Sampling weighted perfect matchings on the square-octagon lattice

Prateek Bhakta, Dana Randall

PII: S0304-3975(17)30063-4
DOI: http://dx.doi.org/10.1016/j.tcs.2017.01.014
Reference: TCS 11046

To appear in: *Theoretical Computer Science*

Received date: 15 June 2016
Revised date: 2 January 2017
Accepted date: 16 January 2017

Sampling Weighted Perfect Matchings on the Square-Octagon Lattice

Prateek Bhakta * Dana Randall †

Abstract

Perfect matchings of the square-octagon lattice, also known as "fortresses" [19], have been shown to have a rich combinatorial structure. We are interested in a natural local Markov chain for sampling from the set of perfect matchings that is known to be ergodic and has been used in practice to discover properties of random fortresses. However, unlike related Markov chains used for sampling domino and lozenge tilings, this Markov chain on the square-octagon lattice appears to converge slowly. To understand why, we introduce a weighted version of the chain and prove that this chain can converge in polynomial time or exponential time depending on the settings of the parameters.

Keywords: Markov Chains, Sampling, Perfect Matchings, Probability, Fortress Model

1 Introduction

Perfect matchings arise in many natural computational contexts, and have been the cornerstone problem underlying many fundamental complexity questions. They are also of specific interest to the statistical physics community, where they are studied in the context of *dimer models*. Here, edges in a matching represent diatomic molecules, or dimers, and perfect matchings of a lattice region correspond to dimer packings. Physicists study the properties of these physical systems by relating fundamental thermodynamic quantities to weighted sums over the set of all configurations of the system, in our case the set of all perfect matchings of the lattice region.

The seminal work of Edmonds established that the decision and construction problems, i.e. efficiently deciding if a given graph has a perfect matching and finding it if so, were in P [6]. Subsequently, Valiant showed that counting perfect matchings is #P-complete, so it is believed that there is no such polynomial time general solution [26]. As a consequence, there has been a great deal of interest in finding both efficient *approximate* counting algorithms, as well as efficient exact counting algorithms in restricted settings. Jerrum, Sinclair and Vigoda showed how to approximately count and sample perfect matchings in any bipartite graph efficiently, although the complexity remains open on general graphs [13]. Alternatively, in 1969, Kasteleyn et al. developed a robust method to exactly count perfect matchings on any planar graph in polynomial time by calculating a Pfaffian on a directed version of the adjacency matrix [14, 25]. In fact, when the underlying graph is a lattice region, determinant-based methods for counting matchings have been shown to be even more efficient [8, 15].

Matchings on lattices arise naturally as well. For example, on finite regions of the hexagonal lattice, perfect matchings correspond to lozenge tilings of the dual region, and on finite regions of

*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.

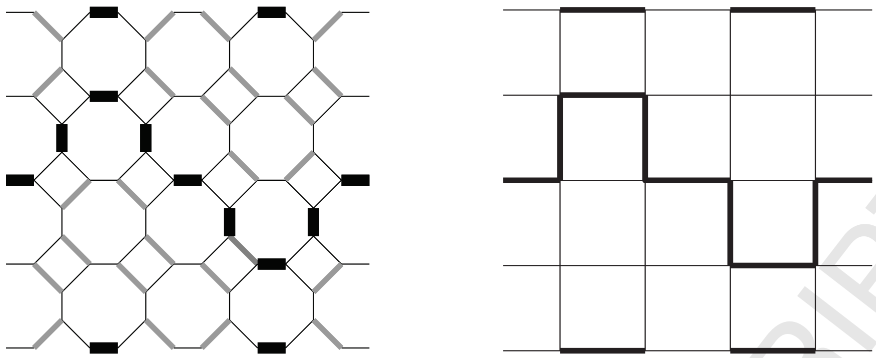


Figure 1: The mapping between (a) perfect matchings of G and (b) turning graphs of G^*.

the Cartesian lattice $\mathbb{Z}^2$, they correspond to domino tilings of the dual. A common Markov chain approach for sampling perfect matchings on these lattices that is popular among experimentalists is based on *rotations*. Specifically, on $\mathbb{Z}^2$ the Markov chain evolves by choosing a unit face uniformly, and if this face contains two edges of the matching on opposite sides, the Markov chain replaces those edges with the other two opposing edges on the face with some probability. A similar approach on the hexagonal lattice replaces three alternating edges around a hexagonal face with the complement set of alternating edges with some probability.

This Markov chain based on dimer rotations was first studied by Propp and Wilson [21]. They innovated their "coupling-from-the-past" algorithm for this problem, and showed that it could be run on dimer covers of the Cartesian lattice $\mathbb{Z}^2$ to generate perfectly uniform samples of perfect matchings, although there were no guarantees that the algorithm would terminate in expected polynomial time. The proof that the expected time to converge is polynomially bounded was provided by Luby et al. [17], Randall and Tetali [22] and further improved by Wilson [28]. Coupling-from-the-past has subsequently been used to study matchings on many other lattices as well, providing perfectly random samples, although not always efficiently. This paradigm gave rise to many conjectures about the 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" [20]. Many remarkable properties of lozenge and domino tilings, such as the existence of frozen regions at equilibrium, are known to hold for fortresses [20].

