TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 357, Number 8, Pages 3013–3029
S 0002-9947(05)03610-X
Article electronically published on March 10, 2005

MIXING TIMES OF THE BIASED CARD SHUFFLING AND THE ASYMMETRIC EXCLUSION PROCESS

ITAI BENJAMINI, NOAM BERGER, CHRISTOPHER HOFFMAN,
AND ELCHANAN MOSSEL

Abstract. Consider the following method of card shuffling. Start with a deck of N cards numbered 1 through N. Fix a parameter p between 0 and 1. In this model a "shuffle" consists of uniformly selecting a pair of adjacent cards and then flipping a coin that is heads with probability p. If the coin comes up heads, then we arrange the two cards so that the lower-numbered card comes before the higher-numbered card. If the coin comes up tails, then we arrange the cards with the higher-numbered card first. In this paper we prove that for all $p \neq 1/2$, the mixing time of this card shuffling is $O(N^2)$, as conjectured by Diaconis and Ram (2000). Our result is a rare case of an exact estimate for the convergence rate of the Metropolis algorithm. A novel feature of our proof is that the analysis of an infinite (asymmetric exclusion) process plays an essential role in bounding the mixing time of a finite process.

1. Introduction

The Metropolis algorithm is a widely-used algorithm for sampling from distributions on large finite sets. There are a variety of techniques which are useful to analyze the convergence rate of the Metropolis algorithm (see [8]). Yet in many problems arising in applications we do not know how to estimate the convergence rate. In this paper we introduce new techniques which allow the analysis of the Metropolis algorithm on a distribution which was not amenable to the standard techniques in the field.

Card shuffling procedures provide a natural family of Markov chains which played a crucial role in the development of the theory of the convergence rate of Markov chains (see e.g. [2, 5, 6]). In this paper we analyze the mixing time of a *biased* card shuffling procedure which has a nonuniform stationary distribution.

1.1. The biased card shuffling chains.

As the biased card shuffling has a finite state space, we formulate it as a discrete time Markov chain.

Definition 1.1. For $0 \leq p \leq 1$, let $\mathcal{CA}_d(N,p)$ denote the following discrete time Markov chain on permutations of N cards labelled $1, \ldots, N$. A step of the chain consists of selecting uniformly at random a pair of adjacent cards and then flipping a coin that is heads with probability p. If the coin comes up heads, then we arrange the two cards so that the lower-numbered card comes before the higher-numbered card. If the coin comes up tails, then we arrange the cards with the higher-numbered card first.

Received by the editors June 10, 2003 and, in revised form, June 24, 2003.
2000 *Mathematics Subject Classification.* Primary 60J10, 60K35.

Note that if $p = 1/2$, then the stationary distribution for $\mathcal{CA}_d(N, p)$ is the uniform distribution on S_N (the set of all permutations on N elements), but if $p \neq 1/2$, then the stationary distribution is not uniform.

A novel feature of our results is that the heart of the proof of the mixing result for the finite card shuffling model involves the analysis of infinite processes on $\mathbb{Z}$. Since infinite processes are naturally defined in continuous time, we use a continuous time version of the biased card shuffling model.

Definition 1.2. For $0 \leq p \leq 1$, let $\mathcal{CA}(N, p)$ denote the following continuous time Markov chain on permutations of N cards labelled $1, \dots, N$. Each pair of adjacent cards i, $i+1$ is picked with rate 1 independently. Then we toss a coin which is heads with probability p. If the coin comes up heads, then we arrange the two cards so that the lower-numbered card comes before the higher-numbered card. If the coin comes up tails, then we arrange the cards with the higher-numbered card first.

Diaconis and Ram [7] were interested in the following slightly different chain:

Definition 1.3. For $0.5 \leq p \leq 1$, let $q = 1 - p$ and let $\theta = q/p$. The **Metropolis biased card shuffling** is the following discrete time Markov chain on permutations of N cards labelled $1, \dots, N$. A step of the chain starts with selecting uniformly at random a pair of adjacent cards. If the two cards are arranged in a decreasing order, then we switch them. If they are arranged in an increasing order, then with probability θ we switch them and with probability $1 - \theta$ we do nothing.

1.2. **Main results.** The **total-variation** distance between measures μ and ν on a finite space X is

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{x \in X} |\mu(x) - \nu(x)| = \sup_{A \subset X} |\mu(A) - \nu(A)|.$$

Now we define the mixing time of a Markov chain σ on a finite state space X. For any $x \in X$ let x_t be the distribution on X at time t under the action of σ. The **mixing time** of the Markov chain is defined by

$$\tau_1 = \inf\bigl\{t \colon \sup_{x, x' \in X} \|x_t - x'_t\|_{TV} \leq e^{-1}\bigr\}.$$

Our main result is

Theorem 1.4. *For all $p \neq 1/2$ there exists a constant $K = K(p)$ such that the mixing time of the discrete time biased card shuffling on N cards is at most KN^2.*

Corollary 1.5. *For all $p > 1/2$ there exists a constant $K = K(p)$ such that the mixing time of the Metropolis biased card shuffling on N cards is at most KN^2.*

Proof. The discrete time card shuffling is a slow down of the Metropolis card shuffling. More precisely, consider the following process on S_N: At every time, with probability $1 - p$ we do nothing, and with probability p we do a step of the Metropolis biased card shuffling. This process is the discrete time card shuffling. □

This verifies a conjecture of Diaconis and Ram [7]. A lower bound for the mixing time of the form N^2 is easy and well known.

As the only difference between the discrete time card shuffling and the continuous time card shuffling is that the continuous time process is "$N - 1$ times faster", the following result is equivalent.

Theorem 1.6. *For all $p \neq 1/2$ there exists a constant $K = K(p)$ such that the mixing time of the continuous time biased card shuffling on N cards is at most KN.*

As our proofs are all done with continuous time processes, we will prove Theorem 1.6 and derive Theorem 1.4 as a corollary.

1.3. **Motivations and related results.** When running the Metropolis algorithm for sampling from distributions on large finite sets, it is important that the algorithm converges rapidly. Various techniques were developed in order to bound the convergence rate of the Metropolis algorithm in different situations (see Subsection 1.5), yet in many cases none of these methods apply. Such an example is the biased card shuffling chain (see Subsection 1.5) which we analyze in a novel way in this paper.

Our result also suggests an interesting comparison between the "systematic scan" and the "random scan" heuristics in sampling (see e.g. [10] or [7]). In [7] Diaconis and Ram studied a different version of biased card shuffling. In their model the selection of the pair of adjacent cards was not random, but done in a prescribed deterministic manner ("systematic scan"). Their discrete time model, like ours ("random scan"), has a mixing time of $O(N^2)$ for $p \neq 1/2$. Our result may be interpreted as saying that the "systematic scan" does not give an improvement over the "random scan".

