Slow Mixing of Markov Chains Using Fault Lines and Fat Contours[*]

Sam Greenberg [†‡]   Dana Randall [†§]

Abstract

We show that local dynamics require exponential time for two sampling problems motivated by statistical physics: independent sets on the triangular lattice (the hard-core lattice gas model) and weighted even orientations of the two-dimensional Cartesian lattice (the 8-vertex model). For each problem, there is a parameter λ known as the *fugacity*, such that local Markov chains are expected to be fast when λ is small and slow when λ is large. Unfortunately, establishing slow mixing for these models has been a challenge, as standard contour arguments typically used to show that a chain has small conductance do not seem to apply. We modify this approach by introducing the notion of *fat contours* that can have nontrivial area, and use these to establish slow mixing of local chains defined for these models.

1   Introduction

Markov chains based on local moves, known as *Glauber dynamics,* are used extensively in practice to sample from large state spaces. For example, consider the following Markov chain used to sample independent sets on $\mathbb{Z}^d$, the so-called "hard-core lattice gas model." Here, the state space is the set of independent sets and the Gibbs (or Boltzmann) distribution is parameterized by a "fugacity" $\lambda > 0$ and defined as $\pi(I) = \lambda^{|I|}/Z$, where Z is the normalizing constant known as the partition function. The local Markov chain starts at any initial state, say the empty set, and repeatedly adds and removes single vertices according to the correct conditional probabilities so that the chain converges to the Gibbs distribution. We are interested in characterizing when simple chains like this converge quickly to equilibrium, so they can be used for efficient Monte Carlo algorithms requiring many samples.

An interesting phenomenon occurs as the parameter λ is varied: for small values of λ, the local Markov chain converges quickly to stationarity, while

for large values of λ, the convergence will be prohibitively slow. When λ is sufficiently large, dense independent sets dominate the stationary distribution π and the chain will take a very long time to move from an independent set that lies mostly on the odd sublattice to one that lies mostly on the even sublattice. The key observation is that local moves that add or remove just one vertex at a time would have to visit a configuration that has roughly half of its vertices on the even sublattice and half on the odd sublattice; if λ is large, then these balanced configurations will have exponentially small Gibbs probability. This means that it will take exponential time to move from a predominantly even configuration to a predominantly odd one, and the chain will take exponential time to converge to equilibirum. This phenomenon, whereby a system behaves very differently at low or high fugacity, is well known in the statistical physics community and characterizes a *phase transition* in the underlying model. Physicists observe such a dichotomy in the context of identifying when there will be a unique limiting distribution on the infinite lattice, known as a Gibbs state; below some critical λ_c there is a unique Gibbs state, while above λ_c there are multiple Gibbs states.

In order to show that a Markov chain is slow, it suffices to show that it has exponentially small *conductance*, i.e., that there is a bottleneck in the state space. To establish small conductance rigorously, it is sufficient to show that the state space can be partitioned into three sets such that the middle set has exponentially small probability compared to the other two, and yet to move from a configuration in one of the big sets to a configuration in the other it is necessary to pass through one in the middle set.

In the context of independent sets on the Cartesian lattice, the large sets can consist of configurations that lie predominantly on each sublattice, and the middle set can contain configurations that are roughly balanced. Balanced independent sets cannot have nearly as many vertices as configurations that lie predominantly on one of the two sublattices, and therefore they have exponentially small probability when the fugacity is larger than one. In fact, it can be shown that when λ is large enough, then the total weight of the middle set will have exponentially small probability, even though it may contain a very large number of configurations. *Peierls arguments* allow us to formalize this intuition by defining "contours" between regions bounded by odd or even vertices in the independent set and constructing injections that map configurations in the middle set to ones with substantially larger stationary probability. For independent sets, the injection is constructed by shifting the interior of such a contour and adding many new vertices to the set (see, e.g., [2, 4, 3, 9]).

The situation is similar for spin models such as the Ising model. In this case, vertices are assigned + or – spins and neighboring vertices prefer to have the same spin. This preference is amplified at lower temperatures. When the temperature is sufficiently small, configurations will be predominantly one spin or the other, and local Markov chains that flip one spin at a time can be shown to converge slowly to stationarity. As before, the proofs of slow mixing are based on Peierls arguments, however in this case contours separate regions that are mostly – and mostly + and, instead of shifting the interior, the injection flips (or complements) the spins on the points within a contour [12].

2

1.1 Non-bipartite independent sets and weighted even orientations

The dichotomy observed for local Markov chains on independent sets in the Cartesian lattice is believed to persist when the underlying lattice is the 2-dimensional triangular lattice. For small λ, the local chain is known to be rapidly mixing [7], but for large enough λ, independent sets will tend to be denser and we again expect any local dynamics to be slow. Notice that on the triangular lattice there are three maximal independent sets instead of two, arising from the natural tri-partition of the lattice, which we color black, white and gray as in Figure 1. The most likely configurations should be largely monochromatic since these will be the densest. To move from an independent set that is mostly black to one that is mostly white, however, it is necessary to visit one that has fewer than half of the maximum number of vertices of each of the colors, and these should be unlikely when λ is large.

Unfortunately, Peierls arguments that succeed on the Cartesian lattice do not seem to generalize readily to non-bipatite graphs such as the triangular lattice. The problem is that a contour surrounding a region whose boundary is black might be adjacent to some vertices that are white and some that are gray; there is no guarantee that shifting or flipping the interior of the contour will allow us to add enough new vertices to sufficiently increase the weight, as we do on bipartite lattices. An example of such a problematic balanced configuration is illustrated in Figure 1.
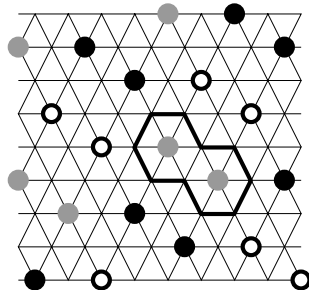

Figure 1: An independent set of the triangle lattice, with a contour separating one region of points from the others.

