On the Sampling Problem for H-Colorings on the Hypercubic Lattice

Christian Borgs* Jennifer T. Chayes* Martin Dyer† Prasad Tetali‡

April 15, 2002

Abstract

We consider the problem of random H-colorings of rectangular subsets of the hypercubic lattice $\mathbb{Z}^d$, with weight $\lambda_i \in (0, \infty)$ for the color i. First, we assume that H is non-trivial in the sense that it is neither the completely looped complete graph nor the complete bipartite graph. We consider quasi-local Markov chains on a periodic box of even side length L, that is, Markov chains that do not change more than a fraction $\rho < 1$ of the sites in the box in any single move. For any finite, connected, non-trivial H, we show that there are weights $\{\lambda_i\}$ such that all quasi-local ergodic Markov chains have slow mixing in the sense that the mixing time is exponential in L^{d-1}. Under the same conditions, we prove phase coexistence in the sense that there are at least two extremal Gibbs states. We also prove that, for a large subclass of graphs H, one can choose weights $\{\lambda_i\}$ such that the corresponding Gibbs measure has exponentially fast spatial mixing.

*Microsoft Research, 1 Microsoft Way, Redmond, WA 98052, borgs@microsoft.com, jchayes@microsoft.com

†School of Computing, University of Leeds, dyer@comp.leeds.ac.uk.

‡School of Mathematics, Georgia Tech, Atlanta, GA 30332, tetali@math.gatech.edu. Research supported in part by the NSF Grant No. DMS–9800351

1 Introduction

Let $H = (W, F)$ be a finite graph without multiple edges, but possibly with loops, and let $G = (V, E)$ be a simple, locally finite graph with countably many vertices. An H-coloring of G is a homomorphism from G to H, i.e., a mapping $\boldsymbol{\omega} : V \to W : x \mapsto \omega(x)$ such that $\langle \omega(x), \omega(y) \rangle \in F$ for all edges $\langle x, y \rangle \in E$. As usual, we denote the set of all H-colorings of G by $\mathrm{Hom}(G, H)$.

When sampling H-colorings, it is often useful to assign different weights to the different colors in W. Given a set of weights $\{\lambda_i\}_{i \in W}$, $0 < \lambda_i < \infty$, and a finite graph G, we thus consider the distribution μ with weights

$$\mu(\omega) = \frac{1}{Z} \prod_{x \in V} \lambda_{\omega(x)}, \tag{1}$$

where Z is the normalization factor

$$Z = \sum_{\boldsymbol{\omega} \in \mathrm{Hom}(G,H)} \prod_{x \in V} \lambda_{\omega(x)}. \tag{2}$$

If G is infinite, we consider the set of Gibbs measures with weights $\{\lambda_i\}_{i \in W}$ on $\mathrm{Hom}(G, H)$. To make this precise, we first define a local observable as a function $f : \mathrm{Hom}(G, H) \to \mathbb{R} : \boldsymbol{\omega} \mapsto f(\boldsymbol{\omega})$ which does not depend on ω_x for all but a finite number of vertices $x \in V$. We then equip $\mathrm{Hom}(G, H)$ with the minimal sigma algebra $\mathcal{F}$ which makes all local observables measurable. As usual, a probability measure μ on $(\mathrm{Hom}(G, H), \mathcal{F})$ is called a Gibbs measure with weights $\{\lambda_i\}_{i \in W}$ if, given any finite set $\Lambda \subset V$, and any function $\boldsymbol{\omega}_{\Lambda^c}$ from $\Lambda^c = V \setminus \Lambda$ into W which can be extended to an H-coloring of G, the conditional measure $\mu(\cdot \mid \boldsymbol{\omega}_{\Lambda^c})$ is given by

$$\mu(\boldsymbol{\omega}_\Lambda \mid \boldsymbol{\omega}_{\Lambda^c}) = \begin{cases} \dfrac{1}{Z(\boldsymbol{\omega}_{\Lambda^c})} \displaystyle\prod_{x \in \Lambda} \lambda_{\omega(x)} & \boldsymbol{\omega} \in \mathrm{Hom}(G, H) \\ 0 & \boldsymbol{\omega} \notin \mathrm{Hom}(G, H). \end{cases} \tag{3}$$

Here $Z(\boldsymbol{\omega}_{\Lambda^c})$ is the normalization factor

$$Z(\boldsymbol{\omega}_{\Lambda^c}) = \sum_{\boldsymbol{\omega}_\Lambda : \boldsymbol{\omega} \in \mathrm{Hom}(G,H)} \prod_{x \in \Lambda} \lambda_{\omega(x)}. \tag{4}$$

As usual, we will often use the term Gibbs state for a Gibbs measure μ with weights $\{\lambda_i\}_{i \in W}$, and call such a measure extremal if it is not possible to write μ as a convex combination of two different Gibbs measures μ_A and μ_B with weights $\{\lambda_i\}_{i \in W}$.

In the past few years, there has been a good deal of work on the H-coloring problem. First, there are a few papers concerning the complexity of the H-coloring problem. An important result of Hell and Nešetřil [21] characterized the complexity of the decision problem, i.e., under what conditions does there exist an H-coloring: if H has no loops and is not bipartite, then they showed that the decision problem is NP-complete; otherwise it is trivially in P. More recently, Dyer and Greenhill [18] characterized the complexity of the counting problem, i.e., the size of $\mathrm{Hom}(G, H)$: If H has a component that is neither the completely looped complete graph, K_n^{loop}, nor the complete bipartite graph, $K_{n,m}$, then the counting problem is $\sharp$P-complete; otherwise, it is trivially in P. The $\sharp$P-completeness remains true even if G has bounded degrees. Henceforth, we will say that

H *is trivial* if all its connected components are completely looped complete graphs or complete bipartite graphs.

There have been many results on almost uniform sampling (and approximate counting), using the Markov chain Monte Carlo (MCMC) method, of H-colorings or weighted H-colorings for specific H, including independent sets, the Widom-Rowlinson model, the beach model, and when H is a tree. Positive (fast mixing) results include those of Jerrum [22], Luby and Vigoda [24], Vigoda [31], Dyer and Greenhill [19] and Cooper, Dyer and Frieze [13]. Negative (slow-mixing) results include those of Thomas [30], Borgs, Chayes, Frieze, Kim, Tetali, Vigoda and Vu [4], Dyer, Frieze and Jerrum [17], and Cooper, Dyer and Frieze [13]. The third and the fourth of these slow mixing results are of the form: for uniform weights on H, there exists a graph G such that the mixing is slow. The first and second results are of a very different nature: if G is the hypercubic lattice, there exist weights on the graph H such that the mixing is slow; the specific H in [30] being an edge with loops on both the vertices, corresponding to the Ising model, and in [4] being an edge with a loop on a single vertex, capturing the independent set model. The slow-mixing results we derive in the present work are a significant generalization of the latter results to all non-trivial coloring graphs H.

The final category of previous results concerns Gibbs states for H-colorings. In 1968, Dobrushin showed that, provided that $\mathrm{Hom}(G, H) \neq \emptyset$, there exists at least one Gibbs measure for any set of activities $\{\lambda_i\}_{i \in W}$. It is then natural to ask for what sets of activities this measure is unique, and when there are non-unique measures. That there are activities for which the measure is unique was shown by van den Berg [2] for $G = \mathbb{Z}^d$ and H being the independent set model or Widom-Rowlinson model, by Burton and Steif [12] if $G = \mathbb{Z}^d$ and H the beach model, by Brightwell and Winkler [8] if G is a tree, and again by Brightwell and Winkler [9] if G is of bounded degree and H satisfies a certain condition called "dismantlability." Burton and Steif [11] also established the existence of weights for which there are non-unique measures if $G = \mathbb{Z}^d$ and H the beach model, thereby proving a phase transition in this case. Similarly, Brightwell and Winkler established the existence of weights for which the Gibbs measure is non-unique if G is a tree and if either H is what they called "fertile" [8] or if H is not dismantlable and satisfies a few additional constraints [9]; in the latter case, the model had infinitely many Gibbs measures. Some other results on non-existence or existence of phase transitions for specific H are described in Brightwell, Häggström and Winkler [7].

