Torpid Mixing of Local Markov Chains on 3-Colorings of the Discrete Torus

David Galvin*
Dana Randall†

Abstract

We study local Markov chains for sampling 3-colorings of the discrete torus $T_{L,d} = \{0, \ldots, L-1\}^d$. We show that there is a constant $\rho \approx .22$ such that for all even $L \geq 4$ and d sufficiently large, certain local Markov chains require exponential time to converge to equilibrium. More precisely, if $\mathcal{M}$ is a Markov chain on the set of proper 3-colorings of $T_{L,d}$ that updates the color of at most ρL^d vertices at each step and whose stationary distribution is uniform, then the convergence to stationarity of $\mathcal{M}$ is exponential in L^{d-1}. Our proof is based on a conductance argument that builds on sensitive new combinatorial enumeration techniques.

1 Introduction

Sampling and counting colorings of a graph are fundamental problems in computer science and discrete mathematics. We consider the problem of sampling uniformly at random from the set $\mathcal{C}_k = \mathcal{C}_k(G)$ of proper k-colorings of a graph $G = (V, E)$. A proper k-coloring χ is a labeling $\chi : V \rightarrow k$ such that all neighboring vertices have different colors. This sampling problem is also fundamental in statistical physics and corresponds to generating configurations from the Gibbs distribution of the zero-temperature antiferromagnetic Potts model [20]. From the physics perspective, the underlying graph is typically taken to be the cubic lattice $\mathbb{Z}^d$ and sampling and counting reveal underlying thermodynamic properties of the corresponding physical system.

Much focus has gone towards solving the sampling problem using rapidly mixing Markov chains. The idea is to design a Markov chain whose stationary distribution is uniform over the set of proper colorings. Then, starting at an arbitrary coloring and simulating a random walk according to this chain for a sufficient number of steps, we get a sample from close to the desired distribution. The number of steps required of this walk is referred to as the *mixing time* (see, e.g., [17]). The chain is called *rapidly mixing* if the mixing time is polynomial in $n = |V|$ (so it converges quickly to stationarity); it is *torpidly mixing* if its mixing time is super-polynomial in n (so it converges slowly). There has been a long history of studying mixing times of various chains in the context of colorings (see, e.g., [1, 6, 10, 11, 12, 15]).

A particular focus of this study has been on *Glauber dynamics*. For proper k-colorings this is any single-site update Markov chain that connects two colorings only if they differ on at most a single vertex. The *Metropolis* chain $\mathcal{M}_k$ on state space $\mathcal{C}_k$ has transition probabilities $P_k(\chi_1, \chi_2)$, $\chi_1, \chi_2 \in \mathcal{C}_k$, given by

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

We may think of $\mathcal{M}_k$ dynamically as follows. From a k-coloring χ, choose a vertex v uniformly from V and a color j uniformly from $\{0, \ldots, k-1\}$. Then recolor v with color j if this is a proper k-coloring; otherwise stay at χ.

When $\mathcal{M}_k$ is ergodic, its stationary distribution π_k is uniform over proper k-colorings. A series of recent papers have shown that $\mathcal{M}_k$ is rapidly mixing provided the number of colors is sufficiently large compared to the maximum degree (see [6] and the references therein). Substantially less is known when the number of colors is small. In fact, for k small it is NP-complete to decide whether a graph admits even one k-coloring.

In this paper we focus on the mixing rate of $\mathcal{M}_k$ on rectangular regions of the cubic lattice $\mathbb{Z}^d$. Observe that the lattice is bipartite, so it always admits a k-coloring for any $k \geq 2$. It is also known that Glauber dynamics connects the state space of k-colorings on any such lattice region [15]. In $\mathbb{Z}^2$ much is known about

*Department of Mathematics, University of Pennsylvania, Philadelphia, PA 19104; dgalvin@math.upenn.edu.
†College of Computing, Georgia Institute of Technology, Atlanta, GA 30332; randall@cc.gatech.edu. Supported in part by NSF grants CCR-0515105 and DMS-0505505.

the mixing rate of $\mathcal{M}_k$. Luby et al. [15] showed that Glauber dynamics for sampling 3-colorings is rapidly mixing on any finite, simply-connected subregion of $\mathbb{Z}^2$ when the colors on the boundary of the region are fixed. Goldberg et al. [10] subsequently showed that the chain remains fast on rectangular regions without this boundary restriction. Substantially more is known when there are many colors: Jerrum [12] showed that Glauber dynamics is rapidly mixing on any graph satisfying $k \geq 2\Delta$, where k is the number of colors and Δ is the maximum degree, thus showing Glauber dynamics is fast on $\mathbb{Z}^2$ when $k \geq 8$. It has since been shown that it is fast for $k \geq 6$ [1, 3]. Surprisingly the efficiency remains unresolved for $k = 4$ or 5.

In higher dimensions much less is known when k is small. Physicists have performed extensive numerical experiments [5, 19] suggesting that Glauber dynamics on 3-colorings is torpidly mixing when the dimension of the cubic lattice is large enough. We prove this conjecture for the first time here by studying the mixing time of the chain on cubic lattices with periodic boundary conditions.

1.1 Results Our focus in this paper is sampling 3-colorings of the even discrete torus $T_{L,d}$. This is the graph on vertex set $\{0, \ldots, L-1\}^d$ (with L even) with edge set consisting of those pairs of vertices that differ on exactly one coordinate and differ by 1 (mod L) on that coordinate. For a Markov chain $\mathcal{M}$ on the 3-colorings of $T_{L,d}$ we denote by $\tau_{\mathcal{M}}$ the mixing time of the chain; this will be formally defined in Section 2. Our main theorem is the following.

THEOREM 1.1. *There is a constant $d_0 > 0$ for which the following holds. For $d \geq d_0$ and $L \geq 4$ even, the Glauber dynamics chain $\mathcal{M}_3$ on $\mathcal{C}_3(T_{L,d})$ satisfies*

$$\tau_{\mathcal{M}_3} \geq \exp\left\{\frac{L^{d-1}}{d^4 \log^2 L}\right\}.$$

Our techniques actually apply to a more general class of chains. A Markov chain $\mathcal{M}$ on state space $\mathcal{C}_3$ is *ρ-local* if, in each step of the chain, at most $\rho|V|$ vertices have their color changed; that is, if

$$P_{\mathcal{M}}(\chi_1, \chi_2) \neq 0$$

implies

$$|\{v \in V : \chi_1(v) \neq \chi_2(v)\}| \leq \rho|V|.$$

These types of chains were introduced in [4], where the terminology *$\rho|V|$-cautious* was employed. We prove the following, which easily implies Theorem 1.1.

THEOREM 1.2. *Fix $\rho > 0$ satisfying $H(\rho) + \rho < 1$. There is a constant $d_0 = d_0(\rho) > 0$ for which the following holds. For $d \geq d_0$ and $L \geq 4$ even, if $\mathcal{M}$ is an ergodic ρ-local Markov chain on $\mathcal{C}_3(T_{L,d})$ with uniform stationary distribution then*

$$\tau_{\mathcal{M}} \geq \exp\left\{\frac{L^{d-1}}{d^4 \log^2 L}\right\}.$$

Here $H(x) = -x \log x - (1-x)\log(1-x)$ is the usual binary entropy function. Note that all $\rho \leq .22$ satisfy $H(\rho) + \rho < 1$.