In [9] the authors introduce a model of computation where each comparison operation has probability $p > 1/2$ of returning the true result and probability $1 - p$ of returning a false result independently of other comparisons. The chain $\mathcal{C}A_d(N, 1-p)$ ($\mathcal{C}A_d(N, p)$) is performing the randomized version of bubble sort in this noisy computation model. Our result shows the robustness to noise of the randomized bubble sort algorithm as the convergence time of $\mathcal{C}A_d(N, 1-p)$ is $O(N^2)$ for all $p > 1/2$.

We would also like to remark that asymmetric exclusion processes, which are the key tool in our proof, also play a crucial role in the study of the quantum Heisenberg model.

We conclude this subsection by discussing some of the history of card shuffling problems. Gilbert, Shannon and Reed began the mathematical study of card shuffling by introducing a good model for how people shuffle cards [11, 13]. The celebrated theorem of Bayer and Diaconis [3] states that for the Gilbert-Shannon-Reed model of card shuffling it takes seven shuffles in order for a standard 52-card deck to be well mixed. More generally, [3] proved that for an N-card deck the mixing time for the Gilbert-Shannon-Reed model is approximately $\frac{3}{2} \log_2 N$.

In the wake of Bayer and Diaconis's result there have been a number of articles analyzing the mixing time for various methods of card shuffling. Most relevant to this paper are results of Wilson as well as Diaconis and Ram. Wilson [17] found that the mixing time for $\mathcal{C}A_d(N, 1/2)$, to within a factor of 2, is $N^3 \log N$. Note the sharp contrast with Theorem 1.4, where we show if $p \neq 1/2$, then the mixing time of $\mathcal{C}A_d(N, p)$ is $O(N^2)$.

1.4. **The asymmetric exclusion process.** Most of the proof of our main result is devoted to analysis of the asymmetric exclusion processes. We now define these processes which are of independent interest (see e.g. [14, 15]). First we define our family of finite exclusion processes. The process $\mathcal{EX}(N, k, p)$ will be an exclusion process with N containers and k particles.

Definition 1.7. Let k and N be integers such that $1 \leq k < N$, and $0 \leq p \leq 1$. Let $\mathcal{EX}(N,k,p)$ be the continuous time Markov process defined on

$$X_{n,k} = \left\{ x \in \{0,1\}^{[1,N]} \colon \sum_{i=1}^{N} x_i = k \right\}$$

in the following way. Given the current state x, each pair of coordinates i, $i+1$ of x is picked at rate 1. If $x_i = x_{i+1}$, then the chain will stay at state x; otherwise, the two coordinates i, $i+1$ will be reassigned as $(x_i, x_{i+1}) = (1,0)$ with probability p, and as $(x_i, x_{i+1}) = (0,1)$ with probability $1-p$.

We would like to emphasize that while the card shuffling process $\mathcal{CA}_d(N,p)$ has state space S_N, the exclusion process $\mathcal{EX}(N,k,p)$ has state space $\{0,1\}^{[1,N]}$.

Definition 1.8. Let $0 \leq p \leq 1$. Let $\mathcal{EX}(\mathbb{Z},p)$ be the continuous time Markov process defined on $\{0,1\}^{\mathbb{Z}}$ in the following way. Given the current state x, each pair of coordinates i, $i+1$ of x is picked at rate 1. If $x_i = x_{i+1}$, then the chain will stay at state x; otherwise, the two coordinates i, $i+1$ will be reassigned as $(x_i, x_{i+1}) = (1,0)$ with probability p, and as $(x_i, x_{i+1}) = (0,1)$ with probability $1-p$.

We are particularly interested in the set

$$A = \left\{ a \colon \sum_{-\infty}^{-1} (1-a_i) = \sum_{0}^{\infty} a_i < \infty \right\}.$$

There is a partial order on the set. We write $a \succeq b$ if for all r

$$\sum_{i=-\infty}^{r} 1 - a_i \leq \sum_{i=-\infty}^{r} 1 - b_i. \tag{1}$$

The maximal state in A is the ground state

$$G_{\mathbb{Z}}(i) = \begin{cases} 1, & i < 0, \\ 0, & i \geq 0. \end{cases} \tag{2}$$

The aspect of asymmetric exclusion processes that we are most interested in is the tail of the hitting time. Given any $x \in A$ (or measure μ on A) the **hitting time**, $H(x)$ (**or** $H(\mu)$), is defined by

$$H(x) = \inf\{t \colon x_t = G_{\mathbb{Z}}\}. \tag{3}$$

In particular we want to consider $H(I_N)$, where

$$I_N(i) = \begin{cases} 1, & i < -N, \\ 0, & i \in [-N,-1], \\ 1, & i \in [0, N-1], \\ 0, & i \geq N. \end{cases} \tag{4}$$

Theorem 1.9. *For all $p > 1/2$ and $\epsilon > 0$ there exists a constant $D = D(p,\epsilon)$ s.t.*

$$\mathbf{P}(H(I_N) < DN) > 1 - \frac{\epsilon}{N}.$$

In Section 2 we show that Theorem 1.9 implies Theorem 1.6. Most of the work in this paper is in proving Theorem 1.9.

1.5. **Remarks on analytic techniques.** When applying standard analytic techniques to the study of the biased card shuffling problem we encounter several problems which prevent us from obtaining sharp bounds for the mixing time. We discuss briefly the difficulties in estimating the mixing time in terms of the spectral gap and the log Sobolev constant. The results of [7] also suggest that a group theoretic approach is hard to apply to this particular problem.

The standard bound for the mixing time in terms of the spectral gap of the generator of the Markov process, $-\lambda_2$ (see [16] for background), yields

$$\tau \geq \frac{2}{-\lambda_2} \log \frac{1}{\min_x \pi(x)}, \tag{5}$$

where π is the stationary distribution. Moreover, combining our reduction from the card shuffling to the exclusion process with the bound on the spectral gap of the (continuous time) exclusion process in [4], it is straightforward to verify that there exists positive c_1 and c_2 such that indeed the spectral gap satisfies $c_1 \leq -\lambda_2(N) \leq c_2$ for all N. However, the probability space contains elements of very small probability, so the term $\log(1/\min_x \pi(x))$ is of order N^2 (see [7], where the stationary distribution for the Metropolis chain is given). Thus (5) yields a bound of order N^2.

A standard way to reduce the dependency on the smallest probability is to use the log Sobolev constant a instead of the spectral gap with the estimate

$$\tau \leq \frac{4}{\alpha} \log_+ \log \frac{1}{\min_x \pi(x)}. \tag{6}$$

However, plugging the indicator of the set which consists of a single element $(N \cdots 1)$ in the variational formula of the log Sobolev constant (see [16]) implies that (for the continuous time model) $\alpha = O(1/N^2)$. (We use the notation $f(N) = O(N)$ $(f(N) = \Omega(N))$ if there exists a constant $c < \infty$ $(c > 0)$ such that $f(N) < cN$ $(f(N) > cN)$.) Thus (6) does not give the right bound of N on the mixing time.