The second model we consider, *weighted even orientations*, also seems resistant to standard contour arguments. Given a rectangular region L in the Cartesian lattice $\mathbb{Z}^2$, the state space Ω_8 is the set of even orientations of the internal edges of L, i.e., orientations of edges incident to at least one vertex of degree 4 so that each of these vertices has even in-degree and out-degree. This is known as the *8-vertex model* in statistical physics, as there are 8 possible orientations of the edges incident to any internal vertex. We call these vertices *sources* if they have out-degree 4 and *sinks* if they have in-degree 4; all other internal vertices are called *Eulerian* as their in-degree and out-degree are

3

both 2. (Vertices on the boundary of L are not sources, sinks, or Eulerian.) For $\sigma \in \Omega_8$, let $S(\sigma)$ be number of sources and sinks in σ. Given $\gamma > 0$, we assign configuration σ probability $\pi(\sigma) = \gamma^{S(\sigma)}/Z$, where Z is the normalizing constant.

The local Markov chain $\mathcal{M}_8$ is defined on the set of even orientations. In each step, the chain chooses a cycle of length four in the lattice (i.e. a face) and either reverses the orientations of all the internal edges around that cycle or keeps them all unchanged, according to the correct conditional probabilities dictated by the Gibbs distribution. This chain can be shown to connect the state space for all finite γ.

When $\gamma = 0$, the only allowable configurations are Eulerian orientations (known as the *6-vertex model*) where every internal vertex has in-degree = out-degree = 2, and the local Markov chain is known to be efficient [6, 10]. When γ is close to 1, we can use simple coupling arguments to show that the chain is again rapidly mixing. However, when γ is sufficiently large, we expect most vertices to be sources or sinks, and the chain should take exponentially long to move from a configuration that has predominantly sources which are on the odd sublattice and sinks which are on the even sublattice, to one with predominantly even sources and odd sinks. While one would expect that configurations that are "balanced" are exponentially unlikely, this does not seem to follow from any standard contour arguments. As demonstrated by Figure 2, it is not always possible to define a map between valid configurations by flipping or shifting the interior of contour that is guaranteed to significantly increase the stationary probability, as required for the Peierls argument.
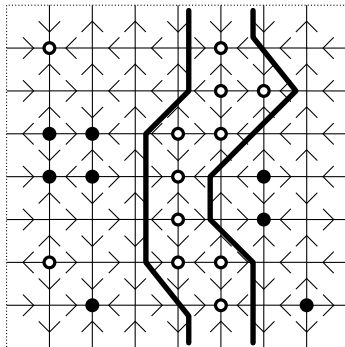

Figure 2: An even orientation of the Cartesian lattice, with a contour separating one region of sources and sinks from the others.

1.2   Our results

We provide the first rigorous proofs that the local Markov chains are slow for independent sets on the triangular lattice and for weighted even orientations on $\mathbb{Z}^2$. For even orientations, we show slow mixing of the local Markov chain on

rectangular regions with fixed boundary conditions, while for independent sets we consider rhomboidal regions with periodic (toroidal) boundary conditions; these turn out to be the simplest regions for our arguments, respectively. Our two main theorems are as follows.

Theorem 1. *Let Λ be an $n \times n$ rhomboidal region of the triangular lattice with periodic boundary conditions, let Ω_{IS} be the set of independent sets on Λ, and let $\mathcal{M}_{IS}$ be the local Markov chain on Ω_{IS}. There exists λ such that for all $\lambda' > \lambda$, the mixing time of $\mathcal{M}_{IS}(\lambda')$ is $\Omega(e^{kn})$ for some constant k.*

Theorem 2. *Let L be an $n \times n$ region in the Cartesian lattice, let Ω_8 be the set of even orientations of L, and let $\mathcal{M}_8$ be the local Markov chain on Ω_8. There exists γ such that for all $\gamma' > \gamma$, the mixing time of $\mathcal{M}_8(\gamma')$ is $\Omega(e^{kn})$ for some constant k.*

Our proofs are based on several innovations. First, we abandon the approach of partitioning the state space so that the middle set contains "balanced" configurations in the sense described above. Instead we expand the approach in [9] of basing the partition of the state space on "topological obstructions." Roughly speaking, the middle set in our partition of the state space is defined by the presence of "fault lines," or paths across the region that pass only through "unfavorable" vertices (in our cases, vacant vertices in the independent sets or Eulerian vertices in the even orientations). The absence of a fault line is characterized by the presence of a pair of monochromatic blocking paths of "favorable" vertices, and the color of these intersecting paths determines which part of the state space a configuration lies in.

To see why this is different from the standard approach, consider an independent set that contains such a pair of perpendicular paths, composed entirely of vertices on the odd sublattice, and then also includes all possible vertices on the even sublattice in the remaining space. This independent set is considered "odd" even though it has $O(n^2)$ even vertices and only $O(n)$ odd ones. This partition of the state space was shown to greatly simplify the combinatorial methods underlying the Peierls argument for bipartite independent sets [9] and can be extended to the models we consider here as well.

It is still the case that the 1-dimensional contours on bipartite lattices, used for independent sets and the Ising model, do not readily generalize to our problems for the reasons outlined above. However, a generalized notion of contours that includes a larger 2-dimensional region can be made to work. Instead of defining a *minimal* connected set of unfavorable vertices, we define *fat contours* to be *maximal* connected sets of unfavorable vertices. We define an injective map from configurations in the "middle set" of the state space by replacing the entire fat contour with favorable vertices (a maximal independent set or a maximal set of sources and sinks). This typically involves shifting parts of the independent set outside of the fat contour in two different directions. We then show that the gain is sufficient to outweigh the amount of information needed to invert the map.

