Fig. 2. The colouring from Figure 1 and one of its corresponding height functions.

It is well known and easy to check that every proper colouring σ has one height function corresponding to each value of $h(v_{1,1})$ satisfying Properties 1 and 2. See [2, 10] for more details. For an example, see Figure 2, where the colouring from Figure 1 is reproduced and is shown with one of its possible related height functions. The rules that define height functions imply that, given a pair of height functions h, h' for the same colouring σ, it is possible to obtain h' by (simultaneously) adding or subtracting a multiple of six from all the values of h.

Let $H(\sigma)$ denote the set of height functions corresponding to colouring σ. The distance between two height functions, h_1 and h_2, is given by

$$d(h_1, h_2) = \sum_{v \in V(G)} |h_1(v) - h_2(v)|.$$

We define a relation S_H on the set of all pairs of height functions. The pair (h_1, h_2) is in S_H if and only if h_1 can be transformed into h_2 by applying one of of two basic *height transformations.* Intuitively, a height transformation either takes a local maximum of one height function, and pushes it down by two, or it takes a local minimum and pushes it up by two. Formally, the height transformations are defined as follows

1. Suppose that $h_1(v)$ is a local maximum in the sense that for all neighbours w of v, $h_1(v) > h_1(w)$. Transform h_1 into h_2, which is defined to be the same as h_1 except that $h_2(v) = h_1(v) - 2$.
2. Suppose that $h_1(v)$ is a local minimum in the sense that for all neighbours w of v, $h_1(v) < h_1(w)$. Transform h_1 into h_2, which is defined to be the same as h_1 except that $h_2(v) = h_1(v) + 2$.

[1]Most applications of height functions do not include Property 1 (parity), but this will be useful for us.

Suppose that the decreasing height transformation transforms h_1 into h_2. Let σ_1 be the colouring corresponding to h_1 and let σ_2 be the colouring corresponding to h_2. Then $(\sigma_1, \sigma_2) \in S$ and $\sigma_2(v) = \sigma_1(v)+1 \pmod 3$. Similarly, if the increasing height transformation transforms h_1 into h_2 and σ_1 is the colouring corresponding to h_1 and σ_2 is the colouring corresponding to h_2 then $(\sigma_1, \sigma_2) \in S$ and $\sigma_2(v) = \sigma_1(v) - 1 \pmod 3$. (Note the seeming disparity, a *decreasing* height transformation corresponds to *increasing* the colour of the vertex by one $\pmod 3$, and an *increasing* height transformation corresponds to *decreasing* the colour of the vertex by one $\pmod 3$. This follows because each height transformation alters the height of a point by two.)

The following observation follows from the first condition in our definition of height function.

Observation 4.2. *For any two height functions, h_1 and h_2, $d(h_1, h_2) \equiv 0 \pmod 2$.*

Observation 4.2 implies that the value k in the next lemma is an integer, so the lemma makes sense.

Lemma 4.3. *Suppose that h and h^* are height functions. Let $k = d(h, h^*)/2$. Then there is sequence of height functions $h = h_0, \ldots, h_k = h^*$ such that for all $i \in \{0, \ldots, k-1\}$, h_{i+1} can be obtained from h_i by a height transformation.*

Proof. The lemma is proved by induction on $d(h, h^*)$. It is vacuously true for the base case when $d(h, h^*) = 0$. In this case h and h^* are identical.

Suppose now that $d(h, h^*) > 0$. We will show that there is a height transformation transforming h into some h' such that $d(h', h^*) = d(h, h^*) - 2$. The lemma then follows by induction.

Suppose, without loss of generality, there is some vertex v such that $h(v) > h^*(v)$ and let us choose such a v such that $h(v)$ is as large as possible. We will show that v is a local maximum of h.

Let w be an arbitrary neighbour of v. If $h(w) > h^*(w)$ then $h(v) > h(w)$ by the maximality of v (and $h(v) \neq h(w)$). Otherwise, $h(w) \leq h^*(w)$. So

$$h(v) \geq h^*(v) + 1 \geq h^*(w) \geq h(w),$$

so $h(v) > h(w)$ and v is a local maximum of h.

Since $h(v) > h^*(v)$ and all height functions satisfy the parity constraint, $h(v) \geq h^*(v) + 2$. Pushing h down at v is the required height transformation of h into h' such that $d(h', h^*) = d(h, h^*) - 2$. ∎

Lemma 4.4. *The transitive closure of S is Ω^2.*

Proof. Let σ and σ^* be colourings in Ω. Select height functions $h \in H(\sigma)$ and $h^* \in H(\sigma^*)$. By Lemma 4.3, there is a sequence of height functions $h = h_0, \ldots, h_k = h^*$ such that for all $i \in \{0, \ldots, k-1\}$, h_{i+1} can be obtained from h_i by a height transformation. Thus, if we look at the sequence $\omega_0, \ldots, \omega_k$ of colourings corresponding to these height functions, we have, for all $i \in \{0, \ldots, k-1\}$, $(\omega_i, \omega_{i+1}) \in S$. Also, $\omega_0 = \sigma$ and $\omega_k = \sigma^*$. ∎

Lemma 4.5. *For any pair of colourings* $(\sigma_1, \sigma_2) \in \Omega^2$,

$$\Phi(\sigma_1, \sigma_2) = \frac{1}{2} \min_{\substack{h_1 \in H(\sigma_1),\\ h_2 \in H(\sigma_2)}} d(h_1, h_2).$$

Proof. First, we show

$$\Phi(\sigma_1, \sigma_2) \geq \frac{1}{2} \min_{\substack{h_1 \in H(\sigma_1),\\ h_2 \in H(\sigma_2)}} d(h_1, h_2).$$

Suppose

$$\Phi(\sigma_1, \sigma_2) = \sum_{i=0}^{k-1} \phi(\omega_i, \omega_{i+1}),$$

where $\sigma_1 = \omega_0$, $\sigma_2 = \omega_k$, and, for all $i \in [0, k-1]$, $(\omega_i, \omega_{i+1}) \in S$. Let h'_0 be any height function in $H(\omega_0)$. For $i \in [0, k-1]$, look at the single-site recolouring needed to transform ω_i into ω_{i+1}. Let h'_{i+1} be the height function derived from h'_i by performing the corresponding height transformation. It follows from the definition of a height function that $h'_{i+1} \in H(\omega_{i+1})$. Now

$$\begin{aligned}
\min_{\substack{h_1 \in H(\sigma_1),\\ h_2 \in H(\sigma_2)}} d(h_1, h_2) &\leq d(h'_0, h'_k) \\
&= \sum_{v \in V(G)} |h'_0(v) - h'_k(v)| \\
&\leq \sum_{i=0}^{k-1} d(h'_i, h'_{i+1}) \\
&= k \cdot 2 = \Phi(\sigma_1, \sigma_2) \cdot 2.
\end{aligned}$$

Next, we show

$$\Phi(\sigma_1, \sigma_2) \leq \frac{1}{2} \min_{\substack{h_1 \in H(\sigma_1),\\ h_2 \in H(\sigma_2)}} d(h_1, h_2).$$

Choose any $h \in H(\sigma_1)$ and $h^* \in H(\sigma_2)$. Let $k = d(h, h^*)/2$. Observation 4.2 guarantees that k is an integer. Lemma 4.3 shows that there is a sequence $h = h_0, \ldots, h_k = h^*$ such that for all $i \in \{0, \ldots, k-1\}$, h_{i+1} can be obtained from h_i by a height transformation. Thus, if we look at the sequence $\omega_0, \ldots, \omega_k$ of colourings corresponding to these height functions, we have, for all $i \in \{0, \ldots, k-1\}$, $(\omega_i, \omega_{i+1}) \in S$. Since $\omega_0 = \sigma_1$ and $\omega_k = \sigma_2$, we have $\Phi(\sigma_1, \sigma_2) \leq k$. Thus, $\Phi(\sigma_1, \sigma_2) \leq d(h, h^*)/2$. ∎

Lemma 4.6. *For every pair* $(\sigma_1, \sigma_2) \in \Omega^2$, $\Phi(\sigma_1, \sigma_2) \leq 2mn^2$.

