The Annals of Probability
2009, Vol. 37, No. 2, 453–477
DOI: 10.1214/08-AOP409

IMPROVED MIXING TIME BOUNDS FOR THE THORP SHUFFLE AND *L*-REVERSAL CHAIN[1]

BY BEN MORRIS

University of California, Davis

We prove a theorem that reduces bounding the mixing time of a card shuffle to verifying a condition that involves only pairs of cards, then we use it to obtain improved bounds for two previously studied models.

E. Thorp introduced the following card shuffling model in 1973: Suppose the number of cards n is even. Cut the deck into two equal piles. Drop the first card from the left pile or from the right pile according to the outcome of a fair coin flip. Then drop from the other pile. Continue this way until both piles are empty. We obtain a mixing time bound of $O(\log^4 n)$. Previously, the best known bound was $O(\log^{29} n)$ and previous proofs were only valid for n a power of 2.

We also analyze the following model, called the *L-reversal chain*, introduced by Durrett: There are n cards arrayed in a circle. Each step, an interval of cards of length at most L is chosen uniformly at random and its order is reversed. Durrett has conjectured that the mixing time is $O(\max(n, \frac{n^3}{L^3}) \log n)$. We obtain a bound that is within a factor $O(\log^2 n)$ of this, the first bound within a poly log factor of the conjecture.

1. Introduction. Card shuffling has a rich history in mathematics, dating back to the work of Markov [11] and Poincaré [15]. A basic problem is to determine the mixing time, that is, the number of shuffles necessary to mix up the deck (see Section 1.1 for a precise definition). A natural first step (used as far back as Borel and Cheron [2] in 1940) is to determinine the number of steps necessary to randomize single cards and pairs. Clearly this is always a lower bound for the mixing time. On the other hand, it is often not far from an upper bound as well; for a number of models of card shuffling (see, e.g., Diaconis and Shahshahani [7], Wilson [16] or Bayer and Diaconis [1]) the the mixing time is only a small factor [e.g., $O(1)$ or $O(\log n)$] larger than the time required to mix pairs. This suggests finding a general method that reduces bounding the mixing time (in the global sense that the distribution on all $n!$ permutations is roughly uniform) to verifying a local condition that involves only pairs of cards. In this paper, we introduce such a method and use it to analyze two previously studied models. In both cases we find an upper bound for the mixing time, that is, within a poly-logarithmic factor of optimal.

Received January 2008; revised April 2008.
[1]Supported in part by Sloan Fellowship and NSF Grant DMS-07-07144.
AMS 2000 subject classification. 60J10.
Key words and phrases. Markov chain, mixing time.

We study card shuffles that can be viewed as generalizations of three card Monte. In three card Monte, the cards are spread out face down on a table. In one step, the dealer chooses two cards, puts them together and then separates them quickly so that an observer cannot tell which is which. We call this operation a *collision*, and model it mathematically as a random permutation, that is, an even mixture of a transposition and the identity. We prove a general theorem that applies to any method of shuffling that uses collisions. The theorem bounds the change in relative entropy after many steps of the chain, based on something, that is, related to the interactions between pairs of cards. Next we use the theorem to analyze two card shuffling models, the Thorp shuffle and Durrett's L-reversal model.

1.1. *Applications.* In this section we describe two applications of our main theorem. First, we give a formal definition of the mixing time. Let $p(x, y)$ be transition probabilities for a Markov chain on a finite state space V with a uniform stationary distribution. For probability measures μ and ν on V, define the total variation distance $\|\mu - \nu\| = \sum_{x \in V} |\mu(x) - \nu(x)|$, and define the mixing time

$$T_{\text{mix}} = \min\{n : \|p^n(x, \cdot) - \mathcal{U}\| \leq \tfrac{1}{4} \text{ for all } x \in V\}, \tag{1}$$

where $\mathcal{U}$ denotes the uniform distribution.

Our first application is the Thorp shuffle, which is defined as follows: Assume that the number of cards, n, is even. Cut the deck into two equal piles. Drop the first card from the left pile or the right pile according to the outcome of a fair coin flip; then drop from the other pile. Continue this way, with independent coin flips deciding whether to drop LEFT-RIGHT or RIGHT-LEFT each time, until both piles are empty.

The Thorp shuffle, despite its simple description, has been hard to analyze. Determining its mixing time has been called the "longest-standing open card shuffling problem" [5]. In [14] the author obtained the first poly-log upper bound, proving a bound of $O(\log^{44} n)$ valid when n is a power of 2. Montenegro and Tetali [13] built on this to get a bound of $O(\log^{29} n)$. In the present paper, we dispense with the power-of-two assumption and get an improved bound of $O(\log^4 n)$.

We also analyze a Markov chain that was introduced by Durrett [9] as a model for evolution of a genome (see [10]). In the *L-reversal chain* there are two parameters, n and L. The cards are located at the vertices of an n-cycle, which we label $0, \ldots, n-1$. Each step, a (nonempty) interval of cards of length at most L is chosen uniformly at random and its order is reversed. By the coupon collector problem, $O(n \log n)$ steps are needed to break adjacencies between neighboring pairs. Furthermore, the mixing time for a single card is on the order $\frac{n^3}{L^3}$, because for each step the probability that a particular card moves is on the order of L/n and each time a card moves it performs a step of a symmetric random walk with typical displacement on the order L. These considerations led Durrett to the following conjecture.

CONJECTURE (Durrett). The mixing time for the L-reversal chain is $O(\max(n, \frac{n^3}{L^3}) \log n)$.

In [9], Durrett proves the corresponding lower bound using Wilson's technique [16] based on eigenfunctions. The spectral gap was determined to be within constant factors of $\max(n, \frac{n^3}{L^3})$ by Cancrini, Caputo and Martinelli [3]. The best previously-known bound for the mixing time, which could be obtained by applying standard comparison techniques, was within a factor $O(n^{2/3})$ of the Durrett's conjecture in the worst case.

Durrett's conjecture has presented a challenge to existing techniques. As shown by Martinelli et al., the log Sobolev constant does not give the conjectured mixing time. Furthermore, the mixing time in L^2 [defined by replacing total variation distance by an appropriate L^2 distance in (1)] can be nearly $n^{1/3}$ times the conjecture, as the following example shows: let $L = n^{2/3}$, so that the conjectured mixing time is $O(n \log n)$. We claim that in this case the L^2 mixing time is at least $cn^{4/3}$ for a constant c. Let A be the event that cards $1, \ldots, n/2$ occupy positions $1, \ldots, n/2$ in any order. If the initial ordering is the identity permutation, then after t shuffles we have

$$\begin{aligned}\mathbf{P}(A) &\geq \mathbf{P}(\text{none of the reversed intervals contained cards 1 or } n/2)\\ &\geq \left(1 - \frac{2L}{n}\right)^t,\end{aligned}$$

which is much larger than $\binom{n}{n/2}^{-1}$ unless $t \geq cn^{4/3}$ for a constant c. Since mixing in L^2 implies convergence of transition probabilities, the L^2 mixing time is at least on the order of $n^{4/3}$, which is higher than the conjecture. This means that in order to prove the conjectured bound on the mixing time in total variation, one cannot use any method for bounding mixing times that gives a bound in L^2.

In the present paper, we prove that the mixing time is $O((n \vee \frac{n^3}{L^3}) \log^3 n)$. This is the first upper bound, that is, within a poly-log factor of the conjecture.

The remainder of this paper is organized as follows: in Section 2 we give some necessary background on entropy and prove some elementary inequalities. In Section 3 we define *Monte shuffles*, the general model of card shuffling to which our main theorem will apply. In Section 4 we prove the main theorem. In Section 5 we analyze the Thorp shuffle and in Section 6 we analyze the L-reversal chain.

2. Background. For a probability distribution $\{p_i : i \in V\}$, define the (relative) entropy of p by $\mathrm{ENT}(p) = \sum_{i \in V} p_i \log(|V| p_i)$, where we define $0 \log 0 = 0$. The following well-known inequality links relative entropy to total variation distance. Let $\mathcal{U}$ denote the uniform distribution over V. Then

$$\|p - \mathcal{U}\| \leq \sqrt{\tfrac{1}{2} \mathrm{ENT}(p)}. \tag{2}$$

If X is a random variable (or random permutation) taking finitely many values, define $\mathrm{ENT}(X)$ as the relative entropy of the distribution of X. Note that if $\mathbf{P}(X = i) = p_i$ for $i \in V$ then $\mathrm{ENT}(X) = \mathbf{E}(\log(|V|p_X))$. We shall think of the distribution of a random permutation in $\mathcal{S}_n$ as a sequence of probabilities of length $n!$, indexed by permutations in $\mathcal{S}_n$. If $\mathcal{F}$ is a sigma-field, then we shall write $\mathrm{ENT}(X \mid \mathcal{F})$ for the relative entropy of the conditional distribution of X given $\mathcal{F}$. Note that $\mathrm{ENT}(X \mid \mathcal{F})$ is a random variable. If π is a random permutation in S_n, then for $1 \le k \le n$, define $\mathcal{F}_k = \sigma(\pi^{-1}(k), \ldots, \pi^{-1}(n))$, and define $\mathrm{ENT}(\pi, k) = \mathrm{ENT}(\pi^{-1}(k) \mid \mathcal{F}_{k+1})$ [where we think of the conditional distribution of $\pi^{-1}(k)$ given $\mathcal{F}_{k+1}$ as being a sequence of length k]. The standard entropy chain rule (see, e.g., [4]) gives the following proposition.

