Probab. Theory Relat. Fields 117, 407–417 (2000)
Digital Object Identifier (DOI) 10.1007/s004400000055

Don Coppersmith · Igor Pak

Random walk on upper triangular matrices mixes rapidly

Received: 25 January 1999 / Revised version: 17 September 1999 /
Published online: 14 June 2000 – © Springer-Verlag 2000

Abstract. We present an upper bound $O(n^2)$ for the mixing time of a simple random walk on upper triangular matrices. We show that this bound is sharp up to a constant, and find tight bounds on the eigenvalue gap. We conclude by applying our results to indicate that the asymmetric exclusion process on a circle indeed mixes more rapidly than the corresponding symmetric process.

Introduction

Consider the following discrete time random process. Let $G = U(n, \mathbb{F}_q)$ be a group of upper triangular $(n \times n)$-matrices over the finite field with q elements with ones on diagonal. At each step we choose a uniform row $i \in \{2, \dots, n\}$, multiply by a uniform element $a \in \mathbb{F}_q$ and add to the previous row. The question is *after how many steps do we get (nearly)-uniform elements in G ?* In other words, we would like to bound the mixing time **mix** of the process (see definitions below).

Variations and advanced generalizations of the random walk defined above have been studied in a number of papers (see [DSC2, P1, P2, S1, S2]). R. Stong showed in [S1] that the second eigenvalue λ_2 satisfies $1-c_1/n < \lambda_2 < 1-c_2/n$ as $n \to \infty$ where $c_1(q), c_2(q)$ are constants depending only on q. This gives an upper bound $\mathbf{mix} = O(n^3)$ for the mixing time. Note that the lower bound on the eigenvalue follows easily since our generators change an element in the upper right corner with probability at most $1/(n-1)$. Now let $q \gg n^2$ in a precise sense we will specify below. Until now the best known upper bound was the recent bound $\mathbf{mix} = O(n^{2.5})$ due to the second author. We will refine the arguments in [P2] to show that

$$\mathbf{mix} = O(n^2)$$

D. Coppersmith: T.J. Watson Research Center, IBM, Yorktown Heights, NY 10598, USA.
e-mail: copper@watson.ibm.com

I. Pak: Department of Mathematics, Yale University, New Haven, CT 06520, USA
e-mail: paki@math.yale.edu

Mathematics Subject Classification (1991): 60J10

Key words and phrases: Random walks on groups – Strong stationary time – Separation distance – Upper triangular matrices – Asymmetric exclusion process

As a byproduct from the proof we obtain the tight bound on the eigenvalue

$$\lambda_2 = 1 - \frac{1+o(1)}{n}$$

The result is remarkable for several reasons. First, it was expected that the walk is slower than a similar random walk where random pairs (i, j) are chosen, $1 \le i < j \le n$, and we add to row i a row j times uniform $a \in \mathbb{F}_q$. For this random walk a sharp bound $\textbf{mix} = O(n^2 \log n)$ is known (see [AP, P1, P2]).

Second, it is easy to see that the obvious lower bound on **mix** is the diameter $\Delta = \binom{n}{2}$. Thus our walk is a rare example of the walk which mixes in $O(\Delta)$, but does not have constant eigenvalues gap, in contrast with expanders.

Our results are too weak to establish or reject the existence of the cutoff, also known as threshold (see [D1, D2]). It was long believed that when n grows there always exists a cutoff (see [D2, S1, P2]). Our results, however, seem to weaken this belief.

The idea of the proof is based on an explicit construction of the stopping rule with a uniform stationary distribution. The distribution of the stopping time of such a rule gives rise to a bound on the rate of convergence for the random walk. Then we employ a probabilistic argument to estimate the mean stopping time. Finally, we combine our results with known bounds on mixing time to prove the claim.

We finish by observing a remarkable similarity between our random walk and a one dimensional asymmetric exclusion process. Consider n particles on a circle with $2n$ states arranged clockwise. At each time randomly choose a state $i \in \{1, 2, \dots, 2n\}$. If state i is occupied and state $i+1$ is empty, move the particle from state i to state $i+1$. It is not hard to see that this defines a discrete Markov chain with uniform stationary distribution. Thus at infinity the probability to move the particle is about $1/4$. We show that already after $O(n^2)$ steps this probability becomes bounded away from 0.

