*Journal of Theoretical Probability, Vol. 2, No. 1, 1989*

# Randomization Time for the Overhand Shuffle

**Robin Pemantle**[1]

*Received November 23, 1987; revised May 25, 1988*

This paper analyzes repeated shuffling of a deck of $N$ cards. The measure studied is a model for the popular *overhand shuffle* introduced by Aldous and Diaconis. It is shown that convergence to the uniform distribution requires at least order $N^2$ shuffles, and that order $N^2 \log(N)$ shuffles suffice. For a 52-card deck, more than 1000 shuffles are needed.

**KEY WORDS:** Shuffle; overhand shuffle; convolution; coupling time; uniform distribution; convergence.

## 1. THE OVERHAND SHUFFLE $\mathcal{O}_p$

The two most common methods of shuffling a deck of playing cards are the riffle shuffle and the overhand shuffle. The riffle shuffle is performed by splitting the deck into two halves and interlacing them. Aldous and Diaconis have shown in Ref. 1 that approximately $2 \log N$ riffle shuffles suffices to randomize a deck (the meaning of this will be made more precise below). In particular, about 7 riffle shuffles are sufficient to randomize a 52-card deck. This paper shows that the overhand shuffle takes order of $N^2$ or more shuffles to randomize a 52-card deck, explaining why it is only the second most popular shuffle.

To perform an overhand shuffle, hold the deck of cards in your right hand and slide a small packet of cards off the top of the deck into your left hand. Repeat this process, putting each successive packet on top of the previous one, until all the cards are in your left hand, so the order of the cards gets reversed in clumps.

The overhand shuffle is frequently used by casual card players. It

[1] M.I.T. 2-230, Cambridge, Massachusetts 02139. Present address: Department of Statistics, University of California, Berkeley, California, 94720.

should be noted here that a version of this shuffle called the "strip shuffle" is sometimes used in Las Vegas. In this version, the deck is flat on the table and successive packets are taken off the top of the deck and dropped on each other in reverse order.

Shuffling a deck of cards induces a permutation on it which is not predictable. Thus a single shuffle corresponds to a probability measure on the symmetric group $\mathbf{S}_N$ of permutations of an $N$-card deck. Following Aldous and Diaconis,[1] successive shuffles can be treated as independent, so repeating a shuffle corresponds to convolving the measure with itself.

When modeling the overhand shuffle, we shall find it convenient to look at the deck backwards after an odd number of shuffles. In this case, each shuffle reverses the order of the cards within any given packet but keeps the relative order of the packets fixed. If $\sigma$ is the order reversing involution, we will have $\sigma\mathcal{O}_p\sigma^{-1} = \mathcal{O}_p$ for all our distributions, so looking at the deck backwards every second shuffle will not change anything.

To choose an overhand shuffle at random, let $e_1, e_2,..., e_{N-1}$ be independent Boolean variables, each being "Yes" with probability $p$. For convenience, we shall let $q = 1 - p$ throughout this paper. Now choose the packets by separating the $i$th card from the $i+1$st exactly when $e_i =$ "Yes."

**Example.** Suppose a 13-card deck is initially in the order

A 2 3—4 5 6 7 8 9—T—J Q—K

Suppose $p = 1/4$ and $(e_1,..., e_{12}) = NNYNNNNNYYNY$. Then the packets are as shown and the new order is

3 2 A 9 8 7 6 5 4 T Q J K

Formally, $\mathcal{O}_p$ is the distribution such that $\mathcal{O}_p(\pi) = p^k q^{n-k-1}$ if $\pi$ is of the form $(r_1, r_1 - 1,..., 1, r_2,..., r_1 + 1,..., N, N-1,..., r_k + 1)$ and 0 otherwise. Note that if $p = q = 1/2$ then $\mathcal{O}_p$ is closest to uniform since every possible resulting permutation is equally likely. $\mathcal{O}_p$ can also be viewed as a Markov chain whose states are arrangements of the deck, where the next state is gotten by applying a random permutation from the prescribed measure.

## 2. THE MAIN RESULTS

Let $\mathbf{P}$ and $\mathbf{Q}$ be probability distributions on $\mathbf{S}_N$ and define

$$|\mathbf{P} - \mathbf{Q}| = 1/2 \sum_{\pi \in \mathbf{S}_N} |\mathbf{P}(\pi) - \mathbf{Q}(\pi)|$$

or equivalently

$$|\mathbf{P} - \mathbf{Q}| = \max\{\mathbf{P}(A) - \mathbf{Q}(A) : A \subseteq \mathbf{S}_N\}$$

Let $\mathbf{U}$ be the uniform distribution on $\mathbf{S}_N$. Let $\mathbf{P}^k$ denote the convolution of $\mathbf{P}$ with itself $k$ times. Then we have the following upper bound theore.

**Theorem 1.** (Upper bound theorem.) For any $p$ there is a $k$ such that for all $r, N \in \mathbf{Z}^+$

$$|\mathcal{O}_p^{kN^2[\log(N)+r\log(2)]} - \mathbf{U}| \leqslant (1/2)^{r-1}$$

Letting $r = \lceil 1 + \log_2(\varepsilon) \rceil$, we can think of this as saying that it takes at most $kN^2[\log(N) + r\log 2]$ shuffles to come within $\varepsilon$ of the uniform distribution. The asymptotic upper bound given by this theorem for fixed $p$ and $\varepsilon$ while $N \to \infty$ is on the order of $N^2 \log N$. If $N$ is fixed but $\varepsilon \to 0$, the bound is logarithmic in $\varepsilon$, which agrees with the Perron-Frobenius theory (see Ref. 1).

Note that the distance function we use is quite unforgiving. For example, let $\mathbf{E}$ be the distribution selecting any even permutation with uniform probability on the subset of even permutations. If you choose a permutation randomly from $\mathbf{E}$ and then look at the cards one by one, you will never have any clue as to the identity of the next card until you reach the penultimate card, at which point you will know the relative order of the last 2 cards with certainty. So you do not have much information, yet $|\mathbf{E} - \mathbf{U}| = 1/2$.

