The Annals of Applied Probability
2000, Vol. 10, No. 4, 1302–1321

ON THE RATE OF MIXING FOR p-SHUFFLES[1]

BY STEVEN P. LALLEY

University of Chicago

The *p-shuffle* is a natural generalization of the *dovetail* shuffle. It is defined as follows. First, the deck is cut into a top stack and a bottom stack so that the distribution of the size of the top stack is Binomial (N, p), where N is the total number of cards in the deck. Then, conditional on the outcome of the cut, the two stacks are "riffled" in such a way that all possible riffles (interleavings) of these two stacks are equally likely. The main result of the paper is an asymptotic ($N \to \infty$) bound on the number of repetitions needed to "randomize" the deck.

1. Introduction. The *p-shuffle* (or GSR-p shuffle, for Gilbert, Shannon, and Reeds) is a simple and natural generalization of the *dovetail* (or GSR shuffle) studied in [3] and [1]. For $0 < p < 1$ and a deck of size N, the p-shuffle is defined as follows: first, "cut" the deck into two stacks in such a way that the cardinality of the top stack has the Binomial (N, p) distribution. Next, "riffle" the two stacks by dropping cards one at a time from either the top stack or the bottom stack, according to the following rule: at any stage of the riffle, if there are A cards remaining in the top stack and B cards remaining in the bottom stack, then the probability that the next card dropped is from the top stack is $A/(A+B)$. The special case $p = 1/2$ is the dovetail shuffle of [3].

Mixing properties of the dovetail shuffle are elegantly set out by Bayer and Diaconis in [3], following earlier work by Aldous [1]. There is a *cutoff phenomenon* for large decks: If the deck has N cards, then for large N the number of repetitions of the shuffle required for the distribution to approach to within total variation distance $1/e$ of the uniform distribution on all permutations is about $(3/2)\log_2 N$, and the transition to uniformity is rapid. [Recall that the *total variation distance* between two probability distributions μ, ν on a finite set $\mathscr{X}$ is defined to be $(1/2)\sum_{x\in\mathscr{X}} |\mu(x) - \nu(x)|$.] In particular, if $d_N(n)$ is the total variation distance to the uniform distance after n repetitions, then for every $\varepsilon > 0$,

$$\lim_{N\to\infty} \inf_{n\le((3/2)-\varepsilon)\log_2 N} d_N(n) = 1 \tag{1}$$

and

$$\lim_{N\to\infty} \sup_{n\ge((3/2)+\varepsilon)\log_2 N} d_N(n) = 0. \tag{2}$$

Received March 1999; revised December 1999.

[1] Supported in part by NSF Grant DMS-96-26590.

AMS 1991 *subject classifications*. Primary 60C05; secondary 60B15, 20B30.

Key words and phrases. Riffle shuffle, cutoff phenomenon.

Bayer and Diaconis show that the transition to uniformity is even more rapid: see [3] for a precise statement.

It is natural to conjecture that a cutoff phenomenon holds in the p-shuffle, for any value of the parameter p. Moreover, it is tempting for the nonexpert to suspect that routine modifications of the Aldous and Bayer–Diaconis analyses would establish the cutoff. Unfortunately, this is not the case. The dovetail shuffle has a special property not shared by any other p-shuffle: the probability distribution on the permutation group induced by one repetition of the dovetail shuffle is *uniform* on all "riffle permutations." (A *riffle permutation* is defined to be a permutation with either one or two rising sequences, that is, a permutation which may result from one repetition of a p-shuffle.) This uniformity permits an exact formula for the distribution of the deck after n repetitions; see [3]. No such exact formula is available for the p-shuffle in general. Exact formulas for the eigenvalues of the transition probability matrices of the p-shuffles have recently been obtained (see [4] and [5]) but it does not seem that these can be used to determine the mixing rates.

The main result of this paper is an explicit bound on the rate of convergence for the p-shuffle. For each $p \in (0,1)$, let $q = 1 - p$, and define constants C_p and C_p^* by

$$C_p := \frac{3+\theta_p}{-4\log(p^2+q^2)} \tag{3}$$

and

$$C_p^* := \frac{2}{-\log(p^2+q^2)}, \tag{4}$$

where $\theta = \theta_p$ is the unique real number such that

$$p^\theta + q^\theta = (p^2+q^2)^2. \tag{5}$$

Observe that $C_p < C_p^*$, and that $C_{1/2} = 3/(2\log 2)$.

THEOREM 1. *There exists an open neighborhood of $p = \frac{1}{2}$ such that for all p in this neighborhood, it takes at least $(C_p - \varepsilon)\log N$ and no more than $(C_p^* + \varepsilon)\log N$ repetitions of the p-shuffle to randomize a deck of size N. In particular, if $d_N(x)$ represents the total variation distance between the uniform distribution and the distribution of the state of the deck after $\lfloor x \rfloor$ repetitions of the shuffle, then for any $\varepsilon > 0$,*

$$\lim_{N\to\infty} d_N((1-\varepsilon)C_p\log N) = 1 \tag{6}$$

and

$$\lim_{N\to\infty} d_N((1+\varepsilon)C_p^*\log N) = 0. \tag{7}$$

The result (7) is due to Jason Fulman [8]. (See also [7] for discussion of a related shuffling model.) A short proof of (7) is given in Section 2.3 below; see Corollary 3. The relation (6) is considerably harder. Its proof is given in Section 3. If the cutoff phenomenon does indeed occur for the p-shuffle, the cutoff must occur considerably before $C_p^* \log N$, at least when $p = 1/2$, since $C_{1/2}^*$ is strictly larger than $(3/2 \log 2)$. Thus, if the cutoff phenomenon occurs, C_p^* is *not* the right constant. We conjecture the right constant is C_p, at least for p sufficiently near 1/2:

CONJECTURE 1. *For p sufficiently near $\frac{1}{2}$, there is a "cutoff phenomenon" at $C_p \log N$, that is* [*in light of* (6)], *for any* $\varepsilon > 0$,

$$\lim_{N\to\infty} d_N((1+\varepsilon)C_p \log N) = 0. \tag{8}$$

A heuristic argument in favor of this conjecture is given in Section 4. It seems likely that for *every* $p \in (0, 1)$ the cutoff phenomenon occurs at $C_p \log N$ for a suitable constant C_p, but our arguments are valid only for p near $\frac{1}{2}$. However we will show (Proposition 1) that for all $p > \frac{1}{2}$,

$$d_N\left((1-\varepsilon)\frac{\log N}{\log p^{-1}}\right) = 1 \tag{9}$$

for every $\varepsilon > 0$. Numerical analysis shows that for $p > 0.71$,

$$\frac{1}{\log p^{-1}} > \frac{3+\theta_p}{-4\log(p^2+q^2)}.$$

Therefore, if for every $p \in [0.5, 1)$ the cutoff phenomenon occurs at $C_p \log N$ for some constant C_p, and if our conjecture above is correct, then the function $p \mapsto C_p$ cannot be analytic in the parameter p.

Let $q = 1 - p$. There is an obvious "duality" between the p-shuffle and the q-shuffle: if one turns the deck upside down, then performs a p-shuffle, then turns the deck upside down again, the result is a q-shuffle. Thus, any result concerning the mixing rate of iterated p-shuffles holds also for q-shuffles. Henceforth, we shall assume that $1/2 \le p < 1$.

2. Preliminaries

2.1. *The associated dynamical system.* Imagine that the p-shuffle is performed infinitely many times, independently, on a deck of size N whose cards are initially labeled $1, 2, \dots, N$ from bottom to top. Each card i will have an *orbit*,

$$x^i = x_1^i, x_2^i, \dots,$$

the sequence of 0's and 1's recording which stack (top = 1, bottom = 0) the card visits at each step. Note that the orbits of distinct cards are different, with probability 1, because on each repetition of the shuffle the conditional probability that cards i and j will be put in different stacks (given the history of the deck up to that repetition) is bounded below by a positive number.