Proof. By Lemma 4.5,

$$\Phi(\sigma_1, \sigma_2) = \frac{1}{2} \min_{\substack{h_1 \in H(\sigma_1),\\ h_2 \in H(\sigma_2)}} d(h_1, h_2).$$

Choose $h_1 \in H(\sigma_1)$ such that $|h_1(v_{1,1})| \leq 2$. This is possible because $h_1(v_{1,1}) \equiv 0 \pmod 2$ from the first property of height functions. Similarly, choose $h_2 \in H(\sigma_2)$ such that $|h_2(v_{1,1})| \leq 2$. Since heights of adjacent vertices differ by at most 1, every vertex v has $h_1(v)$ and $h_2(v)$ in the range $\{-(n+m), n+m\}$ so $|h_1(v) - h_2(v)| \leq 2(n+m)$. The result follows, since there are nm vertices and $nm(n+m) \leq 2mn^2$. ∎

Lemmas 4.4 and 4.6 show that the distance metric Φ is suitable for using path coupling to show rapid mixing of the chain $\mathcal{M}_{\mathrm{Gl}}$ with state space Ω. Unfortunately, we do not know of a direct coupling or path coupling that works for the chain $\mathcal{M}_{\mathrm{Gl}}$. No known coupling satisfies the "contraction" condition in Theorem 4.1. In particular, for known couplings, there are pairs of colourings $(\sigma_1, \sigma_2) \in S$ for which

$$\mathbb{E}\left[\Phi(X_{t+1}, Y_{t+1}) \mid (X_t, Y_t) = (\sigma_1, \sigma_2)\right] > \Phi(\sigma_1, \sigma_2),$$

so Theorem 4.1 does not apply. In the next section, we circumvent this difficulty by defining a new Markov chain $\widetilde{\mathcal{M}}$ with state space Ω. We still use the distance metric Φ we have just established to prove the rapid mixing of $\widetilde{\mathcal{M}}$. Using the comparison method of Diaconis and Saloff-Coste (see Subsection 7.A), we can then infer a bound on the mixing rate of $\mathcal{M}_{\mathrm{Gl}}$.

5. A MARKOV CHAIN WITH ADDITIONAL MOVES

We add some additional transitions to the chain $\mathcal{M}_{\mathrm{Gl}}$ to get a new chain $\widetilde{\mathcal{M}}$. This type of additional move, called a "tower move", is adapted from the work of Luby, Randall, and Sinclair [10] on the fixed-boundary case. Tower moves were also used in Madras and Randall [11].

Definition. *Let σ be a 3-colouring in Ω. A* right-going tower *in σ is a sequence of vertices $v_{i,j}, \ldots, v_{i,j+h}$ satisfying*

- *$h \geq 1$, and*
- *for every vertex $v_{i,j+\ell}$ with $\ell \in [0, h-1]$ (i.e. every vertex other than the last in the sequence) the colour $\sigma(v_{i,j+\ell+1})$ is unique amongst the colours of the neighbours of $v_{i,j+\ell}$ (i.e. $v_{i,j+\ell}$ has no neighbours with colour $\sigma(v_{i,j+\ell+1})$ other than $v_{i,j+\ell+1}$ itself), and*
- *all of the neighbours of $v_{i,j+h}$ have the same colour.*

The start vertex *of this tower is $v_{i,j}$ and the* start colour *of the tower is $\sigma(v_{i,j+1})$. The* end vertex *of this tower is $v_{i,j+h}$. The* length *of the tower is $h+1$. The* tower move *corresponding to this tower generates a new colouring σ' which is the same as σ except that the vertices in the tower are recoloured. In particular, every vertex $v_{i,j+\ell}$ with $\ell \in [0, h-1]$ receives the colour $\sigma'(v_{i,j+\ell}) = \sigma(v_{i,j+\ell+1})$. Also, $\sigma'(v_{i,j+h})$ is defined to be whichever colour is not in $\{\sigma(v_{i,j+h}), \sigma(v_{i,j+h-1})\}$. Towers going in other directions (left, up, and down) are defined analogously.*

Note that there is at most one tower corresponding to a given (start) vertex and colour.

2	0	1	0	2	1
0	1	0	1	0	2
1	0	2	0	2	1
2	1	0	2	1	2
1	0	2	1	2	0
2	1	0	2	0	1

2	0	1	0	2	1
0	1	0	1	0	2
1	0	2	0	2	1
2	1	0	2	1	2
1	2	1	0	2	0
2	1	0	2	0	1

Fig. 3. A (right-going) tower move of length 3.

See Figure 3 for an example of a right-going tower. The tower shown in the left picture has start vertex $v_{5,2}$, start colour 2, end vertex $v_{5,4}$, and length 3. The tower move transforms the left picture into the right one. Figure 4 shows the action of a tower move in one of the height functions that corresponds to the colouring in Figure 3. Recalling our remark from Subsection 4.B, note that the colours (in Figure 3) of the three vertices decrease by one (mod 3), while the heights (in Figure 4) of the vertices in the height function increase by two.

Remark. Note that a tower move can be simulated by a sequence of single-site recolouring steps. In particular, the vertices of a right-going tower $v_{i,j}, \ldots, v_{i,j+h}$ can be recoloured in the order $v_{i,j+h}, \ldots, v_{i,j}$. All of the colourings obtained along the way are proper. (This follows from the definition above.) The same thing holds for towers in other directions. Tower moves can also be "undone" by another tower move. In particular, suppose σ' is formed from σ by applying a right-going tower move $v_{i,j}, \ldots, v_{i,j+h}$. Then σ' has a left-going tower move $v_{i,j+h}, \ldots, v_{i,j}$, and applying this tower move with start vertex $v_{i,j+h}$ and start colour $\sigma'(v_{i,j+h-1})$ transforms σ' back to σ.

Finally, note that the vertices in a tower form a "monotonic" sequence modulo 3, i.e. the sequence of colours in the tower is a subsequence of consecutive elements from $\ldots, 0, 1, 2, 0, \ldots$ or $\ldots, 2, 1, 0, 2, \ldots$.

2	3	4	3	2	1
3	4	3	4	3	2
4	3	2	3	2	1
5	4	3	2	1	2
4	3	2	1	2	3
5	4	3	2	3	4

2	3	4	3	2	1
3	4	3	4	3	2
4	3	2	3	2	1
5	4	3	2	1	2
4	5	4	3	2	3
5	4	3	2	3	4

Fig. 4. The action of the tower move from Figure 3 in one of the corresponding height functions.