1.6. **Road map.** We conclude the Introduction with an overview of the main steps of the proof.

1. In Section 2 we show how Theorem 1.9 implies Theorems 1.6 and 1.4. The reduction follows [17] in using **height functions**, together with coupling arguments. The height functions provide a coupling of the biased card shuffling to the finite exclusion processes. This reduces the problem of bounding the mixing time for the biased card shuffling to bounding the tail of the hitting times of some finite exclusion process. Then we couple the finite exclusion processes with the infinite exclusion process. This reduces the problem to bounding the tail of $H(I_N)$. Now we can use some of the machinery developed for the study of exclusion processes on $\mathbb{Z}$.
2. In Section 3 we define the blocking measure Ψ on $\{0,1\}^{\mathbb{Z}}$. We show that Ψ is an invariant measure for $\mathcal{EX}(\mathbb{Z}, p)$. In Section 5 we will see how to bound the tail of $H(I_N)$ in terms of the tail of $H(\Psi)$.
3. In Section 4 we introduce an asymmetric exclusion process with second-class particles. Second-class particles are a common tool used in the study of exclusion processes (see e.g. [15]). This will be our main tool for proving Theorem 1.9. We will also discuss some of the processes related to the asymmetric exclusion process with second-class particles.

4. In Section 5 we use exclusion processes with second-class particles to bound the tail of $H(I_N)$ in terms of the tail of $H(\Psi)$. Then we bound the tail of $H(\Psi)$. This bound allows us to prove Theorem 1.9, which in turn allows us to bound the mixing time of the biased card shuffling process.

2. Coupling card shuffling to exclusion processes

In this section we show how Theorem 1.9 implies Theorem 1.6. This reduces bounding the mixing time for the biased card shuffling to bounding the tail of $H(I_N)$. Following [17], we use the following collection of height functions to map a permutation of a deck of cards to an exclusion process configuration.

For any k, $1 \le k < N$, consider the map $h_k\colon S_N \to X_{N,k}$ defined by

$$(h_k(\pi))_i = \begin{cases} 1, & \pi_i \le k, \\ 0, & \pi_i > k. \end{cases}$$

It is easy to see that

Claim 2.1. *π is determined by $(h_k(\pi))_{k=1}^{N-1}$.*

For $\pi \in S_N$, we write π_t for the random variable representing the value of the process at time t that starts at π and evolves according to $\mathcal{CA}(N,p)$. Similarly for $x \in X_{N,k}$ (or $\{0,1\}^{\mathbb{Z}}$) we let the random variable x_t represent the configuration at time t for the exclusion process that started at x and evolves according to $\mathcal{EX}(N,k,p)$ ($\mathcal{EX}(\mathbb{Z},p)$).

Claim 2.2. *For all N and k, the processes $h_k(\pi_t)$ are Markovian. Moreover, $h_k(\pi_t)$ evolves according to the exclusion process $\mathcal{EX}(N,k,p)$.*

Throughout the paper we will define a number of couplings. The most important of these we refer to as **canonical couplings**. The idea of the canonical couplings is that we have a collection of initial conditions of permutations (or exclusion process states) and we use one set of clocks and one set of biased coin flips to update all of the processes simultaneously.

We begin by defining a coupling of the process $\mathcal{CA}(N,p)$ for all the configurations $\pi \in S_N$ in the following way: A transition of shuffling is to be performed by choosing a pair of adjacent coordinates i, $i+1$ at rate 1, and tossing a coin X which is heads with probability p. If $X = H$, then we rearrange the cards in coordinates i, $i+1$ in increasing order, while if $X = T$, then we rearrange the cards in coordinates i, $i+1$ in decreasing order. The same pair of coordinates i, $i+1$ and the same coin X are chosen for all $\pi \in S_n$ simultaneously. We call this coupling the **canonical coupling** for $\mathcal{CA}(N,p)$.

We can similarly define a coupling for all configurations in $\mathcal{EX}(N,k,p)$. A transition is to be performed by choosing a pair of adjacent coordinates i, $i+1$ in $\mathbb{Z}$ at rate 1 and tossing a coin X which is heads with probability p. For a state x of $\mathcal{EX}(\mathbb{Z},p)$ we will update x as follows. If $x_i = x_{i+1}$, then we do nothing. Otherwise, if $X = H$, we let $x_i = 0$, $x_{i+1} = 1$, and if $X = T$, we let $x_{i+1} = 0$, $x_i = 1$. Again, the same pair of coordinates and the same coin is chosen for all states of $\mathcal{EX}(N,k,p)$. We call this coupling the **canonical coupling** for $\mathcal{EX}(N,k,p)$. In the same way we define a canonical coupling for $\mathcal{EX}(\mathbb{Z},p)$.

It is immediate to verify that

Claim 2.3. *For all N, k and p, the canonical couplings for $\mathcal{CA}(N,P)$, $\mathcal{EX}(N,k,p)$ and $\mathcal{EX}(\mathbb{Z},p)$ are all well defined and have the right marginals.*

Moreover for all N, k and p, the map h_k maps the canonical coupling of $\mathcal{CA}(N,p)$ to the canonical coupling of $\mathcal{EX}(N,k,p)$, i.e., if $(\pi_t)_{\pi\in S_N,\ t\geq 0}$ evolves according to the canonical coupling for S_N, then the process

$$(\{h_k(\pi_t)\colon \pi \in S_N\})_{t\geq 0}$$

has the same distribution as the process

$$(\{\sigma_t\colon \sigma \in X_{N,k}\})_{t\geq 0},$$

where $(\sigma_t)_{\sigma\in X_{N,k},t\geq 0}$ evolve according to the canonical coupling for $\mathcal{EX}(N,k,p)$.

The analysis of the process $\mathcal{EX}(N,k,p)$ utilizes monotonicity properties of the canonical coupling. For $a,b\in X_{N,k}$ we write $a\succeq b$, if for all r, $\sum_{i=1}^r a_i \geq \sum_{i=1}^r b_i$. The maximal state with respect to this partial order is

$$g_{N,k}(i) = \begin{cases} 1, & i\leq k,\\ 0, & i>k,\end{cases}$$

and the minimal state with respect to this partial order is

$$m_{N,k}(i) = \begin{cases} 0, & i\leq N-k,\\ 1, & i>N-k.\end{cases}$$

We let $H(N,k)$ be the hitting time of the state $g_{N,k}$ for the process $\mathcal{EX}(N,k,p)$ started at $m_{N,k}$.