Lemma 1. *The lexicographic order on the set of orbits $\{x^1, x^2, \dots, x^N\}$ coincides with the original order of the cards in the deck.*

Proof. This follows from the nature of the riffling process. If card i is below card j originally, then it will stay below for as long as the orbits of the two cards coincide, because in each riffle cards are dropped one at a time from the bottoms of the two stacks. Thus, at the first time they are in *different* stacks, card i will be in the bottom stack (0) and card j will be in the top stack (1). Hence, $x^i \le x^j$ in the lexicographic order. □

Lemma 2. *The unordered collection of sequences $x^1, x^2, \dots, x^N$ is a random sample of N independent Bernoulli-p processes. Equivalently, a version of the set $\{x^{(i)}\}_{1\le i\le N}$ of orbits may be obtained by lex-ordering N independent Bernoulli-p sequences $\{x^i\}_{1\le i\le N}$.*

For the proof, see [6], Chapter 4D, for the case $p = \frac{1}{2}$; the general case is essentially the same.

The preceding lemmas imply that a version of the random walk on the permutation group $\mathscr{S}_N$ induced by independent repetitions of the p-shuffle may be (re)constructed from the realizations of N independent Bernoulli-p sequences $x^1, x^2, \dots, x^N$. The position of the card i with orbit $x^{(i)}$ in the deck after n shuffles is determined as follows: apply the shift σ n times to each of the sequences x^j, and then determine the relative order of $\sigma^n x^{(i)}$ in the set $\{\sigma^n x^1, \dots, \sigma^n x^N\}$.

2.2. *Example.* In this example, the deck has $N = 5$ cards labelled a, b, c, d, e. The following table shows the arrangement of cards in the deck (from top to bottom) after $n = 0, 1, 2, 3, 4$ repetitions of the shuffle:

Position	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$
1	e	e	d	d	d
2	d	b	c	a	a
3	c	d	e	c	c
4	b	c	a	b	b
5	a	a	b	e	e.

The next table shows the first seven entries of each orbit:

e	0	0	0	1	0	1	1
d	0	1	0	0	0	0	0
c	0	1	0	1	0	1	0
b	1	0	1	1	0	0	1
a	1	1	1	0	0	0	1.

Observe that the fifth entry of each orbit is 0, indicating that each card is in the “bottom” stack of the fourth riffle. Since the “top” stack is empty, the relative order of the cards is unchanged by the fourth riffle; that is, the fourth riffle permutation is the identity. Note that the same would be true if all the cards were in the “top” stack, that is, if the fifth entries of the orbits were all 1 instead of 0. This illustrates that the sequence of permutations does not uniquely determine the set of orbits.

2.3. *The associated random graph.* The state of the deck after n repetitions of the p-shuffle may be described in terms of a random graph $\mathscr{G} = \mathscr{G}_n^N$ and an independent random permutation Π (uniformly distributed on the permutation group $\mathscr{S}_N$). This representation will facilitate comparison with the uniform distribution.

Let the orbits of the cards be $x^{(i)}$, $1 \le i \le N$, as described in Section 2.1 above. Assume that the cards are labeled so that the original composition of the deck is $1, 2, \dots, N$, from bottom to top; thus, in the lexicographic order,

$$x^{(1)} \le x^{(2)} \le \cdots \le x^{(N)}.$$

Say that cards i, j have the same n-orbit if the orbits $x^{(i)}$ and $x^{(j)}$ coincide in their first n entries. Note that if i and j have the same n-orbit and $i < j$ then for every $i < i' < j$ the cards i and i' have the same n-orbit. Define $\mathscr{G} = \mathscr{G}_n^N$ to be the (random) graph with vertex set $[N] = \{1, 2, \dots, N\}$ and edge set $\mathscr{E}$ consisting of those pairs $(i, i+1)$ such that cards i and $i+1$ have the same n-orbit. Observe that $\mathscr{G}$ is completely determined by the set of n-orbits, that is, by the $N \times n$ Bernoulli Matrix,

$$M = M^{(N,n)} = \begin{pmatrix} x_1^1 & x_2^1 & \cdots & x_n^1 \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ & & \cdots & \\ x_1^N & x_2^N & \cdots & x_n^N, \end{pmatrix}$$

and consequently is independent of the (unordered) set of n-shifted orbits,

$$\{\sigma^n x^{(i)}\colon i \in [N]\}.$$

EXAMPLE. In the example described in Section 2.2 above, the graph $\mathscr{G}_4^5$ has no edges, because the 4-orbits of the five cards are distinct. However, the graph $\mathscr{G}_3^5$ has an edge connecting vertices 2 and 3, because cards c and d have the same 3-orbit.

For a permutation π of $[N]$ define the *$\mathscr{G}$-modification* $\pi_{\mathscr{G}}$ of π to be the permutation obtained by ordering the assignments within each clique of $\mathscr{G}$. (A *clique* of $\mathscr{G}$ is a maximal connected set of vertices.) Thus, for each singleton clique i (i.e., for each card i with a unique n-orbit),

$$\pi_{\mathscr{G}}(i) = \pi(i);$$

for each doubleton clique $i, i+1$,

$$\pi_{\mathscr{G}}(i) = \pi(i) \quad \text{and} \quad \pi_{\mathscr{G}}(i+1) = \pi(i+1) \quad \text{if } \pi(i) < \pi(i+1),$$

but

$$\pi_{\mathscr{G}}(i) = \pi(i+1) \quad \text{and} \quad \pi_{\mathscr{G}}(i+1) = \pi(i) \quad \text{if } \pi(i) > \pi(i+1)$$

and so on. Observe that, on the event that $\mathscr{G}$ has no edges, $\pi_{\mathscr{G}} = \pi$.

LEMMA 3. *Let $X_n = X_n^N$ be the permutation representing the state of the deck after n independent repetitions of the p-shuffle.* (*The ith entry $X_n(i)$ is the position of card i in the deck after n repetitions of the shuffle.*) *Let $\mathscr{G}$ be the random graph associated to an $N \times n$ Bernoulli matrix M, and let Π be a random permutation independent of M with the uniform distribution on $\mathscr{S}_N$. Then the $\mathscr{G}$-modification of Π is a version of X_n.*

PROOF. A version of the random walk $\{X_k\}_{k\geq 1}$ may be constructed from N independent Bernoulli-p processes $x^1, x^2, \dots, x^n$ by the procedure described in Section 2.1 above. Write

$$\begin{aligned} x^i &= z_1^i z_2^i \cdots z_n^i y_1^i y_2^i \cdots, \\ z^i &= z_1^i z_2^i \cdots z_n^i, \\ y^i &= y_1^i y_2^i \cdots. \end{aligned}$$

The orbits of the cards $i = 1, 2, \dots, N$ are obtained by lex-ordering the N sequences x^i. This may be done in two steps: first, rearrange the rows by ordering the finite sequences z^i; this may result in some ties. Second, rearrange the rows by ordering the infinite sequences y^i within any ties left by the first step. Since the rearrangement in the first step depends only on the values of the finite sequences z^i, the ordering of the infinite sequences y^i after the first stage is completely random. In particular, the assignment of row number i to the relative rank of the y-sequence in row i after the first step is a random permutation Π independent of the finite sequences z^i and uniformly distributed on the set $\mathscr{S}_N$. The permutation representing the deck after n P-shuffles is, clearly, obtained from Π by $\mathscr{G}$-modification, where $\mathscr{G}$ is the random graph obtained from the Bernoulli matrix z. □

COROLLARY 1. $d_N(n) \leq P\{\Pi \neq \Pi_{\mathscr{G}}\}$.