Last, in order to show slow mixing in the context of independent sets on the triangular lattice with periodic boundary conditions, it is necessary to talk about

5

multiple non-contractible fault lines, depending on the color of the boundary vertices. On the Cartesian lattice with periodic boundary conditions, it was only necessary to find two non-contractible cycles and to shift (or flip) the configuration between these; on the triangular lattice it is sometimes necessary to find three non-contractible cycles, as there are three types of regions.

Notice that it is not necessary to consider periodic boundary conditions in the context of the 8-vertex model because our Peierls argument is based on a map that just flips the interior of a fat contour. For independent sets, the map requires shifting the interior of a fat contour, and this turns out to be simpler on regions with periodic boundary conditions since we do not lose part of the configuration after the shift as we would if we had non-periodic boundaries.

In Section 2, we cover some background material for basic Markov chain mechanics more formally. In Section 3 we introduce the notion of fat contours and prove Theorem 2 for the 8-vertex model. In Section 4 we extend the argument based on fat contours to the torus and prove Theorem 1 for independent sets on the triangular lattice. In Section 5 we conclude with extensions and open problems.

2 Preliminaries

Let $\mathcal{M}$ be an ergodic (i.e., irreducible and aperiodic), reversible Markov chain with finite state space Ω, transition probability matrix P, and stationary distribution π. Let $P^t(x, y)$ be the t-step transition probability from x to y.

Definition 1. *The* total variation distance *at time t is*

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

Definition 2. *For $\varepsilon > 0$, the* mixing time $\tau = \min\{t : \|P^{t'}, \pi\| \leq \varepsilon, \forall t' \geq t\}$.

We say a Markov chain is *rapidly mixing* if the mixing time is bounded by a polynomial in n and *slowly mixing* if the mixing time is exponential in n.

The *conductance* [5, 11] provides a way to upper and lower bound the mixing time, and is defined as

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

The following theorem relating the conductance and the mixing time is due to Jerrum and Sinclair [11] and is useful for showing a chain is slowly mixing.

Theorem 3. *For any reversible Markov chain with conductance Φ,*

$$\tau(\varepsilon) \geq \frac{1 - 2\Phi}{2\Phi} \ln \varepsilon^{-1}.$$

Thus, to lower bound the mixing time, it is sufficient to show that the conductance is sufficiently small.

3 Weighted Even Orientations

Let L be an $n \times n$ region of the Cartesian lattice $\mathbb{Z}^2$, and let Ω_8 be the set of all orientations of the internal edges of L so that all vertices in the interior of L have even in-degree and even out-degree. Given a fixed constant $\gamma > 0$ representing the fugacity, for each $\sigma \in \Omega_8$, the Gibbs measure is $\pi(\sigma) = \gamma^{S(\sigma)}/Z$, where $S(\sigma)$ is the number of sources and sinks in σ and Z is the normalizing constant.

Let the Markov chain $\mathcal{M}_8$ be the local chain on Ω_8 that chooses a face of L uniformly at random and either flips the orientations of all internal edges incident to that face or does nothing, according to the correct conditional probabilities. More precisely, let s_1 be the number of sources and sinks among the four vertices defining the face in the current configuration and let s_2 be the number of sources and sinks that would surround that face if we were to flip the orientations of the bounding edges. Then we flip the orientations of the bounding edges with probability $\lambda^{s_2}/(\lambda^{s_1} + \lambda^{s_2})$ and we keep the orientation unchanged with probability $\lambda^{s_1}/(\lambda^{s_1} + \lambda^{s_2})$. This chain is ergodic and converges to π. We are interested in the mixing rate of the chain.

When $\lambda = 1$, all even orientations are equally likely and the probability of flipping the orientation of edges around any face is the same. If we define the distance between two configurations to be the number of edges in which their orientations differ, then it is easy to construct a coupling argument to show that the chain is rapidly mixing. (See [1] for the coupling theorem.) In fact, when $\lambda = 1$, all moves occur with probability $1/2$. We can define a coupling so that the distance function never increases during moves of the coupled chain.

When $\lambda \neq 1$ the distance function can increase as well as decrease, but the coupling argument still can be made to work when λ is sufficiently close to 1. We also know that the chain is rapidly mixing when $\lambda = 0$, as this corresponds to Eulerian orientations on Cartesian lattice regions (since there cannot be any sources or sinks)t. The chain is known to be rapidly mixing in this special case [6, 10].

However, when λ is large, the Markov chain behaves quite differently and we verify that the convergence to equilibrium requires exponential time. Before proceeding with our analysis of the mixing time, we present a reinterpretation of Ω_8 as an edge coloring. For every configuration, color an edge white if it points from a vertex on the even sublattice to a vertex on the odd one, and color it black if it points from an odd vertex to an even one. An example of this transformation is shown in Figure 3.

Now Ω_8 can be seen as the set of edge-colorings where every internal vertex has an even number of edges of each color. The sources and sinks are now *monochromatic* vertices, i.e., all incident edges are the same color. We call internal vertices that are incident to both black and white edges *bichromatic*. (As before, the vertices on the boundary are neither monochromatic or bichromatic.) The intuition behind the slow mixing is that it takes a long time to move from a configuration that is predominantly white to one that is predominantly black. This is because it is necessary to pass through configurations that have a large number of bichromatic vertices, and these have much smaller stationary prob-

ability. The goal is to show that there is a partition of the state space that defines a bad cut so that we can use the conductance theorem to show that the chain mixes slowly. Rather than the natural choice of partitioning the state space according to the relative numbers of black and white vertices in the configurations, we instead use the approach of [9] and partition according to "fault lines."
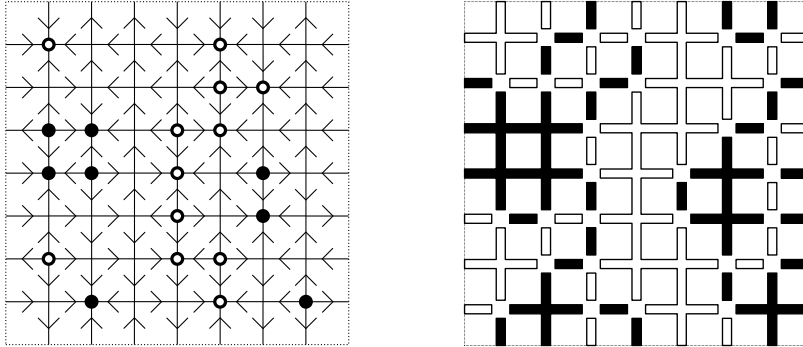

Figure 3: A configuration $\sigma \in \Omega_8$ (with sources and sinks marked) and the corresponding edge-coloring.