For the description of the new Markov chain, we need the following definition: We say that a tower is a *boundary tower* if one of the following holds.

- All of the vertices $v_{a,b}$ in the tower have $a = 1$ (all vertices are in row 1), or

- all of the vertices $v_{a,b}$ in the tower have $a = m$ (all vertices are in the last row), or
- all of the vertices $v_{a,b}$ in the tower have $b = 1$ (all vertices are in column 1), or
- all of the vertices $v_{a,b}$ in the tower have $b = n$ (all vertices are in the last column).

We also use the following notation: If v is a "corner" vertex of G (i.e. $v \in \{v_{1,1}, v_{1,n}, v_{m,1}, v_{m,n}\}$), then $p_v = 3/4$. Otherwise, $p_v = 1$.

With these definitions given above, we are now ready to describe the new Markov chain $\widetilde{\mathcal{M}}$ on Ω. Note that the chain has a self-loop at each step with probability at least $3/4$. This is for technical reasons — we want to be able to apply the comparison method of Theorem 7.1 to the chain, and the self-loops guarantee that all eigenvalues of the transition matrix of this new chain are at least $1/2$.

One step of the Markov chain $\widetilde{\mathcal{M}}$:

1. With probability $3/4$, set $\sigma' = \sigma$.
2. With the remaining probability

- **a.** Pick $x \in_u V(G)$ and $c \in_u \{0, 1, 2\}$.
- **b.** If x has no neighbours coloured c
  - (i) With probability p_x, obtain σ' from σ by colouring x using c.
  - (ii) With the remaining probability, set $\sigma' = \sigma$.
- **c.** If there is a length-t tower with start vertex x and start colour c
  - (i) If it is a boundary tower
    - A. With probability $1/(4t)$ derive σ' from σ by performing the tower move.
    - B. With the remaining probability set $\sigma' = \sigma$.
  - (ii) Otherwise
    - A. With probability $1/(2t)$ derive σ' from σ by performing the tower move.
    - B. With the remaining probability set $\sigma' = \sigma$.
- **d.** Otherwise, set $\sigma' = \sigma$.

Since this chain includes all moves of the Glauber chain $\mathcal{M}_{\mathrm{Gl}}$, we see that it is ergodic on Ω. From the remarks following the definition of towers, we see that $\widetilde{\mathcal{M}}$ is a reversible Markov chain. The transition matrix is symmetric, so from detailed balance we see that the stationary distribution is uniform on Ω.

6. ANALYSIS OF THE MODIFIED CHAIN

In this section, we use path coupling to bound the mixing time of the Markov chain $\widetilde{\mathcal{M}}$.

A. Defining the coupling and proving that Φ does not increase in expectation over coupled pairs

We now define a coupling over all pairs $(\sigma_1, \sigma_2) \in S$. The coupling that we use is the simplest one, namely we pick an ordered pair (x, c), where $x \in V(G)$ and $c \in \{0, 1, 2\}$ (as in Step 2(a) of $\widetilde{\mathcal{M}}$), and attempt to perform a recolouring of x using c in **both** σ_1 and σ_2, according to the transition probabilities of $\widetilde{\mathcal{M}}$. We use the optimal coupling, conditioned on the choice (x, c). The details of the coupling will be given in the lemmas below which consider various possibilities for x. The coupling is set up in such a way that whenever (x, c) corresponds to a transition that is possible in both σ_1 and σ_2 with the same probability then the move is performed jointly in σ_1 and σ_2 or not at all. Any move of this type will not change the value of Φ. Therefore, in the following series of lemmas, we only consider the moves which are different in σ_1 and σ_2.

Suppose that σ_1 and σ_2 differ solely at vertex v, so that $(\sigma_1, \sigma_2) \in S$. Without loss of generality in the analysis that follows, we may assume that $\sigma_1(v)$ has colour 0 and $\sigma_2(v)$ has colour 1. This means that all of the neighbours of v must be coloured with colour 2 in both σ_1 and σ_2. Let W be the set of neighbours of v in G. If v has a neighbour above it in the grid, we refer to this neighbour as w_u; a neighbour below v in the grid will be referred to as w_d; the left and right neighbours, if present, will be referred to as w_ℓ and w_r, respectively, so $W \subseteq \{w_u, w_d, w_\ell, w_r\}$.

We let σ_1' (respectively σ_2') denote the colouring obtained from σ_1 (respectively σ_2) after one step of the coupling. Through a series of lemmas we will prove that $\mathbb{E}[\Phi(\sigma_1', \sigma_2')] \le \Phi(\sigma_1, \sigma_2)$ or, equivalently, $\mathbb{E}[\Delta\Phi(\sigma_1, \sigma_2)] \le 0$ where $\Delta\Phi(\sigma_1, \sigma_2) = \Phi(\sigma_1', \sigma_2') - \Phi(\sigma_1, \sigma_2)$. This satisfies one of the requirements for using Theorem 4.1.

For a vertex x of G, we let $\Pr(x)$ denote the probability that x is chosen in Step 2(a) of $\widetilde{\mathcal{M}}$ and we let $\mathbb{E}[\Delta\Phi(\sigma_1, \sigma_2)|x]$ denote the expected change in the distance given that x was selected. We begin with an important observation about the role of towers in our analysis of the coupling:

Lemma 6.1. *Suppose that $(\sigma_1, \sigma_2) \in S$ and that σ_1 and σ_2 differ at vertex v. Let W be the set of neighbours of v. Then for all $x \notin \{v\} \cup W$,*

$$\Pr(x)\mathbb{E}[\Delta\Phi(\sigma_1, \sigma_2)|x] \le 0.$$

Proof. It is clear that if x is not in the row or column of some vertex in $\{v\} \cup W$ then a choice (x, c) will either correspond to a transition that is allowable in both σ_1 and σ_2 (with the same probability) or to a transition that is not allowable in either. Thus these choices have a zero net effect on the distance.

Now consider a right-going tower coming in towards w_d. This tower would have to end before w_d because of the fact that w_ℓ has the same colour as w_d. (Recall the definition of a tower.) Thus, this tower has the same start and end vertices in both σ_1 and σ_2, hence the corresponding tower move will contribute nothing to the change in distance. By similar arguments, we see that other towers have no effect on Φ (because they have the same start and end vertices and are therefore allowed with identical probabilities in both copies), except possibly those towers that start in the row or column containing v.

Now consider a right-going tower in the colouring σ_1 which starts to the left of w_ℓ in the row of v. If it ends to the left of w_ℓ then it will have the same effect in both copies. It cannot end to the right of v because all neighbours of v have the same colour (again, see the definition of a tower). Thus, it ends either at w_ℓ or at v.

Case 1: Suppose it ends at w_ℓ and has length $t > 1$. Now consider the same start vertex and start colour in copy σ_2. The same tower comes along to the right. It does not stop at w_ℓ because $\sigma_2(v)$ is different from $\sigma_1(v)$. Instead, it stops at vertex v and has length $t + 1$. Now consider what happens in the coupling. With probability $1/(k(t+1))$ the tower move is made in both copies (k is either 2 or 4 depending upon whether it is a boundary tower). In this case, the two copies couple (i.e. $\sigma_1' = \sigma_2'$), so the change to Φ is -1. With probability $1/(kt) - 1/(k(t+1))$ the tower move is made only in copy σ_1 and Φ increases by at most t. Noting that

$$(-1) \cdot \frac{1}{k(t+1)} + t \cdot \left(\frac{1}{kt} - \frac{1}{k(t+1)} \right) = 0$$

we see in this case $\mathbb{E}\left[\Delta\Phi(\sigma_1, \sigma_2)\right] \leq 0$.