PROOF. Recall that $d_N(n)$ is the total variation distance between the uniform distribution and the distribution of the deck after n repetitions of the p-shuffle. Since Π has the uniform distribution and $\Pi_{\mathscr{G}}$ has the same distribution as the random permutation resulting from n repetitions of the p-shuffle, by the preceding lemma, it follows that

$$\begin{aligned} d_N(n) &= \tfrac{1}{2} \sum_{\pi \in \mathscr{S}_N} \left| P\{\Pi_{\mathscr{G}} = \pi\} - P\{\Pi = \pi\} \right| \\ &= \tfrac{1}{2} \sum_{\pi \in \mathscr{S}_N} \left| P\{\Pi_{\mathscr{G}} = \pi \text{ and } \Pi \neq \Pi_{\mathscr{G}}\} - P\{\Pi = \pi \text{ and } \Pi \neq \Pi_{\mathscr{G}}\} \right| \\ &\leq \tfrac{1}{2} \sum_{\pi \in \mathscr{S}_N} P\{\Pi_{\mathscr{G}} = \pi \text{ and } \Pi \neq \Pi_{\mathscr{G}}\} + \tfrac{1}{2} \sum_{\pi \in \mathscr{S}_N} P\{\Pi = \pi \text{ and } \Pi \neq \Pi_{\mathscr{G}}\} \\ &= P\{\Pi \neq \Pi_{\mathscr{G}}\}. \qquad \Box \end{aligned}$$

Define

$$\tau = \tau_N := \min\{n : \text{ no two rows of } M^{(N,n)} \text{ are the same}\}. \tag{10}$$

Since the rows of $M^{(N,\infty)}$ are independent sequences of i.i.d. Bernoulli-p random variables, τ_N is almost surely finite. Clearly, if $n \geq \tau_N$ then the random graph $\mathscr{G} = \mathscr{G}_n^N$ has no edges, and so $\Pi_{\mathscr{G}} = \Pi$. This implies the corollary.

COROLLARY 2. $d_N(n) \leq P\{\tau_N > N\}$.

Observe that τ_N is a *strong stationary time* in the sense of [2]. It is likely that the bounds on $d_N(n)$ obtained via Corollary 2 are loose.

COROLLARY 3. $\lim_{N\to\infty} d_N((1+\varepsilon)C_p^* \log N) = 0$.

PROOF. The entries of the matrix $M^{(N,n)}$ are i.i.d. Bernoulli-p random variables; thus, the probability that any two given rows are identical is $(p^2+q^2)^n$. It follows that the *expected* number of pairs of identical rows is

$$\binom{N}{2}\left(p^2+q^2\right)^n < N^2\left(p^2+q^2\right)^n.$$

If $n \geq (1+\varepsilon)C_p^* \log N$, then $(p^2+q^2)^n \leq N^{-2(1+\varepsilon)}$. Consequently, by the Chebyshev–Markov inequality, for $n = \lceil (1+\varepsilon)C_p^* \log N \rceil$,

$$P\{\tau_N > N\} = P\big\{M^{(N,n)} \text{ has two identical rows}\big\} \leq N^{-2\varepsilon} \longrightarrow 0$$

as $N \to \infty$. □

2.4. *Entropy and information.* The construction of the preceding section shows that the distribution of the permutation X_n representing the state of the deck after n p-shuffles is completely determined by the distribution of the random graph $\mathscr{G}$. When $p \neq \frac{1}{2}$ this distribution may be rather complicated; for instance, when $p > \frac{1}{2}$ cliques are more likely in certain “cold spots” of the deck where n-orbits are more likely to contain more than the average number np of 1’s. This is because when $p \neq \frac{1}{2}$, not all n-orbits are equally likely, and, in fact, the likelihoods of different n-orbits may be on entirely different orders of magnitude. It will ultimately be necessary to estimate the number of different n-orbits with likelihood on a given order of magnitude. These estimates will involve *entropy* and *information* numbers, about which we gather some basic information here.

For an n-orbit $x = x_1x_2\cdots x_n$ (a 0–1 sequence of length n) let $\bar{x} = \sum x_i/n$ be the relative frequency of 1’s. For $\alpha \in (0,1)$, $\varepsilon > 0$ and $k \leq n$ define

$$\Omega_n = \{0,1\}^n$$

and

$$\Omega_n^\varepsilon(\alpha) = \{x \in \Omega_n : |\bar{x} - \alpha| < \varepsilon\}.$$

Define the *Shannon entropy function* $H(\beta)$ and the *Kullback–Leibler information function* $I(p,\beta)$ by

$$H(\beta) = -\beta\log\beta - (1-\beta)\log(1-\beta) \quad \text{for } 0 < \beta < 1,$$

$$I(p,\beta) = -\beta\log p - (1-\beta)\log q \quad \text{for } 0 < \beta,\ p < 1.$$

The following well-known lemma is an easy consequence of Stirling's formula.

LEMMA 4. *For each $\alpha \in (0,1)$ and each $\varepsilon > 0$ small enough that $0 < \alpha - \varepsilon < \alpha + \varepsilon < 1$,*

$$\lim_{n\to\infty} \frac{1}{n}\log\left|\Omega_n^\varepsilon(\alpha)\right| = \max_{|\beta-\alpha|\le\varepsilon} H(\beta).$$

Fix $p \in (\frac{1}{2}, 1)$. For each sequence $x = x_1x_2\cdots x_n \in \Omega_n$, define the *likelihood*

$$\lambda(x) = \prod_{i=1}^{n}(p^{x_i}q^{1-x_i}).$$

COROLLARY 4. *For each $\alpha \in (0,1)$ and $\varepsilon > 0$, and for every fixed integer $k \ge 1$,*

$$\lim_{n\to\infty} \frac{1}{n}\log\left(\sum_{x\in\Omega_n^\varepsilon(\alpha)} \lambda(x)^k\right) = \max_{|\beta-\alpha|\le\varepsilon}(H(\beta) - kI(p,\beta)).$$

This is a routine consequence of Lemma 4 and the continuity of the functions H and I in their arguments.

Fix $p \in (\frac{1}{2}, 1)$, and for $t \ge 0$ define

$$\psi(t) = \log(p^t + q^t), \tag{11}$$

$$p_t = p^t/(p^t + q^t), \tag{12}$$

$$q_t = q^t/(p^t + q^t) = 1 - p_t. \tag{13}$$

Since $p > \frac{1}{2}$, the function $t \mapsto p_t$ is strictly increasing in t, with $p_0 = \frac{1}{2}$ and $p_\infty = 1$. Moreover, $\psi(t)$ is strictly decreasing in t, with $\psi(0) = \log 2$ and $\psi(\infty) = -\infty$. Consequently, since $\psi(2) < 0$, there is a unique $\theta > 0$ such that $\psi(\theta) = 2\psi(2)$; this is the value singled out in the statement of Theorem 1. Notice that for every $t > 0$,

$$\psi'(t) = p_t\log p + q_t\log q = -I(p, p_t), \tag{14}$$

where $I(\cdot,\cdot)$ is the Kullback–Leibler information function. Since p_t is increasing in t, $I(p, p_t)$ is decreasing in t, and therefore $\psi(t)$ is strictly convex in t. The functions H, I and ψ are related by the identities

$$H(p_t) = tI(p, p_t) + \psi(t) = -t\psi'(t) + \psi(t), \tag{15}$$

which shows that $-H(p_t)$, considered as a function of $-I(p, p_t)$, is the Legendre transform of $\psi(t)$.