Call two vertices of L *edge-adjacent* if they share an edge of L and call two vertices of L *face-adjacent* if they lie on a common face of L. ("Face-adjacent" is a looser term, as the vertices can be edge-adjacent or diagonally opposite across a face.) Define a *vertical fault line* to be a connected path of bichromatic face-adjacent vertices where one end is adjacent to the top row of L and the other is adjacent to the bottom. A *horizontal fault line* is defined similarly. Let $\mathcal{F} \subset \Omega_8$ be the set of all configurations containing a fault line.

Now we define a *vertical bridge* to be an edge-connected path of monochromatic vertices where one end is adjacent to the top of L and the other is adjacent to the bottom. A *horizontal bridge* is defined similarly. We say that a configuration has a *cross* if it contains both a horizontal and a vertical bridge (of the same color). Let $\mathcal{W}$ be the set of configurations containing a white cross and $\mathcal{B}$ be the set of configurations containing a black cross. We now show that these three sets $\mathcal{F}$, $\mathcal{W}$, and $\mathcal{B}$ are disjoint and characterize all of Ω_8.

Lemma 1. *We may partition Ω_8 into $\mathcal{F}$, $\mathcal{W}$, and $\mathcal{B}$. That is, every configuration of Ω_8 has either a fault line, a white cross, or a black cross (but no two of these).*

Proof. The sets $\mathcal{W}$ and $\mathcal{B}$ are clearly disjoint, as every configuration in $\mathcal{W}$ has a vertical white bridge and every configuration in $\mathcal{B}$ has a horizontal black bridge; no configuration can have both. Similarly $\mathcal{F}$ is disjoint from $\mathcal{W}$ and $\mathcal{B}$ because fault lines obstruct crosses; if a configuration has a horizontal fault line it can not have a vertical bridge. What remains to be shown is that these three sets cover all of Ω_8, that any configuration without a cross must have a fault line.

Let σ be a configuration with no horizontal bridge and let the vertex set T be the union of all bichromatic face-connected paths with one end adjacent to the top of L. As the top row of internal vertices of L is not a bridge, T is non-empty (although not necessarily connected). An example of T is illustrated in Figure 4. By the definition of T, each vertex face-adjacent to T is monochromatic, or it would be included in T. If T is nowhere adjacent to the bottom of L, then the vertices adjacent to T contain a horizontal bridge (and therefore a contradiction). Hence T must contain a vertex adjacent to the bottom of L, and so contains a vertical fault line.
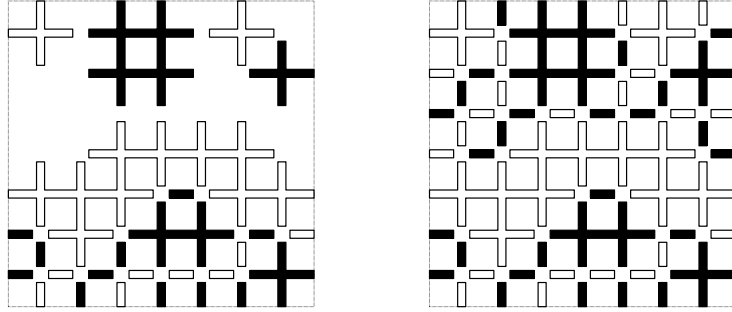

Figure 4: A configuration σ with and without T removed. Notice that, as T does not reach the bottom of L, σ contains a horizontal bridge just below T.

By a similar argument, if σ has no vertical bridge then σ has a horizontal fault line. $\qquad\square$

We now show that for $\mathcal{M}_8$ to pass from $\mathcal{W}$ to $\mathcal{B}$, it must pass through $\mathcal{F}$, so the edges incident to $\mathcal{F}$ define a cut in the state space.

Lemma 2. *For transition probability $P(\cdot, \cdot)$ of $\mathcal{M}_8$, we have $P(\sigma_W, \sigma_B) = 0$ for all $\sigma_W \in \mathcal{W}$ and $\sigma_B \in \mathcal{B}$.*

Proof. The first important thing to note is that a single move of $\mathcal{M}_8$ cannot turn a monochromatic white vertex into a monochromatic black vertex (or vice versa), as the move can only change at most two incident edges. We may now prove the lemma by contradiction.

Assume we do have two configurations $\sigma_W \in \mathcal{W}$ and $\sigma_B \in \mathcal{B}$ that differ by a single move of $\mathcal{M}_8$, say on face f. The configuration σ_W must have a horizontal white bridge, where each edge on the bridge is incident to only white edges. Moreover, each vertex on this bridge must be incident to at least two white edges that are not on f. If we recolor the edges of f black, each of these vertices remain white or become bichromatic, and therefore none of these can be part of a black bridge in σ_B. But σ_B has a vertical black bridge consisting of black vertices, which is a contradiction since every vertical bridge must intersect each horizontal bridge. Therefore we can conclude that $\mathcal{W}$ and $\mathcal{B}$ are not connected by a single move in $\mathcal{M}_8$.