Case 2: Suppose the tower ends at v. Then in copy σ_2 the corresponding tower ends at w_ℓ, so the analysis is the same as Case 1.

Towers going left, up, and down in the rows of columns containing v also have a neutral effect on Φ by similar arguments. ∎

The previous lemma shows us we now only have to examine what happens if we select either v or a neighbour of v in the coupling. We start with the easiest observation:

Lemma 6.2. *Suppose that $(\sigma_1, \sigma_2) \in S$ and that σ_1 and σ_2 differ at vertex v. Then*

$$\Pr(v)\mathbb{E}\left[\Delta\Phi(\sigma_1, \sigma_2)|v\right] = \frac{-p_v}{6mn}.$$

Proof. The choice $(v, 0)$ couples the pair with probability p_v and this reduces Φ by 1. The same thing happens for $(v, 1)$. Each of these choices of (v, c) occurs with probability $\frac{1}{12mn}$, so $\Pr(v)\mathbb{E}\left[\Delta\Phi(\sigma_1, \sigma_2)|v\right] = \frac{-2p_v}{12mn} = \frac{-p_v}{6mn}$. ∎

Lemma 6.3. *Suppose that $(\sigma_1, \sigma_2) \in S$ and that σ_1 and σ_2 differ at vertex v. Suppose that w is a neighbour of v and that w and v are not* both *on the boundary. Then*

$$\Pr(w)\mathbb{E}\left[\Delta\Phi(\sigma_1, \sigma_2)|w\right] \leq \frac{1}{24mn}.$$

Proof. The conditions in the hypothesis of the lemma imply that neither v nor w is a corner node, so both have degree at least 3. Recall that without loss of generality we take $\sigma_1(v) = 0$, $\sigma_2(v) = 1$ and $\sigma_1(w) = \sigma_2(w) = 2$. Also, by symmetry, we may assume that at most half of the neighbours of w (not counting v) are coloured 1. Therefore, at most one neighbour of w, other than v, is coloured 1 in σ_1 and σ_2. There are two cases to consider.

Case 1: All neighbours of w (other than v) are coloured 0.

In this case the only possible change to Φ is when the pair $(w, 1)$ is chosen. The probability that w gets recoloured 1 in copy σ_1 is $p_w = 1$. In copy σ_2 we have a tower with start vertex w and end vertex v and start colour 1. By the hypothesis on v and w, it cannot be a boundary tower. Since the length of this tower is 2, the tower move is accepted with probability $1/(2\cdot 2)$ (in which case we have $\sigma_1' = \sigma_2'$). Thus, Φ decreases by 1 with probability $1/4$, and increases by 1 with probability $3/4$. In this case the expected change to Φ is $1/(12mn) \cdot [(-1) \cdot 1/4 + (+1) \cdot 3/4] = 1/(24mn)$.

Case 2: Exactly one neighbour of w (other than v) is coloured 1.

Case 2A: Suppose w is an internal node of G, i.e. w has degree 4.

In this case, there will be no change to copy σ_2, since w has two neighbours coloured 1 in this colouring of G (and two neighbours coloured 0). The only possible change in σ_1 is if the pair $(w, 1)$ is chosen. There is some tower of length $t > 1$ having w as its start node (pointing towards the neighbour of w that is coloured 1). This tower move is accepted in σ_1 with probability $1/(2 \cdot t)$ as it is not a boundary tower. If performed, the increase in Φ is at most t. Therefore, since $\sigma_2' = \sigma_2$ in this case, we see that

$$\Pr(w)\mathbb{E}\left[\Delta\Phi(\sigma_1, \sigma_2)|w\right] \le \frac{1}{12mn} \cdot (+t) \cdot \frac{1}{2t} = \frac{1}{24mn}.$$

Case 2B: Suppose w is on the boundary. (Since w is on the boundary, but is not a corner vertex, it has degree 3.)

In this case the choice $(w, 1)$ causes a boundary tower of length $t \ge 2$ in σ_1. Performing this tower move contributes an expected increase of $1/(4 \cdot t) \times t = 1/4$. Similarly, the choice $(w, 0)$ causes a boundary tower in σ_2 which also contributes $1/4$. (Note that the $(w, 1)$ tower in σ_1 and the $(w, 0)$ tower in σ_2 both have w as the start vertex, and point in opposite directions to one another.) Thus, overall, we see that if w is on the boundary (and not a corner vertex) we have

$$\Pr(w)\mathbb{E}\left[\Delta\Phi(\sigma_1, \sigma_2)|w\right] \le \frac{1}{12mn} \cdot \left(\frac{1}{4} + \frac{1}{4}\right) = \frac{1}{24mn}.$$

We have therefore shown the lemma holds in all cases when v and w are not both on the boundary of G. ∎

The final two lemmas deal with the case when v and w both lie on the boundary of G.

Lemma 6.4. *Suppose that $(\sigma_1, \sigma_2) \in S$ and that σ_1 and σ_2 differ at boundary vertex v. Suppose that w, a neighbour of v, is on the boundary but is not a corner vertex.*

$$\Pr(w)\mathbb{E}\left[\Delta\Phi(\sigma_1, \sigma_2)|w\right] \le \frac{1}{16mn}.$$

Proof. If both neighbours of w, other than v, are coloured 0 then the only choice that has an effect on Φ is $(w, 1)$. This is exactly the same as in Case 1 of the proof of Lemma 6.3 except that the tower is a boundary tower. Therefore, the tower of length 2 in σ_2 gets recoloured with probability $1/(4 \cdot 2)$. The probability that w gets recoloured in σ_1 is $p_w = 1$. Hence, with probability $1/8$ we recolour both w in σ_1 and the tower in σ_2, so that $\sigma_1' = \sigma_2'$, decreasing Φ by 1. With probability

$1 - 1/8$ we recolour w in σ_1 only, increasing Φ by 1. Thus

$$\Pr(w)\mathbb{E}\left[\Delta\Phi(\sigma_1,\sigma_2)|w\right] \le \frac{1}{12mn} \cdot \left[(-1)\cdot\frac{1}{8} + (+1)\cdot\left(1-\frac{1}{8}\right)\right] = \frac{1}{16mn}.$$

If, on the other hand, one neighbour of w is coloured 1 and the other is coloured 0 then there are two towers, similar to Case 2B of the proof of Lemma 6.3. As in that case, colour 1 corresponds to a tower in σ_1 and no change in σ_2. Similarly, colour 0 corresponds to a tower in σ_2 and no change in σ_1. The difference here is that one of these towers is a boundary tower, and the other is not. The boundary tower (of length $t > 1$) contributes (at most) $(+t)\cdot 1/(4\cdot t) = 1/4$ to the change in Φ. On the other hand, the non-boundary tower (of some length $t^* > 1$) contributes $(+t^*)\cdot 1/(2\cdot t^*) = 1/2$. So for this neighbour w, the net expectation is

$$\Pr(w)\mathbb{E}\left[\Delta\Phi(\sigma_1,\sigma_2)|w\right] \le \frac{1}{12mn} \cdot \left(\frac{1}{4}+\frac{1}{2}\right) = \frac{1}{16mn}.$$

■

The final case to examine, where v is adjacent to a corner and w is a corner vertex, is handled by the following lemma.