1.2 Techniques We show slow mixing via a conductance argument by identifying a "bad cut" in the state space requiring exponential time to cross. Intuitively, in sufficiently high dimension, the set of 3-colorings of the lattice is believed to naturally partition into 6 classes: each class is identified by a predominance of one (of 3) colors on one of the two (even or odd) sublattices. This characterization was recently rigorously verified on the infinite lattice, thereby establishing the existence of 6 distinct "maximal entropy Gibbs states" [9]. That work builds heavily on technical machinery introduced by Galvin and Kahn [8] showing that independent sets partition similarly in sufficiently high dimensions in that they lie primarily on the even or odd sublattices. Specifically, write $\mathcal{E}$ and $\mathcal{O}$ for the sets of even and odd vertices of $\mathbb{Z}^d$ (defined in the obvious way) and set $\Lambda_L = [-L, L]^d$ and $\partial\Lambda_L = [-L, L]^d \setminus [-(L-1), L-1]^d$. For $\lambda > 0$, choose $\mathbb{I}$ from $\mathcal{I}(\Lambda_L)$ (the set of independent sets of the box) with $\Pr(\mathbb{I} = I) \propto \lambda^{|I|}$. Galvin and Kahn showed that for $\lambda > Cd^{-1/4}\log^{3/4} d$ (for a large constant C) and fixed $v \in \Lambda_L \cap \mathcal{E}$

$$\lim_{L\to\infty} \mathbb{P}(v \in \mathbb{I} \mid \mathbb{I} \supseteq \partial\Lambda_L \cap \mathcal{E}) > \lim_{L\to\infty} \mathbb{P}(v \in \mathbb{I} \mid \mathbb{I} \supseteq \partial\Lambda_L \cap \mathcal{O}).$$

In other words, the influence of the boundary on the center of a large box persists as the boundary recedes.

Notice that neither of the results of [8] or [9] establishing the presence of multiple Gibbs states directly implies anything about the behavior of Markov chains on finite lattice regions. However, they do suggest that in the finite setting, typical configurations fall into the distinct classes described in stationarity and that it will be unlikely to move between these classes; the remaining configurations are expected to have negligible weight for large lattice regions, even when they are finite.

Galvin [7] extended the results of [8], showing that in sufficiently high dimension, Glauber dynamics on independent sets mixes slowly in rectangular regions of $\mathbb{Z}^d$ with periodic boundary conditions. Similar results were known previously about independent sets; however, one significant new contribution of [7] was showing that as d increases, the critical λ above which Glauber dynamics mixes slowly tends to 0. In particular, there is some dimension d_0 such that for all $d \geq d_0$, Glauber dynamics will be slow on $\mathbb{Z}^d$ when $\lambda = 1$. This turns out to be the crucial new ingredient allowing us to rigorously verify slow mixing for sampling 3-colorings in high dimensions, as there turns out to be a close connection between the independent set model at $\lambda = 1$ and the 3-coloring model. Note that unlike most statistical physics models, the 3-coloring problem does not have a parameter λ that can be tweaked to establish desired bounds; this makes the proofs here significantly more delicate than the usual slow mixing arguments. Section 3.1 provides a more detailed discussion of the elements of the proof and of some of the difficulties inherent to the sampling problem under discussion.

2 Partitioning the state space

We begin by formalizing some definitions. Given a Markov chain $\mathcal{M}$ on state space Ω with uniform stationary distribution denoted by π, let $P^t(X, \cdot)$ be the distribution of the chain at time t given that it started in state X. The *mixing time* $\tau_{\mathcal{M}}$ of $\mathcal{M}$ is defined to be

$$\tau_{\mathcal{M}} = \min \left\{ t_0 : ||P^t, \pi||_{\text{tv}} \leq \frac{1}{e} \quad \forall t > t_0 \right\}$$

where

$$||P^t, \pi||_{\text{tv}} = \max_{X \in \Omega} \frac{1}{2} \sum_{Y \in \Omega} |P^t(X, Y) - \pi(Y)|,$$

is the *total variation distance*.

We prove Theorem 1.2 via a well-known *conductance argument* [13, 14, 18], using a form of the argument derived in [4]. As above, 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 [4] we have

$$\tau_{\mathcal{M}} \geq \frac{\pi(A)}{8\pi(M)}. \tag{2.1}$$

Let us return to the setup of Theorem 1.2. For even L, $T_{L,d}$ is bipartite with partition classes $\mathcal{E}$ (consisting of those vertices the sum of whose coordinates is even) and $\mathcal{O}$. To show torpid mixing, it is sufficient to identify a single bad cut. We concentrate on the vertices in each 3-coloring that are colored with the first color, 0. The objective of Theorem 1.2 will be to verify that most 3-colorings have an imbalance whereby the vertices colored 0 lie predominantly on $\mathcal{E}$ or $\mathcal{O}$, and those that are roughly balanced on the two sublattices are highly unlikely in stationarity. This is sufficient to show that the conductance is small.

Accordingly let us define the set of "balanced" 3-colorings by

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

and likewise let

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

By symmetry, $\pi_3(\mathcal{C}_3^{\mathcal{E},\rho}) \leq 1/2$. Notice that since $\mathcal{M}$ updates at most ρL^d vertices in each step, we have that if $\chi_1 \in \mathcal{C}_3^{\mathcal{E},\rho}$ and $\chi_2 \in \mathcal{C}_3 \setminus (\mathcal{C}_3^{\mathcal{E},\rho} \cup \mathcal{C}_3^{b,\rho})$ then $P_{\mathcal{M}}(\chi_1, \chi_2) = 0$. Therefore, by (2.1),

$$\tau_{\mathcal{M}} \geq \frac{\pi_3(\mathcal{C}_3^{\mathcal{E},\rho})}{8\pi_3(\mathcal{C}_3^{b,\rho})} \geq \frac{1 - \pi_3(\mathcal{C}_3^{\mathcal{E},\rho})}{16\pi_3(\mathcal{C}_3^{b,\rho})},$$

and so Theorem 1.2 follows from the following critical theorem.

THEOREM 2.1. *Fix $\rho > 0$ satisfying $H(\rho) + \rho < 1$. There is a constant $d_0 = d_0(\rho) > 0$ for which the following holds. For $d \geq d_0$ and $L \geq 4$ even,*

$$\pi_3(\mathcal{C}_3^{b,\rho}) \leq \exp\left\{ \frac{-2L^{d-1}}{d^4 \log^2 L} \right\}.$$

3 Proof of Theorem 2.1

3.1 Setup and overview For a generic $\chi \in \mathcal{C}_3^{b,\rho}$ there are regions of $T_{L,d}$ consisting predominantly of even vertices colored 0 together with their neighbors, and regions consisting of odd vertices colored 0 together with their neighbors. These regions are separated by two-layer "0-free" moats or *cutsets*. In Section 3.2 we describe a procedure that selects a particular collection of these cutsets. Our main technical result, Lemma 3.1, asserts that for each specification of cutset sizes $c_1, \ldots, c_\ell$ and vertices $v_1, \ldots, v_\ell$, the probability that a coloring has among its associated cutsets a collection $\gamma_1, \ldots, \gamma_\ell$ with $|\gamma_i| = c_i$ and with v_i surrounded by γ_i is exponentially small in the sum of the c_i's. This lemma is presented in Section 3.3 and Theorem 2.1 is derived from it in Section 3.4.

