Hang Lung Mathematics Awards
Vol. 11 (2025), pp. 267–297

THE MATHEMATICS OF RIFFLE SHUFFLING FROM SINGLE TO MULTI-DECK CARD GAMES

A RESEARCH REPORT SUBMITTED TO THE SCIENTIFIC COMMITTEE OF THE HANG LUNG MATHEMATICS AWARD

TEAM NUMBER
2596844

TEAM MEMBERS
JUSTIN WANG AND DANIEL WANG

TEACHER
MS. ALEXANDRA STREETER

SCHOOL
KING GEORGE V SCHOOL

Abstract. While significant progress has been made on the question "How many shuffles are required to sufficiently randomize a deck of cards?", the most influential work on the subject remains the 1992 paper by Bayer and Diaconis, which provided an answer for a standard 52-card deck, proving that seven shuffles suffice. However, many variations of this problem have yet to be extensively explored. In particular, multiple decks of cards are commonly used in various card games, such as the popular Chinese game GuanDan (2 decks), making the randomness of a deck after shuffling highly relevant in both recreational and professional settings. In this paper, we provide a mathematical analysis on the randomness of riffle shuffling multiple decks of cards, extending the result of Bayer and Diaconis [2] to the multi-deck setting. A key contribution of ours is the investigation of the randomness of a player's hand after riffle shuffling in a deck, a question that has not been investigated at all previously, from which we are able to give explicit formulae for the separation distance for a player's hand after riffle shuffling for both the single deck and double deck setting.

Contents

1. Introduction

Playing cards trace their origins to the 9th century, first appearing in China during the Tang Dynasty. From there, they spread to Persia, Egypt, and eventually Europe, where the standard 52-card deck emerged in the 13th century. Alongside the rise of playing cards came the practice of shuffling—a range of methods designed to randomize the deck and eliminate predictability in the order of cards drawn. Over time, numerous shuffling techniques have been developed and utilized, including the overhand shuffle and the riffle shuffle, among others. In this paper, we focus exclusively on the riffle shuffle, widely regarded as the most commonly used shuffling method. This technique involves splitting the deck into two piles and alternately interleaving them to produce a shuffled deck. The primary aim of this study is to determine the number of shuffles required to adequately randomize a multi-deck of cards, a question of practical significance due to its relevance in various card games such as casino blackjack and the popular Chinese game of GuanDan. Importantly, we recognize that in the context of card games, the randomness of individual hands dealt to each player is of greater importance than the overall ordering of the deck. We also study this question in this paper, which has not been studied before.

1.1. **Past Research.**

Early work on the mathematics of card shuffling can be traced back to the mid-20th century. In 1955, Ed Gilbert (working with Claude Shannon) in his technical memorandum *Theory of Shuffling* [3], studied this problem using the new science of information theory. In particular, Gilbert defined the *GSR Shuffle* where GSR stands for Gilbert-Shannon-Reeds, which is commonly used in all studies as the model for riffle shuffling.

Following this, various famous mathematicians including Persi Diaconis, David Bayer and David Aldous worked on this problem. This was until 1992, when Bayer and Diaconis published the landmark paper *Trailing the Dovetail Shuffle to its Lair* [2]. They generalized the GSR shuffle to the so-called *a-shuffle*, a form of riffle shuffling that splits the deck into a packets instead of just 2. Starting from the initial ordering of $(1, \ldots, n)$ for the deck of n cards, Bayer and Diaconis proved the probability that permutation π is produced after an a-shuffle is:

$$P_a(\pi) = \frac{\binom{a+n-r}{n}}{a^n}$$

where r is the number of rising sequences in π. Notably, it was proved [2] that performing s GSR shuffles is equivalent to an a-shuffle where $a = 2^s$. Hence, the probability of producing a given permutation π with r rising sequences after performing s GSR shuffles on a deck of n cards is:

$$P_{2^s}(\pi) = \frac{\binom{n+2^s-r}{n}}{2^{sn}}.$$

This formula enables the precise calculation of the total variation distance (which will be defined later) from the uniform distribution for a deck of n cards after an arbitrary number of GSR shuffles [2]. With this, Bayer and Diaconis concluded that 7 GSR shuffles are enough to reasonably shuffle a standard deck of 52 cards, after the total variation distance falls below $\frac{1}{2}$. In general for a deck on n cards, $\sim \frac{3}{2}\log_2 n$ shuffles are needed to sufficiently randomize a deck of n cards as $n \to \infty$. These somewhat surprising results were discussed and published on various newspapers including the New York Times article *In Shuffling Cards, 7 is Winning Number* [4].

Following Bayer and Diaconis [2], a number of authors have further studied and generalized this problem in various ways. For example, in 1999, Trefethen and Trefethen [5] revived Gilbert and Shannon's idea of measuring randomness as information loss. Through repeated Monte Carlo simulations, they demonstrated that about 3.52 percent of information remains after only 5 shuffles, and about 0.92 percent of information remains after 6 shuffles.

Another interesting perspective on the problem is how riffle shuffling impacts specific card games. In 2004, Van Zuylen and Schalekamp [6] examined the card game New Age Solitaire. In this game, a deck of cards is divided into two piles, which we will refer to as Pile 1 and Pile 2. Pile 1 consists of the Ace through King of Hearts, followed by the Ace through King of Clubs. Pile 2 consists of the Ace through King of Spades, followed by the Ace through King of Diamonds. Generalizing the initial deck to $\{1, 2, \ldots, 52\}$, Pile 1 will consist of cards $\{1, 2, \ldots, 26\}$, and Pile 2 will consist of cards $\{27, 28, \ldots, 52\}$. Cards must be placed on these piles in order. To begin, the shuffled deck is placed face down, and the top card is drawn. If the card immediately follows the top card of either Pile 1 or Pile 2 in sequence, it is placed on the corresponding pile. If the card cannot be placed on either Pile 1 or Pile 2, it is placed onto Pile 3. This process continues until all cards from the deck have been drawn and placed. Once Pile 3 is formed, the cards in Pile 3 are run through using the same rules, attempting to build onto Pile 1 or Pile 2. The game ends when either Pile 1 or Pile 2 is completed. Completing Pile 1 constitutes a win, whilst completing Pile 2 constitutes a loss. They investigated the probability of winning when a deck following the new deck order was shuffled using the GSR shuffle, ultimately proving that with 7 GSR riffle shuffles, the probability of winning was 81 percent, far from the expected 50 percent win rate in a perfectly shuffled deck.