The proof of the upper bound theorem uses a coupling argument. If the reader wants a more detailed account of coupling than the one given in this paragraph, see Ref. 2. Define a probability distribution, $\mathscr{C}_p$, on pairs of shuffles of two decks having the property that $\mathscr{C}_p$ restricted to either deck is just $\mathcal{O}_p$. Formally, $\mathscr{C}_p$ is a probability distribution on $\mathbf{S}_N \times \mathbf{S}_N$ such that $\mathscr{C}_p$ restricted to either coordinate is distributed identically to $\mathcal{O}_p$. The joint distribution of $\mathscr{C}_p$ is rigged so that the two decks eventually become identically arranged and stay that way. If the second deck begins in a random arrangement, it is tempting but incorrect to believe that the first deck is random as soon as it agrees with the second deck. What is true, however, is that for any fixed $M$, if the first deck and the initially random second deck agree after $M$ shuffles with probability $1 - \varepsilon$, then $|\mathcal{O}_p^M - \mathbf{U}| \leqslant \varepsilon$.

Most of the work comes in determining $M$. One particularly long calculation is summarized as Lemma 1. Although the truth of Lemma 1 may seem intuitively clear, the author knows of no short or elegant proof. Since the argument we use relies on some calculations done in Section 4 we postpone the proof of Lemma 1 until Section 5. A lower bound is given in Section 4. Asymptotically, this bound is on the order of $N^2$ so it is lower than the upper bound by an asymptotic factor of $\log N$. In Section 6 we present some numerical determinations of the upper and lower bounds for a 52-card deck which show that the constant factors in the two bounds serve to make the bounds very far apart (35 versus 370,000,000,000) even for $N$ as small as 52. Some empirical evidence is presented indicating that the tightest numerical bound in the case $N = 52$ is between 1000 and 3000.

## 3. PROOF OF THE UPPER BOUND THEOREM

We define the process $\mathscr{C}_p$, dependent on the states of the two decks, as follows. Find the set, $S$, of all positions, $j$, such that the $j$th card in deck 1 is the same as the $j$th card in deck 2. We call these the matched cards and matched positions. We will make sure that matched cards remain matched while allowing the two permutations of the decks to be as independent as possible on the unmatched cards.

Pick $j \in S$ and choose Boolean variables $e_j, e_{j+1},\ldots$ one by one, each with probability $p$ of being "Yes," until a "Yes" is reached. Do the same for $e_{j-1}, e_{j-2},\ldots$ as well, so that the boundaries of the packet containing position $j$ have been determined. Pick another $j \in S$ for which $e_j$ or $e_{j-1}$ has not been determined and again choose $e_j, e_{j+1},\ldots$ and $e_{j-1}, e_{j-2},\ldots$ until you reach a "Yes" or a $k$ such that $e_k$ has already been chosen. Continue doing this until $S$ is used up; at this point the packet boundaries for all the matched cards have been chosen. Now choose the remaining variables $e_i^1$ and $e_i^2$ independently. Shuffle deck 1 with an overhand shuffle corresponding to the variables $(e_1^1,\ldots, e_{N-1}^1)$ and use the variables $(e_1^2,\ldots, e_{N-1}^2)$ to shuffle deck 2.

**Remark.** The shuffle does not really depend on the order in which the positions in $S$ are picked; imagine that $e_1,\ldots, e_{N-1}$ are already determined but that you look at as few as possible in order to determine the boundaries of every packet containing a matched card. Later we will use a similar argument to choose a shuffle in $\mathscr{C}_p$ by first choosing some variables, $e_i^1$ and $e_j^2$ in nonoverlapping positions, and then going through the usual process, choosing $e_k = e_k^1$ or $e_k = e_k^2$ whenever possible and the rest independently.

**Example.** The decks begin like this:

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Deck 1: | A | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | T | J | Q | K |
| Deck 2: | 8 | 4 | J | K | 5 | 7 | A | 6 | 9 | T | 3 | Q | 2 |

and the random Boolean variables produce this diagram as explained below:

$$\begin{array}{ccccccccccccc} & & & & \downarrow & & & & \downarrow & \downarrow & & \downarrow & \\ \mathrm{A} & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \mathrm{T} & \mathrm{J} & \mathrm{Q} & \mathrm{K} \\ 8 & 4 & \mathrm{J} & \mathrm{K} & 5 & 7 & \mathrm{A} & 6 & 9 & \mathrm{T} & 3 & \mathrm{Q} & 2 \end{array}$$

$S = \{5, 9, 10, 12\}$, as shown by the ↓'s. Say $p = 1/2$. For $j = 5$ suppose $e_5 = Y$, $e_4 = N$, $e_3 = N$, $e_2 = Y$. For $j = 9$ we find that $e_9 = N$, $e_{10} = N$, $e_{11} = Y$, $e_8 = Y$. For $j = 10$, $e_j$ and $e_{j-1}$ have already been chosen. For $j = 12$, we find $e_{12} = Y$. We let $e_k^1 = e_k^2 = e_k$ for $k = 2, 3, 4, 5, 8, 9, 10, 11, 12$. We now choose the reamaining variables independently and find $e_1^1 = Y$, $e_6^1 = N$, $e_7^1 = Y$, $e_1^2 = N$, $e_6^2 = Y$, $e_7^2 = Y$. The packets shown in boxes are necessarily the same for both decks. The Boolean variables indicated by presence or absence of vertical slashes were chosen independently. The new order is

$$\begin{array}{lcccccc} & & \downarrow\qquad\quad & \downarrow\!\downarrow\qquad & \quad\downarrow\ \ \downarrow & \downarrow & \\ \text{Deck 1:} & \mathrm{A}\,|\,2 & \boxed{5\ \ 4\ \ 3} & 7\ \ 6\,|\,8 & \boxed{\mathrm{J}\ \ \mathrm{T}\ \ 9} & \boxed{\mathrm{Q}} & \mathrm{K} \\ \text{Deck 2:} & 4\ \ 8 & \boxed{5\ \ \mathrm{K}\ \ \mathrm{J}} & 7\,|\,\mathrm{A}\,|\,6 & \boxed{3\ \ \mathrm{T}\ \ 9} & \boxed{\mathrm{Q}} & 2 \end{array}$$