PROPOSITION 1. *For any $i \le n$ we have*

$$\mathrm{ENT}(\pi) = \mathbf{E}(\mathrm{ENT}(\pi \mid \mathcal{F}_i)) + \sum_{k=i}^{n} \mathbf{E}(\mathrm{ENT}(\pi, k)).$$

To compute the relative entropy in the first term on the right-hand side, we think of the distribution of π given $\mathcal{F}_i$ as a sequence of probabilities of length $(i-1)!$.

REMARK. Substituting $i = 1$ into the formula gives $\mathrm{ENT}(\pi) = \sum_{k=1}^{n} \mathbf{E}(\mathrm{ENT}(\pi, k))$.

If we think of π as representing the order of a deck of cards, with $\pi(i) =$ location of card i, then this allows us to think of $\mathbf{E}(\mathrm{ENT}(\pi, k))$ as the portion of the overall entropy $\mathrm{ENT}(\pi)$, that is, attributable to the location k. If $S \subset \{1, \ldots, n\}$ is a set of positions, then we shall refer to the quantity $\sum_{k \in S} \mathrm{ENT}(\pi, k)$ as the entropy, that is, *attributable to* S.

DEFINITION 2. For $p, q \ge 0$, define $d(p, q) = \frac{1}{2} p \log p + \frac{1}{2} q \log q - \frac{p+q}{2} \times \log(\frac{p+q}{2})$.

We will need the following proposition.

PROPOSITION 3. *Fix $p \ge 0$. The function $d(p, \cdot)$ is convex.*

PROOF. A calculation shows that the second derivative is positive. □

Observe that $d(p, q) \ge 0$, with equality iff $p = q$ by the strict convexity of the function $x \to x \log x$. Furthermore, some calculations give

$$d(p, q) = \frac{p+q}{2} f\left(\frac{p-q}{p+q}\right), \tag{3}$$

where $f(\Delta) = \frac{1}{2}(1+\Delta)\log(1+\Delta) + \frac{1}{2}(1-\Delta)\log(1-\Delta)$. If $p = \{p_i : i \in V\}$ and $q = \{q_i : i \in V\}$ are both probability distributions on V, then we can define the "distance" $d(p,q)$ between p and q, by $d(p,q) = \sum_{i \in V} d(p_i, q_i)$. [We use the term *distance* loosely and do not claim that $d(\cdot,\cdot)$ satisfies the triangle inequality.] Note that $d(p,q)$ is the difference between the average of the entropies of p and q and the entropy of the average (i.e., an even mixture) of p and q.

We will use the following projection lemma.

LEMMA 4. *Let X and Y be random variables with distributions p and q, respectively. Fix a function g and let P and Q be the distributions of $g(X)$ and $g(Y)$, respectively. Then $d(p,q) \geq d(P,Q)$.*

PROOF. Let $S_i = \{x : g(x) = i\}$. Then

$$P_i = \sum_{x \in S_i} p_x; \qquad Q_i = \sum_{x \in S_i} q_x.$$

We have

$$d(p,q) = \sum_i \sum_{x \in S_i} d(p_x, q_x) \tag{4}$$

$$= \sum_i \sum_{x \in S_i} \frac{p_x + q_x}{2} f\left(\frac{p_x - q_x}{p_x + q_x}\right) \tag{5}$$

$$= \sum_i \left[\frac{P_i + Q_i}{2}\right] \sum_{x \in S_i} \frac{p_x + q_x}{2} \left[\frac{P_i + Q_i}{2}\right]^{-1} f\left(\frac{p_x - q_x}{p_x + q_x}\right). \tag{6}$$

Note that f has a positive second derivative, hence it is convex. Thus by Jensen's inequality, the quantity (6) is at least

$$\sum_i \left[\frac{P_i + Q_i}{2}\right] f\left(\sum_{x \in S_i} \frac{p_x + q_x}{2} \left[\frac{P_i + Q_i}{2}\right]^{-1} \frac{p_x - q_x}{p_x + q_x}\right) \tag{7}$$

$$= \sum_i \left[\frac{P_i + Q_i}{2}\right] f\left(\frac{P_i - Q_i}{P_i + Q_i}\right)$$

$$= \sum_i d(P_i, Q_i) \tag{8}$$

$$= d(P,Q). \tag{9}$$

□

Let $\mathcal{U}$ denote the uniform distribution on V. Note that if μ is an arbitrary distribution on V, then $\mathrm{ENT}(\mu)$ and $d(\mu, \mathcal{U})$ are both notions of a distance from μ to $\mathcal{U}$. The following lemma relates the two.

LEMMA 5. *For any distribution μ on V we have*

$$d(\mu, \mathcal{U}) \geq \frac{c}{\log |V|} \mathrm{ENT}(\mu)$$

for a universal constant $c > 0$.

PROOF. Let $n = |V|$, define $\widehat{\mu} = n\mu$ and define $g : (0, \infty) \to \mathbf{R}$ by $g(x) = x \log x - (x - 1)$. Then

$$\mathrm{ENT}(\mu) = \sum_{i \in V} \mu(i) \log(n\mu(i)) \tag{10}$$

$$= \frac{1}{n} \sum_{i \in V} \widehat{\mu}(i) \log \widehat{\mu}(i) - \big(\widehat{\mu}(i) - 1\big) \tag{11}$$

$$= \frac{1}{n} \sum_{i \in V} g(\widehat{\mu}(i)), \tag{12}$$

where the second equality holds because $\sum_{i \in V} (\widehat{\mu}(i) - 1) = 0$. Thus it is enough to show for a universal constant c we have

$$d\left(\mu(i), \frac{1}{n}\right) \geq \frac{c}{n \log n} g(\widehat{\mu}(i)) \tag{13}$$

for all $i \in V$. Fix $i \in V$ and let $x = \widehat{\mu}(i)$. Then by (3) we have

$$d\left(\mu(i), \frac{1}{n}\right) = \frac{1}{n} d(x, 1) \tag{14}$$

$$= \frac{1}{n} \left(\frac{x+1}{2}\right) f\left(\frac{x-1}{x+1}\right), \tag{15}$$

where $f(\Delta) = \frac{1}{2}(1 + \Delta) \log(1 + \Delta) + \frac{1}{2}(1 - \Delta) \log(1 - \Delta)$. Thus it remains to show that the function $R(x)$ defined by

$$R(x) = \frac{g(x)}{((x+1)/2) f((x-1)/(x+1))} \tag{16}$$

is at most $c^{-1} \log n$ on the interval $[0, n]$, for a constant $c > 0$. Note that $R(x)$ is bounded on the interval $[0, 2]$. (This can be seen by applying l'Hôpital's rule twice for the point $x = 1$.) Let $x \in [2, n]$. The denominator in (16) is at least

$$\frac{x}{2} f\left(\frac{x-1}{x+1}\right) \geq \frac{x}{2} f\left(\frac{1}{3}\right),$$

since the function $x \to f(x - 1/x + 1)$ is increasing on $[2, \infty)$. The numerator is $g(x) \leq x \log x \leq x \log n$. Thus $R(x) \leq 2 \log n / f(\frac{1}{3})$ on the interval $[2, n]$ and the proof is complete. □

3. General set-up: card shuffles with collisions.

3.1. *Collisions.* We shall now define a *collision*, which is the basic ingredient in all of the card shuffles analyzed in the present paper. If π is a random permuta-

tion in S_n such that

$$\pi = \begin{cases} \text{id}, & \text{with probability } \frac{1}{2}, \\ (a,b), & \text{with probability } \frac{1}{2}, \end{cases}$$

for some $a, b \in \{1, 2, \ldots, n\}$ [where we write id for the identity permutation and (a, b) for the transposition of a and b], then we will call π a *collision*. If π and μ are permutations in S_n, then we write $\pi\mu$ for the composition $\mu \circ \pi$.

A card shuffle can be described as a random permutation chosen from a certain probability distribution. If we start with the identity permutation and each shuffle has the distribution of π, then after t steps the cards are distributed like $\pi_1 \cdots \pi_t$, where the π_i are i.i.d. copies of π. In this paper, we shall consider shuffling permutations π that are written in the form

$$\pi = \nu c(a_1, b_1) c(a_2, b_2) \cdots c(a_k, b_k), \tag{17}$$

where ν is an arbitrary random permutation, the numbers $a_1, \ldots, a_k, b_1, \ldots, b_k$ are distinct, and $c(a_j, b_j)$ is a collision of a_j and b_j. The values of a_j and b_j and the number of collisions (which can be zero) may depend on ν, but conditional on ν the $c(a_j, b_j)$ are independent collisions. We shall call shuffles of this type *Monte*. For $t \geq 1$, define $\pi_{(t)} = \pi_1 \cdots \pi_t$.

REMARK. Of course any shuffle can be written in Monte form in a trivial way (i.e., with no collisions). The trick is to present any given shuffle in Monte form in a useful way.

3.2. *Warm-up lemma*. In this section we prove a simple lemma with a short proof that brings out many of the central ideas of our main theorem (Theorem 9 below). We start with an easy proposition.

PROPOSITION 6. *Suppose that π is any fixed permutation. Then*

$$\text{ENT}(\mu\pi) = \text{ENT}(\mu).$$

PROOF. Up to a re-labeling of indices, the random permutation $\mu\pi$ has the same distribution as μ, hence the same relative entropy. □

If π is random and independent of μ then $\text{ENT}(\mu\pi) \leq \text{ENT}(\mu)$, which follows by using Jensen's inequality (applied to the function $x \to x \log x$) to condition on π, and then applying Proposition 6. It follows that if $\pi_1, \pi_2, \ldots$ are i.i.d. copies of π then $\text{ENT}(\pi_1 \cdots \pi_k)$ is nonincreasing in k. In this section we study the decay of entropy $\text{ENT}(\mu\pi) - \text{ENT}(\mu)$ in the case where the permutation π is a collision.