Prior research has also addressed the problem concerning the effect of shuffling on multi-decks with repeated cards. In 2009, Diaconis, Assaf and Soundararajan [1] analyzed the effect of shuffling on a more general deck of cards. In this n-card deck, there are D_1 cards labelled 1, D_2 cards labelled 2, …, D_m cards labelled m. They were able to present the explicit formulae on the separation distance of the deck when it started from the initial ordering $1, \ldots, 1, 2, \ldots, 2, \ldots, m, \ldots, m$ after an a-shuffle, a concept we will introduce later in this section. They analyzed the effect of shuffling on games where cards are either indistinguishable or where distinguishable cards function in the same way. Notable examples include the blackjack deck, where cards with the same value (regardless of suit) function identically, and all cards valued 10 or above are interchangeable, and the red-black deck, where the sole distinguishing feature is card color (red or black), with no other meaningful differences between cards. They were able to show that the separation distance of a standard 52-card blackjack deck is about 0.970 after 6 riffle shuffles, which is still high, and the separation distance of a 52-card red-black (a deck with only 2 colors) deck is about 0.317 after 6 riffle shuffles.

1.2. **Basics of Shuffling.**

We first start with a single deck of n distinct cards, with the initial order of $(1, 2, \ldots, n)$. Shuffling represents a function that maps one permutation of cards to another. This is done with the aim of randomizing the deck of playing cards. After a shuffle, each permutation is given a fixed probability of occurring.

In this paper, we will consider the riffle shuffle, which can be modeled using the Gilbert-Shannon-Reeds (GSR) model of riffle shuffle. It may be defined in the following way:

The GSR Shuffle: The n-card deck is cut into two packets, Packet 1 and Packet 2, with Packet 1 containing m cards and Packet 2 containing $n-m$ cards. The way that the deck is split is based upon the binomial density, that is the probability of Packet 1 containing exactly k cards is given by

$$\frac{\binom{n}{k}}{2^n}$$

Then, the cards are dropped from either packet one by one in order. The method of dropping the cards follows the following probabilistic procedure: If Packet 1 contains t cards, and Packet 2 contains s cards, the probability of the card being from Packet 1 is $\dfrac{t}{t+s}$ and the probability of the card being from Packet 2 is $\dfrac{s}{t+s}$. This procedure is repeated until all the cards are dropped from both packets.

Another frequently considered alternative to the GSR shuffle is the *inverse riffle shuffle*, which can be described as follows:

The Inverse Riffle Shuffle: This is when a n-card deck is laid out in order, to be split into two packets. Cards are placed into either Packet 1 or Packet 2 one by one, preserving the order, with equal probability. This procedure is done for all the cards. Following this procedure, the deck is reorganized by placing the two packets together, with Packet 2 following Packet 1.

A generalized form of the GSR shuffle called the *a-shuffle* was introduced in [2], where $a \geq 2$. It generalizes the two packets in the GSR shuffle into a packets, performing essentially the same process. More precisely, it can be described as follows:

The a-Shuffle: The n-card deck is cut into a packets $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_a$ of size $k_1, k_2, \dots, k_a$, where $k_i \geq 0$, according to the multinomial distribution of

$$M(k_1, k_2, \dots, k_a) = \frac{1}{a^n} \binom{n}{k_1, k_2, \dots, k_a},$$

where $M(k_1, k_2, \dots, k_n)$ is the probability of obtaining packets with sizes

$$k_1, k_2, \dots, k_a, \text{ with}$$

$\mathcal{P}_1 = (1, \dots, k_1), \mathcal{P}_2 = (k_1 + 1, \dots, k_1 + k_2), \dots, \mathcal{P}_a = (k_1 + \cdots + k_{a-1} + 1, \dots, n)$, maintaining the order of cards in each packet. The cards are then dropped from the packets one by one in order through the following probabilistic procedure: Let $\mathcal{P}_i$ contain t_i cards for all $i \in \{1, 2, \dots, a\}$, the probability that the next card dropped is from $\mathcal{P}_i$ is $\dfrac{t_i}{t_1 + t_2 + \cdots + t_a}$. This procedure is repeated until all n cards have been dropped.

Similarly we may define the inverse a-shuffle as follows:

The Inverse a-Shuffle: This is when an n-card deck is laid out in order, to be split into a packets $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_a$. Each card is pulled into one of $\mathcal{P}_i$ for $i \in \{1, 2, \dots, a\}$ with equal probability while preserving the order. Note here that some packets can be empty. This procedure is done for all n cards. Following this procedure, the deck is reorganized by placing the a packets together, with $\mathcal{P}_{i+1}$ following $\mathcal{P}_i$ for all $i \in \{1, 2, \dots, a-1\}$.

An easier way to look at this shuffle is to assign each element of the initial ordering one of a colors, $C_1, C_2, \dots, C_a$, with each color being equally likely to be assigned. Cards with the same assigned color are then grouped together with the order preserved, and each group of cards is stacked together with cards colored C_{i+1} following cards colored C_i for all $i \in \{1, 2, \dots, a-1\}$, forming the shuffled deck.

Another equivalent definition of the shuffle is where the n card deck is laid out in order, and are to be split into a packets $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_a$ of fixed sizes $k_1, k_2, \dots, k_a$, according to the multinomial distribution probability

$$M(k_1, k_2, \dots, k_a) = \frac{1}{a^n} \binom{n}{k_1, k_2, \dots, k_a}, \text{ where } k_i \geq 0$$

The cards are then pulled one by one into one of $\mathcal{P}_i$ for $i \in \{1, 2, \dots, a\}$ preserving the order, such that the probability of the card being pulled into

$$\mathcal{P}_i \text{ is } \frac{t_i}{t_1 + t_2 + \cdots + t_a}$$

where t_i is the number of remaining unfilled places in $\mathcal{P}_i$ for all $i \in \{1, 2, \dots, a\}$. The packets are then grouped and stacked such that $\mathcal{P}_{i+1}$ follows $\mathcal{P}_i$ for all $i \in \{1, 2, \dots, a-1\}$.