The main thrust of [9] is the proof of a result that is essentially (but not quite) the case $\ell = 1$ of Lemma 3.1. One difficulty we have to overcome in moving from a Gibbs measure argument to a torpid mixing argument is that of going from bounding the probability of a configuration having a single cutset to bounding the probability of it having an ensemble of cutsets. Another difficulty is that the cutsets we consider in these ensembles can be topologically more complex than the connected cutsets that are considered in [9]. In part, both of these difficulties are dealt with by the machinery developed in [7].

We use a "Peierl's argument" to prove Lemma 3.1. By carefully modifying each $\chi \in \mathcal{C}_3^{b,\rho}$ inside its cutsets, we can exploit the fact that the cutsets are 0-free to map χ to a set $\varphi(\chi)$ of many different $\chi' \in \mathcal{C}_3$. If the $\varphi(\chi)$'s were disjoint for distinct χ's, we would essentially be done, having shown that there are many more 3-colorings in total than 3-colorings in $\mathcal{C}_3^{b,\rho}$. To control the possible overlap, we define a flow $\nu : \mathcal{C}_3^{b,\rho} \times \mathcal{C}_3 \to [0,\infty)$ supported on pairs (χ,χ') with $\chi' \in \varphi(\chi)$ in such a way that the flow out of each $\chi \in \mathcal{C}_3^{b,\rho}$ is 1. Any uniform bound we can obtain on the flow into elements of $\mathcal{C}_3$ is then easily seen to be a bound on $\pi_3(\mathcal{C}_3^{b,\rho})$. We define the flow via a notion of approximation modified from [8]. To each cutset γ we associate a set $A(\gamma)$ that approximates the interior of γ in a precise sense, in such a way that as we run over all possible γ, the total number of approximate sets used is small. Then for each $\chi' \in \mathcal{C}_3$ and each collection of approximations $A_1,\ldots,A_\ell$, we consider the set of those $\chi \in \mathcal{C}_3^{b,\rho}$ with $\chi' \in \varphi(\chi)$ and with A_i the approximation to γ_i. We define the flow so that if this set is large, then $\nu(\chi,\chi')$ is small for each χ in the set. In this way we control the flow into χ' corresponding to each collection of approximations $A_1,\ldots,A_\ell$; since the total number of approximations is small, we control the total flow into χ'. In the language of statistical physics, this approximation scheme is a *course-graining* argument. The details appear in Section 4.

The main results of [7] and [8] are proved along similar lines to those described above. One of the difficulties we encounter in moving from these arguments on independent sets to arguments on colorings is that of finding an analogous way of modifying a coloring inside a cutset in order to exploit the fact that it is 0-free. The beginning of Section 4 (in particular Claims 4.1 and 4.2) describes an appropriate modification that has all the properties we desire.

3.2 Cutsets We describe a way of associating with each $\chi \in \mathcal{C}_3^{b,\rho}$ a collection of minimal edge cutsets, following the approaches of [2] and [7]. First we need a little notation.

Write V for the vertex set of $T_{L,d}$ and E for its edge set. For $X \subseteq V$, write $\nabla(X)$ for the set of edges in E that have one end in X and one end outside X; $\overline{X}$ for $V \setminus X$; $\partial_{int}X$ for the set of vertices in X that are adjacent to something outside X; $\partial_{ext}X$ for the set of vertices outside X that are adjacent to something in X; X^+ for $X \cup \partial_{ext}X$; $X^{\mathcal{E}}$ for $X \cap \mathcal{E}$ and $X^{\mathcal{O}}$ for $X \cap \mathcal{O}$. Further, for $x \in V$ set $\partial x = \partial_{ext}\{x\}$. We abuse notation slightly, identifying sets of vertices of V and the subgraphs they induce.

For each $\chi \in \mathcal{C}_3^{b,\rho}$ set $I = I(\chi) = \chi^{-1}(0)$. Note that $I(\chi)$ is an independent set (a set of vertices no two of which are adjacent). For each component R of $(I^{\mathcal{E}})^+$ or $(I^{\mathcal{O}})^+$ and each component C of $\overline{R}$, set $\gamma = \gamma_{RC}(I) = \nabla(C)$ and $W = W_{RC}(I) = \overline{C}$. Evidently C is connected, and W consists of R, which is connected, together with a number of other components of $\overline{R}$, each of which is connected and joined to R, so W is connected also. It follows that γ is a minimal edge-cutset in $T_{L,d}$. Say that γ is *even* if R is a component of $(I^{\mathcal{E}})^+$ and *odd* otherwise. Define $\text{int}\,\gamma$, the *interior* of γ, to be the smaller of C, W (if $|W| = |C|$, take $\text{int}\,\gamma = W$).

The cutsets γ associated to χ depend only on the independent set $I(\chi)$, and coincide exactly with the cutsets associated to an independent set in [7]. We may therefore apply the machinery developed in [7] for independent set cutsets in the present setting. In particular, from [7, Lemmas 3.1 and 3.2] we know that for each $\chi \in \mathcal{C}_3$ there is a collection of associated cutsets $\Gamma(I)$ such that either

$$\begin{aligned} &\text{for all } \gamma,\gamma' \in \Gamma(I), \\ &\gamma,\gamma' \text{ are even with } \text{int}\,\gamma \cap \text{int}\,\gamma' = \emptyset, \\ &\text{and } I^{\mathcal{E}} \subseteq \cup_\gamma \text{int}\,\gamma, \end{aligned} \tag{3.2}$$

or we have the analogue of (3.2) with even replaced by odd. Set $\mathcal{C}_3^{even} = \{\chi \in \mathcal{C}_3 : \chi \text{ satisfies (3.2)}\}$. From here on whenever $\chi \in \mathcal{C}_3^{even}$ is given we assume that I is its associated independent set and that $\Gamma(I)$ is a particular collection of cutsets associated with χ and satisfying (3.2). Numerous properties of $\gamma \in \Gamma(I)$ are established in [7, Lemmas 3.3 and 3.4]. We list some here that will be of use in the sequel. That the cutsets are indeed 0-free regions is established by (3.4).

$$\partial_{int}W \subseteq \mathcal{O} \text{ and } \partial_{ext}W \subseteq \mathcal{E}; \tag{3.3}$$

$$\partial_{int}W \cap I = \emptyset \text{ and } \partial_{ext}W \cap I = \emptyset; \tag{3.4}$$

$$W^{\mathcal{O}} = \partial_{ext}W^{\mathcal{E}} \text{ and } W^{\mathcal{E}} = \left\{y \in \mathcal{E} : \partial y \subseteq W^{\mathcal{O}}\right\}; \tag{3.5}$$

$$\text{for large enough } d, \quad |\gamma| \geq \max\{|W|^{1-1/d}, d^{1.9}\}. \tag{3.6}$$

3.3 The main lemma For $c \in \mathbb{N}$ and $v \in V$ set

$$\mathcal{W}(c,v) = \left\{\gamma : \begin{array}{l} \gamma \in \Gamma(I) \text{ for some } \chi \in \mathcal{C}_3^{even} \\ \text{with } |\gamma| = c,\ v \in W^{\mathcal{E}} \end{array}\right\}$$

and set $\mathcal{W} = \cup_{c,v}\mathcal{W}(c,v)$. A *profile* of a collection $\{\gamma_1,\ldots,\gamma_\ell\} \subseteq \mathcal{W}$ is a vector $\underline{p} = (c_1,v_1,\ldots,c_\ell,v_\ell)$ with $\gamma_i \in \mathcal{W}(c_i,v_i)$ for all i. Given a profile vector $\underline{p}$ set

$$\mathcal{C}_3(\underline{p}) = \left\{\chi \in \mathcal{C}_3^{even} : \begin{array}{l} \Gamma(I) \text{ contains a subset} \\ \text{with profile } \underline{p} \end{array}\right\}.$$

Our main lemma (c.f. [7, Lemma 3.5]) is the following.