The following lemma relates to the case where π is a collision between the jth card and another card of smaller index. The lemma says that the relative entropy is reduced by at least $c\,\text{ENT}(\mu, j)/\log n$, on average (where "on average" means with respect to the different possible choices of indices $i \leq j$).

LEMMA 7. *Let μ be a random permutation. Then for a universal constant c we have*

$$j^{-1}\sum_{i\le j}\mathrm{ENT}(\mu c(i,j)) \le \mathrm{ENT}(\mu) - c\,\mathrm{ENT}(\mu, j)/\log n.$$

PROOF. Using the abuse of notation $\frac{1}{2}\pi_1 + \frac{1}{2}\pi_2$ for a random permutation whose distribution is an even mixture of the distributions of π_1 and π_2, we have

$$\mu c(i,j) = \tfrac{1}{2}\mu + \tfrac{1}{2}\mu(i,j).$$

Let $\mathcal{L}(X \mid \mathcal{F})$ denote the conditional distribution of random variable (or random permutation) X given the sigma field $\mathcal{F}$. Let $\widehat{\mu} = \mu(i,j)$ [i.e., the product of μ and the transposition (i,j)]. Note that $\widehat{\mu}$ and μ are the same, except that $\widehat{\mu}^{-1}(i) = \mu^{-1}(j)$ and $\mu^{-1}(i) = \widehat{\mu}^{-1}(j)$ and recall that $i \le j$. It follows that $\mathrm{ENT}(\widehat{\mu} \mid \mathcal{F}_{j+1}) = \mathrm{ENT}(\mu \mid \mathcal{F}_{j+1})$ and hence $\mathrm{ENT}(\mu c(i,j) \mid \mathcal{F}_{j+1}) - \mathrm{ENT}(\mu \mid \mathcal{F}_{j+1}) = -d(\mathcal{L}(\widehat{\mu} \mid \mathcal{F}_{j+1}), \mathcal{L}(\mu \mid \mathcal{F}_{j+1}))$. But by the projection lemma,

$$\begin{aligned} d(\mathcal{L}(\widehat{\mu} \mid \mathcal{F}_{j+1}), \mathcal{L}(\mu \mid \mathcal{F}_{j+1})) &\ge d(\mathcal{L}(\widehat{\mu}^{-1}(j) \mid \mathcal{F}_{j+1}), \mathcal{L}(\mu^{-1}(j) \mid \mathcal{F}_{j+1})) \\ &= d(\mathcal{L}(\mu^{-1}(i)|\mathcal{F}_{j+1}), \mathcal{L}(\mu^{-1}(j) \mid \mathcal{F}_{j+1})). \end{aligned}$$

Hence

$$\begin{aligned} &j^{-1}\sum_{i\le j}\mathrm{ENT}(\mu c(i,j) \mid \mathcal{F}_{j+1}) - \mathrm{ENT}(\mu \mid \mathcal{F}_{j+1}) \\ &\quad\le -j^{-1}\sum_{i\le j} d(\mathcal{L}(\mu^{-1}(i) \mid \mathcal{F}_{j+1}), \mathcal{L}(\mu^{-1}(j) \mid \mathcal{F}_{j+1})) \\ &\quad\le -d\left(j^{-1}\sum_{i\le j}\mathcal{L}(\mu^{-1}(i) \mid \mathcal{F}_{j+1}), \mathcal{L}(\mu^{-1}(j) \mid \mathcal{F}_{j+1})\right) \qquad (18) \\ &\quad= -d(\mathcal{U}, \mathcal{L}(\mu^{-1}(j) \mid \mathcal{F}_{j+1})) \\ &\quad\le -\frac{c}{\log n}\,\mathrm{ENT}(\mathcal{L}(\mu^{-1}(j) \mid \mathcal{F}_{j+1})), \end{aligned}$$

where the first inequality is by Proposition 3 and the second is by Lemma 5. Here $\mathcal{U}$ denotes the uniform distribution over $\{1,\dots,n\} - \{\mu^{-1}(j+1),\dots,\mu^{-1}(n)\}$. Taking expectations gives

$$\begin{aligned} &j^{-1}\sum_{i\le j}\mathbf{E}(\mathrm{ENT}(\mu c(i,j) \mid \mathcal{F}_{j+1})) - \mathbf{E}(\mathrm{ENT}(\mu \mid \mathcal{F}_{j+1})) \\ &\quad\le -\frac{c}{\log n}\,\mathrm{ENT}(\mu, j). \qquad (19) \end{aligned}$$

Since $\mathrm{ENT}(\mu, k) = \mathrm{ENT}(\mu c(i,j), k)$ for all $k \ge j+1$, Proposition 1 and (19) yield the lemma. □

4. Main theorem. Let π be a random permutation in $\mathcal{S}_n$, that is, Monte [i.e., can be written in the form (17)] and let $\pi_1, \pi_2, \ldots$ be independent copies of π. For $t \geq 1$ let $\pi_{(t)} = \pi_1 \cdots \pi_t$.

CONVENTION. We shall use the following convention throughout: we denote by *card* x the card initially in position x.

For cards x and y, say that x *collides with* y *at time* m if for some i and j we have $\pi_{(m)}^{-1}(i) = x$, $\pi_{(m)}^{-1}(j) = y$, and π_m has a collision of i and j.

We will need the following definition.

DEFINITION 8. For a random variable X, a finite set S and a real number $A \in [0, 1]$, say that the distribution of X is A-uniform over S if

$$\mathbf{P}(X = i) \geq A|S|^{-1}$$

for all $i \in S$.

REMARK. If $A < 1$ then the distribution of X need not be concentrated on S. (But if $A = 1$, then X is uniform over S.)

Our main theorem is a generalization of Lemma 7. It generalizes from a collision to an arbitrary Monte shuffle, and it bounds the loss in relative entropy after many steps.

THEOREM 9. *Let π be a Monte shuffle on n cards. Fix an integer $t > 0$ and suppose that T is a random variable taking values in $\{1, \ldots, t\}$, which is independent of the shuffles $\{\pi_i : i \geq 0\}$. For a card x, let $b(x)$ denote the first card to collide with x after time T [or $b(x) = x$ if there is no such card]. Define the match $m(x)$ of x by*

$$m(x) := \begin{cases} b(x), & \text{if } x = b(b(x)), \\ x, & \text{otherwise.} \end{cases}$$

Suppose that for every card i there is a constant $A_i \in [0, 1]$ such that the distribution of $m(i)$ is A_i-uniform over $\{1, \ldots, i\}$. Let μ be an arbitrary random permutation, that is, independent of $\{\pi_i : i \geq 0\}$. Then

$$\mathrm{ENT}(\mu\pi_{(t)}) - \mathrm{ENT}(\mu) \leq \frac{-C}{\log n} \sum_{k=1}^{n} A_k E_k,$$

where $E_k = \mathbf{E}(\mathrm{ENT}(\mu, k))$ and C is a universal constant.

PROOF. Let $\mathcal{M} = (m(i) : 1 \le i \le n)$. For i and j with $j \le i$, let $c(i, j)$ be a collision of i and j. Assume that all of the $c(i, j)$ are independent of μ, $\pi_{(t)}$ and each other. Note that

$$\left[\prod_{i\,:\,m(i)\le i} c(i, m(i))\right]\pi_{(t)}$$

has the same distribution as $\pi_{(t)}$, so it is enough to bound the relative entropy of the distribution of $\mu[\prod_{i\,:\,m(i)\le i} c(i, m(i))]\pi_{(t)}$. By expressing this as a mixture of conditional distributions given $\mathcal{M}$ and $\pi_{(t)}$, and then using Jensen's inequality applied to $x \to x \log x$, the entropy can be bounded above by the expected value of

$$\begin{aligned}
&\mathrm{ENT}\left(\mu\left[\prod_{i\,:\,m(i)\le i} c(i, m(i))\right]\pi_{(t)} \mid \mathcal{M}, \pi_{(t)}\right) \\
&= \mathrm{ENT}\left(\mu\left[\prod_{i\,:\,m(i)\le i} c(i, m(i))\right] \,\Big|\, \mathcal{M}, \pi_{(t)}\right) \qquad (20)\\
&= \mathrm{ENT}\left(\mu\left[\prod_{i\,:\,m(i)\le i} c(i, m(i))\right] \,\Big|\, \mathcal{M}\right), \qquad (21)
\end{aligned}$$

where the first equality holds by Proposition 6 and the second equality holds because the permutation μ, the product of collisions $c(i, m(i))$ and $\pi_{(t)}$ are conditionally independent given $\mathcal{M}$. For $1 \le k \le n$, let

$$\nu_k = \prod_{i\,:\,m(i)\le i\le k} c(i, m(i)).$$

Note that the right-hand side of (21) is $\mathrm{ENT}(\mu\nu_n \mid \mathcal{M})$ and $\nu_0 = \mathrm{id}$. Since μ is independent of $\mathcal{M}$, we have $\mathrm{ENT}(\mu \mid \mathcal{M}) = \mathrm{ENT}(\mu)$ and hence

$$\mathrm{ENT}(\mu\nu_n \mid \mathcal{M}) - \mathrm{ENT}(\mu) = \sum_{k=1}^{n} \mathrm{ENT}(\mu\nu_k \mid \mathcal{M}) - \mathrm{ENT}(\mu\nu_{k-1} \mid \mathcal{M}).$$

Thus, it is enough to show that for every k we have

$$\mathbf{E}\big(\mathrm{ENT}(\mu\nu_k \mid \mathcal{M}) - \mathrm{ENT}(\mu\nu_{k-1} \mid \mathcal{M})\big) \le \frac{-CA_kE_k}{\log n}. \qquad (22)$$

