Slow Convergence of Ising and Spin Glass Models with Well-Separated Frustrated Vertices

David Gillman
Department of Computer Science, New College of Florida, [Sarasota, FL, 34243], USA
dgillman@ncf.edu

Dana Randall[1]
School of Computer Science, Georgia Institute of Technology, [Atlanta, GA 30332], USA
randall@cc.gatech.edu

Abstract

Many physical models undergo phase transitions as some parameter of the system is varied. This phenomenon has bearing on the convergence times for local Markov chains walking among the configurations of the physical system. One of the most basic examples of this phenomenon is the ferromagnetic Ising model on an $n \times n$ square lattice region Λ with mixed boundary conditions. For this spin system, if we fix the spins on the top and bottom sides of the square to be $+$ and the left and right sides to be $-$, a standard Peierls argument based on energy shows that below some critical temperature t_c, any local Markov chain $\mathcal{M}$ requires time exponential in n to mix.

Spin glasses are magnetic alloys that generalize the Ising model by specifying the strength of nearest neighbor interactions on the lattice, including whether they are ferromagnetic or antiferromagnetic. Whenever a face of the lattice is bounded by an odd number of edges with ferromagnetic interactions, the face is considered *frustrated* because the local competing objectives cannot be simultaneously satisfied. We consider spin glasses with exactly four well-separated frustrated faces that are symmetric around the center of the lattice region under 90 degree rotations. We show that local Markov chains require exponential time for all spin glasses in this class. This argument extends to the ferromagnetic Ising model with mixed boundary conditions described above, which behaves like spin glasses with frustrated faces on the boundary. The standard Peierls argument breaks down when the frustrated faces are on the interior of Λ and yields weaker results when they are on the boundary of Λ but not near the corners. We show that there is a universal temperature T below which $\mathcal{M}$ will be slow for all spin glasses with four well-separated frustrated faces. Our argument shows that there is an exponentially small cut indicated by the *free energy*, carefully exploiting both entropy and energy to establish a small bottleneck in the state space to establish slow mixing.

2012 ACM Subject Classification Theory of computation → Randomness, geometry and discrete structures → Random walks and Markov chains

Keywords and phrases Mixing time, spin glass, Ising model, mixed boundary conditions, frustration

Digital Object Identifier 10.4230/LIPIcs.AofA.2018.24

Acknowledgements The authors would like to thank the anonymous reviewers for many helpful suggestions.

[1] Supported in part by NSF grants CCF-1526900, CCF-1637031 and CCF-1733812.

1 Introduction

The celebrated Ising model on the Cartesian lattice is a fundamental model for ferromagnetism and one of the simplest models demonstrating an order-disorder phase transition. Each configuration σ in the state space $\Omega = \{-1, +1\}^{n^2}$ consists of an assignment of a + or − spin to each of the vertices, and the *Gibbs (or Boltzmann) distribution* assigns weight

$$\pi(\sigma) = e^{-\beta H(\sigma)}/Z(\beta),$$

where

$$H(\sigma) = -\sum_{(i,j)\in E} \sigma_i \sigma_j$$

is the *Hamiltonian* (or energy) of the system, $\beta = 1/T$ is inverse temperature, and $Z(\beta) = \sum_{\sigma\in\Omega} e^{-\beta H(\sigma)}$ is the normalizing constant known as the *partition function*. In Sections 3 and 4 it will be convenient to write the probability of a configuration in terms of $\lambda = e^{2\beta} = e^{2/T}$, where λ can be seen as the weight assigned to edges whose endpoints are assigned like spins.

Physicists characterize when there is a phase transition in a physical model by asking whether there is a unique limiting conditional distribution on finite subregions as the lattice size grows. The Gibbs distribution is defined as any limiting measure, but this limit might not be unique. For example, for the Ising model on $\mathbb{Z}^2$ at sufficiently low temperatures, the probability of an interior vertex being assigned + will be much higher if the boundary vertices were hard-wired to be + than if they were hard-wired to be −, and this difference persists in the limit. The infinite volume Ising model was solved exactly by Onsager in 1944 [23], showing that there is a critical value $\beta_c = \ln(1+\sqrt{2})/2$ such that for $\beta < \beta_c$ (i.e., high temperature), the limiting distribution is unique, and for $\beta > \beta_c$ (i.e., low temperature), spins on the boundary of the region persist and there are multiple limiting distributions. The all-plus and the all-minus boundary conditions are known to be extremal measures [1, 12].

A related effect has been observed in the context of mixing times of local Markov chains for the Ising model on finite lattice regions with free boundaries (i.e., boundary vertices can take on either spin). The *mixing time* $\tau(\mathcal{M})$ of a chain $\mathcal{M}$, i.e., the number of steps required so that the distribution over configurations is close to its stationary distribution, also undergoes a phase change. When β is small, local dynamics are known to be efficient [18, 19, 15], while when β is large, local chains require exponential time to converge to equilibrium [31]. At low enough temperature, the Gibbs distribution strongly favors configurations that have predominantly one spin, and it will take exponential time to move from a mostly + state to a mostly − one using moves that only change $o(n^2)$ sites at a time [17].

Mixing times of Markov chains are known to be sensitive to boundary conditions. For example, local chains on Ising configurations are conjectured to converge in polynomial time at all temperatures for the "all +" boundary condition where all vertices on the boundary are hard-wired to have + spins. While still open, Martinelli [16] showed mixing is indeed sub-exponential at all temperatures with all + boundary conditions and subsequently Lubetsky et al. [15] showed that the chain converges in quasi-polynomial time. However, a standard Peierls argument can be used to show that when there are mixed boundary conditions with 4 connected components of like spins on the boundary, alternating "$+, -, +, -$", then the chain again will be slow at low temperatures. In particular, for *mixed boundary conditions* where we fix the boundary to be + on the vertical sides and − on the horizontal sides, then the chain provably requires time exponential in n at sufficiently low temperature. For "p-shifted mixed boundary conditions" where we rotate the mixed boundary conditions clockwise by p units, We explain this in Section 3. More powerful machinery such as the approach of Dobrushin, Kotecký and Schlosman [6] for the Ising model establish bounds on the temperature below which convergence is slow, but they do not easily extend to other cases we consider.

Similar questions can be examined in the context of *spin glasses*, or magnetic alloys that are a natural generalization of the ferromagnetic and antiferromagnetic Ising models. We are given a graph $G = (V, E)$ and a set of couplings $J_{ij} \in \{-1, +1\}$ for each edge $(i, j) \in E$. The state space is $\Omega = \{-1, +1\}^V$, where a configuration assigns a spin to each vertex in V. For a spin glass configuration $\sigma \in \Omega$, the Hamiltonian is defined as

$$H(\sigma) = -\sum_{(i,j)\in E} J_{ij}\sigma(i)\sigma(j)$$

and the Gibbs distribution is defined as for the Ising model as $\pi(\sigma) = e^{-\beta H(\sigma)}/Z(\beta)$.

When all the $J_{ij} = +1$, this model is precisely the ferromagnetic Ising model on G; when all the $J_{ij} = -1$, it is antiferromagnetic. In general, the behavior of a spin glass is much richer than simple models of magnetism because of the presence of *frustration*, or competition between local interactions. In the case of $G = \Lambda$, a square region in the lattice, a face of Λ is *frustrated* when $J_{ij} = -1$ for an odd number of edges around the face. No setting of the sites on the corners of such a face will satisfy all four edges, i.e., make each $J_{ij}\sigma(i)\sigma(j) = 1$. Even finding the ground states (or most likely configurations) reduces to solving an optimization problem that can be NP-hard (see, e.g., [2]). It will be convenient to refer to the dual lattice $\overline{\Lambda} = (\overline{V}, \overline{E})$ and refer to a frustrated face f of Λ by the frustrated vertex $v = \overline{f}$ in $\overline{V}$.