Lemma 6.5. *Suppose that $(\sigma_1,\sigma_2) \in S$ and that σ_1 and σ_2 differ at vertex v. Suppose that w, a neighbour of v, is a corner vertex. Then*

$$\Pr(w)\mathbb{E}\left[\Delta\Phi(\sigma_1,\sigma_2)|w\right] \le \frac{1}{16mn}.$$

Proof. Without loss of generality, we can assume that the neighbour of w other than v is coloured 0.

The choice $(w,0)$ has no effect in σ_1. In σ_2 it could cause a boundary tower (of some length $t > 1$) with probability $1/(4\cdot t)$, increasing Φ by t. Hence this contributes $(+t)\cdot 1/(4\cdot t) = 1/4$ to the conditional expected change in Φ.

Consider the choice of ordered pair $(w,1)$. What happens in σ_1? With probability $p_w = 3/4$, w gets recoloured 1. With probability $1/4$, it stays the same. What happens in σ_2? There is a length-2 boundary tower ending at v. If it is accepted (with probability $1/8$) then v is recoloured 0 and w is recoloured 1. Otherwise nothing happens.

Consider the coupling. With probability $1/8$, we perform the tower move in σ_2 *and* recolour w in σ_1. This combination means that $\sigma_1' = \sigma_2'$ and, hence, decreases Φ by 1. With probability $3/4 - 1/8$, we recolour w in σ_1 *only.* This increases Φ by $+1$. Thus, the choice of $(w,1)$ contributes $(-1)\cdot 1/8 + (+1)\cdot(3/4-1/8) = 1/2$ to the change in Φ.

Overall, this means if w is a corner vertex, then

$$\Pr(w)\mathbb{E}\left[\Delta\Phi(\sigma_1,\sigma_2)|w\right] \le \frac{1}{12mn} \cdot \left(\frac{1}{4}+\frac{1}{2}\right) = \frac{1}{16mn}.$$

■

We can combine all the previous lemmas, to get the following result.

Theorem 6.6. *Suppose* $(\sigma_1, \sigma_2) \in S$. *After one step of the coupling, we have* $\mathbb{E}[\Delta\Phi(\sigma_1, \sigma_2)] \leq 0$.

Proof. Lemma 6.1 shows us that the contribution of choosing a vertex outside of $\{v\} \cup W$ will not increase Φ in expectation. Lemmas 6.2 and 6.3 deal with the contribution from all cases in which v is an internal node (a degree-4 node). In this case we have $p_v = 1$, so v contributes $-1/(6mn)$ to the change in Φ. Each neighbour of v contributes at most $1/(24mn)$ so the net contribution from v and its four neighbours is zero.

If v lies on the boundary of G, but is not a corner vertex, nor a neighbour of a corner vertex, Lemma 6.2 shows that v still contributes $-1/(6mn)$ to the expected change in Φ. Lemma 6.3 tells us that the internal neighbour of v contributes at most $1/(24mn)$, while Lemma 6.4 tells us that the two neighbours of v that also lie on the boundary contribute at most $1/(16mn)$ to the expected change in Φ. Overall, we have

$$\mathbb{E}[\Delta\Phi(\sigma_1, \sigma_2)] \leq \frac{-1}{6mn} + \frac{1}{24mn} + 2 \cdot \frac{1}{16mn} = 0.$$

If v is a corner vertex, Lemma 6.2 says that v contributes $-1/(8mn)$, since $p_v = 3/4$ in this case. Each of the two neighbours still contributes $1/(16mn)$ from Lemma 6.4 and 6.5, so the net contribution is still zero.

Finally, if v is a neighbour of a corner vertex it again contributes $-1/(6mn)$ (Lemma 6.2 where $p_v = 1$), the internal neighbour of v contributes $1/(24mn)$ (Lemma 6.3), and the other two neighbours contribute $1/(16mn)$ each (Lemmas 6.4 and 6.5). Putting all these observations together we have

$$\mathbb{E}[\Delta\Phi(\sigma_1, \sigma_2)] \leq \frac{-1}{6mn} + \frac{1}{24mn} + \frac{1}{16mn} + \frac{1}{16mn} = 0.$$

So we see that in all cases, the distance will not increase in expectation in one step for a pair $(\sigma_1, \sigma_2) \in S$. ■

Using the method of path coupling (Theorem 4.1) in Subsection 4.A, we can now give a bound on the mixing time of the Markov chain $\widetilde{\mathcal{M}}$. Recall from Lemma 4.6 that the maximum value D that Φ can take is at most $2mn^2$. In Subsection 6.B we show that the value α in Theorem 4.1 can be taken to be $1/(48mn^2)$. Theorem 4.1 then gives us the following.

Theorem 6.7. *The Markov chain* $\widetilde{\mathcal{M}}$ *on* Ω *is rapidly mixing, and*

$$\tau_{\widetilde{\mathcal{M}}}(\varepsilon) \leq \lceil e(2mn^2)^2 48mn^2 \rceil \lceil \log \varepsilon^{-1} \rceil \leq 193em^3n^6 \lceil \log \varepsilon^{-1} \rceil.$$

B. Bounding the variance of the coupling

In Subsection 6.A we defined a coupling on pairs in S. We have shown that the distance Φ does not increase in expectation over pairs in S, and the proof of the path-coupling theorem (Theorem 4.1) shows that Φ also does not increase in expectation over all pairs in Ω^2. This is not quite enough to ensure fast coupling.

In particular, we have not ruled out the possibility that we have rigged up some strange coupling, which keeps Φ fixed. Thus, we must now show that the coupling itself has some variance.

Formally, the path-coupling theorem (Theorem 4.1) requires us to find an $\alpha > 0$ such that for all $(\sigma_1, \sigma_2) \in \Omega^2$, where $\sigma_1 \neq \sigma_2$,

$$\Pr(\Phi(X_{t+1}, Y_{t+1}) \neq \Phi(\sigma_1, \sigma_2) \mid (X_t, Y_t) = (\sigma_1, \sigma_2)) \geq \alpha. \tag{6.1}$$

Note that it is not sufficient to establish Equation (6.1) just for pairs in S — we cannot simply add variances along a path, so we are forced to establish Equation (6.1) for every pair $(\sigma_1, \sigma_2) \in \Omega^2$. Doing so is made difficult by the fact that the coupling of an arbitrary pair (σ_1, σ_2) is very complicated — in particular it depends upon a particular geodesic path $\sigma_1 = \omega_0, \ldots, \omega_k = \sigma_2$ such that

$$\Phi(\sigma_1, \sigma_2) = \sum_{i=0}^{k-1} \phi(\omega_i, \omega_{i+1}),$$

and our coupling of (σ_1, σ_2) is obtained by extending this coupling pairwise along the path.