There is a natural analogue to the dimer-rotating Markov chain for fortress graphs, that 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 this chain may in fact require exponential time on the square-octagon lattice to converge. As an informal explanation of the motivation for why the convergence time of this Markov chain is likely exponential, it is useful to interpret these perfect matchings as "contours" [1, 10, 16]. In particular, given any perfect matching on a simply connected region of the square-octagon lattice, we first contract the four vertices of each square into a single vertex, leaving only the edges bordering 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 Figure 1).

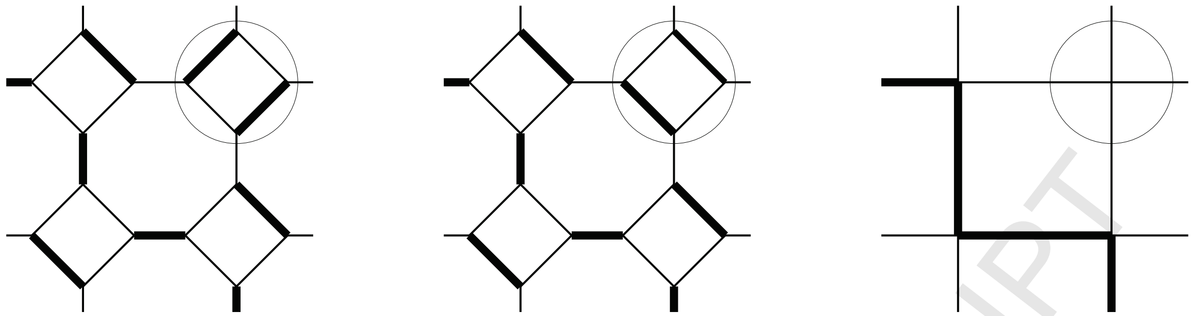


Figure 2: Two possible orientations (a) and (b) for each free vertex in G^* (c).

If we decompose these sets of edges by pairing non-crossing adjacent edges at the degree 4 vertices, we therefore get a collection of "turning paths," that terminate at odd-degree 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 Figure 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 each turning graph σ has weight proportional to $2^{V(\sigma)}$, where $V(\sigma)$ is the number of free vertices in σ. This tells us 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 transformation, as it allows us to use analysis similar to other models in statistical physics, most notably the Ising model of ferromagnetism, that are slowly mixing when long contours are similarly disfavored [16].

As with 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 temperature [12, 16, 23]. 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) [16]. 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 [27], and slow to converge when this parameter is high, favoring denser independent sets [2].

To understand the interaction of these turning paths, we introduce a more general weighted version of the model to expose a phase transition in the mixing time. Such an approach was taken recently in the context of triangulations [5] 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.

We focus our attention on a modified version of the Aztec Diamond graph, which we call the "Decorated Aztec Diamond." The Decorated Aztec Diamond has 4 additional edges attached to the corners of diamond, which cause the corner vertices to have an odd number of incident edges. These four boundary vertices must be connected by a pair of turning paths in one of the two non-crossing ways, and moving between these two classes of configurations requires passing through configurations where the two paths touch. We will show that for this to happen, the paths must be quite long, which is exponentially less likely at equilibrium for appropriate settings of the parameters. We then argue that because it will take exponential time to reach such a configuration, the Markov chain will require exponential time to converge. Formalizing this type of intuition is often challenging, however, and this particular problem has been open since proposed by Jim Propp in 1997 [19, 20].

We now state our results. For simplicity of notation, our terminology throughout the paper is based on weighted turning graphs rather than matchings. Let $G \subset \mathbb{Z}^2$ 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.

2 Preliminaries

We begin by formalizing our model. Let $\Lambda_{so} = (V, E)$ be the infinite square-octagon lattice, and let G_{so} be a finite, simply connected region of Λ_{so}. We are interested in randomly sampling from the set of perfect matchings on G_{so}, which we denote $\Omega_{SO} = PM(G_{so})$. Starting at any perfect matching σ_{so}, the local Markov chain $\mathcal{M}_{so}$ first chooses a face $f \subseteq E(G_{so})$ uniformly from the interior faces of G_{so}. If the edges of the perfect matching σ_{so} restricted to the face f alternate around f, the Markov chain can transition to the configuration $\sigma'_{so} = (\sigma_{so} \oplus f)$, where $\oplus$ represents the symmetric difference operator (i.e., the chain can complement the edges on a single face f if it forms a perfect matching when restricted to f). We call this a "rotation" of the edges about the face.

It will be convenient to introduce an alternate representation of perfect matchings on the square-octagon lattice as *turning contours* on $\mathbb{Z}^2$. We distinguish two types of edges of G_{so}, the edges that border both a square and an octagon, which we call *square* edges, and those that have octagons on both sides, which we call *octagon* edges. Consider the map from Λ_{so} to $\mathbb{Z}^2$ that contracts every square into a single vertex, eliminating self loops. This map eliminates all square edges in Λ_{so}, and leaves the octagon edges in Λ_{so} as edges of $\mathbb{Z}^2$. Let G_C be the image of G_{so} under this map, and let $\sigma \subset E(G_C)$ be the image of perfect matching $\sigma_{so} \in PM(G_{so})$. G_C is a simply connected region of $\mathbb{Z}^2$, formed from the octagon edges of G_{so}, and we define σ_C to be the corresponding octagon edges of σ_{so}.