1. Definitions and main results

Let $\Bbbk$ be any (finite or infinite) compact commutative ring, and let η be an invariant measure on $\Bbbk$. For example, $\Bbbk$ can be the ring of p-adic integers $\mathbb{Z}_p$, or a finite field $\mathbb{F}_q$. Denote by $\beta = \beta(\Bbbk)$ the measure of the noninvertible elements

$$\beta(\Bbbk) = \eta(\{a \in \Bbbk \mid \nexists\, b, a\, b = 1\})$$

For example, $\beta(\mathbb{F}_q) = 1/q$, and $\beta(\mathbb{Z}_p) = 1/p$.

Let $G = U(n, \Bbbk)$ be the group of upper triangular$(n \times n)$-matrices over $\Bbbk$ with ones on diagonal. Denote by μ the invariant measure on G, also known as Haar measure. It is known that μ is given as a product measure

$$\mu = \eta \times \eta \times \cdots \times \eta$$

where the product is taken over all $\binom{n}{2}$ entries above diagonal (see e.g. [H]).

Let $S \subset G$ be a set of matrices with ones on diagonal and zeroes elsewhere but some element right above diagonal. Clearly, S is a generating set of G. Now let Q^k

be a probability distribution of the product $M_1 \cdot \ldots \cdot M_k$ where M_i are independent and uniform in S. We think of Q^k as a probability distribution of the k-th step of a random walk $\mathcal{W} = \mathcal{W}(G, S)$ on G generated by S. By ergodicity, $Q^k(X) \to \mu(X)$ as $k \to \infty$ for all $X \subset G$.

Define the *separation distance* for a random walk $\mathcal{W}$

$$\mathbf{s}(k) = \sup_{A \subset G, \mu(A) > 0} \left(1 - \frac{Q^k(A)}{\mu(A)}\right)$$

One can think of $\mathbf{s}(k)$ as of a one-sided l_∞ distance (see [AD]). It is easy to see that $1 \geq \mathbf{s}(k) \geq 0$ for all $k > 0$. It is known also that the separation distance is nonincreasing: $\mathbf{s}(k+1) \leq \mathbf{s}(k)$, and submultiplicative: $\mathbf{s}(k+m) \leq \mathbf{s}(k) \cdot \mathbf{s}(m)$ (see e.g. [AD, AF, D1]).

It often convenient to consider the following definition of *mixing time*:

$$\mathbf{mix} = \min_k \left\{ \mathbf{s}(k) < \frac{1}{2} \right\}$$

By definition, after $k = \mathbf{mix}$ number of steps we have $Q^k(A) > \mu(A)/2$ for any $A \in G$. For other similar definitions of mixing times see [AF, LW].

The main result of the paper is the following theorem.

Theorem 1.1. *Let $G = U(n, \Bbbk)$, such that $\beta(\Bbbk) = 0$. Then for any $c > 0$ and integer $n \geq 2$ the separation distance $\mathbf{s}(m)$ for a random walk $\mathcal{W}(G, S)$ satisfies*

$$\mathbf{s}(m) < e^{-c/2},$$

where $m > 4(2\ln 2 + 1)\, n^2 + c\, n$. Moreover, the result holds when $\beta(\Bbbk) < C/m^{2+\epsilon}$ for any $\epsilon > 0$ and some universal constant $C = C(\epsilon)$.

The theorem implies that for large n the mixing time is of the order at most $4(2\ln 2 + 1)\, n^2 + O(n)$. In particular, when $m = 10\, n^2$, $n \geq 4$, we have $Q^m(A) \geq \frac{1}{2}\mu(A)$ for all $A \subset G$.

Also, Theorem 1.1 shows that the second largest eigenvalue $\lambda_2 < 1 - 1/2\, n + O(1/n^2)$. Indeed, fix n and let $c \to \infty$. We have $\mathbf{s}(m) < e^{-c/2} = e^{-(m - 4(2\ln 2+1)\, n^2)/2n}$. Since $\mathbf{s}(m) \sim Const \cdot \lambda_2^m$ we obtain $\lambda \leq e^{-1/2n} < 1 - 1/2\, n + 1/8n^2$.

