# Seven Shuffles For The Win

Sajid Rizvi

November 2020

## 1 Introduction

How many times do you have to shuffle a deck of cards in order to mix them reasonably well? The answer is about seven for a deck of fifty two cards, or so claims Persi Diaconis[1].In this paper a mathematical model of card shuffling is constructed, and used to determine how much shuffling is necessary to randomize a deck of cards. The crucial aspect of this model is rising sequences of permutations. The probability of an arrangement of cards occurring under shuffling is a function only of the number of rising sequences in the permutation. This fact makes computation of variation distance, a measure of randomness, feasible; for in an n card deck there are at most n rising sequences but n! possible arrangements. This computation is done exactly for n = 52, and other approximation methods are considered.

## 2 What do we mean by a shuffle?

Before we begin with our exposition, we must define what a shuffle, or method of shuffling, is. It's just a probability density on $S_n$, considering each permutation as a way of rearranging the deck. This means that each permutation is given a certain fixed probability of occurring, and that all such probabilities add upto one.
We would now like to choose a realistic model of how actual cards are physically shuffled by people. A particular one with nice mathematical properties is given by the "riffle shuffle."

**Definition 2.1** (riffle shuffle)**.** A riffle shuffle is a probability density on $S_n$ achieved as follows, first cut the deck into two packets, the first containing $k$ cards, and the other the remaining $n-k$ cards. Choose $k$, the number of cards cut, according to the binomial density, meaning the probability of the cut occurring exactly after $k$ cards is given by $\left(\begin{array}{c} n \\ k \end{array}\right)/2^n$. Once the deck has been cut into two packets, interleave the cards from each packet in any possible way, such that the cards of each packet maintain their own relative order.

This means that the cards originally in positions $1, 2, 3, \ldots k$ must still be in the same order in the deck after it is shuffled, even if there are other cards in-between; the same goes for the cards originally in positions $k+1, k+2, \ldots n$. This requirement is quite natural when you think of how a person shuffles two packets of cards, one in each hand. The cards in the left hand must still be in the same relative order in the shuffled deck, no matter how they are interleaved with the cards from the other packet, because the cards in the left hand are dropped in order when shuffling; the same goes for the cards in the right hand.

Choose among all such interleavings uniformly, meaning each is equally likely. since there are $\left( \begin{array}{c} n \\ k \end{array} \right)$ possible interleavings (as we only need choose $k$ spots among $n$ places for the first packet, the spots for the cards of the other packet then being determined), this means any particular interleaving has probability $1/\left( \begin{array}{c} n \\ k \end{array} \right)$ of occuring. Hence the probability of any particular cut followed by a particular interleaving, with $k$ the size of the cut, is $\left( \begin{array}{c} n \\ k \end{array} \right)/2^n \cdot 1/\left( \begin{array}{c} n \\ k \end{array} \right) = 1/2^n$. Note that this probability $1/2^n$ contains no information about the cut or the interleaving! In other words, the density of cuts and interleavings is uniform every pair of a cut and a possible resulting interleaving has the same probability.

This uniform density on the set of cuts and interleavings now induces in a natural way a density on the set of permutations, i.e. a shuffle, according to our definition. We will call this the riffle shuffle and denote it by $R$. It is defined for $\pi$ in $S_n$ by $R(\pi) =$ the sum of the probabilities of each cut and interleaving that gives the rearrangement of the deck corresponding to $\pi$, which is $1/2^n$ times the number of ways of cutting and interleaving that give the rearrangement of the deck corresponding to $\pi$. In short, the chance of any arrangement of cards occurring under riffle shuffling is simply the proportion of ways of riffling which give that arrangement. Now let $R^{(k)}$ stand for convoluting R with itself k times. This corresponds to the density after k riffle shuffles. For which k does $R^{(k)}$ produce a randomized deck? The next section begins to answer this question.

# 3 A measure of randomness

Before we consider the question of how many times we need to shuffle, we must decide what we want to achieve by shuffling. The answer should be randomness of some sort. What does randomness mean? Simply put, any arrangement of cards is equally likely; no one ordering should be favored over another. This means the uniform density $U$ on $S_n$, each permutation having probability $U(\pi) = 1/|S_n| = 1/n!$

Now it turns out that for any fixed number of shuffles, no matter how large, riffle shuffling does not produce complete randomness in this sense. So when

we ask how many times we need to shuffle, we are not asking how far to go in order to achieve randomness, but rather to get close to randomness. So we must define what we mean by close, or far, i.e. we need a distance between densities. The concept we will use is called variation distance (which is essentially the $L^1$ metric on the space of densities). Suppose we are given two probability densities, $Q_1$ and $Q_2$, on $S_n$. Then the variation distance between $Q_1$ and $Q_2$ is defined to be

$$\|Q_1 - Q_2\| = \frac{1}{2} \sum_{\pi \in S_n} |Q_1(\pi) - Q_2(\pi)|$$

The $\frac{1}{2}$ normalizes the result to always be between 0 and 1. Now the question we really want to ask is: how big must we take $k$ to make the variation distance $\left\|R^{(k)} - U\right\|$ between the riffle and uniform small? This can be best answered by a graph of $\left\|R^{(k)} - U\right\|$ versus $k$.