We observe that for $\sigma_{so} \in PM(G_{so})$, there must be an even number of octagon edges incident to any interior square, and that if there are only two incident octagon edges, then they are not diametrically opposite. This means that σ_C must have an even number of edges incident to every interior vertex, and if only two edges are incident to a vertex, they must form a right angle, or "turn." The squares on the boundary of G_{so} can have an even or odd number of outgoing edges, and are therefore mapped to vertices in G_C that retain an "even" or "odd" designation that indicates the parity of incident edges that any configuration σ_C may have. It follows that any valid perfect matching σ_{so} on G_{so} maps to a σ_C that can be decomposed into cycles and paths that begin and end at odd-parity vertices on the boundary. Since our paths and cycles alternate between horizontal and vertical edges of G_C, we call such configurations *turning graphs* [1].

To correspond with the set Ω_{SO} of all perfect matchings on G_{so}, let Ω_C be the corresponding set of all turning graphs on G_C. There is a well-structured many-to-one map between Ω_{SO} and Ω_C. Let a vertex of G_C be called a *free vertex* of σ_C if it is not incident to any edge of σ_C. For every $\sigma_C \in G_C$ with exactly k free vertices, there are exactly 2^k pre-images $\in G_{so}$, as each "free" vertex in σ corresponds to a square in G_{so} whose edges can freely be matched in exactly two ways, independently of all other vertices (see Figure 2).

2.1 Weighted turning graphs

It will be useful to consider a generalized, weighted model of turning graphs on G_C. First, we introduce a natural weighted model on Ω_{SO}. Given a perfect matching $\sigma_{so} \in \Omega_{SO}$, let $\#N(\sigma_{so})$ be the number of octagon edges. For input parameter $\lambda > 0$, 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 a normalizing constant, also known as the *partition function.*

Projecting this weighting to Ω_C, for a particular turning graph σ_C, it follows that $\pi(\sigma_C)$ is

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

where $k(\sigma_C)$ is the number of free vertices of σ_C. Given a turning contour on G_C sampled according to the prescribed probability distribution, we can easily sample a perfect matching on G_{so} by flipping a bit for each free vertex uniformly at random, and assigning one of the two orientations of the perfect matching at each free vertex accordingly (see Figure 2).

We further generalize this model by introducing a parameter μ, and letting the weight of a configuration

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

where

$$Z \;=\; \sum_{\tau \in \Omega_C} \mu^{k(\tau)}\lambda^{|E(\tau)|}.$$

For convenience, we denote this probability model on Ω_C as $\pi_{\lambda,\mu} : \Omega_C \to \mathbb{R}$. Note that 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 effect of 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.

2.2 A Markov chain on weighted turning graphs

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_C, starting at any initial configuration σ_0. The number of steps t required to produce samples sufficiently close to equilibrium will be discussed subsequently.

The Markov chain $\mathcal{M}$

Repeat for t steps:

- Choose a face x of G_C uniformly at random.
- If no edges of x are in σ_t, let σ be the turning path created by adding the edges of face x.
- If every edge of x is in σ_t, 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 $\sigma_{t+1} = \sigma_t$.

This Markov chain represents precisely the octagon rotating moves of $\mathcal{M}_{so}$ and ignores the square rotating moves. Note that two configurations that differ by only a square rotation will map to the same turning graph. The fact that $\mathcal{M}$ connects the state space Ω_C of turning graphs of $\mathbb{G}_C$ and is aperiodic follows from the ergodicity of $\mathcal{M}_{so}$ on perfect matchings of the square octagon lattice [19].

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

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

We say that a Markov chain is *rapidly mixing* (or *polynomially mixing*) if the mixing time is bounded above by a polynomial in n, the size of a configuration, and $\log(\epsilon^{-1})$. It is *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 $\mu = 1$. In Section 4, we extend these results to the more general model when $\mu > 1$.

The key strategy of our proofs will center on the relationship between the *conductance* of a Markov chain and its mixing time [11, 24]. 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,

$$\Phi_M = \min_{S \subset \Omega} \{\Phi(S) : \pi(S) \leq 1/2\}.$$

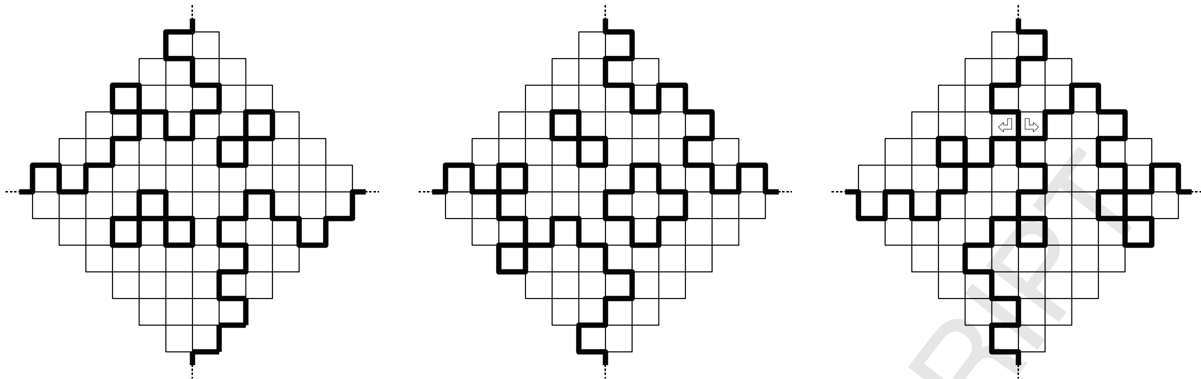
The conductance of a Markov chain is related to its mixing time $\tau(\epsilon)$ as follows:

Theorem 2.1: (Jerrum and Sinclair [11]) The mixing time of a Markov chain with conductance Φ satisfies:

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

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

The proofs that $\mathcal{M}$ is slowly mixing use so-called "Peierls arguments," first introduced to study phase transitions in statistical physics models [18]. The Peierls argument identifies an exponentially small (or unlikely) set in the state space by defining a map from this set to the entire state space with exponential gain in weight. When this unlikely set is a cut in the state space, this implies that the chain will take exponential time to move from one side of the cut to the other. This therefore implies that the chain is slowly mixing.



(a) Disjoint T-L and B-R paths (b) Disjoint T-R and B-L paths (c) A crossing pair of paths

Figure 3: Example configurations in (a) Ω_L, (b) Ω_R, and (c) Ω_C respectively in G_6.

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.

We first focus on the problem of sampling from our weighted distribution over turning graphs in the case when $\mu = 1$. Here, the weight of a contour depends only on the number of edges in the contour. We will show in Sections 3.1 and 3.2 that $\mathcal{M}$ is slowly mixing when λ is sufficiently small (or sufficiently large) by bounding the chain's conductance. Conversely, in Section 3.3 we show when $\mathcal{M}$ is polynomially mixing in the case $\lambda = 1$ by reducing to a related chain on 3-colorings of finite regions of the grid.

3.1 Slow mixing of $\mathcal{M}$ on $\pi_{\lambda,1}$ for small λ

We begin by showing that when $\lambda < 1/2\sqrt{e}$, the Markov chain $\mathcal{M}$ mixes slowly on a decorated version of a certain diamond graph also known as the "Aztec Diamond" graph G [7]. Starting with the standard Aztec Diamond graph of order n, G_n, we add extra edges E_B to the four corners of the diamond, which we denote as "T, R, B, L" to indicate the top, bottom, left and right corners of the diamond respectively. Let this subgraph of $\mathbb{Z}^2$ be G_C, and the corresponding region of Λ_{so} be G_{so}. We force a perfect matching in G_{so} to include our added corner edges, since those edges are degree 1. Therefore in G_C, each of those four corners must be the endpoint of some turning path. By construction, these four corners are in fact the *only* possible endpoints of a turning path, and therefore in any turning graph σ_C of G_C, there must either be turning paths from T to L and B to R, or alternately from T to R and B to L. (See Figure 3).

We will use these paths to define out cut, with the cut Ω_C separating the "left" and "right" sets Ω_L and Ω_R. Let Ω_L to be the set of configurations with paths from T to L and B to R but *not* from T to R or B to L. Similarly, let Ω_R be the set of configurations with paths from T to R and B to L but not from T to L or B to R. Finally, let the cut Ω_C consist of all states where both paths exist. For any state in Ω_C, we can identify a "crossing pair of paths" that are a *connected* set of edges that are the union of a T to L path and a B to R path. This crossing pair of paths has the interesting property that it could alternatively also have been interpreted as the union of a T to R path and a B to L path. 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].

When λ is small, we are favoring configurations where the turning paths are short, i.e. have few total edges. We will show that average configurations in Ω_C have many more edges than those of Ω_L or Ω_R, and thereby we will show that Ω_C has small weight relative to Ω_L and Ω_R. Recall that the conductance of a subset $S \subseteq \Omega$ is defined as $\Phi(S) = \sum_{s_1 \in S, s_2 \in \bar{S}} \pi(s_1)P(s_1, s_2)/\pi(S)$. We will show that the conductance of the chain is exponentially small when $\lambda < 1/(2\sqrt{e}$. This suggests that it will take a long time to transition between configurations in Ω_L and Ω_R if we have to pass through Ω_C, and from Theorem 2.1, we will therefore conclude that $\mathcal{M}$ mixes exponentially slowly.

Theorem 3.1: When $\lambda < 1/(2\sqrt{e})$, the mixing time of the Markov chain $\mathcal{M}$ on $\pi_{\lambda,1}$, weighted turning graphs of the Aztec Diamond G_n, is at least