Analogously, from the proof it follows that $\lambda < 1 - \frac{1-\epsilon}{n} + O(1/n^2)$ for all $\epsilon > 0$. Together with an easy lower bound $\lambda_2 > 1 - 1/n + O(1/n^2)$ this gives us $\lambda_2 = 1 - 1/n + o(1/n)$. This result is much sharper than what can be extracted from [S2] for the present walk.

The proof of Theorem 1.1 is based on the following key lemma. Consider the following *backgammon game* (for just one player). Define a *board* to be an interval $[1, n]$. Place n pieces on the first space. At each step pick a uniform integer i from $[1, n-1]$. If the space i is occupied *and* space $i+1$ is unoccupied, move the piece from i to $i+1$. If $i = 1$ we can move either of the pieces gathered on space 1. The game is over when all spaces are occupied, i.e. no moves are allowed. Denote by τ the stopping time of the game. Think of τ as of a random variable.

Lemma 1.2. *In conditions of Theorem* 1, *the separation distance* $\mathbf{s}(m)$ *satisfies*

$$\mathbf{s}(m) \le \mathbf{Pr}(\tau > m), \quad for\ all\ m > 0$$

Suppose now that $\beta(\Bbbk) > 0$. For example, let $\Bbbk = \mathbb{F}_q$ be a finite field with q elements. While Theorem 1.1 is not applicable in this case, the analog of the result holds for q large enough.

Theorem 1.3. *Let* $G = U(n, \mathbb{F}_q)$, *such that* $n \ge 4$, $q > 2\,n^2$. *Then the separation distance* $\mathbf{s}(m)$ *for a random walk* $\mathcal{W}(G, S)$ *satisfies:*

$$\mathbf{s}(m) \le \left(\frac{3}{4}\right)^c$$

where $m > c \cdot 20\,n^2$, $c > 1$.

Note here that $\beta(\mathbb{F}_q) = 1/q > 0$, which violates the condition $\beta(\Bbbk) = 0$ in Theorem 1.1. This seemingly small change becomes critical in the proof of the theorem. It is also the main reason why the bound in Theorem 1.1 is lower than that in Theorem 1.3.

2. Strong uniform times

Let G be a finite or compact group, and let μ be an invariant measure. Let $\mathcal{W}$ be a random walk on G with uniform stationary distribution. A *randomized stopping rule* is an algorithm which observes the walk and stops it depending on the state passed and, perhaps, additional randomness. Denote by ϱ and τ the *stopping state* and *stopping time* of the randomized stopping rule.

A a stopping time τ is called *strong uniform* if for all $A \subset G$ and $k > 0$ we have

$$\mathbf{Pr}(\varrho \in A\,|\tau = k) = \mu(A)$$

In other words, we need ϱ to be uniform in G and independent of τ. We will need the following classical result of Aldous and Diaconis (see [AD, D1]).

Theorem 2.1. *Let* τ *be a strong uniform time for a random walk* $\mathcal{W}$. *Then*

$$\mathbf{s}(k) \le \mathbf{Pr}(\tau > k), \quad k > 0$$

Now let $G = U(n, \Bbbk)$ and let $\mathcal{W}$ be a random walk defined above. We will present an explicit construction of a stopping rule which defines a strong uniform time τ.

For convenience, number rows of the upper triangular matrices upside-down. Namely, let the bottom row be 1, the next row be 2, ..., the top row be n. We can define the random walk $\mathcal{W}_2$ as follows: Choose a random integer i between 1 and $n - 1$, and add to the $(i + 1)$-th row the i-th row multiplied by a uniform number $a \in \Bbbk$. Recall the backgammon game on a board $[1, n]$ defined in section 1. Let us play, i.e. move pieces on the board, according to the same choices of integer i.

Note that while playing, we disregard the number a we used above. Now, define a stopping rule to stop the walk whenever the backgammon game is over. As before, let τ be the stopping time.

Lemma 2.2. *Let $\beta(\Bbbk) = 0$. Then stopping time τ defined above is strong uniform.*

Observe that Lemma 2.2 together with Theorem 2.1 implies Lemma 1.2. The proof of lemma is done by an elaborate induction.