Our strategy for establishing Equation (6.1) consists of examining several cases. In each case, we will start with a pair $(\sigma_1, \sigma_2) \in \Omega^2$ such that $\sigma_1 \neq \sigma_2$. We will choose $h_1 \in H(\sigma_1)$ and $h_2 \in H(\sigma_2)$ such that

$$\Phi(\sigma_1, \sigma_2) = \frac{1}{2} d(h_1, h_2).$$

We will find a reasonably-likely move that transforms h_1 into h_1'. We will show that, however the coupling transforms h_2 into h_2', we will have $d(h_1', h_2') < d(h_1, h_2)$, so

$$\Phi(\sigma_1', \sigma_2') \leq \frac{1}{2} d(h_1', h_2') < \frac{1}{2} d(h_1, h_2) = \Phi(\sigma_1, \sigma_2),$$

where the first inequality follows from Lemma 4.5.

Alternatively, we might find a reasonably-likely move that transforms h_2 into h_2' such that, however the coupling transforms h_1 into h_1', we have $d(h_1', h_2') < d(h_1, h_2)$.

Without loss of generality, we can assume there is a vertex v such that $h_1(v) > h_2(v)$. Let $H = \max_u(h_1(u) - h_2(u))$ and let

$$R = \{u \mid h_1(u) - h_2(u) = H\}.$$

By construction, R is non-empty.

Case 1: $R = V$

Since $R = V$, we can observe from the definition of R that $h_1(v) - h_2(v)$ is independent of v. In particular, $h_1(v) - h_2(v) = H$ for all v. Furthermore, there is some $\delta \in \{+1, -1\}$ such that every $v \in V$ satisfies $\sigma_1(v) = \sigma_2(v) + \delta \pmod 3$. Let z be any local maximum in h_1 and let c be the colour that is not used at z or at its neighbours in h_1. z is also a local maximum in h_2 (since h_1 and h_2 are parallel) but c is used either at z or at all of its (two or more) neighbours in h_2. (If c were not used then the colour of z would be the same in σ_1 and σ_2, so we would have $\sigma_1 = \sigma_2$, which is not the case.) Now choose (z, c). This choice occurs with probability $1/(12mn)$. With probability $p_z \geq 3/4$, $h'_1(z) = h_1(z) - 2$. But $h'_2(z) = h_2(z)$ since c is used at z or at a neighbour of z in h_2. Thus, $d(h_1, h_2) - d(h'_1, h'_2) = 2$. In this case, it suffices to take any $\alpha \leq 1/(16mn)$.

Notation: For the remaining cases, we pick some $z \in R$ such that z has at least one neighbour in $\overline{R}$. Suppose z has degree $\Delta \in \{2, 3, 4\}$ in G, and let $k \geq 1$ denote the number of neighbours of z in $\overline{R}$. Note that all edges from z to $\overline{R}$ in h_1 go down (that is, the height decreases along these edges). Also, all edges from z to $\overline{R}$ in h_2 go up. Finally, if there is an edge between z and another vertex in R, then this edge goes in the same direction (up or down) in h_1 as in h_2.

Case 2: $k = \Delta$
In this case, z is a local maximum in h_1 and a local minimum in h_2. Let c be the colour that is not used at z or at its neighbours in h_1. Choose (z, c). This choice occurs with probability $1/(12mn)$. With probability $p_z \geq 3/4$, $h'_1(z) = h_1(z) - 2$. Since z is a local minimum in h_2, $h'_2(z) \geq h_2(z)$. Since $z \in R$ we have $h_1(z) > h_2(z)$. Furthermore, $h_1(z) = h_2(z) \pmod 2$. Now if $h_2(z) = h_1(z) - 2$ then the choice (z, c) will not move the height of z in h_2 (since the same choices moved $h_1(z)$ to $h_1(z) - 2 = h_2(z)$) so $h'_1(z) \geq h'_2(z)$. On the other hand, if $h_2(z) \leq h_1(z) - 4$ then certainly $h'_1(z) \geq h'_2(z)$. Thus, in all cases we have $h'_1(z) \geq h'_2(z)$. Therefore, $d(h_1, h_2) - d(h'_1, h'_2) \geq 2$. In this case, it suffices to take any $\alpha \leq 1/(16mn)$.

Case 3: $k = \Delta - 1$
Let r be the neighbour of z in R.

Suppose first that the edge between z and r goes down from z to r. Then z is a local maximum in h_1. Let c be the colour that is not used at z or at its neighbours in h_1. Choose (z, c). With probability $p_z \geq 3/4$, $h'_1(z) = h_1(z) - 2$. Consider the effect of this choice on h_2. The k neighbours of z in $\overline{R}$ are all above z in h_2, so the only way that $h'_2(z) < h_2(z)$ could happen is that a tower move could be made, i.e. if $k = 1$. If the tower move has length t, the probability that it is accepted is at most $1/(2t) \leq 1/4$. We do not know the details of how this is event is correlated with whether or not z is recoloured in h_1 — the details depend upon the coupling along the chosen geodesic chain from σ_1 to σ_2 in the application of the path-coupling theorem — but since the probability of recolouring in h_1 is at least $3/4$, for any correlation, there is a probability at least $1/2$ that z is recoloured in h_1, but the tower is rejected. Thus, we can take any $\alpha \leq 1/(24nm)$.

Suppose instead that the edge between z and r goes up from z to r. Then z is a local minimum in h_2, and the argument is similar.

Case 4: $k = \Delta - 2$
Let r_1, r_2 be the neighbours of z in R. If the edges (z, r_1) and (z, r_2) go in the same direction (either both go up or both go down), then the argument is the same as in Case 3. Therefore, assume that one edge goes down from z to r_1 and the other goes up from z to r_2.

Choose z, and choose the colour of r_2 in h_1. In h_1, this choice corresponds to a tower move starting at z, then going on to r_2 (and possibly continuing from there). Let t be the length of the tower. Recall that in any tower the sequence of colours forms a "monotonic" sequence modulo 3, i.e., it is either a subsequence of $\ldots, 0, 1, 2, 0, \ldots$ or it is a subsequence of $\ldots, 2, 1, 0, 2, \ldots$. Hence, the heights increase in h_1 along this tower. This means that the whole of the tower is above h_2 (since $h_1(z) > h_2(z)$). If the tower move is accepted, every vertex v in the tower will have $h'_1(v) = h_1(v) - 2$. Thus, $d(h_1, h_2) - d(h'_1, h_2) = 2t$.

Consider the effect of this choice on h_2. It is not possible to have $h'_2(z) < h_2(z)$ since z has at least two neighbours above it in h_2. It may be possible that the choice (z, c) induces some tower move in h_2 (via a tower with some length t' involving r_1). If the colour of z is changed in h_2 (via this tower move involving r_1), then $d(h_1, h_2) - d(h_1, h'_2) = 2t'$.

Note that the only vertex which could be involved in both the h_1 tower and the possible tower in h_2 is z. If both of these moves are possible, then z moves to the same colour in both copies, so $h'_1(z) \geq h'_2(z)$. If there is no tower move possible with the choice (z, c) in h_2, then $h'_2(z) = h_2(z)$. In either case, we always have $h'_1(z) \geq h'_2(z)$.

We conclude that, as long as the tower move in h_1 is accepted, $d(h_1, h_2) - d(h'_1, h'_2) \geq 2t$. The probability of accepting this move is at least $1/(4t) \geq 1/(4n)$. Thus, it suffices to take any $\alpha \leq 1/(48mn^2)$.

