Sampling Biased Lattice Configurations using Exponential Metrics

Sam Greenberg* Amanda Pascoe* Dana Randall†

Abstract

Monotonic surfaces spanning finite regions of $\mathbb{Z}^d$ arise in many contexts, including DNA-based self-assembly, card-shuffling and lozenge tilings. We explore how we can sample these surfaces when the distribution is biased to favor higher surfaces. We show that a natural local chain is rapidly mixing with any bias for regions in $\mathbb{Z}^2$, and for bias $\lambda > d^2$ in $\mathbb{Z}^d$, when $d > 2$. Moreover, our bounds on the mixing time are optimal on d-dimensional hyper-cubic regions. The proof uses a geometric distance function and introduces a variant of path coupling in order to handle distances that are exponentially large.

1 Introduction

Random sampling of lattice configurations, including tilings and colorings, is used to study properties of various computational, combinatorial and physical models. In statistical physics, for example, domino tilings of the Cartesian lattice and lozenge (rhombus) tilings of the triangular lattice are natural models of diatomic molecules. Sampling provides insight into the thermodynamics properties of these systems. Similarly, 3-colorings of lattice regions represent states of the zero temperature Potts model, a popular model of ferromagnetism. Local Markov chains that either update a single site of a coloring, or rotate 2 or 3 nested tiles in a tiling, can be used to generate random configurations.

Luby, Randall and Sinclair [11] analyzed a family of nonlocal, Markov chains for each of those sets of configurations and showed that they are rapidly mixing (i.e., converging rapidly to equilibrium). Subsequently, Randall and Tetali [15] showed that bounds on the mixing times of the nonlocal chains imply that the local chains are also rapidly mixing. The key to the analysis was the observation that in each case the planar coloring or tiling could be mapped to a 3-dimensional monotonic surface [11]. Accordingly, steps of the local Markov chains on these planar configurations can be viewed as perturbing a 3-dimensional surface by inverting a locally convex region to a concave one, or vice versa. For example, Figure 1 shows a two-dimensional lozenge tiling. Our eyes naturally interpret this as a 3-dimensional picture where we are viewing the tops of a set of cubes. The local Markov chain that rotates three closely packed lozenges by 180 degrees can be interpreted as adding or removing a single cube from the 3-dimensional structure.

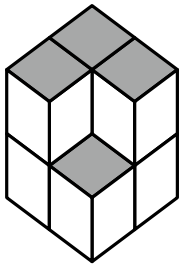

Figure 1: A monotonic surface in three dimensions.

There is a natural generalization of the monotonic surfaces arising in the context of lozenge tilings to any number of dimensions. Informally, in two dimensions the monotonic surface is just a "staircase walk" going to the right and down as in Figure 2. The local Markov chain adds or removes individual squares adjacent to the boundary. It is easy to show that this Markov chain is rapidly mixing using a straight-forward coupling argument (see, e.g., [2] for background on coupling). In higher dimensions, d-dimensional monotonic surfaces correspond bijectively with $(d-1)$-dimensional tilings [14], although the convergence rate of the corresponding local Markov chain is unknown for $d > 3$.

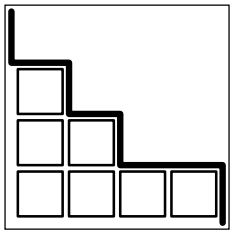

Figure 2: A monotonic surface in two dimensions.

*School of Mathematics, Georgia Institute of Technology, Atlanta GA, 30332-0160.

†School of Computer Science, Georgia Institute of Technology, Atlanta GA, 30332-0765. Supported in part by NSF grants CCR-0515105 and DMS-0505505.

Recently there has been a lot of interest in a *biased* version of this local Markov chain, where it is more likely to add unit cubes than remove them. More precisely, let $P(\sigma, \tau)$ be the transition probability of moving from σ to τ in one move for a pair of configurations σ_t and τ_t that differ by a single move. If τ is the configuration formed by adding one cube to σ, then $\lambda = P(\sigma, \tau)/P(\tau, \sigma)$ is the *bias* of the chain. Using detailed balance, it is easy to see that the stationary probability of a configuration σ will be proportional to $\lambda^{|\sigma|}$, where $|\sigma|$ is the number of unit cubes defining the surface σ. This biased version of the Markov chain arises in the context of nanotechnology [12] and biased card shuffling [3], while a biased version of the chain for sampling 3-colorings comes up in the context of asynchronous cellular automata [5]. We restrict our attention here to the first two examples where the monotonic surfaces correspond to sets of supported cubes, although the methodology also applies in the third setting.

The nanotechnology example that motivates our work is a model for DNA-based self-assembly. Roughly "square" shaped tiles are constructed from strands of DNA so that each side of the square is single-stranded. Specified pairs of tiles are encouraged to line up and attach along edges by encoding their corresponding sides with complementary sequences of base pairs. At appropriately chosen temperatures, these tiles will have a good chance of assembling according to these prescribed rules, although they also have a chance of disassociating and breaking apart. See, e.g., [8, 17, 19, 20] for more details. The model considered by Majumder et al. [12] allows the left column and bottom row of a large square to form first, and then iteratively allows tiles to associate with the large substrate if their left and bottom neighbors are already present. Likewise, tiles can disassociate if their upper and right neighbors are not present, although disassociation happens at a lower rate. The dynamics of this model are precisely captured by the local Markov chain on 2-dimensional monotonic surfaces and the chain must be rapidly mixing if the substrate is to efficiently self-assemble, as required. The 3-dimensional analogue is also used to study self-assembly, where now tiles are shaped like cubes and complementary sequences are used to encourage corresponding faces to attach.

Benjamini et al. [3] studied the 2-dimensional biased chain in order to analyze a biased card shuffling algorithm that favors putting each pair of cards in the lexicographically correct order. They gave optimal $O(n)$ bounds on the mixing rate of the biased chain on $\sqrt{n} \times \sqrt{n}$ regions of $\mathbb{Z}^2$ for any constant bias. In three and higher dimensions, substantially less is known. Majumder et al. [12] showed that the chain mixes quickly when the bias is $\Theta(n)$; apparently the case of large bias is the most interesting for nanotechnology applications. Nothing else is known about the convergence of the biased chain. Both of these results are highly technical and do not readily generalize to other values of the bias or other dimensions.

We make progress in several aspects of the problem of sampling biased surfaces. In two dimensions, we show that the biased chain is rapidly mixing for any bias on a large family of simply-connected regions, even when the bias is arbitrarily close to one. Our proof is significantly simpler than the arguments of Benjamini et al., while achieving the same optimal bounds on the mixing time for square regions when the bias is constant. We also show the chain is rapidly mixing on d-dimensional lattice regions provided the bias $\lambda \geq d^2$, for $d > 2$. Again, our bounds on the mixing time are optimal when the regions are hyper-cubes and we show the chain is rapidly mixing for a large family of simply-connected regions in $\mathbb{Z}^d$.

