Sampling Weighted Perfect Matchings on the Square-Octagon Lattice

Prateek Bhakta *　　Dana Randall †

Abstract

We consider perfect matchings of the square-octagon lattice, also known as "fortresses" [16]. There is a natural local Markov chain on the set of perfect matchings that is known to be ergodic. However, unlike Markov chains for sampling perfect matchings on the square and hexagonal lattices, corresponding to domino and lozenge tilings, respectively, the seemingly related Markov chain on the square-octagon lattice appears to converge slowly. To understand why, we consider a weighted version of the problem. As with domino and lozenge tilings, it will be useful to view perfect matchings on the square-octagon lattice in terms of sets of paths and cycles on a corresponding lattice region; here, the paths and cycles lie on the Cartesian lattice and are required to turn left or right at every step. For input parameters λ and μ, we define the weight of a configuration to be $\lambda^{|E(\sigma)|}\mu^{|V(\sigma)|}$, where $E(\sigma)$ is the total number of edges on the paths and cycles of σ and $V(\sigma)$ is the number of vertices that are not incident to any of the paths or cycles in σ. Weighted paths already come up in the reduction from perfect matchings to turning lattice paths, corresponding to the case when $\lambda = 1$ and $\mu = 2$.

First, fixing $\mu = 1$, we show that there are choices of λ for which the chain converges slowly and another for which it is fast, suggesting a phase change in the mixing time. More precisely, the chain requires exponential time (in the size of the lattice region) when $\lambda < 1/(2\sqrt{e})$ or $\lambda > 2\sqrt{e}$, while it is polynomially mixing at $\lambda = 1$. Further, we show that for $\mu > 1$, the Markov chain $\mathcal{M}$ is slowly mixing when $\lambda < \sqrt{\mu}/(2\sqrt{e})$ or $\lambda > 2\mu\sqrt{e}$. These are the first rigorous proofs explaining why the natural local Markov chain can be slow for weighted fortresses or perfect matchings on the square-octagon lattice.

1 Introduction

Perfect matchings have been the cornerstone problem underlying many fundamental computational complexity questions. The seminal work of Edmonds [5] established that the decision and construction problems were in P, while Valiant [23] showed that counting perfect matchings is #P-complete. Jerrum, Sinclair and Vigoda [11] showed how to approximately count and sample perfect matchings in any bipartite graph efficiently, although the complexity remains open on general graphs.

Likewise, perfect matchings also have captivated the statistical physics community, who study them in the context of *dimer models*. In this setting, edges in a matching represent diatomic molecules, or dimers, and perfect matchings correspond to dimer coverings of a graph. Physicists study the thermodynamic properties of physical systems by relating fundamental quantities to weighted sums over the set of configurations, such as perfect matchings on finite lattice regions. Kasteleyn et al. showed how to exactly count perfect matchings on any planar graph in polynomial time by calculating a Pfaffian on a directed version of the adjacency matrix [12, 22]. When the underlying graph is a lattice region, there are alternative, faster determinant-based methods for counting matchings [7, 13]. For the most common lattices regions on $\mathbb{Z}^2$ and the hexagonal lattice, where perfect matchings correspond to domino and lozenge tilings, respectively, various stochastic approaches have been explored to improve both the efficiency and the simplicity of the algorithms. For example, a common approach for sampling perfect matchings on $\mathbb{Z}^2$ that is popular among experimentalists is based on *dimer rotations*, where we choose a unit face uniformly, and if this 4 cycle contains two edges of the matching, the Markov chain can replace those two with the other two edges around the alternating cycle. A similar approach on the hexagonal lattice replaces, if possible, the three alternating edges around the face with their complement.

The Markov chain based on dimer rotations was first studied by Propp and Wilson [18]. They showed that their "coupling-from-the-past" algorithm could be run on dimer covers of the Cartesian lattice Z^2 to generate perfect uniform samples of perfect matchings, although they supplied no guarantees that the algorithm would terminate in polynomial time. The proof that the expected time to converge is polynomially bounded was provided by Luby et al. [15], Randall and Tetali [19] and further improved by Wilson [25]. Using the paradigm

*College of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0765; Supported in part by NSF CCF-1526900; pbhakta@gatech.edu.

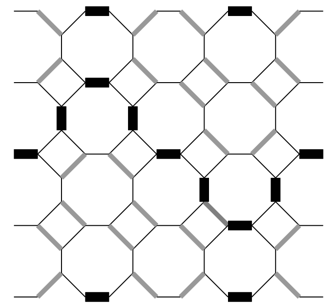
†College of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0765; Supported in part by NSF CCF-156900 and CNS-1544090; randall@cc.gatech.edu.

of coupling-from-the-past, matchings on many other lattices have been explored as well, providing perfectly random samples (although not always efficiently) and generating many conjectures about convergence times and stationary distributions underlying these chains. A compelling example is perfect matchings on the square-octagon lattice, Λ_{so}, where the dual is a dimer problem on a graph of squares and triangles known as "fortresses" [17]. Many remarkable properties of lozenge and domino tilings, such as the existence of frozen regions at equilibrium, are known to hold for fortresses [17].

There is a natural analogue of the dimer-rotating Markov chain as well, which has been used experimentally to study these matchings. This chain is known to connect the state space of perfect matchings [1], but nothing is known rigorously about its convergence time. Although related Markov chains on other lattices are known to converge in polynomial time, including the Cartesian and hexagonal lattices, simulations suggest that on the square-octagon lattice this chain may in fact require exponential time.

An intuitive explanation for why the convergence time of this Markov chain is likely exponential can be seen by interpreting these perfect matchings as a "contour model" [1, 8, 14] Given any perfect matching on a simply connected region of the square-octagon lattice, we can contract all vertices of each square into a single vertex, leaving only the edges that border two octagons. The resulting configuration will be a collection of edges on the Cartesian lattice, where every vertex, except possibly those on the boundary, must have even degree, and where each vertex of degree 2 must be incident to one horizontal and one vertical edge (see Fig. 1). We can decompose these sets of edges into a collection of "turning paths" that connects certain boundary vertices and closed "turning cycles." The turning property refers to the fact that traversals of the edges of a path or cycle are required to turn left or right at *every* step.

It is important to note that this map is not bijective, and each turning graph is the image of 2^k perfect matchings on the square-octagon lattice, where k is the number of degree 0, or *free*, vertices in σ (see Fig. 2). Each free vertex corresponds to a square on Λ_{so} containing two matched edges, and there are exactly two ways this can occur. Thus, in order to generate perfect matchings on the square octagon lattice, it suffices to generate turning graphs with the weight of each configuration σ proportional to $2^{V(\sigma)}$, where $V(\sigma)$ is the number of free vertices in σ; configurations with more free vertices will have greater weight, and this weighting penalizes configurations with long paths and cycles. This is the key insight gained by considering this path representation, letting us use analysis similar to many other models in statistical physics, most notably the Ising model, that are slowly mixing at low temperatures when long contours are similarly disfavored [14].


