J. Appl. Math. & Informatics Vol. **37***(2019), No. 5 - 6, pp. 469 - 480*
https://doi.org/10.14317/jami.2019.469

GENERALIZED DOMINOES TILING'S MARKOV CHAIN MIXES FAST

K.K. KAYIBI, U. SAMEE, MERAJUDDIN AND S. PIRZADA*

Abstract. A generalized tiling is defined as a generalization of the properties of tiling a region of $\mathbb{Z}^2$ with dominoes, and comprises tiling with rhombus and any other tilings that admits height functions which can be ordered into a distributive lattice. By using properties of the distributive lattice, we prove that the Markov chain consisting of moving from one height function to the next by a flip is fast mixing and the mixing time $\tau(\epsilon)$ is given by $\tau(\epsilon) \leq (kmn)^3 (mn \ln k + \ln \epsilon^{-1})$, where mn is the area of the grid Γ that is a k-regular polycell. This result generalizes the result of the authors (T-tetromino tiling Markov chain is fast mixing, Theor. Comp. Sci. (2018)) and improves on the mixing time obtained by using coupling arguments by N. Destainville and by M. Luby, D. Randall, A. Sinclair.

AMS Mathematics Subject Classification : 05C99, 11K99.
Key words and phrases : Dominoes, distributive lattice, tiling, partition functions, height function, mixing time of Markov chain, FPRAS.

1. Introduction

Throughout this paper, a *lattice* P is a set $\mathcal{L}$ and a partial order $\preceq$ such that, for all $x, y, z \in \mathcal{L}$, the following hold.
1. $x \preceq x$.
2. $x \preceq y$ and $y \preceq x$ entails $x = y$.
3. $x \preceq y$ and $y \preceq z$ entails $x \preceq z$.
4. For every pair (x, y), there exit $x \wedge y$ and $x \vee y$, where $x \wedge y$ is the largest element of P that is smaller than both x and y, and $x \vee y$ is the smallest element of P that is bigger than both x and y. A lattice is *distributive* if $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ and $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$. A *grid* in $\mathbb{Z}^r$ is a set of points of $\mathbb{Z}^r$ that are equally spaced. We note that many authors in the tiling literature use the term lattice for what we call grid.

Received January 11, 2019. Revised August 3, 2019. Accepted August 5, 2019. *Corresponding author.

A *tile* is a connected subset of $\mathbb{R}^2$. Let θ be a tile and Γ be a subset of $\mathbb{R}^2$. A *tiling* of Γ with θ is a covering of Γ by copies of θ such that there is no overlap between two tiles. An example of tiling is given in Figure 1 (b). Apart from their recreational values [4], tilings have received much attention because of their connection with the partition function in Statistical Mechanics [2, 6, 7, 12, 11, 18]. A typical example is the modeling of crystal structures using tilings on various grids. A *dimer* is a diatomic molecule. A *matching* of graph G is a set of edges of G such that no vertex is adjacent to two vertices or more. A *complete matching* is a matching that covers all the vertices. The adsorption of dimers on the surface of a crystal can be modelled as a matching on a two dimensional rectangular grid where atoms are sited on vertices and the links between the atoms are the edges. This model is illustrated in Figure 1 (a) and (b).

FIGURE 1. (a) Dimers adsorbed on the surface of a crystal. Each dimer is represented by two vertices joined by an edge in the grid. Notice that this configuration of dimers is just a complete matching on the grid. (b) The tiling with dominoes corresponding to the configuration of dimers in (a).

Now, putting a dimer i in some position on the grid Γ requires some energy, ϵ_i, which may be positive or negative. Hence, if C denotes a configuration of dimers, then C has an energy $E(C)$ associated with it and this energy is given by $E(C) = \sum_i \epsilon_i$, where the sum is over all dimers $i \in C$. An obvious physical requirement is that the most probable configurations should be the ones with the lowest energies. This requirement is only satisfied if the probability to get a configuration of energy level $E(C)$ is of the form

$$Pr(C) = \frac{e^{-\beta E(C)}}{Z}, \tag{1}$$

where $Z = \sum_C e^{-\beta E(C)}$, *the partition function*, is the normalization factor to get a probability distribution, and $\beta = \frac{1}{\kappa T}$, where T is the absolute temperature and κ is the constant of Boltzmann. The partition function is used to calculate various thermodynamic properties of the system. For example, for a system containing N particles, the *internal energy* U, is given by

$$U = -N\frac{\delta \ln Z}{\delta \beta} = N\kappa T^2 \frac{\delta \ln Z}{\delta T}.$$

The entropy of the system, denoted by S, is given by

$$S = \frac{U}{T} + N\kappa \ln Z.$$

From these two quantities, it is easy to find the *Helmholtz free energy*, $F = U - TS$.

