Convergence Rates of Markov Chains for Some Self-Assembly and Non-Saturated Ising Models

Sam Greenberg[*]   Dana Randall[*]

Abstract

Algorithms based on Markov chains are ubiquitous across scientific disciplines as they provide a method for extracting statistical information about large, complicated systems. For some self-assembly models, Markov chains can be used to predict both equilibrium and non-equilibrium dynamics. In fact, the efficiency of these self-assembly algorithms can be related to the rate at which simple chains converge to their stationary distribution. We give an overview of the theory of Markov chains and show how many natural chains, including some relevant in the context of self-assembly, undergo a phase transition as a parameter of the model representing temperature is varied. We illustrate this behavior for the *non-saturated Ising model* in which there are two types of tiles that prefer to be next to other tiles of the same type. Unlike the standard Ising model, we also allow empty spaces that are not occupied by either type of tile. We prove that a local Markov chain that allows tiles to attach and detach from the lattice, the rate of convergence is fast at high temperature and slow at low temperature.

1   Introduction

Markov chain Monte Carlo methods are used in many areas of science as a computational tool for studying large, combinatorial sets. A Markov chain is a process that simulates a random walk moving among configurations in the large set, just like shuffling a deck of cards. Even though each configuration might have only a relatively small number of neighbors, a Markov chain may be designed so that it converges to desirable distributions over the entire space of configurations. The time required for the chain to converge to (close to) this equilibrium distribution, known as the *mixing time*, tells us whether a particular Markov chain is an efficient tool for sampling. For a state space that is exponentially large in the size of the input, we require that the chain comes close to the equilibrium distribution in only polynomial time in order for this sampling method to be effective.

Markov chains provide a natural tool for understanding many DNA-based self-assembly models. One popular model uses DNA strands to construct "tiles" that have particular single-stranded DNA sequences along each of its four edges (see, e.g., [20]). If two tiles have complementary sequences on opposite edges, then the base pairs along these edges encourage the tiles to attach by forming bonds, while tiles that have very few complementary base pairs are less likely to attach. Informally, the likelihood of each configuration

1

(or tiling) is given by the *Gibbs distribution*, a function of the number of complementary base pairs along the edges of all the tiles and the temperature. At high temperature any pair of tiles has similar probability of attaching along an edge, but at low temperature pairs with many complementary base pairs along each edge of the tiling are much more likely to attach along these bonds. Thus, at high temperature tilings look quite random, while at low temperature they tend to look very ordered.

If we put our set of tiles in a closed system where they self-assemble into tilings, we expect that the likelihood of a tiling will be given by the Gibbs distribution. Markov chains simulate the non-equilibrium dynamics of the tiling model as it approaches stationarity. At each step we allow any tile to attach or detach according to the probabilities determined by markings on the edges of the tile and its neighbors. *Glauber dynamics* are a general class of local Markov chains that connect any two configurations that have Hamming distance one (i.e., they differ only at a single tile).

For simplicity, in this paper we shall consider a particular tiling model where there are two tile types, **A** and **B**, that are designed so that each tile prefers to be next to other tiles of its same type. This is very reminiscent of the *Ising model* of statistical physics, where lattice points are assigned a spin of $\{+, -\}$ and similar spins prefer to be next to each other. However, in the tiling model under consideration we also allow lattice squares to be unoccupied, represented by a third symbol, **0**. This corresponds to a *non-saturated Ising model* where tiles **A** and **B** each prefer to be next to tiles of their own type, and in addition, neighbors of any type (**A** or **B**) are preferable to empty space. The relative strength of these preferences are a function of the temperature and are very strong at low temperature and weak at high temperature. We define these models more carefully in Section 2.

The local Markov chain (i.e., the Glauber dynamics) iteratively picks a lattice square at random. If there is a tile present, it can remove it (change it to a **0**). If there is no tile present, it tries to add a tile by changing the **0** to an **A** or a **B**. Note that in the model considered here, we disallow moves that allow a lattice square to change from an **A** tile to a **B** tile in a single step since this is unrealistic in the tiling scenario. However, we note that including such moves would not change the results we derive in any qualitative way.

A critical consideration in the design of a Markov chain is its rate of convergence to stationarity, known as its *mixing time*. A chain is *rapidly mixing* if it converges in polynomial time and *slowly mixing* if it requires exponential time to approach stationarity. The physical interpretation of these self-assembly models provides insights into their equilibrium structure, as well as the efficiency of various Markov chain dynamics. Typically, local chains are found to be rapidly mixing at high temperature and slowly mixing at low temperature. For example, this behavior is known for Glauber dynamics on the standard Ising model. At sufficiently high temperature, Ising configurations will tend to be half + and half –, and the local Markov chain can be shown to be rapidly mixing (see, e.g., [14]). However, at sufficiently low temperature, with high probability, Ising configurations will have an overwhelming predominance of one spin. To move from a mostly + configuration to a mostly - one requires moving through a configuration that is half + and half –, but this set of configurations is exponentially unlikely [23]. This reveals an exponentially small cut (or bottleneck) in the state space that indicates the chain will require exponential time

to converge to equilibrium [8].

We show that similar results hold for the non-saturated Ising model. To simplify the proofs, we consider the model on the $n \times n$ torus, but the same results can be shown for regions without periodic boundary conditions as well. Specifically, we show the following two theorems:

Theorem 1. *The mixing time of the local Markov chain defined for the non-saturated Ising model on an $n \times n$ toroidal lattice region at sufficiently high temperature is at most $p(n, \log(1/\epsilon))$, for some polynomial p.*

Theorem 2. *The mixing time of the local Markov chain defined for the non-saturated Ising model on an $n \times n$ toroidal lattice region at sufficiently low temperature is at least $exp(\psi(\beta)n)$ where $\psi(\beta) > 0$.*

This outline of the paper is as follows. In Section 2 we formalize the self-assembly and statistical physics models discussed. In Section 3 we review the fundamentals of Markov chains and more precisely define the local Markov chain we consider here. Finally, in Sections 4 and 5, we prove Theorems 1 and 2 to show the difference in the behavior of the Markov chain at high and low temperatures.