In this work, we will focus on G being the hypercubic lattice, and show that, for a large class of graphs H, there are weights for which the measure is unique and others for which it is non-unique, thus establishing the existence of a phase transition. While the precise statements of our results are described in the next section, we point out that our main objective here is a complete characterization of H-colorings not only in terms of unique and nonunique Gibbs measures, but fast and slow mixing of the associated Markov chains.

2 Statements of Results

In this paper we consider random H-coloring of the hypercubic lattice $\mathbb{Z}^d$ and the d-dimensional torus $(\mathbb{Z}/L\mathbb{Z})^d$, where L is assumed to be even. As usual, two vertices $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ in $\mathbb{Z}^d$ are joined by an edge whenever there is a j so that $|x_i - y_i| = \delta_{i,j}$ for all i; similarly, $x, y \in (\mathbb{Z}/L\mathbb{Z})^d$ are joined by an edge if there is a j so that $|x_i - y_i| = \delta_{i,j} \mod L$. We denote the corresponding graphs by $G = (V, E)$ and $G_L = (V_L, E_L)$, and the set of all H-colorings of G and G_L by Ω and Ω_L. On Ω_L, we consider the measure

$$\mu_L(\boldsymbol{\omega}) = \frac{1}{Z_L} \prod_{x \in V_L} \lambda_{\omega(x)}, \tag{5}$$

where

$$Z_L = \sum_{\boldsymbol{\omega} \in \Omega_L} \prod_{x \in V_L} \lambda_{\omega(x)} \tag{6}$$

while on Ω, we consider the set of all Gibbs measures defined by (3). As usual, we say that a Gibbs measure μ has exponentially decaying correlations if there exist an $\epsilon > 0$ such that for all local observables f one has

$$\sum_{x \in V} \Big| E_\mu[f\, T_x f] - E_\mu[f] E_\mu[T_x f] \Big| e^{\epsilon |x|} < \infty. \tag{7}$$

Here E_μ denotes expectations with respect to μ, $T_x f$ stands for the translate of f by x, i.e. $[T_x f](\boldsymbol{\omega}) := f(\boldsymbol{\omega}^{(x)})$, where $\boldsymbol{\omega}^{(x)}$ is the "shifted" coloring defined by $\omega^{(x)}(y) := \omega(y - x)$, and $|x|$ denotes the ℓ_1 norm on $\mathbb{Z}^d$.

Let $H = (W, F)$ be bipartite, with $W = W_{\text{even}} \cup W_{\text{odd}}$. We call the vertices in W_{even} even, and the vertices in W_{odd} odd. In a similar way, the sets V and V_L can be split into sets of even and odd vertices. Following the usual convention, we call a vertex x in V or V_L even if the sum of its coordinates $\sum_{i=1}^d x_i$ is even, and odd otherwise, and write V^{even} or V_L^{even} for the set of even, and V^{odd} or V_L^{even} for the set of odd vertices.

If H is bipartite, the measure μ_L can be naturally decomposed as

$$\mu_L = \frac{1}{2}(\mu_L^+ + \mu_L^-), \tag{8}$$

where μ_L^+ lives on the space Ω_L^+ of H-colorings $\boldsymbol{\omega} \in \Omega_L$ that map even vertices into even vertices, and μ_L^- lives on $\Omega_L^- = \Omega_L \setminus \Omega_L^+$. More explicitly, μ_L^+ and μ_L^- are given by

$$\mu_L^\pm(\boldsymbol{\omega}) = \begin{cases} \dfrac{1}{Z_L^\pm} \displaystyle\prod_{x \in V_L} \lambda_{\omega(x)} & \text{if } \boldsymbol{\omega} \in \Omega^\pm \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

where

$$Z_L^\pm = \sum_{\boldsymbol{\omega} \in \Omega_L^\pm} \prod_{x \in V_L} \lambda_{\omega(x)}. \tag{10}$$

In a similar way, any Gibbs measure on Ω decomposes naturally into a convex combination of Gibbs measures on Ω^+ and Ω^-, where Ω^+ is the space of H-colorings $\boldsymbol{\omega} \in \Omega$ that map even vertices into even vertices, and $\Omega^- = \Omega \setminus \Omega^+$.

Given two H-colorings $\omega, \omega' \in \Omega_L$, let $D(\omega, \omega')$ be the number of vertices $x \in V_L$ such that $\omega(x) \neq \omega'(x)$. For a Markov chain $\mathcal{M}_L$ on Ω_L, let $D(\mathcal{M}_L)$ be the maximum of $D(\omega, \omega')$ over all ω and ω' for which the transition probability is non-zero. We say that $\mathcal{M}_L$ is local if $D(\mathcal{M}_L)$ is bounded uniformly in L, and we say that it is ρ-quasi-local if $D(\mathcal{M}_L) \leq \rho L^d$ for some $\rho < 1$

which is independent of L. As usual, the mixing time of an ergodic Markov chain $\mathcal{M}_L$ on Ω_L is defined as the time after which the variational distance from the stationary measure is small enough, say smaller than $1/2e$; see (29) in Section 5 for a precise definition. If H is bipartite, no local or quasi-local Markov chain connects the two components Ω_L^+ and Ω_L^- of Ω_L. We therefore consider the restrictions of $\mathcal{M}_L$ to Ω_L^+ and Ω_L^- separately.

Henceforth, we will consider only connected, non-trivial graphs H, i.e., connected graphs which are neither $H = K_n^{\text{loop}}$ nor $H = K_{n,m}$. Recall that trivial graphs H lead to counting problems in P.

Our first set of results establish the existence of weights that lead to fast spatial mixing (exponentially decaying correlations) and a unique Gibbs state for a large class of constraint graphs.

Theorem 1 *Let H be a non-trivial connected graph, and let $d \geq 1$.*
i) If H contains at least one loop, there are weights $\{\lambda_i\}$ such that the limit

$$\mu = \lim_{L\to\infty} \mu_L \tag{11}$$

exists and describes a translation-invariant, extremal Gibbs measure with exponentially decaying correlations.
ii) If H is bipartite, the analogue of statement (i) holds separately for the limits of each of the measures $\mu_L^\pm$,

$$\mu^\pm = \lim_{L\to\infty} \mu_L^\pm, \tag{12}$$

except that the translation invariance is replaced by periodicity with period 2.

Remark 1 *Assume in addition that H is dismantlable in the sense of [9]. Combining the methods of [9] with those presented here, we can show that the weights $\{\lambda_i\}_{i\in W}$ can be chosen in such a way that the Gibbs state is unique and is given by (11).*

Remark 2 *If W contains a vertex i such that $\langle i,j\rangle \in F$ for all $j \in W$, the single-site Dobrushin criterion is satisfied provided λ_i is chosen large enough. This immediately implies uniqueness of the Gibbs state and fast mixing for the standard single-site Heat bath algorithm.*

Remark 3 *Theorem 1 does not cover all connected, non-trivial graphs H. E.g., it does not apply to loopless graphs with at least one odd cycle, such as K_3. Indeed, if $H = K_3$, it is believed that for d large enough and any weights $\{\lambda_i\}_{i\in W}$, the limiting measure (11) is a convex combination of at least two extremal Gibbs states.*

Our next two results establish the existence of weights that lead to non-unique Gibbs states and slow mixing for all non-trivial connected H.