Note that all matched cards remain matched and some new matches (marked by ↓↓) may be created. It is clear that the distribution of shuffles restricted to either deck 1 or deck 2 is just $\mathcal{O}_p$. Now we have to calculate how many shuffles it takes before the two decks are identically arranged with probability at least $1 - 1/2^{r-1}$ regardless of their initial states.

Suppose that the $j$th card in deck 1 is identical to the $(j+1)$st card in deck 2. Then if $e_{j-1}^1 = e_{j-1}^2 = e_{j+1}^1 = e_{j+1}^2 = Y$, and $e_j^1 \neq e_j^2$, then the identical cards become matched. The probability of this happening is at least $2p^5q$, with equality when all six are chosen independently. Let

$$f(N) = \lceil \log[(1/2^r)N]/\log(1 - 2p^5q) \rceil$$

so that if we wait for a given card to have a position in deck 1 equal or next to its position in deck 2 at least $f(N)$ times, it has probability at most $(1/2^r)N$ of being unmatched. Now we appeal to a lemma whose proof is deferred to Section 5.

**Lemma 1.** For each $p$ there is a constant, $\gamma$, independent of $N$ such that if the coupled shuffle $\mathscr{C}_p$ is repeated $\gamma N^2$ times, the probability that a given card has had a position in deck 1 equal to or within 1 of its position in deck 2 at some point is at least 3/4 for any initial states of the two decks.

Assuming this lemma, we finish proving the theorem. For convenience refer to the event of a card, **c**, having a position in deck 1 equal to or differing by 1 from its position in deck 2 as $E(\mathbf{c})$. Then after $2f(N)\gamma N^2$ shuffles, the probability, $x$, that $E(\mathbf{c})$ has not occurred at least $f(N)$ times can be bounded by considering the shuffles in blocks of $\gamma N^2$. If $E(\mathbf{c})$ has not occurred at least $f(N)$ times then at least of the blocks must contain no such occurrence. The number of ways of selecting at least $f(N)$ blocks from $2f(N)$ is less than $2^{f(N)}$ so

$$x < 2^{f(N)}(1/4)^{f(N)} = (1/2)^{f(N)}$$

But $(1-2p^5q)^{f(N)} \leqslant (1/2^r)N$ and $1/2 \leqslant 1-2p^5q$, so

$$x < 1/2^{f(N)} \leqslant 1/2^r N$$

Then for any card, $\mathbf{c}$, we have

$$\begin{aligned}&\text{prob}(\mathbf{c}\text{ is unmatched after }2f(N)\gamma N^2\text{ shuffles})\\ &\quad\leqslant \text{prob}(E(\mathbf{c})\text{ failed to occur at least }f(N)\text{ times})\\ &\qquad+\text{prob}(E(\mathbf{c})\text{ occurred }f(N)\text{ times but }\mathbf{c}\text{ is still unmatched})\\ &\quad< (1/2^r)N + (1/2^r)N = 1/2^{r-1}N\end{aligned}$$

So the probability that there exists an unatched card after $2f(N)\gamma N^2$ shuffles is less than $N(1/2^{r\ 1}N) = 1/2^{r-1}$. Since $f(N)$ is bounded by $[\log(N) + r\log 2]$ times a constant slightly larger than $1/\log(1-2p^5q)^{-1}$, we can choose a $k \approx 2\gamma/\log(1-2p^5q)^{-1}$ and the theorem is proved. □

## 4. LOWER BOUNDS

Following the position of a single card gives a lower bound on the number of overhand shuffles it takes to randomize a deck. The analysis of the motion of a single card will also be used in the next section to prove Lemma 1. The argument proceeds by getting bounds on the mean and variance.

For any card $\mathbf{c}$, let $a_{\mathbf{c}}$ (respectively, $b_{\mathbf{c}}$) denote the number of cards after (respectively, before) $\mathbf{c}$ in the same packet as $\mathbf{c}$. We will supress the subscripts when no ambiguity arises. If $\mathbf{c}$ is in position $j$ before a shuffle then it is in position $j + a_{\mathbf{c}} - b_{\mathbf{c}}$ afterwards. In a random shuffle, $a_{\mathbf{c}}$ is a random variable with $\text{prob}(a_{\mathbf{c}} = m) = pq^m$, truncated at $N - \text{pos}(\mathbf{c})$, where $\text{pos}(\mathbf{c}) = j$ when $\mathbf{c}$ is in the $j$th position. The variable, $b_{\mathbf{c}}$, is distributed similarly but truncated at $\text{pos}(\mathbf{c}) - 1$.

Define

$$\overline{\text{pos}}(\mathbf{c}) = \text{pos}(\mathbf{c}) - (N+1)/2$$

Here are some easy facts about $a_{\mathbf{c}} - b_{\mathbf{c}}$.

**Lemma 2.** (

(2) $\mathbf{E}(a_{\mathbf{c}} - b$

(3) $|\mathbf{E}(a_{\mathbf{c}} -$

*Proof.* The
$\overline{\text{pos}}(\mathbf{c}) > 0$ (the c
$\min\{a'_{\mathbf{c}}, N - \text{pos}(\mathbf{c}$
$\text{pos}(\mathbf{c}) - 1$ so it is
about zero, so
$2\overline{\text{pos}}(\mathbf{c})$.