2 Self-assembly and pairwise influence models

2.1 DNA-based self-assembly

Self-assembly is a process in which large numbers of simple objects aggregate into larger structures in somewhat predictable ways. One exciting approach that has received much attention is tile-based self-assembly. In tiling models, tiles are designed with markings on each side so that two tiles are more likely to join together along an edge if they have identical markings. Wang studied such tiling systems and showed that they form a universal model of computation [26], making them an appealing object of study.

The primary challenge, then, is to define a set of marked tiles so that tiles are likely to assemble into large aggregates, and so that markings determine which tiles are most likely to line up. The approach taken by Seeman [7, 20], Winfree [24], and others is to use DNA double-crossover molecules in order to construct marked tiles with DNA sequences on each side. We construct these sequences so that pairs of tiles we want next to each other have large numbers of complementary base pairs along their matched edges, with the likelihood of their attaching being given by their hybridization energies. See, e.g., [7, 19, 20, 24, 25] for more details.

Here we are interested in a theoretical abstraction that captures the fundamental features of this model. We imagine we have a large (infinite) supply of rectangular tiles of various types. These tiles have markings on each of their four sides, and there are well-defined energies describing how likely it is for a pair to line up along any of their edges. For example, imagine we have two types of tiles, **A** and **B**. If the sides of **A** are complementary to each other but not to **B** and the sides of **B** are complementary to each other but not to **A**, then tiles of the same type will tend to cluster together. Moreover, this
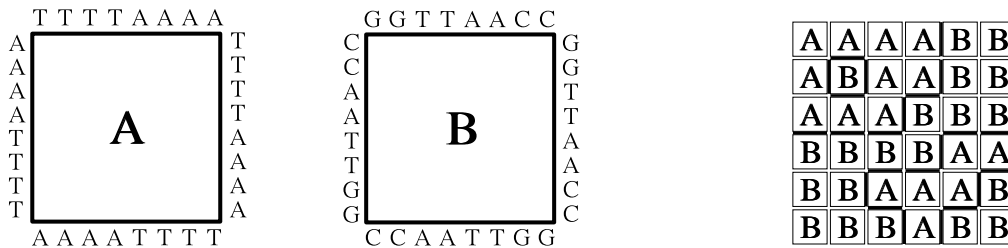
Figure 1: Tiles **A** and **B** designed so that each tile prefers to be adjacent to a tile of the same type.

preference is amplified at low temperatures and dampened at high temperatures. Such a tiling is illustrated in Figure 1.

Note that we may create the DNA sequences along the sides of the tiles so that there are only a small number of complimentary pairs that align when **A** touches **B**, but many (or all) that are complimentary when **A** touches **A** or when **B** touches **B** along any of their sides. For ease of analysis, we restrict to the model where tiles meet only when opposite sides perfectly line up, so we can think of tiles as occupying unit squares in the two-dimensional Cartesian lattice.

2.2   General pairwise influence models

The tile-based self-assembly algorithms are an example of a broad class of models studied in statistical physics that we can describe as *pairwise-influence models*, also known as *Markov Random Fields*. Each nearest-neighbor pair of tiles is given a weight representing how favorable it is for the two tiles to bond to one another along their nearest sides. It will be convenient to think of the arrangement on the dual lattice, so vertices are assigned labels **A**, **B** or **0**. Then the "energy" of an edge in the dual is precisely the bond strength between the two tiles corresponding to the endpoints of the edge.

More precisely, suppose we are given a graph $\Lambda = (V, E)$, say the $n \times n$ Cartesian lattice with periodic boundary conditions. This is just the torus where (x, y) is connected to $(x \pm 1, y)$ and $(x, y \pm 1)$, with all operations taken mod n. Let $\Omega = \{1, ..., q\}^V$ be the state space and define a symmetric set of weights $\{w_{i,j} = w_{j,i}\}$ for each pair $i, j \in \{1, \ldots, q\}$. Then give each configuration $\sigma \in \Omega$ a weight

$$w(\sigma) = \prod_{u,v:(u,v)\in E} w_{\sigma(u),\sigma(v)}.$$

Normalizing, we get a probability distribution

$$\pi(\sigma) = \frac{w(\sigma)}{\sum_{\tau \in \Omega} w(\tau)}.$$

For notation, for a set $S \subseteq \Omega$, we let $\pi(S) = \sum_{\sigma \in S} \pi(\sigma)$.

By adjusting the values for $w_{i,j}$ we can favor certain pairs of nearest neighbors and discourage others. For example, if for $i, j \in \{1, \ldots, q\}$ we let $w_{i,j} = 1$ when $i \neq j$ and let

4

$w_{i,i} = 0$ for all i, the probability distribution arising is precisely the uniform distribution on the set of proper q-colorings.

These models were originally defined to represent simple physical systems. An *energy function* on the space of configurations is defined by a *Hamiltonian* $H(\sigma)$. For models where the energy is determined solely from nearest-neighbor interactions, $H(\sigma) = \sum_{(u,v)\in E} g(\sigma(u), \sigma(v))$, for some function g. Just like a spring relaxing, systems tend to favor configurations that minimize energy, where this preference is controlled by temperature. Each configuration in Ω is given a weight

$$w(\sigma) = e^{-\beta H(\sigma)},$$

where $\beta = 1/T$ is inverse temperature. Thus, for low values of β the differences between the energy of configurations are dampened, while at large β these differences are magnified. Taking $w_{i,j} = e^{g(i,j)}$ reconciles these two definitions.

The likelihood of each configuration is then given by

$$\pi(\sigma) = w(\sigma)/Z,$$

where $Z = \sum_\tau w(\tau)$ is the normalizing constant known as the *partition function*. This probability measure is known as the *Gibbs (or Boltzmann) distribution*. Taking derivatives of the generating function Z (or $\ln Z$) with respect to the appropriate variables allows us to calculate many of the interesting thermodynamic properties of the system, such as the specific heat and the free energy.