From Equation (1), it is obvious that lower the temperature, the higher the probability to find configurations of lowest energy. If $T = 0$, the system would consist only of configurations called *ground states*. At higher temperatures, all the states are equiprobable, and thus, when $T \to \infty$, calculating Z ends up being equivalent to counting the number of different configurations of dimers. That is, counting the number of matchings on the lattice graph. This, in turn, is equivalent to counting the number of tilings of an appropriate lattice by dominoes and hence the connection between tilings and Statistical Mechanics.

In a typical Statistical Mechanics problem, N, the number of particles in the system is of the magnitude of 10^{23}, the Avogadro number. Thus, obtaining Z may involve an intractable number of computations. Although some exact counting has been found for some cases [6, 7, 18], most practical works rely on sampling to calculate the statistics under study. Sampling techniques, such as the Metropolis Sampling, construct a Markov chain and aim at sampling configurations according to some probability distribution (preferably the uniform distribution.) Since successive states of a Markov chain are not independent, an unbiased sample can only be obtained if the chain reaches stationarity, the stage when the probability of sampling a particular configuration is fixed in time. Informally, the *mixing time* of a Markov chain is the number of steps necessary to reach the stationary state. And, knowing the mixing time of a particular Markov chain is crucial to avoid either to get a biased sample (if stationarity is not reached), or to avoid the computational cost of running the chain more than necessary.

1.1. Canonical path and mixing time of Markov Chains. Let $\mathcal{M}$ be a Markov chain on a set of states Ω. Let P be the matrix of transitions from one state to another. One may visualize the Markov chain $\mathcal{M}$ as a weighted directed graph $\mathcal{G}$ where the states are vertices of $\mathcal{G}$ and there is an edge of weight $P(x, y)$ if the transition from state x to state y has probability $P(x, y)$. A Markov chain is *irreducible* if, for all pairs of states x and y, there is an integer n, depending on the pair (x, y), such that $P^n(x, y) > 0$. In terms of the graph $\mathcal{G}$, the Markov chain is irreducible if there is a path between every pair of vertices of $\mathcal{G}$. A Markov chain is *aperiodic* if for all states x, there is an integer n such that for all $n' \geq n$, $P^{n'}(x, x) > 0$. That is, after sufficient number of iterations, the chain has a positive probability to stay on x. This ensures that the return to state x is not periodic. In terms of the graph $\mathcal{G}$, this can be achieved by having a loop at

every vertex. A Markov chain is *ergodic* if it is irreducible and aperiodic. It can be shown that if P is the matrix of transitions of an ergodic Markov chain, then there is an integer n such that for all the pairs of states (x, y), $P^n(x, y) > 0$. (Notice that n does not depends on the pair). It can also be proved that for every ergodic Markov Chain with transition matrix P, the largest eigenvalue of P is equal to 1 (Perron-Frobenius Theorem). Using this, it can be proved that there is a unique probability vector π, such that $\pi P = \pi$. The vector π is the *stationary* distribution of $\mathcal{M}$.

Let $\mathcal{M}$ be an ergodic Markov chain defined on a finite set of states Ω, with a transition matrix P and stationary distribution π. Starting the chain from an initial state x, we would like to measure $\Delta_x(t)$, the distance between the distribution at time t and the stationary distribution. More formally, if $P^t(y|x)$ represents the probability that, at time t, the chain is at state y given initial state x, and $\pi(y)$ represents the probability that the chain is at state y at stationarity, then the *variation distance*, denoted by $\Delta_x(t)$, is defined as

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

A converging chain is one such that $\Delta_x(t) \to 0$ as $t \to \infty$ for all initial states x. The *rate of convergence* is measured by $\tau_x(\epsilon)$, the time required to reduce the variation distance to ϵ given an initial state x.

$$\tau_x(\epsilon) = min\{t : \Delta_x(t) \leq \epsilon, \ for\ all\ t' \geq t\}.$$

The *mixing time* of the chain, denoted by $\tau(\epsilon)$, is defined as $\max_{x \in \Omega}$, the maximum being over all the initial points x. A Markov chain is said to be *rapidly mixing* if its mixing time is bounded above by a polynomial in the size of the input and $\frac{1}{\epsilon}$.