Here, we study spin glasses with exactly four well-separated frustrated faces in order to understand the long-range interactions and their effects on mixing times. We fix the nearest-neighbor interactions around the boundary of Λ to be ferromagnetic, and we assume fixed + sites on the boundary. Similar models with well-separated defects have been explored to understand long-range correlation; for example, in seminal work, Ciucu [4] studied the monomer-dimer model with a constant number of monomers and established a connection with electrical networks, settling a nearly century old conjecture about long-range effects due to isolated monomers. Similar questions arise naturally in the context of spin glasses.

We show that there is a universal temperature T below which the Markov chain $\mathcal{M}$ will be slow for any spin glass with exactly four frustrated vertices that are well-separated and symmetric around the origin under 90 degree rotations. We identify a bottleneck in the state space by looking at the how the *free energy* (i.e., $\ln Z/n^2$) changes as a parameter of the system is varied. The same argument easily extends to the Ising model with p-shifted mixed boundary conditions, which behaves like spin glasses with four symmetric frustrated faces near the boundary (and indeed can be viewed as a special case of the spin glasses we consider if we also fix + spins adjacent to the boundary).

▶ **Theorem 1.1.** *Let Λ be a square lattice region with fixed + sites on the boundary of Λ and a fixed ferromagnetic interaction $J_{ij} = 1$ on each boundary edge (i, j). Suppose Λ has exactly four frustrated faces, $f_1, \ldots, f_4$, that are symmetric around the center of the lattice region under 90 degree rotations and are well-separated (i.e., the shortest lattice path from f_i to f_{i+1} has length $2n$, $i = 1, 2, 3$). Then there is a universal temperature $T = 0.360\ldots$ such that the Glauber dynamics $\mathcal{M}$ for the spin glass model on Λ with $f_1, ..., f_4$ the faces with frustration has mixing time $\tau(\mathcal{M}) \geq e^{cn}$, for some constant $c > 0$, whenever $t < T$.*

The theorem remains true under the additional assumption of fixed + sites adjacent to the boundary. As a corollary this gives a universal bound on the temperature for the Ising model with p-shifted mixed boundary conditions.

The proof of Theorem 1.1 requires several innovations. The standard argument to show slow mixing is based on the *conductance* of the Markov chain. The key is showing that the state space Ω can be partitioned into two sets, S and its complement S^C, such that getting from S to some subset S^C requires passing through a small cutset $\mathcal{C} \subset S^C$, and the stationary weights $\pi(S)$ and $\pi(S^C)$ are both exponentially larger than $\pi(\mathcal{C})$. This establishes that the chain has low conductance, which implies it takes exponential time to converge to equilibrium [13]. The main ingredient is typically a *Peierls argument* [24], which introduces a map Ψ from $\mathcal{C}$ to $S \cup S^C$. Typically Ψ is chosen so that for all $\sigma \in \mathcal{C}$, we have $\pi(\Psi(\sigma)) \geq \pi(\sigma)e^{cn}$, mapping elements of $\mathcal{C}$ to configurations with exponentially larger weight. If we can show that Ψ is nearly injective (i.e., the cardinality of the inverse image of each configuration is bounded by a polynomial), then we can conclude that $\pi(\mathcal{C})$ is exponentially small.

In our setting, there is not always a natural candidate map that increases the probability of a configuration exponentially. In fact, the standard map gives no guaranteed increase to the stationary probability when each side of the boundary has close to an equal number of $+$ and $-$ spins (when $p = 0.5$ and the boundary changes spin at the center of the four sides of the boundary). In this case, we exploit the low *entropy* of $\mathcal{C}$ by defining an injective map from $\mathcal{C} \times 2^{cn} \to \Omega$, for some $c > 0$. The map never decreases the weight of a configuration, so we again can conclude that $\pi(\mathcal{C})$ is exponentially small. As we vary p, the *free energy* of $\mathcal{C}$ remains small compared to the two sides of the cut due to a derease in energy (when p is close to 0) or due to entropy (when p is close to 0.5); all other cases rely on both.

An important technical contribution in our proofs is in the construction of a new injective map. The *contour representation* of a spin glass configuration consists of edges in the dual lattice that cross edges $e = (i, j)$ where $J_{ij}\sigma(i)\sigma(j) = -1$; in this representation the frustrated vertices in the dual lattice have odd degree and all other vertices have even degree. Because of this property the contour representation can be decomposed into a even cycles (closed contours) and two long paths whose endpoints are the four frustrated vertices. In the standard case of the Ising model with mixed side boundary conditions, we can define an injective map that shifts the paths connecting the four frustrated vertices to paths with much shorter length, and therefore much larger probability. The new paths can be added along the boundary by shifting closed contours. In our case we cannot do this since we cannot always construct maps to configurations with larger probability. Therefore we define a map to a *set* of configurations of *at least equal* probability. To complete the proof we require a careful map that allows us to reconstruct the original path, the new path, and the closed contours that are intersected when the new path is added. Verifying that the map is injective now requires a very sensitive combinatorial encoding and decoding that is likely of independent interest.

2 Preliminaries

We review some standard background on Markov chains, convergence times, and the Ising model that are required for our results.

2.1 Markov chains and mixing times

Let $\mathcal{M}$ be an ergodic, reversible Markov chain with arbitrary finite state space $\mathcal{S}$, transition probability matrix P, and stationary distribution π. Let $P^t(x, y)$ be the t-step transition probability from x to y, and let $||\cdot, \cdot||$ denote total variation distance.

▶ **Definition 2.1.** For $\varepsilon > 0$, the *mixing time* is defined as

$$\tau(\epsilon) = \min\{t : \max_{x \in \mathcal{S}} \sum_{y \in \mathcal{S}} ||P^{t'}(x, y), \pi(y)|| \leq \epsilon, \text{ for all t}' \geq \text{t}\}.$$

A Markov chain is *rapidly* (or *polynomially) mixing* if the mixing time is bounded above by a polynomial in $\log \mathcal{S}$, the length of a description of a state in $\mathcal{S}$. A chain is *slowly mixing* if

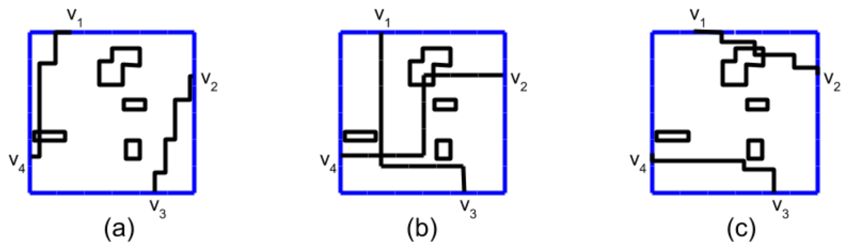


Figure 1 States with (a) positive orientation, (b) orientation 0, (c) negative orientation.

the mixing time is bounded below by an exponential function. The *conductance*, introduced by Jerrum and Sinclair [13], is useful to bound the mixing time [13].

▶ **Definition 2.2.** For a Markov chain with stationary distribution π, the *conductance* Φ is

$$\Phi = \min_{S: 0 < \pi(S) \leq 1/2} \frac{\sum_{x \in S, y \notin S} \pi(x) P(x,y)}{\pi(S)}.$$

▶ **Theorem 2.3.** (Jerrum and Sinclair [13]) *The mixing time of a Markov chain with conductance Φ satisfies:*

$$\tau(\epsilon) \;\geq\; \left(\frac{1-2\Phi}{2\Phi}\right) \ln \epsilon^{-1}.$$

To establish slow mixing, our strategy will be to define a set S along with sets $T \subset S^C$ and $\mathcal{C} \subset S^C \setminus T$ in the state space, such that $\pi(S) = \pi(T)$ and $\pi(\mathcal{C})/\pi(S) < e^{-cn}$ and such that getting from S to S^C in the Markov chain requires going through $\mathcal{C}$.