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

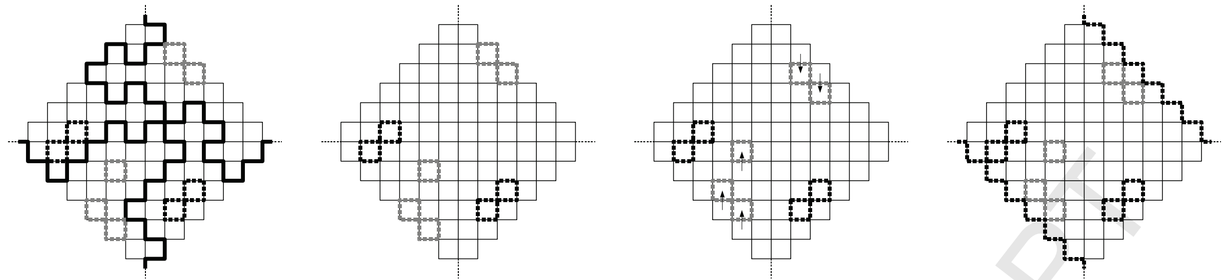
PROOF: For any $\sigma \in \Omega_C$, we first decompose the edges of σ into its crossing pair of paths, the union of a T to L turning path and a B to R turning path that share a vertex, which we call the *crossing vertex*. For any crossing pair of paths, we can uniquely identify the lexicographically first crossing vertex as a special vertex. These paths define four regions of G, one for each diagonal boundary, that can be viewed as a maximal connected component of faces that do not cross any edges of the crossing pair of paths (through they may cross edges of cycles). We will refer to these regions by the two corners of G that they border, e.g. the "T-L" region shown in Figure 3.

We will show that Ω_C is an exponentially small cut in our state space, thereby bounding the conductance. We first describe a map $\phi_r : \Omega_C \to \Omega$ 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 Figure 4). Given a state $\sigma \in \Omega_C$, take a pair of crossing paths in σ with maximal edges, and call the edges in this pair of crossing paths C. We remove the edges of C from G, and then shift all edges in σ from the B-L region up by one edge (we increase the y-coordinate by 1), and all edges from the T-R region down by one edge. Finally, we add in edges along the bottom left and top-right boundaries of G to form a valid turning graph.

It will be convenient to 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 left vertex to x, concatenated with the sub path of the bottom-right path from x to the right vertex. We similarly define the "vertical path" from the top vertex to x to the bottom vertex passing through x. This horizontal path, viewed as a left-to-right path, contains some $h \geq 0$ "backwards" edges from right to left. Similarly the "vertical path" has $v \geq 0$ backwards edges from bottom to top. In this case, 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 contains exactly $2n + 2h$ total horizontal edges. Moreover, it is the union of two turning paths, alternating horizontal and vertical edges, so the number of vertical edges in the horizontal path must be $2n + 2h + \delta_h$, where $\delta_h = 0$ if the edges at the point of intersection are vertical, and $\delta_h = -2$ if they are horizontal. Similarly, the vertical path has exactly $2n+2v$ vertical edges and $2n+2v+\delta_v$ horizontal edges, with $|\delta_v| \leq 2$. We note that $\delta_v = \delta_h = \delta$, as if the horizontal path crosses the point of intersection with horizontal edges, then the vertical path must cross the point of intersection with vertical edges.

If $\delta = 0$, we may encode the horizontal and vertical paths essentially as two separate, equal length interleaved bit sequences, one sequence for the horizontal moves, another for the vertical moves, and a single special symbol x to indicate the location of the crossing. Given the location x, we can also account for the case where $\delta = -2$ by adding a single extra bit.



(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 $\phi : \Omega_C \to \Omega$.

By this encoding, we see that a bound on the number of pre-images of the map ϕ that have left-right paths of type h is at most

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

Similarly, the number of preimages of ϕ that have top-down paths of type v is at most

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

First, we see that by Stirling's approximation,

$$\begin{aligned}\binom{2n+2h}{n+h}\binom{2n+2v}{n+v} &\leq \frac{2^{2n+2h}}{\sqrt{\pi(n+h)}}\frac{2^{2n+2v}}{\sqrt{\pi(n+v)}} \\ &= \frac{2^{4n+2h+2v}}{\pi\sqrt{n+h}\sqrt{n+v}}.\end{aligned}$$

Similarly by Stirling's approximation and the well known approximation for e,

$$\begin{aligned}\binom{2n+2h}{h}\binom{2n+2v}{v} &\leq \frac{(2n+2h)^h}{h!}\frac{(2n+2v)^v}{v!} \\ &\leq \frac{(2e)^{h+v}}{2\pi\sqrt{hv}}\left(1+\frac{n}{h}\right)^h\left(1+\frac{n}{v}\right)^v \\ &\leq \frac{e^{2n}(2e)^{h+v}}{2\pi\sqrt{hv}}.\end{aligned}$$

Finally, it follows that $P(h,v)$, the total number of preimages of type (h,v), is therefore at most

$$\begin{aligned}P(h,v) &\leq 2n^2\binom{2n+2h}{h}\binom{2n+2h}{n+h}\binom{2n+2v}{v}\binom{2n+2v}{n+v} \\ &\leq 2\frac{n^2 2^{4n+2h+2v}(2e)^{h+v}(e)^{2n}}{2\pi^2\sqrt{hv}\sqrt{n+h}\sqrt{n+v}}.\end{aligned}$$

We map from a configuration with $8n + 4v + 4h$ edges in the crossing paths to a configuration with exactly $4n$ new edges, so 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:

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

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

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

By Theorem 2.1, it follows that $\tau(\epsilon)$, the mixing time of $\mathcal{M}$, satisfies