LEMMA 3.1. *There are constants $c, d_0 > 0$ such that the following holds. For all even $L \geq 4$, $d \geq d_0$ and profile vector $\underline{p}$,*

$$\pi_3(\mathcal{C}_3(\underline{p})) \leq \exp\left\{-\frac{c\sum_{i=1}^{\ell} c_i}{d}\right\}. \tag{3.7}$$

We will derive Theorem 2.1 from Lemma 3.1 in Section 3.4 before proving the lemma in Section 4. From here on we assume that the conditions of Theorem 2.1 and Lemma 3.1 are satisfied (with d_0 sufficiently large to support our assertions).

3.4 Proof of Theorem 2.1 assuming the main lemma We begin with an easy count that dispenses with colorings where $|I(\chi)|$ is small. Set

$$\mathcal{C}_3^{small} = \left\{\chi \in \mathcal{C}_3^{b,\rho} : \min\{|I^{\mathcal{E}}|, |I^{\mathcal{O}}|\} \leq L^d/4d^{1/2}\right\}.$$

LEMMA 3.2. $\pi_3(\mathcal{C}_3^{small}) \leq \exp\left\{-\Omega(L^d)\right\}$.

Proof: For any $A \subseteq \mathcal{E}$ and $B \subseteq \mathcal{O}$, let $\text{comp}(A, B)$ be the number of components in $V \setminus (A \cup B \cup \partial^\star A \cup \partial^\star B)$, where for $T \subseteq \mathcal{E}$ (or $\mathcal{O}$),

$$\partial^\star T = \{x \in \partial_{ext} T : \partial x \subseteq T\} \ (= \{x \in V : \partial x \subseteq T\}).$$

We begin by noting that by $\mathcal{E}$-$\mathcal{O}$ symmetry

$$|\mathcal{C}_3^{small}| \leq 2 \sum 2^{|\partial^\star A| + |\partial^\star B| + \text{comp}(A,B)}, \tag{3.8}$$

where the sum is over all pairs $A \subseteq \mathcal{E}$, $B \subseteq \mathcal{O}$ with no edges between A and B and satisfying $|A| \leq L^d/4d^{1/2}$ and $|B| \leq (\rho + 1/2d^{1/2})L^d/2$. Indeed, once we have specified that the set of vertices colored 0 is $A \cup B$, we have a free choice between 1 and 2 for the color at $x \in \partial^\star A \cup \partial^\star B$, and we also have a free choice between the two possible colorings of each component of $V \setminus (A \cup B \cup \partial^\star A \cup \partial^\star B)$.

A key observation is the following. For A and B contributing to the sum in (3.8),

$$\text{comp}(A, B) \leq L^d/2d. \tag{3.9}$$

To see this, let C be a component of $V \setminus (A \cup B)$. If $C = \{v\}$ consists of a single vertex, then (depending on the parity of v) we have either $\partial v \subseteq A$ or $\partial v \subseteq B$ and so $v \in \partial^\star A \cup \partial^\star B$. Otherwise, let vw be an edge of C with $v \in \mathcal{E}$ (and so $w \in \mathcal{O}$). If v has k edges to B and u has ℓ to A, then (since there are no edges from A to B) we have $(k-1) + (\ell - 1) \leq 2d - 2$ or $k + \ell \leq 2d$. (Here we are using that in $T_{L,d}$, if $uv \in E$ then there is a matching between all but one of the neighbors of u and v.) Since v has $2d - 1 - k$ edges to $\mathcal{O} \setminus (B \cup \{w\})$ and w has $2d - 1 - \ell$ edges to $\mathcal{E} \setminus (A \cup \{v\})$ we have that $|C| = 4d - (k + \ell) \geq 2d$. From this, (3.9) follows.

Inserting (3.9) into (3.8) and bounding $|\partial^\star A|$ and $|\partial^\star B|$ by the maximum values of $|A|$ and $|B|$ (valid since $T \subseteq \mathcal{E}$ (or $\mathcal{O}$) satisfies $|T| \leq |\partial_{ext} T|$, so $|\partial^\star T| \leq |T|$) and with the remaining inequalities justified below, we have

$$\begin{aligned} |\mathcal{C}_3^{small}| &\leq 2^{\frac{L^d}{2}\ \rho + \frac{1}{d^{1/2}} + \frac{1}{d}} \cdot \sum_{i \leq L^d/4d^{1/2}} \binom{L^d/2}{i} \\ &\qquad \cdot \sum_{j \leq (\rho + 1/2d^{1/2})L^d/2} \binom{L^d/2}{j} \\ &\leq 2^{\frac{L^d}{2}\ \rho + \frac{1}{d^{1/2}} + \frac{1}{d} + H\ \frac{1}{2d^{1/2}}\ + H\ \rho + \frac{1}{2d^{1/2}}} && (3.10) \\ &\leq 2^{\frac{L^d}{2}(1 - \Omega(1))} && (3.11) \end{aligned}$$

for sufficiently large $d = d(\rho)$. In (3.10) we use the Chernoff bound $\sum_{i=0}^{[\beta M]} \binom{M}{i} \leq 2^{H(\beta)M}$ for $\beta \leq \frac{1}{2}$; in (3.11) we use $H(\rho) + \rho < 1$. Using $2^{L^d/2} \leq |\mathcal{C}_3|$, the lemma follows. □

We now consider

$$\mathcal{C}_3^{large,even} := (\mathcal{C}_3^{b,\rho} \setminus \mathcal{C}_3^{small}) \cap \mathcal{C}_3^{even}.$$

By Lemma 3.2 and $\mathcal{E}$-$\mathcal{O}$ symmetry, Theorem 2.1 reduces to bounding (say)

$$\pi_3(\mathcal{C}_3^{large,even}) \leq \exp\left\{-\frac{3L^{d-1}}{d^4 \log^2 L}\right\}. \tag{3.12}$$

Let $\mathcal{C}_3^{large,even,nt}$ be the set of $\chi \in \mathcal{C}_3^{large,even}$ such that there is a $\gamma \in \Gamma(I)$ with $|\gamma| \geq L^{d-1}$ (we think of such cutsets as being topologically non-trivial ("nt"); see [7] for an explanation of this) and also let $\mathcal{C}_3^{large,even,triv} = \mathcal{C}_3^{large,even} \setminus \mathcal{C}_3^{large,even,nt}$. We assert that

$$\pi_3(\mathcal{C}_3^{large,even,nt}) \leq \exp\left\{-\Omega\left(\frac{L^{d-1}}{d}\right)\right\} \tag{3.13}$$

and

$$\pi_3(\mathcal{C}_3^{large,even,triv}) \leq \exp\left\{-\frac{4L^{d-1}}{d^4 \log^2 L}\right\}; \tag{3.14}$$

this gives (3.12) and so completes the proof of Theorem 2.1. Both (3.13) and (3.14) are corollaries of Lemma 3.1, and the steps are identical to those that are used to bound the measures of "$\mathcal{I}_{large,even}^{non-trivial}$" and "$\mathcal{I}_{large,even}^{trivial}$" in [7, Section 3.3].