# 4 Rising Sequences

Before we determine $R^{(k)}$ we must consider a fundamental concept, that of a rising sequence.

**Definition 4.1** (Rising Sequence)**.** A rising sequence of a permutation is a maximal consecutively increasing subsequence.

It turns out that a deck breaks down as a disjoint union of its rising sequences, since the union of any two consecutively increasing subsequences containing a given element is also a consecutively increasing subsequence that contains that element. The following example sheds light on how rising sequences occur in a permutation.

**Example 1.** Suppose we know that the order of an eight card deck after shuffling the natural order is 45162378. Start with any card, say 3. We look for the next card in value after it, 4, and do not find it. So we stop looking after and look before the 3. We find 2, and then we look for 1 before 2 and find it. So one of the rising sequences is given by 123. Now start again with 6. We find 7 and then 8 after it, and 5 and then 4 before it. So another rising sequence is 45678. We have accounted for all the cards, and are therefore done. Thus this deck has only two rising sequences.

# 5 Generalizing the riffle shuffle

We now consider a more general model of shuffling, to be referred to as an $a-$shuffle.

**Definition 5.1** (a-shuffle)**.** An $a$ -shuffle is a probability density on $S_n$, achieved as follows. Let $a$ stand for any positive integer. Cut the deck into $a$ packets, of nonnegative sizes $p_1, p_2, \ldots, p_a$, with the probability of this particular packet

structure given by the multinomial density: $\binom{n}{p_1, p_2, \ldots, p_a} / a^n$. Note we must have $p_1 + \cdots + p_a = n$, but some of the $p_i$ may be zero. Now interleave the cards from each packet in any way, so long as the cards from each packet maintain their relative order among themselves.

With a fixed packet structure, consider all interleavings equally likely. Let us count the number of such interleavings. We simply want the number of different ways of choosing, among $n$ positions in the deck, $p_1$ places for things of one type, $p_2$ places for things of another type, etc. This is given by the multinomial coefficient $\binom{n}{p_1, p_2, \ldots, p_a}$. Hence the probability of a particular rearrangement, i.e. a cut of the deck and an interleaving, is

$$\binom{n}{p_1, p_2, \ldots, p_a} / a^n \cdot \binom{n}{p_1, p_2, \ldots, p_a} = \frac{1}{a^n}$$

## 5.1 Another Way To Look At It

There is a useful code that we can construct to specify how a particular $a$ -shuffle is done. This is done through $n$ digit base $a$ numbers. Let $A$ be any one of these $n$ digit numbers. Count the number of 0 's in $A$. This will be the size of the first packet in the $a$ -shuffle, $p_1$. Then $p_2$ is the number of 1 's in $A$, and so on, up through $p_a$ = the number of $(a-1)$ 's. This cuts the deck cut into $a$ packets. Now take the beginning packet of cards, of size $p_1$. Envision placing these cards on top of all the 0 digits of $A$, maintaining their relative order as a rising sequence. Do the same for the next packet, $p_2$, except placing them on the 1's. Again, continue up through the $(a-1)'$ 's. This particular way of rearranging the cards will then be the particular cut and interleaving corresponding to $A$.

Here is an example, with the deck starting in natural order.

**Example 1.** Let $A = 23004103$ be the code for a particular 5 -shuffle of the 8 card deck. There are three 0 's, one 1, one 2, two 3 's, and one 4. Thus $p_1 = 3, p_2 = 1, p_3 = 1, p_4 = 2$, and $p_5 = 1$. So the deck is cut into 123—4—5—67— 8 So we place 123 where the 0 's are in $A, 4$ where the 1 is, 5 where the 2 is, 67 where the 3 's are, and 8 where the 4 is. We then get a shuffled deck of 56128437 when $A$ is applied to the natural order.