Let $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of the transition matrix P. The *spectral gap* of the matrix P is defined as $max\{1 - \lambda_2, 1 - |\lambda_n|\}$. If $P(x, x) < \frac{1}{2}$ for all x, then $\lambda_n > 0$, and therefore smaller than λ_2. The spectral gap of P is then the real number $\lambda_1 - \lambda_2$, that is, $1 - \lambda_2$. It can be shown that largest the gap, the faster is the mixing time of the chain. The analysis of the mixing time of a Markov chain is based on the intuition that a random walk on the graph $\mathcal{G}$ mixes fast (i.e., reaches all the states quickly) if $\mathcal{G}$ has no bottleneck. That is, there are no cuts between any set of vertices S to its complement, which blocks the flow of the Markov chain and thus prevents the Markov chain from reaching easily some states. See [5, 17] for a better exposition on the topic. To make this more formal, we need some preliminary definitions which conforms with [17].

The analysis of the mixing time of a Markov chain is based on the intuition that a random walk on the graph $\mathcal{G}$ mixes fast (i.e., reaches all the states quickly) if $\mathcal{G}$ has no bottleneck. That is, there are no cuts, between any set of vertices S to its complement, that blocks the flow of the Markov chain and thus prevents

the Markov Chain from reaching easily some states.

Denoting the probability of x at stationarity by $\pi(x)$ and the probability of moving from x to y by $P(x,y)$, the *capacity* of the arc $e=(x,y)$, denoted $c(e)$, is given by

$$c(e)=\pi(x)P(x,y).$$

Let $\mathcal{P}_{x,y}$ denote the set of all simple paths p from x to y (paths that contain every vertex at most once). A *flow* in $\mathcal{G}$ is a function ϕ, from the set of simple paths to the reals, such that

$$\sum_{p\in\mathcal{P}_{x,y}}\phi(p)=\pi(x)\pi(y),$$

for all vertices x,y of $\mathcal{G}$ with $x\neq y$. A flow along an arc e is then defined as

$$\phi(e)=\sum_{p\ni e}\phi(p).$$

For a flow ϕ, a measure of existence of an overload along an arc is given by the quantity $\rho(e)$, where

$$\rho(e)=\frac{\phi(e)}{c(e)},$$

and the *cost* of the flow ϕ, denoted by $\rho(\phi)$, is given by

$$\rho(\phi)=\max_{e}\rho(e).$$

If a network $\mathcal{G}$ representing a Markov chain can support a flow of low cost, then it can not have any bottlenecks, and hence its mixing time should be small. This intuition is confirmed by the following Theorem [14].

Theorem 1.1. [14] *Let $\mathcal{M}$ be an ergodic reversible Markov chain with holding probabilities $P(x,x)\geq\frac{1}{2}$ at all states x. The mixing time of $\mathcal{M}$ satisfies*

$$\tau_x(\epsilon)\leq\rho(\phi)|p|\left(ln\frac{1}{\pi(x)}+ln\frac{1}{\epsilon}\right),$$

where $|p|$ is the length of a longest path carrying non-zero flow in ϕ.

Literature abounds with analysis of mixing times of different tilings [3, 8, 9, 14, 17]. Luby et al. [14] use a coupling argument to prove that the mixing time of tilings in planar grids is given by $\tau(\epsilon)\leq 12n^4(1+\ln\epsilon^{-1})$, where n is the area of the grid. This applies to dominoes, rhombus and tiling that can be represented as Eulerian orientations on a graph. This bound is confirmed in [3], for the mixing time of plane rhombus tilings, through another heuristic coupling argument. By taking an approach based on the structure of the lattice of the height functions, the present paper recovers the seminal result [14] while improving on the bound on the mixing times. Our approach uses the canonical