3. Proof of Theorem 1. We have already proved the easy inequality (7) in Corollary 3 above. It remains to prove the hard inequality (6). Fix $p > \frac{1}{2}$ and $C < C_p$, and let

$$n = \lfloor C \log N \rfloor \tag{16}$$

be the number of repetitions of the p-shuffle. As usual, N is the cardinality of the deck. In this section we will show that if $p - \frac{1}{2}$ is sufficiently small then

$$\lim_{N \to \infty} d_N(n) = 1. \tag{17}$$

Since the total variation distance $d_N(n)$ is, for fixed N, nonincreasing in n, we may assume in proving (17) that C is arbitrarily close to C_p. Recall that $C_{1/2} = 3/(2 \log 2)$, so for p near $\frac{1}{2}$, C_p is close to $3/(2 \log 2)$. Consequently, if $\beta > 0$ is small then for p near $\frac{1}{2}$ and C near C_p,

$$p^n < N^{-1-\beta}. \tag{18}$$

3.1. *Clumping.* There are (at least) two obstructions to mixing for p-shuffles, *clumping* and *cold spots*. The term *clumping* refers to a tendency for a large numbers of cards at the top or bottom of the deck to remain in their original order. This may occur in the p-shuffle for p near 0 or 1 because the deck is, with high probability, so unevenly divided that cards in one of the two stacks do not have a chance to be mixed with cards from the other. The term *cold spots* refers to the existence of predictable areas of the deck in which unusually large numbers of neighboring *pairs* $i, i+1$ tend to remain in their original order. For p near $\frac{1}{2}$, the existence of cold spots is the obstruction that persists for large values of n, while for $p > 0.71$ clumping is the persistent obstruction. The following lemma shows that clumping will occur when $C < 1/\log p^{-1}$.

PROPOSITION 1. *For any* $p \in [\frac{1}{2}, 1)$, *if* $C < 1/\log p^{-1}$ *then* $\lim_{N \to \infty} d_N(n) = 1$.

PROOF. If $C < 1/\log p^{-1}$, then there exists $\varepsilon > 0$ such that $p^n \geq N^{-1+\varepsilon}$ for all N. But p^n is the probability that a randomly chosen card will have n-orbit $111 \cdots 1$. Since the N orbits are independent, it follows that, with probability approaching 1 as $N \to \infty$, the number of cards with n-orbit $111 \cdots 1$ is at least $N^{\varepsilon/2}$. On this event, the $\lfloor N^{\varepsilon/2} \rfloor$ topmost cards of the deck remain in their original order after n shuffles. But for a completely random (that is uniformly distributed) permutation, the length of the longest run of consecutive cards that remain in their original order is, with probability approaching 1 as $N \to \infty$, smaller than $(\log N)^2$. (To see this, observe that the probability that $2k$ consecutive cards remain in order is smaller than 2^{-k}, because, in a completely random permutation, the probability that two neighboring cards remain in order is 1/2, and these events are independent for nonoverlapping pairs.) □

3.2. *Cold spots.* The existence or nonexistence of "cold spots" in the deck depends on the statistics of the random graph $\mathscr{G}$ associated to the set of n-orbits (see Section 2.3). Let $\mathscr{E}$ be the edge set of $\mathscr{G}$; define

$$\mathscr{K} = \{i \in [N]\colon (i, i+1) \in \mathscr{E}\}.$$

For a subset $H \subset [N]$ let ∂H be the (inner) *boundary* of H in $[N]$, that is, the subset of H consisting of all points at distance 1 from $[N] - H$.

PROPOSITION 2. *Suppose that there exist nonrandom subsets $H = H_N$ of $[N]$ such that for some $\varepsilon > 0$ all three of the following conditions are met as $N \to \infty$:*

$$|H| \longrightarrow \infty; \tag{19}$$

$$|\partial H| = O(|H|^{1/2}); \tag{20}$$

and

$$P\{|\mathscr{K} \cap H| \geq |H|^{1/2+\varepsilon}\} \to 1. \tag{21}$$

Then

$$\lim_{N\to\infty} d_N(n) = 1.$$

Note. The sets $H = H_N$ are the "cold spots." Aldous [1] states the special case $p = \frac{1}{2}$ and $H = [N]$, but his proof appears to have serious gaps.

To prove Proposition 2, we will show that under the hypotheses (19), (20) and (21), a recognizably large number of pairs of neighboring cards in H_N remain in their original order after n repetitions of the shuffle. For any permutation π and an index $i \in [N-1]$ define $\zeta_i = \zeta_i(\pi)$ by

$$\begin{aligned} \zeta_i &= 1 \quad \text{if } \pi(i) < \pi(i+1), \\ \zeta_i &= 0 \quad \text{if } \pi(i) > \pi(i+1). \end{aligned} \tag{22}$$

The number of *alignments* in H is defined to be $\sum_{i\in H} \zeta_i$. Let Π be a completely random permutation, and let $\Pi_{\mathscr{G}}$ be the $\mathscr{G}$-modification of Π. Observe that the random variables $\zeta_i(\Pi)$ are Bernoulli-$\frac{1}{2}$, so the expected number of alignments in H is $|H|/2$.

LEMMA 5. *For any interval $J = \{a, a+1, \dots, b\} \subset [N-1]$,*

$$E\left(\sum_{i\in J} \zeta_i(\Pi_{\mathscr{G}}) \middle| \mathscr{G}\right) \geq \tfrac{1}{2}|J| + \tfrac{1}{6}|\mathscr{K} \cap J| - 3.$$

PROOF. For a fixed graph $\mathscr{G} = ([N], \mathscr{E})$ whose vertex set is $[N]$ and whose edge set $\mathscr{E}$ is contained in $\{(i, i+1)\colon 1 \leq i \leq N-1\}$, define a sequence of subgraphs $\mathscr{G}_k = ([N], \mathscr{E}_k)$ by taking for the edges set $\mathscr{E}_k$ the set of all edges $(i, i+1) \in \mathscr{E}$ such that $i \leq k$. The edges sets of these graphs are nested, with $\mathscr{G}_0 = ([N], \varnothing)$ and $\mathscr{G}_N = \mathscr{G}$, and $\mathscr{G}_{k+1}$ is obtained from $\mathscr{G}_k$ by adjoining *at most* the single edge $(k+1, k+2)$ to the edge set of $\mathscr{G}_k$.

A completely random permutation Π may be constructed by attaching independent, identically distributed uniform-(0,1) random variables U_i to the indices $i = 1, 2, \ldots, N$ and defining $\Pi(i)$ to be the relative rank of U_i in the set $\{U_j\}_{1 \leq j \leq N}$. The $\mathscr{G}_k$-modification of Π is obtained by first permuting the r.v.'s U_i within $\mathscr{G}_k$-cliques so that the U's within cliques are ordered, then setting $\Pi_{\mathscr{G}_k}(i)$to be the relative rank of the U attached to the index i. Consider first the graph $\mathscr{G}_0$. Since its edge set is empty, the $\mathscr{G}_0$-modification of Π is identical to Π; consequently.

$$E\left(\sum_{i \in J} \zeta_i(\Pi_{\mathscr{G}_0}) \middle| \mathscr{G}_0\right) = \tfrac{1}{2}|J|.$$

Now consider the change in $\zeta_i(\Pi_{\mathscr{G}_k})$ when k is incremented to $k+1$. For those k such that $k+2 < a$ (a is the leftmost index in J), none of the values ζ_i, $i \in J$, is affected, since adding (at most) the single edge $(k+1, k+2)$ does not change the U attached to any $i \in J$. Hence, if $k+2 < a$, then

$$E\left(\sum_{i \in J} \zeta_i(\Pi_{\mathscr{G}_{k+1}}) \middle| \mathscr{G}_{k+1}\right) = E\left(\sum_{i \in J} \zeta_i(\Pi_{\mathscr{G}_k}) \middle| \mathscr{G}_k\right).$$