Theorem 2 *Let H be a non-trivial connected graph, and let $d \geq 2$. If H is not bipartite, then there are weights $\{\lambda_i\}$ such that the limit (11) exists and is a convex combination*

$$\mu = \frac{1}{2}(\mu_A + \mu_B), \tag{13}$$

where μ_A and μ_B are extremal Gibbs states with exponentially decaying correlations. If H is bipartite, the analogous statement holds separately for the measures $\mu^\pm$, so that there are at least four extremal Gibbs states.

Theorem 3 *Let H be a non-trivial connected graph, let $\rho < 1$ and let $d \geq 2$. If H is not bipartite, then there are weights $\{\lambda_i\}$ such that for all L sufficiently large, the mixing time τ_L of any ρ-quasi-local ergodic Markov chain with stationary distribution μ_L is exponentially large, i.e.,*

$$\tau_L \geq e^{K_1 L^{d-1}/(\log L)^2}, \tag{14}$$

where K_1 is a constant that depends on d, ρ and the weights $\{\lambda_i\}$. If H is bipartite, the analogous statement holds for any ρ-quasi-local ergodic Markov chains on Ω_L^+ and Ω_L^-.

These results are proved using expansion methods. Theorem 1 is a so-called "high-temperature" or "disordered phase" result, and the proof is relatively straightforward. Namely, we find weights $\{\lambda_i\}$ so that the H-coloring problem maps into what is called a dilute polymer model, and then use standard Mayer expansions [29, 10, 16, 25] for the dilute polymer model to prove Theorem 1. This is sketched in Section 3. The details may be found in [5].

Theorems 2 and 3 are the so-called "low-temperature" or "ordered phase" results, the proofs of which are much more involved than those of the high-temperature results. First, we find a suitable classification of all non-trivial coloring graphs H. Within each class, we find weights $\{\lambda_i\}$ so that the H-coloring problem maps into what is called a dilute contour model. Here the contours separate different ordered phases. We then use rather involved expansion methods, namely Pirogov-Sinai theory [27, 28] in the form developed by Borgs and Imbrie [6], to control this contour model and prove Theorem 2. This is sketched in Section 4. Finally, in Section 5, we sketch how these Pirogov-Sinai methods can be combined with the conductance bounds of [4] to prove Theorem 3. Again, the details will be given in the full version of this paper; we also hope that the full version introduces and further illustrates the usefulness of the statistical physics techniques (such as Mayer expansions and Pirogov-Sinai theory) in the context of Markov chain Monte-Carlo algorithms.

3 Polymer Expansions and Fast Spatial Mixing

In this section, we sketch the proof of Theorem 1. To this end, we map the partition function of our model to an abstract polymer system with sufficiently small weights. As usual, an abstract polymer system is a triple $\mathbf{\Gamma} = (\Gamma, \leftrightarrow, z(\cdot))$, where Γ is a finite set, $\leftrightarrow$ is a symmetric, reflexive relation on Γ, and $z(\cdot)$ is a complex-valued function on Γ. The elements of Γ are called polymers. Two polymers $\gamma, \gamma' \in \Gamma$ are said to be incompatible if $\gamma \leftrightarrow \gamma'$, and compatible otherwise. Finally, $z(\gamma)$ is called the weight or activity of the polymer γ. The partition function of the polymer system $\mathbf{\Gamma}$ is defined as

$$\mathcal{Z}(\mathbf{\Gamma}) = \sum_{\tilde{\Gamma} \subset \Gamma} \phi(\tilde{\Gamma}) \prod_{\gamma \in \tilde{\Gamma}} z(\gamma), \tag{15}$$

where $\phi(\tilde{\Gamma}) = 0$ whenever there is a pair of polymers $\gamma, \gamma' \in \tilde{\Gamma}$ such that $\gamma \leftrightarrow \gamma'$, and $\phi(\tilde{\Gamma}) = 1$ otherwise. In other words $\phi(\tilde{\Gamma}) = 1$ whenever the polymers in $\tilde{\Gamma}$ are pairwise compatible, and $\phi(\tilde{\Gamma}) = 0$ otherwise.

In order to prove Theorem 1, we will choose weights λ_i such that the partition functions (6) and (10) can be written in terms of a polymer system with small weights. Together with a similar representation for the measures (5) and (9), the general theory of Mayer expansions for abstract polymer systems [29, 10, 16, 25] then gives Theorem 1. For the case where H has at least one loop, this will be sketched in Subsection 3.1, and for the case where H is bipartite this will be sketched in Subsection 3.2.

3.1 At least one loop present

Let H be a graph with at least one loop, and without loss of generality let us assume that the vertices in W are labelled in such a way that the loop $\ell_1 = \langle 1, 1\rangle \in F$. We then set

$$\lambda_i = \begin{cases} \lambda & \text{if } i = 1 \\ 1 & \text{otherwise.} \end{cases} \tag{16}$$

If λ is large, the configuration $\boldsymbol{\omega}$ with largest weight is the one where every site has color $\omega(x) = 1$. For a general configuration $\boldsymbol{\omega}$, we define a vertex $x \in V_L$ to be excited whenever $\omega(x) \neq 1$, and call the connected components of the set of excited vertices the polymers corresponding to $\boldsymbol{\omega}$. A configuration $\boldsymbol{\omega}$ with polymers $\gamma_1, \dots, \gamma_n$ then has weight $\lambda^{|V_L|} \prod_{i=1}^n \lambda^{-|\gamma_i|}$, where $|V_L|$ and $|\gamma_i|$, $i = 1, \dots, n$, denotes the number of vertices in V_L and γ_i, respectively. This motivates the following definition of a polymer system $\boldsymbol{\Gamma}_1$: The set of polymers, Γ_1, is defined as the set of connected subsets $\gamma \subset V_L$. Two polymers γ and γ' are called incompatible if $\gamma \cup \gamma'$ is a connected subset of V_L. The weight $z(\gamma)$ of a polymer γ is defined as $\lambda^{-|\gamma|}$ times the number of H-colorings $\boldsymbol{\omega}$ with $\boldsymbol{\omega}^{-1}(\{1\}) = V_L \setminus \gamma$.

Lemma 1 *Let $\boldsymbol{\Gamma}_1$ be the polymer system defined above, and let Z_L be the partition function defined in* (6)*. Then*

$$Z_L = \lambda^{|V_L|} \mathcal{Z}(\boldsymbol{\Gamma}_1). \tag{17}$$

Proof: Given a subset $W \subset V_L$, let $N_{V_L}(W)$ be the number of H-colorings $\boldsymbol{\omega}$ such that W is the set of excited sites corresponding to $\boldsymbol{\omega}$, i.e., $N_{V_L}(W) = |\{\boldsymbol{\omega} \colon \boldsymbol{\omega}^{-1}(\{1\}) = V_L \setminus W\}|$. By definition, the polymers corresponding to an H-coloring $\boldsymbol{\omega}$ form a set of pairwise compatible polymers $\tilde{\Gamma} \subset \boldsymbol{\Gamma}_1$. Since the weight of an H-coloring with polymers $\gamma_1, \dots, \gamma_n$ is $\lambda^{|V_L|} \prod_{i=1}^n \lambda^{-|\gamma_i|}$, the left hand side of (17) is therefore equal to

$$\lambda^{|V_L|} \sum_{n \geq 0} \sum_{\{\gamma_1, \dots, \gamma_n\}} N_{V_L}(\gamma_1 \cup \dots \cup \gamma_n) \prod_{i=1}^n \lambda^{-|\gamma_i|}, \tag{18}$$

where the second sum goes over sets of pairwise compatible polymers. Since the the weight $z(\gamma)$ of a polymer γ can be rewritten as $\lambda^{-|\gamma|} N_{V_L}(\gamma)$, the proof of the lemma reduces therefore to the fact that $N_{V_L}(W) = \prod_{i=1}^n N_{V_L}(W_i)$ whenever $W_1, \dots, W_n$ are the connected components of W. The proof of this fact is elementary and is left to the reader. □