$\qquad\square$

Next, we proceed to show that the stationary probability of $\mathcal{F}$ is exponentially small. We define a *fat contour* to be a maximally face-connected set of bichromatic vertices containing a fault line. As a fault line is a face-connected set of bichromatic vertices, any configuration with a fault line has a fat contour. To bound $\pi(\mathcal{F})$, we define a mapping $\psi : \mathcal{F} \to \Omega_8$ which takes $\sigma \in \mathcal{F}$ and recolors edges incident to a fat contour so that the contour contains only white (or black) vertices. Although ψ is not one-to-one, we will show that for every $\sigma' \in \text{Img}(\psi)$, the stationary probability of the pre-image $\psi^{-1}(\sigma') = \{\sigma \in \mathcal{O} : \psi(\sigma) = \sigma'\}$ is exponentially less than the stationary probability of σ'.

Our definition of $\psi(\sigma)$ proceeds as follows. First, choose an arbitrary fat contour of σ, F (e.g. farthest to the left or top). Define an *island* to be a connected component of vertices after the removal of F and the boundary of L. Notice that, while an island might contain bichromatic vertices, the vertices of an island adjacent to F are either entirely white or entirely black. Our mapping reverses the color of every edge with at least one vertex in a white island. (We leave black islands as they are.) Now the edges incident to F are entirely black and F can be recolored completely with black monochromatic vertices. An example of this modification is in Figure 5. The resulting configuration has $|F|$ more monochromatic vertices than the pre-image, corresponding to additional sources and sinks in the original even orientation.
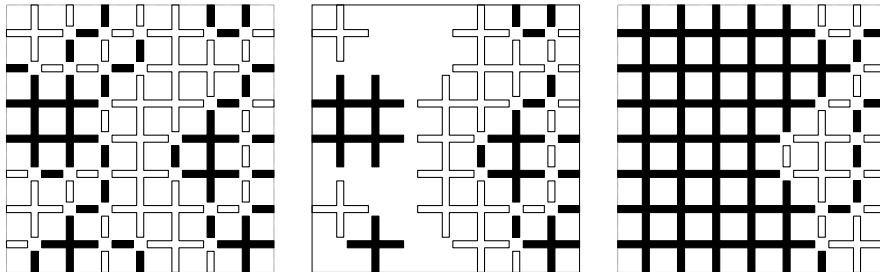

Figure 5: A coloring σ, σ with the fat contour removed, and $\psi(\sigma)$.

Notice that to find the inverse of ψ, we need only the location of F and the colorings of the edges incident to vertices of F. The edges on the boundary of the fat contour will then define whether a island was originally white or black, and we can recover the colors of edges in those islands accordingly. In Lemma 3 below, we bound the number of possible fat contours and their edge colorings. This allows us to bound the size of $\mathcal{F}$ in Lemma 4.

Lemma 3. *If F is a fat contour such that $|F| = \ell$, then there are fewer than $2n \cdot 64^\ell$ possible choices for the vertices of F and fewer than 6^ℓ colorings of the edges incident to those vertices.*

Proof. Without loss of generality, we assume F is a vertical fat contour. This at most halves the total choices for F. The fat contour must then include some vertex adjacent to the top of L. The location of this vertex together with

10

a description of a DFS traversal of F starting at this vertex is sufficient to reconstruct the vertices in F. As there are at most eight choices at each step in the DFS (four adjacencies along edges and four diagonals across faces) there are at most $8^{2\ell}$ such traversals. With n choices for the starting vertex, we have the bound on F.

Finally, for each vertex in F, there are at most 6 possible Eulerian orientations of the incident edges, so 6^ℓ is an immediate (albeit weak) upper bound on the number of colorings of the edges incident to F. □

Lemma 4. *There exist constants* $\gamma_0, n_0, c > 1$ *such that, if* $\gamma > \gamma_0$ *and* $n > n_0$, *then* $\pi(\mathcal{F}) < c^{-n}$.

Proof. For each $\ell \in [n, n^2]$, let $\mathcal{F}_\ell$ be the edge-colorings in $\mathcal{F}$ where the fat contour chosen by ψ is of size ℓ. Then, for each $\sigma' \in \mathrm{Img}(\psi)$,

$$\pi(\psi^{-1}(\sigma')) = \sum_{\ell=n}^{n^2} \sum_{\sigma \in \mathcal{F}_\ell : \psi(\sigma)=\sigma'} \pi(\sigma)$$

$$= \sum_{\ell=n}^{n^2} \sum_{\sigma \in \mathcal{F}_\ell : \psi(\sigma)=\sigma'} \pi(\sigma') \cdot \gamma^{-\ell}$$

$$\leq \sum_{\ell=n}^{n^2} 2n 64^\ell 6^\ell \cdot \pi(\sigma') \cdot \gamma^{-\ell}$$

$$< 2n^3 \left(\frac{384}{\gamma} \right)^n \pi(\sigma').$$

This bound on the pre-image allows us to bound $\pi(\mathcal{F})$ as follows:

$$\pi(\mathcal{F}) = \sum_{\sigma' \in \mathrm{Img}(\psi)} \pi(\psi^{-1}(\sigma'))$$

$$< \sum_{\sigma' \in \mathrm{Img}(\psi)} 2n^3 \left(\frac{384}{\gamma} \right)^n \pi(\sigma')$$

$$= 2n^3 \left(\frac{384}{\gamma} \right)^n \pi(\mathrm{Img}(\psi))$$

$$< 2n^3 \left(\frac{384}{\gamma} \right)^n.$$

Taking $\lambda > 384$ yields the lemma. □

The final step of the proof of Theorem 2 is to show that this lemma is sufficient to bound the conductance. Consider the cut defined by the set $\mathcal{W}$. By symmetry $\pi(\mathcal{W}) = \pi(\mathcal{B}) = (1 - \pi(\mathcal{F}))/2$. Therefore, as $\mathcal{W}$ satisfies the restrictions of the set S in the definition of Φ,

$$\Phi \leq \frac{\pi(\mathcal{F})}{\pi(\mathcal{W})} \leq \frac{c^{-n}}{(1 - c^{-n})/2},$$

where $c > 1$. Appealing to Theorem 3, this proves Theorem 2.