In this paper, we will focus on the simplest local Markov chain $\mathcal{M}$ for the Ising and spin glass models, known as *Glauber dynamics*, which connects pairs of configurations whose spins differ on at most one vertex. In a given step, the chain picks any vertex $v \in \Lambda$ at random and changes the spin with the appropriate transition probabilities so that the chain converges to the Gibbs distribution π. For our models, the transition probabilities of $\mathcal{M}$ are defined as

$$P(\sigma, \tau) \;=\; \frac{1}{2n^2} \;\min\left(1, \frac{\pi(\tau)}{\pi(\sigma)}\right),$$

if $|\{i : \sigma_i \neq \tau_i\}| = 1$, and with all remaining probability stay at the current configuration.

2.2 The Contour representation of the Ising and spin glass models

It will be convenient to view Ising and spin glass configurations in terms of *contours*. For every configuration $\sigma \in \Omega$, there is a contour representation $\Gamma(\sigma)$ in $\overline{\Lambda}$, the planar dual to Λ. We define $\overline{\Lambda} = (\overline{V}, \overline{E})$ by letting $\overline{V}$ correspond to the centers of unit squares in Λ and edges $\overline{E}$ connect any two vertices whose corresponding squares share an edge in Λ. An edge $e' \in \overline{E}$ that is dual to $e = (i,j) \in E$ is in $\Gamma(\sigma)$ if $J_{ij}\sigma(i)\sigma(j) = -1$ and we omit it if $J_{ij}\sigma(i)\sigma(j) = +1$. For the Ising model where all the $J_{ij} = +1$, the contour representation $\Gamma(\sigma)$ is precisely the set of edges separating $+$ and $-$ components in σ. Note that we can reconstruct the spin configuration σ from the contour representation (given a single spin) if we know the values of $\{J_{ij}\}$. The weight of a configuration σ is determined by $\Gamma(\sigma)$, and there is a weight-preserving bijection between the configurations of any two spin glasses with the same set of frustated vertices.

For the spin glass model considered here, all vertices of $\overline{V} \setminus \{v_1, ..., v_4\}$ have even degree in $\Gamma(\sigma)$ and the frustrated vertices $\{v_1, ..., v_4\}$ have odd degree. It follows that $\Gamma(\sigma)$ must be the union of two paths terminating at the frustated vertices, along with even cycles. (Note that these paths and cycles can intersect each other, and therefore are not necessarily unique.) In all that follows, it will be convenient to shift the primal lattice Λ by $(-1/2, -1/2)$ so that the vertices of $\overline{\Lambda}$ are integral. Now, recall that we assume that the four frustrated vertices lie on the boundary of a $2n \times 2n$ square S within $\overline{\Lambda}$ centered at (n, n), and they are symmetric under rotations by 90 degrees. Without loss of generality, we label these so that v_1 lies on the top side of S and is the i^{th} vertex from the upper left corner for some $0 \leq i \leq n$. Setting $p = i/2n$, v_1 is at a distance of $2pn$ from the upper left corner, v_2 is on the right side of S a distance of $2pn$ from the upper right corner, v_3 is on the bottom of S a distance of $2pn$ from the lower right corner, and v_4 is on the left side of S a distance of $2pn$ from the lower left corner. The key to all of our arguments is how the two long paths in $\Gamma(\sigma)$ pair up these frustrated vertices. Let $\alpha(\sigma)$ be the length of the shortest path in $\overline{\Lambda}$ from the connected component of $\Gamma(\sigma)$ containing v_1 to the connected component containing v_4 (if v_1 and v_4 are connected, $\alpha(\sigma) = 0$). Likewise, let $\beta(\sigma)$ be the length of the shortest path between the component containing v_1 and the component containing v_2. Let $\gamma(\sigma) = \beta(\sigma) - \alpha(\sigma)$ be the *orientation* of the configuration σ. We partition the state space Ω into a disjoint union $\Omega = \cup_{i \in \mathbb{Z}} \Omega_i$, where $\sigma \in \Omega_i$ if $\gamma(\sigma) = i$.

The partition of Ω into $\cup_i \Omega_i$ allows us to define a cut in the state space in order to bound the conductance. In particular, we let $\Omega^- = \cup_{i<0} \Omega_i$ and $\Omega^+ = \cup_{i>0} \Omega_i$, and we observe that $\Omega = \Omega^- \cup \Omega_0 \cup \Omega_+$. We specify a subset of $\mathcal{C} \subset \Omega_0$ that will be critical to defining the cut as $\mathcal{C} = \{\sigma \in \Omega_0 : \alpha(\sigma) = \beta(\sigma) = 0\}$ (i.e., the configurations in which v_1 is connected to both v_2 and v_4). See Figure 1. Finally, we define $\mathcal{C}^* = \mathcal{C} \cup \Omega_{-1} \cup \Omega_1$ to be the configurations where the paths connecting the frustrated vertices are within distance 1 of each other. Following [25], for configurations in $\mathcal{C}$, we partition the cross into two paths, one from v_1 to v_3 and a one from v_2 to v_4; we do the same for configurations in Ω_{-1} and Ω_1, although it may be necessary to add a single "defect" that encodes where one or both of these paths incurs a jump by one unit. To move from a configuration in Ω^- to one in Ω^+ using Glauber dynamics, we must pass through a configuration in $\mathcal{C}^*$. We will show that the probability of $\mathcal{C}^*$ is exponentially small, and this will allow us to argue that the Glauber dynamics requires exponential time to converge to equilibrium.

3 Slow Mixing for the Ising model with Mixed Boundaries

We start with the standard approach used to show slow mixing when the boundary conditions alternate spins on the boundary of a $(2n+1) \times (2n+1)$ lattice region Λ. Here $\overline{\Lambda}$ is the $2n \times 2n$ lattice region centered in Λ. This will motivate the approach used in the general spin glass setting (when the frustrated vertices are not necessarily on the boundary of Λ) and will elucidate the difficulties in generalizing this simpler result.

Fix $0 \leq p \leq 1/2$ and let $q = 1 - p$. We define $v_1 = (2pn, 2n), v_2 = (2n, 2qn), v_3 = (2qn, 0)$ and $v_4 = (0, 2pn)$. Recall that all vertices on the boundary between v_1 and v_2 and between v_3 and v_4 are assigned $+$ and the others are assigned $-$. The vertices $v_1, ..., v_4$ define the endpoints of a pair of paths in each configuration. (There may be more than one choice of paths.) Using the strategy outlined in Section 2.2, we recall that $\mathcal{C}$ consists of those configurations where there are paths from v_1 to both v_2 and v_4 (and therefore also to v_3). Using the notion of "fault lines" introduced in [25], we note that this is the set of configurations that contain a *horizontal fault line*, i.e.,. a path from v_2 to v_4, and a *vertical fault line*, i.e., a path from v_1 to v_3. When both fault lines are present (and intersect) we call their union a *cross*. We define the cross so that it is a maximal component of the contour representation of the configuration.

Let C be a cross in $\overline{\Lambda}$. As we will show in Lemma 4.1, the minimum length of C is $L = 6n - 4np$. We write the length as $|C| = L + \ell$, for some $\ell \geq 0$. Let $\mathcal{C}_C$ be the set of configurations in $\mathcal{C}$ that have C as their cross.

We will write the weight of a configuration σ as $\lambda^{-H(\sigma)}$, $\lambda = e^{2\beta} = e^{2/T}$, and note that the energy $H(\sigma)$ is the number of edges in the contour representation of σ.