Note that if $m(k) > k$, then $\nu_k = \nu_{k-1}$. If $m(k) \le k$, then $\nu_k = \nu_{k-1}c(k, m(k))$. We can now proceed in a way, that is, analogous to the proof of Lemma 7. Note that when $m(k) \le k$ we have

$$\mu\nu_k = \tfrac{1}{2}\mu\nu_{k-1} + \tfrac{1}{2}\mu\nu_{k-1}(k, m(k)).$$

Fix $i \le k$, let $\lambda = \mu\nu_{k-1}$ and let $\widehat{\lambda} = \lambda(k, i)$. Note that $\widehat{\lambda}$ and λ are the same, except that $\widehat{\lambda}^{-1}(k) = \lambda^{-1}(i)$ and $\lambda^{-1}(k) = \widehat{\lambda}^{-1}(i)$. Note also that ν_{k-1} has $k +$

$1, \dots, n$ as fixed points, so $(\lambda^{-1}(k+1), \dots, \lambda^{-1}(n)) = (\mu^{-1}(k+1), \dots, \mu^{-1}(n))$. Let

$$\begin{aligned}\mathcal{F}_{k+1} &= \sigma\big(\mu^{-1}(k+1), \dots, \mu^{-1}(n)\big) \\ &= \sigma\big(\lambda^{-1}(k+1), \dots, \lambda^{-1}(n)\big)\end{aligned}$$

and define $\widehat{\mathcal{F}}_{k+1} = \sigma(\mathcal{F}_{k+1}, \mathcal{M})$. Then we have $\mathrm{ENT}(\widehat{\lambda} \mid \widehat{\mathcal{F}}_{k+1}) = \mathrm{ENT}(\lambda \mid \widehat{\mathcal{F}}_{k+1})$ and hence

$$\mathrm{ENT}(\lambda c(k,i) \mid \widehat{\mathcal{F}}_{k+1}) - \mathrm{ENT}(\lambda \mid \widehat{\mathcal{F}}_{k+1}) = -d(\mathcal{L}(\widehat{\lambda} \mid \widehat{\mathcal{F}}_{k+1}), \mathcal{L}(\lambda \mid \widehat{\mathcal{F}}_{k+1})).$$

But by the projection lemma,

$$\begin{aligned}d(\mathcal{L}(\widehat{\lambda} \mid \widehat{\mathcal{F}}_{k+1}), \mathcal{L}(\lambda \mid \widehat{\mathcal{F}}_{k+1})) &\geq d(\mathcal{L}(\widehat{\lambda}^{-1}(k) \mid \widehat{\mathcal{F}}_{k+1}), \mathcal{L}(\lambda^{-1}(k) \mid \widehat{\mathcal{F}}_{k+1})) \\ &= d(\mathcal{L}(\lambda^{-1}(i) | \widehat{\mathcal{F}}_{k+1}), \mathcal{L}(\lambda^{-1}(k) \mid \widehat{\mathcal{F}}_{k+1})).\end{aligned}$$

Thus, since $m(k)$ is $\widehat{\mathcal{F}}_{k+1}$-measurable, on the event that $m(k) \leq k$ we have

$$\begin{aligned}&\mathrm{ENT}(\mu\nu_k \mid \widehat{\mathcal{F}}_{k+1}) - \mathrm{ENT}(\mu\nu_{k-1} \mid \widehat{\mathcal{F}}_{k+1}) \\ &\quad= \mathrm{ENT}(\lambda c(k, m(k)) \mid \widehat{\mathcal{F}}_{k+1}) - \mathrm{ENT}(\lambda \mid \widehat{\mathcal{F}}_{k+1}) \\ &\quad\leq -d(\mathcal{L}(\lambda^{-1}(m(k)) | \widehat{\mathcal{F}}_{k+1}), \mathcal{L}(\lambda^{-1}(k) \mid \widehat{\mathcal{F}}_{k+1})) \\ &\quad= -\sum_{i \leq k} \mathbf{1}(m(k) = i) d(\mathcal{L}(\mu^{-1}(i) \mid \mathcal{F}_{k+1}), \mathcal{L}(\mu^{-1}(k) \mid \mathcal{F}_{k+1})),\end{aligned}$$

where in the third line we replaced λ by μ because ν_{k-1} does not contain the collision $c(k, m(k))$ and hence has k and $m(k)$ as fixed points, and we replaced the sigma field $\widehat{\mathcal{F}}_{k+1}$ by $\mathcal{F}_{k+1}$ because μ is independent of $\mathcal{M}$. Taking expectations gives

$$\begin{aligned}&\mathbf{E}\big(\mathrm{ENT}(\mu\nu_k \mid \widehat{\mathcal{F}}_{k+1}) - \mathrm{ENT}(\mu\nu_{k-1} \mid \widehat{\mathcal{F}}_{k+1})\big) \\ &\quad\leq -\mathbf{E}\left(\sum_{i \leq k} \mathbf{P}\big(m(k) = i\big) d(\mathcal{L}(\mu^{-1}(i) \mid \mathcal{F}_{k+1}), \mathcal{L}(\mu^{-1}(k) \mid \mathcal{F}_{k+1}))\right) \\ &\quad\leq -\mathbf{E}\left(A_k k^{-1} \sum_{i \leq k} d(\mathcal{L}(\mu^{-1}(i) \mid \mathcal{F}_{k+1}), \mathcal{L}(\mu^{-1}(k) \mid \mathcal{F}_{k+1}))\right) \\ &\quad\leq -\mathbf{E}\left(A_k d\left(k^{-1} \sum_{i \leq k} \mathcal{L}(\mu^{-1}(i) \mid \mathcal{F}_{k+1}), \mathcal{L}(\mu^{-1}(k) \mid \mathcal{F}_{k+1})\right)\right),\end{aligned} \tag{23}$$

where the second inequality follows by the A_k-uniformity of $m(k)$ and the independence of $m(k)$ and μ, and the third inequality is by Proposition 3. The first argument of $d(\cdot, \cdot)$ in the right-hand side of (23) is the uniform distribution over

$\{1, \dots, n\} - \{\mu^{-1}(k+1), \dots, \mu^{-1}(n)\}$. Thus the right-hand side of (23) is

$$-A_k \mathbf{E}(d(\mathcal{U}, \mathcal{L}(\mu^{-1}(k) \mid \mathcal{F}_{k+1}))) \tag{24}$$

$$\leq -\frac{CA_k}{\log n}\mathbf{E}(\mathrm{ENT}(\mathcal{L}(\mu^{-1}(k) \mid \mathcal{F}_{k+1}))) = -\frac{CA_kE_k}{\log n}, \tag{25}$$

where the inequality holds by Lemma 5. Since $\mu\nu_k$ and $\mu\nu_{k-1}$ agree in positions $k+1, \dots, n$, the portion of their respective entropies that are attributable to those positions coincides, hence Proposition 1 and (25) yield the theorem. □

REMARK. Since for any distribution p we have $d(p, p) = 0$, (23) is still true if $m(k)$ is only A_k-uniform over $\{0, \dots, k-1\}$. So the assumptions of the theorem can be relaxed so that there is no lower bound necessary on the probability that $m(k) = k$.

5. Thorp shuffle. In this section we show that Theorem 9 implies an improved bound for the Thorp shuffle. Recall that the Thorp shuffle has the following description: assume that the number of cards, n, is even. Cut the deck into two equal piles. Drop the first card from the left pile or the right pile according to the outcome of a fair coin flip; then drop from the other pile. Continue this way, with independent coin flips deciding whether to drop LEFT-RIGHT or RIGHT-LEFT each time, until both piles are empty.

We will actually work with the time reversal of the Thorp shuffle, which is well known and easily shown to have the same mixing time. Suppose that we label the positions in the deck $0, 1, \dots, n-1$. Note that the Thorp shuffle can be described in the following way: with each step, for x with $0 \leq x \leq \frac{n}{2} - 1$, the cards at positions x and $x + n/2$ collide and are moved to positions $2x \bmod n$ and $2x + 1 \bmod n$. Thus, the time reversal can be described as follows: each step, for even numbers $x \in \{0, \dots, n-2\}$, the cards in positions x and $x+1$ collide and are moved to positions $x/2$ and $x/2 + n/2$. More precisely, a reverse Thorp shuffle can be written in Monte form as

$$\nu c(0, n/2)c(1, n/2+1) \cdots c(n/2-1, n-1),$$

where ν is the (deterministic) permutation that moves the card in position x either to position $x/2$ or to position $x/2 + n/2$, according to whether x is even or odd.

We write $\pi_{(t)}$ for a product of t i.i.d. copies of the reverse Thorp shuffle. Our main lemma is the following:

LEMMA 10. *Let $t = \lceil \log_2 n \rceil$. There is a universal constant C such that for any random permutation μ we have*

$$\mathrm{ENT}(\mu\pi_{(t)}) \leq (1 - C/\log^2 n)\,\mathrm{ENT}(\mu).$$

PROOF. Partition the locations $0, \ldots, n-1$ into intervals I_m as follows: let $I_0 = \{0\}$, and for $m = 1, 2, \ldots, \lceil \log_2 n \rceil$, define $I_m = \{2^{m-1}, \ldots, 2^m - 1\} \cap \{0, \ldots, n-1\}$.