The key observation underlying these results is that we can define an exponential metric on the state space such that the distance between pairs of configurations is always decreasing in expectation. We then show how to modify the path coupling theorem to handle the case when the distances are exponentially large and the expected change in distance is small during moves of the coupled chain. We believe that this new theorem is of independent interest.

We note that our method for bounding the mixing rate of the biased chain can also be generalized to Markov chains for other families of monotonic surfaces. A particularly intriguing example is "biased 3-colorings." Three-colorings of finite regions of the d-dimensional Cartesian lattice are known to correspond to a family of $(d+1)$-dimensional surfaces (see [11] for the bijection in two dimensions; the higher dimensional mapping follows similarly). Again, there is a local Markov chain that perturbs each surface locally. This chain is known to be rapidly mixing in two and three dimensions in the unbiased case [11, 15]. However, it is also known that in sufficiently high dimensions, the unbiased chain requires exponential time to converge [9, 10]. The mixing time of the unbiased chain is unknown in all intermediate dimensions. A biased version of this Markov chain is motivated by a model for asynchronous cellular automata and favors moves that increase the height of the surface [5]. The arguments introduced in this paper can be extended to this second model of monotonic surfaces, with nearly identical results. We leave the details for the full version of the paper.

The remainder of this paper is organized as follows. In Section 2 we formalize the model and the Markov chain. In Section 3 we review the path coupling method and introduce the modified path coupling theorem that can be more appropriate when distances are exponentially large. Finally, in Section 4, we show how we can use our new path coupling theorem to conclude that the chain on biased surfaces is rapidly mixing.

2 Monotonic surfaces

Let $R \subset \mathbb{Z}^d$ be the union of the corners of a (face) connected set of unit cubes. Given R, we let $\widehat{R} \in \mathbb{R}^d$ be the real region that is the union of all the points in the corresponding real cubes. If $\widehat{R}$ is simply-connected, then we say the discrete set R is also. The set of edges of R is just the union of edges of the cubes defining the region and the boundary of R is the set of vertices in R that lie on the boundary of the real region $\widehat{R}$.

We first restrict our attention to 2-dimensional regions. A *monotonic surface (or path)* in R is a path starting and ending on the boundary of R that only takes steps down and to the right and is composed entirely of edges in R. Such a path is illustrated in Figure 2 when R is a 4×4 square. Notice that any monotonic surface can be interpreted as the upper boundary of a set of unit squares, where each square in the set is supported below or to the left by other squares in the set or the boundary of R.

Likewise, in three dimensions a monotonic surface in R is the union of two-dimensional faces such that any cross-section along an axis-aligned plane is a two-dimensional monotonic surface. Such a surface is illustrated in Figure 1 when R is a $2 \times 2 \times 2$ region.

In general, we consider simply-connected regions $R \subset \mathbb{Z}^d$ composed of unit hyper-cubes. A d-dimensional monotonic surface is a set of $(d-1)$-dimensional faces such that any cross-section along an axis-aligned $(d-1)$-dimensional hyper-plane is a $(d-1)$-dimensional monotonic surface.

We will restrict our focus to a family of simply-connected regions that have favorable properties for the purposes of sampling monotonic surfaces. The following definitions will allow us to define these regions.

DEFINITION 2.1. *Let $\overline{u}^* = (1, 1, ..., 1) \in \mathbb{Z}^d$. For $v \in \mathbb{Z}^d$, we define the* real ray *$\widehat{r}(v)$ to be the set*

$$\widehat{r}(v) = \{v + k\bar{u}^* : k \in \mathbb{R}\},$$

and we define the discrete ray *$r(v)$ to be the set*

$$r(v) = \widehat{r}(v) \bigcap \mathbb{Z}^d.$$

DEFINITION 2.2. *A d-dimensional simply-connected region $R \subset \mathbb{Z}^d$ is* nice *if, for all $v \in R$, the set $\widehat{R} \bigcap \hat{r}(v)$ is connected.*

Note that all hyper-rectangular regions are nice.

2.1 The biased Markov chain In order to formalize the interpretation of monotonic surfaces in terms of sets of supported cubes, it will be convenient to represent the surfaces in terms of *downsets*. Let $\overline{u}_i = (0, 0, \ldots, 1, \ldots, 0)$ be the unit vector in the ith direction, and recall that $\overline{u}^* = (1, \ldots, 1)$. Given a nice region R, we let $R_L = \{v \in R \text{ such that } v - \overline{u}^* \notin R\}$ be the lower envelope of the region.

DEFINITION 2.3. *Let $R \subset \mathbb{Z}^d$ be a simply-connected region consisting of the union of unit cubes. A* downset *is a subset $\sigma \subseteq R$, with $R_L \subseteq \sigma$, such that for any i, if $v \in \sigma$ and $v - \overline{u}_i \in R$, then $v - \overline{u}_i \in \sigma$.*

For a nice region R, we define the state space Ω_{mon} to be the set of all downsets of R. The following definition helps us formalize the Markov chain we will be using to sample from Ω_{mon}.

DEFINITION 2.4. *Let R be any nice region and let σ be any downset of R. We say the* upper boundary *of σ is $\partial(\sigma) = \{v \in \sigma \text{ such that } v + \overline{u}^* \notin \sigma\}$.*

Notice that, just as the downset represents the vertices of the cubes lying below a monotonic surface, the upper boundary is the set of vertices that lie on the surface.

We are now ready to describe the Markov chain on Ω_{mon}. For simplicity, we start by defining the *unbiased* chain $\widehat{\mathcal{M}}_{mon}$ that converges to the *uniform distribution* over monotonic surfaces Ω_{mon}. Start at an arbitrary downset, e.g., $\sigma_0 = R_L$ and repeat the following steps. If we are at a downset σ_t at time t, pick a point $v \in \partial(\sigma)$ and an integer $b \in \pm 1$ uniformly at random. If $b = +1$, let $\sigma_{t+1} = \sigma_t \cup (v + \overline{u}^*)$, if this is a valid downset. If $b = -1$, let $\sigma_{t+1} = \sigma_t \setminus \{v\}$, if this is a valid downset. In all other cases, keep σ_t unchanged so that $\sigma_{t+1} = \sigma_t$.