**Lemma 3.**

*Proof.* $a_{\mathbf{c}}$ a
$\text{Var}(b_{\mathbf{c}})$. The varia
of the correspondi
geometric variabl
truncated at $R$ ha

Hence $\text{Var}(a_{\mathbf{c}} - b_{\mathbf{c}}$
Let $\overline{\text{pos}}(\mathbf{c})'$ d

$\mathbf{E}(\overline{\text{pos}}(\mathbf{c})$

After $M$ shuffles, i

E

This is true for an
middle of the deck
on whether $N$ is o
this is linear in $M$
In the unifc
tribution, $\mathbf{Q}$, wit

**Lemma 2.** (1) $\mathbf{E}(a_c - b_c)$ has opposite sign to $\overline{\mathrm{pos}}(\mathbf{c})$

(2) $\mathbf{E}(a_c - b_c) = 0$ if and only if $\overline{\mathrm{pos}}(\mathbf{c}) = 0$

(3) $|\mathbf{E}(a_c - b_c)| \leqslant |\overline{\mathrm{pos}}(\mathbf{c})|$

*Proof.* The first two facts are obvious. To prove the third, assume $\overline{\mathrm{pos}}(\mathbf{c}) > 0$ (the other case is identical). Then $a_c$ can be written as $\min\{a'_c, N - \mathrm{pos}(\mathbf{c})\}$ where $a'_c$ is a geometric variable truncated at $\mathrm{pos}(\mathbf{c}) - 1$ so it is distributed identically to $b_c$. Then $a'_c - b_c$ is symmetric about zero, so $|\mathbf{E}(a_c - b_c)| = \mathbf{E}(a'_c - a_c) \leqslant \mathrm{pos}(\mathbf{c}) - 1 - [N - \mathrm{pos}(\mathbf{c})] = 2\,\overline{\mathrm{pos}}(\mathbf{c})$. □

**Lemma 3.** $\mathrm{Var}(a_c - b_c) \leqslant 2p/q^2$.

*Proof.* $a_c$ and $b_c$ are independent, so $\mathrm{Var}(a_c - b_c) = \mathrm{Var}(a_c) + \mathrm{Var}(b_c)$. The variances of the truncated variables are less then the variances of the corresponding untruncated variables. More precisely, an untruncated geometric variable has variance $q/p^2$, while a geometric variable $X$ truncated at $R$ has variance

$$\mathrm{Var}(X) = \frac{q}{p^2}\left[1 - q^R p\left(2R + 1 + \frac{q^{R+1}}{p}\right)\right]$$

Hence $\mathrm{Var}(a_c - b_c) \leqslant 2q/p^2$. □

Let $\overline{\mathrm{pos}}(\mathbf{c})'$ denote the value of $\overline{\mathrm{pos}}(\mathbf{c})$ after one more shuffle. Then

$$\begin{aligned}
\mathbf{E}(\overline{\mathrm{pos}}(\mathbf{c})')^2 &= \mathbf{E}(\overline{\mathrm{pos}}(\mathbf{c}) + a_c - b_c)^2\\
&= \mathbf{E}(\overline{\mathrm{pos}}(\mathbf{c}))^2 + 2\,\overline{\mathrm{pos}}(\mathbf{c})\,\mathbf{E}(a_c - b_c) + \mathbf{E}(a_c - b_c)^2\\
&\leqslant \overline{\mathrm{pos}}(\mathbf{c})^2 + \mathbf{E}(a_c - b_c)^2 - |\mathbf{E}(a_c - b_c)|^2\\
&\quad \text{by facts (1)–(3) above}\\
&= \overline{\mathrm{pos}}(\mathbf{c})^2 + \mathrm{Var}(a_c - b_c)\\
&\leqslant \mathbf{E}(\overline{\mathrm{pos}}(\mathbf{c}))^2 + 2q/p^2
\end{aligned}$$

After $M$ shuffles, induction gives

$$\mathbf{E}(\overline{\mathrm{pos}}(\mathbf{c}))^2 - [\text{original } \overline{\mathrm{pos}}(\mathbf{c})]^2 < 2qM/p^2$$

This is true for any card, $\mathbf{c}$. In particular, choosing $\mathbf{c}$ to be as near to the middle of the deck as possible gives $\overline{\mathrm{pos}}(\mathbf{c}) = 0$ or $1/2$ originally, depending on whether $N$ is odd or even, so $\mathbf{E}(\overline{\mathrm{pos}}(\mathbf{c}))^2 < 2qM/p^2 + 1/4$. The fact that this is linear in $M$ will imply the quadratic lower bound as follows.

In the uniform distribution, $\mathbf{E}(\overline{\mathrm{pos}}(\mathbf{c}))^2 = N^2 - 1/12$. For a distribution, $\mathbf{Q}$, with $|\mathbf{Q} - \mathbf{U}| \leqslant \varepsilon$, it is clear that $\int (x - N + 1/2)^2\, d\mathbf{Q}$ is minimized when **Q** agrees with **U** except that the $1/2N\varepsilon$ lowest values and the $1/2N\varepsilon$ highest values are changed to $\lfloor N/2 \rfloor$. In this case $\int (x-N+1/2)^2\, d\mathbf{Q} = 1/12N^2(1-\varepsilon)^2$ plus an error of less than $N/2$ introduced when $\varepsilon$ is not a multiple of $2/N$. So

$$|\mathcal{O}_p^M - \mathbf{U}| \leqslant \varepsilon$$

$$\Rightarrow \int (x-N+1/2)^2\, d(\mathcal{O}_p^M) \geqslant (1-\varepsilon)^2\,(N^2/12 - N/2)$$

$$\Rightarrow M > (1-\varepsilon)^2\,(N^2/6 - N/2)\,p^2/q$$

## 5. PROOF OF LEMMA 1

In order to show that the event $E(\mathbf{c})$ is likely to have occurred after a certain number of shuffles we use an "uncoupling" argument. Begin with a second pair of decks, decks 3 and 4, arranged exactly the same as decks 1 and 2 respectively. Decks 3 and 4 will be shuffled in a manner that is easier to analyze, but that usually gives the same result as the shuffle of decks 1 and 2. The argument then consists of establishing that $E(\mathbf{c})$ is likely to have occurred in decks 3 and 4 by the allotted time, and that if decks 1 and 2 have ever deviated from decks 3 and 4 then $E(\mathbf{c})$ was likely to have occurred in decks 1 and 2 at the time of the first deviation.