2.3   Saturated and non-saturated Ising models

The *Ising model* is a standard model of ferromagnetism studied in statistical physics. Given a graph Λ on n vertices, the state space is defined by the 2^n ways of assigning $+$ or $-$ *spins* to each of the vertices. In the *ferromagnetic* Ising model, the Hamiltonian is defined so as to favor configurations which tend to have equal spins on the endpoints of its edges; the *antiferromagnetic* Ising model prefers unequal spins on the endpoints. For $\sigma \in \{\pm\}^n$, the Hamiltonian of the ferromagnetic Ising model is given by

$$H_S(\sigma) = \sum_{(u,v)\in E} J_{\sigma(u),\sigma(v)},$$

where $J_{+,+} = J_{-,-} > 0$ and $J_{+,-} = J_{-,+} = 0$. We define $\beta = 1/T$ as the inverse temperature. Then the stationary probability of σ is

$$\pi_S(\sigma) = \frac{e^{\beta H_S(\sigma)}}{Z_S},$$

where Z_S is the normalizing constant. It is common to have a slightly different definition of the Hamiltonian, but we note that this definition can be seen to be equivalent with a change in variables. For an informal introduction to the Ising model, see [5].

The Ising model can be thought of as a fully-packed, or *saturated*, self-assembly model with two tiles **A** and **B** (corresponding to $+$ and $-$) occupying lattice squares, where
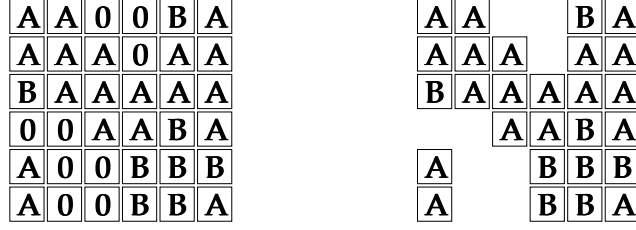
Figure 2: The non-saturated Ising model.

each tile prefers to be next to others of the same type (regardless of direction). In this abstraction, we imagine that every square is occupied by one of the two tile types. To define the model, we create tiles with DNA sequences along the sides so that the bond strengths satisfy $w_{\mathbf{A},\mathbf{A}} = w_{\mathbf{B},\mathbf{B}} > w_{\mathbf{A},\mathbf{B}} = w_{\mathbf{B},\mathbf{A}}$ over all edges. Setting $w_{\mathbf{A},\mathbf{A}} = e^{\beta J}$ and $w_{\mathbf{A},\mathbf{B}} = 1$, we have precisely the Ising model as defined above.

A simple modification of the standard Ising model makes it more fitting for a tile-based self-assembly modes in which configurations can include empty spaces. In the *non-saturated Ising model*, empty spaces are represented by a third tile type **0**. Then for $\sigma \in \{\mathbf{A}, \mathbf{B}, \mathbf{0}\}^n$, we define the Hamiltonian as

$$H_U(\sigma) = \sum_{(u,v) \in E} J_{\sigma(u),\sigma(v)},$$

where $J_{\mathbf{A},\mathbf{A}} = J_{\mathbf{B},\mathbf{B}} > J_{\mathbf{A},\mathbf{B}} = J_{\mathbf{B},\mathbf{A}} > J_{\mathbf{0},x} = J_{x,\mathbf{0}} = 0$ for $x \in \{\mathbf{A}, \mathbf{B}, \mathbf{0}\}$. As before, we take β to be inverse temperature, and define the Gibbs distribution as

$$\pi_U(\sigma) = \frac{e^{\beta H_U(\sigma)}}{Z_U},$$

where Z_U is the normalizing constant. This generalization helps capture the way that tiles will assemble and pull apart at equilibrium according to the temperature. Thus, a tile of either type most prefers to be next to another tile of the same type, but it also prefers to have its neighbors occupied rather than unoccupied, even if it is with a tile of the other type. This is demonstrated in Figure 2.

For simplicity, we can let $w_{x,y} = e^{\beta J_{x,y}}$ for all $x, y \in \{\mathbf{A}, \mathbf{B}, \mathbf{0}\}$. This gives us symmetric pairwise interactions $w_{\mathbf{A},\mathbf{A}} = w_{\mathbf{B},\mathbf{B}} = \lambda > w_{\mathbf{A},\mathbf{B}} = \mu > w_{\mathbf{A},\mathbf{0}} = w_{\mathbf{B},\mathbf{0}} = w_{\mathbf{0},\mathbf{0}} = 1$. We can now describe the stationary probability of a configurations σ as

$$\pi(\sigma) = \pi_U(\sigma) = \frac{\lambda^{\#\mathrm{match}} \mu^{\#\mathrm{differ}}}{Z},$$

where #match is the number of nearest neighbor pairs that are both assigned **A** or both assigned **B**, and #differ is the number of pairs where one tile is **A** and the other is **B**. It is this representation we will be using in the proofs that bound the mixing times.

6

3 The basics of sampling

We are going to study the non-saturated Ising model by simulating a Markov chain whose states correspond to allowable configurations. We design a local chain that adds or removes individual tiles in each step, where the transition probabilities are defined so that the chain converges to the Gibbs distribution. Our interest will be in the mixing time of this chain, or the time it takes for the chain to get close to equilibrium. In this section we review the basics of Markov chains, including the main techniques we use to bound the mixing time of our chain at high and low temperatures.

3.1 The local Markov chain for the non-saturated Ising model

We first define a graph K that connects the state space, known as the *Markov kernel*. We will consider local or *Glauber dynamics* that connect pairs of configurations that have Hamming distance one. This means that in a given step, the chain can pick any position and change the type that is there. For the non-saturated Ising model it might make the most sense to connect two configurations of Hamming distance one only if one of the configurations has a tile of type **0** in the position of disagreement — recall that **0** represents an empty position, so these moves correspond to tiles of type **A** or **B** attaching or detaching from the larger configuration. Notice that the Markov chain connects the state space even if we disallow directly changing a tile from type **A** to type **B**.

To define the transition probabilities of the Markov chain $\mathcal{M}$ on the edges of K, we refer to the desired stationary distribution π. In the case of the non-saturated Ising model, π is just the *Gibbs distribution*. The *Metropolis-Hastings algorithm* dictates transition probabilities which force the Markov chain to converge to π [15]:

$$P(\sigma, \tau) = \frac{1}{2\Delta} \min\left(1, \frac{\pi(\tau)}{\pi(\sigma)}\right),$$

for all (σ, τ) neighbors in K, where Δ is the maximum degree of Λ.