(a)

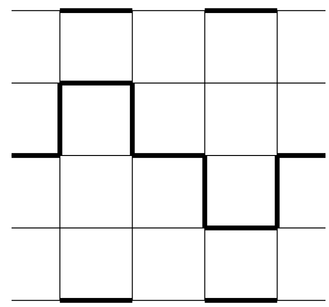

(b)

Figure 1: The mapping between (a) perfect matchings of G and (b) turning graphs of G^*.

To make this more concrete, we designate 4 odd degree vertices on the boundary, which then must be connected by paths in one of the two non-crossing ways; moving between these two classes of configurations requires passing through configurations where the two paths touch (or nearly touch). For this to happen, the paths must be quite long, which is exponentially less likely at equilibrium. One should expect that for appropriate settings of the parameters, it will take exponential time to reach such a configuration, implying that the Markov chain will require exponential time to get close to stationarity. Formalizing this type of intuition is often challenging, however, and this particular problem has been open since proposed by Jim Propp in 1997 [16, 17].

1.1 Weighted models and phase transitions

For many statistical physics models, we see a relationship between the rate of convergence of local Markov chains and an underlying *phase transition* in the physical model itself. For the Ising model, a fundamental model of ferro-magnetism, local algorithms are known to converge in polynomial time (in the diameter of the region) at high temperature, but require exponential time at low

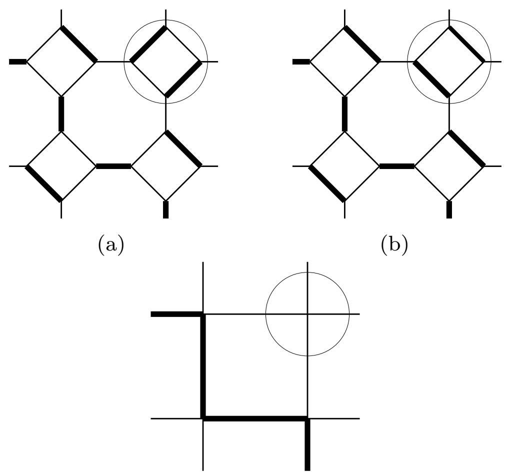

(a) (b) (c)

Figure 2: Two possible orientations (a) and (b) for each free vertex in G^* (c).

temperature [10, 14, 20]. On $\mathbb{Z}^2$, there is a sharp phase transition: there is a critical temperature below which the chain is slowly mixing (requiring exponential time), and at and above which it is rapidly mixing (converging in polynomial time) [14]. A similar behavior is seen for weighted independent sets on $\mathbb{Z}^2$ as we change the "activity" (or "fugacity"), a parameter that controls the expected density of an independent set. Local sampling algorithms for independent sets are known to be rapidly mixing when this parameter is small, favoring sparse independent sets [24], and slow to converge when this parameter is high, favoring denser independent sets [2].

A natural approach to many such problems is to introduce an activity (or weight) in order to understand these phase transitions by learning when a Markov chain converges in polynomial or exponential time. Such an approach was taken recently in the context of triangulations [4] and rectangular dissections and dyadic tilings [3], revealing similar dichotomies. A similar approach was previously considered to study a different, nonlocal Markov chain on sets of perfect and near-perfect matchings on the square-octagon lattice [1], but the behavior of the more natural local dimer-rotating Markov chain studied here remains open.

1.2 Our results

We can now state our results. For simplicity of notation, our terminology throughout the paper is based on weighted turning graphs rather than matchings. Let $G \subset Z^2$ be a finite region on the Cartesian lattice, and let T be the set of turning graphs on G. (i.e., all vertices $v \in G \setminus \partial(G)$ have even degree, and any traversal must "turn" at each vertex.) For input parameters $\lambda > 0$ and $\mu > 0$, we define the distribution as follows. Let $\sigma \in T$ be a turning graph. Then

$$\pi_{\lambda,\mu}(\sigma) = \lambda^{|E(\sigma)|}\mu^{|V(\sigma)|}/Z,$$

where $E(\sigma)$ are the edges in σ, $V(\sigma)$ are the "free" vertices in σ, those that are *not* incident with any edge. $Z = \sum_{\tau \in T} \lambda^{|E(\tau)|}\mu^{|V(\tau)|}$, is a normalizing constant known as the *partition function.*

When $\mu = 1$, we weight configurations $\sigma \in T$ by $\lambda^{|E(\sigma)|}$, favoring shorter contours when $\lambda < 1$ and longer ones when $\lambda > 1$. We show that when $\mu = 1$ and $\lambda < 1/(2\sqrt{e})$ or $\lambda > 2\sqrt{e}$, the Markov chain $\mathcal{M}$ mixes slowly. (A duality in the lattice implies that when $\mu = 1$, if the chain is slow for $\lambda = \lambda^*$ then it is also slow for $\lambda = 1/\lambda^*$.) For $\mu > 1$, we show that if $\lambda < \sqrt{\mu}/2\sqrt{e}$ or $\lambda > 2\mu\sqrt{e}$, the Markov chain $\mathcal{M}$ again mixes slowly.

The proofs that $\mathcal{M}$ is slowly mixing use so-called "Peierls arguments" to identify exponentially small cuts in the state space, implying that the chain will take exponential time to move from one side of the cut to the other. It is fairly simple to show that the chain mixes exponentially slowly when $\lambda > 4$ or $\lambda < 1/4$. We improve this by using a more careful combinatorial analysis, thereby showing slow mixing when $\lambda < 1/[2\sqrt{e}]$ or $\lambda > 2\sqrt{e}$. The proof that $\mathcal{M}$ is rapidly mixing when $\mu = 1$ and $\lambda = 1$ relies on a novel bijection between turning graphs and 3-colorings of the grid.