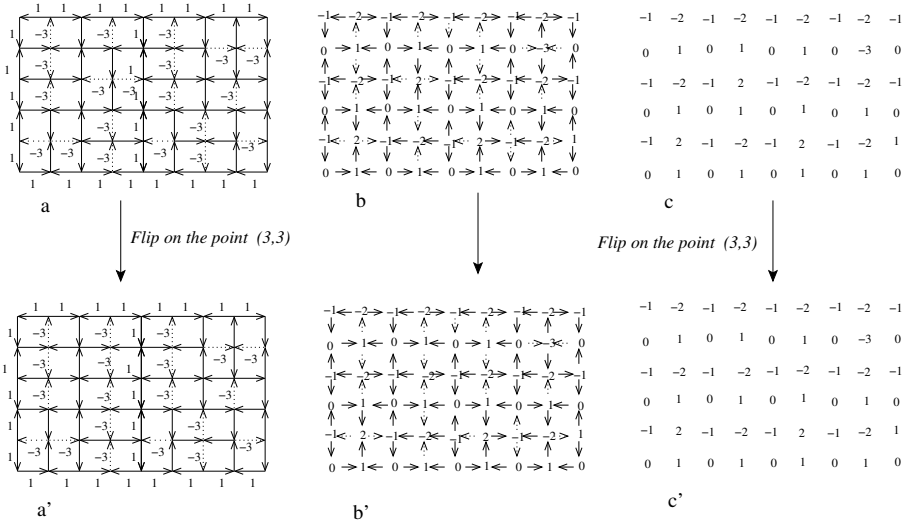

FIGURE 2. (a) 4-regular polycell and weight on its edges. To avoid overcrowding, we have only the tension on edges representing tiles and some edges of tension 1. (b) The difference of potential on vertices. (c) The height function of the tiling.

path analysis, as in [8, 10], and is based on the notion of generalized tilings introduced in [1].

2. Generalized tiling

Bodini and Latapy [1] defined a generalised tiling as follows. Let G be a simple directed graph. A *cell* is an elementary oriented circuit of G and a *polycell*, denoted by Γ, is a set of cells. The *vertices* of Γ are vertices of its cells. A polycell is *k-regular* if every one of its cells is a circuit of length k. A *boundary* of Γ is an arbitrary partial subgraph of P. Given an edge $e \in \Gamma$ that is not on the boundary, a *tile* is the set of all cells that have e in common. A *tiling* of Γ is a set of tiles that constitutes a partition of the cells of Γ. See Figure 2 and 3 for an illustration, where the polycells are 4-regular and 3-regular, respectively. The 4-regular polycell is tiled with dominos while the 3-regular is tiled with rhombuses.

Now, let $\mathcal{T}$ be a tiling of a k-regular polycell Γ. A flow, $\phi_{\mathcal{T}}$, is associated to $\mathcal{T}$ in the following way: $\phi_{\mathcal{T}}(e) = 1 - k$ if e is a tile of $\mathcal{T}$, and $\phi_{\mathcal{T}}(e) = 1$, otherwise.

The *height function* of $\mathcal{T}$, denoted by h, is a function from the set of vertices of Γ to the integers, defined recursively as follows. In Γ, choose a vertex, v, (preferably the one with the least cartesian coordinates.) Let $h(v) = 0$. Then, starting from v, if $e = (v_i, v_j)$, let $h(v_j) = h(v_i) + \phi_{\mathcal{T}}(e)$ if e is directed towards v_j and $h(v_j) = h(v_i) - \phi_{\mathcal{T}}(e)$ if e is directed away from v_j. Using the definition