LEMMA 2.1. *For any nice region R, the Markov chain $\widehat{\mathcal{M}}_{mon}$ connects the state space Ω_{mon}.*

Proof. Let σ be any downset and let v^{max} be any point in σ such that $\sum_i v_i^{max}$ is maximized. We can always remove v^{max} and move to $\sigma' = \sigma \setminus v^{max}$ without violating the downset condition because the maximality of v^{max} tells us that $v^{max} + u_i \notin \sigma$, for all i. Thus, from any valid downset σ we can always remove points and get to the "lowest" downset R_L. Also, such a sequence of steps can be reversed to move from R_L to any other downset ρ.

It is important to notice that for any downset σ and point $v \notin \sigma$, if $\sigma \cup v$ is a valid downset, then

$|\partial(\sigma)| = |\partial(\sigma \cup v)|$. This is because $\partial(\sigma \cup \{v\}) = \partial(\sigma) \cup \{v\} \setminus \{v - \overline{u}^*\}$. It follows that for any nice region R, the size of the boundary of each valid downset is fixed. This observation motivates the following two definitions that will be convenient when we state the mixing time of our Markov chain.

DEFINITION 2.5. *The* span *of a nice region R is $\alpha = |\partial(\sigma)|$, where α is any downset of R.*

DEFINITION 2.6. *Let R be any nice region. The* stretch *of R is*

$$\max_{x \in R} x \cdot \overline{u}^* - \min_{y \in R} y \cdot \overline{u}^*,$$

where $\cdot$ is the dot product.

Consider, for example, when R is an $h \times \cdots \times h$ region in $\mathbb{Z}^d$. Then the span is $\alpha = dh^{d-1}$ and the stretch is $\gamma = dh$.

Since we have shown that the moves of $\widehat{\mathcal{M}}_{mon}$ connect the state space and all valid moves have the same transition probabilities, we can conclude that the chain converges to the uniform distribution over downsets in Ω_{mon}. We now define the *biased* Markov chain by using Metropolis-Hastings probabilities [13] to modify the transition probabilities so that we converge to the desired distribution on biased surfaces. This new chain connects the state space by the same argument as in Lemma 2.1.

The Markov chain $\mathcal{M}_{mon}$ with bias $\lambda > 1$.

`Starting at any` σ_0`, iterate the following:`

- `Choose` (v, b, p) `uniformly at random from` $\partial(\sigma_t) \times \{+1, -1\} \times (0, 1)$.
- `If` $b = +1$`, let` $\sigma_{t+1} = \sigma \cup \{v + \overline{u}^*\}$ `if it is a valid downset.`
- `If` $b = -1$ `and` $p \leq \frac{1}{\lambda}$`, let` $\sigma_{t+1} = \sigma \setminus \{v\}$ `if it is a valid downset.`
- `Otherwise let` $\sigma_{t+1} = \sigma_t$.

The biased Markov chain $\mathcal{M}_{mon}$ converges to the correct distribution on Ω_{mon} by the detailed balance condition. Moreover, notice that for any given v there is at least one choice of b, corresponding to adding or removing a vertex, that does not result in a valid downset. Therefore, for all $\sigma_t \in \Omega_{mon}$, we have $\mathbb{P}[\sigma_{t+1} = \sigma_t] \geq 1/2$. A chain like this with a self-loop probability of at least $1/2$ at each step is called *lazy*, and we will use this property later in our analysis.

2.2 Our results We give conditions under which we can prove $\mathcal{M}_{mon}$ converges quickly to the stationary distribution. First, we review some Markov chain terminology.

Let $\mathcal{M}$ be a Markov chain with transition matrix $P(\cdot, \cdot)$, state space Ω, and stationary distribution π. The *mixing time* is defined as follows.

DEFINITION 2.7. *For $\varepsilon > 0$, the* mixing time $\tau(\varepsilon)$ *is*

$$\tau(\varepsilon) = \min\{t : \|P^{t'}, \pi\|_{tv} \leq \varepsilon, \forall t' \geq t\},$$

where

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

We say a Markov chain is *rapidly mixing* if the mixing time is bounded above by a polynomial in n and $\log \epsilon^{-1}$, where n is the size of each configuration in the state space.

We may now present our main results.

THEOREM 2.1. *Let R be any nice d-dimensional region with volume n, span α, stretch γ, and bias λ. Let $\widehat{\lambda}_d = \left(\frac{2}{d - \sqrt{d^2 - 4}}\right)^2$.*

1. *If $\lambda - \widehat{\lambda}_d > c$, for some positive constant c, then the mixing time of $\mathcal{M}_{mon}$ satisfies*

$$\tau(\varepsilon) = O\left(\alpha(\gamma + \ln n) \ln \varepsilon^{-1}\right).$$

2. *If $\lambda - \widehat{\lambda}_d > 0$ is less than any constant, then the mixing time of $\mathcal{M}_{mon}$ satisfies*

$$\tau(\varepsilon) = O\left(\alpha(\gamma + \ln n)^2) \ln(\varepsilon^{-1})\right).$$

Note that for all nice regions, we have $\alpha, \gamma \leq n$, so the mixing time of $\mathcal{M}_{mon}$ is always polynomially bounded for the values of λ given in the theorem.

When $d = 2$, Theorem 2.1 tells us the biased chain is rapidly mixing for all $\lambda > 1$. In general, for $d \in \mathbb{Z}^+$, we have $\left(\frac{2}{d - \sqrt{d^2 - 4}}\right)^2 < d^2$, so in any dimension $d > 2$ we find that the chain is rapidly mixing when $\lambda \geq d^2$. We conjecture the chain is rapidly mixing for all values of $\lambda > 1$, in all dimensions d, but do not yet have a proof for small values of λ in dimensions higher than 2.

When R is the $h \times h \times \cdots \times h$ hyper-cube, with $h^d = n$, Theorem 2.1 gives optimal bounds on the mixing time. To see this, recall that in this case $\alpha = dh^{d-1}$ and $\gamma = dh$. This gives the following corollary.

COROLLARY 2.1. *Let R be the d-dimensional $h \times h \times \ldots \times h$ hyper-cube with $h^d = n$. If $d = 2$ and $\lambda > 1$ is constant or if $d \geq 3$ and $\lambda \geq d^2$, then the mixing time of $\mathcal{M}_{mon}$ satisfies*

$$\tau(\varepsilon) = O(d^2 n \ln \varepsilon^{-1}).$$

This matches the bounds of Benjamini et al. for square regions in $\mathbb{Z}^2$, and generalizes their result to higher dimensions by giving optimal bounds on the mixing time for large enough λ.

3 Path Coupling

We will use a coupling argument to prove Theorem 2.1 bounding the mixing rate of the biased Markov chain. A *coupling* of a chain $\mathcal{M}$ is a joint Markov process on $\Omega \times \Omega$ such that the marginals each agree with $\mathcal{M}$ and, once the two coordinates coalesce, they move in unison thereafter. The coupling lemma bounds the mixing time in terms of the expected time of coalescence of any coupling.