Proof of Lemma 2.2. We claim that at any time t, given pieces are positioned in spaces $1 < i_1 < i_2 < \cdots$, then the corresponding rows $i_1, i_2, \ldots$ of the obtained upper triangular matrix are uniform and independent. Call rows $i_1, i_2, \ldots$, *marked rows*. Use induction. The claim is trivial when $t = 0$. Suppose true when $t = m$. Say at the next step we choose i. We can either move a game piece or not. If not, and we do not add anything to any of the marked rows ($i + 1 \neq i_l$ for any l), there is nothing to check. Suppose we add a row i to a marked row $i + 1$. The row $i + 1$ will remain uniform. Also, if both rows are marked, which means uniform and independent, then afterwards they are still uniform and independent. Indeed, call these rows X_1, X_2 and think of them as n-vectors over $\Bbbk$. Clearly, "(X_1, X_2) are uniform and independent" is *equivalent* to "$(X_1, X_2 + a\, X_1)$ are uniform and independent" for *any* $a \in \Bbbk$.

Now, suppose we add marked row i to an unmarked row $i + 1$. Row i was $(0, \ldots, 0, 0, 1, x_{n-i+2}, \ldots, x_n)$, with the x_j uniform and independt of other marked rows. Row $i+1$ was $(0, \ldots, 0, 1, y_{n-i+1}, y_{n_i+2}, \ldots, y_n)$, with the y_j arbitrary. After the addition, row $i+1$ is now $(0, \ldots, 0, 1, y_{n-i+1}+a, y_{n_i+2}+ax_{n-i+2}, \ldots, y_n+ ax_n)$. Regardless of the initial y_j of row $i+1$, the fact that a and x_j are uniform and independent of the other marked rows, implies that the new row $i+1$ is uniform and independent of the other marked rows. Indeed, this is clear for any invertible $a \in \Bbbk$, and the result follows from $\beta(\Bbbk) = 0$. Now note that row i is not *independent* on row $i + 1$. Therefore by removing row i from consideration, we confirm the step of induction and prove the claim.

Now consider what happens when the game is over. By definition, this is the stopping time τ, or, equivalently, the first time when all rows are marked. By the claim above, when the game is over we get all the rows uniform and independent on each other. Recall that the measure on $U(n, \Bbbk)$ is a product measure. We have $\mathbf{Pr}\big(\varrho = (X_1, \ldots, X_{n-1}) \in (A_1, \ldots, A_n) \,|\, \tau = k\big) = \mu_1(A_1) \times \cdots \times \mu_{n-1}(A_{n-1}) = \mu(A)$ where X_i, $i = 1, \ldots, n-1$, are rows of the resulting matrix $\varrho \in G$ and μ_i are the corresponding measures. This finishes the proof. □

3. Proof of Theorem 1.1

The proof is based on Lemma 1.2 and the analysis of the backgammon game described in section 1. We shall modify it first, solve it in this case, and then come back to the original problem.

Consider the following random process. Place k pieces on an integer line, in positions $1, 2, \ldots, k$. By (1) denote the rightmost piece, by (2) denote the next one, and so on. At each step, choose a random piece and move it to the right, if the next place is not occupied. The question is how far will (k) move after t steps?

Lemma 3.1. *Let $c > 0$. Then after t steps the piece (k) will move at least $t/(2ck)$ steps with probability*

$$> 1 - e^{(-1/4+(1/2+\ln 2)/c)t/k}$$

Let us start with some easy special cases. When $k = 1$ we have only one piece which moves freely to the right. When $k = 2$, the piece (1) clearly moves freely when chosen, while piece (2) is lagging behind. Since the distance between (1) and (2) behaves like a reflecting random walk on a line (see e.g. [F]), after t steps the piece (2) is at expected $t/2 - O(\sqrt{t})$ distance from 0. Of course, the complexity of the problem grows with k.

Proof of Lemma 3.1. Consider what happens after $t = 2\,c\,k^2$ steps. Divide time t into $2k$ intervals, each of length $c\,k$. Work backwards in time. Let t_{k-1} be the last time (k) is right behind $(k-1)$. Analogously, let t_{k-2} be the last time before t_{k-1} that $(k-1)$ is right behind $(k-2)$. In general, let t_i be the last time before t_{i+1} that $(i+1)$ is right behind (i), $1 \le i \le k-2$. Call these t_i *breaking points*.