For $i \in \{0, \ldots, n-1\}$, define $E_i = \mathrm{ENT}(\mu, i)$. We can write the entropy of μ as

$$\mathrm{ENT}(\mu) = \sum_m \sum_{i \in I_m} E_i.$$

Let m^* be the value of m that maximizes $\sum_{i \in m} E_i$. Then

$$\sum_{j \in I_{m^*}} E_j \geq \frac{c}{\log n} \mathrm{ENT}(\mu)$$

for a constant c. Since the reverse Thorp shuffle is in Monte form, we may use Theorem 9. We will also use the remark immediately following Theorem 9, which says that the distribution of the card matched with i need only be A_i uniform over $\{j : j < i\}$ in order for the conclusions of the theorem to hold. Fix m with $1 \leq m \leq \lceil \log_2 n \rceil$. We will show that the assumptions of the theorem hold with $t = \lceil \log_2 n \rceil$,

$$A_i = \begin{cases} 1/4, & \text{if } i \in I_m, \\ 0, & \text{otherwise,} \end{cases}$$

and the random variable T defined as follows: let T be any random variable that satisfies

$$\mathbf{P}(T = r) \geq 2^{r-m-1} \tag{26}$$

for $r = 0, \ldots, m$.

Fix $i \in I_m$. We shall show that for any $j < i$ we have $\mathbf{P}(m(i) = j) \geq 1/4i$. Define $f : \mathbf{Z} \to \mathbf{Z}$ by $f(t) = \lfloor t/2 \rfloor$. Note that if $X_s(j)$ denotes the position of card j at time s, then

$$X_s(j) = f(X_{s-1}(j)) + Z_s(j), \tag{27}$$

where $Z_s(j)$ is a random "offset" whose distribution is uniform over $\{0, n/2\}$. Note that in step of the shuffle, the distance between a pair of cards is cut roughly in half if they have the same offsets. More precisely, if $x > y$ then

$$f(x) - f(y) \leq \begin{cases} (x-y)/2, & \text{if } x \text{ is odd or } y \text{ is even,} \\ (x-y)/2 + \frac{1}{2}, & \text{otherwise.} \end{cases} \tag{28}$$

It follows that $\lceil \log_2(f(x) - f(y)) \rceil \leq \lceil (\log_2(x-y)) \rceil$ and $\lceil \log_2(f(x) - f(y)) \rceil \leq \lceil (\log_2(x-y)) \rceil - 1$ unless $x = y + 1$ and x is even.

Say that two positions x and y are neighbors if $|x - y| = 1$ and $\min(x, y)$ is even. (Note that in each step of the reverse Thorp shuffle, the neighbors collide.) Since n is even we can write $n/2 = 2^k l$ for some $k \geq 0$ and odd integer l. Fix i and j with $j \leq i$. We shall show that $\mathbf{P}(m(i) = j) \geq 1/4i$.

First, we claim that $\mathbf{P}(X_m(j) \text{ is even}) \geq \frac{1}{2}$. To see this for the case $m \leq k$ we shall use the following lemma. Let f^r denote the r-fold iterate of f.

LEMMA 11. *If $s \le k$, then $X_s(j) = f^s(j) + \sum_{r=0}^{s-1} 2^{-r} Z_{s-r}(j)$.*

PROOF. The proof will be by induction on s. The base case $s = 1$ follows from (27). Now suppose that $s < k$ and

$$X_s(j) = f^s(j) + \sum_{r=0}^{s-1} 2^{-r} Z_{s-r}(j). \tag{29}$$

Since $n/2 = 2^k l$, and each of the terms Z_{s-r} is either 0 or $n/2$, the sum in (29) is even. It follows that

$$\begin{aligned} X_{s+1}(j) &= f(X_s(j)) + Z_{s+1} \\ &= \left(f^{s+1}(j) + \tfrac{1}{2} \sum_{r=0}^{s-1} 2^{-r} Z_{s-r}(j) \right) + Z_{s+1} \\ &= f^{s+1}(j) + \sum_{r=0}^{s} 2^{-r} Z_{s+1-r}(j). \end{aligned}$$

□

Since $f^m(j) = 0$, Lemma 11 implies that if $m \le k$, then $X_m(j) = \sum_{r=0}^{m-1} 2^{-r} \times Z_{m-r}(j)$. It follows that $X_m(j)$ is even since each of the $Z_{m-r}(j)$ is either 0 or $2^k l$, so each term in the sum is even.

Assume now that $m > k$. Suppose that the value of $Z_{m-k}(j)$ (which is either 0 or $n/2$) is determined by an unbiased coin flip. For $m - k \le s \le m$, let $X'_s(j)$ be what the position of card j at time s would have been if the outcome of the coin flip determining Z_{m-k} had been different. Since $f(x) - f(y) = \frac{1}{2}(x - y)$ if $x - y$ is even, it follows that $|X'_s(j) - X_s(j)| = 2^{m-s} l$ for $m - k \le s \le m$. Thus $|X'_m(j) - X_m(j)| = l$, which is odd. So one of $X'_m(j)$ and $X_m(j)$ is odd and the other is even. Since they have the same distribution, they are each even with probability $\frac{1}{2}$.

Let $y_0 = X_0(i)$, and for $s \ge 1$ let $y_s = f(y_{s-1}) + Z_s(j)$, that is, where card i would be located after s steps if its offsets were the same as those for j. Let $\tau = \min\{s : |y_s - X_s(j)| = 1 \text{ and } X_s(j) \text{ is even}\}$. Since $|i - j| \le 2^m$ (28) and the sentence immediately following it imply that there must be a value of $s \le m$ such that $|y_s - X_s(j)| = 1$. Combining this with the fact that $X_m(j)$ is even with probability at least $\frac{1}{2}$ gives $\mathbf{P}(\tau \le m) \ge \frac{1}{2}$. Furthermore, given $\tau = r$, the conditional probability that $X_s(i) = y_s$ for $0 \le s \le r$ (and hence i and j collide at time τ) is 2^{-r}. Finally, since assumption (26) gives $\mathbf{P}(T = r) \ge 2^{r-m-1}$, it follows that $\mathbf{P}(m(i) = j) \ge 2^{-m-2} \ge \frac{1}{4i}$.

We have shown that the assumptions of Theorem 9 are met with $t = \lceil \log_2 n \rceil$ and $A_i = 1/4$ for $i \in I_m$. Applying this with $m = m^*$ shows that for any permutation μ, we have $\mathrm{ENT}(\mu\pi_{(t)}) \le (1 - C/\log^2 n)\,\mathrm{ENT}(\mu)$, for a universal constant C.

It follows that for any $B \in \{1, 2, \ldots\}$ we have

$$\begin{aligned}\operatorname{ENT}\bigl(\pi_{(Bt\log^3 n)}\bigr) &\le (1 - C/\log^2 n)^{B\log^3 n} \operatorname{ENT}(\mathrm{id}) \\ &\le n^{1-CB}\log n,\end{aligned}$$

since $\operatorname{ENT}(\mathrm{id}) = \log n! \le n\log n$ and $1 - u \le e^{-u}$ for all u. If B is large enough so that $n^{1-CB}\log n \le \frac{1}{8}$ for all n, then $\operatorname{ENT}(\pi_{(Bt\log^3 n)}) \le \frac{1}{8}$ and hence $\|\pi_{(Bt\log^3 n)} - \mathcal{U}\| \le \frac{1}{4}$ by (2). It follows that the mixing time is at most $Bt\log^3 n = O(\log^4 n)$. $\square$

6. L-reversal chain. In this section we analyze Durrett's L-reversal chain. Recall that the L-reversal chain has two parameters, n and L. The cards are located at the vertices of an n-cycle, which we label $\{0, \ldots, n-1\}$. For each step, a vertex v and a number $l \in \{0, \ldots, L\}$ is chosen independently and uniformly at random. Then the interval of cards $v, v+1, \ldots, v+l$ is reversed, where the numbers are taken mod n. Equivalently, for each step a (nonempty) interval of length at most L (i.e., of size between 1 and $L+1$) is chosen uniformly at random and reversed. We shall assume that $L > L_0$ for a suitable value of L_0 and $n \ge 4L$. The cases where L is constant and where $n \le cL$ for a constant c were both treated in [9].

We put the shuffle in Monte form as follows: let $\mu_{i,j}$ denote the permutation that reverses the cards in positions $i, i+1, \ldots, j$ and leaves the rest unchanged. Let Z be uniform over $\{1, \ldots, L\}$. Choose v uniformly at random from $\{0, \ldots, n-1\}$ and let

$$\pi = \begin{cases} \mu_{v,v+L}, & \text{with probability } \dfrac{1}{2(L+1)}, \\ \mu_{v,v+L-1}, & \text{with probability } \dfrac{1}{2(L+1)}, \\ \mu_{v,v+Z}c(v,Z), & \text{with probability } \dfrac{L}{(L+1)}. \end{cases} \tag{30}$$

It will be convenient to think of the permutations $\mu_{v,v+2}(v, v+2)$ and $\mu_{v,v+1}(v, v+1)$ as being $\mu_{v,v}$. (We can do this since they are all just the identity permutation.) Using this convention, for any $r \in \{0, \ldots, n-1\}$ we have

$$\begin{aligned}&\mathbf{P}(\pi \text{ is assigned the value } \mu_{r,r}) \\ &\quad = \mathbf{P}(v = r)\mathbf{P}(Z \in \{1,2\})\mathbf{P}\bigl(c(v,Z) = (v,Z)\bigr)\frac{L}{L+1} \\ &\quad = \frac{1}{n(L+1)}.\end{aligned}$$

The following remark shows that for any s with $1 \le s \le L$ the probability that π is assigned the value $\mu_{r,r+s}$ is also $\frac{1}{n(L+1)}$, and hence π has the distribution of an L-reversal shuffle.