Given Lemma 1, the partition function (6) can be analyzed using the general theory of Mayer expansions for abstract polymer systems. In order to apply this theory, one has to show that the weights $z(\gamma)$ are small enough. A sufficient condition is that

$$\sum_{\gamma \in \Gamma : \gamma \leftrightarrow \gamma'} z(\gamma) e^{|\gamma|} \leq |\gamma'| \tag{19}$$

for all $\gamma' \in \Gamma$, see [10, 16, 25]. Since the number of connected sets $\gamma \subset V_L$ of size s that have distance 1 or less from a given point $x \in V_L$ can be bounded by $(2de)^s$, while $N_{V_L}(\gamma)$ is obviously at most $(|W|-1)^{|\gamma|}$, the condition (19) is satisfied whenever

$$\lambda > 4de^2(|W|-1). \tag{20}$$

For $\lambda > 4de^2(|W|-1)$, the partition function Z_L can therefore be analyzed with the help of the general theory of Mayer expansion for abstract polymer systems. Together with a similar representation for the measure $\mu_L(\cdot)$, one obtains a proof of Theorem 1 (i); see [5] for details.

3.2 Bipartite

Assume without loss of generality that the vertices in W are labelled in such a way that F contains the edge $\langle 1, 2\rangle$. We then set

$$\lambda_i = \begin{cases} \lambda & \text{if } i = 1 \text{ or } 2 \\ 1 & \text{otherwise.} \end{cases} \tag{21}$$

If λ is large, the configurations with maximal weight are the two checkerboard configurations $\boldsymbol{\omega}_A$ and $\boldsymbol{\omega}_B$: in $\boldsymbol{\omega}_A$, every even site in V_L has color 1 and every odd site has color 2, and vice versa in $\boldsymbol{\omega}_B$. Restricting ourselves to the configurations in Ω_L^+ rules out either $\boldsymbol{\omega}_A$ or $\boldsymbol{\omega}_B$, so that we are left with only one configuration of maximal weight. For a general configuration $\boldsymbol{\omega} \in \Omega_L^+$, we define the set of excited vertices as the set of vertices x with color $\omega(x) \in W \setminus \{1,2\}$. The connected components of the set of excited vertices are again called the polymers corresponding to $\boldsymbol{\omega}$.

As in the argument in the last subsection, this leads to a polymer representation of the partition function which can be analyzed using convergent Mayer expansions if λ is sufficiently large. The only difference here is the fact that the weights $z(\gamma)$ are no longer translation invariant. Again, see [5] for details.

4 Contour Representations and Phase Coexistence

In this section we derive a contour representation for random H-colorings which will allow us to sketch the proofs of Theorems 2 and 3. To this end, we first note that all connected, non-trivial graphs H fall into one of the following four classes:

1) all loops present
2) not all loops, but at least one loop, present
3) no loops present, but at least one odd cycle
4) bipartite (and thus loopless)

Our contour model will be defined differently in each case, but in all cases contours will be pairs $\gamma = (\operatorname{supp}\gamma, \boldsymbol{\omega}_\gamma)$, where $\operatorname{supp}\gamma$ is a $*$-connected subset of W, and $\boldsymbol{\omega}_\gamma$ is an H-coloring of $\operatorname{supp}\gamma \cup \partial\operatorname{supp}\gamma$. Here $\Lambda \subset V_L$ is called $*$-connected if for every pair of vertices $x, y \in \Lambda$ there is a path of vertices $x_1, \dots, x_k \in \Lambda$ with $x_1 = x$ and $x_k = y$ such that for all $i = 1, \dots, k-1$, x_i and x_{i+1} have Euclidean distance $\sqrt{2}$ or less. As usual, the boundary $\partial\Lambda$ of a set $\Lambda \subset V_L$ is the set of all vertices in $V_L \setminus \Lambda$ that are connected to Λ via an edge in E_L. Throughout this section, two contours γ and γ' will be called compatible if the Euclidean distance between $\operatorname{supp}\gamma$ and $\operatorname{supp}\gamma'$ is strictly larger than $\sqrt{2}$, and a set $\{\gamma_1, \dots, \gamma_n\}$ of contours will be called compatible if γ_i and

γ_j are compatible for all pairs $i \neq j$. Finally the size $|\gamma|$ of a contour $\gamma = (\operatorname{supp}\gamma, \boldsymbol{\omega}_\gamma)$ will be defined as the number of vertices in $\operatorname{supp}\gamma$.

4.1 All loops present

Since by assumption, H is not the completely looped complete graph, there must be vertices i and j in W, such that $\langle i, j\rangle$ is not an edge in H; without loss of generality we assume that the vertices in W are labelled in such a way that $\langle 1, 2\rangle \notin F$. We then set

$$\lambda_i = \begin{cases} \lambda_A & \text{if } i = 1 \\ \lambda_B & \text{if } i = 2 \\ 1 & \text{otherwise.} \end{cases} \tag{22}$$

If λ_A and λ_B are large, the dominant configurations $\boldsymbol{\omega} \in \Omega_L$ will be the constant configurations $\boldsymbol{\omega}_A \equiv 1$ and $\boldsymbol{\omega}_B \equiv 2$, with weights $\lambda_A^{|V_L|}$ and $\lambda_B^{|V_L|}$, respectively.

For a general configuration $\boldsymbol{\omega}$, we define the "ground state regions" $V_A(\boldsymbol{\omega})$ and $V_B(\boldsymbol{\omega})$ as $V_A(\boldsymbol{\omega}) = \boldsymbol{\omega}^{-1}(\{1\})$ and $V_B(\boldsymbol{\omega}) = \boldsymbol{\omega}^{-1}(\{2\})$, and the "set of excited vertices" $V^*(\boldsymbol{\omega})$ as $V^*(\boldsymbol{\omega}) = V_L \setminus (V_A(\boldsymbol{\omega}) \cup V_B(\boldsymbol{\omega}))$. The contours corresponding to the H-coloring $\boldsymbol{\omega}$ are then defined as the pairs $\gamma_1 = (\operatorname{supp}\gamma_1, \boldsymbol{\omega}_{\gamma_1})$, $\dots$, $\gamma_n = (\operatorname{supp}\gamma_n, \boldsymbol{\omega}_{\gamma_n})$, where $\operatorname{supp}\gamma_1$, $\dots$, $\operatorname{supp}\gamma_n$ are the $*$-connected components of $V^*(\boldsymbol{\omega})$, and $\boldsymbol{\omega}_{\gamma_1}, \dots, \boldsymbol{\omega}_{\gamma_n}$ are the restrictions of $\boldsymbol{\omega}$ to $\operatorname{supp}\gamma_1 \cup \partial\operatorname{supp}\gamma_1$, $\dots$, $\operatorname{supp}\gamma_n \cup \partial\operatorname{supp}\gamma_n$, respectively. The weight of an H-coloring $\boldsymbol{\omega}$ can then be written as a product of suitable weights for the contours and ground state regions corresponding to $\boldsymbol{\omega}$,