▶ **Lemma 3.1.** *For any cross C, we have*

$$\pi(\mathcal{C}_C) \leq \lambda^{-(2n-4pn+\ell)}.$$

Proof. We define the injective map $\psi_C : \mathcal{C}_C \to \Omega$ so that $\pi(\psi_C(\sigma)) = \pi(\sigma)\lambda^{(L-4n+\ell)}$ for any fixed C. Given this map, we find

$$1 \;=\; \pi(\Omega) \geq \sum_{\sigma \in \mathcal{C}_C} \pi(\psi_C(\sigma)) \;=\; \sum_{\sigma \in \mathcal{C}_C} \pi(\sigma)\lambda^{(L-4n+\ell)} \;=\; \lambda^{(2n-4pn+\ell)}\pi(\mathcal{C}_C).$$

The map ψ_C is defined by removing C; then, along the upper-left boundary of Λ between v_1 and v_4 we add each edge not in σ and remove each edge in σ; then, along the lower-right boundary of Λ between v_3 and v_2 we add each edge not in σ and remove each edge in σ. ◀

▶ **Theorem 3.2.** *Let $\Lambda \subset \mathbb{Z}^2$ be an $(2n+1) \times (2n+1)$ lattice region and $0 \leq p \leq 1/2$ define a family of balanced mixed boundary conditions on Λ. Let Ω be the set of all Ising configurations and let $\mathcal{C}$ be the Ising configurations containing a cross. Then*

$$\pi(\mathcal{C}) \leq f(n)e^{-cn},$$

for some polynomial $f(n)$ and constant $c > 0$, whenever $\lambda^{(1-2p)} > 3^{(3-2p)}$.

Proof. By Lemma 3.1,

$$\pi(\mathcal{C}) \;\leq\; \sum_{C} \lambda^{-(2n-4pn+\ell)} \;\leq\; \sum_{\ell \geq 0} \lambda^{-(2n-4np+\ell)} 3^{(6n-4np+\ell)} \leq\; 4n^2(3^{(3-2p)}\lambda^{-(1-2p)})^{2n},$$

which is exponentially small when $\lambda^{(1-2p)} > 3^{(3-2p)}$. The second inequality holds because there are at most $3^{(6n-4np+\ell)}$ ways to choose a cross of length $6n - 4np + \ell$. ◀

Thus, when $\lambda^{(1-2p)} > 3^{(3-2p)}$ we have that the size of the cut is exponentially small, and therefore the conductance of the graph is also exponentially small. By Theorem 2.3, this implies that the chain takes exponential time to mix.

▶ **Corollary 3.3.** *Glauber dynamics for the Ising model on Λ with balanced mixed boundary conditions takes time at least e^{cn} to mix, for some constant $c > 0$, when $\lambda^{(1-2p)} > 3^{(3-2p)}$.*

Notice that this gives $\lambda > 27$ when $p = 0$ and $\lambda > 3^{(2^{(k+1)}+1)}$ when $p = 1/2 - 1/2^k$ and when $p = 1/2$ this fails to give any useful bound.

4 Slow Mixing for Frustrated Spin Glasses Using Free Energy

We will now proceed to extend the result in Section 3 by establishing slow mixing below some temperature for spin glasses with four well-separated frustrated vertices.

In this setting we define Λ as a $kn \times kn$ lattice region, $k \geq 2$. Four distinguished faces are symmetric around the center of the lattice region under 90 degree rotations. The centers of these faces are four vertices $v_1, .., v_4$ in $\overline{\Lambda}$. As in Section 2.2 we define $\mathcal{C}$ to be the set of

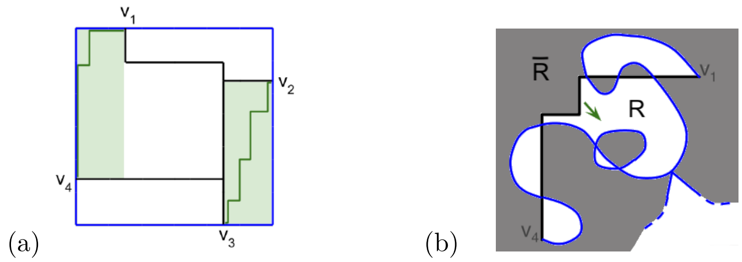


Figure 2 (a) A minimal cross is shown in black, with two possible monotone paths in green. Any monotone path in either shaded region is possible. (b) A staircase is shown in black, together with the part of a cross containing a path from v_1 to v_4. The green arrow shows the direction edges of σ are shifted in the region bounded by the middle section of the staircase and the cross.

contour configurations in which v_1 is connected to both v_2 and v_4, and we define the cross C in such a configuration as the component containing v_1. The argument in Section 3 fails when $p = 1/2$, in particular when $\ell = o(n)$. The length of the cross C in that case is $4n + \ell$, and our injective map ψ_C removes C and replaces it with two paths of total length $4n$. The difference in energy, $H(\sigma) - H(\psi_C(\sigma)) = \ell$, is too small to show that σ has exponentially small probability.

The remedy comes from noticing that in exactly the case $\ell = o(n)$, C is nearly a minimal cross and there are many alternative choices of ψ_C. We will allow any monotone path that, in order to ensure loss of energy, does not intersect C. The set of possible paths is illustrated in Figure 2(a). We have the following lemma, whose proof appears in the Appendix.

▶ **Lemma 4.1.** *Let S_n be the $2n \times 2n$ axis-aligned square whose sides contain $v_1, .., v_4$. For some $\ell \geq 0$, $|C| = 6n - 4pn + \ell$. If $\ell < 2pn$ there are two $(2n - 2pn - \ell) \times (2pn - \ell)$ rectangular regions on opposite corners of the interior of S_n that contain no edges of C.*

Our new strategy is to use *all* possible choices of ψ_C, thereby defining an exponential family of images. We will define a function Ψ_C that involves mapping a configuration $\sigma \in \mathcal{C}_C$ to the union of possible $\psi_C(\sigma)$ defined by different pairs of monotone paths. Figure 2(a) also shows the tradeoff between energy and entropy for our method. As p decreases, the energy loss due to the map increases. As the width of each shaded area decreases, the number of possible paths, $\binom{2n}{2np}$, also decreases. This is what we mean by a decrease in entropy.

Just as we needed ψ_C to be injective in Section 3, we would like our new map to have the property that two different configurations map to disjoint sets of configurations. Instead, we define Ψ_C to pass a small amount of "side information," and with this definition we will get a disjointness property that serves our purpose. The side information is in the form of tokens placed on certain edges along each of the two paths that define the configuration σ is mapped to. Formally, for each path this information is encoded as a binary string of length $2n$: `0` for any plain edge, `1` for an edge with a token. The nice property that will make this side information small is that no two adjacent edges of a path are occupied by tokens.

Let $B(m)$ be the set of binary strings of length m with no consecutive `1`'s. Let $B = B_C = B(2n - \ell)$. Formally, we will define a function $\Psi_C : \mathcal{C}_C \to 2^{\Omega \times B \times B}$ that has the nice properties in the following lemma. To get our hands on the set of mapped configurations minus the tokens, we define the projection operator $\Pi : 2^{\Omega \times B \times B} \to 2^{\Omega}$, so that $\Pi(\{\sigma_i, b_i, b'_i\}) = \{\sigma_i\}$. Formally, $\Pi \circ \Psi_C$ is the map from one configuration to a set of configurations.

In the following lemmas, fix $0 \leq p \leq 1/2$ and let $L = 6n - 4np$.

▶ **Lemma 4.2.** *Let C be a maximal cross of length $|C| = L + \ell$. There exists a function $\Psi_C : \mathcal{C}_C \to 2^{\Omega\times B\times B}$ such that $\forall \sigma, \sigma' \in \mathcal{C}_C$, $\sigma'' \in \Pi \circ \Psi_C(\sigma)$,*

$$\begin{aligned} \Psi_C(\sigma) \cap \Psi_C(\sigma') \ &= \ \emptyset, \\ |\Psi_C(\sigma)| \ &= \ \binom{2n-2\ell}{2pn-\ell}^2, \\ \text{and} \quad H(\sigma'') \ &\leq \ H(\sigma) - (2n - 4np + \ell). \end{aligned}$$

We postpone constructing the function Ψ_C (and proving Lemma 4.2) until the next subsection. Theorem 4.5 is an analogue of Theorem 3.2 that gives an exponential bound for all p, $0 \leq p \leq 1/2$. As a corollary of Theorem 4.5, we will prove our main result, Theorem 1.1, asserting slow mixing for spin glasses with frustration.