REMARK. Fix $r \in \{0, \dots, n-1\}$ and s with $1 \le s \le L$. Note that for any value of s there are two ways for π to be $\mu_{r,r+s}$. The first way is to have $v = r$, $Z = s$, $c(v, Z) = \mathrm{id}$, and to make the bottom choice for π in (30). This happens with probability $\frac{1}{2n(L+1)}$. The second way depends on the value of s. If $s < L-1$, then the second way to have $\pi = \mu_{r,r+s}$ is to have $v = r-1$, $Z = s+1$, and $c(v, Z) = (v, Z)$; this occurs with probability $\frac{1}{2n(L+1)}$. If $s = L$ or $s = L-1$, then the second way to have $\pi = \mu_{r,r+s}$ is to have $v = r$ and to make the top or middle choice, respectively, for π in (30); this occurs with probability $\frac{1}{2n(L+1)}$. It follows that $\mathbf{P}(\pi = \mu_{r,r+s}) = \frac{1}{n(L+1)}$ for every r and s; furthermore, given that $\pi = \mu_{r,r+s}$, the conditional probability that π contains the collision $c(r, s)$ is $\frac{1}{2}$. This fact will be used in the sequel.

We write $\pi_{(t)}$ for a product of t i.i.d. copies of the L-reversal shuffle. Our main technical lemma is the following:

LEMMA 12. *There is a universal constant C such that for any random permutation μ there is a value of $t \in \{1, \dots, \frac{Cn^3}{L^3}\}$ such that*

$$\mathrm{ENT}(\mu\pi_{(t)}) \le \big(1 - f(t)\big)\,\mathrm{ENT}(\mu),$$

where $f(t) = \frac{\gamma}{\log^2 n}(\frac{t}{n} \wedge 1)$, for a universal constant γ.

Before proving Lemma 12, we first show how it gives the claimed mixing time bound.

LEMMA 13. *The mixing time for the L-reversal chain is $O((n \vee \frac{n^3}{L^3}) \log^3 n)$.*

PROOF. Let t and f be as defined in Lemma 12. Then

$$\frac{t}{f(t)} = \gamma^{-1}(\log^2 n)t\left(\frac{n}{t} \vee 1\right) = \gamma^{-1}\log^2 n(n \vee t) \le T, \tag{31}$$

where $T = \gamma^{-1}\log^2 n[n \vee \frac{Cn^3}{L^3}]$. Note that $1/T$ is a bound on the long run rate of entropy loss per unit of time. Lemma 12 implies that there is a $t_1 \in \{1, \dots, \frac{Cn^3}{L^3}\}$ such that

$$\mathrm{ENT}(\pi_1 \cdots \pi_{t_1}) \le \big(1 - f(t_1)\big)\,\mathrm{ENT}(\mathrm{id})$$

and a $t_2 \in \{1, \dots, \frac{Cn^3}{L^3}\}$ such that

$$\mathrm{ENT}(\pi_1 \cdots \pi_{t_1+t_2}) \le \big(1 - f(t_2)\big)\,\mathrm{ENT}(\pi_1 \cdots \pi_{t_1}),$$

etc. Continue this way to define t_3, t_4, and so on. For $j \geq 1$ let $\tau_j = \sum_{i=1}^{j} t_i$. Then

$$\mathrm{ENT}(\pi_{(\tau_j)}) \leq \left[\prod_{i=1}^{j}(1 - f(t_j))\right] \mathrm{ENT}(\mathrm{id}) \tag{32}$$

$$\leq \exp\left(-\sum_{i=1}^{j} f(t_j)\right) \mathrm{ENT}(\mathrm{id}). \tag{33}$$

But since $t_j \leq T f(t_j)$ by (31), we have

$$\tau_j = \sum_{i=1}^{j} t_j \leq T \sum_{i=0}^{j} f(t_j).$$

It follows that

$$\mathrm{ENT}(\pi_{(\tau_j)}) \leq \exp\left(\frac{-\tau_j}{T}\right) \mathrm{ENT}(\mathrm{id}). \tag{34}$$

Since $\mathrm{ENT}(\mathrm{id}) = \log n! \leq n \log n$, it follows that if $\tau_j \geq T \log(8n \log n)$ we have $\mathrm{ENT}(\pi_{(\tau_j)}) \leq \frac{1}{8}$ and hence $\|\pi_{(\tau_j)} - \mathcal{U}\| \leq \frac{1}{4}$ by (2). It follows that the mixing time is $O(T \log(8n \log n)) = O((n \vee \frac{n^3}{L^3}) \log^3 n)$. $\square$

We shall now prove Lemma 12.

PROOF OF LEMMA 12. Let $m = \lceil \log_2(n/L) \rceil$. Then we can partition the set of locations $\{0, \ldots, n-1\}$ into $m+1$ intervals as follows: let $I_0 = \{0, \ldots, L\}$, and for $1 \leq k \leq m$ define $I_k = \{2^{k-1}L + 1, \ldots, 2^k L\} \cap \{0, \ldots, n-1\}$. Define $E_k = \mathbf{E}(\mathrm{ENT}(\mu, k))$. Note that we can write the entropy of μ as

$$\mathrm{ENT}(\mu) = \sum_{k=0}^{m} \sum_{j \in I_k} E_j. \tag{35}$$

Thus, if k^* maximizes $\sum_{j \in I_k} E_j$, then

$$\sum_{j \in I_k^*} E_j \geq \frac{1}{m+1} \mathrm{ENT}(\mu).$$

Suppose first that $k^* = 0$. In this case we can take $t = 1$, as we now show. Let π be a random permutation corresponding to one move of the L-reversal chain. Let E be the event that π reverses $a, a+1, \ldots, b$ for $a, b \in \{0, \ldots, L\}$. Then (using an abuse of notation similar to that in Section 3.2) we can write π as

$$\pi = \alpha \pi_1 + (1 - \alpha)\pi_2,$$

where $\alpha = \mathbf{P}(E)$, π_1 is π conditioned on E, and π_2 is π conditioned on E^c. Then $\mu\pi = \alpha\mu\pi_1 + (1-\alpha)\mu\pi_2$ and hence

$$\mathrm{ENT}(\mu\pi) = \mathrm{ENT}\big(\alpha\mu\pi_1 + (1-\alpha)\mu\pi_2\big) \tag{36}$$

$$\leq \alpha\,\mathrm{ENT}(\mu\pi_1) + (1-\alpha)\,\mathrm{ENT}(\mu\pi_2) \tag{37}$$

$$\leq \alpha\,\mathrm{ENT}(\mu\pi_1) + (1-\alpha)\,\mathrm{ENT}(\mu), \tag{38}$$

where both inequalities follow from the convexity of $x \to x\log x$. It follows that

$$\mathrm{ENT}(\mu\pi) - \mathrm{ENT}(\mu) \leq \alpha[\mathrm{ENT}(\mu\pi_1) - \mathrm{ENT}(\mu)]. \tag{39}$$

Note that π_1 does not move any of the cards in locations $\{L+1,\dots,n\}$. Hence by Proposition 1, the entropy difference $\mathrm{ENT}(\mu\pi_1) - \mathrm{ENT}(\mu)$ is the expected loss in entropy attributable to positions $\{0,\dots,L\}$, that is, $\mathbf{E}(\mathrm{ENT}(\mu\pi_1 \mid \mathcal{F}_{L+1}) - \mathrm{ENT}(\mu \mid \mathcal{F}_{L+1}))$, where $\mathcal{F}_{L+1} = \sigma(\mu^{-1}(L+1),\dots,\mu^{-1}(n-1))$. The permutation π_1 is a step of a modified L-reversal chain on the $L+1$ cards in the line graph $\{0,\dots,L\}$, reversing an interval of the form $a, a+1,\dots,b$ for $0 \leq a \leq b \leq L$.

In Theorem 6 of [9], it is shown (by comparison with shuffling through random transpositions [7]; see [6] for background on comparison techniques) that the log Sobolev constant for the L-reversal chain on n cards is at most $B\frac{n^3}{L^2}\log n$ for a constant B. This remains true if we consider the modified L-reversal process on the line graph. Thus π_1 has a log Sobolev constant, that is, at most $2BL\log L$, and hence (by the well-known relationship between the log Sobolev constant and decay of relative entropy; see, e.g., [12]) multiplying μ by π_1 reduces the relative entropy by at least $1/B'L\log L$ times the entropy attributable to positions $\{0,\dots,L\}$, for a constant B'. Thus the right-hand side of (39) is at most

$$-\alpha(B'L\log L)^{-1}\sum_{j\in I_1} E_j \leq -(8B'n\log L)^{-1}\sum_{j\in I_1} E_j \tag{40}$$

$$= -(8B'n\log^2 n)^{-1}\,\mathrm{ENT}(\mu), \tag{41}$$

where the second line follows from the fact that $\alpha \geq \frac{L}{8n}$.

Next we shall consider the case where $k^* \geq 1$, so that the interval is of the form $\{2^{k-1}L+1,\dots,2^kL\} \cap \{0,1,\dots,n-1\}$. We will use Theorem 9 to get a decay of entropy in this case. We make the following claim.

CLAIM 14. *Fix $k \geq 1$. There are universal constants C and $\alpha > 0$ such that if $t = 4^kCn/L^3$, $T = t/2$ and*

$$A_y = \begin{cases} \alpha\left(\dfrac{t}{n} \wedge 1\right), & \text{if } y \in I_k, \\ 0, & \text{otherwise}, \end{cases}$$

then the assumptions of Theorem 9 are satisfied by t, T and the A_y.

In order to prove this claim, it is helpful to know that the L-reversal chain enjoys certain monotonicity properties. Roughly speaking, the closer two cards are together, the more likely they are to collide after a given number of steps. Before proving Claim 14, we shall verify these monotonicity properties.