$$\prod_{x \in V_L} \lambda_{\omega(x)} = e^{-e_A|V_A(\boldsymbol{\omega})|} e^{-e_B|V_B(\boldsymbol{\omega})|} \prod_{i=1}^{n} \rho(\gamma_i), \tag{23}$$

with $e_A = -\log\lambda_A$, $e_B = -\log\lambda_B$ and $\rho(\gamma) = \prod_{x \in \operatorname{supp}\gamma} \lambda_{\omega(x)} = 1$ for all contours γ (recall that $\lambda_i = 1$ for $i \neq 1, 2$). As is usual in cluster expansions, we will need a Peierls' bound of the form

$$\rho(\gamma) \leq e^{-\tau|\gamma|} e^{-e_0|\gamma|}, \tag{24}$$

where $e_0 = \min\{e_A, e_B\}$ is the ground state energy and $\tau > 0$ is a "suppression factor" used to control the expansion. Here, since $\rho(\gamma) = 1$ and $e_0 = -\max\{\log\lambda_A, \log\lambda_B\}$, the contour weights trivially have an exponential suppression with respect to the corresponding ground state weights provided λ_A and λ_B are large, and (24) holds with $\tau = \max\{\log\lambda_A, \log\lambda_B\}$.

A general configuration thus consists of regions in one of the two ground states, separated by regions which have much less weight. If the vertices 1 and 2 are "isomorphic" vertices (i.e., if there is an automorphism which transposes 1 and 2) in H, we set $\lambda_A = \lambda_B = \lambda$. For large λ, a Peierls argument then implies the existence of at least two translation invariant Gibbs states μ_A and μ_B related to each other by the symmetry $1 \leftrightarrow 2$. Here μ_A consists of small fluctuations around the ground state $\boldsymbol{\omega}_A$, and μ_B consists of small fluctuations around the ground state $\boldsymbol{\omega}_B$, in close analogy to the two low temperature states of the Ising model which are small perturbation of the ground states where all spins are up and down, respectively.

If 1 and 2 are not isomorphic, the fluctuation about the ground states $\boldsymbol{\omega}_A$ and $\boldsymbol{\omega}_B$ will in general favor one of the two, and only one of them will contribute to the limiting state (11) for $\lambda_A = \lambda_B$.

To correct for this, we will set $\lambda_A = \lambda e^h$ and $\lambda_B = \lambda e^{-h}$. For λ large and a suitable choice of h (which will in general depend on λ) the difference in the weight for the ground state $\boldsymbol{\omega}_A$ and $\boldsymbol{\omega}_B$ will exactly compensate for the difference induced by the fluctuations, so that we again get two different, translation invariant Gibbs states μ_A and μ_B which are small perturbations of the ground states $\boldsymbol{\omega}_A$ and $\boldsymbol{\omega}_B$. The precise argument is rather complicated; it uses the version of Pirogov-Sinai theory [27] developed in [6] and is given in [5]. These methods also imply exponential decay of correlations and extremality for μ_A and μ_B, whether or not the vertex 1 and 2 are related by symmetry. To prove that the Gibbs state obtained via the limit (11) is a convex combination of μ_A and μ_B with equal weight for μ_A and μ_B (see (13)), we use the methods of Section 5 of [6]; see again [5] for details.

4.2 Not all, but at least one loop present

Consider two vertices i and j in W such that the loop $\ell_i \notin F$ and the loop $\ell_j \in F$. Since H is connected, there must be a pair of vertices $\tilde{i}$ and $\tilde{j}$ in W such that $\ell_{\tilde{i}} \notin F$, $\ell_{\tilde{j}} \in F$ and $\langle \tilde{i}, \tilde{j} \rangle \in F$. Without loss of generality we assume that H has been labelled in such a way that $\tilde{i} = 1$ and $\tilde{j} = 2$, so that $\ell_1 \notin F$, $\ell_2 \in F$ and $\langle 1, 2 \rangle \in F$. We then set

$$\lambda_i = \begin{cases} \lambda & \text{if } i = 1 \\ 1 & \text{otherwise.} \end{cases} \tag{25}$$

If λ is large, one would like to set the color of all vertices to 1. But since $\ell_1 \notin F$, this is not a configuration in $\Omega_L = \mathrm{Hom}(G_L, H)$. The best one can do is to put color 1 on every second vertex, i.e., to either color all even or all odd vertices with the color 1.

In general, there are many configurations $\boldsymbol{\omega} \in \Omega_L$ with $\omega(x) = 1$ for all $x \in V_L^{\text{even}}$ or $\omega(x) = 1$ for all $x \in V_L^{\text{odd}}$, namely $|\mathcal{N}(1)|^{|V_L|/2}$, where $\mathcal{N}(1)$ is the neighborhood of 1 in H, $\mathcal{N}(1) = \{i \in W \colon \langle 1, i \rangle \in F\}$. In the language of statistical physics, this means that these states have ground state entropy $S_0 = s_0 |V_L|$, with $s_0 = \frac{1}{2} \log |\mathcal{N}(1)|$. Note that in contrast to the situation in the last subsection, a "ground state" of the system is not a single configuration, but rather a measure supported on many configurations. Here the two ground state measures, ν_A and ν_B, are defined by

$$\nu_A(\boldsymbol{\omega}) = \prod_{x \in V_L^{\text{even}}} \mathbb{I}(\omega(x) = 1) \prod_{y \in V_L^{\text{odd}}} \mathbb{I}(\omega(y) \in \mathcal{N}(1)) \tag{26}$$

and a similar equation for ν_B with the roles of V_L^{even} and V_L^{odd} switched. Here $\mathbb{I}(\mathcal{E})$ stands for the indicator function of the event $\mathcal{E}$.

To define the contours corresponding to a configuration $\boldsymbol{\omega} \in \Omega_L$, we consider again three regions V_A, V_B and V^*. The set V_A consists of all even vertices x with color $\omega(x) = 1$ and all odd vertices which have only neighbors of color 1, the set V_B consists of all odd vertices x with color $\omega(x) = 1$ and all even vertices which have only neighbors of color 1, and the set V^* consists of the remaining vertices in V_L. For the ground state configurations described above, the set V^* is empty, and either V_A or V_B consists of all vertices in V_L. The contours corresponding to a general configuration $\boldsymbol{\omega} \in \Omega_L$ are again defined as the pairs $\gamma_1 = (\operatorname{supp} \gamma_1, \boldsymbol{\omega}_{\gamma_1})$, $\dots$, $\gamma_n = (\operatorname{supp} \gamma_n, \boldsymbol{\omega}_{\gamma_n})$, where $\operatorname{supp} \gamma_1$, $\dots$, $\operatorname{supp} \gamma_n$ are the $*$-connected components of V^*, and $\boldsymbol{\omega}_{\gamma_1}$, $\dots$, $\boldsymbol{\omega}_{\gamma_n}$ are the restrictions of $\boldsymbol{\omega}$ to $\operatorname{supp} \gamma_1 \cup \partial \operatorname{supp} \gamma_1$, $\dots$, $\operatorname{supp} \gamma_n \cup \partial \operatorname{supp} \gamma_n$, respectively.

The weight of a set of contours $\Gamma = \{\gamma_1, \dots, \gamma_n\}$ can again be written as a product of suitable weights for the contours and ground state regions V_A and V_B; however, the explicit formula is slightly more complicated due to the existence of the extra entropy factors $|\mathcal{N}(1)|$ for the sites in V_A and V_B whose color is not fixed. Nevertheless, it is possible to rewrite the weight of an H-coloring $\boldsymbol{\omega}$ with contours $\gamma_1, \dots, \gamma_n$ and ground state regions V_A and V_B in the form (23), with $e_A = e_B = -\log(\lambda^{1/2}|\mathcal{N}(1)|^{1/2})$ and weights $\rho(\gamma)$ for the contours which obey a bound of the form (24) with $\tau = \Theta\big(\frac{1}{d}\log\lambda\big)$; see [5] for details.