Formally, the process $\mathcal{U}_p$ will be defined on four decks. The card that we are interested in will be called $\mathbf{c}^{(i)}$, $i = 1, 2, 3, 4$, depending on which deck it is in. Assume that decks 1 and 3 start out identically arranged and that decks 2 and 4 do as well. At the start, consider the decks "coupled." Now follow these rules:

> If the decks have become uncoupled then shuffle decks 1 and 2 with a random shuffle from the distribution $\mathscr{C}_p$. In this case decks 3 and 4 receive independent random shuffles from the distribution $\mathcal{O}_p$.
>
> If the decks are still coupled then $\mathrm{pos}(\mathbf{c}^{(1)}) = \mathrm{pos}(\mathbf{c}^{(3)})$ and $\mathrm{pos}(\mathbf{c}^{(2)}) = \mathrm{pos}(\mathbf{c}^{(4)})$. Choose Boolean variables $e_i^3$ and $e_i^4$ until the packets containing $\mathbf{c}^{(3)}$ and $\mathbf{c}^{(4)}$ are determined. If the packets containing $\mathbf{c}^{(3)}$ and $\mathbf{c}^{(4)}$ both contain position $j$ for some $j$ then say the decks have become uncoupled. In this case decks 1 and 2 receive a random shuffle from $\mathscr{C}_p$ by letting $e_i = e_i^3$ for all defined values of $e_i^3$ letting $e_i = e_i^4$ when $e_i^4$ is defined but $e_i^3$ is not, and using any method of completing this to a coupled shuffle from $\mathscr{C}_p$ (the remark near the beginning of Section 3 tells us that we can choose a shuffle from $\mathscr{C}_p$ beginning this way without prejudicing it). Decks 3 and 4 use all the defined values of $e_i^3$ and $e_i^4$, but complete the shuffles to independent shuffles from $\mathcal{O}_p$. If the positions in the packets containing $\mathbf{c}^{(3)}$ and $\mathbf{c}^{(4)}$ do not overlap then the decks are still considered coupled. The Boolean variables $e_i$, $e_i^1$, and $e_i^2$ are chosen to give a shuffle with distributions $\mathscr{C}_p$ according to the provision choose $e_i = e_i^1 = e_i^3$ whenever the

> latter is defined
> remark in Secti
> more, complete

Note that $\mathcal{U}_p$
decks 3 and 4 is $\mathcal{O}$
of $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$ in t
"acted independen
let $T$ denote the fi
first time $E(\mathbf{c})$ occ
in decks 3 and 4; l
sign. Now use the

**Lemma 4.**
prob($T_{\mathrm{pos}} \leqslant 16p^2N$

**Lemma 5.** F

**Lemma 6.** F

Assuming the
and hence of the
Combining Lemma
for $N$ sufficiently la

Now $T > t$ and $T_{3}$

prob

$[p^4/($

where the first ine
the second is ar
$p^6/(q + 2p^4)(4p^2 +$
prob($T_{12} > 16p^2N^2$
to a total of $\gamma N^2$ sl
just adjust $\gamma$ for th
hypotheses of Lem
Lemmas 4, 5, and

latter is defined and choose $e_i = e_i^2 = e_i^4$ whenever the latter is defined. (Again the remark in Section 3 says this is a valid way of chosing a shuffle from $\mathscr{C}_p$.) Once more, complete the shuffles on decks 3 and 4 to independent shuffles from $\mathscr{O}_p$.

Note that $\mathscr{U}_p$ restricted to decks 1 and 2 is just $\mathscr{C}_p$ while $\mathscr{U}_p$ restricted to decks 3 and 4 is $\mathscr{O}_p \times \mathscr{O}_p$. This construction allows us to analyze the motion of $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$ in two cases, depending on whether or not they have always "acted independently." Fix a card $\mathbf{c}$ and define the following random times: let $T$ denote the first time the process $\mathscr{U}_p$ becomes uncoupled; let $T_{12}$ be the first time $E(\mathbf{c})$ occurs in decks 1 and 2; let $T_{34}$ be the first time $E(\mathbf{c})$ occurs in decks 3 and 4; let $T_{\text{pos}}$ be the first time that $\text{pos}(\mathbf{c}^{(3)}) - \text{pos}(\mathbf{c}^{(4)})$ changes sign. Now use the following lemmas.

**Lemma 4.** If $pq^{(N-1)/2}(1+N) < 1/2$ and $q^{(N-1)/2} < p$, then $\text{prob}(T_{\text{pos}} \leqslant 16p^2N^2/q) \geqslant 1/4$.

**Lemma 5.** For any $t$, $\text{prob}(T_{34} \leqslant t) \geqslant [p^2/(p^2+q)]\ \text{prob}(T_{\text{pos}} \leqslant t)$.