2 Preliminaries

We start by formalizing our terminology. Let $G_{so} = (V, E)$ be a finite region of the square-octagon lattice, Λ_{so}. We are interested in sampling from the set of random perfect matchings on G_{so}, which we denote $PM(G_{so})$. It is important to distinguish two types of edges of G_{so}, the edges that border both a square and an octagon, which we denote *square* edges, and those that have octagons on both sides, which we refer to as *octagon* edges. Given a perfect matching $\sigma_{so} \in PM(G_{so})$, let $\#N(\sigma_{so})$ be the number of octagon edges. Given an input parameter $\lambda > 0$, we let the weight of σ_{so} be defined as $\lambda^{\#N(\sigma_{so})}/Z$, where $Z = \sum_{\sigma_{so} \in PM(G_{so})} \lambda^{\#N(\sigma_{so})}$ is the normalizing constant, also known as the *partition function.* The local Markov chain $\mathcal{M}_{fl}$, at any initial perfect matching σ_{so}, chooses a face $\in G_{so}$ uniformly at random and attempts to "rotate" the edges of σ_{so} about the face. For example if it chose an octagonal face with all 4 square edges present, the Markov chain would attempt to "rotate" to the configuration with all 4 octagon edges present instead. About a square face, the Markov chain would attempt to rotate between one pair of opposite edges of the square and the other pair.

All of our arguments follow more naturally in the context of paths and cycles, so we now formalize the representation of perfect matchings on Λ_{so} as *turning contours* on $\mathbb{Z}^2$. Let the graph $G = (V', E')$ be the graph formed when each square in G_{so} is contracted to a single vertex, eliminating all self-loops. The graph G is the region of $\mathbb{Z}^2$ formed from the remaining octagon edges of G_{so}. Likewise, the edges of perfect matching $\sigma_{so} \in PM(G_{so})$ maps to configuration σ that corresponds to the remaining octagon edges of σ_{so}.

We observe that in any $\sigma_{so} \in PM(G_{so})$, there must be an even number of octagon edges incident to any interior square. Therefore, the any vertex in the corresponding σ must have even degree. Furthermore, if only two edges are incident to a vertex, they must form a right angle, and cannot be both be vertical or horizontal edges. The vertices on the boundary of G_{so} can have even or odd parity depending depending on which edges of the square are included and excluded. The squares on the boundary therefore can be mapped to vertices with an "even" or "odd" designation that indicates the allowed parity of incident edges. It follows that any valid perfect matching on G_{so} corresponds to edges in G that can be decomposed into cycles and paths that begin and end at an odd-parity vertex on the boundary. These cycles and paths are composed of edges that alternate between horizontal and vertical edges of G, and for this reason we call such configurations *turning graphs* [1].

There is a well-structured many-to-one map between the set of perfect matchings of G_{so} and turning graphs of G. For every $\sigma \in G$ with k vertices with degree zero (i.e., with no incident edges), there are exactly 2^k pre-images $\in G_{so}$. This follows from the fact that each "free" vertex in σ corresponds to a square in G_{so} whose edges can be matched in exactly two ways, independently of all other vertices.

It will be convenient to consider a generalized model of weighted turning graphs on G. Let Ω be the set of all turning graphs σ of G. A free vertex of a turning graph configuration σ is a vertex that is not incident to any edge present in σ. A *free face* of σ is a unit square of G whose four edges are not present in σ, while an *occupied face* of σ is similarly a unit square of G whose four edges are all present in σ. In order to be a useful representation of perfect matchings, a particular turning graph σ should be sampled with probability

$$\pi(\sigma) \ = \ \pi_\lambda(\sigma) \ = \ \frac{2^k \lambda^{|\sigma|}}{Z},$$

where k is the number of free vertices in the graph with σ. Given a turning contour on G sampled according to the prescribed probability distribution, we can easily sample a perfect matching on G_{so} by choosing one of the two orientations of the perfect matching at each free vertex. We generalize this model by introducing a parameter μ, and letting the weight of a configuration

$$\pi(\sigma) \ = \ \pi_{\lambda,\mu}(\sigma) \ = \ (\mu^k \lambda^{|C|})/Z,$$

where

$$Z \ = \ \sum_{\tau \in T} \lambda^{|E(\tau)|} \mu^{|V(\tau)|}.$$

For convenience, we denote this probability model on Ω as $T_{\lambda,\mu}$ of G. The case where $\mu = 2$ corresponds to perfect matchings of the square-octagon lattice.

By setting $\mu = 1$ in this model, we effectively ignore the free vertices and the weight of a configuration is more directly influenced by the underlying geometry of the turning graphs. We show that techniques used to analyze this special case can be extended to the general case of arbitrary μ. A natural Markov chain that has been considered in the context of perfect matchings on the square-octagon lattice iteratively took a square or octagon face and rotated all the edges present if this resulted in a valid configuration. Rotations on square faces did not affect the weight of a configuration, while rotations of an octagonal face could increase or decrease the weight of a configuration multiplicatively by λ^4.

We define the local Markov chain $\mathcal{M}$ on turning graphs Ω, starting at any initial configuration σ_0. The number of steps t required to produce samples sufficiently close to equilibrium will be discussed subsequently.

<u>The Markov chain $\mathcal{M}$</u>

Repeat for t steps:

- Choose a face x of G uniformly at random.
- If x is empty, let σ be the turning path created by adding the edges of face x.
- If x is occupied, let σ be the turning path created by removing the edges of face x.
- With probability $\min(1, \frac{\pi(\sigma)}{\pi(\sigma_t)})$, let $\sigma_{t+1} = \sigma$, and with the remaining probability, let $\sigma_{t+1} = \sigma_t$.

Note that this Markov chain represents precisely the octagon rotating moves of $\mathcal{M}_{\int\wr}$, an ignores the square rotating moves. The fact that the chain $\mathcal{M}$ connects the state space Ω of turning graphs and is aperiodic follows from the ergodicity of $\mathcal{M}_{\int\wr}$ on perfect matchings of the square octagon lattice [16].

For all $\epsilon > 0$, the *mixing time* $\tau(\epsilon)$ of a Markov chain $\mathcal{M}$ is defined as

$$\tau(\epsilon) = \min\{t : \max_{x \in \Omega} \frac{1}{2} \sum_{y \in \Omega} |P^t(x,y) - \pi(y)| \leq \epsilon, \forall\, t' \geq t\}.$$

We say that a Markov chain is *rapidly mixing* (or *polynomially mixing*) if the mixing time is bounded above by a polynomial in n and $\log(\epsilon^{-1})$ and *slowly mixing* if it is bounded below by an exponential function. In Section 3, we bound the mixing time of the Markov chain $\mathcal{M}$ on the turning graph model at various input parameters λ when $\mu = 1$, and in Section 4, we extend these results to the more general model where $\mu > 1$.

3 Mixing of the Markov Chain $\mathcal{M}$ on $T_{\lambda,1}$

We first consider turning graphs when $\mu = 1$, so contours are weighted by the lengths of the contours, independent of the number of free vertices. We will show in Sections 3.1 and 3.2 that $\mathcal{M}$ is slowly mixing on certain graphs by bounding its *conductance*. In Section 3.3 we show when $\mathcal{M}$ is polynomially mixing by reducing to a chain on 3-colorings of the grid.