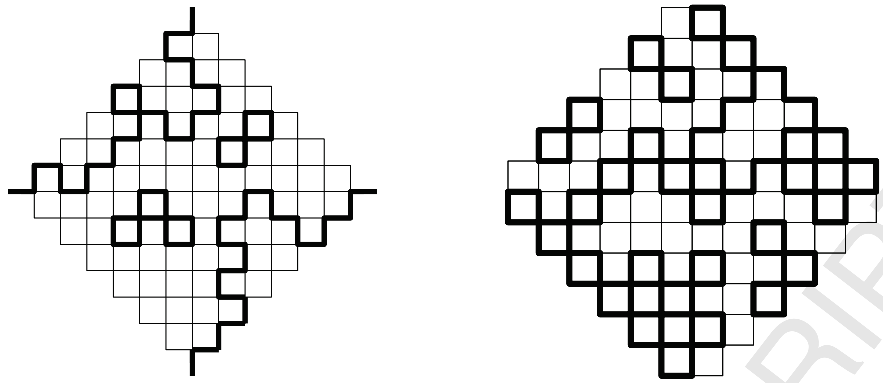
$$
\tau(\epsilon) \geq \frac{1}{8n}(2\lambda\sqrt{e})^{-4n} \ln \epsilon^{-1},
$$

so we require exponentially many steps to converge when $\lambda < 1/(2\sqrt{e})$. ∎

3.2 Slow mixing of $\mathcal{M}$ on $\pi_{\lambda,1}$ for large λ

Perhaps surprisingly, we can also show that the chain is slowly mixing when λ is large. This actually follows from a duality between the edges and non-edges in a turning graph on the grid.

Specifically, 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 a slightly different version of the Aztec Diamond graph. 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 $\lambda' = 1/\lambda$.

Where G is the modified Aztec Diamond graph described in the previous section, let G' be the corresposonding unmodified Aztec Diamond graph without the corner edges added. For each vertex v in G with parity requirement $p(v)$, set the parity of the corresponding vertex in G' to $deg(v) - p(v)$ in G'. It follows that for any turning graph σ_C of G, the complementary turning graph $\sigma'_C = E \backslash \sigma_C$ 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' (see Figure 5).

Corollary 3.2: When $\lambda > 2\sqrt{e}$, the mixing time of the Markov chain $\mathcal{M}$ on $\pi_{\lambda,1}$, weighted turning graphs of the Aztec Diamond G_n, is at least

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

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 $sigma'_C$ with parameter $\lambda' = 1/\lambda$ is exactly

$$G' = \lambda'^{|\sigma'_C|} = \lambda'^{|E| - |\sigma_C|} = \lambda'^{|E|}\lambda'^{-|\sigma_C|} = \lambda'^{|E|}\lambda^{|\sigma_C|}.$$

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

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

Thus, the normalized probability $\pi(\sigma'_C) = a\pi(\sigma_C)$.

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 3.1. ∎

3.3 Polynomial mixing of $\mathcal{M}$ on $\pi_{\lambda,1}$ when $\lambda = 1$

On the positive side, we now show that the chain $\mathcal{M}$ does converge to equilibrium efficiently when $\mu = \lambda = 1$. Our proof relies on the fact that a corresponding Markov chain on proper three colorings of finite regions of $\mathbb{Z}^2$ is known to be polynomially mixing [17, 22]. We will then describe a novel bijection between turning paths of a region of the grid $G = (V, E)$ with three-colorings of the faces of G, both subject to certain boundary conditions. This bijection will allow us to infer that the Markov chain $\mathcal{M}$ is polynomially mixing on any region G of $\mathbb{Z}^2$ when $\mu = \lambda = 1$.

A proper three-coloring of a region G of the grid is a labeling of each vertex v in G with a color chosen from $\{0, 1, 2\}$ such that no edge of G has two ends with the same color. Proper 3-colorings of graphs (and more generally k-colorings) are natural combinatorial structures that arise in numerous contexts across mathematics and computer science. They are studied in statistical physics, notably in the context of the zero-temperature anti-ferromagnetic Potts model for general graphs and the "6 vertex ice model" when restricted to regions of $\mathbb{Z}^2$. In many of these settings, the following natural, local Markov chain, known as the single-site *Glauber Dynamics*, is of key interest.

Let G be a finite, connected region of $\mathbb{Z}^2$, and Ω_3 be the space of 3-colorings of G. We define $\mathcal{M}'$ on Ω_3 as follows:

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

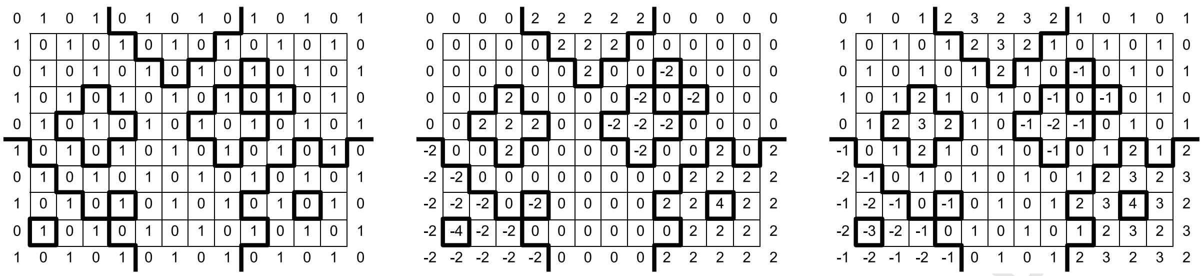
Beginning at any given initial coloring σ_0, repeat for t steps:

- Choose a vertex v of G uniformly at random, and $b \in 0, 1, 2$ uniformly at random.
- Let σ_{t+1} be σ_t with v colored b if this is a valid coloring.
- Otherwise $\sigma_{t+1} = \sigma_t$.