Again, a general configuration consists of regions in one of the two ground states, separated by regions which have much less weight. For large λ, a modified Peierls argument like the one used by Dobrushin in his proof of a phase transition for the independent set model [15] then implies the existence of at least two Gibbs states μ_A and μ_B related to each other by the symmetry of shifting each contour by one lattice unit. Here μ_A consists of small fluctuations around the ground state ν_A, and μ_B consists of small fluctuations around the ground state ν_B. While the emerging picture is very similar to that of the independent set model on the level of contours, it is slightly more complicated on the level of configurations, since the ground states we perturb around are now product measures, not delta functions on a single configuration. Nevertheless, the techniques developed for the independent set model [15] can be generalized to this case. Combined with the methods of Section 5 of [6], this leads to the proof of Theorem 2 in the case where H has at least one loop, but does not have all loops. Again see [5] for details.

4.3 No loop present, at least one odd cycle

This case is completely analogous to the case with at least one, but not all loops present. Indeed, let i be a vertex contained in an odd cycle, without loss of generality the vertex labelled by 1. If we set

$$\lambda_i = \begin{cases} \lambda & \text{if } i = 1 \\ 1 & \text{otherwise,} \end{cases} \tag{27}$$

there are again two ground state measures for large λ, namely the measure ν_A defined in (26) and its analog ν_B. While the precise contour weights will be different from those for the case where at least one, but not all loops present, the general features of the resulting contour model will be very similar. In particular, we can again use a modified Peierls argument to infer the existence of at least two Gibbs states μ_A and μ_B related to each other by the symmetry of shifting each contour by one lattice unit, with μ_A consisting of small fluctuations around the ground state ν_A, and μ_B consisting of small fluctuations around the ground state ν_B.

4.4 Bipartite

We recall that the vertex set of a bipartite graph H naturally splits into two subsets W_{odd} and W_{even}, such all edges in F are of the form $\langle i, j\rangle$, $i \in W_{\text{odd}}$, $j \in W_{\text{even}}$. Since H is not the complete bipartite graph by the assumption that it is non-trivial, there exist two vertices $\widetilde{i} \in W_{\text{odd}}, \widetilde{j} \in W_{\text{even}}$ such that $\langle \widetilde{i}, \widetilde{j}\rangle \notin F$. Without loss of generality, assume $\widetilde{i} = 1$ and $\widetilde{j} = 2$. Obviously, neither ℓ_1

nor ℓ_2 are in F by the assumption that H is bipartite. We set

$$\lambda_i = \begin{cases} \lambda_A & \text{if } i = 1 \\ \lambda_B & \text{if } i = 2 \\ 1 & \text{otherwise.} \end{cases} \tag{28}$$

Restricting ourselves to colorings in Ω_L^+, we obtain two ground states: the ground state ν_A, where every odd vertex has color 1, while the colors of the even vertices are chosen independently at random from $\mathcal{N}(1)$, and the ground state ν_B, where every even vertex has color 2, while the colors of the odd vertices are chosen independently at random from $\mathcal{N}(2)$. The weights of these ground states are $\lambda_A^{|V_L|/2}|\mathcal{N}(1)|^{|V_L|/2}$ and $\lambda_B^{|V_L|/2}|\mathcal{N}(2)|^{|V_L|/2}$, respectively.

For a general configuration $\boldsymbol{\omega} \in \Omega_L^+$, the ground state region V_A now consists of all vertices x with color $\omega(x) = 1$ and all vertices y which have only neighbors of color 1, and the ground state region V_B consists of all vertices x with color $\omega(x) = 2$ and all vertices which have only neighbors of color 2. Setting again $V^* = V_L \setminus (V_A \cup V_B)$ we define contours as before, obtaining again a representation of the form (23), with $e_A = -\log\left(\lambda_A^{1/2}|\mathcal{N}(1)|^{1/2}\right)$, $e_B = -\log(\lambda_B^{1/2}|\mathcal{N}(2)|^{1/2})$, and contour weights $\rho(\gamma)$ which obey a bound of the form (24) provided λ_A and λ_B are large enough. If 1 and 2 are isomorphic vertices in H, then a Peierls argument proves that typical configurations are either small perturbations of ground state ν_A, or small perturbations of ground state ν_B if $\lambda_A = \lambda_B$. If there is no symmetry, one has to adjust the ratio of λ_A and λ_B to correct for entropic preference of one of the two states. The proof again uses Pirogov-Sinai theory, and is carried out in [5]. Of course each of the above Gibbs states has a ghost sister supported in Ω^- where the even and odd sublattices of V_L exchange their roles, leading to a total of at least four extremal Gibbs states.

5 Mixing Time and Conductance Bounds

In this section, we sketch the key ideas of the proof of Theorem 3. The proof uses the notion of conductance, first introduced to the field of MCMC by Jerrum and Sinclair in [23]. We start with a few general definitions.

Let $\mathcal{M}$ be an ergodic Markov chain on a finite state space Ω, with transition probabilities $P(\boldsymbol{\omega}, \widetilde{\boldsymbol{\omega}})$, $\boldsymbol{\omega}, \widetilde{\boldsymbol{\omega}} \in \Omega$. Let π denote the stationary distribution of $\mathcal{M}$. For $\boldsymbol{\omega}_0 \in \Omega$, we denote by $P_{t,\boldsymbol{\omega}_0}(\boldsymbol{\omega})$ the probability that the system is in the state $\boldsymbol{\omega}$ at time t given that $\boldsymbol{\omega}_0$ is the initial state. The mixing time of the Markov chain $\mathcal{M}$ is defined as

$$\tau = \min\left\{t : \max_{\boldsymbol{\omega}\in\Omega} d(P_{t,\boldsymbol{\omega}}, \pi) \le \frac{1}{2e}\right\}, \tag{29}$$

where $d(P_{t,\boldsymbol{\omega}}, \pi)$ is the *variational distance* between $P_{t,\boldsymbol{\omega}}$ and π,

$$d(P_{t,\boldsymbol{\omega}}, \pi) = \max_{S\subseteq\Omega} |P_{t,\boldsymbol{\omega}}(S) - \pi(S))|\,. \tag{30}$$

The conductance of a set of states $\emptyset \neq S \subset \Omega$ is

$$\Phi_S = \sum_{\boldsymbol{\omega}\in S} \sum_{\widetilde{\boldsymbol{\omega}}\in\Omega\setminus S} \frac{\pi(\boldsymbol{\omega})P(\boldsymbol{\omega}, \widetilde{\boldsymbol{\omega}})}{\pi(S)\pi(\Omega \setminus S)}, \tag{31}$$

and the conductance of the chain itself is simply $\Phi_{\mathcal{M}} = \min_{S\neq\emptyset} \Phi_S$.