**Lemma 6.** For any $t$ $\text{prob}(T_{12} \leqslant t) \geqslant [p^4/(p^4+q)]\ \text{prob}(T \leqslant t)$.

Assuming these lemmas for the moment, the proof of Lemma 1, and hence of the Upper Bound Theorem, can be completed as follows. Combining Lemmas 4 and 5 with $t = 16p^2N^2/q$ in Lemma 5, it follows that for $N$ sufficiently large,

$$\text{prob}(T_{34} \leqslant 16p^2N^2/q) \geqslant p^2/(4p^2+4q) \tag{5.1}$$

Now $T > t$ and $T_{34} \leqslant t$ together imply $T_{12} < t$. So

$$\begin{aligned} &\text{prob}(T_{12} \leqslant 16p^2N^2/q) \\ &\qquad \geqslant \max\{p^2/(4p^2+4q) - \text{prob}(T \leqslant 16p^2N^2/q) \\ &[p^4/(q+p^4)]\ \text{prob}(T \leqslant 16p^2N^2/q)\} \\ &\qquad \geqslant p^6/(q+2p^4)(4p^2+4q) \end{aligned}$$

where the first inequality is a consequence of (5.1) and Lemma 6 and the second is an instance of $\max\{a-\gamma, \gamma b\} \geqslant ab/(1+b)$. Let $\delta = p^6/(q+2p^4)(4p^2+4q)$ and $\gamma = (16p^2/q)\lceil \log(1/4)/\log(1-\delta) \rceil$, so that $\text{prob}(T_{12} > 16p^2N^2/q) < 1-\delta$. By repeating blocks of $16p^2N^2/q$ shuffles up to a total of $\gamma N^2$ shuffles in all, we can force $\text{prob}(T_{12} \leqslant \gamma N^2) \geqslant 3/4$. Now just adjust $\gamma$ for the finitely many small values of $N$ that do not fit the hypotheses of Lemma 4 and Lemma 1 is proved. It remains to prove Lemmas 4, 5, and 6.

*Proof of Lemma* 4. Recall that $\mathscr{U}_p$ restricted to decks 3 and 4 gives independent overhand shuffles. So the two cards $\mathbf{c}^{(3)}$ and $\mathbf{c}^{(4)}$ undergo a series of independent overhand shuffles. Assume without loss of generality that initially $\text{pos}(\mathbf{c}^{(3)}) < \text{pos}(\mathbf{c}^{(4)})$. Recall from Section 4 the quantities $a_{c^{(3)}}$ and $b_{c^{(3)}}$. Call these $a$ and $b$ and let $a'$ and $b'$ denote the correspnding quantities for deck 4. From a given initial position we want to track the motion of $\mathbf{c}^{(3)}$ and $\mathbf{c}^{(4)}$ so put subscripts in the foregoing variables that increase with each move. In other words let $\text{pos}_i(\mathbf{c})$ denote the position of $\mathbf{c}$ after $i$ shuffles from the given position; then $a$ $b$ $a'$ and $b'$ are independent random variables gotten by truncating some geometric random variables $\tilde{a}$ $\tilde{b}$, $\tilde{a}'$ and $\tilde{b}'$ at $\text{pos}_{i-1}(\mathbf{c}^{(3)}) - 1$, $N - \text{pos}_{i-1}(\mathbf{c}^{(3)})$, $\text{pos}_{i-1}(\mathbf{c}^{(4)}) - 1$, and $N - \text{pos}_{i-1}(\mathbf{c}^{(4)})$, respectively. We want to determine when $\text{pos}_i(\mathbf{c}^{(3)}) - \text{pos}_i(\mathbf{c}^{(4)})$ becomes non-negative.

Now define a new sequence of variables $\{\mathbf{d}_i : i \geqslant 0\}$, that give us a lower bound on $\text{pos}_i(\mathbf{c})$ and are easier to work with. Let $\mathbf{d}_0 = \text{pos}_0(\mathbf{c}^{(3)}) - \text{pos}_0(\mathbf{c}^{(4)})$ and $\mathbf{d}_i = \mathbf{d}_{i-1} + z_i$ where $z_i = a_i - b_i - a_i'' + b_i''$, $a_i''$ is $\tilde{a}'$ truncated at $\text{pos}(\mathbf{c}^{(3)}) - 1$ and $b_i''$ is $\tilde{b}'$ truncated at $N - \text{pos}(\mathbf{c}^{(3)})$, so the truncation is according to deck 3 instead of deck 4. This makes $a_i - a_i''$ and $b_i - b_i''$ and hence $z_i$ symmetric about zero. Since $a_i'' \leqslant a_i'$ and $b_i'' \geqslant b_i'$ when $\text{pos}_{i-1}(\mathbf{c}^{(3)}) \leqslant \text{pos}_{i-1}(\mathbf{c}^{(3)})$, it is clear that $\mathbf{d}_i \leqslant \text{pos}_i(\mathbf{c}^{(3)}) - \text{pos}_i(\mathbf{c}^{(4)})$ as long as $\text{pos}_j(\mathbf{c}^{(3)}) \leqslant \text{pos}_j(\mathbf{c}^{(4)})$ for all $j < i$.

Clearly, $\mathbf{E}(\mathbf{d}_i) = \mathbf{d}_0$ for all $i$. Since the variables $z_i$ are all independent, the variance of $\mathbf{d}_i$ is given by

$$\begin{aligned}\sum_{j=0}^{i-1} \text{Var}(z_j) &= \sum_{j=0}^{i-1} [\text{Var}(a_j - a_j'') + \text{Var}(b_j - b_j'')] \\ &> \sum_{j=0}^{i-1} \max\{\text{Var}(a_j - a_j''), \text{Var}(b_j - b_j'')\}\end{aligned}$$

For each $j$, either $a_j$ and $a_j''$ are truncated at a value of at least $(N-1)/2$, or else $b_j$ and $b_j''$ are truncated at a value of at least $(N-1)/2$. Then the maximum is at least

$$\frac{q}{p^2} 1 - q^{(N-1)/2} p \left[ 2(N-1)/2 + 1 + \frac{q^{(N+1)/2}}{p} \right]$$

so the variance of $\mathbf{d}_i$ is at least $iq/2p^2$ by the hypotheses of the lemma.