These definitions will be used in later proofs depending on which one is more convenient for the given situation.

The following results for a single deck were proved [2].

Proposition 1.1. Consider the two shuffle models, the a-shuffle and the inverse a-shuffle, we have the following properties:

1 In both shuffle models, an a-shuffle followed by a b-shuffle is equivalent to an ab-shuffle. Hence s GSR shuffles are equivalent to a 2^s-shuffle.
2 The two shuffle models, the a-shuffle model and the inverse a-shuffle model generate the same permutation distribution. In other words, the probability of generating π under an a-shuffle is equal to the probability of generating π^{-1} under an inverse a-shuffle.
3 The probability that an a-shuffle will result in the permutation π is

$$\frac{\binom{a+n-r}{n}}{a^n}$$

where r is the number of rising sequences in π.

To measure the randomness of a deck after a shuffle, the concept of total variation distance was introduced in [2]. Let S_n be the symmetric group containing all the permutations of n elements. Let Q be a probability distribution on S_n and U be the uniform distribution on S_n, then the *total variation distance* of Q from U is defined as

$$\|Q-U\| = \max_{A\subset S_n} |Q(A)-U(A)|$$

where $U(A) = \sum_{\pi\in A} U(\pi)$ and $Q(A) = \sum_{\pi\in A} Q(\pi)$. It is easy to see that

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

In [2] the paper provided estimates for the total variation distance of a deck of n cards after s GSR shuffles. In particular, after 7 shuffles on a 52-card deck, the total variation distance falls below $\frac{1}{2}$.
Another metric for measuring randomness is the *separation distance*, which was introduced in [1]. The separation distance of a deck after an a-shuffle is defined as

$$\text{SEP}(a) = \max_{\pi\in S_n}\left(1-\frac{P_a(\pi)}{U(\pi)}\right) = 1-\frac{\min P_a(\pi)}{U(\pi)}.$$

where $P_a(\pi)$ is the probability of obtaining the permutation π after an a-shuffle. Obviously this definition can be applied to other problems involving the measurement of randomness, including other types of shuffles such as inverse a-shuffle. In essence it measures the "worst case" scenario so it is a much more stringent measurement. This could be seen from the results in [1]. For example, while 7 shuffles will take the total variation distance to below 0.5 for a deck of 52 cards, it will take 11 shuffles to lower the separation distance to below 0.5. In fact, after 7 shuffles the separation distance remains to be $0.999999\cdots$, meaning that the least likely permutation of the deck has essentially probability 0 to occur. Although there is no direct correlation between the two measures, it is fairly obvious from computation that separation distance takes many more shuffles to lower.
In Section 2 of this paper, we shall compute both the total variation distance and the separation distance for the shuffling of 2-decks. In Section 3 we tackle a completely new question, namely the randomness of a player's hand after shuffling.

To our surprise this question hasn't been studied before, as it seems to be a very natural question. We shall compute the separation distance of the distribution of a player's hand after shuffling, for both the single and 2-deck setting. The results show that randomization of a hand takes fewer shuffles to achieve from the angle of separation distance.

2. Multi-Deck Shuffling: Two Decks

We begin by introducing the concept of a *multiset*, which is a set that allows for multiple instances of each element. For example, $\{1, 1, 2, 3, 3, 3\}$ is a multiset. In this paper, we will focus on the multiset $Z_{n,m}$, which is defined as

$$\{1, 1, \ldots, 1\,(m \text{ times}), 2, 2, \ldots, 2\,(m \text{ times}), \ldots, n, n, \ldots, n\,(m \text{ times})\}.$$

In this section, we will focus on the case of $m = 2$, i.e.

$$Z_{n,2} = \{1, 1, 2, 2, \ldots, n, n\}.$$

Let $S_{n,m}$ be the set of all permutations (ordered sequence) of $Z_{n,m}$. For example,

$$S_{2,2} = \Big\{(1, 1, 2, 2), (1, 2, 1, 2), (1, 2, 2, 1), (2, 1, 1, 2), (2, 1, 2, 1), (2, 2, 1, 1)\Big\}.$$

It is easy to see that the cardinality of $S_{n,m}$ is

$$|S_{n,m}| = \frac{(mn)!}{(m!)^n}$$

It is sometimes more concise to write a permutation without parentheses, e.g. $(1, 1, 2, 2)$ as 1122. In this paper we will freely interchange the two notations. Therefore, under this more concise notation, we would write

$$S_{2,2} = \Big\{1122, 1212, 1221, 2112, 2121, 2211\Big\}.$$

In this section, we study the effect of shuffling on the randomnness of the multi-deck $Z_{n,m}$, with $m = 2$. We will focus exclusively on inverse a-shuffles, as it is an easier model to analyze. We analyze the randomness using total variation distance. Interestingly, unlike in a single deck, the total variation distance of a shuffled multideck is dependent on the initial ordering of the deck, which we will illustrate later in the section. In our section, we will assume that the initial ordering of the deck $Z_{n,2}$ before shuffling is

$$\pi_0^* = (1, 1, 2, 2, \ldots, n, n).$$

To tackle this problem, we start by introducing two key concepts: *descents*, which are when an element in the permutation is immediately followed by a smaller element. and *sticky pairs*, which are when two non-distinguishable cards are adjacent in a deck. For example, suppose we have the permutation 34241123 in $S_{4,2}$. Here there will be 2 descents, 34|24|1123, marked by the two bars respectively and 1 sticky pair, 3424**11**23.

Theorem 2.1. Suppose we start with the initial ordering π_0^*, and let $\pi \in S_{n,2}$. Then the probability that an inverse a-shuffle results in the permutation π of $Z_{n,2}$ is only dependent on the number of sticky pairs and descents in π, and this probability is

$$P_a(\pi) = Q_{n,d,k} = \frac{1}{a^{2n}} \sum_{r=0}^{k} 2^{n-k+r} \binom{k}{r} \binom{2n+a-d-k-1}{a-d-r-1} \tag{2.1}$$

where $a > 1$, k is the number of sticky pairs in π and d is the number of descents in π.