DEFINITION 3.1. *For initial states x, y let*

$$T^{x,y} = \min\{t : X_t = Y_t \mid X_0 = x,\, Y_0 = y\},$$

and define the coupling time *to be $T = \max_{x,y} \mathbb{E}[T^{x,y}]$.*

The following result relates the mixing time and the coupling time.

THEOREM 3.1. **(Aldous [2])** *For any coupling with coupling time T, the mixing time satisfies*

$$\tau(\epsilon) \leq \lceil T \mathrm{e} \ln \epsilon^{-1} \rceil.$$

The goal, then, is to define a good coupling and show that the coupling time is polynomially bounded. Path coupling [4, 7] is a convenient way of establishing such a bound by only considering a subset of the joint state space $\Omega \times \Omega$. The following version is convenient.

THEOREM 3.2. **(Dyer and Greenhill [7])** *Let φ be an integer valued metric defined on $\Omega \times \Omega$ which takes values in $\{0, \ldots, B\}$. Let U be a subset of $\Omega \times \Omega$ such that for all $(x_t, y_t) \in \Omega \times \Omega$ there exists a path $x_t = z_0, z_1, \ldots, z_r = y_t$ between x_t and y_t such that $(z_i, z_{i+1}) \in U$ for $0 \leq i < r$ and*

$$\sum_{i=0}^{r-1} \varphi(z_i, z_{i+1}) = \varphi(x_t, y_t).$$

Let $\mathcal{M}$ be a Markov chain on Ω with transition matrix P. Consider any random function $f : \Omega \to \Omega$ such that $\mathbb{P}[f(x) = y] = P(x, y)$ for all $x, y \in \Omega$, and define a coupling of the Markov chain by $(x_t, y_t) \to (x_{t+1}, y_{t+1}) = (f(x_t), f(y_t))$.

1. *If there exists $\beta < 1$ such that*

$$\mathbb{E}[\varphi(x_{t+1}, y_{t+1})] \leq \beta \varphi(x_t, y_t),$$

for all $(x_t, y_t) \in U$, then the mixing time satisfies

$$\tau(\epsilon) \leq \frac{\ln(B\epsilon^{-1})}{1 - \beta}.$$

2. *If $\beta = 1$ (i.e., $\mathbb{E}[\Delta\varphi(x_t, y_t)] \leq 0$, for all $x_t, y_t \in U$), let $\alpha > 0$ satisfy $Pr[\varphi(x_{t+1}, y_{t+1}) \neq \varphi(x_t, y_t)] \geq \alpha$ for all t such that $x_t \neq y_t$. The mixing time of $\mathcal{M}$ then satisfies*

$$\tau(\epsilon) \leq \left\lceil \frac{\mathrm{e} B^2}{\alpha} \right\rceil \lceil \ln \epsilon^{-1} \rceil.$$

To understand why it is difficult to use coupling to prove Theorem 2.1, we first examine the straightforward coupling of (σ_t, ρ_t) with a natural distance function. Recall that a move of $\mathcal{M}_{mon}$ is defined by choosing (v, b, p), where $v \in \partial(\sigma_t)$. Equivalently, we can choose v by selecting a ray v^* uniformly, since for each σ_t each ray has a unique intersection with $\partial(\sigma_t)$. The coupling simply chooses the same (v^*, b, p) to generate both σ_{t+1} and ρ_{t+1}. The natural distance metric on $\Omega_{mon} \times \Omega_{mon}$ is the Hamming distance, where $h(\sigma_t, \rho_t) = |\sigma_t \oplus \rho_t|$, and $\oplus$ is the symmetric difference. Unfortunately, we encounter difficulties using this coupling and metric that will be useful to illustrate.

Examine the pair of downsets in Figure 3. They differ at a single point, so $h(\sigma_t, \rho_t) = 1$. In order to use Theorem 3.2, we need for the distance to be non-increasing in expectation, i.e., that $\mathbb{E}[h(\sigma_{t+1}, \rho_{t+1})] \leq h(\sigma_t, \rho_t)$. However, for this pair of downsets this will not be the case. There are two moves that increase the distance, each occurring with probability $1/(2\alpha)$, and two moves that decrease the distance, one occurring with probability $1/(2\alpha)$ and the other occurring with probability $1/(2\alpha\lambda)$. This is sufficient for a coupling argument when $\lambda = 1$ since the expected change in distance would be 0. Unfortunately, when $\lambda > 1$ the expected change in distance is positive.

In higher dimensions, the situation becomes even worse. For the pair of 3 dimensional downsets in Figure 4, there are *three* moves which increase the Hamming distance with probability $1/(2\alpha)$ and only two moves that potentially decrease the distance, one occurring with probability $1/(2\alpha)$ and the other with probability $1/(2\alpha\lambda)$. In d dimensions there are pairs of configurations at Hamming distance one so that d moves that increase the Hamming distance and only two moves that bring the configurations together.

One promising remedy is to alter the distance metric. When two downsets σ_t and ρ_t differ at a point x,

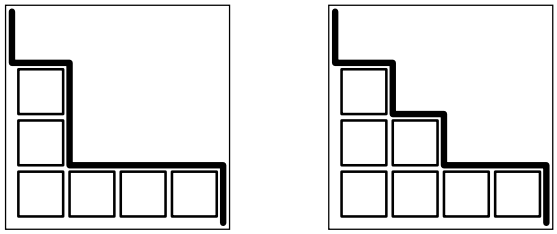

Figure 3: A pair of downsets σ_t (left) and ρ_t (right) where $\rho_t = \sigma_t \cup \{(1,1)\}$.

with $\rho_t = \sigma_t \cup \{x\}$, the two moves which decrease the distance involve adding x to σ_t or removing x from ρ_t. The moves that increase the distance involve adding $x + \overline{u}_i$ to ρ_t, for some i or removing $x - \overline{u}_i$ from σ_t, although the later case occurs with smaller probability. Therefore, we consider a distance metric that counts the distance between two sets that differ on x as greater than the distance between two sets that differ on $x+\overline{u}_i$.

We find the following distance metric does achieve the desired contraction property for appropriate values of λ in order to bound the coupling time. First, let $m = \max_{x \in R} x \cdot \overline{u}^*$. Then, for two downsets σ and ρ, we define the distance function ϕ to be

$$\phi(\sigma, \rho) = \sum_{x \in \sigma \oplus \rho} (\sqrt{\lambda})^{m - x \cdot \overline{u}^*}.$$

We present the proof that this metric is indeed decreasing in expectation at every step in Section 4. Unfortunately, however, this definition of the distance metric introduces new difficulties, especially when λ is very close to 1. First, the distances might now take on non-integer values, while the Path Coupling Theorem requires integer valued metrics. In fact, if this restriction is merely removed, then the theorem is no longer true as the distances might get smaller and smaller without coalescence occurring in a polynomial number of steps. However, it turns out to be enough to add an additional condition requiring that no pairs of configurations have a distance within the open interval $(0, 1)$. Notice that our definition of ϕ satisfies this restriction as the distance between any two downsets is either 0 or at least 1.