In the context of sampling Eulerian rotations, Luby et al. [17] proved that the chain $\mathcal{M}'$ is rapidly mixing for any region G of $\mathbb{Z}^2$ with fixed colors on the boundary of G. Goldberg et al. [9] proved that $\mathcal{M}'$ is also rapidly mixing for rectangular regions G with free (unrestricted) boundary conditions. Cannon and Randall [4] subsequently generalized these results to more complicated regions G with hybrid fixed and free boundary conditions. Importantly, these results made use of the *height representation* of a 3-coloring of G.

Since we will be constructing a coloring of the faces of G, we will describe what follows in the context of coloring faces, although these results were originally described in the context of coloring vertices. Given a region G of $\mathbb{Z}^2$, a *height function* on the faces $F(G)$ of G is an assignment $h : F(G) \to \mathbb{Z}$ such that for any two neighboring faces v, w on G, $|h(v) - h(w)| = 1$. Every height function is assigned a canonical 3-coloring of G simply by assigning color $h(f) \mod 3$ to each face f. As each 3-coloring has multiple possible height functions, we may fix an arbitrary face f_0 of G and declare its height to be 0. There is then a bijection between all 3 colorings of G where f_0 is colored 0 and all valid height functions of G where $h(f_0) = 0$ [17] (see Figure 6). Note that as the graph of the faces of G is bipartite and is also a region of $\mathbb{Z}^2$, all vertices with even height must lie on the same side of the bipartition as f_0. We call these "even" faces, and all other vertices "odd" faces.

We are now ready to describe a novel bijection between turning graphs of a region of the grid $G = (V, E)$ with three-colorings of the faces of G, both subject to fixed boundary conditions. Informaly, we think of a turning graph as every other level curve of the height function of some 3-coloring. As we shall see, a face rotation transition of $\mathcal{M}$ in the turning graph G then corresponds exactly to changing the color assigned to a single face of G, which is a move of $\mathcal{M}'$. This will allow



(a) The parity function h_p (b) Twice the level function $2h_l$ (c) The height $h = h_p + 2h_l$

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

us to translate the proofs of rapid mixing of $\mathcal{M}'$ in the context of three colorings to a proof of rapid mixing of $\mathcal{M}$ in the context of turning graphs.

Theorem 3.3: Given a simple, connected region G of $\mathbb{Z}^2$, along with boundary conditions representing the parity of incident edges allowed in any turning graph of G at each boundary vertex v, we can assign colors to the faces on the external boundary of G such that there is a bijection between the set of turning graphs on G satisfying the given vertex boundaries and the set of 3-colorings on the faces of G satisfying the constructed coloring boundaries.

Proof: It will be convenient to first understand the properties of the mapping from height functions to turning paths. Let f_0 be any face on the exterior boundary, and without loss of generality we consider only height functions that set $h(f_0) = 0$. For any integer k, we say that the faces at level k are those with height value either $2k$ or $2k+1$. By properties of the height function, faces at level k can only be adjacent to faces at level $k-1, k$, or $k+1$.

Consider the (possibly non-simple) path of edges along the boundary between maximally connected regions of faces at height levels k and $k+1$. By construction, this path must have odd faces on one side, and even faces on the other side. It is easy to see that this can only happen if this path has the turning property. By construction, the paths between levels k and $k+1$ cannot touch the paths between $k-1$ and k (or any other levels), as this would imply that there are vertices connected by a path of two edges with at least a height difference of at least three.

We define the map ρ of the height function h to be the set of all edges between levels of h, forming a configuration σ which, by the above argument, is a valid turning graph on G. This turning graph σ can be decomposed into paths and cycles. The paths all end at edges that separate boundary faces from two different levels. Since the locations of these endpoints depend only on the values of the height function on the boundaries of σ, it follows that any σ' with the same values of the height function on the boundaries must map to a turning graph with paths that end in the same locations as σ.

It is easy to show that this map from height functions to turning paths is injective, as if two height functions differ in their assignment at any face f, the difference in assigned height values must be at least 2, by parity. Therefore that vertex would belong to different levels in the map from each height function. As we fixed face f_0 to have color 0 in both maps, the number of turning paths crossed in any path from f_0 to f would be different in the two maps, and therefore the two turning graphs must be different as well.

To show that this map is surjective, we consider the following reverse map ρ^{-1} from the set of turning graphs to height functions. Say we are given a simply connected region G of $\mathbb{Z}^2$ with parities $p(v)$ on each vertex on the boundary. We begin by identifying the vertices for which $p(v)$ is odd; these are the starting points for the turning paths of G. These vertices separate the boundary into disjoint sections of edges, which we identify as $B_1, B_2, \ldots$. We may identify each boundary section as either a set of edges or exterior faces as convenient. For any turning graph σ on G, the edges of σ separate the faces of G into maximal, connected regions $R_1, R_2, \ldots$, and each boundary section connects to exactly one (possibly empty) region.