There is a well known relationship between the conductance of a Markov chain and its mixing time (see, e.g., [9, 21]) that will be the basis of the proofs of slow mixing. For an ergodic Markov chain $\mathcal{M}$ with stationary distribution π, the conductance of a subset $S \subseteq \Omega$ is defined as $\Phi(S) = \sum_{s_1 \in S, s_2 \in \bar{S}} \pi(s_1)P(s_1, s_2)/\pi(S)$. The conductance of the chain $\mathcal{M}$ is then the minimum conductance of all subsets, $\Phi = \min_{S \subset \Omega}\{\Phi(S) : \pi(S) \leq 1/2\}$. The conductance of a Markov chain is related to its mixing time $\tau(\epsilon)$ as follows:

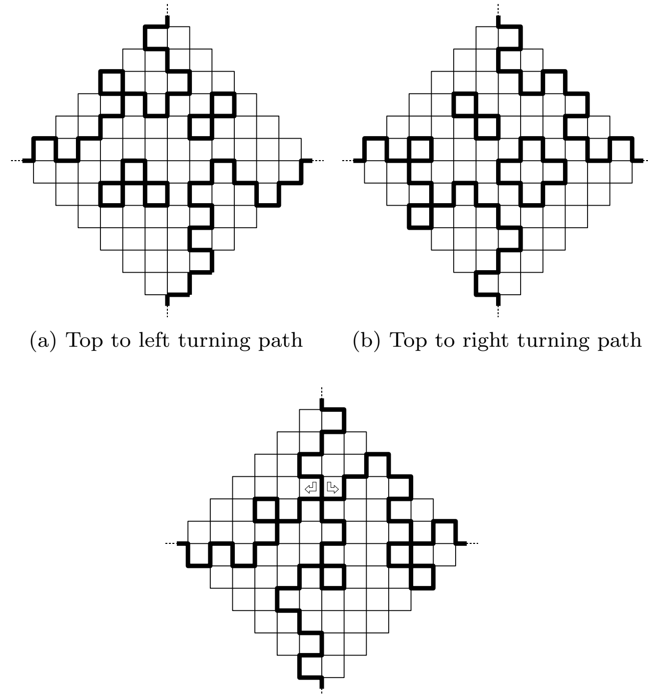
Theorem 3.1. (Jerrum and Sinclair [9]) *The mixing time of a Markov chain with conductance Φ satisfies:*

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

Our strategy in the proofs of slow mixing is to identify three sets, Ω_L, Ω_R and Ω_C, that partition the state space with the middle set Ω_C being a cut that has exponentially smaller weight than the left and right sets Ω_L and Ω_R. From Theorem 3.1 it follows that the conductance of the chain is exponentially small, and therefore we can conclude that $\mathcal{M}$ mixes in time $\Omega(\frac{1}{n}\,(2\lambda\sqrt{e})^{-4n})$.

3.1 Slow mixing of $\mathcal{M}$ on $T_{\lambda,1}$ for $\lambda < 1/2\sqrt{e}$

We start by showing that when $\lambda < 1/2\sqrt{e}$, the Markov Chain $\mathcal{M}$ mixes slowly. Specifically, we show slow mixing on the so called "Aztec Diamond" graph G [6]. Starting with G_n, the standard Aztec Diamond graph of order n, we add extra edges E_B to the four corners of the graph, as shown in Fig 3.1. Because a perfect matching in the original G_{so} must include those edges, and this serves to set the "parity" of the boundary vertices. For convenience, we do not consider the extra edges E_B as part of G_n, and treat them as boundary conditions on the graph G_n. We see that with the given boundary conditions, we see that these four corners are the only



(a) Top to left turning path (b) Top to right turning path

(c) Crossing pair of paths

Figure 3: Configurations in (a) Ω_L, (b) Ω_R, and (c) Ω_C.

vertices with odd parity in G_n. Therefore, in any turning graph of G_n, there must be either a turning path from the top vertex to the left vertex and a turning path from the bottom vertex to the right, or vice-versa.

We define Ω_L to be the set of configurations where these two paths do not cross each other and the top vertex has a path to the left vertex. Similarly, we let Ω_R be the set of configurations where these paths do not cross and the top vertex connects to the right vertex. The cut Ω_C consists of all other states where these two paths do cross (or touch), and in order to pass from configurations in Ω_L to Ω_R, the Markov chain $\mathcal{M}$ must pass through a crossing configuration in Ω_C [1]. For any $\sigma \in \Omega_C$, we decompose the edges into a *crossing pair of paths*, the union of a top-left turning path and a bottom-right turning path that share a vertex, which we call a *crossing vertex*. For any crossing pair of paths, we can uniquely identify the lexicographically first crossing vertex as a special vertex. Note that a crossing pair of paths can simultaneous be interpreted as the union of a top-right and bottom-left path that also pass through the same crossing vertex. These paths partition G into four regions, one for each diagonal boundary; we will refer to these regions by the side of G that they border.

Theorem 3.2. *When $\lambda < 1/(2\sqrt{e})$, the Markov*

Chain $\mathcal{M}$ on $T_{\lambda,1}$, weighted turning graphs of the Aztec Diamond G_n, is at least

$$\tau(\epsilon) \;\geq\; n\left(2\lambda\sqrt{e}\right)^{-4n}\ln\epsilon^{-1}.$$

Proof. Our goal is to show that Ω_C is an exponentially small cut in our state space by exhibiting a mapping $\phi_r : \Omega_C \to \Omega$ such that for any $\sigma \in \Omega_C$, the weight of the image $\pi(\phi_r(\sigma))$ is exponentially larger in n than $\pi(\sigma)$. We construct $\phi_r(\sigma)$ for $\sigma \in \Omega_C$ as follows (see Fig. 4). Given a state $\sigma \in \Omega_C$, take a maximal pair of crossing paths in σ. We then remove the edges of this crossing path from G. We then shift all edges in σ from the bottom left region up by one and all edges from the top right region down by one. Finally, we add in edges along the bottom left and top right boundaries of G to form a valid turning graph.

We first partition Ω_C into sets $\Omega_{C,h,v}$ for $h, v \geq 0$ as follows. Given $\sigma \in \Omega_C$, consider the lexicographically first pair of crossing paths in σ. We separate these into "top-left" and "bottom right" turning paths that meet at their lexicographically first crossing point x. We define the "horizontal path" as the sub path of the top-left path from the leftmost vertex to x, concatenated with the sub path of the bottom-right path from x to the rightmost vertex. We similarly define the "vertical path" from the topmost vertex to x to the bottommost vertex passing through x. This horizontal path, when viewed as a path from the left vertex to the right vertex, contains some number $h \geq 0$ "backwards" edges from right to left. Similarly the "vertical path" has some $v \geq 0$ backwards edges from bottom to top. We say that $\sigma \in \Omega_{C,h,v}$.