The second, more serious concern is that the maximum distance between two configurations can be exponentially large in n while we are only guaranteed that the distance is decreasing in expectation. If the magnitude of the expected change is small (polynomial), then we cannot expect the distance to be zero in only a polynomial number of steps. Consider, for example, when λ is very close to one in $\mathbb{Z}^2$. We can find configurations x_t and y_t so that $\mathbb{E}[\varphi(x_{t+1}, y_{t+1})] \leq (1 - 2^{-n})\varphi(x_t, y_t)$, so

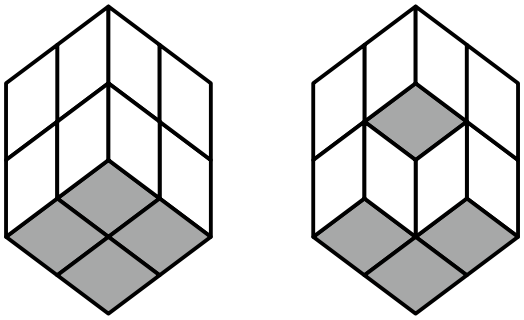

Figure 4: A pair of downsets σ_t (left) and ρ_t (right) where $\rho_t = \sigma_t \cup \{(0,0,0)\}$.

$1 - \gamma \leq 2^{-n}$. Therefore the expected change is too small to apply the first part of Theorem 3.2. Moreover, the maximum distance B is very large, so we cannot get a good bound on the mixing time using the second part of Theorem 3.2 either since the bound on the mixing time depends quadratically on B.

3.1 Path coupling with exponential metrics

The following modification of the Path Coupling Theorem allows us to handle cases when the distances can be exponentially large and the expected change in distance is small (or even zero). We show that it suffices to show that the expected change in the absolute value of the distance is proportional to the current distance, and with this condition the mixing time is polynomially bounded. We apply this new theorem to the biased Markov chain $\mathcal{M}$ in Section 4.

THEOREM 3.3. *Let $\phi : \Omega \times \Omega \to \mathbb{R}^+ \cup \{0\}$ be a metric that takes on finitely many values in $\{0\} \cup [1, B]$. Let U be a subset of $\Omega \times \Omega$ such that for all $(X_t, Y_t) \in \Omega \times \Omega$ there exists a path $X_t = Z_0, Z_1, \ldots, Z_r = Y_t$ such that $(Z_i, Z_{i+1}) \in U$ for $0 \leq i < r$ and $\sum_{i=0}^{r-1} \phi(Z_i, Z_{i+1}) = \phi(X_t, Y_t)$.*

Let $\mathcal{M}$ be a lazy Markov chain on Ω and let (X_t, Y_t) be a coupling of $\mathcal{M}$, with $\phi_t = \phi(X_t, Y_t)$. Suppose there exists $\beta \leq 1$ such that, for all $(X_t, Y_t) \in U$,

$$\mathbb{E}[\phi_{t+1}] \leq \beta \phi_t.$$

1. *If $\beta < 1$, then the mixing time satisfies*

$$\tau(\varepsilon) \leq \frac{\ln(B\varepsilon^{-1})}{1 - \beta}.$$

2. *If there exists $\kappa, \eta \in (0, 1)$ such that $\mathbb{P}\left[|\phi_{t+1} - \phi_t| \geq \eta \phi_t\right] \geq \kappa$ for all t provided that $X_t \neq Y_t$, then*

$$\tau(\varepsilon) \leq \left\lceil \frac{e \ln^2(B)}{\ln^2(1+\eta)\kappa} \right\rceil \left\lceil \ln \varepsilon^{-1} \right\rceil.$$

Again, there are two main differences between Theorem 3.2 and Theorem 3.3. The first is that Theorem 3.3 allows for non-integer metrics, provided that ϕ takes on a finite set of real values and we have the condition that for all $X, Y \in \Omega$, $\phi(X,Y) < 1$ implies $\phi(X,Y) = 0$. This requires merely a minor change to the proof of Theorem 3.2 [7]. We use the fact that $\phi_t \notin (0,1)$ to ensure that

$$\mathbb{E}[\phi_t] \geq \mathbb{P}(X_t \neq Y_t),$$

even though the values of ϕ_t might not be integral.

The second difference is that β can be close to (or equal to) one while B can be exponentially large, the case in which both parts of Theorem 3.2 are insufficient for deriving polynomial bounds. We modify the original proof of the second half of the theorem, essentially replacing the original distance $\phi(X_t, Y_t)$ with $\ln(\phi(X_t, Y_t))$. There are some technical lemmas concerning the expectation and variance of the logarithm, which we present below, but the novelty of Theorem 3.3 is more in the statement of the result than a new method of proof.

Note that including this case of $\beta = 1$ and exponential B requires a strong bound on the variance of ϕ_t. Without this bound on variance, Theorem 3.3 is not true; if $\phi_0 = 2^n$ and $\phi_{t+1} = \phi_t - 1$ for all $t \geq 1$, then clearly it will take time exponential in n for $\phi_t = 0$.

3.2 Proof of Theorem 3.3 In order to prove the new path coupling theorem to handle exponential metrics, we define a new variable ψ, which is essentially $\ln(\phi)$. If we hope to prove a chain is rapidly mixing by looking at $\ln(\phi)$, we need to bound the time to reach $\ln(0) = -\infty$, and the expected time could be unbounded. In particular, in order to bound the convergence time, we need that the sequence $\{\psi_t\}$ has bounded differences. We fix this using the assumption that $\phi_t \notin (0,1)$, so we need only to bound the time until we reach a negative value for $\ln(\phi_t)$. Accordingly, we define

$$\psi_t = \begin{cases} \ln(\phi_t) & \text{if } \phi_t > 0, \\ -2\ln 2 & \text{if } \phi_t = 0. \end{cases}.$$

This means that $\psi_t \in [-2\ln 2, \ln(B)]$. The particular value for ψ_t when $\phi_t = 0$ is chosen to satisfy the following lemmas. In Lemma 3.1 below, we show that if the expected distance is non-increasing, then the expected value for ψ_{t+1} is also non-increasing. In Lemma 3.2, we show that if that the variance of the distance is proportional to the current distance, then the variance of ψ_t is at least a constant. Together, these lemmas suffice to prove Theorem 3.3 following the arguments in the proof of [11] exactly.