We first bound the probability of the set of configurations containing a given cross C.

▶ **Lemma 4.3.** *For any maximal cross C of length $|C| = L + \ell$ we have*

$$\pi(\mathcal{C}_C) \leq \pi(\Pi \circ \Psi_C(\mathcal{C}_C))\lambda^{-(2n-4np+\ell)}\phi^{4n-2\ell+1} \Big/ \binom{2n-2\ell}{2np-\ell}^2, \tag{1}$$

where $\phi = (1+\sqrt{5})/2$.

Proof. It is well known that $|B(m)|$ is the m^{th} Fibonacci number, which is within 1 of ϕ^m. Each $\sigma'' \in \Pi \circ \Psi_C(\sigma)$ appears in at most $|B|^2 \leq \phi^{4n-2\ell+1}$ elements of $\Psi_C(\sigma)$. The bound on $H(\sigma'')$ in Lemma 4.2, gives $\pi(\sigma'') \geq \pi(\sigma)\lambda^{-(2n-4np+\ell)}$ and the two equalities imply

$$\pi(\Pi \circ \Psi_C(\mathcal{C}_C)) \geq \sum_{\sigma \in \mathcal{C}_C} \pi(\sigma)\lambda^{(2n-4np+\ell)}\phi^{-(4n-2\ell+1)}\binom{2n-2\ell}{2np-\ell}^2. \tag{2}$$

The inequality follows by replacing $\sum \pi(\sigma)$ with $\pi(\mathcal{C}_C)$. ◀

Our main theorems establishing slow mixing of Glauber dynamics for spin glasses with well-separated frustrated vertices (Theorems 4.5 and 1.1) depend on the following technical lemma regarding the set $\mathcal{C}_\ell$ of configurations containing maximal crosses of fixed length $L+\ell$: $\mathcal{C}_\ell = \cup\{\mathcal{C}_C : |C| = L + \ell\}$. The idea of the lemma is to show that $\pi(\mathcal{C}_\ell)$ is exponentially small, where the constant in the exponent is independent of ℓ. This also means that the free energy $\ln \pi(\mathcal{C}_\ell)/n$ is less than some negative constant. Since there are polynomially many values of ℓ, it will follow that the whole set $\mathcal{C}$ is exponentially small.

▶ **Lemma 4.4.** *Let $\mathcal{C}_\ell$ be the spin glass configurations where $v_1, .., v_4$ are all connected by a maximal cross of length $L + \ell$. Then for $\lambda \geq 256$ we have*

$$\pi(\mathcal{C}_\ell) \leq 2^{-0.2n} \text{ poly}(n). \tag{3}$$

Proof. Let $s = 1/2 - p$ and $r = \ell/n$. We will actually prove that

$$\pi(\mathcal{C}_\ell) \leq \lambda^{-8sn} \ (3/\lambda)^{rn} \ 2^{n[(4-2r)\log_2\phi + \mathcal{L}(r,s) + \mathcal{P}(r,s) - \mathcal{T}(r,s)]} \text{ poly}(n) \ , \tag{4}$$

where

$$\mathcal{L}(r,s) = (2+4s+r)h(\frac{r}{2+4s+r}),$$

$$\mathcal{P}(r,s) = (2+4s)h(\frac{2s}{1+2s}),$$

$$\mathcal{T}(r,s) = \max(0, 4-4r)h(\frac{1}{2} - \frac{s}{1-r}),$$

and $h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$. Then we will show that the right-hand side of Equation 4 is less than $2^{-0.2n}$.

First, we establish Equation 4. Each C consists of vertical path connecting v_1 to v_3 and a horizontal path connecting v_2 to v_4. The vertical path contains a minimal vertical path of $2n$ vertical edges and $2n-4pn$ horizontal edges. There are $\binom{4n-4pn}{2n-4pn} = \binom{2n+4sn}{4sn}$ choices of minimal vertical path. There is one choice of minimal horizontal path, which contains only horizontal edges connecting v_2 and v_4 to the vertical path. Then there are $\binom{6n-4np+\ell}{\ell} = \binom{2n+4sn+rn}{rn}$ ways to choose the locations of the ℓ extra edges, and 3 possible directions for each extra edge. Applying Lemma 4.3 and Stirling's formula,

$$\begin{aligned}
\pi(\mathcal{C}_\ell) &\le \binom{6n-4np+\ell}{\ell}\binom{4n-4pn}{2n-4pn}3^\ell \max_{|C|=L+\ell} \pi(\mathcal{C}_C) \\
&\le 2^{(2n+4sn+rn)h(r/(2+4s+r))}2^{(2n+4sn)h(2s/(1+2s))}3^{rn} \\
&\qquad \cdot\ \lambda^{-(8sn+rn)}\phi^{4n-2rn+1}2^{-2(2n-2rn)h((1-r-2s)/(2-2r))}.
\end{aligned}$$

Equation 4 follows immediately by collecting the terms in the exponents.

By taking logs and dividing by n it follows that $\log_2 \pi(\mathcal{C}_\ell)/n \le \mathcal{F}(r,s)$, where

$$\mathcal{F}(r,s) = (-r-8s)\log_2\lambda + r\log_2 3 + (4-2r)\log_2\phi + \mathcal{L}(r,s) + \mathcal{P}(r,s) - \mathcal{T}(r,s)$$

It remains to show that $\mathcal{F}(r,s) \le -0.2$, for all s, r, $0 \le s \le 1/2$, $r > 0$, and large enough λ.

$\mathcal{L}(r,0)$ is concave as a function of r, $\mathcal{L}(r,s)$ and $\mathcal{P}(r,s)$ are concave as functions of s, and $-\mathcal{T}(r,s)$ is convex as a function of s. We numerically approximate the concave functions with a tangent line and the convex function with a secant, yielding these results:

$$\begin{aligned}
&\mathcal{L}(r,0) \le 0.5 + 2.9r; && \mathcal{P}(r,s) \le 0.5 + 12s; \\
&\mathcal{L}(r,s) \le 0.5 + 2.9r + 2rs \ \le\ 0.5 + 3.9r; && -\mathcal{T}(r,s) \le -4 + 4r + 8s.
\end{aligned}$$

Also, $r\log_2 3 < 1.5r$ and $(4-2r)\log_2\phi < (2.8-1.4)r$. Adding terms, for $\lambda \ge 256$, we get

$$\mathcal{F}(r,s) \ \le\ (-r-8s)\log_2\lambda + 8r + 20s - 0.2 \ \le\ -0.2.$$

◀

We now state the key theorem bounding the probability of the set $\mathcal{C}$ of configurations containing crosses.