To implement moves of the Markov chain $\mathcal{M}$ for the non-saturated Ising model, we start at any initial configuration and repeat the following. If we are currently at configuration σ, we choose a triple (F, S, R) uniformly, where F is a face in the lattice, S is one of the three tile types **A**, **B** or **0**, and R is a real number in $[0, 1]$. Define τ to be the new configuration with the tile at face F in σ changed to S. If the move from σ to τ is allowable, then we accept the move if $R \leq \frac{\pi(\tau)}{\pi(\sigma)}$. (Recall that in our model tiles cannot change directly from **A** to **B** without first changing to a **0**.) Otherwise we remain at σ.

Note that in the ratio $p = \frac{\pi(\tau)}{\pi(\sigma)}$, the individual Gibbs probabilities might be very difficult to calculate due to the normalizing constant, but the ratio of probabilities depends solely on S, F and its neighbors. This allows us to efficiently calculate the transition probabilities. For example, if we have a **0** tile surrounded by four **A** tiles, and we try to change it to an **A**, then we accept this move with probability $\min(1, \lambda^4) = 1$; if we then try to reverse the move by changing an **A** surrounded by four **A** tiles to a **0**, then we accept the move with probability $\min(1, 1/\lambda^4) = \lambda^{-4}$.
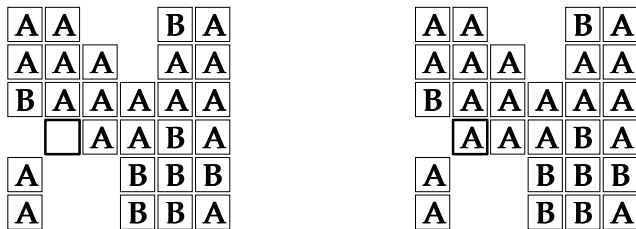
Figure 3: A move in the Markov chain $\mathcal{M}$ for the non-saturated Ising model.

3.2 Mixing times

It is easy to verify for a finite Markov chain that if the kernel is ergodic (connected and aperiodic), then π is the unique stationary distribution. This means that if we start at any vertex in K and perform a random walk according to the transition probabilities, we will eventually converge to the desired distribution π. For this approach to be useful, we need the chain to converge sufficiently rapidly to π so that after a small, polynomial number of steps, our samples will be generated according to a distribution which is provably close to π. With this property we have a method for efficiently sampling.

More formally, let $P^t(x, y)$ denote the t-step transition probability from x to y.

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

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

Definition 2. *For any $\varepsilon > 0$, the* mixing time $\tau(\varepsilon)$ *is*

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

A Markov chain is *rapidly mixing* if the mixing time is bounded above by a polynomial in the size of the state space and $\log \frac{1}{\varepsilon}$. When the mixing time is exponential in the size of the space, we say the chain is *slowly mixing*.

3.3 Tools for bounding mixing times

Our proofs of Theorems 1 and 2 rely on two fundamental techniques in the analysis of Markov chains, namely coupling and conductance. We use a coupling argument to show that our Markov chain is rapidly mixing at high temperature, and we bound the conductance to show that the chain is slowly mixing at low temperatures. We start by outlining these methods.

One of the most popular methods for bounding mixing times is coupling, both because of its elegance and its simplicity.

Definition 3. *A* coupling *is a Markov chain on $\Omega \times \Omega$ defining a stochastic process $(X_t, Y_t)_{t=0}^{\infty}$ with the properties:*

1. *Each of the processes X_t and Y_t is individually a faithful copy of the Markov chain (given initial states $X_0 = x$ and $Y_0 = y$).*

2. *If $X_t = Y_t$, then $X_{t+1} = Y_{t+1}$.*

Condition 1 means that each process, viewed in isolation, is just simulating the probabilities as prescribed by the original chain – yet the coupling updates them simultaneously so that they will tend to coalesce, or move closer together according to some notion of distance. Once the pair of configurations agree, condition 2 guarantees they agree from that time forward.

The coupling (or expected coalescence) time can provide a good bound on the mixing time of $\mathcal{M}$ if it is a carefully chosen coupling.

Definition 4. *For initial states x, y let*

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

and let $ET^{x,y}$ denote its expectation. Then we define the coupling time to be $T = \max_{x,y} ET^{x,y}$.

The following result relates the mixing time and the coupling time (see, e.g., [1]).

Theorem 3. *The mixing time satisfies $\tau(\epsilon) \leq \lceil Te \ln \epsilon^{-1} \rceil$, where e is the base of the natural logarithm.*

While coupling is potentially a powerful technique, it is often prohibitively cumbersome to measure the expected change in distance between two arbitrary configurations. The method of *path coupling*, introduced by Bubley and Dyer, greatly simplifies this approach by showing that we really need only consider pairs of configurations that are close according to some metric on the state space [3].

The idea behind path coupling is to consider a small set $U \subseteq \Omega \times \Omega$ of pairs of configurations that are "close" according to some distance metric φ. Suppose that we have shown that the expected change in distance is decreasing for all of the pairs in U. To now reason about arbitrary configurations $X, Y \in \Omega$, we define a shortest path $z_0, z_1, ..., z_r$ of length r from $X = z_0$ to $Y = z_r$, where $(z_i, z_{i+1}) \in U$ for all $0 \leq i < r$. If we define U correctly, then $\varphi(X, Y) = \sum_{i=0}^{r-1} \varphi(z_i, z_{i+1})$. Then by linearity of expectation, the expected change in distance between X and Y is the sum of the expected change between the pairs (z_i, z_{i+1}). Each of these has been shown to be at most zero, so the total distance is at most zero. Of course, after the update there might be a shorter path between the new configurations X' and Y', but this just causes the new distance to be even smaller.

The following version of the path coupling theorem is convenient [6].

Theorem 4. *If there exists $\beta < 1$ such that*

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

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

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

where $B = \max \varphi$.

9

A second method for estimating the convergence time of a Markov chain is bounding the *conductance* of the chain [8, 11]. The conductance can be used to upper or lower bound the mixing time of a chain, although here we will use it only to show that our Markov chain is slowly mixing at low temperature.