For every interval I_j, $0 \le j \le 2k$, define $\psi(j)$ to be $j - \max\{i \mid t_i < ckj\}$. In other words, $\psi(i)$ is equal to j minus the number of breaking points that occur before the end of interval I_j. Clearly, $\psi(0) = 0$, $\psi(2k) = k+1$ and $\psi(j+1) \le \psi(j)+1$. Also, if $\psi(j+1) = \psi(j)+1$ then no breaking points occurred during I_j. Call such intervals *empty*. The are at least $k+1$ empty intervals. Now consider a sequence $j_1, j_2, \ldots, j_{k+1}$, such that $\psi(j_m) = m > \psi(j')$ for all $m = 1, \ldots, k+1$ and $j' < j_m$. Observe that intervals $I_{j_1}, I_{j_2}, \ldots, I_{j_{k+1}}$ are all different and empty. For a moment, fix these intervals.

Now, while running the random process follow (1) until time t_1, then follow (2) until $t_2, \ldots$, and finally follow (k) until t. Notice that in the interval I_{j_m} we follow piece $(j_m - m + 1)$ because $t_{j_m - m}$ must occur before the end of I_{j_m}, while $t_{j_m - m}$ must not. Observe that in the $k+1$ intervals we fixed, the piece we follow moves freely, i.e. always moves to the right when chosen.

Denote by r the position of (k) after t steps. By l_m denote the displacement of the piece we follow during the interval I_{j_m}, $1 \le m \le k+1$. Let $l = l_1 + \cdots + l_{k+1}$. Clearly, $r \ge l$ and $E(l) = \sum_m E(l_m) = (k+1)(c\,k)/k > c\,k$. Since l is the sum of $(k+1)(c\,k)$ independent Bernoulli trials, each with probability of success $1/k$, the Chernoff bound gives us:

$$\mathbf{Pr}\big(l < (1-\delta)E(l)\big) < e^{-\delta^2 c\,k/2}$$

for any $\delta > 0$. And therefore for any fixed intervals $i_1, \ldots, i_{k+1}$ we obtain

$$\mathbf{Pr}\big(r > c\,(1-\delta)k \mid i_1, \ldots, i_{k+1}\big) \ge \mathbf{Pr}(l > (1-\delta)E(l)) > 1 - e^{-\delta^2 c\,k/2}$$

On the other hand, the total number of ways to choose $k+1$ intervals out of $2\,k$ is given by $\binom{2k}{k+1} < 2^{2k}$. Summing over all the possibilities of choosing intervals, we have

$$\mathbf{Pr}(r > c\,(1-\delta)k) > 1 - 2^{2k} e^{-\delta^2 c\,k/2}$$

Now let $1-\delta = 1/c$. We obtain

$$\mathbf{Pr}(r > k) > 1 - e^{2k\ln 2 - (1-1/c)^2 c\,k/2} > 1 - e^{(2\ln 2 + 1 - c/2)\,k}$$

Recall now that $t = 2\,c\,k^2$. We have

$$\mathbf{Pr}(r > t/2\,c\,k) > 1 - e^{f(c)\,t/k},$$

where $f(c) = (2\ln 2 + 1 - c/2)/2\,c = -1/4 + (1/2 + \ln 2)/c = -1/4 + O(1/c)$. This finishes the proof. □

Proof of Theorem 1.1. In the original backgammon game we have only n pieces which are placed on the board one by one. Suppose instead they are all positioned to the left of 1. Also, instead of choosing a uniform space i, let us choose a uniform piece, and move it to the right if possible. This can only slow the game. Indeed, the piece that are still to the left of 1 correspond to the empty spaces in $[1, \ldots, n]$ and thus the pieces in this interval move with right probabilities. On the other hand, now it is not true that whenever space 1 is chosen, and 2 is unoccupied the piece always moves there. The reason is that the pieces to the left of 1 may lag in getting there.