Reflection shows that this code gives a bijective correspondence between n digit base a numbers and the set of all ways of cutting and interleaving an n card deck according to the a-shuffle. In fact, if we put the uniform density on the set of n digit base a numbers, this transfers to the correct uniform probability for cutting and interleaving in an a-shuffle, which means the correct density is induced on $S_n$, i.e. we get the right probabilities for an a-shuffle.

### 5.2 Relation with Rising Sequences

**Theorem 1.** *The probablity of achieving a permutation $\pi$ when doing an $a$-shuffle is given by $\binom{n+a-r}{n}/a^n$, where $r$ is the number of rising sequences in $\pi$.*

*Proof.* First note that if we establish and fix where the $a-1$ cuts occur in an $a$ -shuffle, then whatever permutations can actually be achieved by interleaving the cards from this cut/packet structure are achieved in exactly one way; namely, just drop the cards in exactly the order of the permutation. Thus the probability of achieving a particular permutation is the number of possible ways of making cuts that could actually give rise to that permutation, divided by the total number of ways of making cuts and interleaving for an $a$ -shuffle.

So let us count the ways of making cuts in the naturally ordered deck that could give the ordering that results when $\pi$ is applied. If we have $r$ rising sequences in $\pi$, we know exactly where $r-1$ of the cuts have to have been; they must have occurred between pairs of consecutive cards in the naturally ordered deck such that the first card ends one rising sequence of $\pi$ and the second begins another rising sequence of $\pi$. This means we have $a-1-(r-1)=a-r$ unspecified, or free, cuts. These are free in the sense that they can in fact go anywhere. So we must count the number of ways of putting $a-r$ cuts among $n$ cards. This can easily be done by considering a sequence of $(a-r)+n$ blank spots which must be filled by $(a-r)$ things of one type (cuts) and $n$ things of another type (cards). There are $\binom{(a-r)+n}{n}$ ways to do this, i.e. choosing $n$ places among $(a-r)+n$

This is the numerator for our probability expressed as a fraction; the denominator is the number of possible ways to cut and interleave for an $a$ -shuffle. By considering the encoding of shuffles we see there are $a^n$ ways to do this, as there are this many $n$ digit base $a$ numbers. Hence our result is true. □

# 6 Multiplication Theorem

**Theorem 2.** *An $a$-shuffle followed by a $b$-shuffle is equivalent to a single $ab$-shuffle, in the sense that both processes give exactly the same resulting probability density on the set of permutations.*

*Proof.* Let us use the previously described code for shuffles. Suppose that $A$ is an $n$ digit base $a$ number, and $B$ is an $n$ digit base $b$ number. Then first doing the cut and interleaving encoded by $A$ and then doing the cut and interleaving encoded by $B$ gives the same permutation as the one resulting from the cut and interleaving encoded by the $n$ digit base $ab$ number given by $A^B\&B$. $A^B$ is defined to be the code that has the same base $a$ digits as $A$, but rearranged according to the permutation specified by $B$. The symbol & in $A^B\&B$ stands for digit-wise concatenation of two numbers, meaning treat the base $a$ digit $A_i^B$ in the $i$ th place of $A^B$ together with the base $b$ digit $B_i$ in the $i$ th place of $B$

as the base $ab$ digit given by $A_i^B \cdot b + B_i$. In other words, treat the combination $A_i^B \& B_i$ as a two digit number, the right-most place having value 1, and the left-most place having value $b$, and then treat the result as a one digit base $ab$ number.

Why this formula holds is better shown by an example than by general formulas. Suppose $A = 012210$ is the code for a particular 3 shuffle, and $B = 310100$ is the code for a particular 4 -shuffle. (Again we are abusing terminology slightly.) Let $\pi_A$ and $\pi_B$ be the respective permutations. Then in the tables below note that $\pi_A \circ \pi_B$, the result of a particular 3 -shuffle followed by a particular 4 -shuffle, and $\pi_{A^B \& B}$, the result of a particular 12 -shuffle, are the same permutation.

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $\pi_A(i)$ | 1 | 3 | 5 | 6 | 4 | 2 |
| $\pi_B(i)$ | 6 | 4 | 1 | 5 | 2 | 3 |
| $\pi_A \circ \pi_B(i)$ | 2 | 6 | 1 | 4 | 3 | 5 |

| $A$ | 0 | 1 | 2 | 2 | 1 | 0 |
|---|---|---|---|---|---|---|
| $B$ | 3 | 1 | 0 | 1 | 0 | 0 |
| $A^B$ | 0 | 2 | 0 | 1 | 1 | 2 |
| $B$ | 3 | 1 | 0 | 1 | 0 | 0 |
| $A^B \& B$ | 3 | 9 | 0 | 5 | 4 | 8 |