Proof. We shall be using the last equivalent definition of an inverse a-shuffle here, i.e. the one involving using the multinomial distribution to first determine the size of each packet. An inverse a-shuffle results in a packets of ordered elements of $Z_{n,2}$ being placed one after another to form a permutation of $Z_{n,2}$. To obtain π, we will cut π into a packets and examine the probability of getting these packets after the shuffle. Before proving the general case, let us consider an example, where $a = 3$, $n = 3$ and $\pi = 123312$.

If we cut π into $1|23|312$, we will quickly realise that this cannot be obtained because the numbers within each packet must be non-descending due to the initial permutation being non-descending. More generally, any cut with a descent within a packet is impossible to obtain. Thus the only possible cuts are

$$|1233|12, \quad 1|233|12, \quad 12|33|12, \quad 123|3|12,$$
$$1233||12, \quad 1233|1|2, \quad 1233|12|,$$

where each set of cuts corresponds to three packets, with each packet having a different probability of being obtained.

To examine the probability of obtaining each packet, we first make elements of $Z_{n,2}$ distinguishable, which can be done by adding 0.5 to each of the second duplicated cards. So now our initial ordering is

$$(1, 1.5, 2, 2.5, 3, 3.5, \ldots, n, n + 0.5)$$

We consider the cut $1|233|12$. Note that if the duplicate elements are in different packets then adding 0.5 to either of the duplicate elements does not result in a new descent. Therefore the cut $1|233|12$ corresponds to 4 distinct orderings

$$\big(1|2\,3\,3.5|1.5\,2.5\big),\ \big(1.5|2\,3\,3.5|1\,2.5\big),\ \big(1|2.5\,3\,3.5|1.5\,2\big),\ \big(1.5|2.5\,3\,3.5|1\,2\big).$$

Also note that if the two duplicate elements are in the same packet, they must form a sticky pair. In this case, such as here when the two 3's are in the same packet, we can only make the second 3 into 3.5. In other words we cannot interchange them. Take any one of the above cuts, it will have packet sizes 1, 3 and 2, which has probability

$$\frac{\binom{6}{1,3,2}}{3^6}$$

of occurring by the multinomial distribution. As shown in [2], there will be $\binom{6}{1,3,2}$ ways to drop the (distinguishable) cards in their respective packets, hence the probability of this particular arrangement occurring from this cut is

$$\frac{\binom{6}{1,3,2}}{3^6} \cdot \frac{1}{\binom{6}{1,3,2}} = \frac{1}{3^6}.$$

Now since there are 4 cases, the probability of obtaining 1|233|12 for the multi-deck setting is $\frac{4}{3^6}$.
We will see that this result stands for all cuts of 123312 with 2 pairs of duplicated elements residing in different packets.
Now let's consider another cut $\pi = 123|3|12$. Since adding 0.5 to elements in different packets does not result in a descent, the cut 123|3|12 will correspond to 8 distinct orderings

$$\begin{gathered}(1.5\,2.5\,3.5|3|1\,2),\ (1.5\,2.5\,3|3.5|1\,2),\ (1.5\,2\,3.5|3|1\,2.5),\ (1.5\,2\,3|3.5|1\,2.5),\\ (1\,2.5\,3.5|3|1.5\,2),\ (1\,2.5\,3|3.5|1.5\,2),\ (1\,2\,3.5|3|1.5\,2.5),\ (1\,2\,3|3.5|1.5\,2.5).\end{gathered}$$

Take any one of the above cuts, it will have packet sizes 3, 1 and 2, which have probability

$$\frac{\binom{6}{3,1,2}}{3^6}$$

of occurring by the multinomial distribution. As shown in [2], there will be $\binom{6}{3,1,2}$ ways to drop the (distinguishable) cards in their respective packets, hence the probability of this particular arrangement occurring from this cut is

$$\frac{\binom{6}{3,1,2}}{3^6} \cdot \frac{1}{\binom{6}{3,1,2}} = \frac{1}{3^6}.$$

Since there are 8 cases, the probability of obtaining 123|3|12 in the multi-deck setting is $\frac{8}{3^6}$.
Hence the overall probability of obtaining the permutation $\pi = 123312$ is

$$P_3(\pi) = \frac{4}{3^6} \cdot 6 + \frac{8}{3^6} = \frac{32}{3^6}.$$

The more general case works in the exact same way as the above example. Let π be a permutation in the set $S_{n,2}$ with d descents and k sticky pairs. We can find the probability of π occurring after an inverse a-shuffle almost identically. As seen in the example, the numbers in each packet must be non-descending, hence each of the d descents must be split by a cut. This takes up a total of d cuts, leaving $a-d-1$ cuts that are free. However, as seen in the example, we must also take into account sticky pairs, as a sticky pair split by a cut increases the number of counts in comparison to those that are not split. Let r be the number of sticky pairs split by cuts, this leaves $a-1-d-r$ cuts that are free. Since $k-r$ sticky pairs cannot have cuts in between them, this allocates $a-d-r-1$ cuts to $2n-k+r$ cards. Using the standard stars and bars argument, it gives us $\binom{2n+a-d-k-1}{a-d-r-1}$ ways to allocate the $a-d-r-1$ cuts. Note that we have $\binom{k}{r}$ ways to choose the r sticky pairs, this gives us a total of $\binom{k}{r}\binom{2n+a-d-k-1}{a-d-r-1}$ ways to allocate cuts with r sticky pairs being split. Exactly the same as in the previous examples, for each of the cut arrangements with r sticky pairs being split, the probability of cards dropping into the packets formed by the cuts is

$$\frac{2^{n-k+r}}{a^{2n}},$$

where $n-k+r$ represents the number of duplicated cards that can be interchanged. Therefore, the probability of obtaining the permutation π after an inverse a-shuffle is

$$\frac{1}{a^{2n}}\sum_{r=0}^{k} 2^{n-k+r}\binom{k}{r}\binom{2n+a-d-k-1}{a-d-r-1}.$$

■

Corollary 2.2. We have

$$\lim_{a\to\infty}\frac{1}{a^{2n}}\sum_{r=0}^{k} 2^{n-k+r}\binom{k}{r}\binom{2n+a-d-k-1}{a-d-r-1} = \frac{2^n}{(2n)!}.$$

Therefore as we increase a, the probability distribution tends to the uniform distribution on $S_{n,2}$.