It is immediate to see that

Claim 2.4. *The canonical couplings for $\mathcal{EX}(N,k,p)$ and $\mathcal{EX}(\mathbb{Z},p)$ are monotone. That is, for both processes if $x\succeq y$, then for all t it holds that $x_t\succeq y_t$.*

Since $g_{N,k}$ and $m_{N,k}$ are the maximal and minimal elements with respect to $\succeq$ it follows that

Claim 2.5. *Under the canonical coupling for $\mathcal{EX}(N,k,p)$ it holds that*

$$\mathbf{P}(\exists x,y\in X_{N,k} \text{ s.t. } x_t\neq y_t) = \mathbf{P}(H(N,k)>t).$$

Lemma 2.6. *Under the canonical coupling for $\mathcal{CA}(N,p)$ it holds that*

$$\mathbf{P}(\exists \sigma,\tau\in S_N \text{ s.t. } \sigma_t\neq \tau_t) \leq \sum_{k=1}^{N-1}\mathbf{P}(H(N,k)>t).$$

Proof.

(7)

$$\mathbf{P}(\exists \sigma,\tau\in S_N \text{ s.t. } \sigma_t\neq\tau_t) = \mathbf{P}(\exists 1\leq k\leq N-1, \sigma,\tau\in S_N \text{ s.t. } h_k(\sigma_t)\neq h_k(\tau_t))$$

(8)

$$\leq \sum_{k=1}^{N-1}\mathbf{P}(\exists x,y\in X_{N,k} \text{ s.t. } x_t\neq y_t) = \sum_{k=1}^{N-1}\mathbf{P}[H(N,k)>t],$$

where (7) follows from Claim 2.1 and (8) follows from Claims 2.3 and 2.5. □

The remaining coupling step is to couple the finite processes to an infinite process.

Lemma 2.7. *For all $p \geq 1/2$, all N, all $1 \leq k < N$, and all $t > 0$ the processes $\mathcal{EX}(N,k,p)$ and $\mathcal{EX}(\mathbb{Z},p)$ satisfy that*

$$\mathbf{P}(H(N,k) > t) \leq \mathbf{P}(H(I_N) > t).$$

Proof. Consider the map $X_{N,k} \to \{0,1\}^{\mathbb{Z}}$ sending $x \in X_{N,k}$ to $\hat{x}$ with

$$\hat{x}(i) = \begin{cases} 1, & i < -k, \\ x(i+k+1), & i \in [-k, N-k-1], \\ 0, & i \geq N-k. \end{cases} \tag{9}$$

We will now couple $\mathcal{EX}(N,k,p)$ to the process $\mathcal{EX}(\mathbb{Z},p)$. More formally, we will couple the canonical coupling of $\mathcal{EX}(N,k,p)$ with the canonical coupling of $\mathcal{EX}(\mathbb{Z},p)$.

For the process $\mathcal{EX}(\mathbb{Z},p)$ we pick a pair of coordinates i, $i+1$ at rate 1 and then use a coin X to rearrange the coordinates i, $i+1$ in the usual manner. Now, if $1 \leq i+k+1 \leq i+k+2 \leq N$, then we use the same coin X to rearrange the coordinates $i+k+1$, $i+k+2$ for the process $\mathcal{EX}(N,k,p)$.

Clearly this coupling is well defined and has the right marginals. It is easy to see that if $p \geq 1/2$, then this coupling has the following important property: For all $x \in X_{n,k}$ and all $t \geq 0$, it holds that $\widehat{(x_t)} \succeq (\hat{x})_t$.

Writing $I = N_n$, $g = g_{N,k}$, and $m = m_{N,k}$, and since $\hat{m} \succeq I$, it follows that under this coupling, if $I_t = G_{\mathbb{Z}}$, then

$$\widehat{(m_t)} \succeq (\hat{m})_t \succeq I_t = G_{\mathbb{Z}},$$

and therefore $m_t = g$. The claim of the lemma follows. □

We end this section by noting that the canonical coupling of the biased card shuffling shows that with high probability by time DN all of the processes agree.

Lemma 2.8. *Theorem* 1.9 *implies that the canonical copuling for $\mathcal{CA}(N,p)$ has*

$$\mathbf{P}(\exists \sigma, \tau \in S_N \text{ s.t. } \sigma_{DN} \neq \tau_{DN}) < \epsilon. \tag{10}$$

In particular Theorem 1.9 *implies Theorems* 1.6 *and* 1.4.

Proof. By Theorem 1.9, $\mathbf{P}[H(I_N) > DN] < \epsilon/N$, and therefore by Lemma 2.7 for all k it holds that $\mathbf{P}[H(N,k) > DN] < \epsilon/N$. It now follows from Lemma 2.6 that

$$\mathbf{P}(\exists \sigma, \tau \in S_N \text{ s.t. } \sigma_{DN} \neq \tau_{DN}) < (N-1)\epsilon/N - \epsilon,$$

so we obtain (10). Taking $\epsilon = e^{-1}$, we deduce Theorem 1.6 which immediately implies Theorem 1.4. □

3. The blocking measure

In this section we define a distribution Ψ on $\{0,1\}^{\mathbb{Z}}$ which is invariant under the action of $\mathcal{EX}(\mathbb{Z},p)$. Ψ is known as the **blocking measure** (see e.g. [14]). In Section 5 we will bound the tail of $H(\Psi)$ and show that the tail of $H(I_N)$ can be bounded in terms of the tail of $H(\Psi)$.

Fix any $p > 1/2$. Define $\mu = \mu(p)$ on $\{0,1\}^{\mathbb{Z}}$ to be the product measure with probabilities

$$\mu(\eta(i) = 1) = \left(\frac{1-p}{p}\right)^i \Big/ \left(1 + \left(\frac{1-p}{p}\right)^i\right). \tag{11}$$

Lemma 3.1 ([14]). *The measure μ is stationary for $\mathcal{EX}(\mathbb{Z}, p)$.*

Proof. This is proven on page 381 of [14]. □

Notice that μ is supported on configurations η s.t. there exists a C_η s.t. $\eta(i) = 1$ for every $i < -C_\eta$ and $\eta(i) = 0$ for every $i > C_\eta$. There are only countably many configurations of this type, and each of them has a positive measure. We have already defined

$$A = \left\{a\colon \sum_{-\infty}^{-1}(1 - a_i) = \sum_{0}^{\infty} a_i < \infty\right\}.$$

Definition 3.2. The blocking measure Ψ on $\{0,1\}^{\mathbb{Z}}$ is defined by

$$\Psi = \mu|_A. \tag{12}$$

Corollary 3.3. *Ψ is stationary and ergodic for the exclusion process.*

By Poincaré's recurrence theorem and the fact that $\Psi(G_{\mathbb{Z}}) > 0$, we get the following lemma.