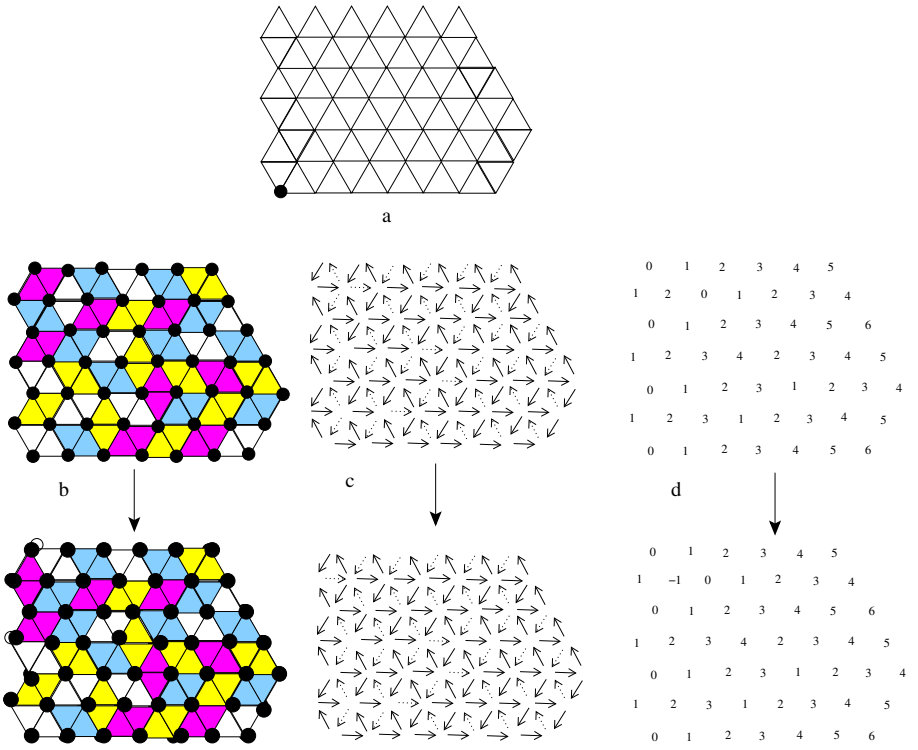

FIGURE 3. (a) illustrates a triangular grid tiled with rhombuses, (b) represents the orientation associated with the tiling. The edges represented by broken lines are the ones representing the tiles as illustrated at (c). The height function of the tiling is given at (d). Below (b), (c) and (d) are their equivalents for the new tiling obtained from the tiling (a) by flipping the top left tile.

of the orientations of edges of γ, it is easy to check that for a given tiling $\mathcal{T}$, h is well defined, in the sense that it does not depend on the trajectory taken. Moreover, it is easy to check that there is a bijection between tilings of Γ and the set of height functions.

A *flip* on a vertex v consists of changing the positions of all the tiles that have v in common. In terms of the oriented edges, a flip consists of permuting the roles of edges incident to v: it changes the edges representing tiles into normal edges and normal edges to edges representing tiles in such a way as to respect the orientations of the cells. A point v is *flippable* if a flip can be performed on it. From the definition of height function on a k-regular polycell, it is easy to notice that a point v is flippable if and only if its height is exactly one unit away

(lower or higher) from that of half of its adjacent neighbors, and is $k-1$ units away (lower or higher) from the other half. From this fact, it is easy to notice that if Γ is a k-regular polycell, and a flip at v changes the height function h into h', then $h'(v) = h(v) \pm k$, where $+$ applies if the height of v is lower than its neighbors and $-$ applies otherwise. Let the *distance* between two vertices v and w in the height function h be defined as $|h(v) - h(w)|$. Notice that a flip preserves the set of distances between v and its adjacent neighbors (while there is a permutation of distances within the set of neighbors of v). Now, we give properties of the lattice of height functions that are required to prove the main theorem of this paper. Let h and h' be two tilings of Γ. An ordering $\prec$ is defined on the set of height functions by $h \prec h'$ if $h(w) \leq h'(w)$ for all vertices w of Γ. Let $|h - h'|$, the *distance between* h and h' be defined as the number of vertices v such that $h(v) \neq h'(v)$.

Lemma 2.1. *In a generalized tiling, every flip is an involution, that is, performing a flip twice in succession at the same point brings the point back to its original height.*

Proof. Let Γ be a k-regular polycell and let the operator g be a flip at a point v having height $h(v)$. It is obvious that $g(h)$, the height function obtained after the flip is also a height function. Now, if v is flippable, then $|h(v)-h(w)| = k-1$ or $|h(v)-h(w)| = 1$ if w is a neighbor of v. We know that the operation flipping consists only of permuting the distances between v and its neighbors and raises or lowers the height of v by k. That is, $g(h(v)) = h(v) \pm k$. Since $g(h)$ is also a height function, the set of distances between v and its neighbors is preserved in $g(h)$. And, by induction, the operation $gg(h) = g^2(h)$ also preserves the distance between v and its neighbors, since $g^2(h)$ is also a height function. But the only way to permute the distance between v and its neighbors while preserving the set of these distances is that $g^2(h) = h$. For, if $g^2(h) \neq h$, then, for some vertex w adjacent to v, $|g^2(h(v)) - g^2(h(w))| \geq 2k$. Hence $g^2(h)$ is not a height function. This is a contradiction. Therefore $g^2(h) = I$, where I is the identity operator. □