For any set $S \subset \Omega$, let

$$\Phi_S = \frac{\sum_{x \in S, y \notin S} Q(x, y)}{\pi(S)},$$

where $Q(x, y) = \pi(x)P(x, y)$ is regarded as the "capacity" of the edge (x, y) and $\pi(S) = \sum_{x \in S} \pi(x)$ is the weight of the set. Note that for any reversible chain, $Q(x, y) = Q(y, x)$. We now define the *conductance* as

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

If a Markov chain has low conductance, then there is a small cut in the state space that will cause a bottleneck in the mixing time. The conductance theorem due to Jerrum and Sinclair [8] and Lawler and Sokal [11] characterizes when a Markov chain mixes rapidly or slowly. We will use the use the lower bound from this theorem.

Theorem 5. *For any reversible Markov chain with conductance* Φ,

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

Thus, to show that a chain is slowly mixing, it suffices to demonstrate that there is a "bad cut" in the state space that requires exponential time to cross. This is the approach we will take in Section 5.

4   Rapid mixing at high temperature

In this section we consider the convergence rate of the Markov chain at high temperature. Using the notation of the previous section, we can now state Theorem 1 more precisely.

Recall that $\mu = e^{\beta J_{\mathbf{A}, \mathbf{B}}}$ and $\lambda = e^{\beta J_{\mathbf{A}, \mathbf{A}}} = e^{\beta J_{\mathbf{B}, \mathbf{B}}}$ are the edge weights of a configuration, with $\mu < \lambda$. We are given fixed values for the $J_{x,y}$ which are determined by the number of base pairs in our tiling model as well as the number of complementary base pairs on opposite sides of the x and y tiles. Then at high enough temperature (or small β), μ and λ will both be close to one, while at low temperature (large β), μ and λ will both be large. We show the following theorem.

Theorem 6. *If* $1 < \mu < \lambda < 8/7$, *then the mixing time* $\tau(\epsilon)$ *of* $\mathcal{M}$ *on the non-saturated Ising model on the* $n \times n$ *square toroidal lattice is* $O(n^2 \ln 1/\epsilon)$.

Notice that Theorem 1 follows as a corollary.

To prove $\mathcal{M}$ is rapidly mixing for λ and μ close to 1, we use a coupling argument; we create a new Markov chain which acts on *pairs* of configurations in $\Omega \times \Omega$, and show that the time until any two configurations become equal under this coupling is polynomial in n. To define the coupling, we choose a triple (F, S, R) exactly as before, but now, given

10

a pair of configurations (σ_t, ρ_t) at time t, we apply the move defined by he triple to both configurations at once, getting a new pair $(\sigma_{t+1}, \rho_{t+1})$.

We define the distance $d(\sigma, \rho)$ as the minimum number of moves of $\mathcal{M}$ to go from σ to ρ. (Note that this is slightly different than the Hamming distance, as changing a single face from $\mathbf{A}$ to $\mathbf{B}$ requires *two* moves.) Using Theorem 4, the path coupling theorem, we may restrict our examination to the pairs that are distance 1. If we show that, for any such pair, the expected distance decreases after a single move, this will show rapid mixing.

Lemma 1. *Let σ_t and ρ_t be configurations such that $d(\sigma_t, \rho_t) = 1$. Then, for every $\lambda < \frac{5}{4}$,*

$$\mathrm{E}[d(\sigma_{t+1}, \rho_{t+1})] \leq 1 - \frac{1}{36n^2}.$$

Proof. Configurations σ_t and ρ_t differ at a single face F^*. Without loss of generality, we may assume $\sigma_t(F^*) = \mathbf{A}$ and $\rho_t(F^*) = \mathbf{0}$.

When performing the move prescribed by the triple (F, S, R) to the pair (σ_t, ρ_t), the distance only changes if F is F^* or a neighbor of F^*. In all other cases, the Markov chain performs the same change to σ_t and ρ_t, and $d(\sigma_t, \rho_t) = d(\sigma_{t+1}, \rho_{t+1})$.

On the other hand, if F is F^* or one of its neighbors, the distance might increase or decrease by 1 after the move. To show that the expected distance is decreasing, we first lower bound the probability of it decreasing, and then upper bound the probability of it increasing.

The two moves on which we may decrease the distance are when $F = F^*$ and S is either $\mathbf{A}$ or $\mathbf{0}$. In the first, ρ_t is changed to σ_t while σ_t remains the same. Therefore $\rho_{t+1} = \sigma_t = \sigma_{t+1}$. In the second, σ_t is changed to ρ_t while ρ_t stays the same, so $\sigma_{t+1} = \rho_t = \rho_{t+1}$. The probability of choosing an appropriate F and S, and an R high enough that the move succeeds, is

$$\frac{1}{n^2}\frac{1}{3}\min\left(1, \frac{\pi(\sigma_t)}{\pi(\rho_t)}\right) + \frac{1}{n^2}\frac{1}{3}\min\left(1, \frac{\pi(\rho_t)}{\pi(\sigma_t)}\right).$$

As $\pi(\sigma_t) \geq \pi(\rho_t)$, this probability is lower bounded by $\frac{1}{3n^2} + \frac{1}{3n^2\lambda^4}$.

There are many moves that potentially increase the distance. The first is when $F = F^*$ and $S = \mathbf{B}$. Then the move automatically does nothing to σ_t, but might further change ρ_{t+1}. This occurs with at most probability $\frac{1}{n^2}\frac{1}{3}$.

Another move that may increase the distance is choosing a face F' adjacent to F^*. Here, the range of acceptable R might be larger for one configuration than the other, so one configuration's move succeeds while the other's fails. Such a difference occurs if $\sigma_t(F') = \rho_t(F') \neq \mathbf{0}$. Then, when $F = F'$ and $S = \mathbf{0}$, the Markov chain changes $\pi(\sigma_t)$ by a factor of (at most) λ more than it changes $\pi(\rho_t)$. The probability of succeeding on ρ_t but not σ_t is the difference of their two chances of success, which is at most $\frac{1}{n^2}\frac{1}{3}(1 - \frac{1}{\lambda})$.