Two types of monotonocity. Fix x and y in $\{0,\dots,n\}$ and let x_m and y_m denote the positions of cards x and y, respectively, at time m. Define $Z_m = |x_m - y_m|$, that is, the graph distance between x_m and y_m in the n-cycle. Note that Z_m is a Markov chain. We shall need the following lemma.

LEMMA 15. *Let $\widehat{P}$ denote the transition matrix of Z_m. Then $\widehat{P}$ is monotone, that is, if $b \geq a$ then $\widehat{P}(b,\cdot) \succeq \widehat{P}(a,\cdot)$, where $\succeq$ denotes stochastic domination.*

PROOF. Fix positions u and a with $a \leq n/2$, and let $N(a,u)$ denote the number of legal intervals (i.e., intervals of length at most L) that move the card in position a to position u without moving the card in position 0. Then

$$N(a,u) = \begin{cases} \min\bigl(u, \lfloor \frac{1}{2}(L-a+u)+1 \rfloor\bigr), & \text{if } u < a, \\ \min\bigl(a, \lfloor \frac{1}{2}(L-u+a)+1 \rfloor\bigr), & \text{if } u > a. \end{cases}$$

(Recall that we assume that $n \geq 4L$.) Suppose that $|x_m - y_m| = a$. For $u \leq n/2$, let $M(a,u)$ denote the number of legal intervals whose reversal at time m would make $|x_{m+1} - y_{m+1}| = u$. If $a \neq u$ then $M(a,u)$ counts intervals that move x but not y and intervals that move y but not x. Thus we have $M(a,u) = 2(N(a,u) + N(a,n-u))$. It is easily verified that $M(a,u)$ is nonincreasing in a for $u < a \leq n/2$ and nondecreasing in a for $0 < a < u$. It follows that Z_m is monotone. $\square$

We now prove that Z_m has another type of monotonicity property. Note that in each move of the L-reversal process, there are exactly four cards that are adjacent to a different pair of cards after the move than they were before. We say that those cards are *cut* and write, for example, "card i is cut at time m." We say that a location is cut if the card in that location is cut.

The cut-stopped process. It will be convenient to consider a modified version Z'_m of Z_m, where we introduce two absorbing states 0 and ∞, and have the following occur when either x or y is cut. If x and y are within a distance L of each other, then Z'_m transitions to 0; otherwise, it transitions to ∞.

We shall call this modified process the *cut-stopped process*. We can impose an order on the state space of $\{Z'_m : m \geq 0\}$ based on the order of the positive integers, with the additional states 0 and ∞ as the minimum and maximum states, respectively.

Our next lemma says that the cut-stopped process Z'_m is monotone with respect to this order. Let Q denote the transition matrix of the cut-stopped process.

LEMMA 16. *The cut-stopped process is monotone, that is, if $b \geq a$, then $Q(b,\cdot) \succeq Q(a,\cdot)$, where $\succeq$ denotes stochastic domination.*

PROOF. The proof is a slight modification of the proof of Lemma 15. Suppose that $Z'_m = z$. Note that the probability of absorbing in 0 in the next step is a nonincreasing function of z, and the probability of absorbing in ∞ in the next step is a nondecreasing function of z. The rest of the argument is almost identical to the proof of Lemma 15. Fix positions u and a with $a \leq n/2$, and let $N'(a,u)$ denote the number of intervals of length at most L that move the card in position a to position u, but neither move the card in position 0, cut position 0, nor cut position a. Then

$$N'(a,u) = \begin{cases} \max\bigl(0, \min\bigl(u-2, \lfloor \frac{1}{2}(L-a+u)\rfloor\bigr)\bigr), & \text{if } u < a, \\ \max\bigl(0, \min\bigl(a-2, \lfloor \frac{1}{2}(L-u+a)\rfloor\bigr)\bigr), & \text{if } u < a. \end{cases}$$

Suppose that $|x_m - y_m| = a$. For $u \leq n/2$, let $M'(a,u)$ denote the number of legal intervals that do not cut x or y and whose reversal at time m would make $|x_{m+1} - y_{m+1}| = u$. If $a \neq u$, then $M'(a,u) = 2(N'(a,u) + N'(a,n-u))$. It is easily verified that $M'(a,u)$ is nonincreasing in a for $u < a \leq n/2$ and nondecreasing in a for $0 < a < u$. It follows that Z'_m is monotone. $\square$

We are now ready to prove Claim 14. For the convenience of the reader, we state the claim again. Recall that $I_k = \{2^{k-1}L + 1, \ldots, 2^k L\} \cap \{0, \ldots, n-1\}$.

CLAIM 14. *There are universal constants C and $\alpha > 0$ such that if $t = 4^k Cn/L^3$, $T = t/2$ and*

$$A_y = \begin{cases} \alpha\left(\dfrac{t}{n} \wedge 1\right), & \text{if } y \in I_k; \\ 0, & \text{otherwise}, \end{cases}$$

then the assumptions of Theorem 9 are satisfied by t, T and the A_y.

PROOF. Let $y \in I_k$. We need to show that if $x \leq y$, then with probability at least A_y, cards x and y collide between time T and time t, and this is the first collision in which either is involved after time T.

Fix $y \in I_k$ and x with $x < y$. Let τ be the first time after time T that either x or y is cut. Note that if x and y collide at time τ and $\tau \leq t$ then $m(x) = y$. Thus, given that $|x_\tau - y_\tau| \leq L$ and $\tau \leq t$ the conditional probability that $m(x) = y$ is at least $1/8L$. This is because the number of intervals that cut either x or y is at most $4L$, so the conditional probability that x and y are at the endpoints of the interval, that is, reversed at time τ is at least $1/4L$. The conditional probability that x and y collide is at least half of this by the remark following (30).

Thus it is enough to show that for a universal constant α we have

$$\mathbf{P}(|x_\tau - y_\tau| \le L, \tau \le t) \ge \alpha\left(\frac{t}{n} \wedge 1\right)L/y. \tag{42}$$

For $m \ge 0$ let $Z_m = |x_m - y_m|$. Let $\beta > 0$ be a constant and suppose that $L > 2\beta$. We claim that with probability bounded away from 0 we have $|Z_m| \le \beta L$ for some $m < T$. To see this, let $M = \min\{m : Z_m \le L\}$. First, we will show that with probability bounded away from zero we have $M \le T'$, where $T' = T/2$. Suppose that $Z_0 > L$. Note that we can write $x_m = x + W_1^x + \cdots + W_m^x (\mathrm{mod}\, n)$, where $W_j^x \in \{-L, \ldots, L\}$ is the displacement of card x at time j. Define $x'_m = x + W_1^x + \cdots + W_m^x$ (i.e., like x_m, but without the mod n), with a similar definition for y'_m. Define $z'_m = y'_m - x'_m$. Note that the conditional distribution of $z'_{m+1} - z'_m$, given that $|x_m - y_m| = r$, does not depend on r if $r > L$. Let X be a random variable with this common distribution, and let $X_1, X_2, \ldots$ be i.i.d. copies of X. Then the random variable $z'_{T'} - z'_0$ can be coupled with the X_i in such a way that $z'_{T'} - z'_0 = \sum_{1=1}^{T'} X_i$ on the event that $M > T'$. Since $M \le T'$ whenever $z'_{T'} - z'_0 \le -Z_0$, it follows that $\mathbf{P}(M \le T') \ge \mathbf{P}(\sum_{i \le T'} X_i \le -Z_0)$. But since, when X is nonzero (which happens with probability on the order of L/n), it has a typical value on the order of L, it has second and third moments satisfying $\sigma^2 \ge C_2 L^3/n$ and $\rho \le C_3 L^4/n$, respectively. Furthermore, we have $\mathbf{E}(X) = 0$. Thus, Berry-Esséen bounds (see, e.g., [8]) imply that for a universal constant C_B we have

$$|F_{T'}(x) - \Phi(x)| \le \frac{C_B \rho}{\sigma^3 \sqrt{T'}} \le \frac{C' L}{Cy}, \tag{43}$$

where $F_{T'}$ is the cumulative distribution function (c.d.f.) of $\frac{1}{\sigma\sqrt{T'}} \sum_{i \le T'} X_i$; Φ is the standard normal c.d.f.; C' is a constant that incorporates C_2, C_3 and C_B; and C is the constant appearing in the definition of t. For the final inequality we use the fact that $t = 4T'$ is within constant factors of $Cy^2 n/L^3$, since $y \in I_k$.

Since $y \ge L$, the quantity (43) can be made arbitrarily close to zero for sufficiently large C. It follows that $\sum_{i \le T'} X_i$ is roughly normal with standard deviation a large constant times y, hence is less than $-Z_0$ with probability bounded away from zero. (Recall that $Z_0 = y - x \le y$.) It follows that with probability bounded away from zero we have $Z_m \le L$ for some $m \le T/2$. Now note that if x and y are within distance L, then given that one of them moves in the next step, the conditional probability that they are brought to within a distance βL is bounded away from zero. Since t is much larger than n/L, there is probability bounded away from zero that either x or y is moved between time m and $m + T/2$. This verifies the claim.

The above claim and the strong Markov property imply that in order to show (42), it is enough to show that if $|i - j| \le \beta L$, $m' \le T/2$ and τ is the first time that i or j is cut after time m', then for a universal constant $\alpha > 0$ we have $\mathbf{P}(|i_\tau - j_\tau| \le L, \tau \le m' + t/2) \ge \alpha(\frac{t}{n} \wedge 1)L/y$.