4   Independent Sets on the Triangular Lattice

A similar approach to the one used in Section 3 can be applied to the context of independent sets on the triangular lattice. The primary challenge in extending the argument to this context is that the map now needs to be based on shifting the interior of a contour, and here the lattice regions have periodic boundary conditions.

Let Λ be a $3n \times 3n$ rhomboidal region of the triangular lattice with periodic boundary conditions and let Ω_{IS} be the set of all independent sets on Λ. Given a constant $\lambda > 0$, define a probability distribution $\pi(I) = \lambda^{|I|}/Z$, where Z is the normalizing constant.

The chain $\mathcal{M}_{IS}$ is defined as follows: choose a vertex of Λ uniformly at random and add or remove that vertex from I with the correct conditional probabilities, if possible. That is, if the vertex could be included without violating the independence condition, add it with probability $\lambda/(1 + \lambda)$ and remove it with probability $1/(1 + \lambda)$. Otherwise do nothing.

The lattice has a natural tri-partition, which we color black, white, and gray. Call a face of Λ *empty* if it is not incident to any vertex of I, as illustrated in Figure 6. Call two faces of Λ adjacent if they share at least one vertex. We then define a *fault line* to be a non-contractible cycle of empty faces. We let $\mathcal{F} \subset \Omega_{IS}$ be the set of all independent sets with at least one fault line.
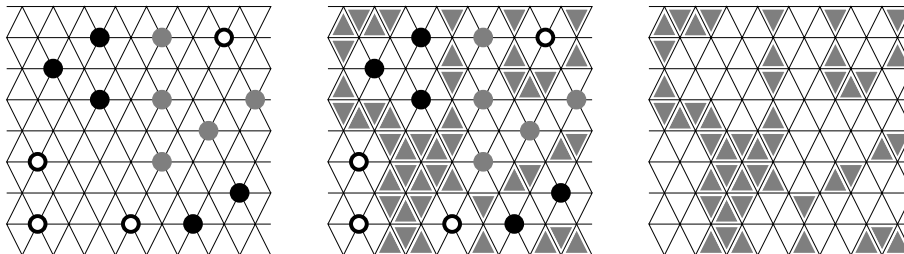

Figure 6: An independent set on Λ and the corresponding empty faces.

The obstruction preventing a fault line must be a set of tightly packed vertices of I. Call two vertices of an independent set *touching* if they are incident to faces which share an edge. Note that touching vertices must have the same color. We define a monochromatic *bridge* to be a non-contractible cycle of touching vertices of I. For any non-contractible cycle, the *winding vector* is an ordered pair of integers (w_x, w_y), where w_i represents the net number of times the cycle intersects an elementary loop in the ith lattice direction. For instance, the elementary loops have winding vector $(0, 1)$ and $(1, 0)$. We say a configuration has a monochromatic *cross* if it contains two bridges with different winding vectors. (As cycles with different winding vectors must intersect, these two bridges are

automatically of the same color.) Let $\mathcal{B}$, $\mathcal{W}$, and $\mathcal{G}$ be the set of independent sets containing black, white, and gray crosses, respectively.

We first show that these sets define a partition of Ω_{IS}.

Lemma 5. *Every independent set in Ω_{IS} has a fault line or a white, black, or gray cross, but no two of these.*

Proof. This proof is similar to the argument in [9] that shows that independent sets on the Cartesian lattice without fixed boundary conditions must have horizontal and vertical bridges of one color or there must be a fault line. A key point we use here is that any two non-contractible cycles of different winding vectors must intersect.

The sets $\mathcal{B}$, $\mathcal{W}$, and $\mathcal{G}$ are disjoint, as an independent set cannot have two crosses of different colors, as that would involve two bridges of different winding vectors and different colors, whose intersection would lead to a contradiction. Similarly, $\mathcal{F}$ is disjoint from $\mathcal{B}$, $\mathcal{W}$, and $\mathcal{G}$, as no set can have both a cross and a fault line; the fault line must intersect at least one of the bridges, which is impossible.

To see that there must be either a cross or a fault line, examine the torus after the removal of a bridge. The remaining space is a non-contractible strip of the torus of the same winding vector as the bridge. If there exists a path of touching vertices across this strip, we find a bridge of a different winding vector and therefore a cross. This is illustrated in Figure 7. On the other hand, if no such path exists, then there must exist a fault line along the strip.
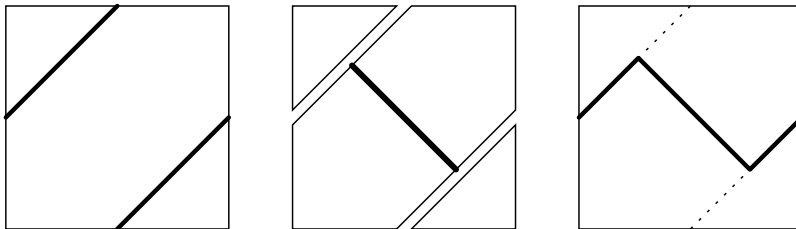

Figure 7: A cycle of winding vector $(1,1)$, the strip remaining after its removal, and the cycle of winding vector $(0,1)$ created using the path across the strip.

$\square$

We now show that removing $\mathcal{F}$ disconnects $\mathcal{W}$, $\mathcal{B}$, and $\mathcal{G}$. Let $P(\cdot, \cdot)$ be the transition probabilities of $\mathcal{M}_{IS}$.