Lemma 3.4. $\lim_{T\to\infty} \mathbf{P}(H(\Psi) > T) = 0$.

In Lemma 5.11 we will show that $\mathbf{P}(H(\Psi) > T) = e^{-\Omega(\sqrt{T})}$.

Lemma 3.5.

$$\Psi(\exists_{i>N}(\eta(i) = 1)) = \Psi(\exists_{i<-N}(\eta(i) = 0)) = O\left(\left(\frac{1-p}{p}\right)^N\right). \tag{13}$$

For any $T > 0$

$$\mathbf{P}(\exists t \in (T, T+N) \text{ and } i > 2N \text{ such that } \eta_t(i) = 1) = e^{-\Omega(N)}. \tag{14}$$

Proof. In the product measure, the μ probability that there exists an occupied site right of position N is bounded by

$$\sum_{i=N+1}^{\infty} \frac{((1-p)/p)^i}{1 + ((1-p)/p)^i} \leq \sum_{i=N+1}^{\infty} \left(\frac{1-p}{p}\right)^i = O\left(\left(\frac{1-p}{p}\right)^N\right).$$

Since Ψ is obtained from the product measure by conditioning on an event of positive probability, the first part of the lemma is true. For the second part, if there exists $t \in (T, T+N)$ and $i > 2N$ such that $\eta_t(i) = 1$, then either

1. there exists $i \geq N$ such that $\eta_T(i) = 1$, or
2. for some t

$$\max\{i\colon \eta_t(i) = 1\} - \max\{i\colon \eta_T(i) = 1\} > N.$$

By the first part of the lemma the probability of the first event is decreasing exponentially in N. The second event happens only if the right most particle moves to the right N times in a period of length N. As moves to the right happen with rate $1 - p < \frac{1}{2}$ the probability that this happens is also decreasing exponentially in N. □

4. Exclusion processes with second-class particles

The main tool that we will use in the rest of the paper is adding second-class particles to our exclusion process. We now describe some of the basics about exclusion processes with second-class particles. For a more rigorous treatment of exclusion processes with second-class particles see [15].

Definition 4.1. Let $0 \le p \le 1$. Let $\mathcal{EX}_2(\mathbb{Z}, p)$ be the continuous time Markov process defined on $\{0,1,2\}^{\mathbb{Z}}$ in the following way. Given the current state x, each pair of coordinates i, $i+1$ of x is picked at rate 1. If $x_i = x_{i+1}$, then the chain will stay at state x. If the two coordinates i, $i+1$ initially are $(0,1)$ or $(1,0)$, then they will be reassigned as $(x_i, x_{i+1}) = (1,0)$ with probability p, and as $(x_i, x_{i+1}) = (0,1)$ with probability $1-p$. If initially they are $(0,2)$ or $(2,0)$, then they will be reassigned as $(2,0)$ with probability p, and as $(0,2)$ with probability $1-p$. If initially they are $(1,2)$ or $(2,1)$, then they will be reassigned as $(1,2)$ with probability p, and as $(2,1)$ with probability $1-p$.

If $x_i = 1$, then we say that there is a first-class particle in position i, if $x_i = 2$, then we say that there is a second-class particle in position i, and if $x_i = 0$, then we say that the site i is empty.

It is helpful to have in mind the following ordering of 0, 1 and 2. Particle 1 has priority over 0 and 2 in moving to the left. Particle 2 has priority over 0 (but not over 1) in moving to the left. Particle of type 2 is therefore ranked in-between particle of type 0 and particle of type 1.

Is is therefore natural to consider to the following two projections. In the first projection $\delta^{2\to 1}$, 2's are projected to 1's, while in the second projection $\delta^{2\to 0}$, 2's are projected to 0's. More formally,

$$\delta_t^{2\to 1}(i) = \begin{cases} 0, & \delta_t(i) = 0, \\ 1, & \delta_t(i) > 0, \end{cases}$$

and

$$\delta_t^{2\to 0}(i) = \begin{cases} 0, & \delta_t(i) \ne 1, \\ 1, & \delta_t(i) = 1. \end{cases}$$

Claim 4.2. *Both $\delta^{2\to 1}$ and $\delta^{2\to 0}$ evolve according to $\mathcal{EX}(\mathbb{Z}, p)$.*

The next process we consider represents the dynamics between particles of type 1 and particles of type 2 and eliminates all the information on the 0's.

To define $\delta_t^{\otimes}$ we first eliminate all of the zeroes from δ_t, and then change all of the twos to zeroes. This is only well defined up to translation, so we must also decide which translate we want. We do this by tagging one particle in δ_t and having $\delta_t^{\otimes}(0)$ correspond to the tagged particle.

We now make this more formal. Let

$$u_0(0) = \begin{cases} \sup\{i \colon \delta_0(i) = 1\}, & \text{if } \sup\{i \colon \delta_0(i) = 1\} < \infty, \\ \sup\{i < 0 \colon \delta_0^1(i) = 1\}, & \text{otherwise.} \end{cases}$$

We refer to the particle which is in position $u_0(0)$ at time 0 as the **tagged particle**. Let $u_t(0)$ be the location of the tagged particle at time t. For $n = 1, 2, \dots$, let

$$u_t(n) = \min\{i > u_t(n-1) \colon \delta_t(i) > 0\}$$

and

$$u_t(-n) = \max\{i < u_t(-n+1) \colon \delta_t(i) > 0\}.$$

Thus $u_t(n)$ represents the location of the nth particle at time t. Let

$$\delta_t^{\otimes}(i) = \begin{cases} 0, & \delta_t(u_t(i)) = 2, \\ 1, & \delta_t(u_t(i)) = 1. \end{cases} \tag{15}$$

Lemma 4.3. *If δ evolves according to $\mathcal{EX}_2(\mathbb{Z}, p)$ and has initial distribution $\delta_0^{\otimes}$ stochastically dominating Ψ, then $\delta_t^{\otimes}$ stochastically dominates Ψ for all t.*

Proof. In [14] it is shown that for all $i \in \mathbb{Z}$, Ψ is invariant under the Markov operator on $\{0,1\}^{\mathbb{Z}}$ which at rate 1 tosses a coin X which heads with probability p. Then if $x_i \neq x_{i+1}$ and $X = H$, then x_i and x_{i+1} are updated as $x_i = 1$, $x_{i+1} = 0$, and if $x_i \neq x_{i+1}$ and $X = T$, then x_i and x_{i+1} are updated as $x_i = 0$, $x_{i+1} = 1$. This implies that Ψ is an invariant measure for the process $\delta^{\otimes}$. It now immediately follows that if $\delta_0^{\otimes}$ stochastically dominates Ψ, then $\delta_t^{\otimes}$ stochastically dominates Ψ for all t. □

5. Proof of the main results

Let $\{Y_i\}_{i\in\mathbb{Z}}$ be i.i.d. random variables s.t. $\mathbf{P}(Y_i = 0) = \mathbf{P}(Y_i = 1) = 1/2$, and let $Z_i = 2Y_i$. The main tool to prove Theorem 1.9 is to study $\mathcal{EX}_2(\mathbb{Z}, p)$ with initial conditions

$$\sigma_0(i) = \begin{cases} 1, & i < -N, \\ 0, & i \in [-N, -1], \\ 1, & i \in [0, N-1], \\ Z_i, & i \geq N. \end{cases} \tag{16}$$

This is useful for proving Theorem 1.9 because $\sigma_t^{2\to 0} = (I_N)_t$.