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $\pi_{A^B \& B}(i)$ | 2 | 6 | 1 | 4 | 3 | 5 |

We now have a formula $A^B \& B$ that is really a one-to-one correspondence between the set of pairs, consisting of one $n$ digit base $a$ number and one $n$ digit base $b$ number, and the set of $n$ digit base $ab$ numbers; further this formula has the property that the cut and interleaving specified by $A$, followed by the cut and interleaving specified by $B$, result in the same permutation of the deck as that resulting from the cut and interleaving specified by $A^B \& B$. since the probability densities for $a$, $b$, and $ab-$ shuffles are induced by the uniform densities on the sets of $n$ digit base $a, b$, or $ab$ codes, respectively, the properties of the one-to one correspondence imply the induced densities on $S_n$ of an $a$ -shuffle followed by a $b$ -shuffle and an $ab$ -shuffle are the same. Hence our result is true. □

# 7 Connecting the dots

Let us now combine our two major results of the last section to get a formula for $R^{(k)}$, the probability density for the riffle shuffle done $k$ times. This is just $k$ 2-shuffles, one after another. So by the multiplication theorem, this is equivalent to a single $2 \cdot 2 \cdot 2 \cdots 2 = 2^k$ -shuffle. Hence in the $R^{(k)}$ density,

there is a $\binom{2^k+n-r}{n}/2^{nk}$ chance of a permutation with $r$ rising sequences occurring, by our rising sequence formula. This now allows us to work on the variation distance $\|R^k - U\|$. For a permutation $\pi$ with $r$ rising sequences, we see that

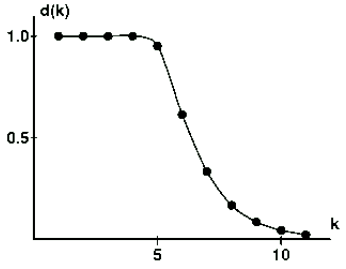
$$\left|R^k(\pi) - U(\pi)\right| = \left|\binom{2^k+n-r}{n}/2^{nk} - \frac{1}{n!}\right|$$

We must now add up all the terms like this, one for each permutation. We can group terms in our sum according to the number of rising sequences. If we let $A_{n,r}$ stand for the number of permutations of $n$ cards that have $r$ rising sequences, each of which have the same probabilities, then the variation distance is given by

$$\|R^k - U\| = \frac{1}{2}\sum_{r=1}^{n} A_{n,r}\left|\binom{2^k+n-r}{n}/2^{nk} - \frac{1}{n!}\right|$$

The only thing unexplained is how to calculate the $A_{n,r}$. These are called the Eulerian numbers[2], and various formulas are given for them. One recursive one is $A_{n,1} = 1$ and $A_{n,r} = r^n - \sum_{j=1}^{r-1}\binom{n+r-j}{n}A_{n,j}$

Now the expression for variation distance may seem formidable, and it is. But it is easy and quick for a computer program to calculate and graph $\|R^k - U\|$ versus $k$ for any specific, moderately sized $n$. Even on the computer, however, this computation is tractable because we only have $n$ terms, corresponding to each possible number of rising sequences. If we did not have the result on the invariance of the probability when the number of rising sequences is constant, we would have $|S_n| = n!$ terms in the sum. For $n = 52$, this is approximately $10^{68}$ which is much larger than any computer could handle. Here[3] is the graphical result of a program that does the calculations for $n = 52$. The horizontal axis is the number of riffle shuffles, and the vertical axis is the variation distance to uniform. The answer is finally at hand. It is clear that the




graph makes a sharp cutoff at $k = 5$, and gets reasonably close to 0 by $k = 11$. A good middle point for the cutoff seems to $k = 7$, and this is why seven shuffles are said to be enough for the usual deck of 52 cards.

# References


[1] Dave Bayer and Persi Diaconis. "Trailing the Dovetail Shuffle to its Lair". In: *Ann. Appl. Probab.* 2 (May 1992), pp. 294–313. DOI: 10.1214/aoap/1177005705. URL: https://doi.org/10.1214/aoap/1177005705.

[2] Wikipedia contributors. *Eulerian number — Wikipedia, The Free Encyclopedia*. [Online; accessed 28-November-2020]. 2020. URL: https://en.wikipedia.org/w/index.php?title=Eulerian_number&oldid=985518776.

[3] David Austin. *How Many Times Do I Have to Shuffle This Deck?* 2010. URL: http://www.ams.org/publicoutreach/feature-column/fcarc-shuffle.