Lemma 6. *Let $I_1 \in \mathcal{W}$, $I_2 \in \mathcal{B}$ and $I_3 \in \mathcal{G}$ be three independent sets. Then $P(I_i, I_j) = 0$ for all $i \neq j$.*

Proof. Individual moves of $\mathcal{M}_{IS}$ either add or remove a single vertex. Clearly it requires multiple moves to eliminate one cross and complete another. $\square$

We now show that the stationary probability of $\mathcal{F}$ is exponentially small. In doing so, we again extend fault lines to 2-dimensional regions. Define a *fat contour* to be a maximally connected set of empty faces containing a fault line. We define a mapping $\psi : \mathcal{F} \to \Omega_{IS}$ which eliminates at least one fat contour. Although the mapping is not one-to-one, we will show that each $I' \in \mathrm{Img}(\psi)$ has a pre-image whose total weight is exponentially smaller.

To bound the number of sets in this pre-image, we bound the number of fat contours.

Lemma 7. *If F is a fat contour with ℓ faces, then there are at most $2n^2 36^\ell$ choices for the locations of those faces.*

Proof. First, we limit the notion of adjacencies in F. Define two adjacent faces to be *edge-adjacent* if they share an edge. Call them *point-adjacent* if they share a single vertex and yet are not both adjacent to a common face. As illustrated in Figure 8, not all pairs of adjacent faces are edge- or point-adjacent. However, note that edge- and point-adjacencies suffice to connect F, as a vertex of I removes a complete hexagon from F. We can therefore find a traversal of F that uses only edge- and point-adjacencies.

There are $2n^2$ choices for a face to start the DFS of F. Then each step of the DFS has six possible directions (three edge-adjacencies and three point-adjacencies), so there are at most 6^{2l} possible traversals starting at f. $\square$

Figure 8: Adjacent faces which are point-adjacent, edge-adjacent, and neither.

Our definition of ψ is slightly more complicated than in Section 3 because we are considering toroidal regions. Suppose first that F contains two fault lines with different winding vectors. Then the complement of F contains only regions whose boundaries are contractible. By the maximality of F, each of these regions has a monochromatic boundary, so we may refer to these connected components (or "islands") by the colors of their boundaries. Note that if the lattice partition is colored as in Figure 1, if we shift all white islands one space East so that their boundaries become gray, and shift all black islands one space to the North-East so that their boundaries also become gray (and leave gray islands as they are), we form a new independent set of the same size. After this shift, all vertices incident to F are gray, so ψ may fill the entire fat contour with gray vertices. If F has ℓ faces, ψ adds exactly $\ell/6$ vertices to I, each the center of a vacant hexagon. Such a transformation is illustrated in Figure 9.

Unfortunately, fat contours need not contain multiple fault lines with differing winding vectors and, indeed the complement of F can contain regions whose boundaries are non-contractible. If this is the case, then F has a bridge on each side. Define a *bulge* to be a maximal set of touching vertices of I which contain
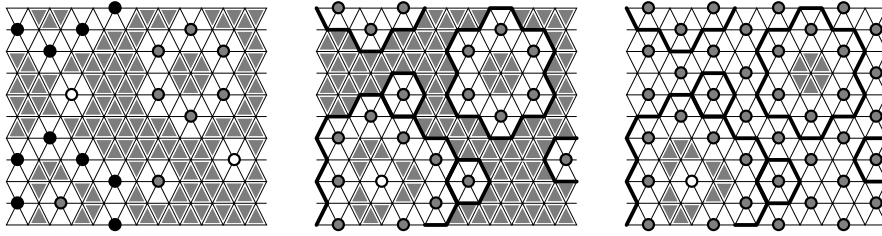
Figure 9: A set σ containing a pair of fault lines with differing winding vectors, and $\psi(\sigma)$.

a bridge. In this case F must be incident to a bulge on each side. If these are of the same color, ψ can shift all islands within F to that color and fill F, as in Figure 10.
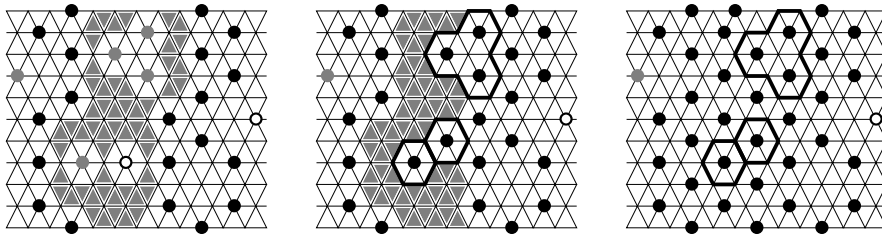

Figure 10: A set σ containing a fault line with a black bulge on each side, and $\psi(\sigma)$.

Complicating matters further, there may be no fat contour incident to two bulges of the same color. For instance, let F_1 be a fat contour incident to a white bulge on the left and a black one on the right. To the right of the black bulge there must be another fat contour, F_2. If F_2 is incident to a white bulge, ψ can shift the islands of F_1 and F_2 and the black bulge all to white. Then ψ can fill both fat contours with white vertices. If F_1 and F_2 contain ℓ_1 and ℓ_2 faces (respectively), then ψ adds $(\ell_1 + \ell_2)/6$ vertices to I.

In one final case, suppose we have no pair of neighboring fat contours bordered by bulges of the same color. Then there must then be a third fat contour F_3 which is incident to still another bulge. Luckily we only have three colors of bulges; at some point these colors must repeat. For example, if we have, in order, a black bulge, F_1, a white bulge, F_2, a gray bulge, F_3, and then a black bulge, then ψ can shift the white bulge, the gray bulge, and all islands of the fat contours to black. We may then fill all three fat contours with black vertices.

To find the inverse of ψ, note that we need only the faces of the fat contour(s); the colors of the neighboring vertices can be inferred from the shape of F, and these colors define the direction of the shift.

We may now bound the stationary probability of $\mathcal{F}$.