Now define a stopping time $\tau = \min\{t : |\mathbf{d}_t - \mathbf{d}_0| \geqslant 2|\mathbf{d}_0|\}$. Note that $\tau < \infty$ almost surely. Then $\mathbf{E}(\mathbf{d}_\tau - \mathbf{d}_0)^2 \leqslant 4N^2$ since $|\mathbf{d}_\tau - \mathbf{d}_0| \leqslant |\mathbf{d}_{\tau-1} - \mathbf{d}_0| + |z_\tau| \leqslant N + N = 2N$. But $\mathbf{E}(\mathbf{d}_\tau^2) \geqslant q/2p^2 \mathbf{E}(\tau)$, so $\mathbf{E}(\tau) \leqslant 8p^2N^2/q$. Then $\text{prob}(\tau \leqslant 16p^2N^2/q) \geqslant 1/2$. Since $z_i$ is symmetric around 0, $\text{prob}(\mathbf{d}_j \geqslant 0$ for some $j \leqslant 16p^2N^2/q) \leqslant 1/4$. Since $\text{pos}_i(\mathbf{c}^{(3)}) - \text{pos}_i(\mathbf{c}^{(4)}) \geqslant \mathbf{d}_i$ until it changes sign, Lemma 4 is proved. □

*Proof of Lemma*
probability of $\mathbf{E}(\mathbf{c})$ in
times the probability t
Assuming this for the
$t$th shuffle and let $B_t \subseteq$

$\text{prob}(T_{34} = t) \geqslant$

$\geqslant$

$=$

and summing over
$p^2/(p^2+q) \geqslant \text{prob}(T_{\text{po}}$

To show the clair
and define $a$, $a'$, $b$, a
$(b'+a) = 0$, 1, or $-1$.
$\text{pos}(\mathbf{c}^{(3)}) - \text{pos}(\mathbf{c}^{(4)})$ ch
and only if $s + b + a' -$
by the truncation; it is
random variables s
$\text{prob}(D(\mathbf{c}) \mid s, b, a')$ is s

The probability t
$a' - 1$, the ratio prob(

$\frac{(t+1)p^2}{(t+4}$

$= \frac{(t}{}$

$\geqslant \frac{}{\Sigma}$

$= \frac{}{(t}$

$= \frac{}{(t}$

$= \frac{}{(t}$

$\geqslant \frac{3(}{(}$

*Proof of Lemma* 5. I claim that from any initial arrangement the probability of $\mathbf{E}(\mathbf{c})$ in decks 3 and 4 on the next shuffle is at least $p^2/q$ times the probability that $\operatorname{pos}(\mathbf{c}^{(3)}) - \operatorname{pos}(\mathbf{c}^{(4)})$ changes sign on that shuffle. Assuming this for the moment, let $\mathscr{A}_t$ be the $\sigma$ algebra of events up to the $t$th shuffle and let $B_t \subseteq \mathscr{A}_t$ be the event $T_{34} \wedge T_{\text{pos}} > t$. Then for any $t$,

$$\begin{aligned}\operatorname{prob}(T_{34} = t) &\geqslant \int_{B_{t-1}} \operatorname{prob}(T_{34} = t \mid \mathscr{F}_{t-1})\\ &\geqslant [p^2/(p^2+q)] \int_{B_{t-1}} \operatorname{prob}(T_{34} \wedge T_{\text{pos}} = t \mid \mathscr{F}_{t-1})\\ &= [p^2/(p^2+q)] \operatorname{prob}(T_{34} \wedge T_{\text{pos}} = t)\end{aligned}$$

and summing over $t$ gives $\operatorname{prob}(T_{34} \leqslant t) \geqslant \operatorname{prob}(\min\{T_{34}, T_{\text{pos}}\} \leqslant t)\, p^2/(p^2+q) \geqslant \operatorname{prob}(T_{\text{pos}} \leqslant t)\, p^2/(p^2+q)$.

To show the claim, suppose $\mathbf{c}^{(3)}$ and $\mathbf{c}^{(4)}$ are in positions $j$ and $j+s$ and define $a$, $a'$, $b$, and $b'$ as before. Then $\mathbf{E}(\mathbf{c})$ occurs iff $s+b+a'-(b'+a)=0$, 1, or $-1$. Similarly, the event $D(\mathbf{c})$, defined to occur when $\operatorname{pos}(\mathbf{c}^{(3)}) - \operatorname{pos}(\mathbf{c}^{(4)})$ changes sign without the occurrence of $E(\mathbf{c})$, occurs if and only if $s+b+a'-(b'+a) < -1$. Now $\operatorname{prob}(E(\mathbf{c}) \mid s, b, a')$ is unaffected by the truncation; it is just the probability that two independent geometric random variables sum to $s+b+a'$ or to $s+b+a' \pm 1$. But $\operatorname{prob}(D(\mathbf{c}) \mid s, b, a')$ is strictly reduced by the truncation.

The probability that $b'+a=r$ is $(r+1)p^2q^r$, so letting $t=s+b+a'-1$, the ratio $\operatorname{prob}(E(\mathbf{c}) \mid s, b, a')/\operatorname{prob}(D(\mathbf{c}) \mid s, b, a')$ is at least

$$\begin{aligned}&\frac{(t+1)p^2q^t + (t+2)p^2q^{t+1} + (t+3)p^2q^{t+2}}{(t+4)p^2q^{t+3} + (t+5)p^2q^{t+4} + \cdots}\\ &\quad= \frac{(t+1)+(t+2)q+(t+3)q^2}{\sum_{i\geqslant 3}(t+1+i)q^i}\\ &\quad\geqslant \frac{3(t+1)q^2}{\sum_{i\geqslant 3}(t+4)q^i + \sum_{i\geqslant 4}(i-3)q^i}\\ &\quad= \frac{3(t+1)q^2}{(t+4)p^{-1}q^3 + p^{-2}q^4}\\ &\quad= \frac{3(t+1)p^2}{(t+4)pq+q^2}\\ &\quad= \frac{3(t+1)p^2}{(t+4)q-(t+4)q^2+q^2}\\ &\quad\geqslant \frac{3(t+1)p^2}{(t+4)q} \geqslant p^2/q \qquad \text{sincc} \quad t \geqslant s-1 \geqslant 1\end{aligned}$$