We prove our lower bounds on mixing time by showing that $\Phi_{\mathcal{M}}$ is small and then using the well-known bound [1] (see [14] or Claim 2 of the journal version of [17] for the non-reversible case):

$$\tau^{-1} = O(\Phi_{\mathcal{M}}). \tag{32}$$

Here, the finite state space is the space of all H-colorings Ω_L (or the spaces $\Omega_L^{\pm}$ if H is bipartite). Our goal is to decompose Ω_L as a disjoint union of three sets $\Omega_{L,A}$, $\Omega_{L,B}$ and Ω_L^* such that $\pi(\Omega_L^*) = O(e^{-KL^{d-1}/(\log L)^2})$, while both $\pi(\Omega_{L,A})$ and $\pi(\Omega_{L,B})$ are $\Theta(1)$, and such that *for any* ρ-quasi-local Markov chain, the transition probability $P(\boldsymbol{\omega}, \widetilde{\boldsymbol{\omega}})$ is zero whenever $\boldsymbol{\omega} \in \Omega_{L,A}$ and $\widetilde{\boldsymbol{\omega}} \in \Omega_{L,B}$. Taking $S = \Omega_{L,A} \cup \Omega_L^*$, we then have

$$\Phi_{\mathcal{M}} \leq \Phi_S = \sum_{\boldsymbol{\omega}\in\Omega_L^*} \sum_{\widetilde{\boldsymbol{\omega}}\in\Omega_{L,B}} \frac{\pi(\boldsymbol{\omega})P(\boldsymbol{\omega},\widetilde{\boldsymbol{\omega}})}{\pi(\Omega_{L,A}\cup\Omega_L^*)\pi(\Omega_{L,B})} \leq \frac{\pi(\Omega_L^*)}{\pi(\Omega_{L,A})\pi(\Omega_{L,B})} = O(e^{-KL^{d-1}/(\log L)^2}), \tag{33}$$

which combined with (32) gives Theorem 3.

More precisely, we define $\Omega_{L,A}$, $\Omega_{L,B}$ and Ω_L^* in such a way that $|V_A(\boldsymbol{\omega})| > (1-\epsilon)|V_L|$ for all $\boldsymbol{\omega} \in \Omega_{L,A}$ and $|V_B(\boldsymbol{\omega})| > (1-\epsilon)|V_L|$ for all $\boldsymbol{\omega} \in \Omega_{L,B}$ and then combine the methods developed in [4] with Pirogov-Sinai theory to show that for λ sufficiently large $\mu_L(\Omega_L^*) \leq e^{-KL^{d-1}/(\log L)^2}$, $\mu_L(\Omega_{L,A}) = \Theta(1)$ and $\mu_L(\Omega_{L,B}) = \Theta(1)$; see [5] for details. If we choose ϵ as $\frac{1}{2}(1-\rho)$, then we have that for all ρ-quasi-local Markov chains the transition probability $P(\boldsymbol{\omega}, \widetilde{\boldsymbol{\omega}})$ is zero whenever $\boldsymbol{\omega} \in \Omega_{L,A}$ and $\widetilde{\boldsymbol{\omega}} \in \Omega_{L,B}$, and (33) implies the statement of Theorem 3.

In the rest of this section, we proceed to give the proof of a weakened form of (33) in a particularly simple case. As already noted in [4], it is significantly easier to prove a bound of the form (33) in which the exponential factor of $L^{d-1}/(\log L)^2$ is replaced by $L/(\log L)^2$. Here we sketch a proof of this weakened inequality for a subclass of the graphs H considered in Subsection 4.1.

Suppose that H has all loops present, so that W contains two vertices (w.l.o.g., we assume they are labelled 1 and 2) such that $\langle 1,2\rangle \notin F$; see Section 4.1. Suppose further that 1 and 2 are isomorphic vertices in H, so that the statements of Theorem 2 hold whenever the weights $\{\lambda_i\}_{i\in W}$ are chosen according to (22) with $\lambda_A = \lambda_B = \lambda$ sufficiently large. In particular this latter assumption allows us to use a Peierls' argument rather than the considerably more involved methods of Pirogov-Sinai theory.

In a first step, we introduce the space $\widetilde{\Omega}_L^*$ as the set of H-colorings $\boldsymbol{\omega}$ such that at least one of the contours corresponding to $\boldsymbol{\omega}$ has size L or larger. In order to bound $\pi(\widetilde{\Omega}_L^*)$, we use the fact that the probability that a given contour γ is among the contours corresponding to a configuration $\boldsymbol{\omega}$ can be bounded by $\rho(\gamma)$. (Since vertices 1 and 2 are isomorphic, so that the states A and B are related by a symmetry, this can be proved via a standard Peierls argument; see [26], [20].) The weight of the space $\widetilde{\Omega}_L^*$ can therefore be bounded by

$$\mu_L(\widetilde{\Omega}_L^*) \leq \sum_{\gamma:|\gamma|\geq L} \tilde{\rho}(\gamma), \tag{34}$$

where the sum goes over all contours of size $|\gamma| \geq L$ and $\tilde{\rho}(\gamma) := \rho(\gamma)e^{e_0|\gamma|}$ measures the suppression with respect to the ground state energy (recall that $e_0 = e_A = e_B$ due to our symmetry assumptions). To estimate this sum, we note that the number of vertices $y \in V_L$ that have distance at most $\sqrt{2}$ from a given vertex $x \in V_L$ is equal to $2d+2d(2d-2)$, implying that the number of $*$-connected sets of size s in V_L can be bounded by $|V_L|(2d(2d-1)e)^s$. As a consequence, the number of contours γ that have size $|\gamma| = s$ can be bounded by $L^d C(d,|W|)^s$ for some constant $C(d,|W|)$ depending on d and the size of W. Together with the bound (24) and the fact that $\tau = \log\lambda$, we therefore get that whenever λ is sufficiently large, the weight of $\widetilde{\Omega}^*_L$ is bounded by

$$\mu_L(\widetilde{\Omega}^*_L) \leq L^d \sum_{s\geq L} \Big(C(d,|W|)\lambda^{-1}\Big)^s = O\Big(L^d\Big(C(d,|W|)\lambda^{-1}\Big)^L\Big). \tag{35}$$

Consider now a contour γ corresponding to a configuration $\boldsymbol{\omega} \in \Omega_L \setminus \widetilde{\Omega}^*_L$, i.e. a contour γ with $|\gamma| < L$. Since the diameter of the torus G_L is L, such a contour can be embedded in $\mathbb{Z}^d$. Defining $V(\gamma)$ as the union of $\operatorname{supp}\gamma$ with all finite components of $\mathbb{Z}^d \setminus \operatorname{supp}\gamma$, we then introduce the exterior of γ as $V_L \setminus V(\gamma)$. This in turn allows us to define the exterior $\operatorname{Ext}(\boldsymbol{\omega})$ of a configuration $\boldsymbol{\omega} \in \Omega_L \setminus \widetilde{\Omega}^*_L$ as the intersection of the exteriors of all contours corresponding to $\boldsymbol{\omega}$. Using our definition of contours as pairs $(\gamma, \operatorname{supp}\gamma)$, where $\operatorname{supp}\gamma$ is a $*$-connected component of $V^*_L(\boldsymbol{\omega})$, it is not hard to verify that the exterior $\operatorname{Ext}(\boldsymbol{\omega})$ is a connected set, and that on $\operatorname{Ext}(\boldsymbol{\omega})$, $\boldsymbol{\omega}$ is constant and either equal to 1 or equal to 2. This allows us to decompose $\boldsymbol{\omega} \in \Omega_L \setminus \widetilde{\Omega}^*_L$ into two sets: one, denoted by $\widetilde{\Omega}_{L,A}$, for which $\operatorname{Ext}(\boldsymbol{\omega}) \subset V_A(\boldsymbol{\omega})$, and one, denoted by $\widetilde{\Omega}_{L,B}$, for which $\operatorname{Ext}(\boldsymbol{\omega}) \subset V_B(\boldsymbol{\omega})$.