15

Lemma 8. *There exist constants $\lambda_0, n_0, c > 1$ such that, if $\lambda > \lambda_0$ and $n > n_0$, then $\pi(\mathcal{F}) < c^{-n}$.*

Proof. For each $\ell \in [n, 2n^2]$, let $\mathcal{F}_\ell \subset \mathcal{F}$ be the independent sets where the fat contours chosen by ψ contain a total of ℓ faces. Given ℓ_1, ℓ_2, ℓ_3 such that $\ell_1 + \ell_2 + \ell_3 = \ell$, Lemma 7 shows that there are at most $\prod_{i=1}^{3} 2n^2 36^{\ell_i} = 8n^6 36^\ell$ choices of faults such that $|F_i| = \ell_i$. Then, for each $I' \in \text{Img}(\psi)$,

$$\pi(\psi^{-1}(I')) = \sum_{\ell=n}^{2n^2} \sum_{I \in \mathcal{F}_\ell : \psi(I)=I'} \pi(I)$$

$$= \sum_{\ell=n}^{2n^2} \sum_{I \in \mathcal{F}_\ell : \psi(I)=I'} \pi(I') \cdot \lambda^{-\frac{\ell}{6}}$$

$$< \sum_{\ell=n}^{2n^2} l^3 8n^6 36^\ell \cdot \pi(I') \cdot \lambda^{-\frac{\ell}{6}}$$

$$< 64n^{12} \left(\frac{36^6}{\lambda} \right)^{\frac{2n^2}{6}} \pi(I').$$

The bound on the pre-image then allows us to bound $\pi(\mathcal{F})$ as follows:

$$\pi(\mathcal{F}) = \sum_{I' \in \text{Img}(\psi)} \pi(\psi^{-1}(I'))$$

$$< \sum_{I' \in \text{Img}(\psi)} 64n^{12} \left(\frac{36^6}{\lambda} \right)^{2n^2} \pi(I')$$

$$= 64n^{12} \left(\frac{36^6}{\lambda} \right)^{\frac{2n^2}{6}} \pi(\text{Img}(\psi))$$

$$< 64n^{12} \left(\frac{36^6}{\lambda} \right)^{\frac{2n^2}{6}}.$$

Taking $\lambda > 36^6$ completes the proof. $\qquad\square$

Note that by symmetry the sets $\mathcal{W}, \mathcal{B}$ and $\mathcal{G}$ have equal stationary probability. Observing now that their total weight is at least $1 - c^{-n}$, the conductance arguments of Sections 2 allow us to complete the proof of Theorem 1. Thus, we have shown that local Markov chains for sampling independent sets on the triangular lattice converge slowly to equilibrium for large λ.

5   Conclusions

We introduced a new tool for showing slow mixing of a local Markov chain via simple combinatorial arguments. The main contribution is the use of fat

contours that consist of large, full-dimensional regions of unfavorable vertices. We then demonstrated how to use fat contours in the context of weighted even orientations on the Cartesian lattice and independent sets on regions of the triangular lattice with toroidal boundary conditions. In this presentation our goal was presenting basic arguments and we have not made an attempt to optimize constants. It would be interesting to see how these arguments can be improved to show slow mixing for a wider range of values of λ for each model.

Using essentially the same arguments, the slow mixing results can be extended to weighted even orientations on regions with periodic boundary conditions, as well as Cartesian lattice regions in higher dimensions. In the latter case, fat contours have non-trivial volume and are again encoded by a depth-first search of the components connecting large regions of "unfavorable" vertices.

Finally, we conclude with an open problem. We have shown that local chains are slow for weighted independent sets when λ is large enough and and we know that the chain is rapidly mixing when $\lambda = 0$ or very close to 1. We conjecture that there is a unique value for λ below which the chain is rapidly mixing and above which it is slowly mixing. This implies that we have polynomial time convergence for all λ between 0 and 1, although this remains unknown.

References

[1] D. Aldous. Random walks on finite groups and rapidly mixing Markov chains. *Séminaire de Probabilités* XVII, *Springer Lecture Notes in Mathematics* **986**:243–297, 1981/82.

[2] C. Borgs, J.T. Chayes, A. Frieze, J.H. Kim, P. Tetali, E. Vigoda, and V.H. Vu. Torpid mixing of some MCMC algorithms in statistical physics. *Proceedings of the 40th IEEE Symposium on Foundations of Computer Science*, 218–229, 1999.

[3] D. Galvin and J. Kahn. On phase transitions in the hard-core model on Z^d. *Combinatorics, Probability and Computing*, **13**: 137-164, 2004.

[4] D. Galvin and D. Randall. Sampling 3-colorings of the discrete torus. *Proceedings of the 17th ACM/SIAM Symposium on Discrete Algorithms*, 376–384, 2007.

[5] G.F. Lawler and A.D. Sokal, A.D. Bounds on the L2 spectrum for Markov chain and Markov processes: a generalization of Cheeger's inequality. *Transactions of the American Mathematical Society*, **309**: 557-580, 1988.

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

[7] M. Luby and E. Vigoda. Fast convergence of the Glauber dynamics for sampling independent sets. *Random Structures and Algorithms*, **15**: 229–241, 1999.

[8] N. Metropolis, A. W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, **21**: 1087–1092, 1953.

[9] D. Randall. Slow Mixing of Glauber Dynamics via Topological Obstructions. *Proceedings of the 16th ACM/SIAM Symposium of Discrete Algorithms (SODA)*, 870–879, 2006.

[10] D. Randall and P. Tetali. Analyzing Glauber dynamics by comparison of Markov chains. *Journal of Mathematical Physics*, **41**: 1598–1615, 2000.

[11] A. Sinclair and M. Jerrum. Approximate Counting, Uniform Generation and Rapidly Mixing Markov Chains *Information and Computation*, 82: 93-133, 1989.

[12] L.E. Thomas. Bound on the mass gap for finite volume stochastic ising models at low temperature. *Communications in Mathematical Physics*, **126:** 1–11, 1989.