For any $a \in \{0,1\}^{\mathbb{Z}}$ such that $\lim_{i\to-\infty} a_i = 1$ we set

$$L(a) = \min\{i \colon a_i = 0\}. \tag{17}$$

This indicates the left most empty position. In the same way, for a s.t. $\lim_{i\to\infty} a_i = 0$, we indicate the right most particle by

$$R(a) = \max\{i \colon a_i = 1\}.$$

For a constant C we define three events:

$$A_1(C,N) = \{\forall t \in (CN, (C+1)N) L(\sigma_t^{2\to 1}) > 2N\}, \tag{18}$$

$$A_2(C,N) = \{\forall t \in (CN, (C+1)N) R(\sigma_t^{\otimes}) < 2N\}, \tag{19}$$

and

$$A_3(C,N) = \{\exists t \in (CN, (C+1)N) \text{ such that } \sigma_t^{\otimes} = G_{\mathbb{Z}}\}. \tag{20}$$

Lemma 5.1.

$$\begin{aligned} &\mathbf{P}(H(I_N) \leq (C+1)N) \\ &\geq 1 - \mathbf{P}(A_3^c(C,N) | A_1(C,N), A_2(C,N)) - \mathbf{P}(A_1^c(C,N)) - \mathbf{P}(A_2^c(C,N)). \end{aligned}$$

Proof. Recall that by $G_{\mathbb{Z}}$ we denote the ground state (see (2)). Notice that $\sigma_t^{2\to 0} = G_{\mathbb{Z}}$ if $\sigma_t^{\otimes} = G_{\mathbb{Z}}$ and $L(\sigma_t^{2\to 1}) > 0$. Thus if $A_1(C,N)$ and $A_3(C,N)$ both occur, then there exists $t \leq (C+1)N$ such that $\sigma_t^{2\to 0} = (I_N)_t = G_{\mathbb{Z}}$. □

Although the previous lemma did not depend on the definition of $A_2(C,N)$, we use it because it is easy to bound

$$\mathbf{P}(A_3^c(C,N)|A_1(C,N),A_2(C,N))$$

in terms of the tail of $H(\Psi)$.

Lemma 5.2.

$$\mathbf{P}(A_c^3(C,N)|A_1(C,N),A_2(C,N)) \le \mathbf{P}(H(\Psi) > N).$$

Proof. If $A_1(C,N)$ and $A_2(C,N)$ both happen, then $\sigma^{\otimes}$ behaves according to $\mathcal{EX}(\mathbb{Z},p)$ conditioned on the event that there is never a particle to the right of $2N$. By Lemma 4.3 the distribution of $\sigma^{\otimes}_{CN}$ stochastically dominates Ψ. Putting these two facts together gives us $\mathbf{P}(A_3^c(C,N)|A_1(C,N),A_2(C,N)) \le \mathbf{P}(H(\Psi) > N)$. □

It is also not difficult to bound $\mathbf{P}(A_2^c(C,N))$.

Lemma 5.3. $\mathbf{P}(A_2^c(C,N)) = e^{-\Omega(N)}$.

Proof. This follows from Lemmas 3.5 and 4.3. □

Our next goal is to bound $\mathbf{P}(A_1^c(C,N))$. Then we will bound the tail of $H(\Psi)$.

In order to bound $\mathbf{P}(A_1^c(C,N))$ we first bound $\mathbf{P}(\widetilde{A}_1^c(C,N))$, where

$$\widetilde{A}_1(C,N) = \{L(\sigma_{CN}^{2\to 1}) > 3N\},$$

and use the following lemma.

Lemma 5.4. $\mathbf{P}(A_1^c(C,N)) \le \mathbf{P}(\widetilde{A}_1^c(C,N)) + e^{-\Omega(N)}$.

Proof. If $\widetilde{A}_1^c(C,N)$ happens but $A_1^c(C,N)$ does not, then the left most container without a particle moves to the left at least N times in time N. For that to happen, the clock left of this container has to ring N times. But since its rate is smaller than 1, by simple large deviation estimates for Poisson variables, the probability of this happening is decreasing exponentially in N. □

To bound the probability of $\widetilde{A}_1^c(C,N)$ we study the process β which has initial distribution

$$\beta_0(i) = \begin{cases} Y_i, & i \le 0, \\ Z_i = 2Y_i, & i > 0, \end{cases} \tag{21}$$

i.e. the initial distribution of β is the following: Every place contains a particle with probability 1/2, and the places are independent of each other. Left of the origin the particles are first-class particles, while right of the origin the particles are second-class particles.

We are interested in the processes $\beta^{2\to 0}$ and $\beta^{2\to 1}$. The process $\beta^{2\to 1}$ is the stationary i.i.d. process. The process $\beta^{2\to 0}$, on the other hand, is the process that starts with no particles on the right half of the line, and an i.i.d. measure on its left half.

Let $x(t)$ be the location of the tagged particle in $\beta^{2\to 0}$ at time t and let $x'(t)$ be the location of the tagged particle in $\beta^{2\to 1}$ at time t. We will bound the expectation and variance of $x(t)$.

The following lemma follows from the proof of the shock wave phenomenon in [15]. For the convenience of the reader, we prove it here again.

Lemma 5.5. *There exists $\varrho < 1$ such that under the canonical coupling, for every n and for every time t,*

$$\mathbf{P}(|x(t) - x'(t)| > n) < \varrho^n.$$

Proof. We define the processes β^ℓ and β^d (ℓ stands for locations and d stands for distances):

$\beta^\ell_t(i)$ is the location of the ith particle in $\beta^{2\to 1}$. To be more precise, $\beta^\ell_t(0) = x'(t)$, and, inductively, for positive i we take

$$\beta^\ell_t(i) = \min(j\colon j > \beta^\ell_t(i-1) \text{ and } \beta^{2\to 1}_t(j) = 1)$$

and for negative i, equivalently, we take

$$\beta^\ell_t(i) = \max(j\colon j < \beta^\ell_t(i+1) \text{ and } b^{2\to 1}_t(j) = 1).$$

$\beta^d_t(i)$ is defined to be $\beta^\ell_t(i) - \beta^\ell_t(i-1)$.