With the sum below running over all vectors $\underline{p}$ of the form (c, v) with $v \in V$ and $c \geq L^{d-1}$, and with the inequalities justified below, we have

$$\begin{aligned}
\pi_3(\mathcal{C}_3^{large,even,nt}) &\le \sum_{\underline{p}} \pi_3(\mathcal{C}_3(\underline{p})) \\
&\le L^{2d} \exp\left\{-\Omega\left(\frac{L^{d-1}}{d}\right)\right\} \qquad (3.15)\\
&\le \exp\left\{-\Omega\left(\frac{L^{d-1}}{d}\right)\right\},
\end{aligned}$$

giving (3.13). We use Lemma 3.1 in (3.15). The factor of L^{2d} is for the choices of c and v.

The verification of (3.14) involves finding an $i \in [\Omega(\log d), O(d\log L)]$ and a set $\Gamma_i(I) \subseteq \Gamma(I)$ of cutsets with the properties that $|\Gamma_i(I)| \approx L^d/2^i$, $|\gamma| \approx 2^i$ for each $\gamma \in \Gamma_i(I)$ and $\sum_{\gamma\in\Gamma_i(I)} |\gamma| \approx L^{d-1}$. The measure of $\mathcal{C}_3^{large,even,triv}$ is then at most the product of a term that is exponentially small in L^{d-1} (from Lemma 3.1), a term corresponding to the choice of a fixed vertex in each of the interiors, and a term corresponding to the choice of the collection of lengths. The second term will be negligible because $\Gamma_i(I)$ is small and the third will be negligible because all $\gamma \in \Gamma_i(I)$ have similar lengths.

More precisely, for $\chi \in \mathcal{C}_3^{large,even,triv}$ and $\gamma \in \Gamma(I)$ we have $|\gamma| \ge |\text{int}\,\gamma|^{1-1/d}$ (by (3.6)) and so

$$\sum_{\gamma\in\Gamma(I)} |\gamma|^{d/(d-1)} \ge \sum_{\gamma\in\Gamma(I)} |\text{int}\,\gamma| \ge |I^{\mathcal{E}}| \ge L^d/4d^{1/2}.$$

The second inequality is from (3.2) and the third follows since $\chi \notin \mathcal{C}_3^{small}$.

Set $\Gamma_i(I) = \{\gamma \in \Gamma(I) : 2^{i-1} \le |\gamma| < 2^i\}$. Note that $\Gamma_i(I)$ is empty for $2^i < d^{1.9}$ (again by (3.6)) and for $2^{i-1} > L^{d-1}$ so we may assume that

$$1.9\log d \le i \le (d-1)\log L + 1. \qquad (3.16)$$

Since $\sum_{m=1}^{\infty} 1/m^2 = \pi^2/6$, there is an i such that

$$\sum_{\gamma\in\Gamma_i(I)} |\gamma|^{\frac{d}{d-1}} \ge \Omega\left(\frac{L^d}{d^{1/2} i^2}\right). \qquad (3.17)$$

Choose the smallest such i set $\ell = |\Gamma_i(I)|$. We have $\sum_{\gamma\in\Gamma_i(I)} |\gamma| \ge \Omega(\ell 2^i)$ (this follows from the fact that each $\gamma \in \Gamma_i(I)$ satisfies $|\gamma| \ge 2^{i-1}$) and

$$O\left(\frac{dL^d}{2^i}\right) \ge \ell \ge \Omega\left(\frac{L^d}{2^{\frac{id}{d-1}} i^2 d^{1/2}}\right). \qquad (3.18)$$

The first inequality follows from that fact that $\sum_\gamma |\gamma| \le dL^d = |E|$; the second follows from (3.17) and the fact that each γ has $|\gamma|^{d/(d-1)} \le 2^{di/(d-1)}$. We therefore have $\chi \in \mathcal{C}_3(\underline{p})$ for some $\underline{p} = (c_1, v_1, \ldots, c_\ell, v_\ell)$ with ℓ satisfying (3.18), with

$$\sum_{j=1}^{\ell} c_j \ge O(\ell 2^i), \qquad (3.19)$$

with

$$c_j \le 2^i \qquad (3.20)$$

for each j and with i satisfying (3.16). With the sum below running over all profile vectors $\underline{p}$ satisfying (3.16), (3.18), (3.19) and (3.20) we have

$$\pi_3(\mathcal{C}_3^{large,even,triv}) \le \sum_{\underline{p}} \pi_3(\mathcal{C}_3(\underline{p})). \qquad (3.21)$$

The right-hand side of (3.21) is, by Lemma 3.1, at most

$$d\log L \max_{i \text{ satisfying (3.16)}} 2^{\ell i} \binom{L^d}{\ell} \exp\left\{-\Omega\left(\frac{\ell 2^i}{d}\right)\right\}.$$

The factor of $d\log L$ is an upper bound on the number of choices for i; the factor of $2^{\ell i}$ is for the choice of the c_j's; and the factor $\binom{L^d}{\ell}$ is for the choice of the ℓ (distinct) v_j's. By (3.16) and the second inequality in (3.18) we have (for d sufficiently large)

$$\begin{aligned}
2^{\ell i}\binom{L^d}{\ell} &\le 2^{\ell i}\left(\frac{L^d}{\ell}\right)^{\ell} \\
&\le 2^{\ell i}\left(O\left(2^{\frac{id}{d-1}} i^2 d^{1/2}\right)\right)^{\ell} \\
&\le 2^{4\ell i} \\
&= \exp\left\{o\left(\frac{2^i}{d}\right)\right\},
\end{aligned}$$

so that in fact the right-hand side of (3.21) is at most

$$d\log L \max_i \exp\left\{-\Omega\left(\frac{2^i \ell}{d}\right)\right\}.$$

Taking ℓ as small as possible we see that this is at most

$$d\log L \max_i \exp\left\{-\Omega\left(\frac{2^i L^d}{d 2^{\frac{id}{d-1}} i^2 d^{1/2}}\right)\right\}$$

and taking i as large as possible we see that this is at most $\exp\{-4L^{d-1}/d^4\log^2 L\}$. Putting these observation together we obtain (3.14).

4 Proof of Lemma 3.1

Much of what follows is modified from [7] and [9]. Our strategy is as follows. Let $\underline{p} = (c_1, v_1, \ldots, c_\ell, v_\ell)$ be given. Set $\underline{p'} = (c_2, v_2, \ldots, c_\ell, v_\ell)$. We will show

$$\frac{\pi_3(\mathcal{C}_3(\underline{p}))}{\pi_3(\mathcal{C}_3(\underline{p'}))} \le \exp\left\{-\Omega\left(\frac{c_1}{d}\right)\right\} \qquad (4.22)$$

from which the lemma follows by a telescoping product. To obtain (4.22) we define a one-to-many map φ from

$\mathcal{C}_3(\underline{p})$ to $\mathcal{C}_3(\underline{p'})$. We then define a flow $\nu : \mathcal{C}_3(\underline{p}) \times \mathcal{C}_3(\underline{p'}) \to [0,\infty)$ supported on pairs (χ, χ') with $\chi' \in \varphi(\chi)$ satisfying

$$\forall \chi \in \mathcal{C}_3(\underline{p}), \sum_{\chi' \in \varphi(\chi)} \nu(\chi, \chi') = 1 \tag{4.23}$$

and

$$\forall \chi' \in \mathcal{C}_3(\underline{p'}), \sum_{\chi \in \varphi^{-1}(\chi')} \nu(\chi, \chi') \leq \exp\left\{-\Omega\left(\frac{c_1}{d}\right)\right\}. \tag{4.24}$$

This easily gives (4.22).