Proof. Taking the limit to each term, we evaluate

$$\lim_{a\to\infty}\frac{1}{a^{2n}}2^{n-k+r}\binom{k}{r}\binom{2n+a-d-k-1}{a-d-r-1}.$$

Note that

$$\begin{aligned}\binom{2n+a-d-k-1}{a-d-r-1} &= \binom{2n+a-d-k-1}{2n-k+r}\\ &= \frac{1}{(2n-k+r)!}\prod_{j=1}^{2n-k+r}(a+2n-d-k-j).\end{aligned}$$

As n, d and k are all constants,

$$\lim_{a\to\infty}\frac{\prod_{j=1}^{2n-k+r}(a+2n-d-k-j)}{a^{2n}} = \lim_{a\to\infty}\frac{a^{2n-k+r}}{a^{2n}}.$$

It is easy to see that for all $r<k$, this term has value 0, hence

$$\lim_{a\to\infty}\frac{1}{a^{2n}}\sum_{r=0}^{k} 2^{n-k+r}\binom{k}{r}\binom{2n+a-d-k-1}{a-d-r-1} = \frac{2^n}{(2n)!}$$

■

In the proof of Theorem 2.1, we have split a permutation into a packets by inserting $a-1$ bars into the permutation. We shall refer to each bar as a cut, and the sequence of cuts that divides the permutation into a packets (hence there are $a-1$ bars) as an *a-cut.* Given a permutation $\pi \in S_{n,2}$, we call an a-cut of π *eligible* if for every descent in π, there is a cut (bar) separating the descent. In other words, if (c_1, c_2) are adjacent elements in π with $c_1 > c_2$, then there must be a cut between c_1 and c_2. Note that a $\pi \in S_{n,2}$ can be obtained through an inverse a-shuffle if and only if it has an eligible a-cut.
We now compute the total variation distance on the permutation distribution on $Z_{n,2}$ from the uniform distribution after an inverse a-shuffle. Let $\mathcal{A}(n,d,k)$ be the set of all permutations in $S_{n,2}$ with d descents and k sticky pairs and $F(n,d,k) = |\mathcal{A}(n,d,k)|$ be the cardinality of $\mathcal{A}(n,d,k)$.

Lemma 2.3. Starting with the initial ordering of π_0^*, the total variation distance of the permutation distribution on the multi-deck $Z_{n,2}$ from the uniform distribution after an inverse a-shuffle is

$$\frac{1}{2}\sum F(n,d,k)\left|Q_{n,d,k} - \frac{2^n}{(2n)!}\right|$$

where $Q_{n,d,k}$ is given in (2.1).

Proof. Again let $P_a(\pi)$ be the probability of obtaining the permutation π after an a-shuffle on the initial ordering π_0^*. Total variation distance is defined as

$$\frac{1}{2}\sum_{\pi\in S_{n,2}}\left|P_a(\pi) - \frac{2^n}{(2n)!}\right|.$$

The lemma clearly follows from (2.1). ■

We now only need to compute $F(n,d,k)$, which can be calculated from the following recursive formula.

Theorem 2.4. The function $F(n,d,k)$ satisfies:

$$\begin{aligned} F(n,d,k) &= C_1F(n-1,d,k) + C_2F(n-1,d,k-1)\\ &\quad + C_3F(n-1,d-1,k+1) + C_4F(n-1,d-1,k)\\ &\quad + C_5F(n-1,d-1,k-1) + C_6F(n,d-2,k+2)\\ &\quad + C_7F(n-1,d-2,k+1) + C_8F(n-1,d-2,k), \end{aligned} \tag{2.2}$$

where:

$$\begin{aligned} &C_1 = \binom{d+1}{2}, \quad C_2 = d+1, \quad C_3 = d(k+1),\\ &C_4 = d(2n-d-k-1)+k, \quad C_5 = 2n-d-k,\\ &C_6 = \binom{k+2}{2}, \quad C_7 = (k+1)(2n-k-d-1),\\ &C_8 = \binom{2n-d-k}{2}. \end{aligned}$$

Proof. We examine ways to add two n's into permutations in $S_{n-1,2}$ to make them into permutations in $S_{n,2}$ with d descents and k sticky pairs.
For this proof, we will introduce the term *generalized descent position*, which refers to the very last position, or a position between a descent i.e. between two elements c_1, c_2 s.t. $c_1 > c_2$. Therefore, the number of generalized descent positions for all elements of $\mathcal{A}(n, d, k)$ is $d + 1$. When the element n is inserted into a generalized descent position, no additional descent will form. Similarly, the term *generalized ascent position* refers to the very first position, or a position between an ascent i.e. between two elements c_1, c_2 s.t. $c_1 < c_2$. This denotes a position where after an n is inserted, a descent is formed without splitting a sticky pair. Hence, the number of generalized ascents for all permutations of $\mathcal{A}(n, d, k)$ is $2n - d - k$ as it includes all the ascents as well as the position at the start but doesn't include the positions between the k sticky pairs.

To find C_1, note that after the two n's are inserted, no extra sticky pairs or descents are created. This means the two n's must be inserted in separate generalized descent positions. Hence
$$C_1 = \binom{d+1}{2}.$$
To find C_2, note that after the two n's are inserted, one extra sticky pair is created without an extra descent forming. This means the two n's must form a sticky pair and be inserted in a generalized descent position. Hence
$$C_2 = d + 1.$$
To find C_3, note that after the two n's are inserted, one sticky pair is split while one extra descent is formed. This means one n must be inserted between a sticky pair, and the other n must be inserted in a generalized descent position. Hence
$$C_3 = d(k + 1).$$
To find C_4, note that after the two n's are inserted, one extra descent is formed. This means either one n is in a generalized descent position and one is in a generalized ascent position, or the two n's form a sticky pair between another sticky pair. Hence
$$C_4 = d(2n - d - k - 1) + k.$$
To find C_5, note that after the two n's are inserted, one extra descent and one extra sticky pair are created. This means the two n's must form a sticky pair and inserted in a generalized ascent position. Hence
$$C_5 = 2n - d - k.$$
To find C_6, note that after the two n's are inserted, two sticky pairs are lost and two extra descents are formed. This means the two n's must separately split two distinct sticky pairs. Hence
$$C_6 = \binom{k+2}{2}.$$
To find C_7, note that after the two n's are inserted, one sticky pair is lost and two extra descents are formed. This means one n must split a sticky pair, and the other