LEMMA 3.1. *With ψ_t defined in terms of ϕ_t as above, if ϕ satisfies*

$$\mathbb{E}[\phi_{t+1} - \phi_t] \leq 0,$$

then ψ satisfies

$$\mathbb{E}[\psi_{t+1} - \psi_t] \leq 0.$$

Proof. If $\phi_t = 0$, then $\phi_{t+1} = 0$, and the lemma is trivially true. Therefore we assume $\phi_t \neq 0$.

Given the value of ϕ_t, let $\{r_0, r_1, r_2, \ldots, r_N\}$ be the possible values for ϕ_{t+1}, each occurring with probability $\{\zeta_0, \zeta_1, \zeta_2, \ldots, \zeta_N\}$. That is, $\mathbb{P}[\phi_{t+1} = r_i | \phi_t] = \zeta_i$, with $\sum_{i=0}^{N} \zeta_i = 1$. Assume $r_0 = 0$.

As our chain is lazy, $\mathbb{P}[\phi_{t+1} = \phi_t] \geq 1/2$. Therefore we find that $\zeta_0 \leq 1/2$. Now,

$$\begin{aligned}
\mathbb{E}[\psi_{t+1}|\psi_t] &= \zeta_0(-2\ln 2) + \sum_{i=1}^{N} \zeta_i \ln_2(r_i) \\
&= -2\ln 2\zeta_0 + \ln\left(\prod_{i=1}^{N} r_i^{\zeta_i}\right) \\
&\leq -2\ln 2\zeta_0 + \ln\left(\frac{\sum_{i=1}^{N} \zeta_i r_i}{1-\zeta_0}\right) \\
&= \ln\left(\mathbb{E}[\phi_{t+1}|\phi_t]\right) - 2\ln 2\zeta_0 - \ln(1-\zeta_0) \\
&\leq \ln\left(\mathbb{E}[\phi_{t+1}|\phi_t]\right) \\
&\leq \ln \phi_t = \psi_t,
\end{aligned}$$

where the first inequality is by the Arithmetic-Geometric Mean Inequality and the second follows from the fact that $(\ln(1-\zeta_0))/\zeta_0 \geq -2\ln 2$ for $\zeta_0 \in (0, \frac{1}{2})$.

LEMMA 3.2. *If there exist constants $\kappa, \eta \in (0,1)$ such that*

$$\mathbb{P}[|\phi_{t+1} - \phi_t| \geq \eta\phi_t] \geq \kappa,$$

for $\phi_t \neq 0$, then

$$\mathbb{E}[(\psi_{t+1} - \psi_t)^2] \geq \ln^2(1+\eta)\kappa.$$

Proof. Let $\zeta_0 = \mathbb{P}[\phi_{t+1} = 0]$ and let A be the event that $\phi_{t+1} \neq 0$. From our assumption, we have that

$$\begin{aligned}
\kappa &\leq \mathbb{P}[|\phi_{t+1} - \phi_t| \geq \eta\phi_t] \\
&= 1 \cdot \zeta_0 + \mathbb{P}[|\phi_{t+1} - \phi_t| \geq \eta\phi_t \mid A\ ]\ (1-\zeta_0).
\end{aligned}$$

Now, because $\psi_{t+1} = \ln(\phi_{t+1})$ when $\phi_{t+1} \neq 0$, condi-

tioning on A we find

$$\begin{aligned}
&\mathbb{P}[|\phi_{t+1}-\phi_t| \geq \eta\phi_t \mid A\,] \\
&\quad = \mathbb{P}\left[\frac{\phi_{t+1}}{\phi_t} - 1 \geq \eta \mid A\right] \;+ \\
&\qquad\qquad \mathbb{P}\left[\frac{\phi_{t+1}}{\phi_t} - 1 \leq -\eta \mid A\right] \\
&\quad = \mathbb{P}\left[\psi_{t+1}-\psi_t \geq \ln(1+\eta) \mid A\,\right] \;+ \\
&\qquad\qquad \mathbb{P}\left[\psi_{t+1}-\psi_t \leq \ln(1-\eta) \mid A\,\right] \\
&\quad \leq \mathbb{P}\left[|\psi_{t+1}-\psi_t| \geq \ln(1+\eta) \mid A\,\right].
\end{aligned}$$

Since $\phi_t \geq 1$, we have $\psi_t \geq 0$, so $|-2\ln 2 - \psi_t| \geq 2\ln 2$. Now let $m = \ln^2(1+\eta)$, so $m < \ln^2(2)$ since $\eta < 1$. This yields

$$\begin{aligned}
&\mathbb{E}[(\psi_{t+1}-\psi_t)^2] \\
&\quad = (-2\ln 2 - \psi_t)^2\zeta_0 \;+ \\
&\qquad \sum_{\ell\in\Omega,\ell\neq 0} \left(\ln(\ell) - \ln\phi_t\right)^2 \mathbb{P}[\phi_{t+1}=\ell] \\
&\quad \geq (2\ln 2)^2\zeta_0 \;+ \\
&\qquad m\mathbb{P}\left[(\psi_{t+1}-\psi_t)^2 \geq m \mid A\,\right](1-\zeta_0) \\
&\quad \geq (2\ln 2)^2\zeta_0 + m\left(\frac{\kappa-\zeta_0}{1-\zeta_0}\right)(1-\zeta_0) \\
&\quad = \left((2\ln 2)^2 - m\right)\zeta_0 + m\kappa \\
&\quad > m\kappa + 3\zeta_0\ln^2 2 \\
&\quad > \ln^2(1+\eta)\kappa.
\end{aligned}$$

We are now ready to prove Theorem 3.3 by following the method of [11], replacing ϕ_t with ψ_t.

Proof of Theorem 3.3. Part (1) follows directly from the proof of Theorem 3.2 in [7], while allowing for non-integer valued metrics.