Case 5: $k = \Delta - 3$
In this case, $\Delta = 4$. Let r_1, r_2, and r_3 be the neighbours of z in R.

If the edges (z, r_1), (z, r_2) and (z, r_3) go in the same direction, then the argument is the same as in Case 3.

Suppose that the edges (z, r_1) and (z, r_2) go down from z and that the edge (z, r_3) goes up from z. As in Case 4 there is a tower move in h_1 involving z and r_3. Accepting the tower move moves the tower in h_1 down towards h_2, so if the tower has length t, we get $d(h_1, h_2) - d(h'_1, h_2) = 2t$. The colour of z cannot be changed in h_2 (because it has two neighbours above and two below), so it suffices to take any $\alpha \leq 1/(48mn^2)$.

A similar argument applies to the case in which edges (z, r_1) and (z, r_2) go up from z and that the edge (z, r_3) goes down from z. Here, h_2 contains a tower involving z and r_3. Raising this tower reduces the height-function distance. But the colour of z cannot change in h_1.

Cases 1–5 together show us that we may take $\alpha = 1/(48mn^2)$ in Theorem 4.1, giving the bound stated in Theorem 6.7 for the mixing time of the Markov chain $\widetilde{\mathcal{M}}$.

7. BACK TO GLAUBER DYNAMICS

In the last section, we showed that the "augmented" chain $\widetilde{\mathcal{M}}$ is rapidly mixing. Our original goal was to prove a result about the chain $\mathcal{M}_{\mathrm{Gl}}$ that only uses single vertex recolouring steps, not the larger tower moves that $\widetilde{\mathcal{M}}$ allows. The comparison method of Diaconis and Saloff-Coste [5] allows us to do this. We start by giving the background to the method.

A. The Comparison Method

The comparison method of Diaconis and Saloff-Coste [5] is useful for relating the mixing times of two similar Markov chains. Typically, this method is used in cases where we want to bound the mixing time of a Markov chain $\mathcal{M}$, but it is easier to analyse a new Markov chain $\widetilde{\mathcal{M}}$, which is obtained from $\mathcal{M}$ by adding some additional transitions, or by modifying some transition probabilities. The comparison method allows us to bound the mixing time of $\mathcal{M}$ using a bound for the mixing time of $\widetilde{\mathcal{M}}$.

The following is Proposition 4 of Randall and Tetali [15]. In the proposition, we use the notation $E(P)$ to refer to the edge-set of the Markov kernel of the chain with transition matrix P. That is, $E(P) = \{(x, y) \in \Omega^2 \mid x \neq y \text{ and } P(x, y) > 0\}$. We use the notation $|\gamma|$ to denote the length of a sequence γ of states.

Theorem 7.1. *[15, Proposition 4] Suppose that P and $\widetilde{P}$ are the transition matrices of two reversible Markov chains, $\mathcal{M}$ and $\widetilde{\mathcal{M}}$, both with state space Ω and stationary distribution π. For each pair $(x, y) \in E(\widetilde{P})$, define a path γ_{xy} which is a sequence of states $x = x_0, x_1, \ldots, x_k = y$ with $(x_i, x_{i+1}) \in E(P)$ for all i. For $(z, w) \in E(P)$, let $\Gamma(z, w) = \{(x, y) \in E(\widetilde{P}) : (z, w) \in \gamma_{xy}\}$. Let*

$$A = \max_{(z,w)\in E(P)} \left\{ \frac{1}{\pi(z)P(z, w)} \sum_{(x,y)\in\Gamma(z,w)} |\gamma_{xy}|\pi(x)\widetilde{P}(x, y) \right\}.$$

Suppose that the second-largest eigenvalue, λ_1, of $\widetilde{P}$ satisfies $\lambda_1 \geq 1/2$. Then for any $0 < \varepsilon < 1$

$$\tau_{\mathcal{M}}(\varepsilon) \leq \frac{4\log(1/(\varepsilon\pi_*))}{\log(1/(2\varepsilon))} A\tau_{\widetilde{\mathcal{M}}(\varepsilon)}$$

where $\pi_* = \min_{x\in\Omega} \pi(x)$.

The key point in using the comparison method is to choose paths between pairs $(x, y) \in E(\widetilde{P})$ so that the parameter A is “small”, i.e. A is bounded by a polynomial function of n, where as before, n is a measure of the size of each configuration in the state space Ω. Doing this implies that if $\widetilde{\mathcal{M}}$ is rapidly mixing, then so is $\mathcal{M}$.

B. Comparing the two chains

In this section, we will prove our main theorem, which is re-stated here for convenience.

Theorem 3.1. *The Markov chain $\mathcal{M}_{Gl}$ is rapidly mixing. In particular, the mixing time $\tau_{\mathcal{M}_{Gl}}(\varepsilon)$ of $\mathcal{M}_{Gl}$ satisfies*

$$\tau_{\mathcal{M}_{Gl}}(\varepsilon) \in \mathrm{O}\big(m^4 n^9 \log(\varepsilon^{-1})\big).$$

Proof. The self-loops in the definition of $\widetilde{\mathcal{M}}$ guarantee that its second-largest eigenvalue, λ_1, is at least $1/2$, thus Theorem 7.1 applies.

We first note that $|\Omega|$ (and hence $1/\pi_*$) is bounded above by 3^{mn} as this is the number of all (proper and improper) 3-colourings of the $m \times n$ region G. Thus, by Theorem 7.1 and Theorem 6.7,

$$\tau_{\mathcal{M}_{\mathrm{Gl}}}(\varepsilon) \leq \frac{4\,(\log(1/\varepsilon) + mn\log 3)}{\log(1/(2\varepsilon))} \times A \times \left(193em^3n^6\lceil\log\varepsilon^{-1}\rceil\right),$$

where A is the quantity defined in Theorem 7.1.

In particular, for each pair $(\sigma_1, \sigma_2) \in E(\widetilde{P})$, we must define a path $\gamma_{\sigma_1\sigma_2}$ which is a sequence of states $\sigma_1 = \omega_0, \ldots, \omega_k = \sigma_2$ such that, for all $i \in [0, k-1]$, $(\omega_i, \omega_{i+1}) \in S$. Then we let $\Gamma(z, w) = \{(\sigma_1, \sigma_2) \in E(\widetilde{P}) : (z, w) \in \gamma_{\sigma_1\sigma_2}\}$. Finally, A is defined to be the maximum over all $(z, w) \in S$ of the quantity

$$A_{z,w} = \frac{1}{\pi(z)P(z,w)} \sum_{(x,y)\in\Gamma(z,w)} |\gamma_{xy}|\pi(x)\widetilde{P}(x,y) = \frac{1}{P(z,w)} \sum_{(x,y)\in\Gamma(z,w)} |\gamma_{xy}|\widetilde{P}(x,y).$$

We will show that $A_{z,w} < (n+1)^2/8$, which proves the theorem. Recall from the remark following the definition of a tower in Section 5 that a tower move of length t may be simulated by a sequence of t single-site recolourings. (We recolour the vertices in order from the *end* vertex to the *start* vertex of the tower.) We will use this sequence to define the path $\gamma_{\sigma_1\sigma_2}$ for $(\sigma_1, \sigma_2) \in E(\widetilde{P})$. For $t \geq 2$, let

$$\Gamma^t(z, w) = \{(\sigma_1, \sigma_2) : \sigma_1 \text{ and } \sigma_2 \text{ differ by a length-}t\text{ tower move and } (z, w) \in \gamma_{\sigma_1\sigma_2}\}.$$

Also, let

$$\Gamma^1(z, w) = \{(\sigma_1, \sigma_2) : \sigma_1 \text{ and } \sigma_2 \text{ differ at a single vertex and } (z, w) \in \gamma_{\sigma_1\sigma_2}\},$$

so $\Gamma^1(z, w) = \{(z, w)\}$.