▶ **Theorem 4.5.** *Let Ω be the set of all spin glass configurations in a $kn \times kn$ square lattice Λ centered at (n,n), $k \ge 2$. Suppose that four distinguished vertices $v_1,..,v_4$ lie on the boundary of an axis-aligned $2n \times 2n$ square S centered in $\overline{\Lambda}$, and these four vertices form the corners of a (not necessarily axis-aligned) square (i.e., they are shifted by $2p$ around the boundary of S). Let $\mathcal{C}$ be the set of configurations in which v_1 is connected to both v_2 and v_4. Then for $\lambda \ge 256$ we have*

$$\pi(\mathcal{C}) \le 2^{-0.2n}\mathrm{poly}(n). \tag{5}$$

Proof. Since ℓ has at most $(cn)^2$ values, $\pi(\mathcal{C}) \le (cn)^2 \max_\ell \pi(\mathcal{C}_\ell) \le 2^{-0.2n}\mathrm{poly}(n)$. ◀

Proof of Theorem 1.1. Set $T = 2/\ln 256 = 0.360...$. Let $t < T$. The state space Ω contains the two disjoint subsets Ω_- and Ω_+, separated by a cut set $\mathcal{C}^*$ consisting of all configurations

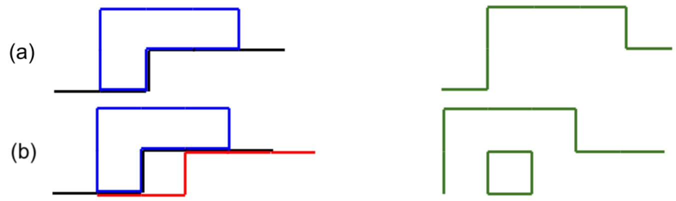


Figure 3 (a) A staircase and patch that share edges (left), and an encoding that loses information (right). (b) A staircase and patch with the default path (left), and an encoding that preserves information (right).

within two steps of $\mathcal{C}$. We have $\pi(\mathcal{C}^*) < \pi(\mathcal{C})\mathrm{poly}(n)$ and by symmetry $\pi(\Omega_-) = \pi(\Omega_+)$. The conductance Φ satisfies

$$\Phi \;\leq\; \frac{\sum_{\sigma\in\Omega_-,\sigma'\in\Omega_{\mathcal{C}}} \pi(\sigma)\Pr(\sigma,\sigma')}{\pi(\Omega_-)} \;\leq\; 4\cdot\pi(\mathcal{C}^*) \;\leq\; 2^{-0.1n}, \quad \text{for large enough } n. \tag{6}$$

Therefore the Markov chain mixes slowly. ◀

4.1 Construction of the Map

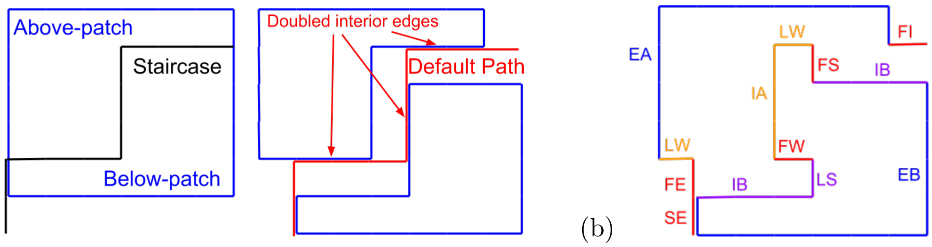
In this section we will construct the map Ψ_C using pairs of paths as shown in Figure 2(a). An *upper staircase* with respect to a cross C of length $L+\ell$ is a path of $\min(\ell, 2pn)$ west edges starting at v_1 followed by zero or more west and south edges, followed by $\min(\ell, 2pn)$ south edges ending at v_4. We refer to the section of west and south edges as the "middle $2n - 2\min(\ell, 2pn)$ edges." We define a *lower staircase* to be a path v_3 to v_2, which, when the configuration is rotated 180°, becomes an upper staircase. Note that the edges on a staircase need not be edges of a particular configuration. Given upper and lower staircases, we will map $\sigma \in \mathcal{C}_C$ to some $\sigma' \in \Omega$, marking certain edges with tokens. We will show that one can reconstruct σ from C, σ', and the marked edges, that no two marked edges are adjacent, and $H(\sigma') \leq H(\sigma) - |C| + 4n$, implying Lemma 4.2.

Our map is motivated by the map ψ_C in the proof of Lemma 3.1. In fact, the construction is the same along the first $\min(\ell, 2pn)$ edges and last $\min(\ell, 2pn)$ edges: we add each edge not in σ and remove each edge in σ. Along the middle section of the staircase that contains west and south edges, our map must encode the locations of the staircase edges in σ' without increasing $H(\sigma')$. The basic strategy is to remove C, shift edges in σ away from the staircase, toward the removed edges of C, then add the edges of the staircase.

Let S_U be an upper staircase and S_L be a lower staircase. The simple regions in the interior of $C \cup S_U \cup S_L$ may be two-colored gray and white, with the exterior, denoted $\overline{R}$, colored gray. Regions separated by an edge in $C \cap S_U$ or $C \cap S_L$ will have the same color. We assume in what follows that $\ell < 2pn$. In particular, S_U and S_L do not both contain edges in any one region boundary. When $\ell \geq 2pn$, S_U and S_L are contained in the boundary of the $2n \times 2n$ square S_n, and the proof of Lemma 3.1 applies.

By Lemma 4.1 there is one white simple region R whose boundary contains the middle $2n - 2\ell$ edges of S_U. The map will shift edges of σ in R southeast, and it will shift the corresponding region bounded by the middle $2n - 2\ell$ edges of S_L northwest. See Figure 2(b).

We may assign a + or − to each site in $R \cup \overline{R}$ so that the sites adjacent to C are + and the edges of σ restricted to $R \cup \overline{R}$ are exactly those edges between two neighboring sites of opposite sign. We define a *patch* to be a connected set of − sites in $R \cup \overline{R}$. The outer



(a) (b)

Figure 4 (a) The components to encode. (b) The contour pieces defining the map.

boundary of a patch is the unique cycle of edges in the configuration that, when traversed counterclockwise, has sites inside to the left of each edge and sites outside to the right.

A naive map would remove C from the configuration and add the upper staircase and lower staircase to the configuration. The flaw in this approach is that σ cannot always be reconstructed when part of a staircase coincides with part of the boundary of a patch. Figure 3(a) shows an upper staircase in black that shares edges with a patch, shown in blue. Adding the staircase creates double edges. The natural recourse is removing double edges while preserving degrees, but shared edges are no longer recoverable from such a map.

Our map modifies the naive approach by shifting the staircases before adding them to the configuration, and shifting edges that are between the staircases toward the empty space left behind after the removal of C. Let S be the maximal contiguous section of S_U that forms part of the boundary of R and contains the middle $2n - 2\ell$ edges of S_U. We define the *default path* to be S shifted one step east. It consists of alternating west and south sections. The *first south* (northernmost) edge of each south section, and the *first west* edge of each west section, are each incident to S at just one vertex (with the exception of the first edge of S if it is a south edge preceded in S_U by a south edge). All other edges on the default path are on S or not incident to it. The *last south* and *last west* edges are defined accordingly.

Figure 3(b) shows the same staircase and patch, with the default path in red. σ is mapped to σ' by starting with the union of the patch and the default path, and removing double edges. The default path can be reconstructed from σ', because it contains the first-south and first-west edges of the default path. This is the information that was missing from the previous mapping. The mapping contains no more energy than the original.

A subtler problem of lost information arises when the staircase enters the interior of a patch. We define an *interior edge* of S to be one that bounds two $-$ sites. Each maximal contiguous segment of interior edges of S divides a patch into two patches, which we refer to as the *above-patch* (or *A-patch*) and the *below-patch* (or *B-patch*).

To solve the problem of interior segments, we triple each interior edge of the staircase, shifting the staircase and the B-patch one step east, and shifting the B-patch one step south. The drawing on the left of Figure 4(a) shows the staircase in black and the patch in blue before the two shifting steps, and the drawing on the right shows the default path in red and the two patches after the shifts. After the shifts, our mapping removes all double edges.

The doubled interior edges of the default path consist of all interior west edges of the A-patch except the *last-west edge of each west section*, and all interior south edges of the below patch except the *last-south edge of each south section*.

This mapping has the one final problem that it increases the energy of the configuration. This problem can be illustrated by labeling the edges as in Figure 4(b). Edges EA and EB (blue) are exterior of the A-patch and B-patch, respectively. Edges IA and LW (orange) are south and last-west interior edges of the A-patch, resp. Edges IB and LS (purple) are west and last-south interior edges of the B-patch, resp. Edges *FI*, *FW*, and *FS* (red) are the first interior edge and all first-west and first-south edges of the default path, resp. Edges *FE* and *SE* (red) are the first and second "exterior" edges of the default path following this segment of interior edges. The first exterior edge will not be interior to any patch, but the second exterior edge may be interior to this or another patch.