For part (2), define the process

$$Z(t) := (\ln B - \psi_t)^2 - Qt,$$

where $Q = \ln(1+\eta)^2\kappa - \frac{1}{2}$. Examining the expected difference between $Z(t)$ and $Z(t+1)$, we have

$$\begin{aligned}
&\mathbb{E}[Z(t+1) - Z(t)] \\
&\quad = \mathbb{E}[(\ln B - \psi_{t+1})^2 - (\ln B - \psi_t)^2] - Q \\
&\quad = \mathbb{E}[-2\ln B(\psi_{t+1}-\psi_t) + \psi_{t+1}^2 - \psi_t^2] - Q \\
&\quad = -2\ln B\mathbb{E}[\psi_{t+1}-\psi_t] \\
&\qquad + 2\mathbb{E}[\psi_{t+1}\psi_t] + \mathbb{E}[(\psi_{t+1}-\psi_t)^2] - Q \\
&\quad \geq -2\ln B\cdot 0 + 2\mathbb{E}[\psi_{t+1}\psi_t] + (\ln(1+\eta)^2\kappa) - Q,
\end{aligned}$$

where the last inequality follows from Lemmas 3.1 and 3.2. Since the chain is lazy, we have

$$\begin{aligned}
\mathbb{E}[\psi_{t+1}\psi_t] &\geq \frac{1}{2}(-2\ln 2)\psi_t + \frac{1}{2}\psi_t^2 \\
&\geq \frac{\ln 2}{2}(-2\ln 2 + \ln 2) \\
&\geq -\frac{1}{4}
\end{aligned}$$

Hence $E[Z(t+1)-Z(t)] \geq -\frac{1}{2} + (Q+\frac{1}{2}) - Q \geq 0$. Also, since the differences $Z(t+1)-Z(t)$ are bounded, we can conclude that $\{Z(t)\}$ is a submartingale. Let

$$\begin{aligned}
T^{x,y} &= \min\{t : \phi_t = 0\} \\
&= \min\{t : \psi_t = -2\ln 2\}.
\end{aligned}$$

Then $T^{x,y}$ is a stopping time for $Z(t)$, so we may apply the Optional Stopping Theorem for submartingales to deduce that

$$\begin{aligned}
\mathbb{E}[T^{x,y}] &\leq \frac{1}{Q}\left[\psi_0(2\ln B - \psi_0) + 4\ln B + 2\right] \\
&\leq \frac{2(\ln B)^2}{Q}.
\end{aligned}$$

It follows from Theorem 3.1 that

$$\tau(\epsilon) \leq \lceil\frac{2e(\ln B)^2}{Q}\rceil\lceil\ln(\epsilon^{-1})\rceil,$$

and this proves the theorem.

4 Rapid mixing of the biased Markov chain

To prove Theorem 2.1 showing that the biased Markov chain $\mathcal{M}_{mon}$ is rapidly mixing, we return to the coupling of (σ_t, ρ_t) that simply supplies the same (v^*, b, p) to both σ_t and ρ_t. We let U be the set of downsets that differ on a single tile. However, instead of the Hamming distance, we consider

$$\phi(\sigma,\rho) = \sum_{x\in\sigma\oplus\rho} (\sqrt{\lambda})^{m-x\cdot\overline{u}^*},$$

where $m := \max_{y\in R} y\cdot\overline{u}^*$.

We will show that this distance metric satisfies non-negative contraction in ϕ_t, which is one of the requirements for Theorem 3.3. However, before we can prove that the distances decrease on average, we examine those moves which can *increase* the distance.

For a pair $(\sigma_t, \rho_t) \in U$, there are two different ways the distance can increase in $(\sigma_{t+1}, \rho_{t+1})$. If $\sigma_t = \rho_t \cup \{x\}$, we can increase the distance by attempting to add a v that succeeds in σ_t but fails in ρ_t. This occurs when $v = x + \overline{u}_i$ for some i, so v is "supported" in σ_t but not ρ_t. The other way to increase the distance

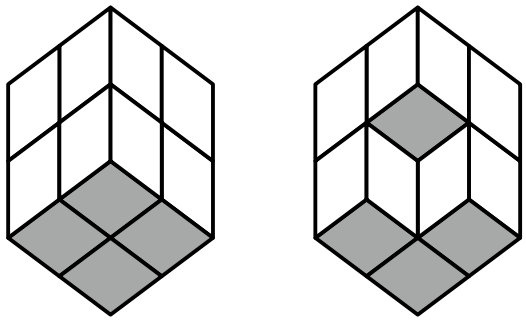

Figure 5: Downsets that differ on x, where $\mathcal{M}_{mon}$ increases ϕ_t by adding $x + \overline{u}_i$, for any i.

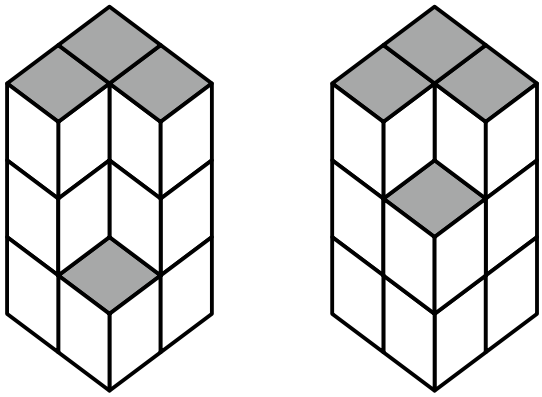

Figure 6: Downsets that differ on x, where $\mathcal{M}_{mon}$ increases ϕ_t by adding the vertex above x or removing the vertex below x.

between σ_t and ρ_t is to remove a v that succeeds in ρ_t but not in σ_t. This occurs when $v = x - \overline{u}_i$ for some i, as the move creates a valid downset in ρ but not in σ. The following lemma bounds the number of places this can occur.

LEMMA 4.1. *Let $(\sigma_t, \rho_t) \in \Omega_{mon} \times \Omega_{mon}$ be neighboring configurations such that $\sigma_t = \rho_t \cup \{x\}$. Then there are at most d choices of (v^*, b) which could cause ϕ_t to increase.*

Proof. We prove the lemma by showing that for dimensions $i \neq j$, if $\mathcal{M}_{mon}$ would increase the distance by choosing $v = x + \overline{u}_i$, then it would not increase the distance by choosing $v = x - \overline{u}_j$. This follows from a proof by contradiction. Suppose that $\mathcal{M}_{mon}$ could increase the distance with $x + \overline{u}_i$. If follows that $\rho_{t+1} = \rho_t \cup \{x + \overline{u}_i\}$ is a valid downset, which means that $x + \overline{u}_i - \overline{u}_j \in \rho_t$. On the other hand, if $\mathcal{M}_{mon}$ would increase the distance by choosing $x - \overline{u}_j$, it is because $\sigma_{t+1} = \sigma_t \backslash \{x - \overline{u}_j\}$ is a valid downset, and this is only true if $x - \overline{u}_j + \overline{u}_i \notin \sigma_t$. But this contradicts the fact $\sigma_t \oplus \rho_t = \{x\}$, thereby justifying our claim.

Now it follows that in order to increase the distance, the coupled chain may do one of three things. It can attempt to add vertices of the form $x + \overline{u}_i$, in each dimension i, as in Figure 5; or, symmetrically, it may attempt to remove vertices of the form $x - \overline{u}_i$, in each dimension i; or it may attempt to add $x + \overline{u}_i$ or remove $x - \overline{u}_i$ in a single designated dimension i, as in Figure 6. In each of these cases, there are at most d choices of v that increase the distance. All other choices of moves have the same effect on σ_t and ρ_t, and therefore do not change the distance between them.