Next, consider $k \geq b$ (b is the rightmost index in J), Adding the edge $(k+1, k+2)$ changes only the rightmost clique; consequently, the only value ζ_i, $i \in J$, that may be affected is at the i immediately to the left of the rightmost clique in J. Hence,

$$\left| E\left(\sum_{i \in J} \zeta_i(\Pi_{\mathscr{G}_N}) \middle| \mathscr{G}_N\right) - E\left(\sum_{i \in J} \zeta_i(\Pi_{\mathscr{G}_{b-1}}) \middle| \mathscr{G}_{b-1}\right) \right| \leq 1.$$

Similarly, if $k = a-2$ or $k = a-1$, the only values of ζ_i, $i \in J$, that may be affected are ζ_a and ζ_{a+1}, and so

$$\left| E\left(\sum_{i \in J} \zeta_i(\Pi_{\mathscr{G}_{a-1}}) \middle| \mathscr{G}_{a-1}\right) - E\left(\sum_{i \in J} \zeta_i(\Pi_{\mathscr{G}_0}) \middle| \mathscr{G}_0\right) \right| \leq 2.$$

Finally, consider the changes in the values ζ_i when k is incremented to $k+1$ for $a \leq k < b$. If the edge $(k+1, k+2)$ is *not* added, then obviously nothing is changed. If $(k+1, k+2)$ *is* added, then at most three values ζ_i are affected: ζ_{k+1}, ζ_{k+2} and ζ_i for the index i immediately to the left of the clique containing the vertex $k+1$. Let c be the size of the $\mathscr{G}_K$-clique containing $k+1$, and let c' be the size of the $\mathscr{G}_K$-clique containing i. Conditional on $\mathscr{G}_K$ and $\mathscr{G}_{K+1}$, respectively,

$$\begin{aligned}
\zeta_i(\Pi_{\mathscr{G}_k}) &\sim \text{Bernoulli-}\bigl(1/\tbinom{c+c'}{c}\bigr), & \zeta_i(\Pi_{\mathscr{G}_{k+1}}) &\sim \text{Bernoulli-}\bigl(1/\tbinom{c+c'+1}{c+1}\bigr),\\
\zeta_{k+1}(\Pi_{\mathscr{G}_k}) &\sim \text{Bernoulli-}(1/(c+1)), & \zeta_{k+1}(\Pi_{\mathscr{G}_{k+1}}) &\sim \text{Bernoulli-}1,\\
\zeta_{k+2}(\Pi_{\mathscr{G}_k}) &\sim \text{Bernoulli-}(1/2), & \zeta_{k+2}(\Pi_{\mathscr{G}_{k+1}}) &\sim \text{Bernoulli-}(1/(c+2)).
\end{aligned}$$

Thus, the net change in (conditional) expectation is

$$\begin{aligned}
&E\left(\sum_{i\in J}\zeta_i(\Pi_{\mathscr{G}_{k+1}})\middle|\mathscr{G}_{k+1}\right) - E\left(\sum_{i\in J}\zeta_i(\Pi_{\mathscr{G}_k})\middle|\mathscr{G}_k\right)\\
&\qquad = 1-\frac{1}{2}+\frac{1}{c+2}-\frac{1}{c+1}+\frac{1}{\binom{c+c'+1}{c+1}}-\frac{1}{\binom{c+c'}{c}} \qquad (23)\\
&\qquad \ge \frac{1}{2}-\frac{1}{6}-\frac{1}{6}=\frac{1}{6}.
\end{aligned}$$

Consequently, the conditional expectation is incremented by at least 1/6 for every edge $(k+1, k+2)$ of $\mathscr{G}$ in J, so

$$E\left(\sum_{i\in J}\zeta_i(\Pi_{\mathscr{G}_{b-1}})\middle|\mathscr{G}_{b-1}\right) - E\left(\sum_{i\in J}\zeta_i(\Pi_{\mathscr{G}_{a-1}})\middle|\mathscr{G}_{a-1}\right) \ge \frac{1}{6}|J\cap\mathscr{K}|. \qquad \square$$

LEMMA 6. *For any set $H \subset [N-1]$,*

$$\operatorname{var}\left(\sum_{i\in H}\zeta_i(\Pi_{\mathscr{G}})\middle|\mathscr{G}\right) \le 3|H|.$$

PROOF. Let $\Gamma_1, \Gamma_2, \dots, \Gamma_m$ be the cliques of $\mathscr{G}$, listed in order from left to right. Each index i is in a unique clique Γ_j, so

$$\sum_{i\in H}\zeta_i = \sum_{j=1}^{m} Z_j,$$

where

$$Z_j = \sum_{i\in H\cap\Gamma_j}\zeta_i.$$

Within any clique, all ζ_i must be 1 except for the rightmost index i, so $Z_j - |H\cap\Gamma_j| + 1$ is a Bernoulli random variable. Furthermore, the values of ζ_i's in different cliques are independent unless the cliques are neighbors, so Z_j and Z_k are uncorrelated unless $|k-j| \le 1$. It follows that the conditional variance of $\sum Z_j$ is no larger than three times the number of cliques that intersect H. $\square$

Proof of Proposition 2. We will compare the distributions of $\sum_{i\in H}\zeta_i(\Pi)$ and $\sum_{i\in H}\zeta_i(\Pi_{\mathscr{G}})$ for a completely random permutation Π independent of $\mathscr{G}$. Recall that the random variables $\zeta_i(\Pi)$ are Bernoulli-$\frac{1}{2}$, and that $\zeta_i(\Pi)$ and $\zeta_j(\Pi)$ are independent unless $|i-j|\le 1$. Consequently, $\sum_{i\in H}\zeta_i(\Pi)$ has expectation $|H|/2$ and variance $O(|H|)$, and hence by Chebyshev's inequality,

$$P\left\{\sum_{i\in H}\zeta_i(\Pi)\ge |H|/2+|H|^{1/2+\varepsilon/4}\right\}\longrightarrow 0$$

as $N\to\infty$.

Now consider the distribution of $\sum_H \zeta_i(\Pi_{\mathscr{G}})$. The set H may be partitioned into nonoverlapping intervals $J_1, J_2, \dots, J_m$; by hypothesis (20), $m = O(|H|^{1/2})$. Consequently, by hypotheses (19) and (21) and Lemmas 5 and 6, with probability approaching 1 as $N\to\infty$,

$$E\left(\sum_{i\in H}\zeta_i(\Pi_{\mathscr{G}})\,\middle|\,\mathscr{G}\right)\ge \frac{1}{2}|H|+|H|^{1/2+\varepsilon/2}$$

and

$$\operatorname{var}\left(\sum_{i\in H}\zeta_i(\Pi_{\mathscr{G}})\,\middle|\,\mathscr{G}\right)=O(|H|).$$

Thus, by Chebyshev's inequality,

$$P\left\{\sum_{i\in H}\zeta_i(\Pi_{\mathscr{G}})\le |H|/2+|H|^{1/2+\varepsilon/4}\right\}\longrightarrow 0$$

as $N\to\infty$. This proves that the total variation distance between the distributions of $\sum_{i\in H}\zeta_i(\Pi_{\mathscr{G}})$ and $\sum_{i\in H}\zeta_i(\Pi)$ converges to 1 as $N\to\infty$, so $d_N(n)\to 1$. □

3.3. *Existence of cold spots for $C < C_p$.* Proposition 2 implies the convergence (17) holds provided there are predictable "cold spots" $H = H_N$ of the deck satisfying relations (19)–(21) as $N\to\infty$. In this section we will show that such cold spots exist for p near $\frac{1}{2}$ and $n=\lfloor C\log N\rfloor$ for a constant $C\in(0,C_p)$ with C close to C_p.