At this point, the proof is a straightforward application of the methods developed in [4]: we define

$$\Omega_{L,A} = \{\boldsymbol{\omega} \in \widetilde{\Omega}_{L,A} \colon \operatorname{Ext}(\boldsymbol{\omega}) > (1-\epsilon)|V_L|\}, \tag{36}$$

and similarly for $\Omega_{L,B}$. Since $\operatorname{Ext}(\boldsymbol{\omega}) \subset V_A(\boldsymbol{\omega})$ if $\boldsymbol{\omega} \in \Omega_{L,A} \subset \widetilde{\Omega}_{L,A}$, we obviously have $|V_A(\boldsymbol{\omega})| > (1-\epsilon)|V_L|$ if $\boldsymbol{\omega} \in \Omega_{L,A}$, and similarly for $\boldsymbol{\omega} \in \Omega_{L,B}$. Defining $\Omega^*_L = \Omega_L \setminus (\Omega_{L,A} \cup \Omega_{L,B})$, we thus are left with a proof of the inequality $\mu(\Omega^*_L) \leq O(e^{-KL/(\log L)^2})$, for some absolute constant $K > 0$. To this end, we consider a configuration $\boldsymbol{\omega} \in \widetilde{\Omega}_{L,A} \setminus \Omega_{L,A}$, and the set of contours Γ corresponding to $\boldsymbol{\omega}$. Using the isoperimetric inequality of Bollobás and Leader [3], we then have

$$\epsilon L^d \leq |V_L \setminus \operatorname{Ext}(\boldsymbol{\omega})| \leq \sum_{\gamma\in\Gamma} |V(\gamma)| \leq \Theta(1) \sum_{\gamma\in\Gamma} |\gamma|^{d/(d-1)} \leq \Theta(1) \Big(\sum_{\gamma\in\Gamma} |\gamma|\Big)^{d/(d-1)}, \tag{37}$$

and hence $\sum_{\gamma\in\Gamma} |\gamma| \geq \Theta(L^{d-1})$. Proceeding with a multi-scale Peierls argument as in the proof of Lemma 8 in [4], one gets the bound

$$\mu_L(\widetilde{\Omega}_{L,A} \setminus \Omega_{L,A}) \leq O(e^{-KL^{d-1}/(\log L)^2}), \tag{38}$$

and similarly for $\mu_L(\widetilde{\Omega}_{L,B} \setminus \Omega_{L,B})$. Combining (35) with (38) and the fact that $\mu_L(\Omega^*_L) \leq \mu(\widetilde{\Omega}^*_L) + \mu_L(\widetilde{\Omega}_{L,A} \setminus \Omega_{L,A}) + \mu_L(\widetilde{\Omega}_{L,B} \setminus \Omega_{L,B})$, we thus get the desired bound $\mu_L(\Omega^*_L) \leq O(e^{-KL/(\log L)^2})$.

References

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

[2] J. van den Berg: *A uniqueness condition for Gibbs measures with appliction to the 2-dimensional Ising antiferromagnet.* Commun. Math. Phys. **152** (1993) 161–163.

[3] B. Bollobás and I. Leader: *Edge-isoperimetric inequalities in the grid*, Combinatorica **11** (1991) 299-314.

[4] C. Borgs, J. T. Chayes, A. Frieze, J. H. Kim, P. Tetali, E. Vigoda, V. Vu: *Torpid mixing of some MCMC algorithms in statistical physics.* Proc. 40^{th} IEEE Symp. on Found. of Comp. Sc. (1999) 218–229.

[5] C. Borgs, J. T. Chayes, M. Dyer, P. Tetali, Full version of this paper, in preparation.

[6] C. Borgs, J. Imbrie: *A Unified approach to phase diagrams in field theory and statistical mechanics.* Commun. Math. Phys. **123** (1989) 305–328.

[7] G.R. Brightwell, O. Häggström, P. Winkler: *Nonmonotonic behavior in hard-core and Widom-Rowlinson models.* J. Stat. Phys. **94** (1999) 415–435.

[8] G.R. Brightwell, P. Winkler: *Graph homomorphisms and phase transitions.* J. Combin. Theory **B77** (1999) 221–262.

[9] G.R. Brightwell, P. Winkler: *Gibbs measures and dismantlable graphs.* J. Combin. Theory **B78** (2000) 141–166.

[10] D. Brydges: *A short course on cluster expansions.* Proc. of the 1984 Les Houches Summer School. North Holland, Amsterdam, 1986, 129 – 1893.

[11] R. Burton, J. Steif: *Nonuniqueness of measures of maximal entropy for subshifts of finite type.* Ergodic Theory and Dynamical Systems **14** (1994) 213–235.

[12] R. Burton, J. Steif: *New results on measures of maximal entropy.* Israel J. of Math. **89** (1995) 275–300.

[13] C. Cooper, M. Dyer, A. Frieze: *On Markov chains for randomly H-colouring a graph.* preprint (2000).

[14] F. Chen, L. Lovász, I. Pak: Unpublished appendix for *Lifting Markov chains to speed up mixing.* Proc. 31^{st} ACM Symp. on Theory of Comp. (1999) 275–281.

[15] R.L. Dobrushin: *The problem of uniqueness of a Gibbsian random field and the problem of phase transitions.* (Russian) Funkcional. Anal. i Priložen. **2** (1968) 44–57.

[16] R.L. Dobrushin: *Estimates on semi-invariants for the Ising model at low temperatures.* In: Topics in Statistical and Theoretical Physics, Amer. Math. Soc. Trans., Ser. 2 **177** (1996) 59–81.

[17] M.E. Dyer, A. Frieze, M. Jerrum: *On counting independent sets in sparse graphs.* Proc. 40^{th} IEEE Symp. on Found. of Comp. Sc. (1999) 210–217.; journal version (submitted).

[18] M.E. Dyer, C.S. Greenhill: *The complexity of counting graph homomorphisms.* In Proc. 11^{th} ACM/SIAM Symp. on Disc. Alg. (2000) 246–255.

[19] M.E. Dyer, C.S. Greenhill: *On Markov chains for independent sets.* J. Algorithms **35** (2000) 17–49.

[20] R. Griffiths: *Peierls' proof of spontaneous magnetization in a two-dimensional Ising ferromagnet.* Phys. Rev.**136A** (1964) 437.

[21] P. Hell, J. Nešetřil: *On the complexity of H-colouring.* J. Combin. Theory B **48** (1990) 92–110.

[22] M. Jerrum: *A very simple algorithm for estimating the number of k-colorings of low-degree graphs.* Rand. Struc. Alg. **7** (1995) 157–165.

[23] M. Jerrum, A. Sinclair: *Approximate counting, uniform generation and rapidly mixing Markov chains.* Information and Computation **82** (1989) 93–133.

[24] M. Luby, E. Vigoda: *Approximate counting up to four.* Proc. 29^{th} ACM Symp. on Theory of Comp. (1997) 682–687.

[25] S. Miracle-Solé: *On the convergence of cluster expansions.* Physica **A279** (2000) 244–249.

[26] R. Peierls: *Ising's model of ferromagnetism.* Proc. Cambridge Phil. Soc. **32** (1936) 477.

[27] S.A. Pirogov, Ya.G. Sinai: *Phase diagrams of classical lattice systems.* Theor. Math. Phys. **25** (1975) 1185–1192.

[28] S.A. Pirogov, Ya.G. Sinai: *Phase diagrams of classical lattice systems. Continuation.* Theor. Math. Phys. **26** (1976) 39–49.

[29] E. Seiler: *Gauge Theories as a Problem of Constructive Quantum Field Theory and Statistical Mechanics.* Lecture Notes in Physics **159**. Springer, New York, 1982.

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

[31] E. Vigoda: *Improved bounds for sampling colorings.* Proc. 40^{th} IEEE Symp. on Found. of Comp. Sc. (1999) 51–59.