n must reside in a generalized ascent position. Hence

$$C_7 = (k+1)(2n-k-d-1).$$

To find C_8, note that after the two n's are inserted, two descents are obtained. This means both n's must be inserted in two distinct generalized ascent positions. Hence

$$C_8 = \binom{2n-d-k}{2}.$$

This proves the theorem. ■

Theorem 2.1 and Theorem 2.4 allow us to easily compute the total variation distance for the permutation distribution of $Z_{n,2}$ after an inverse a-shuffle for any a and n. Table 1 shows the total variation distance for a deck of n cards after s consecutive inverse riffle shuffles (inverse 2-shuffles) for various n and s, with $a = 2^s$.

TABLE 1. Total variation distance after s inverse riffle shuffles on double deck $Z_{n,2}$ for $n = 13, 16, 26, 39, 52, 54, 104$ and 156.

s \ $2n$	26	32	52	78	104	108	208	312
3	0.9990	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0.7718	0.9112	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0.4411	0.5775	0.9165	0.9994	1.0000	1.0000	1.0000	1.0000
6	0.2261	0.3087	0.5987	0.8833	0.9862	0.9913	1.0000	1.0000
7	0.1150	0.1573	0.3234	0.5597	0.7677	0.7954	0.9995	1.0000
8	0.0574	0.0791	0.1641	0.2990	0.4479	0.4710	0.9105	0.9984
9	0.0286	0.0395	0.0828	0.1516	0.2333	0.2464	0.6007	0.8806
10	0.0143	0.0197	0.0414	0.0765	0.1176	0.1243	0.3265	0.5619

One can compare the difference between the total variation distance between shuffling a single deck (shown in Table 2, from [2]) and a 2-deck. Notice that with the same number of cards, the total variation distance for 2-deck is always smaller than that of for a single deck, which is not surprising. For example, for a single deck of 52 cards, the total variation distance after 7 shuffles is 0.334, while for a 2-deck containing 52 cards ($n = 26$), the total variation distance after 7 shuffles is 0.323. Of particular interest is the case of $n = 54$, which is the case for the popular Chinese card games such as Guandan, it takes 8 shuffles to lower the total variation distance to below 0.5.

Theorem 2.5. Let $\pi_1, \pi_2 \in S_{n,2}$. If π_1 and π_2 have the same number of sticky pairs, but π_1 has fewer descents than π_2, then $P_a(\pi_1) > P_a(\pi_2)$.

Proof. Let π_1 be a permutation with d descents and k sticky pairs. Let π_2 be a permutation with $d+1$ descents and k sticky pairs. We prove that $P_a(\pi_1) > P_a(\pi_2)$

TABLE 2. Total variation distance after s inverse riffle shuffles on single deck $Z_{n,1}$ for $n = 25, 32, 52, 78, 104, 208$ and 312 distinct cards [2].

s \ $2n$	25	32	52	78	104	208	312
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0.775	0.929	1.000	1.000	1.000	1.000	1.000
5	0.437	0.597	0.924	1.000	1.000	1.000	1.000
6	0.231	0.322	0.614	0.893	0.988	1.000	1.000
7	0.114	0.164	0.334	0.571	0.772	1.000	1.000
8	0.056	0.084	0.167	0.307	0.454	0.914	1.000
9	0.028	0.042	0.085	0.153	0.237	0.603	0.883
10	0.014	0.021	0.043	0.078	0.119	0.329	0.565

According to Theorem 2.1,

$$P_a(\pi_1) = \frac{1}{a^{2n}} \sum_{r=0}^{k} 2^{n-k+r} \binom{k}{r} \binom{2n+a-k-d-1}{2n-k+r}.$$

Similarly,

$$P_a(\pi_2) = \frac{1}{a^{2n}} \sum_{r=0}^{k} 2^{n-k+r} \binom{k}{r} \binom{2n+a-k-d-2}{2n-k+r}.$$

Since

$$\binom{2n+a-k-d-1}{2n-k+r} > \binom{2n+a-k-d-2}{2n-k+r},$$

it follows that

$$\frac{1}{a^{2n}} \sum_{r=0}^{k} 2^{n-k+r} \binom{k}{r} \binom{2n+a-k-d-1}{2n-k+r}$$
$$> \frac{1}{a^{2n}} \sum_{r=0}^{k} 2^{n-k+r} \binom{k}{r} \binom{2n+a-k-d-2}{2n-k+r}.$$

Thus, $P_a(\pi_1) > P_a(\pi_2)$. ■

Theorem 2.6. Let $\pi_1, \pi_2 \in S_{n,2}$. If π_1 and π_2 have the same number of descents, but π_1 has fewer sticky pairs than π_2, then $P_a(\pi_1) > P_a(\pi_2)$.

Proof. Let π_1 be a permutation with d descents and k sticky pairs, which we call $SP_1, SP_2, \ldots, SP_k$. Let π_2 be a permutation with d descents and $k+1$ sticky pairs, which we call $SP'_1, SP'_2, \ldots, SP'_{k+1}$. We prove that $P_a(\pi_1) > P_a(\pi_2)$.
We first consider an eligible a-cut C on π_1, which splits the permutation into packets of sizes $x_1, \ldots, x_a$, where $\sum_{i=1}^{a} x_i = 2n$. The probability of obtaining π_1

through C after an a-shuffle is

$$\frac{A_C(\pi_1)}{\binom{2n}{x_1, x_2, \ldots, x_a}},$$

where $A_C(\pi_1)$ is the number of ways to drop all $2n$ cards into the packets such that it forms the permutation π_1. Note that $A_C(\pi_1) = 2^{n-k+r}$ where r is the number of sticky pairs split by C. The probability of getting C is

$$\frac{\binom{2n}{x_1, \ldots, x_a}}{a^{2n}}$$

so the total probability of π_1 via C is $\dfrac{A_C(\pi_1)}{a^{2n}}$.
Thus, the overall probability of π_1 is

$$P_a(\pi_1) = \frac{1}{a^{2n}} \sum_{C \in \mathcal{F}_{\pi_1}} A_C(\pi_1),$$

where $\mathcal{F}_{\pi_1}$ is the set of eligible a-cuts for π_1. Partition $\mathcal{F}_{\pi_1}$ into subsets $\mathcal{E}_1, \ldots, \mathcal{E}_k$, where

$$\mathcal{E}_r = \{C \in \mathcal{F}_{\pi_1} \mid C \text{ splits } r \text{ sticky pairs in } \pi_1\}.$$