For each $s \in \{\pm 1, \ldots, \pm d\}$, define σ_s, the *shift in direction* s, by $\sigma_s(x) = x + e_s$, where e_s is the sth standard basis vector if $s > 0$ and $e_s = -e_{-s}$ if $s < 0$. For $X \subseteq V$ write $\sigma_s(X)$ for $\{\sigma_s(x) : x \in X\}$. For $\gamma \in \mathcal{W}$ set $W^s = \{x \in \partial_{int} W : \sigma_{-s}(x) \notin W\}$.

Let $\chi \in \mathcal{C}_3(\underline{p})$ be given. Arbitrarily pick $\gamma \in \Gamma(I) \cap \mathcal{W}(c_1, v_1)$ and set $W = \operatorname{int} \gamma$. Write f for the map from $\{0,1,2\}$ to $\{0,1,2\}$ that sends 0 to 0 and transposes 1 and 2. For each $s \in \{\pm 1, \ldots, \pm d\}$ and $S \subseteq W$ define the function $\chi_S^s : V \to \{0,1,2\}$ by

$$\chi_S^s(v) = \begin{cases} 0 & \text{if } v \in S \\ \chi(v) & \text{if } v \in (W^s \setminus S) \cup (V \setminus W) \\ f(\chi(\sigma_{-s}(v))) & \text{if } v \in W \setminus W^s \end{cases}$$

and set $\varphi_s(\chi) = \{\chi_S^s : S \subseteq W^s\}$.

Claim 4.1. *$\varphi_s(\chi) \subseteq \mathcal{C}_3(\underline{p'})$.*

Proof: An easy case analysis verifies $\varphi_s(\chi) \subseteq \mathcal{C}_3$. We begin with the observation that the graph $\partial_{int} W \cup \partial_{ext} W$ is bipartite with bipartition $(\partial_{int} W, \partial_{ext} W)$. This follows from (3.3). By (3.4), $I \cap (\partial_{int} W \cup \partial_{ext} W) = \emptyset$ and so for each component U of $\partial_{int} W \cup \partial_{ext} W$, χ is constant on $U \cap \partial_{int} W$ and on $U \cap \partial_{ext} W$ and in neither case does it take on the value 0.

Fix $S \subseteq W^s$. We show that if $\{u, v\}$ is an edge of $T_{L,d}$ then $\chi_S^s(u) \neq \chi_S^s(v)$. We consider five cases.

If $u, v \notin W$ then $\chi_S^s(u) = \chi(u)$ and $\chi_S^s(v) = \chi(v)$. But $\chi(u) \neq \chi(v)$, so $\chi_S^s(u) \neq \chi_S^s(v)$ in this case.

If $u \in W$ and $v \notin W$ then $\chi_S^s(v) = \chi(v)$ and $\chi_S^s(u) \in \{0, \chi(u)\}$ (we will justify this in a moment). Since $v \in \partial_{ext} W$ we have $\chi(v) \neq 0$ and we cannot ever have $\chi(v) = \chi(u)$, so $\chi_S^s(u) \neq \chi_S^s(v)$ in this case. To see that $\chi_S^s(u) \in \{0, \chi(u)\}$, we consider subcases. If $u \in S$ then $\chi_S^s(u) = 0$. If $u \in W^s \setminus S$ then $\chi_S^s(u) = \chi(u)$. Finally, if $u \in W \setminus W^s$ then $\chi_S^s(u) = f(\chi(\sigma_{-s}(u)))$; and $f(\chi(\sigma_{-s}(u)))$ is either 0 or $\chi(u)$ depending on whether $\chi(\sigma_{-s}(u))$ equals 0 or $\chi(v)$ ($\chi(\sigma_{-s}(u))$ cannot equal $\chi(u)$).

If $u, v \in W \setminus W^s$ then $\chi_S^s(u) = f(\chi(\sigma_{-s}(u)))$ and $\chi_S^s(v) = f(\chi(\sigma_{-s}(v)))$. Since f is a bijection and $\chi(\sigma_{-s}(u)) \neq \chi(\sigma_{-s}(v))$ we have $\chi_S^s(u) \neq \chi_S^s(v)$ in this case.

If $u \in W \setminus W^s$ and $v \in W^s \setminus S$ then $\chi_S^s(u) \in \{0, \chi(u)\}$ (as in the second case above) and $\chi_S^s(v) = \chi(v)$. Since $\chi(v) \neq 0$, we have $\chi_S^s(u) \neq \chi_S^s(v)$.

Noting that it is not possible to have both $u, v \in W^s$, we finally treat the case where $u \in W \setminus W^s$ and $v \in S$. In this case $\chi_S^s(v) = \chi(v) = 0$. Suppose (for a contradiction) that $\chi_S^s(u) = 0$. This can only happen if $\chi(\sigma_{-s}(u)) = 0$. If $\sigma_{-s}(u) = v$, we have a contradiction immediately. Otherwise, we have $\sigma_{-s}(v) \notin W$ and so (since $\sigma_{-s}(u)\sigma_{-s}(v) \in E$) $\sigma_{-s}(u) \in \partial_{int} W$, also a contradiction.

This verifies $\varphi_s(\chi) \subseteq \mathcal{C}_3$. Because $\operatorname{int} \gamma$ is disjoint from the interiors of the remaining cutsets in $\Gamma(I)$ and the operation that creates the elements of $\varphi_s(\chi)$ only modifies χ inside W it follows that $\varphi_s(\chi) \subseteq \mathcal{C}_3(\underline{p'})$. □

Claim 4.2. *Given $\chi' \in \varphi_s(\chi)$, χ can be uniquely reconstructed from W and s.*

Proof: Following [9], we may reconstruct χ as follows.

$$\chi(v) = \begin{cases} \chi'(v) & \text{if } v \in V \setminus W \\ f(\chi'(\sigma_s(v))) & \text{if } v \in W. \end{cases}$$

□

We define the one-to-many map φ from $\mathcal{C}_3(\underline{p})$ to $\mathcal{C}_3(\underline{p'})$ by setting $\varphi(\chi) = \varphi_s(\chi)$ for a particular direction s. To define ν and s, we employ the notion of approximation also used in [8] and based on ideas introduced by Sapozhenko in [16]. For $\gamma \in \mathcal{W}$, we say $A \subseteq V$ is an *approximation* of γ if

$$A^{\mathcal{E}} \supseteq W^{\mathcal{E}} \quad \text{and} \quad A^{\mathcal{O}} \subseteq W^{\mathcal{O}},$$

$$d_{A^{\mathcal{O}}}(x) \geq 2d - \sqrt{d} \ \text{ for all } x \in A^{\mathcal{E}}$$

and

$$d_{\mathcal{E} \setminus A^{\mathcal{E}}}(x) \geq 2d - \sqrt{d} \ \text{ for all } y \in \mathcal{O} \setminus A^{\mathcal{O}},$$

where $d_X(x) = |\partial x \cap X|$. Note that from (3.3) and (3.5), $W(\gamma)$ is an approximation of γ.