Two points v and w of Γ are *independent* if flipping at v does not effect as whether w is flippable or not. That is, flipping at v does not change the height of the point w and the heights of its nearest neighbors. A height function h *covers* a height function h' if $h \succ h'$ in the ordering $\prec$. In an obvious way, *flipping on a set* means flipping on all the points of the set.

Lemma 2.2. *Let h and h' be two height functions (tilings), and let $D = \{v_1, \ldots, v_t\}$ be the set of points where h differs from h'. That is, $h(v_i) \neq h'(v_i)$ for all $i \in \{1, \ldots, t\}$.*

(1) For $i \in D$, if $|h(v_i) - h'(v_i)| \geq 2k$, then all the neighbors of v_i belong to D.

(2) For any v, $|h(v) - h'(v)| \leq mk$.

Proof. (1) Let w be a neighbor of v_i. From the definition of the height function, we have $|h(v_i)-h(w)| \in \{1,(k-1)\}$. If $|h(v_i)-h'(v_i)| \geq 2k$ and $|h(w)-h'(w)| = 0$, then $|h'(v_i) - h'(w)| \notin \{1,(k-1)\}$. Hence h' is not a height function, which is a contradiction.

(2) If v is on the boundary, then $h(v) = h'(v)$. That is, $|h(v) - h'(v)| \in \{0\}$. If w is a neighbor of a boundary point but w is not on the boundary, then $|h(w) - h(v)| \in \{1,(k-1)\}$ and $|h'(w) - h'(v)| \in \{1,(k-1)\}$, since h and h' are both height functions. Hence $|h(w) - h'(w)| \in \{0,2,k,(k-2),2(k-1)\}$. Using the same argument, if u is a neighbor of w but not a neighbor of any boundary point, then $|h(u) - h'(u)|$ can be a sum or (subtraction) of any pair of elements of $\{0,2,k,(k-2),2(k-1)\}$. Starting from a boundary point v, the chain $(v,w,u,\dots)$ must terminate at another boundary point v'. But starting from v', such a chain must also meet the chain starting from v. And the longest such chain would be if one started from leftmost point and moved diagonally towards the mid point of the grid Γ. Such a chain would consist of $\frac{m+n}{2}$ vertices. Hence, if r is a mid point of the grid, $|h(r) - h'(r)| \in \{0,\cdots,k\frac{m+n}{2}\}$. If $m \geq n$, we have $k\frac{m+n}{2} \leq mk$. □

The following lemma (Theorem 3.8 in [1]) is required in the proof of our main result.

Lemma 2.3. [1] *Let $\mathcal{G}$ be the graph whose vertices are height functions of a polycell P and vertex x is adjacent to vertex y if there is a flip changing x to y. The graph $\mathcal{G}$ associated with the ordering $\prec$ is a distributive lattice.*

The following lemma is a re-edition of a result that is implicit in [1]. We mention it for the sake of completeness and since it is used in the proof of the main result.

Lemma 2.4. *There is a path between any two height functions x and y of a generalized tiling.*

Proof. Take the path from x to y passing through $x \wedge y$. □

Lemma 2.5. *If $(x \succ a_1 \succ a_2 \succ .... \succ x \wedge y \prec b_1 \prec b_2 \prec ... \prec y)$, and $(x \succ c_1 \succ c_2 \succ .... \succ x \wedge y \prec d_1 \prec d_2 \prec ... \prec y)$ are two different paths from x to y passing through $x \wedge y$, then they have the same length. Moreover, this length is at most km^2n, where mn is the area of the grid Γ.*

Proof. We know that the path $x, ..., x \wedge y, ..., y$ consists of flipping the point v, where the height of x differs from y. Hence any two such different paths consist of the same points, in different order. Hence they have the same length. To put a bound on the length of the path, it is enough to know that there are at most mn different points in the path. Moreover, by Lemma 2.2, every point can undergo at most km flips. Hence the result. □

Our main result consists of showing that Markov chains of generalized tilings is fast mixing. We use the canonical path argument for this purpose.