We begin by fixing a face f_0 of the external boundary of G. We will say that the *height parity* $h_p(f) = 0$ for even faces and $h_p(f) = 1$ for odd faces. We then assign a *height level* to each region as follows. Let R_0 be the region that contains f_0. We first set $h_l(R_0) = 0$. We require that if regions R_1 and R_2 share a boundary with odd faces on the side of R_1 and even faces on the side of R_2, then $h_l(R_1) + 1 = h_l(R_2)$. We can assign height levels to every region by performing a depth-first search beginning at R_0 and assigning each region a height relative to its parent following the above requirement. For any face f in R_1, we say that $h_l(f) = h_l(R_1)$.

We can then define $h(f) = h_p(f)+2h_l(f)$. Let f_1, f_2 be two two adjacent faces of G. If they are in the same region, $|h(f_1)-h(f_2)| = |h_p(f_1)-h_p(f_2)| = 1$. If they are in differing regions, then without loss of generality let f_1 be the odd face. Then $|h(f_1)-h(f_2)| = |h_p(f_1)-h_p(f_2)+2(h_l(f_1)-h_l(f_0))| = |(1) + 2(-1)| = 1$. Therefore, h is a valid height function. The corresponding turning graph $\rho(h)$ consists of the edges between every other level of the h, which by construction are exactly the edges of σ. Therefore $\sigma = \rho(h)$, and we have shown that ρ is a bijection. ■

Corollary 3.4: The Markov chain $\mathcal{M}$ when $\lambda = \mu = 1$ is polynomially mixing on finite regions of the grid G.

Proof: The Markov chain $\mathcal{M}$ that adds or removes a single square of edges around a face f would create or destroy a region that consists of a single vertex. This would either increase or decrease the height function at that vertex by exactly two. This 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 on a related chain [17]. ■

4 Mixing of $\mathcal{M}$ on $\pi_{\lambda,\mu}$ 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 $\pi_{\lambda,\mu}$, weighted turning graphs of the Aztec Diamond G_n, is at least

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

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 3.1, keeping both the structure of the proof and the map ϕ.

Let σ be a configuration in $\Omega_{C,v,h}$. As in Theorem 3.1, 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 3.1, this implies by Theorem 2.1 that $\mathcal{M}$ mixes exponentially slowly when $\lambda/\sqrt{\mu} < 1/2\sqrt{e}$, or more simply $\lambda < \sqrt{\mu}/2\sqrt{e}$. ∎

We now similarly analyze the case where $\lambda > 1$ and obtain a result analogous to Corollary 3.2 for this more general case. Following the bijection in Corollary 3.2 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 3.2.

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

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

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 3.2 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 rapidly mixing for some settings of μ and λ and slow to converge to equilibrium for other settings. This suggests the presence of a phase transition in these parameters, although our proofs do not identify the critical point, if it exists. An important feature of our results is the strategy of adapting a proof that $\mathcal{M}$ is slowly mixing for the simpler setting when $\mu = 1$ to settings with 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 original goal. Although our specific methods do not seem sufficient for extending the proofs to this setting, we gain new insights on why experiments suggest that the unweighted chain on perfect matchings of the square-octagon lattice is slow. 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] Nayantara Bhatnagar, Sam Greenberg, and Dana Randall. The effect of boundary conditions on mixing rates of markov chains. *Proceedings of APPROX/RANDOM*, pages 280–291, 2006.

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

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

[4] Sarah Cannon and Dana Randall. Sampling on lattices with free boundary conditions using randomized extensions. *Proceedings of the 27th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1952–1971, 2016.

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

[6] Jack Edmonds. Paths, trees, and flowers. *Canad. J. Math.*, 17:449–467, 1965.

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

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

[9] Leslie Ann Goldberg, Russell Martin, and Mike Paterson. Random sampling of 3-colorings in z^2. *Random Structures & Algorithms*, pages 279–302, 2004.

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

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

[12] M. R. Jerrum and A. J. Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM Journal on Computing*, 1993.

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

[14] P. W. Kasteleyn. The statistics of dimers on a lattice : I. The number of dimer arrangements on a quadratic lattice. *Physica*, pages 1209–1225, 1961.

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

[16] S. Lubetzy and A. Sly. Critical Ising on the square lattice mixes in polynomial time. *Communications in Mathematical Physics*, pages 815–836, 2012.

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

[18] R. Peierls. On ising's model of ferromagnetism. *Proceedings of the Cambridge Philosophical Society*, page 477, 1936.

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

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

[21] James Gary Propp and David Bruce Wilson. Exact sampling with coupled markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 9(1-2):223–252, 1996.

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

[23] D. Randall and D. B. Wilson. Sampling spin configurations of an Ising system. In *Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA, pages 959–960, 1999.

[24] A. J. Sinclair. *Algorithms for random generation and counting.* Progress in theoretical computer science. Birkhäuser, 1993.

[25] H. N. V. Temperley and Michael E. Fisher. Dimer problem in statistical mechanics-an exact result. *Philosophical Magazine*, 1961.

[26] L.G. Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, pages 189 – 201, 1979.

[27] J.C. Vera, E. Vigoda, and L. Yang. Improved bounds on the phase transition for the hard-core model in 2-dimensions. *Proceedings of APPROX/RANDOMs*, pages 699–713, 2013.

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