Since for every t, $\{\beta^{2\to 1}_t(i)\}_{i\in\mathbb{Z}}$ is distributed according to the $(1/2, 1/2)$ product measure, we get that for every t, $\{\beta^d_t(i)\}_{i\in\mathbb{Z}}$ are i.i.d. geometric variables with parameter $1/2$.

Let

$$s(t) = \sup\{i\colon \beta^{\otimes}(i) = 1\}.$$

Then for all t

$$x(t) - x'(t) = \sum_{i=1}^{s(t)} \beta^d_t(i).$$

By Lemmas 3.5 and 4.3

$$\mathbf{P}(s(t) > n/3) = \Psi(\exists i > n/3\colon \eta(i) = 1) = O\left(\frac{1-p}{p}\right)^{n/3}.$$

By the distribution of $\{\beta^d_t(i)\}_{i\in\mathbb{Z}}$ there exists $\alpha < 1$ such that

$$\mathbf{P}\left(\sum_1^{n/3} \beta^d_t(i) > n\right) = O(\alpha^n).$$

Thus there exists $\varrho < 1$ such that

$$\mathbf{P}(|x(t) - x'(t)| > n) \le \mathbf{P}(s(t) > n/3) + \mathbf{P}\left(\sum_1^{n/3} \beta^d_t(i) > n\right) < \varrho^n. \qquad \square$$

The following lemma is proved in [12]:

Lemma 5.6 (Kipnis)**.**

$$\lim_{t\to\infty} \frac{\mathbf{E}(x'(t))}{t} = s' = -\frac{1}{2}(2p-1),$$
$$\lim_{t\to\infty} \frac{\mathbf{Var}(x'(t))}{t} = v' \le \frac{6}{2p-1} < \infty.$$

Remark. Note that $s < 0$.

Combining Lemmas 5.6 and 5.5 and the fact that $\mathbf{E}(x(t))$ and $\mathbf{Var}(x(t))$ are continuous in t, we get

Lemma 5.7. *There exists t_0 and there exist $s' < s < 0$ and $v' < v < \infty$ s.t. for every $t \geq t_0$,*

$$\frac{\mathbf{E}(x(t))}{t} < s$$

and

$$\frac{\mathbf{Var}(x(t))}{t} < v.$$

Consider the exclusion process γ which has the initial distribution

$$\gamma_0(i) = \begin{cases} 1, & i < 0, \\ 1 - Y_i, & i \geq 0. \end{cases} \tag{22}$$

One can couple a copy γ' of γ with $\beta^{2\to 0}$ such that for all t and i

$$\gamma'_t(i) = 1 - \beta_t^{2\to 0}(-i).$$

$x(t)$ corresponds to the location of the right most particle in $\beta_t^{2\to 0}$, so the processes $-x(t)$ and $L(\gamma_t)$ have the same law.

Applying Chebychev's inequality we get the following estimate:

Lemma 5.8. *For any $\delta > 0$ and any $t > t_0$*

$$\mathbf{P}\left(x(t) \leq \left(s + \sqrt{\frac{v}{\delta}} \right) t \right) > 1 - \delta/t$$

and

$$\mathbf{P}\left(L(\gamma_t) \geq -\left(s + \sqrt{\frac{v}{\delta}} \right) t \right) > 1 - \delta/t.$$

For any l we define γ^l to be the process starting at

$$\gamma_0^l(i) = \begin{cases} 1, & i < -l, \\ 1 - Y_i, & i \geq -l. \end{cases}$$

We get from Lemma 5.8 that for any $t > t_0$,

$$\mathbf{P}\left(L(\gamma_t^l) \geq -l - \left(s + \sqrt{\frac{v}{\delta}} \right) t \right) = \mathbf{P}\left(L(\gamma_t) \geq -\left(s + \sqrt{\frac{v}{\delta}} \right) t \right) > 1 - \delta/t. \tag{23}$$

Now we are ready to bound $\mathbf{P}(\widetilde{A}_1^c)$.

Lemma 5.9. *For any $\epsilon > 0$, there exists a constant $C = C(p, \epsilon)$ such that for all $N \geq t_0 + 1$*

$$\mathbf{P}(\widetilde{A}_1^c(C, N)) < \frac{\epsilon}{N}.$$

Proof. As the canonical coupling preserves domination, if

$$\gamma_0^l \succeq \sigma_0^{2\to 0} \text{ and } L(\gamma_t^l) > 3N,$$

then $L(\sigma_t^{2\to 0}) > 3N$. This gives us that for any l and t

$$\mathbf{P}(L(\sigma_t^{2\to 0}) > 3N) \geq 1 - \mathbf{P}(\gamma_0^l \not\succeq \sigma_0^{2\to 0}) - \mathbf{P}(L(\gamma_t^l) \leq 3N). \tag{24}$$

Choose j such that for all N

$$\mathbf{P}\left(\sum_{i=1}^{jN} Y_i < N\right) < \epsilon/2N.$$

Then the lemma follows from (23) and (24) with l, δ, C and t chosen such that $l = jN$, $\delta = 4v/s^2$, $C > \max(-2(3+j)/S, 2\delta/\epsilon, 1)$, and $t = CN$.

This is because

$$\mathbf{P}(\gamma_0^{jN} \not\succeq \sigma_0^{2\to 0}) = \mathbf{P}\left(\sum_{i=1}^{jN} Y_i < N\right) < \epsilon/2N$$

and by Lemma 5.8 and (23),

$$\begin{aligned}
\mathbf{P}(L(\gamma_{CN}^{jN}) > 3N) &= \mathbf{P}(L(\gamma_{CN}) > (3+j)N) \\
&> \mathbf{P}(L(\gamma_{CN}) > -\frac{s}{2}CN) \\
&> \mathbf{P}(L(\gamma_{CN}) > -\left(s + \sqrt{\frac{v}{\delta}}\right) CN) \\
&> \delta/CN \\
&> 1 - \epsilon/2N.
\end{aligned}$$

□

Lemma 5.10. *For every $N > t_0 + 1$ and $\epsilon > 0$,*

$$\mathbf{P}(H(I_N) < (C+1)N) > 1 - \epsilon/N - \mathbf{P}(H(\Psi) > N). \tag{25}$$

Proof. This follows from Lemmas 5.1, 5.2, 5.3, 5.4, and 5.9. □

In order to prove Theorem 1.9 we first prove the following lemma:

Lemma 5.11.