We now bound the expected increase in ϕ_t.

LEMMA 4.2. *For each choise of v^* and b that can cause ϕ_t to increase during a move of the coupling, the expected increase ot the distance is $\phi_t \lambda^{-1/2}$.*

Proof. If the move is of the form $v = x + \overline{u}_i$ for some i, then the increase in distance is $(\sqrt{\lambda})^{m - v \cdot \overline{u}^*} = \phi_t \lambda^{-1/2}$.

If the move is of the form $v = x - \overline{u}_i$ for some i, then the increase in distance is $\lambda^{-(\|v\|_1)/2} = \phi_t \lambda^{1/2}$, and the chance of choosing an appropriate p is $1/\lambda$. Therefore the expected increase is again $\phi_t \lambda^{-1/2}$.

This allows us to show that the expected change in distance is at most zero.

LEMMA 4.3. *For $\phi_t = \phi(\sigma_t, \rho_t)$,*

$$\mathbb{E}[\phi_{t+1} - \phi_t] \leq 0.$$

Proof. As shown in the proofs of Lemmas 4.1 and 4.2, there are at most d choices of (v^*, b) that can increase ϕ_t, each giving an expected increase of $\phi_t / \lambda^{1/2}$. There are also two choices of (v^*, b) that can *decrease* ϕ_t, corresponding to adding x to σ_t or removing x from ρ_t. These moves each decrease the distance by ϕ_t and they succeed with probability 1 and $1/\lambda$, respectively. Therefore the expected change in distance is

$$\begin{aligned}
&\mathbb{E}[\phi_{t+1} - \phi_t] \\
&\leq d \cdot \frac{\phi_t}{\sqrt{\lambda}} - \left(1 + \frac{1}{\lambda}\right) \phi_t \\
&= -\phi_t \left(\frac{1}{\sqrt{\lambda}} - \frac{d + \sqrt{d^2 - 4}}{2}\right)\left(\frac{1}{\sqrt{\lambda}} - \frac{d - \sqrt{d^2 - 4}}{2}\right).
\end{aligned}$$

When $\sqrt{\lambda} > 2/(d - \sqrt{d^2 - 4})$, the above quantity is negative, thereby proving the lemma.

We may now prove the main theorem showing that the biased chain converges quickly.

Proof of Theorem 2.1. We need only verify that the requirements for Theorem 3.3 are all satisfied using the distance metric ϕ defined above.

For arbitrary $\sigma, \rho \in \Omega_{mon}$, if $x \in \sigma \oplus \rho$ for some x, then

$$\phi(\sigma, \rho) \geq (\sqrt{\lambda})^{m - x \cdot \overline{u}^*} \geq 1.$$

Therefore if $\phi(\sigma, \rho) < 1$, we must have that $\phi(\sigma, \rho) = 0$.

We let U be the pairs of downsets that differ by a single vertex. For arbitrary $\sigma, \rho \in \Omega_{mon}$, we can connect σ to ρ by simply adding or removing the vertices in $\sigma \oplus \rho$ one by one, where $\phi(\sigma, \rho)$ is the sum of the distances between neighbors along this path.

Since the volume of $\widehat{R}$ is n, there are at most n possible vertices in $\sigma \oplus \rho$, so $\phi(\sigma, \rho) \leq n\sqrt{\lambda}^{\gamma}$ for all σ and ρ, where γ is hte stretch of R. For any pair of σ, ρ such that $|\sigma \oplus \rho| = 1$, $\mathcal{M}_{mon}$ can always add the vertex at which they differ. The appropriate v^* is chosen with probability $1/\alpha$ and the appropriate b is chosen with probability $1/2$ (and every p succeeds when adding). Therefore there is a $1/(2\alpha)$ chance of decreasing the distance by ϕ_t, so

$$P(|\phi_{t+1} - \phi_t| \geq \phi_t) \geq \frac{1}{2\alpha}.$$

Together, these prove Theorem 2.1 using path coupling by appealing to Theorem 3.3.

References

[1] L. Adleman, Q. Cheng, A. Goel, M.D. Huang and H. Wasserman. Linear self-assemblies: Equilibria, entropies and convergence rates. *6th International Conference on Difference Equations and Applications*, 2001.

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

[3] I. Benjamini, N. Berger, C. Hoffman and E. Mossel. Mixing times of the biased card shuffling and the asymmetric exclusion process. *Transactions of the American Mathematics Society* **357:** 3013–3029, 2005.

[4] R. Bubley and M. Dyer. Faster random generation of linear extensions. *Discrete Math.*, **201**:81–88, 1999.

[5] M. Cook, P. Gacs and E. Winfree. Self-stabilizing synchronization in 3 dimensions. Preprint, 2008.

[6] R. Durrett. Probability: Theory and Examples, 2ed. Duxbury, 1996.

[7] M. Dyer and C. Greenhill. A more rapidly mixing Markov chain for graph colorings. *Random Structures and Algorithms*, **13**:285–317, 1998.

[8] T.-J. Fu and N.C. Seeman. DNA double-crossover molecules. *Biochemistry,* **32:** 3211–3220, 1993.

[9] D. Galvin, J. Kahn, D. Randall and G. Sorkin. In preparation, 2008.

[10] D. Galvin and D. Randall. Torpid mixing of local Markov chains on 3-colorings of the discrete torus. *17th Symposium on Discrete Algorithms (SODA)*, 376–384, 2007.

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

[12] U. Majumder, S. Sahu and J. Reif. Stochastic analysis of reversible self-assembly. To appear in *the Journal of Computational and Theoretical Nanoscience*, 2008.

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

[14] J. Linde, C. Moore and M. Nordahl. An n-Dimensional Generalization of the Rhombus Tiling. *Discrete Mathematics and Theoretical Computer Science Proceedings AA,* 23-42, 2001.

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

[16] J.H. Reif, S. Sahu and P.Yin Compact Error-Resilient Computational DNA Tiling Assemblies. *10th International Meeting on DNA Computing*, 293–307, 2004

[17] N.C. Seeman. DNA in a material world. *Nature,* **421:** 427–431, 2003.

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

[19] E. Winfree. Simulations of Computing by Self-Assembly. *4th DIMACS Meeting on DNA Based Computers*, University of Pennsylvania, 1998.

[20] E. Winfree, X. Yang and N.C. Seeman. Universal computation via self-assembly of DNA: Some theory and experiments. In *DNA Based Computers II,* L.F. Landweber and E.B. Baum, editors, *DIMACS series* **44:** 191–213, 1996.