Similarly, partition $\mathcal{F}_{\pi_2}$ into $\mathcal{D}_1, \ldots, \mathcal{D}_k$, where

$$\mathcal{D}_r = \{C \in \mathcal{F}_{\pi_2} \mid C \text{ splits } r \text{ pairs among } SP'_1, \ldots, SP'_k\}.$$

Hence,

$$P_a(\pi_1) = \frac{1}{a^{2n}} \sum_{i=1}^{k} \sum_{C \in \mathcal{E}_i} A_C(\pi_1), \quad P_a(\pi_2) = \frac{1}{a^{2n}} \sum_{i=1}^{k} \sum_{C \in \mathcal{D}_i} A_C(\pi_2).$$

We compare $\sum_{C \in \mathcal{E}_i} A_C(\pi_1)$ and $\sum_{C \in \mathcal{D}_i} A_C(\pi_2)$. By Theorem 2.1,

$$\sum_{C \in \mathcal{E}_i} A_C = 2^{n-k+i} \binom{k}{i} \binom{2n+a-d-k-1}{a-d-i-1}.$$

For $\sum_{C \in \mathcal{D}_i} A_C$, consider two cases:

Case 1: SP'_{k+1} is not split. Here, i of $k+1$ pairs are split, yielding:

$$2^{n-k+i-1} \binom{k}{i} \binom{2n+a-d-k-2}{a-d-i-1}.$$

Case 2: SP'_{k+1} is split. Now, $i+1$ pairs are split, giving:

$$2^{n-k+i} \binom{k}{i} \binom{2n+a-d-k-2}{a-d-i-2}.$$

Combining both cases:

$$\sum_{C \in \mathcal{E}_i} A_C(\pi_1) = 2^{n-k+i-1} \binom{k}{i} \left[\binom{2n+a-d-k-2}{a-d-i-1} + 2\binom{2n+a-d-k-2}{a-d-i-2} \right]$$

$$= 2^{n-k+i-1}\binom{k}{i}\left[\binom{2n+a-d-k-1}{a-d-i-1}+\binom{2n+a-d-k-2}{a-d-i-2}\right].$$

Since

$$\binom{2n+a-d-k-1}{a-d-i-1} > \binom{2n+a-d-k-2}{a-d-i-2},$$

we have:

$$2\binom{2n+a-d-k-1}{a-d-i-1} > \binom{2n+a-d-k-1}{a-d-i-1}+\binom{2n+a-d-k-2}{a-d-i-2},$$
$$\Longrightarrow 2^{n-k+i}\binom{k}{i}\binom{2n+a-d-k-1}{a-d-i-1} > \sum_{C\in\mathcal{D}_i} A_C(\pi_2).$$

Thus, $\sum_{C\in\mathcal{E}_i} A_C(\pi_1) > \sum_{C\in\mathcal{D}_i} A_C(\pi_2)$, and therefore $P_a(\pi_1) > P_a(\pi_2)$. ■

Another important concept involves modifying elements of a sticky pair to determine whether a bar should pass through it. This can be achieved by adding 0.5 to one element. If the first element becomes larger than the second by 0.5, the pair functions as a descent, thus a cut must be placed in between them. Conversely, if the second element becomes larger, the pair does not constitute a descent, therefore no cut is necessary to be placed in between. This property is useful in the proof below.

Theorem 2.7. Let $\pi_1, \pi_2 \in S_{n,2}$. If π_1 and π_2 have the same total number of descents and sticky pairs, but π_1 has fewer descents than π_2, then $P_a(\pi_1) > P_a(\pi_2)$.

Proof. Let π_1 be a permutation with d descents and $k+1$ sticky pairs, namely $SP_1, SP_2, \ldots, SP_{k+1}$. Let π_2 be a permutation with $d+1$ descents and k sticky pairs, namely $SP_1', SP_2', \ldots, SP_k'$. We prove that $P_a(\pi_1) > P_a(\pi_2)$.
By Theorem 2.1, we have

$$P_a(\pi_2) = \frac{1}{a^{2n}}\sum_{r=0}^{k} 2^{n-k+r}\binom{k}{r}\binom{2n+a-d-k-2}{a-d-r-2}.$$

To solve for $P_a(\pi_1)$, without loss of generality, let the values of SP_{k+1} be mm. We are able to turn one of the elements making up the sticky pair SP_{k+1} into $m+\frac{1}{2}$, giving two cases.
Case 1: $m+\frac{1}{2}$ comes before m. In this case, $m+\frac{1}{2} \to m$ acts as a descent, giving a total of $d+1$ descents and k non-fixed sticky pairs. By Theorem 2.1, this gives the probability

$$\frac{1}{a^{2n}}\sum_{r=0}^{k} 2^{n-k+r-1}\binom{k}{r}\binom{2n+a-d-k-2}{a-d-r-2}.$$

Case 2: m comes before $m + \frac{1}{2}$. In this case, there will be d descents and k non-fixed sticky pairs. This gives the probability

$$\frac{1}{a^{2n}} \sum_{r=0}^{k} 2^{n-k+r-1} \binom{k}{r} \binom{2n+a-d-k-1}{a-d-r-1}.$$

Combining the cases, we have

$$\begin{aligned} P_a(\pi_1) &= \frac{1}{a^{2n}} \sum_{r=0}^{k} 2^{n-k+r-1} \binom{k}{r} \binom{2n+a-d-k-2}{a-d-r-2} \\ &+ \frac{1}{a^{2n}} \sum_{r=0}^{k} 2^{n-k+r-1} \binom{k}{r} \binom{2n+a-d-k-1}{a-d-r-1}. \end{aligned}$$