Before stating our main approximation lemma, which is a slight modification of [8, Lemma 2.18], it will be convenient to further refine our partition of cutsets. To this end set

$$\mathcal{W}(w_e, w_o, v) = \left\{ \gamma : \begin{array}{l} \gamma \in \Gamma(I) \text{ for some } \chi \in \mathcal{C}_3^{even} \\ \text{with } |W^{\mathcal{O}}| = w_o,\ |W^{\mathcal{E}}| = w_e \\ \text{and } v \in W^{\mathcal{E}} \end{array} \right\}.$$

Note that by (3.5) we have $|\gamma| = 2d(|W^{\mathcal{O}}| - |W^{\mathcal{E}}|)$ so $\mathcal{W}(w_e, w_o, v) \subseteq \mathcal{W}((w_o - w_e)/2d, v)$.

Lemma 4.1. *For each w_e, w_o and v there is a family $\mathcal{A}(w_e,w_o,v)$ satisfying*

$$|\mathcal{A}(w_e,w_o,v)| \leq \exp\left\{O\left((w_o-w_e)d^{-\frac{1}{2}}\log^{\frac{3}{2}} d\right)\right\}$$

and a map $\pi : \mathcal{W}(w_e,w_o,v) \to \mathcal{A}(w_e,w_o,v)$ such that for each $\gamma \in \mathcal{W}(w_e,w_o,v)$, $\pi(\gamma)$ is an approximation for γ.

Proof: See [7, Lemma 4.2]. □

We are now in a position to define ν and s. Our plan for each fixed $\chi' \in \mathcal{C}_3(\underline{p}')$ is to fix w_e, w_o and $A \in \mathcal{W}(w_e,w_o,v)$ and to consider the contribution to the sum in (4.24) from those $\chi \in \varphi^{-1}(\chi')$ with $\pi(\gamma)=A$ (where for each χ, γ is a particular $\gamma \in \Gamma(I)\cap\mathcal{W}(c_1,v_1)$). We will try to define ν in such a way that each of these individual contributions to (4.24) is small; to succeed in this endeavor we must first choose s with care. To this end, given $\gamma \in \mathcal{W}(w_e,w_o,v)$, set

$$Q^{\mathcal{E}} = A^{\mathcal{E}} \cap \partial_{ext}(\mathcal{O}\setminus A^{\mathcal{O}}) \text{ and } Q^{\mathcal{O}} = (\mathcal{O}\setminus A^{\mathcal{O}})\cap \partial_{ext}A^{\mathcal{E}},$$

where $A=\pi(\gamma)$ in the map guaranteed by Lemma 4.1. To motivate the introduction of $Q^{\mathcal{E}}$ and $Q^{\mathcal{O}}$, note that for $\gamma \in \pi^{-1}(A)$ we have (by (3.3) and (3.5))

$$A^{\mathcal{E}} \setminus Q^{\mathcal{E}} \subseteq W^{\mathcal{E}},$$

$$\mathcal{E}\setminus A^{\mathcal{E}} \subseteq \mathcal{E}\setminus W^{\mathcal{E}},$$

$$A^{\mathcal{O}} \subseteq W^{\mathcal{O}},$$

and

$$\mathcal{O}\setminus (A^{\mathcal{O}}\cup Q^{\mathcal{O}}) \subseteq \mathcal{O}\setminus W^{\mathcal{O}}.$$

It follows that for each $\gamma \in \pi^{-1}(A)$, $Q^{\mathcal{E}}\cup Q^{\mathcal{O}}$ contains all vertices whose location in the partition $T_{L,d} = W\cup\overline{W}$ is as yet unknown. We choose $s(\chi)$ to be the smallest s for which both of $|W^s| \geq .8(w_o-w_e)$ and $|\sigma_s(Q^{\mathcal{E}})\cap Q^{\mathcal{O}}| \leq 5|W^s|/\sqrt{d}$ hold. This is the direction that minimizes the uncertainty to be resolved when we attempt to reconstruct χ from the partial information provided by $\chi' \in \varphi^{-1}(\chi)$, s and A. (That such an s exists is established in [8, (49) and (50)] by an easy averaging argument). Note that s depends on γ but not I.

Now for each $\chi \in \mathcal{C}_3(\underline{p})$ let $\gamma \in \Gamma(I)$ be a particular cutset with $\gamma \in \mathcal{W}(c_1,v_1)$. Let $\varphi(\chi)$ be as defined before, with s as specified above. Define

$$C = W^s \cap A^{\mathcal{O}} \cap \sigma_s(Q^{\mathcal{E}})$$

and

$$D = W^s \setminus C,$$

and for each $\chi' \in \varphi(\chi)$ set

$$\nu(\chi,\chi') = \left(\frac{1}{4}\right)^{|C\cap I(\chi')|}\left(\frac{3}{4}\right)^{|C\setminus I(\chi')|}\left(\frac{1}{2}\right)^{|D|}.$$

Note that for $\chi \in \varphi^{-1}(\chi')$, $\nu(\chi,\chi')$ depends on W but not on χ itself.

Since $C\cup D$ partitions W we easily have (4.23). To obtain (4.22) and so (3.7) we must establish (4.24).

Fix w_e, w_o such that $2d(w_o-w_e)=c_1$. Fix $A \in \mathcal{A}(w_e,w_o,v_1)$ and $s \in \{\pm1,\dots,\pm d\}$. For χ with $\gamma \in \mathcal{W}(w_e,w_o,v_1)$ write $\chi \sim_s A$ if it holds that $\pi(\gamma)=A$ and $s(\chi)=s$. We claim that with A, s, w_o and w_e fixed, for $\chi' \in \mathcal{C}_3(\underline{p}')$
(4.25)

$$\sum\{\nu(\chi,\chi') : \chi \sim_s A,\ \chi \in \varphi^{-1}(\chi')\} \leq \left(\frac{\sqrt{3}}{2}\right)^{w_o-w_e}.$$

We could extract this directly from [9], but for the convenience of the reader we describe a proof below.

Write $\mathcal{C}_3(\underline{p})(w_e,w_o,s,A,\chi')$ for the set of all $\chi \in \mathcal{C}_3(\underline{p})$ such that $W \in \mathcal{W}(w_e,w_o,v_1)$, $\pi(\gamma)=A$, $s(\chi)=s$ and $\chi' \in \varphi(\chi)$ and set $U = Q^{\mathcal{E}}\cap\sigma_{-s}(\chi')$. Say that a triple (K,L,M) is *good* for χ if it satisfies the following conditions.

$$K\cup L\cup M \text{ is a minimal vertex cover of } Q^{\mathcal{E}}\cup Q^{\mathcal{O}},$$

$$K \subseteq Q^{\mathcal{O}},\ L\subseteq U \text{ and } M \subseteq Q^{\mathcal{E}}\setminus U$$

and

$$K = \partial_{ext}(U\setminus L).$$

We begin by establishing that $\chi \in \mathcal{C}_3(\underline{p})(w_e,w_o,s,A,\chi')$ always has a good triple.