Now, in the original backgammon game whenever piece (i) gets to the space $n - i + 1$, it stays there. In this new version we can as well forget about them and disregard their movements to the right of $n - i + 1$. All we need is to have piece (n) move to space 1, which in the notations above is equivalent to having $r > n$. Then we are done. By Lemma 2.1, if $k = n$ and $t = 2\,c\,n^2$, we obtain

$$\mathbf{Pr}(\tau > t) \le 1 - \mathbf{Pr}(r > n) < e^{(2\ln 2 + 1 - c/2)\,n}$$

By Lemma 1.2 we conclude $\mathbf{s}(t) \le \mathbf{Pr}(\tau > t) < e^{(2\ln 2+1-c/2)\,n}$. Take $c = 4\ln 2 + 2 + c'/n$. Then $t = 2\,c\,n^2 = (8\ln 2 + 4)n^2 + c'n$, and $\mathbf{s}(t) < e^{-c'/2}$. This finishes the proof of the first part of the theorem. We give the proof of the second part in the next section. □

4. Proof of Theorem 1.3

We shall deduce Theorem 1.3 from the same stopping time τ defined in section 2. Observe that τ can no longer be shown to be strong uniform. The proof breaks when we move a piece and claim that if i-th row is uniform, then the next row is uniform. This is no longer true since if we multiply by 0, then *we are not adding* while claim that we obtain uniformity. Of course, this event has probability $\beta(\mathbb{F}_q) = 1/q > 0$.

Still, consider the distribution of the stopping state ϱ we obtain. Since at each addition as above we can "mess up" at most $(1/q)$ portion of the row, after $\binom{n}{2} < n^2/2$ additions we obtain a distribution Q^τ which will be equal to $1/|G|$ on at least $f = (1-1/q)^{\binom{n}{2}}$ fraction of elements. Now, if $q > 2\,n^2$, we have $f > (1 - 1/2\,n^2)^{n^2/2} > 3/4$ for all $n \ge 2$. In other words, we just showed that there exist a subset $A \subset G$ such that $|A| > \frac{3}{4}|G|$ and $Q^\tau(a) = 1/|G|$ for all $a \in A$.

Consider a distribution $Q^{2\tau}$. Since every element can be decomposed as a product of two elements in A by at least $|G|/2$ ways, we have

$$Q^{2\tau}(g) \geq \frac{1}{2|G|}$$

for all $g \in G$. We claim that $\mathbf{s}(20\,n^2) \leq \frac{3}{4}$. Indeed, recall that for $n \geq 4$ we have $\mathbf{Pr}(\tau < 10\,n^2) > 1/2$. This follows from the remark after Theorem 1.1 and proof in section 4.

Therefore after $t = 20\,n^2$ the probability to get each element

$$Q^t(g) \geq \sum_{k=1}^{t/2} \mathbf{Pr}(\tau = k)\frac{1}{2|G|} = \frac{\mathbf{Pr}(\tau < t/2)}{2|G|} \geq \frac{1}{4|G|}$$

This proves the claim. Now by submultiplicativity we get the result. □

Similar computations give a proof of the second part of Theorem 1.1. Noninvertible elements in $\Bbbk$ will play a role the zero element in $\mathbb{F}_q$. We omit the obvious details. □

5. Applications: Asymmetric exclusion process on a circle

The asymmetric exclusion process on a circle is defined as follows. For every n, k, $1 \leq k < n$ define discrete Markov chain $\mathcal{M} = \mathcal{M}(n, k)$ as follows. Let the state space be a set of configurations of k particles on a circle with n spaces. Let the step of a chain to consist of choosing a uniform particle and moving it to the right (clockwise) if the next space is empty. This is a discrete time version of the more general continuous time exclusion process introduced by Spitzer in [Sp] (see also [Li]).

We claim that $\mathcal{M}(n, k)$ is an ergodic Markov chain. Indeed, observe that the number of ways to get into any configuration of particles is equal to the number of consecutive intervals of particles, and therefore is equal to the number of way to leave the configuration. Since the set of configurations is strongly connected under the moves, and each of the moves has probability $1/k$, this immediately implies ergodicity.