The Markov Chain on the distributive lattices of generalized tilings

Let $\mathcal{G}$ be the digraph whose vertices are all the height functions of Γ, and there is an arc $e = (x, y)$ in $\mathcal{G}$ if there is a flip changing x to y. The *Flip Random Walk* on $\mathcal{G}$ is defined as follows. Let x be a vertex of $\mathcal{G}$ of degree d_x and let the vertex y be adjacent to x. Starting at x, move to y with probability $\frac{1}{mn}$, where mn is the dimension of the grid, or remain at x, otherwise. (For $m, n \geq 2$, one may notice that the chain stays at x with probability greater that $1/2$). By Lemma 2.4, the chain is irreducible and the matrix of transitions is symmetric and aperiodic. Hence this random walk is an ergodic and reversible Markov chain, which converges to a uniform distribution. In the sequel of this paper, we show that it is also fast mixing. But, first we need to show that no arc of $\mathcal{G}$ would be congested. Indeed, we have the following result.

Lemma 2.6. *If $e = (z, y)$ is an arc of $\mathcal{G}$, then the number of different canonical paths passing through e is at most N, where N is the number of vertices of $\mathcal{G}$.*

Proof. Among all the edges in $\mathcal{G}$, the ones through which the largest number of canonical paths pass are the edges incident to the vertex $\hat{0}$. Indeed, all the canonical path $(x, ..., \hat{0})$ have to pass through one of these edges incident to $\hat{0}$, because $\hat{0} = x \wedge \hat{0}$. But there are at most $N - 1$ such points x. □

The following is the main theorem.

Theorem 2.7. *The Flip Random Walk on the set of height functions of a Generalized Tiling is a Markov chain that is rapidly mixing and the mixing time $\tau(\epsilon)$ is given by*

$$\tau(\epsilon) \leq (kmn)^3 (mn \ln k + \ln \epsilon^{-1}).$$

Proof. Let x and y be two tilings of Γ. We define the *canonical path* from x to y as the path going from x to y and passing through $x \wedge y$. Such a path necessarily exists, by Lemma 2.3. If there are many such paths, the canonical one is chosen arbitrarily amongst them. The proof uses Theorem 1.1 and consists of showing that there is a flow ϕ such that $\rho(\phi)$ is polynomial in the size of the grid Γ. Indeed, if x and y are two vertices of $\mathcal{G}$, let the canonical path be as defined above. If there are many such paths, the canonical path is chosen arbitrary amongst them. By Lemma 2.5, they have the same length, so the length of the canonical path is well defined. So let $\hat{p}_{xy}$ denote a canonical path from x to y and let ϕ be a flow consisting of injecting $\pi(x)\pi(y)$ units of flow along $\hat{p}_{xy}$. Then, for all arcs e, we have

$$\phi(e) = \sum \pi(x)\pi(y),$$

where the sum is over all the pairs $\{x, y\}$ such that $e \in \hat{p}_{xy}$. Now, since by Lemma 2.6, there are at most N canonical paths through e, and $\pi(x) = \pi(y) = \frac{1}{N}$, since the distribution π is uniform, we have

$$\phi(e) \leq N\pi(x)\pi(y) \leq \frac{1}{N}.$$

Moreover,

$$c(e) = \pi(x)P_{x,y} \geq \frac{1}{Nmn}.$$

Hence

$$\rho(\phi) \leq \frac{max_e \phi(e)}{min_e c(e)} \leq mn. \tag{2}$$

Now, by Lemma 2.5, Theorem 1.1 and Equation (2), we get

$$\tau_x(\epsilon) \leq (mn)(kmn)^2(\ln \frac{1}{\pi(x)} + \ln \frac{1}{\epsilon}).$$

Moreover, it is routine to check that for any k regular tilings, $N \leq k^{mn}$. Indeed, for an mn grid, every k- regular cell can be oriented in at most k different ways. Thus N, the number of different height functions can not exceed k^{mn}. Thus, we have

$$\tau_x(\epsilon) \leq (mn)^3(\ln k^{mn} + \ln \frac{1}{\epsilon}) \leq (kmn)^3(mn \ln k + \ln \epsilon^{-1}).$$

□

Aknowledgements The research of the first author is supported by the School of Mathematics, University of Bristol and the research of second author is supported by SERB-DST, New Delhi.