As there are only four neighbors of F^*, we have

$$\mathrm{E}[d(\sigma_{t+1}, \rho_{t+1}) - d(\sigma_t, \rho_t)] \leq \frac{1}{3n^2} + \frac{4}{3n^2}\left(1 - \frac{1}{\lambda}\right) - \frac{1}{3n^2} - \frac{1}{3n^2\lambda^4}.$$

Substituting for λ, we have the lemma. $\qquad\square$

We may now prove Theorem 6, that the chain is rapidly mixing when λ and μ are sufficiently close to 1.

Proof. We appeal to the path coupling theorem, Theorem 4. The set $U \subset \Omega \times \Omega$ is the pairs of configurations with distance 1. Notice that for any arbitrary pair of configurations with distance d, there is a path of length d that connects them where nearest neighbors along the path have distance 1. We now know from Lemma 1 that we can set $B = 1 - \frac{1}{36n^2} < 1$. The theorem follows. □

5   Slow Mixing at low temperature

We now turn our attention to the low temperature case, when λ and μ are both large and λ is much greater than μ. When λ/μ is large, there is a large penalty to the stationary probability each time an **A** tile is placed next to a **B** tile, or when we have a **0**. We will show that when this ratio is large enough, typical configurations will be dense and will have a predominance of **A** tiles or **B** tiles. This allows us to show there is a bottle-neck in the state space, implying that the chain is slowly mixing due to small conductance. Specifically, we show the following theorem.

Theorem 7. *If $\frac{\lambda}{\mu} > 28$, then there exists constant $c > 1$ such that the mixing time of $\mathcal{M}$ on the non-saturated Ising model on the $n \times n$ square toroidal lattice is $\Omega(c^n)$.*

The proof will be facilitated by a careful partition of the state space to define the bad cut. To move from a configuration which is dominated by type **A** tiles to a configuration which is dominated by type **B** tiles, $\mathcal{M}$ must pass through a configuration where there is no dominating type. Intuitively, these configurations without a dominating type are heavily penalized when **A** and **B** interact often. We therefore partition the state space into three sets, and show that the set of configurations with no dominating type forms a bottle-neck.

However, in defining these sets, instead of dividing the state space according to, say, the relative number of tiles of type **A** or **B**, we take the approach of [17] and focus on finding the simplest possible description of the cut to simplify the subsequent analysis. We define middle set Ω_F as configurations that have a "fault line": a long path consisting entirely of "unfavorable" edges, where either a tile is unoccupied or where **A** meets **B**. We formalize these definitions momentarily, but any configuration with such a path will be penalized for each edge, and therefore have very low stationary probability.

The key idea is that the state space Ω now can be partitioned into three sets Ω_A, Ω_F and Ω_B. Any configuration that does not have a fault line can be shown to have a large connected component of a single type of tile crossing the lattice region; these configurations (which actually comprise the majority of Ω) lie in Ω_A or Ω_B according to the tile type.

In the proof, we show that it is necessary for the Markov chain to pass through configurations in Ω_F to move from Ω_A to Ω_B, and yet Ω_F has exponentially smaller Gibbs measure than each of the other two sets; this guarantees that the chain is slowly mixing since it is enough to show that the chain has exponentially small conductance. This idea

of using "topological obstructions" based on fault lines to partition the state space for a proof of slow mixing has been applied previously in the context of independent sets [17] and the standard Ising model [23].

We need to define several terms to formalize the argument. We call an edge *good* if it lies between **A** and **A** or **B** and **B**, and call it *bad* otherwise. We use the term *noncontractible cycle* to refer to any cycle on the torus that is not contractible to a point. A *fault line* is a noncontractible cycle of bad edges. We say two faces of the lattice are *adjacent* if they share an edge. Define an **A**-*bridge* to be a noncontractible cycle of adjacent **A** faces. For any non-contractible cycle, the *winding vector* is an ordered pair of integers (w_x, w_y), where w_i represents the net number of times the cycle intersects an elementary loop in the ith lattice direction. Define an **A**-*cross* to be a pair of **A**-bridges with different winding vectors. Define a **B**-*bridge* and **B**-*cross* similarly.

Let $\Omega_{\mathbf{F}}$ be the set of configurations that contain a fault line. Let $\Omega_{\mathbf{A}}$ and $\Omega_{\mathbf{B}}$ be the configurations which have an **A**-cross and a **B**-cross, respectively. We now show that these three sets partition Ω.

Lemma 2. *The sets $\Omega_{\mathbf{A}}$, $\Omega_{\mathbf{F}}$, and $\Omega_{\mathbf{B}}$ are disjoint, with $\Omega_{\mathbf{A}} \cup \Omega_{\mathbf{F}} \cup \Omega_{\mathbf{B}} = \Omega$. Furthermore, for any $\sigma \in \Omega_{\mathbf{A}}$ and $\rho \in \Omega_{\mathbf{B}}$, $P(\sigma, \rho) = 0$.*

Proof. A key point we use here is that any two non-contractible cycles of different winding vectors must intersect. Therefore a configuration could not have both an **A**-cross and a **B**-cross, as their intersection would have both types of tiles. Similarly, a configuration can't have both a cross and a fault line, as the fault line would put a bad edge across a bridge, contradicting the notion of a bad edge. Therefore $\Omega_{\mathbf{A}}$, $\Omega_{\mathbf{F}}$, and $\Omega_{\mathbf{B}}$ are disjoint.

To show that the three sets make up all of Ω, examine the torus after the removal of an **A**-bridge. The remaining space is a non-contractible strip of the torus of the same winding vector as the bridge. If there exists a path of adjacent **A** faces across this strip, we find a new **A**-bridge of a different winding vector, and therefore a cross. If this path doesn't exist, then there must exist a fault line along the strip. Hence every configuration is in either $\Omega_{\mathbf{A}}$, $\Omega_{\mathbf{F}}$, or $\Omega_{\mathbf{B}}$.

Every move of the Markov chain adds or removes a tile. Moving from σ to ρ would entail first breaking an **A**-cross and then creating a **B**-cross; this requires at least two moves. Hence Ω_F divides Ω_A from Ω_B. $\qquad\square$

Clearly $\pi(\Omega_{\mathbf{A}}) = \pi(\Omega_{\mathbf{B}})$ by symmetry. To show that the Markov chain has small conductance then, we need only show that for large n, $\pi(\Omega_{\mathbf{F}}) < C^n$, for some fixed constant $C < 1$. To this end, we define a function ψ that maps configurations in Ω_F to new configurations with exponentially larger stationary probability. Although this function will not be one-to-one, we will show that, for each configuration in Ω, the total stationary probability of the preimage is exponentially less than the weight of the configuration itself. This will allow us to show that Ω_F is exponentially smaller than Ω, and therefore significantly smaller than each of Ω_A and Ω_B.