While from a different prospective, the process $\mathcal{M}$ is well studied in statistical physics literature (see [De] for the references). It has long been known that in the *steady state* the velocity of each particle becomes

$$v = \lim_{t\to\infty} \frac{E(Y_t)}{t} = \frac{(n-k)}{k(n-1)}$$

where Y_t is the number of moves of a fixed particle after t steps. Analogously, for the *diffusion* Δ it was recently obtained:

$$\Delta = \lim_{t\to\infty} \frac{E(Y_t^2) - (E(Y_t))^2}{t} = \frac{1}{2k^2(n-1)} \frac{\binom{2(n-1)}{2k-1}}{\binom{n-1}{k}}$$

(see [De, DEHP]).

The main problem is to find a mixing time of $\mathcal{M}(n, k)$. The most interesting case is when the ratio of particles is constant: $k/n \to \rho$, $0 < \rho \leq 1/2$. We will concentrate on that case.

We propose the following conjecture.

Conjecture 5.1. *There exist a universal constant C such that for all $\rho > 0$ we have*

$$\textbf{mix} < C\, n^2$$

where **mix** *is the mixing time of the Markov chain $\mathcal{M}(n, \lfloor \rho n \rfloor)$.*

While we are unable to prove this conjecture let us give partial results in its support. First, one can consider "reversibilization" of $\mathcal{M}$. This is often done by considering a Markov chain with transition matrix $\frac{1}{2}(P + P^T)$ where P is the original transition matrix. In the case of $\mathcal{M}$ this leads to the well understood exclusion process on a circle $\widetilde{\mathcal{M}}(n, k)$, which is defined as follows. Let the state space be as before, but now we move either to right or to the left with equal probability. It is known then (see [DSC1, DSC3]) that this random walk mixes after $O(n^3 \log n)$ steps (cf. [LeY, LuY]). Without going into technical details, let us note that one can obtain this bound in the asymmetric case as well:

Proposition 5.2. (see [DSC3, p. 739]). *There exist a universal constant C such that for all $\rho > 0$ we have*

$$\textbf{mix} < C\, n^3 \log n$$

where **mix** *is the mixing time of the Markov chain $\mathcal{M}(n, \lfloor \rho n \rfloor)$.*

Second, one can introduce a partial measure of how well the particles are mixed. Call a *success* an event of actually moving a particle when it is chosen. At a steady state, the probability of success is about ρ. A variation of the argument given in the proof of Lemma 3.1 gives the probability of success $\rho(1 - \epsilon)$ after $5\, n^2 + O(n)$ steps. It is easy to see that result is sharp up to a constant. Indeed, if we start with configuration when all particles are in one cluster, it would take $C \cdot n^2$ time just to *move* either of the last $k/2$ particles. Formally, we obtain the following result.

Proposition 5.3. *Let $p(t)$ be the probability of success at time t. Then there exist is a universal constant $C_1 > 0$ such that for any n and any starting configuration*

$$p(t) > \rho(1 - \epsilon)$$

where $t > 10\rho\, n^2 + C_1 n \log(1/\epsilon)$. In the other direction, for any n, $\lambda > 0$ we have

$$p(t) < \frac{C_2}{\lambda^2},$$

where C_2 is a universal constant, and $t < 1/2\rho n^2 - \lambda n^{3/2}$.

Note that Conjecture 5.1 implies Proposition 5.3. Let us also remark that it takes about $O(n^3)$ steps for a symmetric exclusion process on a circle to have a constant probability of success.

Finally, there is additional evidence in favor of the conjecture in the statistical physics literature. In [GS] authors use Bethe ansatz to study eigenvalues of the transition matrix of the process. They observe that $Re(\lambda_2) < 1 - C\, n^{-3/2}$ for some universal constant $C > 0$. While the proof is not rigorous from the mathematical point of view, the results in the field are usually correct. Formally, the authors compute a number of eigenvalues, largest of which is as above, without proving the *completeness*, which is that these are all the eigenvalues ([Spo]). If completeness is assumed, one can make an argument that there exists a universal constant C such that for all $\rho > 0$ we have

$$\mathbf{mix} < C\, n^{2.5},$$

where **mix** is the mixing time of the Markov chain $\mathscr{M}(n, \lfloor \rho n \rfloor)$.

One should note that authors in [GS] claim that the third largest eigenvalue satisfies $Re(\lambda_3) = 1 - C'/n$. It is not hard to see that if this claim is true, and assuming the multiplicity of the λ_2 is $2^{O(\sqrt{n})}$, this would prove the conjecture. We hope to return to this problem in the future.