Lemma 4.2. *For each $\chi \in \mathcal{C}_3(\underline{p})(w_e,w_o,s,A,\chi')$ the triple*

$$(\hat{K},\hat{L},\hat{M}) := (W\cap Q^{\mathcal{O}}, U\setminus W, (Q^{\mathcal{E}}\setminus U)\setminus W)$$

is good for χ.

Proof: [8, around discussion of (54)]. □

In view of Lemma 4.2 there is a triple (K,L,M) that is good for χ and which has $|K|+|L|$ as small as possible. Choose one such, say $(K_0(\chi),L_0(\chi),M_0(\chi))$. Set $K'(\chi)=K_0\setminus\hat{K}$ and $L'(\chi)=L_0\setminus\hat{L}$. Lemma 4.3 below establishes an upper bound on $\nu(\chi,\chi')$ in terms of $|K_0|$, $|L_0|$, $|K'|$ and $|L'|$, and Lemma 4.4 shows that for each choice of K', L' there is at most one χ contributing to the sum in the lemma. These two lemmas combine to give (4.25).

Lemma 4.3. *For each $\chi \in \mathcal{C}_3(\underline{p})(w_e,w_o,s,A,\chi')$,*

$$\begin{aligned}\nu(\chi,\chi') &\leq \left(\frac{\sqrt{3}}{2}\right)^{w_o-w_e}\frac{2^{|K_0|}}{3^{|K_0|+|L_0|}2^{|K'|-|L'|}}\\ &:= B(K',L').\end{aligned}$$

Proof: We follow [8, from just before (55) to just after (60)], making superficial changes of notation. □

The inequality in Lemma 4.3 is the 3-coloring analogue of the main inequality of [8]. The key observation that makes this inequality useful is the following.

LEMMA 4.4. *For each w_e, w_o, s, A, χ', K' and L', there is at most one χ with $\chi \in \mathcal{C}_3(\underline{p})(w_e, w_o, s, A, \chi')$, $K' = K'(\chi)$ and $L' = L'(\chi)$.*

Proof: In [8, (56) and following] it is shown that K' and L' determine $W^{\mathcal{O}}$ via

$$\hat{K} = (K_0 \setminus K') \cup (\partial_{ext} L' \cap Q^{\mathcal{O}})$$

and so W (via $W^{\mathcal{E}} = \{v \in \mathcal{E} : \partial v \subseteq W^{\mathcal{O}}\}$). But then by Claim 4.2 K' and L' determine χ. □

Lemmas 4.3 and 4.4 together now easily give (4.25):

$$\begin{aligned} \sum_{\chi \in \mathcal{C}_3(\underline{p})(w_e, w_o, s, A, \chi')} \nu(\chi, \chi') &\leq \sum_{K' \subseteq K_0,\ L' \subseteq L_0} B(K', L') \\ &\leq \left(\frac{\sqrt{3}}{2}\right)^{w_o - w_e}. \end{aligned}$$

We have now almost reached (4.24). With the steps justified below we have that for each $\chi' \in \mathcal{C}_3(\underline{p'})$

$$\begin{aligned} \sum_{\chi \in \varphi^{-1}(\chi')} \nu(\chi, \chi') &\leq {\sum}' \left\{ \nu(\chi, \chi') \colon \begin{array}{l} \chi \sim_s A, \\ \chi \in \varphi^{-1}(\chi') \end{array} \right\} \\ &\leq 2dc_1^{\frac{2d}{d-1}} |\mathcal{A}(w_e, w_o, v_1)| \left(\frac{\sqrt{3}}{2}\right)^{\frac{c_1}{2d}} && (4.26) \\ &\leq 2dc_1^{\frac{2d}{d-1}} \exp\{-\Omega(c_1/d)\} && (4.27) \\ &\leq \exp\{-\Omega(c_1/d)\}, && (4.28) \end{aligned}$$

completing the proof of (4.24). In the first inequality, $\sum'$ is over all choices of w_e, w_o, s and A. In (4.26), we note that there are $|\mathcal{A}(w_e, w_o, v_1)|$ choices for A, $2d$ choices for s and $c_1^{d/(d-1)}$ choices for each of w_e, w_o (this is because $c_1 \geq (w_e + w_o)^{1-1/d}$, by (3.6)), and we apply (4.25) to bound the summand. In (4.27) we use Lemma 4.1. Finally in (4.28) we use $c_1 \geq d^{1.9}$ (again by (3.6)) to bound $2dc_1^{2d/(d-1)} = \exp\{o(c_1/d)\}$.

References

[1] D. Achlioptas, M. Molloy, C. Moore, and F. Van Bussel, Sampling grid colourings with fewer colours, *Proc. of LATIN '04*, 80–89.

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

[3] R. Bubley, M. Dyer and C. Greenhill, Beating the 2Δ bound for approximately counting colourings: a computer-assisted proof of rapid mixing, *Proc. of the 9th ACM-SIAM SODA '98*, 355–363.

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

[5] S.J. Ferreira and A.D. Sokal, Antiferromagnetic Potts model on the square lattice: a high precision Monte Carlo study, *J. Stat. Phys.* **96** (1999), 461–530.

[6] A. Frieze and E. Vigoda, A survey on the use of Markov chains to randomly sample colourings, to appear in *Combinatorics, Complexity and Chance.*

[7] D. Galvin, Sampling independent sets from the discrete torus, submitted (preprint available on the web at `www.math.upenn.edu/~dgalvin/pdf/indmixZd.pdf`).

[8] D. Galvin and J. Kahn, On phase transition in the hard-core model on $\mathbb{Z}^d$, *Comb. Prob. Comp.* **13** (2004), 137–164.

[9] D. Galvin, J. Kahn, D. Randall and G. Sorkin, On phase transition in the 3-coloring model on $\mathbb{Z}^d$, in preparation.

[10] L.A. Goldberg, R. Martin and M. Paterson, Random sampling of 3-colourings in $\mathbb{Z}^2$, *Random Structures and Algorithms* **24** (2004), 279–302.

[11] T. Hayes and E. Vigoda, Coupling with the stationary distribution and improved sampling for colorings and independent sets, *Proc. of the 16th Annual ACM-SIAM SODA '05*, 971-979.

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

[13] M. Jerrum and A. Sinclair, The Monte Carlo Markov chain method: an approach to approximate counting and integration, in *Approximation Alorithms for NP-hard problems*, PWS, 1996.

[14] G.F. Lawler and A.D. Sokal, Bounds on the L_2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality, *Trans. Amer. Math. Soc.* **309** (1988), 557-580.

[15] M. Luby, D. Randall, and A.J. Sinclair, Markov Chains for Planar Lattice Structures, *SIAM Journal on Computing* **31** (2001), 167–192.

[16] A.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.)

[17] A.J. Sinclair, *Algorithms for random generation & counting: a Markov chain approach,* Birkhäuser, Boston, 1993.

[18] L.E. Thomas, Bound on the mass gap for finite volume stochastic Ising models at low temperature, *Commun. Math. Phys.* **126** (1989), 1-11.

[19] J.S. Wang, R.H. Swendsen, and R. Kotecký, Three-state antiferromagnetic Potts models: A Monte Carlo study, *Phys. Rev. B* **42** (1990) 2465–2474.

[20] D. Welsh, Complexity: Knots, Colourings, and Counting, *London Math. Soc. Lec. Note Series* **186**, 1993.