Note that since the horizontal path ends exactly $2n$ edges to the right of its origin, it must contain exactly $2n+2h$ total horizontal edges. Since it is the union of two turning paths, which alternate horizontal and vertical edges, the number of vertical edges in the horizontal path must be $2n + 2h + \delta_h$, where $|\delta_h| \leq 2$. Similarly, the vertical path has exactly $2n + 2v$ vertical edges and $2n + 2v + \delta_v$ horizontal edges, for some $|\delta_v| \leq 2$. We will see that the values of δ_h and δ_v will only affect our final bounds by a constant, and we may safely ignore them for convenience. For both the horizontal and vertical paths, we may encode the entire path essentially as two separate interleaved bit sequences, one for the horizontal moves, another for the vertical moves, and a single special symbol x to indicate the location of the crossing.

To bound the number of pre-images of this map, we note the number of left-right paths of type h is at most

$$n\binom{2n+2h}{h}\binom{2n+2h}{n+h}.$$

Similarly, the number of top-down paths of type v is at most

$$n\binom{2n+2v}{v}\binom{2n+2v}{n+v}.$$

Then $P(h,v)$, the total number of paths of type (h,v), is therefore at most

$$P(h,v) \;=$$

$$n^2\binom{2n+2h}{h}\binom{2n+2h}{n+h}\binom{2n+2v}{v}\binom{2n+2v}{n+v}$$

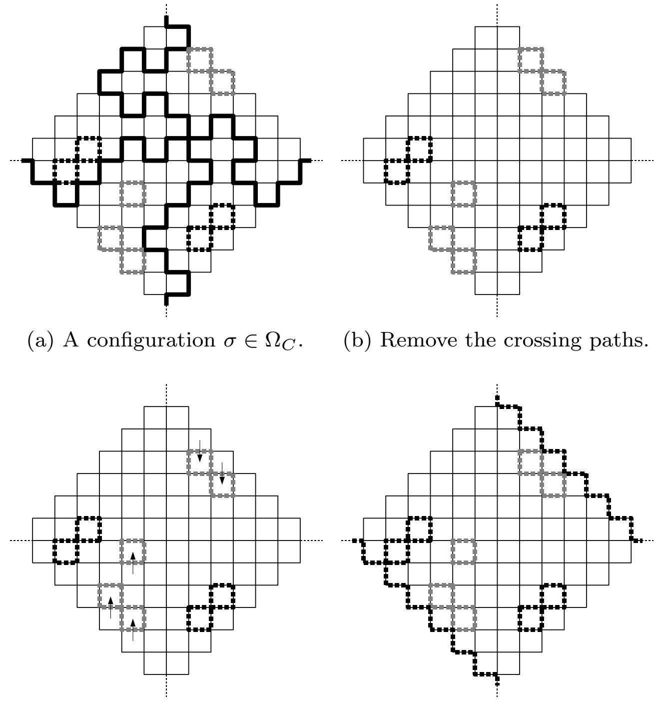
$$< \;\frac{n^2 2^{4n+2h+2v}}{\pi\sqrt{n+h}\sqrt{n+v}}\binom{2n+2h}{h}\binom{2n+2v}{v}$$

$$< \;\frac{n^2 2^{4n+2h+2v}}{\pi\sqrt{n+h}\sqrt{n+v}}\frac{(2n+2h)^h}{h!}\frac{(2n+2v)^v}{v!}$$

$$< \;\frac{n^2 2^{4n+2h+2v}}{\pi\sqrt{hv}\sqrt{n+h}\sqrt{n+v}}\left(\frac{e(2n+2h)}{h}\right)^h\left(\frac{e(2n+2v)}{v}\right)^v$$

$$< \;\frac{n^2 2^{4n+2h+2v}(2e)^{h+v}}{\pi\sqrt{hv}\sqrt{n+h}\sqrt{n+v}}\left(1+\frac{n}{h}\right)^h\left(1+\frac{n}{v}\right)^v$$

$$< \;\frac{n^2 2^{4n+2h+2v}(2e)^{h+v}(e)^{2n}}{\pi\sqrt{hv}\sqrt{n+h}\sqrt{n+v}},$$



(a) A configuration $\sigma \in \Omega_C$. (b) Remove the crossing paths.

(c) Shift away from boundary. (d) Add the boundary to get $\phi(\sigma)$.

Figure 4: The mapping $\phi : \Omega_C \to \Omega$.

where the first inequality follows from Stirling's formula and the last uses the well known fact that $\left((1+a)^b \leq e^{ab}\right)$ for all positive b.

Since we go from a configuration with $8n+4v+4h$ total edges in the crossing path to a configuration with exactly $4n$ new edges, it follows that for all $\sigma \in \Omega_{C,h,v}$, the gain in weight $\pi(\phi_r(\sigma))/\pi(\sigma) = \lambda^{-(4n+4h+4v)}$. Summing over all possible $0 \leq h, v \leq n^2$, we conclude:

$$\begin{aligned}
\pi(\Omega_C) &= \sum_{h,v} \pi(\Omega_{C,h,v})) \\
&\leq \sum_{h,v} \sum_{\sigma \in \Omega_{C,h,v})} \pi(\phi(\sigma)) \frac{\pi(\sigma)}{\pi(\phi(\sigma))} \\
&\leq \sum_{h,v} \sum_{\sigma \in \Omega_{C,h,v})} \pi(\phi(\sigma)) \lambda^{(4n+4h+4v)} \\
&\leq \sum_{h,v} \lambda^{(4n+4h+4v)} \frac{n^2 2^{4n+2h+2v} (2e)^{h+v} (e)^{2n}}{\pi\sqrt{hv}\sqrt{n+h}\sqrt{n+v}} \\
&\leq \sum_{h,v} \frac{n^2}{\pi\sqrt{hv}\sqrt{n+h}\sqrt{n+v}} \frac{(2\lambda\sqrt{e})^{4n+4h+4v}}{(2e)^{h+v}} \\
&\leq n\left(2\lambda\sqrt{e}\right)^{4n}.
\end{aligned}$$

For any constant $\lambda < 1/2\sqrt{e}$, we see that $\pi(\Omega_C)$ is exponentially small in n. We conclude that the conductance $\Phi_{\mathcal{M}}$ of the Markov chain $\mathcal{M}$ must be bounded by

$$\begin{aligned}
\Phi_{\mathcal{M}} &\leq \sum_{s_1 \in \Omega_R, s_2 \in \overline{\Omega_R}} \pi(s_1) P(s_1, s_2)/\pi(\Omega_R) \\
&\leq \pi(\Omega_C)/\pi(\Omega_R) \\
&= 2\pi(\Omega_C)/(1-\pi(\Omega_C)) \\
&\leq 2n\left(2\lambda\sqrt{e}\right)^{4n}.
\end{aligned}$$