Acknowledgements. The authors are grateful to Persi Diaconis, Jim Fill, László Lovász, Alexander Lubotzky, Gregory Margulis, and Gregory Sorkin for helpful remarks. Special thanks to Peter Winkler for pointing out connection with interactive particle systems, and to Bernard Derrida, Joel Lebowitz, Nikolaus Rajewsky, Gene Speer and Herbert Spohn for their invaluable help in understanding results in statistical physics. The second author was partially supported by the NSF Mathematical Sciences Postdoctoral Research Fellowship.

References

[A1] Aldous, D.: Random walks on finite groups and rapidly mixing Markov chains, Lecture Notes in Math. **986** (1983)

[AD] Aldous, D., Diaconis, P.: Strong uniform times and finite random walks, Adv Appl. Math. **8**, 69–97 (1987)

[AF] Aldous, D., Fill, J.: Reversible Markov Chains and Random Walks on Graphs, monograph in preparation, 1996

[AP] Astashkevich, A., Pak, I.: Random walks on nilpotent and supersolvable groups, preprint, 1997

[De] Derrida, B.: An exactly soluble non-equilibrium system: the asymmetric sample exclusion process. Fundamental problems in statistical mechanics (Altenberg, 1997), Phys. Rev. **301**, 65–83 (1998)

[DEHP] Derrida, B., Evans, M.R., Hakim, V., Pasquier, V.: Exact solution of a 1d asymmetric exclusion model using a matrix formulation, J. Phys. A **26**, 1493–1517 (1993)

[D1] Diaconis, P.: Group Representations in Probability and Statistics, IMS, Hayward, California, 1988

[D2] Diaconis, P.: The cutoff phenomenon in finite Markov chains, Proc. Nat. Acad. Sci. U.S.A. **93**, 1659–1664 (1996)

[DF] Diaconis, P., Fill, J.A.: Strong stationary times via new form of duality, Ann. Prob. **18**, 1483–1522 (1990)

[DSC1] Diaconis, P., Saloff–Coste, L.: Comparison theorems for reversible Markov chains, Ann. Appl. Prob. **3**, 696–730 (1993)

[DSC2] Diaconis, P., Saloff–Coste, L.: Moderate growth and random walk on finite groups, Geom. Funct. Anal. **4**, 1–36 (1994)

[DSC3] Diaconis, P., Saloff–Coste, L.: Logarithmic Sobolev inequalities for finite Markov chains, Ann. Appl. Prob. **6**, 695–750 (1996)

[F] Feller, W.: An introduction to Probability theory and its applications (third edition), John Wiley, New York, 1970

[GS] Gwa, L.H., Spohn, H.: Bethe Solution for the Dynamical Scaling Exponent of the Noisy Burgers Equation, Phys. Rev. A **46**, 844–854 (1994)

[H] Humphreys, J.: Linear algebraic groups, Springer, Berlin, 1975

[LeY] Lee, T.-Y., Yau, H.-T.: Logarithmic Sobolev inequality for some models of random walks, Ann. Prob. **26**, 1855–1873 (1998)

[Li] Liggett, T.M.: Interacting Particle Systems, Springer, New York, 1985

[LW] Lovász, L., Winkler, P.: Mixing Times Microsurveys in Discrete Probability (ed. D. Aldous and J. Propp), DIMACS Series, AMS, 1998

[LuY] Lu, S.L., Yau, H.-T.: Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics, Comm. Math. Phys. **156**, 399–433 (1993)

[P1] Pak, I.: Random walks on groups: strong uniform time approach, Ph.D. Thesis, Harvard U. 1997

[P2] Pak, I.: Two random walks on upper triangular matrices, preprint, 1998

[PV] Pak, I.; Vu, V.H.: On finite geometric random walks, preprint, 1998

[Sp] Spitzer, F.: Interaction of Markov processes, Adv. Math. **5**, 246–290 (1970)

[Spo] Spohn, H.: personal communication

[S1] Stong, R.: Random walk on the upper triangular matrices, Ann. Prob. **23**, 1939–1949 (1995)

[S2] Stong, R.: Eigenvalues of random walks on groups, Ann. Prob. **23**, 1961–1981 (1995)