Then

$$A_{z,w} = \sum_{t=1}^{n} \frac{1}{P(z,w)} \sum_{(x,y)\in\Gamma^t(z,w)} |\gamma_{xy}|\widetilde{P}(x,y) = \sum_{t=1}^{n} \frac{1}{P(z,w)} \sum_{(x,y)\in\Gamma^t(z,w)} t\widetilde{P}(x,y).$$

First we obtain a bound on the ratio $\widetilde{P}(x,y)/P(z,w)$. We know that $P(z,w) = 1/(3mn)$ since P is the transition matrix of the Glauber-dynamics chain.

When $t \geq 2$, for $(x,y) \in \Gamma^t(z,w)$ we have $\widetilde{P}(x,y) \leq 1/(4 \cdot 3mn \cdot 2t)$ and hence $\widetilde{P}(x,y)/P(z,w) \leq 1/(8t)$. In the special case when $t = 1$, we have $(x,y) = (z,w)$ and $\widetilde{P}(z,w) = p_z/(4 \cdot 3mn)$. Hence $\widetilde{P}(x,y)/P(z,w) \leq 1/4$ in this case.

Substituting, we have

$$A_{z,w} \leq \frac{1}{4} + \sum_{t=2}^{n} \sum_{(x,y)\in\Gamma^t(z,w)} \frac{1}{8} = \frac{1}{4} + \frac{1}{8}\left|\bigcup_{t=2}^{n} \Gamma^t(z,w)\right|.$$

We must determine a bound for $|\bigcup_{t=2}^{n} \Gamma^t(z,w)|$, the number of pairs (x,y) that differ by a tower move of length at least 2 that use the transition (z,w) to "unwind" this tower.

Suppose that z and w differ at vertex (i,j). We can get a crude upper bound on the number of towers that pass through vertex (i,j) as follows. Consider first any such right-going towers: the start and end vertices are (i, a) and (i, b) respectively, where $1 \le a \le j$ and $j \le b \le n$, and $a \ne b$. The number of such pairs (a, b) is $j(n-j+1)-1$. Accounting for towers in all four directions, we have

$$\left|\bigcup_{t=2}^{n} \Gamma^t(z, w)\right| \le 2j(n-j+1) + 2i(m-i+1) - 4 \le (n+1)^2 - 4,$$

with the maximum occurring at $m = n$, $i = j = \lceil n+1 \rceil/2$. Hence $A < (n+1)^2/8$, giving the conclusion of the theorem. ■

(Note: A more detailed analysis shows that the transition (z, w) gives some extra information about the direction of the towers. As a result the bound $A \le n^2/32 + \mathrm{O}(n)$ is obtainable.)

8. CONCLUSIONS AND OPEN PROBLEMS

We have shown that Glauber dynamics mixes rapidly on the set of 3-colourings of an $m \times n$ rectangle. However, our bound on the mixing time is likely to be far from optimal. Tightening up the comparison method would give a slight improvement. In particular, the comparison added an additional factor of mn to the mixing time because of the term involving $1/\pi_*$ in Theorem 7.1. Randall and Tetali [15] discuss some situations in which this extra factor might be avoided by considering log-Sobolev constants of the Markov chains. There are similar results to Theorem 7.1 which involve the logarithmic Sobolev constants of two related Markov chains. The difficulty in using these results lies in finding a known chain for which the log-Sobolev constant (or a bound on it) is also known. The interested reader should consult [15] for more details on this method, and [6] and the references therein, for definitions and more general results about logarithmic Sobolev related results.

It is still unknown whether the Glauber-dynamics chain is rapidly mixing when the state space is the set of all 4-colourings (or 5-colourings or 6-colourings) of the $m \times n$ rectangle. It is also unknown whether our result extends to other regions of $\mathbb{Z}^2$.

ACKNOWLEDGMENTS

The authors are very grateful to Mary Cryan, Martin Dyer, Mark Jerrum, Dana Randall, and Peter Winkler for useful conversations during the course of this work.

Randall [14] has recently given an argument which, in some sense, explains our choice of transition probabilities in the Markov chain $\widetilde{\mathcal{M}}$. This argument relies on expanding the region slightly, and colouring the new parts of this expanded region in one of 16 prescribed possible ways. By considering the Markov chain from [10] on regions with fixed boundaries, the transition probabilities of $\widetilde{\mathcal{M}}$ can be seen to be an "average" of transition probabilities over all of the 16 possible ways to extend colourings of G to the larger region.

REFERENCES

[1] D. Aldous. *Random walks on finite groups and rapidly mixing Markov chains,* in A. Dold and B. Eckmann, eds., *Séminaire de Probabilités XBII 1981/82*, vol. 986 of *Springer-Verlag Lecture Notes in Mathematics,* Springer-Verlag, New York, 1983, pp. 243–297.

[2] R.J. Baxter. *Exactly Solved Models in Statistical Mechanics.* Academic Press, New York, 1982.

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

[4] R. Bubley, M. Dyer and C. Greenhill. Beating the 2Δ bound for approximately counting colourings: a computer-assisted proof of rapid mixing. *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms* **9** (1998), pp. 355–363.

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

[6] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Annals of Applied Probability* **6** (1996), pp. 695–750.

[7] M. Dyer, L.A. Goldberg and M. Jerrum. Systematic scan Metropolis algorithms, in preparation.

[8] M. Dyer and C. Greenhill. *Random walks on combinatorial objects*, in J.D. Lamb and D.A. Preece, eds., *Surveys in Combinatorics 1999*, vol. 267 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge, 1999, pp. 101–136.

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

[10] M. Luby, D. Randall, and A.J. Sinclair. Markov chain algorithms for planar lattice structures. *SIAM J. Computing* **31** (2001), pp. 167–192.

[11] N. Madras and D. Randall. Factoring graphs to bound mixing rates. *Proc. 37th Annual IEEE Symposium on Foundations of Computer Science* (1996), pp. 194–203.

[12] N. Madras and D. Randall. Markov chain decomposition for convergence analysis. *Annals of Applied Probability* **12** (2002), pp. 581–606.

[13] M. Mihail and P. Winkler. On the number of Eulerian orientations of a graph. *Algorithmica* **4-5** (1996), pp. 402–414.

[14] D. Randall. Personal communication (2002).

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

[16] J. Salas and A.D. Sokal. Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem. *J. Statistical Physics* **86** (1997), pp. 551–579.

[17] E. Vigoda. Improved bounds for sampling colorings. Probabilistic techniques in equilibrium and nonequilibrium statistical physics. *J. Mathematical Physics* **41** (2000), pp. 1555–1569.

Received
Accepted