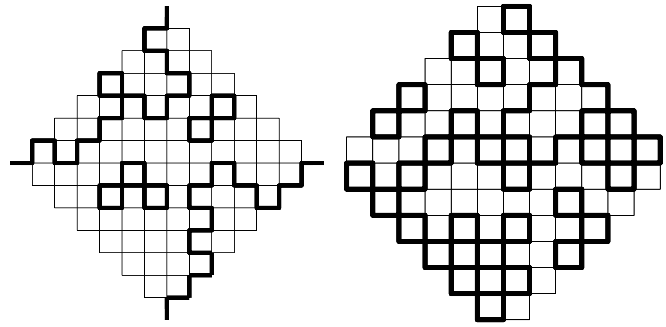
By Theorem 3.1, it follows that $\tau(\epsilon)$, the mixing time of $\mathcal{M}$, satisfies

$$\tau(\epsilon) \;\geq\; \frac{1}{4n}\left(2\lambda\sqrt{e}\right)^{-4n} \ln \epsilon^{-1},$$

which is exponentially large in n.

3.2 Slow mixing of $\mathcal{M}$ on $T_{\lambda,1}$ for $\lambda > 2\sqrt{e}$

Next, we show that when each edge present in the turning contour is given weight at least $\lambda > 2\sqrt{e}$, the Markov Chain $\mathcal{M}$ also mixes slowly on some graphs. Rather than prove this case directly, we exhibit a bijection between the model on graph G for any $\lambda < 1$, and the complimentary model on $G' = G$ with altered boundary conditions for $\lambda' = 1/\lambda > 1$.



(a) G, λ (b) $G', \lambda' = 1/\lambda$

Figure 5: Weight-preserving bijection between $\sigma \subset G$ at parameter λ and $\sigma' \subset G'$ at $\lambda' = 1/\lambda$.

For each vertex v in G with parity $p(v)$, set the parity of that vertex to $deg(v) - p(v)$ in G'. It follows that the complementary turning graph $C' = E \backslash C$ on G' will be a valid turning graph that, by construction, will satisfy the parity boundary conditions of G', as a vertex v with k incident edges in G corresponds to a vertex with $deg(v) - k$ incident edges in E'. Let G be the Aztec diamond graph described in the previous section, and let G' be the Aztec Diamond graph with boundary conditions modified as in Fig. 5.

Corollary 3.1. *When $\lambda > 2\sqrt{e}$, the mixing time of the Markov Chain $\mathcal{M}$ on $T_{\lambda,1}$, weighted turning graphs of the Aztec Diamond G_n, is at least*

$$\tau(\epsilon) \;\geq\; n\left(2\lambda\sqrt{e}\right)^{-4n} \ln \epsilon^{-1}.$$

Proof. We show that the missing edges in this model G' behave exactly like the present edges in G, and will form turning paths of missing edges between vertices that have difference in parity between its degree and other parity requirement. It follows then that the unnormalized weight of a turning graph C' with parameter $\lambda' = 1/\lambda$ is exactly

$$G' = {\lambda'}^{|C'|} = {\lambda'}^{|E|-|C|} = {\lambda'}^{|E|}{\lambda'}^{-|C|} = {\lambda'}^{|E|}\lambda^{|C|}.$$

Since this is exactly the weight of the corresponding turning path C of G multiplied by $\lambda^{|E|}$, it follows that the normalization

$$Z' \;=\; \sum_{C' \in G'} {\lambda'}^{|C'|} = \lambda^{|E|} Z.$$

Thus, the normalized probability $\pi(C') = \pi(C)$.

The Markov chain $\mathcal{M}$ behaves exactly the same on both models, by adding or removing edges with probabilities depending on the relative weights of the current and proposed next state. Thus $\mathcal{M}$ on G with parameter λ behaves exactly the same as $\mathcal{M}$ on G' with parameter $1/\lambda$. The corollary then follows immediately from Theorem 3.2.

3.3 Polynomial mixing of $\mathcal{M}$ on $T_{\lambda,1}$ for $\lambda = 1$

We start by describing a novel bijection between turning paths of a region of the grid $G = (V, E)$ subject to given boundary conditions with three-colorings of G subject to certain corresponding boundary conditions. We then infer that the Markov chain $\mathcal{M}$ is polynomially mixing when $\lambda = 1$ on any region of the grid G from the fact that three colorings on finite regions of $\mathbb{Z}^2$ with fixed boundary conditions are polynomially mixing on any such G [1].

THEOREM 3.3. *Given vertex boundary conditions on the grid G, we can construct coloring boundary conditions on G such that there is a bijection between the Turning graphs on G satisfying the given vertex boundaries and the set of 3-colorings satisfying the corresponding coloring boundaries.*

Proof. Say we are given an instance of the turning path model G with given parities $p(v)$ on each vertex on the boundary. We describe an assignment of faces on the external boundary of G that corresponds with a given $p(v)$. The vertices for which $p(v)$ is odd are starting points for turning paths, and divide the boundary into an even number of regions. We say that region R_1 and region R_2 share a border if the corresponding faces both share a vertex v such that $p(v)$ is odd. Each region of the boundary will be colored by exactly two colors in alternating fashion. We begin by fixing a face f_s of the external boundary of G and assigning it color 1.

Given σ that satisfies the above turning path boundary conditions, we show how to construct $f(\sigma)$, a 3-coloring of G that satisfies the corresponding coloring boundary conditions.

- Beginning at face f_s, we color every face in its region with the same parity as f_s with color 1, and all other faces in this region with color 2.
- In general, say without loss of generality that a region R_1 is colored a, b, and the third color not present in the region is c. If R_1 borders R_2 with a face colored a, then we color R_2 with colors c, b. Otherwise, we color R_2 with colors c, a. After this, we have colored every face of G, and we call the resulting 3-coloring is $f(\sigma)$.