Define m to be the (least) integer nearest to the solution of

$$\left(p^{p_\theta}q^{q_\theta}\right)^m = N^{-1/2+\delta}$$

for a small constant $\delta>0$ to be specified later. Here $\theta=\theta_p$ where θ_p is defined by (5). Observe that m is the (least) integer nearest to the solution m' of

$$m' I(p,p_\theta)=\left(\tfrac{1}{2}-\delta\right)\log N. \tag{24}$$

Observe that $m<n$, because $p>\frac{1}{2}$ and $p^n<1/N$, by relation (18). Loosely speaking, the cold spot will be those intervals of the deck in which cards have m-orbits with relative frequency of 1's approximately p_θ. Unfortunately, the precise definition of $H = H_N$ must be somewhat more complicated, because Proposition 2 requires that the sets H be nonrandom.

For each sequence $x = x_1 x_2 \cdots x_m \in \Omega_m$, set $t_x = P\{\xi_1 \xi_2 \cdots \xi_m \le x_1 x_2 \cdots x_m\}$, where $\xi_1, \xi_2, \ldots$ are i.i.d. Bernoulli-p and $\le$ is the lexicographic order. The numbers t_x, for $x \in \Omega_m$, partition the unit interval into subintervals $J_x = [t_x, t_x + \lambda(x))$ of varying lengths $\lambda(x)$. These intervals are indexed so that if U is uniformly distributed on [0,1] and $x(U)$ is the index of the interval J_x containing U then the entries x_i of $x(U)$ are i.i.d. Bernoulli-p. Consequently, the length $\lambda(x)$ of the interval J_x coincides with likelihood $\lambda(x) = \Pi_{i=1}^m p^{x_i} q^{1-x_i}$. Fix $\varepsilon > 0$ (small), and define

$$H = H_N = \left\{ i \in [N] \colon i/N \in \bigcup_{x \in \Omega_m^\varepsilon} J_x \right\},$$

where

$$\Omega_m^\varepsilon = \Omega_m^\varepsilon(p_\theta)$$

is the set of all m-orbits x for which the relative frequency of 1's is within ε of p_θ. Note that the definition involves the as-yet unspecified constants $\delta > 0$ and $\varepsilon > 0$.

The next two propositions state that for suitably chosen constants $\delta, \varepsilon > 0$, the hypotheses of Proposition 2 are satisfied provided that p is sufficiently near $\frac{1}{2}$ and C is close to C_p. Consequently, their proofs will complete the proof of the assertion (17).

PROPOSITION 3. *For sufficiently small* $\varepsilon > 0, |H| \to \infty$ *and* $|\partial H| = O(|H|^{1/2})$ *as* $N \to \infty$.

PROOF. The set H consists of all integers contained in the nonoverlapping intervals NJ_x, where $x \in \Omega_m^\varepsilon$. Hence, the boundary ∂H has cardinality no larger than $2|\Omega_m^\varepsilon|$. For each $x \in \Omega_m^\varepsilon$, the relative frequency of 1's in x is within ε of p_θ, so the likelihood $\lambda(x)$ (which is also the length of J_x) satisfies

$$\lambda(x) \ge p^{mp_\theta - m\varepsilon} q^{mq_\theta + m\varepsilon} \asymp N^{-1/2+\delta}(p/q)^{-m\varepsilon}.$$

Consequently, the length of each interval NJ_x contained in H is at least a positive multiple (independent of N) of $N^{1/2+\delta}(p/q)^{-m\varepsilon}$. For sufficiently small $\varepsilon > 0$, this is at least $N^{(1+\delta)/2}$. Thus, $|H| \to \infty$ as $n \to \infty$. Since $N \ge |H| \ge |\Omega_m^\varepsilon| N^{(1+\delta)/2}$, the cardinality of Ω_m^ε cannot be larger than $N^{(1-\delta)/2}$; it follows that the cardinality of ∂H cannot be larger than $2N^{(1-\delta)/2}$. Since each interval NJ_x contained in H has length at least $N^{(1+\delta)/2}$, this implies that $|\partial H| = o(|H|^{1/2})$. □

PROPOSITION 4. *Assume that* $p - \frac{1}{2}$ *is small. For sufficiently small* $\delta, \varepsilon > 0$, *there exists* $\alpha > 0$ *such that*

$$\lim_{N \to \infty} P\{|\mathscr{K} \cap H| \ge |H|^{1/2+\alpha}\} = 1. \tag{25}$$

The remainder of this section is devoted to the proof of this proposition. Since $\mathscr{K} \cap H$ is difficult to deal with directly, we will begin by finding a random set H', defined solely in terms of orbit properties, for which the size of the intersection $\mathscr{K} \cap H'$ is more easily estimated, and such that $H' \subset H$ with probability approaching 1 as $N \to \infty$. Fix $0 < \varepsilon' < \varepsilon$; define

$$H' = \{i\colon x(i) \in \Omega_m^{\varepsilon'}(p_\theta) \cap \Gamma\},$$

where $x(i)$ denotes the m-orbit of the ith card and Γ is the set of all finite 0–1 sequences whose last two entries are, in order, 01.

LEMMA 7. *Fix $\delta > 0$. For $0 < \varepsilon' < \varepsilon$ and $\varepsilon > 0$ sufficiently small,*

$$\lim_{N\to\infty} P\{H' \subset H\} = 1.$$

PROOF. Recall that the assignment of m-orbits to cards may be made as follows: start with N independent Bernoulli-p sequences, and attach labels $i = 1, 2, \dots, N$ to the sequences so that the natural order on the set of labels i coincides with the lexicographic order on the attached sequences; then discard all but the first m entries of each sequence. A stochastically equivalent way to make the assignment is as follows: drop N points independently from the uniform distribution on [0,1], and label the points $U_{(i)}$ in accordance with their relative order in [0,1] (thus, $U_{(1)}, U_{(2)}, \dots, U_{(N)}$ are the order statistics of a random sample of N uniforms). Then replace each $U_{(i)}$ by the (unique) sequence $x = x(i) \in \Omega_m$ such that $U_{(i)} \in J_x$. That these constructions are equivalent follows from the definition of the intervals J_x.

Let $x \in \Omega_m^{\varepsilon'}(p_\theta) \cap \Gamma$. Since $x \in \Gamma$, its last two entries are 01. Consequently, the sequences $x^-, x^+ \in \Omega_m$ immediately to the left and right of x in the lexicographic order are the sequences obtained from x by replacing the last two entries 01 by 00 and 10, respectively. These replacements change the relative frequency of 1's by at most $1/N$ which, for sufficiently large N, is less than $\varepsilon - \varepsilon'$; hence, $x^-, x^+ \in \Omega_m^{\varepsilon}$. Thus, if $i \in H'$, the intervals $NJ_{x(i)}$, $NJ_{x(i)^-}$ and $NJ_{x(i)^+}$ are all contained in H.