For every pair of cards i and j, let $T(i,j)$ be the first time that either i or j is cut after time m'. Define $t' = \min(t/2, n)$. Let $A(i,j)$ be the event that $T(i,j) \le m' + t'$ and at time $T(i,j)$ the distance between i and j is most L. Let $f(i,j) = \mathbf{P}(A(i,j))$. Since $t' \le t/2$, it is enough to prove that if $|i-j| \le \beta L$, then

$$f(i,j) \ge \alpha\left(\frac{t}{n} \wedge 1\right) L/y. \tag{44}$$

Since the probability that either i or j is involved in a cut on any given step is at most $8/n$, we have

$$f(i,j) \le \min(1, 8t'/n). \tag{45}$$

Also, note that

$$\begin{aligned}\sum_{i,j} f(i,j) &= \sum_{l=m'+1}^{m'+t'} \sum_{i,j} \mathbf{P}\big(T(i,j) \ge l, |i_l - j_l| \le L, \text{either } i \text{ or } j \text{ is cut at time } l\big) \\ &= \sum_{k=1}^{t'} \sum_{\substack{u<v \\ |u-v|\le L}} g(u,v,k),\end{aligned}$$

where $g(u,v,k)$ is the probability that cards in locations u and v are cut at time $m'+k$, but neither had been cut since time m'. Since the L-reversal process is symmetric it is its own time-reversal. Thus, $g(u,v,k)$ is the probability that either location u or v is cut in the first move, but neither the card in location u at time 1 nor the card in location v at time 1 is cut in the next $k-1$ moves. This probability is at least $\frac{1}{n}(\frac{n-8}{n})^{t'-1}$. Since there are nL such pairs (u,v), summing over u, v and k gives

$$\begin{aligned}\sum_{i,j} f(i,j) &\ge t' n L \frac{1}{n}\left(\frac{n-8}{n}\right)^{t'-1} \\ &\ge c' L t'\end{aligned}$$

for a universal constant c', where the second inequality holds because $t' \le n$. It follows that for any i we have

$$\sum_{j} f(i,j) = \frac{1}{n}\sum_{i,j} f(i,j) \ge c' L t'/n. \tag{46}$$

Let $g(i,j) = \mathbf{P}(A(i,j) \cap B(i,j))$ where $B(i,j)$ is the event that at no time before time $T(i,j)$ was the distance between i and j greater than Dy, where the constant D is to be specified below. We shall now show that $g(i,j)$ is nonincreasing in $|i-j|$. Let Z_m and Z'_m be as defined in Lemmas 15 and 16. Let $\{U_m\}$ be a (not

time-homogeneous) Markov chain whose transition rule is the same as Z_m when $m < m'$ and the same as Z'_m, when $m > m'$. Then

$$g(i,j) = \mathbf{P}\Big(U_{m'+t'} = 0, \max_{m \le m'+t'} U_m \le Dy | U_0 = |i-j|\Big). \tag{47}$$

The quantity (47) is nonincreasing in $|i-j|$ since, by Lemmas 15 and 16, the process $\{U_m\}$ is monotone (in the sense that if $r < s$, then the conditional distribution of U_m given $U_0 = s$ stochastically dominates the conditional distribution of U_m given $U_0 = r$).

Note also that

$$\sum_j g(i,j) = \sum_j f(i,j) - \mathbf{P}\big(A(i,j) \cap B^c(i,j)\big), \tag{48}$$

where $B^c(i,j)$ denotes the complement of $B(i,j)$. We claim that $\sum_j g(i,j) \ge cLt'/n$ for a universal constant c. To see this, fix a card i and $k \le t'$ and say that a card u is *bad* if $|i_0 - u_0| \le L$, and $\max_{0 \le r \le m'+k} |i_r - u_r| > Dy$. Since the L-reversal process is symmetric, and the probability that i or u is cut in any given step is at most $8/n$, we have

$$\sum_j \mathbf{P}\big(A(i,j) \cap B^c(i,j) \cap [T(i,j) = m' + k]\big) \le \frac{8}{n}\mathbf{E}(B), \tag{49}$$

where B is the number of bad cards. Let u be a card initially within distance L of card i. If u_m is the position of card u at time m, then we can write $u_m = u + W_1 + \cdots + W_m (\mathrm{mod}\, n)$, where $W_j \in \{-L, \ldots, L\}$ is the displacement of card u at time j. Define $u'_m = u + W_1 + \cdots + W_m$ (i.e., like u_m, but without the mod n), with a similar definition for i'_m. Then u'_m is a symmetric random walk on the integers. For each step there is a jump with probability on the order of L/n and the sizes of jumps are at most L. It follows that for sufficiently large A, the probability that $\max_{1 \le m \le k} |u'_m - u'| > A(\frac{kL}{n})^{1/2} L$ can be made arbitrarily close to zero. Since k is at most a constant times $\frac{y^2 n}{L^3}$, we have $A(\frac{kL}{n})^{1/2} L \le A' y$ for a constant A'. A similar argument applies to i'_m. Finally, since $|i_m - u_m| \le |i'_m - u'_m|$ (where the first $|\cdot|$ refers to distance in the n-cycle), it follows that for any $\varepsilon > 0$, if D is large enough, then $\mathbf{P}(\max_{1 \le m \le k} |i_m - u_m| > Dy) < \varepsilon$. Thus, since there are at most $2L$ cards initially within a distance L of card i, we have $E(B) \le 2L\varepsilon$. Hence, summing (49) over $k \le t'$ gives

$$\sum_j \mathbf{P}\big(A(i,j) \cap B^c(i,j)\big) \le 16 L \varepsilon t'/n. \tag{50}$$

Combining this with (48) and (46) gives

$$\sum_j g(i,j) \ge cLt'/n \tag{51}$$

for a constant c, if ε is small enough. We now define β to be a constant smaller than $c/32$. Since for any j we have $g(i,j) \le f(i,j) \le 8t'/n$ [by (45)], we have $\sum_{j:|i-j|\le\beta L} g(i,j) \le 16\beta Lt'/n \le cLt'/2n$, and hence

$$\sum_{j:|i-j|>\beta L} g(i,j) \ge cLt'/2n,$$

by (51). Since $g(i,j) = 0$ for $|j-i| > Dy$, the average value of $g(i,j)$ where j ranges over values such that $\beta L < |i-j| \le Dy$ must be at least $cLt'/4Dyn \ge \alpha L(\frac{t}{n} \wedge 1)/y$ for a constant α. Since $g(i,j)$ is nonincreasing in $|i-j|$ [as shown in the discussion following (46)], it follows that $g(i,j) \ge \alpha L(\frac{t}{n} \wedge 1)/y$ if $|i-j| \le \beta L$. Since $g \le f$, this verifies (44), which completes the proof of Claim 14. □

Using Claim 14 with $k = k^*$ and applying Theorem 9 gives

$$\mathrm{ENT}(\mu\pi_{(t)}) - \mathrm{ENT}(\mu) \le \frac{-C}{\log^2 n}\left(\frac{t}{n} \wedge 1\right)\mathrm{ENT}(\mu)$$

for a universal constant C and the proof of Lemma 12 is complete. □

Acknowledgments. I am grateful to A. Soshnikov for many valuable conversations.

REFERENCES

[1] BAYER, D. and DIACONIS, P. (1992). Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.* **2** 294–313. MR1161056

[2] BOREL, E. and CHÉRON, A. (1940). *Théorie mathématique du bridge à la portée de tous*. Gauthier-Villars, Paris. MR0070896

[3] CANCRINI, N., CAPUTO, P. and MARTINELLI, F. (2006). Relaxation time of L-reversal chains and other chromosome shuffles. *Ann. Appl. Probab.* **16** 1506–1527. MR2260071

[4] COVER, T. M. and THOMAS, J. A. (1991). *Elements of Information Theory*. Wiley, New York. MR1122806

[5] DIACONIS, P. Personal Communication.

[6] DIACONIS, P. and SALOFF-COSTE, L. (1993). Comparison theorems for reversible Markov chains. *Ann. Appl. Probab.* **3** 696–730. MR1233621

[7] DIACONIS, P. and SHAHSHAHANI, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete* **57** 159–179. MR626813

[8] DURRETT, R. (2003). *Probability*. Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, CA. MR1068527

[9] DURRETT, R. (2003). Shuffling chromosomes. *J. Theoret. Probab.* **16** 725–750. MR2009200

[10] DURRETT, R., YORK, T. and RASMUS, N. (2007). Dependence of paracentric inversion rate on tract length. *BioMedCentral Bioinformatics* **8** 115.

[11] MARKOV, A. A. (1906). Extension of the law of large numbers to dependent events (Russian). *Bull. Soc. Math. Kazan* (2) **15** 135–156.

[12] MICLO, L. (1996). Sur les problèmes de sortie discrets inhomogènes. *Ann. Appl. Probab.* **6** 1112–1156. MR1422980

[13] MONTENEGRO, R. and TETALI, P. (2006). Mathematical aspects of mixing times in Markov chains. In *Foundations and Trends in Theoretical Computer Science* **1** 237–354. Now Publishers, Hanover, MA.

[14] MORRIS, B. (2005). The mixing time of the Thorp shuffle. In *STOC'*05: *Proc*. 37*th Ann*. *ACM Symp*. *Theor. Comput*. 403–412. ACM, New York. MR2181642

[15] POINCARÉ, H. (1912). *Calcul des probabilités*, 2nd ed. Gauthier Villars, Paris.

[16] WILSON, D. B. (2004). Mixing times of Lozenge tiling and card shuffling Markov chains. *Ann*. *Appl*. *Probab*. **14** 274–325. MR2023023

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, DAVIS
DAVIS, CALIFORNIA 95616
USA
E-MAIL: morris@math.ucdavis.edu