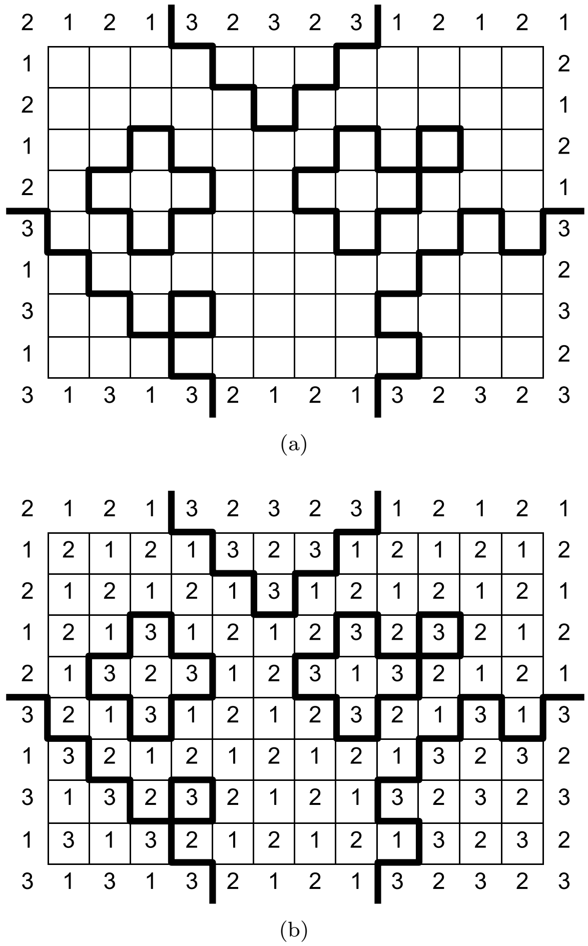


Figure 6: Coloring representations of boundary conditions and turning graphs.

For any given turning path σ on G, we construct the corresponding 3-coloring similarly, using the turning path itself as the boundary between adjacent regions. This procedure is well-defined due to the turning property of the paths in σ - all vertices in a single region that are directly adjacent to any specific turning path must have the same parity in $\mathbb{Z}^2$, and thus will have the same color. By construction, the colors chosen on the boundary form a valid 3-coloring of the tiles of the boundary.

We now show that f is a bijection by constructing the inverse f^{-1}. Beginning at face f_s, we take the maximally connected region of faces that are colored 1 or 2. On the boundary, this will include all faces on both sides of f_s up to any external boundary vertex colored 3. We claim that all faces on the boundary of this maximal region must have the same color - this follows immediately from the fact that by construction, all faces on the boundary of this region must border a face colored 3. Since none of these faces colored 3 can be adjacent, it follows that none of the faces bordering these 3-faces on "our side" of the boundary can be adjacent. Thus, all such faces must be diagonal to each other, and will therefore have the same parity in $\mathbb{Z}^2$ and be colored the same. Thus the edges on the border of this region must be a turning path.

Since all steps above are done deterministically, it follows that this procedure is well defined for all 3-colorings. There is a known bijection between 3-colorings of the grid and height functions [1], where the height at two neighboring faces may only differ by $\{-1, 1\}$. We observe that these turning paths correspond exactly to alternating level curves of the height function. That is, the edges in the turning graph are exactly the set of edges between the vertices at height $2k$ and $2k+1$ for all integer k. The uniqueness of the height representation of a 3-coloring [1] implies that this reverse map f^{-1} must be injective. Therefore, f must be a bijection.

Corollary 3.2. *The Markov chain $\mathcal{M}$ when $\lambda = \mu = 1$ is polynomially mixing on finite grid regions G.*

Proof. The Markov chain $\mathcal{M}$ that adds or removes a single square of edges corresponds to the local Markov chain on 3-colorings that changes the color at a single square at a time. This chain was shown to be polynomially mixing on all subsets of $\mathbb{Z}^2$ with any fixed boundary conditions by a coupling argument [1].

4 Mixing of $\mathcal{M}$ on $T_{\lambda,\mu}$ for general $\mu > 1$

We extend our analysis of the special case when $\mu = 1$ to the general model for any $\mu > 1$ by considering an amortized "cost" for each non-free vertex in σ, distributed among its incident edges.

Theorem 4.1. *When $\lambda < \sqrt{\mu}/(2\sqrt{e})$, the mixing time of the Markov Chain $\mathcal{M}$ on $T_{\lambda,\mu}$, weighted turning graphs of the Aztec Diamond G_n, is at least*

$$\tau(\epsilon) \;\geq\; n \left(2\lambda\sqrt{e}\right)^{-4n} \ln \epsilon^{-1}.$$

Proof. To handle the case where $\mu > 1$, we need to consider the change in the number of vertices used by the turning graph. We follow the structure of Theorem 3.2, keeping both the structure of the proof and the map ϕ.

Let σ be a configuration in $\Omega_{C,v,h}$. As in Theorem 3.2, we see that σ has $8n + 4h + 4v$ edges in some pair of crossing paths. It follows that the sum of all degrees of all vertices incident to these edges must add to $16n + 8h + 8v$. This pair of crossing paths includes at least the topmost and bottommost vertex at each x coordinate, and thus must contain at least $4n$ vertices of degree 2. The degrees of the remaining vertices therefore sum to $8n + 8h + 8v$. Since the maximum degree of any vertex is 4, there must be at least $2n + 2h + 2v$ other vertices used by this pair of crossing paths.

The map $\phi(\sigma)$ removes this pair of crossing paths, and adds two paths of exactly $4n$ edges and $4n$ vertices. Thus, we have a net gain of at least $2n+2h+2v$ vertices between σ and $\phi(\sigma)$. Thus, the change in weight for any $\sigma \in \Omega_{C,h,v}$ will be

$$\begin{aligned} \pi(\phi_r(\sigma))/\pi(\sigma) \;&\geq\; \mu^{2n+2h+2v}\lambda^{-(4n+4h+4v)} \\ &=\; (\lambda/\sqrt{\mu})^{-(4n+4h+4v)}. \end{aligned}$$

As in Theorem 3.2, this directly implies that $\mathcal{M}$ mixes slowly when $\lambda/\sqrt{\mu} < 1/2\sqrt{e}$, or more simply $\lambda < \sqrt{\mu}/2\sqrt{e}$.

We now analyze the case where $\lambda > 1$ similarly, and obtain a result analogous to Corollary 3.1 for this more general case. Following the bijection in Corollary 3.1 that maps a turning graph in G with the complementary graph in G', we could immediately conclude from Theorem 4.1 that for any $\mu < 1$, $\mathcal{M}$ is slowly mixing whenever $\lambda > 2\sqrt{e}/\sqrt{\mu}$. However, we are chiefly interested in the case when $\mu > 1$, especially when $\mu = 2$. In this case, we can reason directly from Corollary 3.1.