References

1. O. Bodini and M. Latapy, *Generalized tilings with height function*, Morfismos **7** (2003) 47–68.
2. H. Cohn, R. Kenyon and J. Propp, *A variational principle for domino tilings*, J. Amer. Math. Soc. **14** (2001) 297–346.
3. N. Destainville, *Mixing times of plane random rhombus tilings*, Disc. Math. Theo. Comp. Sci. Proc. AA (DM-CCG), (2001) 1–22.
4. Solomon W. Golomb, *Polyominoes*, Princeton University Press, (1994).
5. M. Jerrum and A. Sinclair, *Approximating the permanent*, SIAM J. Comput. **18** (1989) 1149-1178.
6. P.W. Kasteleyn, *The statistics of dimers on a lattice. I. The number of dimer arrangements on a quadratic lattice*, Physica **27** (1961) 1209–1225.
7. P.W. Kasteleyn, *Graph Theory and crystal physics, in Graph Theory and Theoretical Physics*, Academic Press, New York, (1967) 43–110.
8. K.K. Kayibi and S. Pirzada, *T-tetromino tiling Markov chain is fast mixing*, Theor. Comp. Sci. (2018) **714** (2018) 1–14.
9. K.K. Kayibi and S. Pirzada, *Planarity, symmetry and counting tilings*, Graphs and Comb. **28** (2012) 483-497.

10. K.K Kayibi, S. Pirzada and T.A. Chishti, *Sampling contengency tables*, AKCE Int. J. Graphs Comb. **15**, 3 (2018) 298–306.
11. R. Kenyon, *The planar dimer model with a boundary: A survey*, (1998) Preprint.
12. R. Kenyon, A. Okounkov and S. Sheffield, *Dimers and Amoebae*, Ann. Math. **163** (2006) 1019-1056.
13. M. Korn and I. Pak, *Tiling of rectangles with T-tetrominoes*, Theor. Comp. Sci. **319** (2007) 3-27.
14. M. Luby, D. Randall and A. Sinclair, *Markov chain algorithms for planar lattice structures*, SIAM J. Comput. **31** (2001) 167-192.
15. C. Merino, *On the number of tilings of the rectangular board with T-tetrominoes*, Aust. J. Comb. **14** (2008) 107-114.
16. I. Pak and J. Yang, *Tiling simply connected regions with rectangles*, J. Comb. Theory, Series A **120** (2013) 1804-1816.
17. A. Sinclair, *Convergence rates for Monte Carlo Experiments*, S.G. Whittington (eds.), Numerical Methods for Polymeric Systems, Springer-Verlag, New York, (1998) 1-17.
18. H.N.V. Temperley and M.E. Fisher, *Dimer problem in statistical mechanics-an exact result*, Phi. Mag. **6** (1961) 1061–1063.
19. R. Thomas, *A survey of Pfaffian orientations of graphs*, Proc. Int. Congress of Mathematicians, Madrid (2006) 963-984.
20. D.W. Walkup, *Covering a rectangle with T-tetrominoes*, Amer. Math. Monthly **72** (1965) 986-988.

K.K. Kayibi received his Ph. D. from University of Oxford in 2002.

School of Mathematics, University of Bristol, UK
e-mail: kokokayibi@yahoo.co.uk

U. Samee received her M.Sc. and Ph. D. from University of Kashmir.

Department of Mathematics, Islamia College of Sciences and Commerce, Srinagar, Kashmir, India.
e-mail: drumatulsamee@gmail.com

Merajuddin received his M.Sc. and Ph. D. from Indian Institute of Technology, Kanpur, India.

Department of Applied Mathematics, AMU, Aligarh, India
e-mail: meraj1957@rediffmail.com

S. Pirzada received his M.Sc. from University of Kashmir (1990) and Ph. D. from Aligarh Muslim University, India (1995).

Department of Mathematics, University of Kashmir, Srinagar, Kashmir, India.
e-mail: pirzadasd@kashmiruniversity.ac.in