Let F_N be the empirical distribution function of the sample $U_1, U_2, \dots, U_N$. The construction of the sequences $x(i)$ forces the inequalities

$$F_N(t_{x(i)^-}) \le \frac{i}{N} \le F_N(t_{x(i)} + \lambda(x)).$$

Now let $i \in H'$. By the last paragraph, each of the intervals $NJ_{x(i)}$, $NJ_{x(i)^-}$ and $NJ_{x(i)^+}$ is contained in H. By the proof of Proposition 3, if $\varepsilon > 0$ is sufficiently small then each interval J_x such that $NJ_x \subset H$ has length at least $|J_x| \ge N^{(-1+\delta)/2}$. Thus, if $i \in H'$, *either* i/N is in one of the three intervals $J_{x(i)}, J_{x(i)^-}, J_{x(i)^+}$ and hence $i \in H$, *or*

$$\max_{t\in[0,1]} \sqrt{N}|F_N(t) - t| \ge N^{\delta/2}.$$

But the Kolmogorov–Smirnov theorem implies that

$$\lim_{N\to\infty} P\left\{\max_{t\in[0,1]} \sqrt{N}|F_N(t)-t| \geq N^{\delta/2}\right\} = 0. \qquad \square$$

LEMMA 8. *Define τ to be the number of* triplets, *that is, the number of indices $i \in [N]$ such that cards $i, i+1$, and $i+2$ all have the same n-orbit. Similarly, define κ to be the number of* quadruplets, *that is, the number of indices $i \in [N]$ such that cards $i, i+1, i+2$ and $i+3$ all have the same n-orbit. For each $\rho > 0$, if p is sufficiently near $\frac{1}{2}$ and if $C_p - C$ is sufficiently small, then as $N \to \infty$,*

$$E\tau = O(N^{\rho}) \tag{26}$$

and

$$E\kappa = O\big(N^{-1/2+\rho}\big) = o(1). \tag{27}$$

PROOF. The (unordered) set of n-orbits has the same distribution as a set of N independent, identically distributed Bernoulli-p sequences of length n. There are $\binom{n}{3}$ possibilities for triplets; for each, the probability that their n-orbits coincide is $(p^3+q^3)^n$. Similarly, there are $\binom{N}{4}$ possibilities for quadruplets; and for each possibility the probability that their n-orbits coincide is $(p^4+q^4)^n$. Consequently,

$$E\tau = \binom{N}{3}(p^3+q^3)^n \leq N^3 \exp\{\psi(3)C\log N\}$$

and

$$E\kappa = \binom{N}{4}(p^4+q^4)^n \leq N^4 \exp\{\psi(4)C\log N\}.$$

As $p \to \frac{1}{2}$,

$$\psi(3) \longrightarrow -2\log 2;$$
$$\psi(4) \longrightarrow -3\log 2;$$

and

$$C_p \longrightarrow 3/(2\log 2).$$

Consequently, if $p - \frac{1}{2}$ is small and $C_p - C$ is small, then relations (26) and (27) will hold as $N \to \infty$. □

LEMMA 9. *Assume that p is near $\frac{1}{2}$. Then for every $\delta' \in (0, \delta)$,*

$$\lim_{N\to\infty} P\{|\mathscr{K} \cap H'| \geq N^{1+2\delta'} \exp\{mH(p_\theta) + (n-m)\psi(2)\}\} = 1. \tag{28}$$

Note. Here $H(p_\theta)$ denotes the Shannon entropy function evaluated at p_θ.

PROOF. Recall that $i \in \mathscr{K}$ if and only if $(i, i+1) \in \mathscr{E}$, that is, if and only if the cards i and $i+1$ have the same n-orbit. In addition, $i \in H'$ if and only if its m-orbit is an element of $\Omega_m^{\varepsilon'} = \Omega_m^{\varepsilon'}(p_\theta)$. Consequently, the cardinality of $\mathscr{K} \cap H'$ is bounded below by the number of distinct n-orbits $x = x_1 x_2 \cdots x_n$ such that (i) at least two cards have n-orbit x and (ii) $x_1 x_2 \cdots x_m \in \Omega_m^{\varepsilon'}$. Thus, by the inclusion–exclusion formula,

$$|\mathscr{K} \cap H'| \geq \sum_{x \in \Omega_m^{\varepsilon'}} (\xi_2(x) - \xi_3(x) - \xi_4(x) - \cdots),$$

where $\xi_k(x)$ is the number of k-sets of cards with the same n-orbit y and $y_1 y_2 \cdots y_m = x$. Now

$$\sum_{x \in \Omega_m^{\varepsilon'}} \xi_3(x) \leq \tau \quad \text{and} \quad \sum_{x \in \Omega_m^{\varepsilon'}} \xi_k(x) \leq \kappa \qquad \forall\, k \geq 4,$$

where τ and κ are the number of triplets and quadruplets, respectively. By Lemma 8, for each $\rho > 0$ there exists $\beta = \beta_\rho > 0$ such that if $p - \frac{1}{2} < \beta$ and $C_p - C < \beta$ then $E\tau < N^\rho$ and $E\kappa \leq N^{-(1/2)+\rho}$ for all sufficiently large N. Hence, by the Markov inequality,

$$\lim_{N\to\infty} P\{\tau > N^{2\rho}\} = 0 \quad \text{and} \quad \lim_{N\to\infty} P\{\kappa \geq 1\} = 0.$$

We will prove below that $Y := \sum_{x\in\Omega_m^{\varepsilon'}} \xi_2(x)$ is, with probability approaching 1 as $N \to \infty$, much larger than $N^{2\rho}$. Therefore, to prove Lemma 9 it suffices to prove that (28) holds with $|\mathscr{K} \cap H'|$ replaced by Y.

We shall estimate the expectation and variance of the random variable Y by appealing to Corollary 4 of Section 2.4. First,

$$EY = E \sum_{x \in \Omega_m^{\varepsilon'}} \xi_2(x) = \binom{N}{2} (p^2 + q^2)^{n-m} \sum_{x \in \Omega_m^{\varepsilon'}} \lambda(x)^2.$$

By Corollary 4, there exist constants $\gamma = \gamma(\varepsilon')$ satisfying $\gamma \to 0$ as $\varepsilon' \to 0$ such that all sufficiently large N,

$$\sum_{x \in \Omega_m^{\varepsilon'}} \lambda(x)^2 \geq \exp\{mH(p_\theta) - 2mI(p, p_\theta) - m\gamma\};$$

hence, by (24), if $\varepsilon' > 0$ is sufficiently small then

$$\begin{aligned} EY &\geq \frac{1}{2}\left(1 - \frac{1}{N}\right) N^{1+2\delta} \exp\{mH(p_\theta) - m\gamma + (n-m)\psi(2)\} \\ &\geq N^{1+2\delta''} \exp\{mH(p_\theta) + (n-m)\psi(2)\} \end{aligned} \tag{29}$$

for some $0 < \delta' < \delta'' < \delta$. Observe that, if p is near $\frac{1}{2}$ and C is near C_p, this is at least $N^{(1/2)-\rho}$ for arbitrarily small $\rho > 0$, and so is of larger order of magnitude than $E\tau$.