Theorem 4.2. *When $\lambda > 2\mu\sqrt{e}$, the mixing time of the Markov Chain $\mathcal{M}$ on $T_{\lambda,\mu}$, weighted turning graphs of the Aztec Diamond G_n, is at least*

$$\tau(\epsilon) \;\geq\; n \left(2\lambda\sqrt{e}\right)^{-4n} \ln \epsilon^{-1}.$$

Proof. We proceed similarly to the proof of Theorem 4.1, but with one important difference. By the nature of the bijection, the mapping ϕ in this context doesn't remove edges and add shorter ones, it removes *non edges*, and adds a shorter path of unchosen edges, potentially increasing the total number of chosen vertices in the process.

However, as in the argument of Theorem 4.1, the sum of the degrees of vertices incident to these edges, other than the boundaries, adds to $8n + 8h + 8v$. It follows then that *at most* $4n + 4h + 4v$ vertices will be added by the map ϕ in the complementary context.

Thus, as before, the change in weight for any $\sigma \in \Omega_{C,h,v}$ is

$$\begin{aligned} \pi(\phi_r(\sigma))/\pi(\sigma) \;&\geq\; \mu^{2n+2h+2v}\lambda^{-(4n+4h+4v)} \\ &=\; (\lambda/\mu)^{-(4n+4h+4v)}. \end{aligned}$$

Treating μ as a constant, by Corollary 3.1 we conclude that $\mathcal{M}$ mixes slowly whenever $(\lambda/\mu) > 2\sqrt{e}$, or when $\lambda > 2\mu\sqrt{e}$.

5 Conclusions

Our arguments verify that, in the weighted setting, the Markov chain on perfect matchings of the square-octagon lattice can be either fast or slow to converge to equilibrium, suggesting the presence of a phase transition. An important facet of our results is the strategy of adapting a slow mixing result for the simpler setting when $\mu = 1$ to arbitrary μ. Note that if we could show that $\mathcal{M}$ is slowly mixing when $\mu = 1$, and $\lambda < 1/\sqrt{2} + \varepsilon$, then this would imply that $\mathcal{M}$ is slowly mixing when $\mu = 2$ and $\lambda = 1$. This setting of the parameters precisely corresponds to uniformly sampling perfect matchings on regions of the square-octagon lattice. Our specific methods do not seem to be sufficient for extending the proofs to this setting. Hence it is still open whether the chain converges slowly when on the set of unweighted perfect matchings, as simulations suggest. We believe that a more refined analysis of the tradeoffs between the "energy" (the weighting that discourages long turning paths) and the "entropy" (bounding the number of configurations with turning paths of different lengths) will be the key to extending these arguments.

Acknowledgments

The authors thank Jim Propp for introducing this problem and for several useful discussions.

References

[1] Bhatnagar, N., Greenberg, S., Randall, D.: The effect of boundary conditions on mixing rates of markov chains. Proceedings of APPROX/RANDOM pp. 280–291 (2006)

[2] Blanca, A., Galvin, D., Randall, D., Tetali, P.: Phase coexistence and slow mixing for the hard-core model on $\mathbb{Z}^2$. Proceedings of APPROX/RANDOM pp. 379–394 (2013)

[3] Cannon, S., Miracle, S., Randall, D.: Phase transitions in random dyadic tilings and rectangular dissections. Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms pp. 1573–1589 (2015)

[4] Caputo, P., Martinelli, F., Sinclair, A.J., Stauffer, A.: Random lattice triangulations: Structure and algorithms. In: Proceedings of the Forty-fifth annual ACM Symposium on Theory of Computing. pp. 615–624 (2013)

[5] Edmonds, J.: Paths, trees, and flowers. Canad. J. Math. 17, 449–467 (1965), `www.cs.berkeley.edu/~christos/classics/edmonds.ps`

[6] Elkies, N., Kuperberg, G., Larsen, M., Propp, J.G.: Alternating-sign matrices and domino tilings. Journal of Algebraic Combinatorics 2, 111–132 (1992)

[7] Gessel, I., Viennot, X.: Binomials determinants, paths and hook length formulae. Advances in Mathematics 58, 300–321 (1985)

[8] Greenberg, S., Randall, D.: Slow mixing of markov chains using fault lines and fat contours. Algorithmica 58, 911–927 (2010)

[9] Jerrum, M.R., Sinclair, A.J.: Approximate counting, uniform generation and rapidly mixing markov chains. Information and Computation 82, 93–133 (1989)

[10] Jerrum, M.R., Sinclair, A.J.: Polynomial-time approximation algorithms for the Ising model. SIAM Journal on Computing (1993)

[11] Jerrum, M.R., Sinclair, A.J., Vigoda, E.: A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. Journal of the ACM 41, 671–697 (2006)

[12] Kasteleyn, P.W.: The statistics of dimers on a lattice : I. The number of dimer arrangements on a quadratic lattice. Physica pp. 1209–1225 (1961)

[13] Lindström, B.: On the vector representations of induced matroids. Bulletin of the London Mathematical Society 5, 85–90 (1973)

[14] Lubetzy, S., Sly, A.: Critical Ising on the square lattice mixes in polynomial time. Communications in Mathematical Physics pp. 815–836 (2012)

[15] Luby, M., Randall, D., Sinclair, A.J.: Markov chain algorithms for planar lattice structures. SIAM Journal on Computing 31, 167–192 (2001)

[16] Propp, J.G.: Diabolo tilings of fortresses. http://www.math.wisc.edu/ propp/diabolo.ps.gz

[17] Propp, J.G.: Mixing time for dimers on the square-octagon graph. http://mathoverflow.net/questions/168637/mixing-time-for-dimers-on-the-square-octagon-graph

[18] Propp, J.G., Wilson, D.B.: Exact sampling with coupled markov chains and applications to statistical mechanics. Random Structures and Algorithms 9(1-2), 223–252 (1996)

[19] Randall, D., Tetali, P.: Analyzing glauber dynamics by comparison of Markov chains. Journal of Mathematical Physics 41, 1598–1615 (2000)

[20] Randall, D., Wilson, D.B.: Sampling spin configurations of an Ising system. In: Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms. pp. 959–960. SODA (1999)

[21] Sinclair, A.J.: Algorithms for random generation and counting. Progress in theoretical computer science, Birkhäuser (1993)

[22] Temperley, H.N.V., Fisher, M.E.: Dimer problem in statistical mechanics-an exact result. Philosophical Magazine (1961)

[23] Valiant, L.: The complexity of computing the permanent. Theoretical Computer Science pp. 189 – 201 (1979)

[24] Vera, J., Vigoda, E., Yang, L.: Improved bounds on the phase transition for the hard-core model in 2-dimensions. Proceedings of APPROX/RANDOMs pp. 699–713 (2013)

[25] Wilson, D.B.: Mixing times of lozenge tiling and card shuffling markov chains. The Annals of Applied Probability 1, 274–325 (2004)