Summing over all triples $s$, $b$, $a'$ gives the inequality for unconditional probabilities. □

*Proof of Lemma* 6. The method is analogous to the method of proving Lemma 5. We will show that from any position that is still coupled, the probability of $E(\mathbf{c})$ occurring in decks 1 and 2 within one shuffle is at least $p^4/q$ times as great as the probability that the decks uncouple. It follows that the probability of the former event occurring before any time $t$ is at least $p^4/(1+p^4)$ times the probability of the latter occurring before time $t$.

Assume without loss of generality that $\text{pos}(\mathbf{c}^{(1)}) = \text{pos}(\mathbf{c}^{(3)}) = j$ and $\text{pos}(\mathbf{c}^{(2)}) = \text{pos}(\mathbf{c}^{(4)}) = j+s$, $s > 0$. If the decks uncouple then either

case (a): $e^3_j = e^3_{j+1} = \cdots = e^3_{j+s-1} = \text{No}$; or

case (b): For some $k \geqslant 0$, $e^3_j = e^3_{j+1} = \cdots = e^3_{j+s-2-k} = \text{No}$, $e^3_{j+s-1-k} = \text{Yes}$ and $e^4_{j+s-1} = e^4_{j+s-2} = \cdots = e^4_{j+s-k-1} = \text{No}$.

So prob (uncoupling) $\leqslant$ prob(case a) + prob(case b) $= q^s + spq^s$. The probability of $E(\mathbf{c})$ is at least probability that $\mathbf{c}$ is the first card in its packet and moves forward to some position, $j+r$ $(0 \leqslant r \leqslant s)$, while $\mathbf{c}^{(2)}$ is the last card in its packet and moves backwards to position $j+r+1$. For this to occur it is necessary that $e^1_{j-1} = \text{Yes}$, $e^1_j = \cdots = e^1_{j+r-1} = \text{No}$, $e^1_{j+r} = \text{Yes}$, $e^2_{j+r+1} = \text{Yes}$, $e^2_{j+r+2} = \cdots = e^2_{j+s-1} = \text{No}$ and $e^2_{j+s} = \text{Yes}$. The probability of this is therefore $sp^4q^{s-2}$. So the ratio of the two probabilities is at least $sp^4q^{s-2}/(q^s + spq^s) = p^4/q[pq + (q/s)] \geqslant p^4/q(p+q) = p^4/q$, as desired. □

## 6. NUMERICAL BOUNDS

A very fine overhand shuffle of a 52-card deck might have $p = 1/2$. For a more typical shuffle, $p$ might be 1/4. (In fact I have often seen people shuffle with $p \approx 1/15$. You do not need any fancy calculations to see that 5 or 6 such shuffles do virtually nothing to a 52-card deck.) Upper and lower bounds are given for the number of shuffles needed to come within $\varepsilon$ of uniform for various $\varepsilon$.

For the upper bound use $\varepsilon = 1/2$ and verify that the hypotheses of Lemma 4 hold. Then calculate the following values:

| $p$ | $f(N)$ | $\delta$ | $16p^2/q$ | $\gamma N^2$ | $(N^2/6 - N/2)\,p^2/q$ |
|---|---|---|---|---|---|
| 1/2 | 170 | 1/120 | 8 | $3.5 \times 10^6$ | 210 |
| 1/4 | 3700 | 1/10000 | 4/3 | $5 \times 10^7$ | 35 |

So the bounds are

| | Lower bounds | | | Upper bound |
|---|---|---|---|---|
| | $\varepsilon = 1/2$ | $\varepsilon = 1/4$ | $\varepsilon \to \infty$ | $\varepsilon = 1/2$ |
| $p = 1/2$ | 50 | 120 | 210 | $1.2 \times 10^9$ |
| $p = 1/4$ | 9 | 20 | 35 | $3.7 \times 10^{11}$ |

A small amount of numerical evidence indicates that the actual number of shuffles needed to make $|\mathcal{O}_p^k - \mathrm{U}| \leqslant 1/2$ for a 52-card deck with $p = 1/2$ is between 1000 and 3000. The bound of 1000 comes from counting the number of inversions in the permutation and comparing it to the median number under the uniform distribution; even after 1000 shuffles, the chance is less than 0.2 that the number of inversions exceeds the median. The bound of 3000 comes from running $\mathscr{C}_p$ on a home computer; the median coupling time seems to be less than 3000 shuffles.

An interesting question not addressed in this paper is what happens when $p$ is allowed to vary with $N$. It is not known, for example, how to choose $p$ as an asymptotic function of $N$ in order to achieve fastest mixing. The lower bound in Section 4 generalizes to order of $(pN)^2$, but in addition there is a lower bound of $N/p$ gotten by seeing when most pairs of adjacent cards have been separated. When $p = N^{-1/3}$ these bounds coincide at order of $N^{4/3}$ shuffles. Whether this is the fastest possible mixing (possibly up to factors of $\log N$) is still open.

## ACKNOWLEDGMENTS

This work was supported by a National Science Foundation Graduate Fellowship. I would like to thank the referee for two careful readings resulting in the correction of several nontrivial mistakes.

## REFERENCES

1. Aldous, D., and Diaconis, P. (1986). Shuffling cards and stopping times. *Am. Math. Monthly* **93** (5), 333.
2. Aldous, D. (1983). Random walks on finite groups and rapidly mixing Markov chains. In: *Séminaire de Probabilités XVII*, Springer Lecture Notes 986, Springer-Verlag, New York.