We know that any configuration $\sigma \in \Omega_F$ has at least one fault line, and possibly many. The function ψ will select one and then will remove it by adding type **A** tiles everywhere that is currently unoccupied, thereby forcing the new configuration to have much greater Gibbs weight. This would be enough to show that $\pi(\Omega_F)$ were small if ψ were injective, but
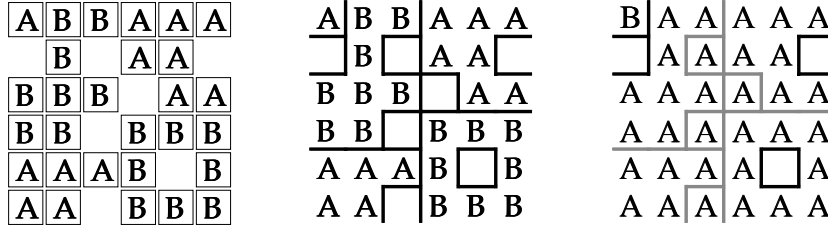
Figure 4: An example of σ, σ with bad edges highlighted, and $\psi(\sigma)$.

of course it typically is not. Therefore we need to argue that for each $\sigma \in \Omega_F$, the increase in weight of $\psi(\sigma)$ is large enough to account for all the preimages in Ω_F, to conclude that $\pi(\Omega_F)$ is indeed small. In addition, we need to generalize the notion of a fault line to a maximal connected component, in order to define a map that has the described properties. This generalization is not necessary for the proofs in [17] and [16], for independent sets and the Ising model, respectively.

Accordingly, define a *fat fault* to be a maximal connected set of adjacent bad edges containing a fault line. We choose an arbitrary fat fault T of σ, say the lexicographically first one. The function ψ will replace all of the bad edges of T with good edges.

To define ψ, first notice that the regions of the torus separated by the edges of T are either individual faces of $\mathbf{0}$ or larger collections of faces that are bordered with entirely $\mathbf{A}$ or entirely $\mathbf{B}$, or $\mathbf{A}$-*bordered* or $\mathbf{B}$-*bordered*, respectively. To find $\psi(\sigma)$, fill in every $\mathbf{0}$ face with a $\mathbf{A}$, and flip the sign of every $\mathbf{A}$ or $\mathbf{B}$ face in the $\mathbf{B}$-bordered regions. Therefore every edge that was bad in σ is between two $\mathbf{A}$ tiles in $\psi(\sigma)$. Such a change is illustrated in Figure 4.

Now suppose that we are given $\sigma' \in \text{Img}(\psi)$. Notice that in order to invert ψ to recover σ, it suffices to have an encoding of T and a labelling of the distinct regions of T dictating for each whether they were $\mathbf{0}$s, $\mathbf{A}$-bordered, or $\mathbf{B}$-bordered before the map made them all $\mathbf{A}$-bordered. In what follows we will continue to call the faces of T *regions* to distinguish them from the faces of the underlying lattice.

We now proceed to prove several lemmas that provide the necessary bounds to put these pieces together.

Lemma 3. *Let $\mathcal{T}_m$ be the set of fat faults with m bad edges. The number of fat faults in $\mathcal{T}_m$ is at most $n^2 16^m$.*

Proof. For some $T \in \mathcal{T}_m$, a depth-first search of the edges of T takes at most $2m$ steps. There are at most n^2 possible starting points, and 4^{2m} possible depth-first search traversals. $\square$

Lemma 4. *For a given $T \in \mathcal{T}_m$, the number of labellings of the regions of T as $\mathbf{A}$-bordered, $\mathbf{B}$-bordered, or $\mathbf{0}$ is at most $\sqrt{3}^m$.*

Proof. Every region of T is bounded by at least four edges. Each edge is on the boundary of two regions. Therefore the number of regions is at most $m/2$. Hence there are at most $3^{\frac{m}{2}}$ choices of labelings for each region. $\square$

14

We may now use Lemmas 3 and 4 to bound the weight of the preimage of ψ.

Lemma 5. *For each $\sigma' \in \text{Img}(\psi)$,*

$$\pi(\psi^{-1}(\sigma')) < n^4 \frac{28\mu}{\lambda}^n.$$

Proof. We first partition preimages based on the size of the fat fault defining ψ. Let $\psi_m^{-1}(\sigma')$ be the set of all σ such that the fat fault chosen by ψ has m edges and $\psi(\sigma) = \sigma'$. The function ψ turns every bad edge of T into an edge separating $\mathbf{A}$ and $\overline{\mathbf{A}}$, thereby increasing the weight of each by at least a factor of $\frac{\lambda}{\mu}$. So, for each $\sigma \in \psi_m^{-1}(\sigma')$,

$$\pi(\sigma) \le \left(\frac{\mu}{\lambda}\right)^m \pi(\sigma').$$

On the other hand, $|\psi_{m,\ell,s}^{-1}(\sigma')|$ is bounded by the number of possible fat faults T and labellings of the regions of T. By Lemmas 3 and 4,

$$|\psi_m^{-1}(\sigma')| < n^2 16\sqrt{3}^m < n^2 28^m.$$

Putting these pieces together, we have

$$\pi(\psi^{-1}(\sigma')) = \sum_{m=n}^{n^2} \pi(\psi_m^{-1}(\sigma'))$$

$$= \sum_{m=n}^{n^2} \sum_{\sigma \in \psi_m^{-1}(\sigma')} \pi(\sigma)$$

$$= \sum_{m=n}^{n^2} |\psi_{m,\ell,s}^{-1}(\sigma')| \frac{\mu}{\lambda}^m \pi(\sigma')$$

$$< \sum_{m=n}^{n^2} n^2 28^m \frac{\mu}{\lambda}^m \pi(\sigma')$$

$$< n^2 n^2 \frac{28\mu}{\lambda}^n \pi(\sigma').$$

$\square$

Lemma 5 allows us to now prove our main Theorem, that at high temperatures the Markov chain mixes slowly.