$$\mathbf{P}(H(\Psi) \geq N) = e^{-\Omega(\sqrt{N})}.$$

Proof. For every N large enough, we wish to estimate the probability that $H(\Psi) \geq (C+1)N^2$, where C is the constant from Lemma 5.10. We take N large enough so that the probability in (25) is bigger than $\frac{1}{2}$. Such N exists by Lemma 3.4. Recall that

$$I_N(i) = \begin{cases} 1, & i < -N, \\ 0, & i \in [-N, -1], \\ 1, & i \in [0, N-1], \\ 0, & i \geq N. \end{cases}$$

Now, for every $j = 0, 1, 2, \ldots, N$, let $P_j = \mathbf{P}(H(\Psi) \geq (C+1)Ni)$. Of course, $P_0 = 1$. Now, we proceed inductively. Let

$$U_N = 1 - \mathbf{P}(\eta_t \succeq I_N).$$

Notice that by Lemma 4.3 it does not depend on t, and, by (13),

$$U_N = e^{-\Omega(N)}.$$

For every t,

$$\mathbf{P}(H(\Psi) \geq t + (C+1)N | \eta_t \succeq I_N) < \frac{1}{2}.$$

Therefore, $P_i \le \frac{P_{i-1}}{2} + U_N$ for every $i > 0$. Therefore,

$$\mathbf{P}(H(\Psi) \ge (C+1)N^2) = P_N \le 2^{-N} + \sum_{i=1}^{N} 2^{-i} U_N = e^{-\Omega(N)}.$$

By monotonicity we can interpolate and get that for every t

$$\mathbf{P}(H(\Psi) \ge (C+1)t) = e^{-\Omega(\sqrt{t})}$$

and thus

$$\mathbf{P}(H(\Psi) \ge t) = e^{-\Omega(\sqrt{t})}. \qquad \square$$

We can now prove our main results.

Proof of Theorems 1.4, 1.6, *and* 1.9. By Lemma 5.11, $H(\Psi) = e^{-\Omega(\sqrt{N})} = o(N^{-1})$. Therefore, by (25),

$$\mathbf{P}(H(I_N) < (C(p,\epsilon)+1)N) > 1 - \frac{\epsilon}{N} - O(N^{-1}).$$

Taking $D = C(p, \frac{\epsilon}{2}) + 1$, Theorem 1.9 is satisfied for all arbitrarily large N. Thus we can choose D so that it is true for all N. Theorems 1.4 and 1.6 follow by Lemma 2.6. $\square$

We conclude the paper with a brief comment about how $D = D(p)$ depends on p. We see in Lemma 5.6 that $s = -\frac{1}{2}(2p-1)$ and $v \le \frac{6}{2p-1}$. For large N, in Lemma 5.7 using $\epsilon = 1/e$ we can choose

$$C = \frac{8ve}{s^2} \le \frac{1024e}{(2p-1)^3}.$$

For large N we can choose

$$D = 2C \le \frac{2048e}{(2p-1)^3}.$$

It is easy to show that D must be chosen bigger than $1/(2p-1)$. The discrepancy in the power of $2p-1$ comes from the use of Chebychev's inequality in Lemma 5.8. We believe that a more careful analysis would allow one to choose D s.t. $D = \theta(1/(2p-1))$.

Acknowledgment

We thank Persi Diaconis for important comments on a previous version of this paper. We thank Dror Weitz, Prasad Tetali, Thomas Liggett and David Aldous for many interesting discussions. Part of the work was done while Noam Berger was an intern at Microsoft Research. Christopher Hoffman was partially supported by NSF grant #0100445.

References

1. D. Aldous and J. A. Fill, *Reversible Markov chains and random walks on graphs*, book in preparation, 2002.
2. D. Aldous and P. Diaconis, *Shuffling cards and stopping times*, Amer. Math. Monthly **93** (1986), no. 5, 333–348. MR0841111 (88a:60021)
3. D. Bayer and P. Diaconis, *Trailing the dovetail shuffle to its lair*, Annals of Applied Probability **2** (1992), 294–313. MR1161056 (93d:60014)
4. P. Caputo and F. Martinelli, *Relaxation time of anisotropic simple exclusion processes and quantum Heisenberg models*, Ann. Appl. Probab. **13** (2003), 691–721. MR1972083

5. P. Diaconis, *Group representations in probability and statistics*, Institute of Mathematical Statistics Lecture Notes—Monograph Series **11** (1988). MR0964069 (90a:60001)
6. P. Diaconis, *From shuffling cards to walking around the building: an introduction to modern Markov chain theory*, Proceedings of the International Congress of Mathematicians, Vol. I, Berlin, 1998, 187–204. MR1648031 (99e:60154)
7. P. Diaconis and A. Ram, *Analysis of systematic scan Metropolis algorithms using Iwahori-Hecke algebra techniques*, Michigan Math. J. **48** (2000), 157–190. MR1786485 (2001j:60132)
8. P. Diaconis and L. Saloff-Coste, *What do we know about the Metropolis algorithm*? 27th Annual STOC, J. Comput. System Sci. **57** (1998), 20–36. MR1649805 (2000b:68094)
9. U. Feige, D. Peleg, P. Raghavan and E. Upfal, *Computing with unreliable information*, STOC **90** (1990), 128–137.
10. G. Fishman, *Coordinate selection rules for Gibbs sampling*, Ann. Appl. Probab. **6** (1996), 444–465. MR1398053 (97k:62058)
11. E. Gilbert, Bell Lab technical report, 1955.
12. C. Kipnis, *Central limit theorems for infinite series of queues and applications to simple exclusion*, Ann. Probab. **14**, no. 2, (1986), 397–408. MR0832016 (88a:60173)
13. J. Reeds, unpublished manuscript, 1981.
14. T. M. Liggett, *Interacting particle systems*, Springer, 1985, Grundlehren der mathematischen Wissenschaften, 276. MR0776231 (86e:60089)
15. T. M. Liggett, *Stochastic interacting systems: contact, voter, and exclusion processes*, Springer, New York, 1999. Grundlehren der mathematischen Wissenschaften, 324. MR1717346 (2001q:60247)
16. L. Saloff-Coste, *Lectures on finite Markov chains*, Lectures on probability theory and statistics (Saint-Flour, 1996), 301–413, Lecture Notes in Math., 1665, Springer, 1996. MR1490046 (99b:60119)
17. D. Wilson, *Mixing times of lozenge tiling and card shuffling Markov chains*, Ann. Appl. Probab. **14**, no. 1, (2004), 274–325. MR2023023

Department of Mathematics, Weizmann Institute, Rehovot 76100, Israel

Department of Mathematics, California Institute of Technology, Pasadena, California 91125-0050

Department of Mathematics, University of Washington, Seattle, Washington 98195

Department of Computer Science, University of California, Berkeley, California 94720