The variance of Y is also easily estimated. Write

$$Y = \sum_{x \in \Omega_m^{\varepsilon'}} \xi_2(x) = \sum_{i,j \in [N]:\ i<j} Y_{ij},$$

where Y_{ij} is the indicator of the event that the n-orbits x^i and x^j coincide and their common truncations fall in $\Omega_m^{\varepsilon'}$ (here the n-orbits are taken to be randomly ordered). The random variables Y_{ij} are identically distributed Bernoullis. If i, j, i', j' are distinct indices then Y_{ij} and $Y_{i'j'}$ are independent; if i, i', j are distinct indices, Y_{ij} and $Y_{i'j}$ are not independent, but $\sum_{i<i'<j} Y_{ij}Y_{i'j} = \tau$, which has expectation $O(N^\rho)$. Consequently,

$$\text{var}(Y) \le E \sum_{x\in\Omega_m^{\varepsilon'}} \xi_2(x) + O(N^\rho). \tag{30}$$

The result now follows from relations (29) and (30), by Chebyshev's inequality. □

Proof of Proposition 4. Recall that m is the (least) integer nearest the solution m' of (24). Recall that H consists of the nonoverlapping intervals NJ_x, where $x \in \Omega_m^\varepsilon$. Since the length of J_x is $\lambda(x)$, Corollary 4 implies that, for some $\gamma = \gamma(\varepsilon) > 0$ satisfying $\gamma \to 0$ as $\varepsilon \to 0$,

$$|H| = N \sum_{x\in\Omega_m^\varepsilon} \lambda(x) \le Ne^{m'H(p_\theta)-m'I(p,p_\theta)+m'\gamma} = N^{-1/2+\delta}e^{m'H(p_\theta)+m'\gamma}.$$

Therefore, by Lemmas 7 and 9, it suffices to prove that for some $\alpha > 0$ and some $0 < \delta' < \delta$,

$$\left(N^{1/2+\delta}\exp\{mH(p_\theta)+m\gamma\}\right)^{1/2+\alpha} \le N^{1+2\delta'}\exp\{mH(p_\theta)+(n-m)\psi(2)\}.$$

Each side of this inequality may be expressed as N raised to a power, using equations (16) and (24) relating n, m, and N. It suffices to show that the power on the left side is less than the power on the right side. Since γ, δ and α may be taken arbitrarily small, they may be set equal to zero before comparing the powers. Thus, it suffices to prove that

$$1 + \frac{H(p_\theta)}{I(p, p_\theta)} < 4 + 2\frac{H(p_\theta)}{I(p, p_\theta)} + 2\left(2c - \frac{1}{I(p, p_\theta)}\right)\psi(2).$$

Now recall, by (15), that $H(p_\theta) = \theta I(p, p_\theta) + \psi(\theta)$; hence, the last inequality reduces to

$$3 + \theta + \frac{\psi(\theta)}{I(p, p_\theta)} + 2\psi(2)\left(2c - \frac{1}{I(p, p_\theta)}\right) > 0.$$

But $\psi(\theta) = 2\psi(2)$, so this is equivalent to

$$3 + \theta + 4\psi(2)C > 0.$$

Since $0 = 3+\theta+4\psi(2)C_p$, by definition of C_p, and since $C < C_p$ and $\psi(2) < 0$, the inequality $3 + \theta + 4\psi(2)C > 0$ is in fact true. □

4. A heuristic argument.

In this section we give a heuristic argument in favor of Conjecture 1. This argument is based on a rough analysis of the Radon–Nikodym derivative dQ/dP, where P and Q are, respectively, the uniform distribution and the distribution of the deck after n repetitions of the p-shuffle.

Recall (Lemma 3) that a version $\Pi_{\mathscr{G}}$ of the random permutation induced by n repetitions of the p-shuffle may be constructed as follows: start with a completely random (uniformly distributed) permutation Π. Build a random graph $\mathscr{G}$ from an $(N \times n)$ Bernoulli matrix X, independent of Π, by putting edges between neighboring indices $i, i+1$ such that the ith and $(i+1)$th largest (lexicographically) rows of X are the same. Then use the random graph $\mathscr{G}$ to modify Π by ordering assignments in $\mathscr{G}$-cliques. Let Q and P be the distributions of the random permutations $\Pi_{\mathscr{G}}$ and Π, respectively.

Lemma 8 above implies that if $p - 1/2$ and $C - C_p$ are small (here $n = \lfloor C \log N \rfloor$) then (i) the expected number of "triplets" (vertices i such that the graph $\mathscr{G}$ contains edges connecting i to $i+1$ and $i+1$ to $i+2$) is of order no more than N^{ρ}, where $\rho < 1/2$; and (ii) the expected number of "quadruplets" is $o(1)$. If $\mathscr{G}$ had neither triplets nor quadruplets, then for any permutation π, the only information about which of the distributions P, Q might have "generated" π in the values $\pi(i)$ and $\pi(i+1)$ would be in the value of $\zeta_i(\pi)$, the indicator of the event $\pi(i) < \pi(i+1)$. Under P, this random variable is Bernoulli-1/2, while under Q, it is Bernoulli-q_i. Here $q_i = (1/2)(1+\rho_i)$, and

$$\rho_i = \rho_i^{N,n} = P(K_i),$$

where $K_i = K_i^{N,n}$ is the event that the orbits of cards cards i and $i+1$ coincide up to time n. If the information in the random variables ζ_i were independent, then the Radon–Nikodym derivative (dQ/dP) would be the product of the likelihood ratios for the Bernoulli random variables ζ_i,

$$\frac{dQ}{dP}(\pi) = \prod_{i=1}^{N-1} (1+\rho_i)^{\zeta_i}(1-\rho_i)^{1-\zeta_i}. \tag{31}$$

If n were sufficiently large that all of the coincidence probabilities ρ_i were small, then (31) could be rewritten as

$$\frac{dQ}{dP}(\pi) = 1 + \sum_{i=1}^{N-1} \xi_i \rho_i + \cdots, \tag{32}$$

where $\xi_i = 2\zeta_i - 1$ and $\cdots$ indicates higher order terms. Now the random variables $\xi_1, \xi_3, \ldots$ are independent under P, as are the random variables $\xi_2, \xi_4, \ldots$, and each ξ_i has mean zero and variance 1 under P. Consequently, if $n = \lfloor C \log N \rfloor$ for some $C > C_p$, then Proposition 5 below would imply that, under P, as $N \to \infty$,

$$\frac{dQ}{dP} \xrightarrow{P} 1, \tag{33}$$

which would, in turn, imply the truth of Conjecture 1.

PROPOSITION 5. *Let $n = \lfloor C \log N \rfloor$. Then*

$$C > C_p \Longrightarrow \lim_{N\to\infty} \sum_{i=1}^{N} \rho_i^2 = 0; \tag{34}$$

$$C < C_p \Longrightarrow \lim_{N\to\infty} \sum_{i=1}^{N} \rho_i^2 = \infty. \tag{35}$$

We shall omit the proof of this result, since we have not been able to make the rest of the argument rigorous. It seems likely to us that, although the expression (31) for the Radon–Nikodym derivative dQ/dP is not exact, it is nevertheless close enough to the true value that the remainder of the argument remains valid.

Acknowledgments. Thanks to Persi Diaconis for his encouragement, and to Jim Fill for his incredibly thorough reading of the paper and numerous helpful suggestions.

REFERENCES

[1] ALDOUS, D. (1983). Random walks on finite groups and rapidly mixing Markov chains. *Seminar on Probability XVII. Lecture Notes in Math*. **986** 243–297. Springer, Berlin.

[2] ALDOUS, D. and DIACONIS, P. (1987). Strong uniform times and finite random walks. *Adv. in Appl. Math.* **8** 69–97.

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

[4] BIDIGARE, P., HANLON, P. and ROCKMORE, D. (1999). A combinatorial description of the spectrum of the Tsetlin library and its generalization to hyperplane arrangments. *Duke Math. J.* **99** 135–174.

[5] BROWN, K. and DIACONIS, P. (1998). Random walks and hyperplane arrangements. *Ann. Probab.* **26** 1813–1854.

[6] DIACONIS, P. (1998). *Group Representation in Probablity and Statistics*. IMS, Hayward, CA.

[7] DIACONIS, P., FILL, J. and PITMAN, J. (1992). Analysis of top to random shuffles. *Combin. Probab. Comput.* **1** 135–155.

[8] FULMAN, J. (1998). The combinatorics of biased riffle shuffles. *Combinatorica* **18** 173–184.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
ECKHART HALL
5734 UNIVERSITY AVENUE
CHICAGO, ILLINOIS 60637
E-MAIL: lalley@galton.uchicago.edu