Proof of Theorem 7. We will show that when $\lambda > 28\mu$, $\Omega_{\mathbf{F}}$ has exponentially low stationary probability, and thus the chain has exponentially small conductance.

15

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

$$< \sum_{\sigma' \in \mathrm{Img}(\psi)} n^4 \frac{28\mu^n}{\lambda} \pi(\sigma')$$

$$\leq n^4 \frac{28\mu^n}{\lambda} \sum_{\sigma' \in \Omega} \pi(\sigma')$$

$$= n^4 \frac{28\mu^n}{\lambda}.$$

By Theorem 5, this implies that the mixing time is exponential in n. $\qquad\square$

6   Conclusions

We defined the non-saturated Ising model to represent a simple DNA-based self-assumbly model consisting of two types of tiles. We then showed that the assembly is efficient at high temperatures, where the stationary distribution allows many "errors" or unfavorable neighbors, and the assembly is slow at low temperatures, where errors are much less likely. This suggests that the temperature must be very carefully chosen in order to have a meaningful stationary distribution that converges in a reasonable amount of time. Although we prove this behavior for just this model, this type of phase transition is very common for pairwise influence models that are widely studied in statistical physics.

In addition to helping us appreciate the dichotomy underlying the behavior of Markov chains based on local updates, the physical models also aid our design of better sampling algorithms. One approach that has received a lot of attention is *simulated annealing* and its sampling analogue *simulated tempering*. Madras and Zhong [13] showed that although sampling Ising configurations at low temperatures is known to be slow, a version of tempering in which the temperature is raised and lowered during the simulation can be shown to be fast and can provide samples at all temperatures; they prove this when the underlying graph is the complete graph instead of the lattice. However, it has been shown that for the ferromagnetic Potts model, a closely related pairwise influence model in which there are a larger number of spins, tempering can also require exponential time [2]. This shows that annealing algorithms must be very carefully analyzed to determine whether they can be used for sampling.

Since it is known that locally defined chains cannot be efficient for sampling Ising configurations at low temperature, this suggests that perhaps we should look for alternate algorithms based on nonlocal chains. Jerrum and Sinclair found such an algorithm for estimating the partition function of any Ising model at any temperature [9], and Randall and Wilson showed that their algorithm can be used to sample Ising configurations [18]. These results demonstrate that the interplay between statistical physics and computer science can bring to light new ideas in the world of sampling. It is not known whether such results can be extended to the non-saturated model. Moreover, the sampling algorithm

of [18] is very indirect and does not provide a natural algorithm in the context of self-assembly. It seems worthwhile to continue to look for alternative approaches that would be more useful in that context.

References

[1] D. Aldous. Random walks on finite groups and rapidly mixing Markov chains. *Séminaire de Probabilités* XVII, *Springer Lecture Notes in Mathematics* **986**:243–297, 1982.

[2] N. Bhatnagar and D. Randall. Torpid mixing of simulated tempering on the Potts model. *Proc. 15th ACM/SIAM Symposium on Discrete Algorithms (SODA)*, 478-487, 2004.

[3] R. Bubley and M. Dyer. Faster random generation of linear extensions. *Discrete Mathematics*, **201:**81-88, 1999.

[4] A. Broder. Generating random spanning trees. Proc. 30th IEEE Symp. on Foundations of Computer Science 442-447, 1989.

[5] B.A. Cipra. An introduction to the Ising model. *American Mathematical Monthly* **94**, pp. 937–959, 1987.

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

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

[8] M.R. Jerrum and A.J. Sinclair. The Markov chain Monte Carlo method: an approach to approximate counting and integration. In *Approximation Algorithms for NP-Hard Problems,* D.S. Hochbaum, ed., PWS Publishing, Boston: 482–520, 1997.

[9] M.R. Jerrum and A.J. Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM J. on Computing,* **22**: 1087–1116, 1993.

[10] R. Kannan. Markov chains and polynomial time algorithms. *Proc. 35th IEEE Symp. on Foundations of Computer Science (FOCS),* 656–671, 1994.

[11] G.F. Lawler and A.D. Sokal, A.D. Bounds on the L2 spectrum for Markov chain and Markov processes: a generalization of Cheeger's inequality. *Trans. Amer. Math. Soc.*, 309: 557-580, 1988.

[12] L. Lovász and P. Winkler. Mixing times. *Microsurveys in Discrete Probability*, D. Aldous and J. Propp, eds., *DIMACS Series in Discrete Math. and Theoretical Computer Science,* **41**: 85-134, 1998.

[13] N. Madras and Z. Zheng. On the swapping algorithm. *Random Structures and Algorithms*, 22: 6697, 2003.

[14] F. Martinelli. Lectures on Glauber dynamics for discrete spin models. *Lectures on probability theory and statistics* (Saint-Flour, 1997) in *Lecture Notes in Math.*, **1717**: 93–191. Springer, Berlin, 1999.

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

[16] D. Randall. Mixing. *Proc. 44th IEEE Symp. on Foundations of Computer Science (FOCS),* 4–15, 2003.

[17] D. Randall. Slow mixing of Glauber dynamics via topological obstructions. *Proc. 17th ACM/SIAM Symposium on Discrete algorithms (SODA)*, 870–879, 2006.

[18] D. Randall and D.B. Wilson. Sampling Spin Configurations of an Ising System. *Proc. 10th ACM/SIAM Symposium on Discrete Algorithms (SODA)*, S959–960, 1999.

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

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

[21] A.J. Sinclair. Convergence rates for Monte Carlo experiments. *Numerical Methods for Polymeric Systems,* S.G. Whittington, ed., *IMA Volumes in Mathematics and its Applications,* 1-18, 1997.

[22] A.J. Sinclair. *Algorithms for random generation & counting: a Markov chain approach.* Birkhäuser, Boston, 1993.

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

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

[25] E. Winfree, X. Yang and N.C. Seeman. Universal compuation 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.

[26] H. Wang. Dominoes and the AEA case of the decision problem. In *Proc. Symp. Math. Theory of Automata* 23–55, 1963.