The increase in energy is caused by the "detours" at *FS-LW* and *FW-LS*. The final mapping step is to flip the signs of sites bounded by corners of those two types and to place a token at each such site. The Appendix presents the map steps in detail.

4.2 Reconstruction

The default path (and hence σ restricted to R_L) can be reconstructed from σ', before token-placing, as it contains all of the first-west and first-south edges. Starting from the FI edge, the default path continues until it encounters the first FS or FW edge. Then it changes direction and the FS or FW edge inductively plays the role of the FI edge. The rest of the interior segment is reconstructed by induction on the number of south and west sections.

Reconstructing the default path in the presence of tokens is the same recursive process, except we look ahead one step. If the next edge has a token, we flip the adjacent site before proceeding. The adjacent site is unambiguous because it is between the A- and B-patches. The Appendix presents the reconstruction steps in detail.

4.3 Energy loss:

Before token-placing and sign-flipping, σ' has more energy than $H(\sigma) - |C| + 4n$. The EA and EB naturally correspond 1-1 to the edges of the original patch. The IA and IB edges correspond 1-1 to the interior segment of the staircase. The excess energy consists of one pair of edges, FS-LW or FW-LS, for each corner of the interior segment, plus two more edges, the FI edge and one LW or LS edge incident to FE.

The mapping solves this problem by short-circuiting the corners. Each FS-LW pair occurs as part of a segment FS-LW-IA that form three sides of a site, and each FW-LS pair occurs as part of a segment FW-LS-IB that also form three sides. The mapping flips the sign of each such site, replacing three edges with one, and places a token at the flipped site.

Two such sites may be adjacent. This happens when an IA or IB section is one edge long. Then one of the two sites is bounded by an FS-LW-IA-FW segment or an FW-LS-IB-FS segment. In either case the mapping replaces four edges with zero. One sign-flip in the first traversal removes the excess energy of both sites, and one token is placed. It also flips one edge of the adjacent site, ensuring that no two tokens will be adjacent. (See Figure 5(a) steps (d)-(f).) Each sign-flip in the second traversal converts three edges to one, canceling excess energy due to this site. In this case, this site will not be adjacent to another token site.

The two remaining excess edges are the FI edge and one LW or LS edge. Suppose it is LW (the case of LS is similar). If FE and a LW form a double edge or SE and an EB form a double edge (the case pictured), the mapping removes the double edge, cancelling the excess energy. In the remaining case FE is an FS edge, SE is an FW edge, and the segment LW-FE-SE forms three sides of a site. The mapping flips the sign of this site and places a token. No token is placed on a site adjacent to this site. SE is not an interior edge of any patch, because this site is on the exterior side of FE. The first interior edge of a patch does not bound a site with a token.

References

1. M. Aizenman. Translation invariance and instability of phase coexistence in the two-dimensional Ising system. *Comm. Math. Phys.*, **73:** 83-94, 1980.
2. F. Barahona. On the computational complexity of Ising spin glass models. *Journal of Physics A: Math. and Gen.*, **15:** 3241–3253, 1982.
3. D. Chelkak and S. Smirnov. Universality in the 2D Ising model and conformal invariance of fermionic observables *Inventiones Mathematicae*, **189:** 1–66, 2009.
4. M. Ciucu. Dimer packings with gaps and electrostatics, *Proc. Natl. Acad. Sci.*, **105:** 2766–2772, 2008.
5. L. Coquille and Y. Velenik. A finite-volume version of Aizenman-Higuchi Theorem for the 2D Ising model. *Probab. Theory Relat. Fields,* **153:** 25–44, 2012.
6. R.L. Dobrushin, R. Kotecký and S. Shlosman. Wulff Construction: a Global Shape from Local Interaction. *AMS translations series,* **104**, Providence R.I.: AMS, 1992.
7. R. Fernandez, P.A. Ferrari, and N.L. Garcia. Loss network representation of Ising contours. *Annals of Probability* **29:** 902–937, 2001.
8. R. Fitzpatrick. Thermodynamics and Statistical Mechanics. Preprint at https://farside.ph.utexas.edu/teaching/sm1/Thermal.pdf.
9. S. Friedli and Y. Velenik. *Statistical Mechanics of Lattice Systems: a Concrete Mathematical Introduction.* Cambridge University Press, 2017, to appear (http://www.unige.ch/math/folks/velenik/smbook/).
10. H.O. Georgii. *Gibbs measures and phase transitions.* de Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin, 1988.
11. S. Greenberg and D. Randall. Slow mixing of Glauber dynamics on perfect matchings of the square-octagon lattice. Preprint, 2006.
12. Y. Higuchi. On the absence of non-translation invariant Gibbs states for the two-dimensional Ising model. *Colloq. Math. Soc.*, J'anos Bolyai, 517–534, 1981.
13. M.R. Jerrum and A.J. Sinclair. Approximate counting, uniform generation and rapidly mixing Markov chains. *Information and Computation* **82**: 93–133, 1989.
14. G. Lawler. Scaling limits and the Schramm-Loewner evolution. *Probab. Surveys,* **8:** 442–495, 2011.
15. E. Lubetzky, F. Martinelli, A. Sly and F. Toninelli. Quasi-polynomial mixing of the 2D stochastic Ising model with "plus" boundary up to criticality. *J. Eur. Math. Soc.*, **15**: 339–386, 2013.
16. F. Martinelli. On the two-dimensional dynamical Ising model in the phase coexistence region. *Journal of Statistical Physics*, **76:** 1179–1246, 1994.
17. F. Martinelli. Lectures on Glauber dynamics for discrete spin models. *Lectures on Probability Theory and Statistics (Saint-Flour, 1997)*, Lecture notes in Mathematics **1717**: 93–191, Springer, Berlin, 1998.
18. F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case. *Comm. Math. Phys.*, **161:** 447–486, 1994.
19. F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case. *Comm. Math. Phys.*, **161:** 487–514, 1994.
20. F. Martinelli, A. Sinclair, D. Weitz, The Ising model on trees: Boundary conditions and mixing time. *Comm. Mathematical Physics* **250**: 301–334, 2004.
21. A. Messager and S. Miracle-Sole. Equilibrium states of the two-dimensional Ising model in the two-phase region. *Comm. Math. Phys.*, **40:**187–196, 1975.
22. E. Mossel and A. Sly. Exact thresholds for Ising – Gibbs samplers on general graphs. *Ann. Probab.*, **41**:294-328, 2013.
23. L. Onsager. Crystal statistics: a two-dimensional model with an order- disorder transition. *Phys. Rev.*, **65:** 117–149, 1944.

24 R. Peierls, On Ising's model of ferromagnetism. *Proc. Camb. Philos. Soc.*, **32**: 477–481, 1936.

25 D. Randall. Slow mixing of Glauber dynamics via topological obstructions. *17th Symposium on Discrete Algorihtms (SODA)*, 2006.

26 D. Randall and D.B. Wilson. Sampling Spin Configurations of an Ising System. *Proc. 10th ACM/SIAM Symposium on Discrete Algorithms*, S959–960, 1999.

27 O. Schramm. Scaling limits of loop-erased random walks and uniform spanning trees. *Israel J. Math.*, **118**: 221–288, 2000.

28 S. Smirnov. Critical percolation in the plane. *Comptes Rendus de l'Académie des Sciences*. **333**: 239–244, 2001.

29 S. Smirnov and H. Duminil-Copin. Conformal invariance in lattice models, in Lecture notes of the 2010 Clay summer school, Buzios, 2010.

30 N. Sun. Conformally invariant scaling limits in planar critical percolation. *Probability Surveys*, **11:** 155–209, 2011.

31 L. Thomas. Bound on the mass gap for the finite volume stochastic Ising models at low temperature. *Comm. Mathematical Physics* **126:** 1–11, 1989.

Appendix

Proof of Lemma 4.1. The minimal cross contains a path from v_1 to v_3 and a path from v_2 to v_4. First let's assume that each of these is minimal. Then they each have length $4n - 4pn$ and the total length of the cross is $8n - 8pn - |o|$, where o is the length of the overlapping segments. Orient the edges along each path from v_1 to v_3 so that the edges all go right or down, and orient the path from v_2 to v_4 so that they go down or left. Then the overlapping segments are oriented the same way in both paths if the edge is vertical and in opposite directions if the edge is horizontal. But all horiztonal edges on the path from v_2 to v_4 after this shared edge are left of the edge, and those on the path from v_1 to v_3 are to the right; similarly, if they share a horizonal edge, all subsequent vertical edges must be to the left of the edge on one path and to the right on the other. Therefore, the overlapping segment must all be vertical or all horizontal. Furthermore, all the vertical edges that overlap have to lie between v_2 and v_4 and have length at most $2n - 4pn$; likewise if the horizontal edges that overlap since they lie between v_1 and v_3. It follows that when the two paths are minimal $|o| \leq 2n - 4pn$ and the length of the cross is at least $6n - 4pn$.

If either of the paths from v_1 to v_3 and v_2 to v_4 is not minimal, then the overlap can contain both horizontal and vertical edges. Notice that the overlapping segments must be contiguous along either path or the cross would contain a cycle, contradicting minimality. If this overlapping segment contains edges oriented both left and right (or down and up), then it can be shortened, again violating minimality. Therefore the overlapping segment must go down and left or down and right. If down and left, then the path from v_1 to v_3 has an extra edge to the right for each horizontal edge in the overlapping segment; if down and right then the path from v_2 to v_4 has an extra edge for each horiztonal edge in the overlap. Finally, if the number of vertical edges in the overlap exceeds the vertical distance between v_2 and v_4, then the path between them must contain at least that many additional vertical edges. Summing all of these up, we find that if there are $2n - 4pn + k$ edges in the overlap, then the sum of the lengths of the two paths must be at least $8n - 8pn + k$. Subtracting the length of the overlapping segment, we again find that the length of the cross is at least $6n - 4pn$.

If the cross is nearly minimal, with length $6n - 4pn + \ell$, the picture is similar. The paths from v_1 to v_3 and v_2 to v_4 must also be nearly minimal, each having length at most $4n - 4pn + \ell$ and the length of the overlapping segments must be at least $2n - 4pn - \ell$. It

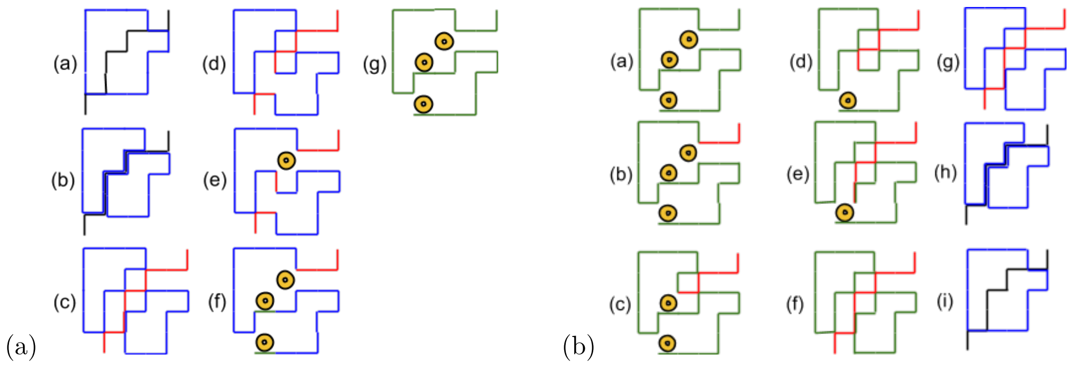


Figure 5 (a) The map: blue edges are the patch boundary, black edges are the staircase, red edges are the default path, and green edges are the final mapping. (b) Reconstruction steps: blue edges are the patch boundary, green edges are the mapping, black edges are the staircase, and red edges are the default path.

follows that the path from v_1 to v_3 lies in a $(2n - 4pn + \ell) \times 2n$ rectangle, the path from v_2 to v_4 lies in a $2n \times (2n - 4pn + \ell)$ rectangle, and the overlapping segments lie in the center $(2n - 4pn + \ell) \times (2n - 4pn + \ell)$ square. The overlapping segments do not have to be contiguous, but the distance between segments is at most ℓ. We find, by a similar argument to before, that all but ℓ edges on the overlap must have the same orientation, horiztonal or vertical. If the overlap is mostly vertical, then the $(2pn - \ell) \times (2n - 2pn - \ell)$ rectangles adjacent to the upper-left and lower-right corners of the region cannot contain any edges from the cross. Similiarly, if the overlapping segments are mostly horizontal, then there cannot be any edges from the cross in the $(2n - 2pn - \ell) \times (2pn - \ell)$ rectangles incident to the upper-right and bottom-left corners of the region. ◀

Map Steps:

Given $\sigma \in \mathcal{C}_C$ pick an upper staircase S_U and a lower staircase S_L. Remove C from σ. Along the initial segment of ℓ edges and final segment of ℓ edges of S_U, add each edge not in σ and remove each edge in σ. Let S be the middle $2n - 2\ell$ edges of S_U.

1. Add S. If this doubles an edge, label one copy *on the staircase* and the other *above* (*below*) the staircase if it is on the boundary of an A-patch (B-patch).
2. Triple each interior edge of S. Label one copy *on the staircase*, the second *above the staircase*, and the third *below the staircase*. (Figure 5(a) step (b).)
3. Shift every edge on or below the staircase one step east.
4. Shift every edge below the staircase one step south. (Figure 5(a) step (c).)
5. Remove every double edge. (After the two shifts there are no triple edges.) (Figure 5(a) step (d).)
6. Traverse the default path twice from start to end (Figure 5(a) steps (e), (f)):
   a. First traversal: if the current edge and the next edge are interior FW or FS edges, then put a token on the site bounded by these two edges and flip its sign.
   b. Second traversal: if the current edge is either an interior FS or FW edge that is part of an FS-LW-IA or FW-LS-IB segment, or an SE edge that is FW or FS and is the third leg of an LW-FS-FW or LS-FW-FS segment, then flip the site bounded on three sides by the segment and place a token on it.

For S_L, rotate the configuration 180°, repeat steps 1-6, and rotate back.

Reconstruction steps:

Given σ', the following steps reconstruct σ. For subpaths of the upper staircase that bound a white region to the left,

1. Infer and traverse the edges of the default path from start to end, but do not add them to the configuration. The first edge will be a west edge. Inductively, at a current edge, the next edge will be one of two possible edges that we'll call *straight*, for the edge that continues in the current direction, and *turning*, for the other edge.
   a. if there is a token by the next edge, flip the sign of the token site. (Figure 5(b) steps (c), (d), (f).)
   b. if the turning edge exists in the configuration (possibly after flipping), it is the next edge. (Figure 5(b) steps (b), (c), (d), (f).)
   c. otherwise the straight edge is the next edge; add it to the configuration if it doesn't exist. (Figure 5(b) steps (b), (e).)
2. Shift every edge in the white region to the left one step north.
3. Shift every edge in the white region to the left one step west.

For subpaths of the upper staircase that bound a white region to the right, reflect σ across the line $y = x$, apply steps 1-3, and reflect back. For the lower staircase, rotate the configuration 180 degrees, repeat the process, and rotate back.

4. Remove all double edges.
5. Add C to the configuration.