Since

$$\binom{2n+a-d-k-1}{a-d-r-1} > \binom{2n+a-d-k-2}{a-d-r-2},$$

we conclude

$$\begin{aligned} P_a(\pi_1) &> \frac{2}{a^{2n}} \sum_{r=0}^{k} 2^{n-k+r-1} \binom{k}{r} \binom{2n+a-d-k-2}{a-d-r-2} \\ &= \frac{1}{a^{2n}} \sum_{r=0}^{k} 2^{n-k+r} \binom{k}{r} \binom{2n+a-d-k-2}{a-d-r-2} \\ &= P_a(\pi_2). \end{aligned}$$

■

Theorem 2.8. For the multi-deck $Z_{n,2}$ of $2n$ cards with the initial ordering π_0^*, the least likely permutation to occur after an inverse a-shuffle is

$$(n, n-1, n-2, \ldots, 1, n, n-1, n-2, \ldots, 1)$$

Proof. The permutation $(n, n-1, n-2, \ldots, 1, n, n-1, n-2, \ldots, 1)$ has $2n-2$ descents. By Theorem 2.7, all permutations with less than or equal to $2n-2$ combined descents and sticky pairs are excluded. This leaves the permutation $(n, n, n-1, n-1, \ldots, 1, 1)$, with $n-1$ descents and n sticky pairs. Let

$$\begin{aligned} \pi_1 &= (n, n-1, \ldots, 1, n, n-1, \ldots, 1) \\ \pi_2 &= (n, n, n-1, n-1, \ldots, 1, 1) \end{aligned}$$

.

We prove $P_a(\pi_1) < P_a(\pi_2)$.
By Theorem 2.1,

$$P_a(\pi_1) = \frac{2^n}{a^{2n}} \binom{a+1}{a-2n+1} = \frac{2^n}{a^{2n}} \binom{a+1}{2n},$$

and

$$P_a(\pi_2) = \frac{1}{a^{2n}} \sum_{r=0}^{n} 2^r \binom{n}{r} \binom{a}{a-n-r} = \frac{1}{a^{2n}} \sum_{r=0}^{n} 2^r \binom{n}{r} \binom{a}{n+r}.$$

Analyzing the nth and $(n-1)$th terms of $P(\pi_2)$:

$$\begin{aligned}
&2^n\binom{a}{2n} + n\cdot 2^{n-1}\binom{a}{2n-1}\\
&= 2^n\binom{a}{a-2n} + 2^n\binom{a}{2n-1} + (n-2)\cdot 2^{n-1}\binom{a}{2n-1}\\
&= 2^n\binom{a+1}{2n} + (n-2)\cdot 2^{n-1}\binom{a}{2n-1}\\
&= P_a(\pi_1) + (n-2)\cdot 2^{n-1}\binom{a}{2n-1} \geq P_a(\pi_1).
\end{aligned}$$

Hence, $P_a(\pi_2) \geq P_a(\pi_1)$, proving that π_1 is the least likely permutation. ■

We should point out that the same result holds for the a-shuffle, which was already proved in [1]. However, as we will show later, for multi-decks inverse a-shuffle is different from a-shuffle.

Conjecture 2.1. For the multi-deck $Z_{n,2}$ of $2n$ cards with the initial ordering π_0^*, the most likely permutation to occur after an inverse a-shuffle is

$$(1, 2, 3, \ldots, n, 1, 2, 3, \ldots, n)$$

By Theorem 2.5 and Theorem 2.6, a permutation with 0 descents and 0 sticky pairs would have the largest probability. However, such a permutation is impossible: the only permutation with 0 descents is $(1, 1, 2, 2, \ldots, n, n)$, which contains n sticky pairs. The next candidate is the permutation with only one descend and no sticky pairs, which is $(1, 2, 3, \ldots, n, 1, 2, 3, \ldots, n)$. By Theorem 2.7, the most likely permutation must be among these two permutations. However, to determine which one requires some complicated combinatorial analysis, which we have not been able to overcome. However, numerical analysis clearly gives edge to $(1, 2, 3, \ldots, n, 1, 2, 3, \ldots, n)$. Incidentally, Assaf, Diaconis and Soundararajan in [1] also did not give the most likely permutation after an a-shuffle. So it is possible that this problem is quite hard.

Remark: It is interesting to note that unlike in a single deck case, both the total variation distance and the separation distance after an a-shuffle or an inverse a-shuffle are dependent on the initial permutation we start with. This is seen in Table 4 for the 2-deck $Z_{2,2}$ after an inverse 2-shuffle, which lists all probability distribution after shuffling with various different starting permutations. Note there are only 6 permutations in total, and the three starting ones listed are exhaustive as the other three can be obtained by interchange 1 and 2. The analysis of distribution after shuffling for other initial conditions appear to be much harder, which we plan to study in the future. The relationship between an a-shuffle and an inverse a-shuffle is not yet well understood for multi-deck setting, this is in stark contrast to the single deck setting, where the probability distribution of both shuffles are identical.

2.1. Multi-Deck Shuffling: Generalization to m-Deck.

While our main focus of this section has been on 2-deck, some results can also be generalized to m-deck for general m. Here in this subsection we present one result.

TABLE 3. Separation distance after s inverse riffle shuffles on double deck $Z_{n,2}$ for $n = 13, 16, 26, 39, 52, 54, 104$ and 156.

s \ $2n$	26	32	52	78	104	108	208	312
3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0.9956	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	0.9181	0.9808	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8	0.7010	0.8488	0.9952	1.0000	1.0000	1.0000	1.0000	1.0000
9	0.4476	0.6034	0.9238	0.9976	1.0000	1.0000	1.0000	1.0000
10	0.2550	0.3673	0.7178	0.9466	0.9951	0.9968	1.0000	1.0000
11	0.1363	0.2036	0.4659	0.7646	0.9263	0.9403	1.0000	1.0000
12	0.0705	0.1073	0.2683	0.5126	0.7256	0.7525	0.9950	1.0000
13	0.0359	0.0551	0.1443	0.3010	0.4747	0.5010	0.9275	0.9974
14	0.0181	0.0279	0.0749	0.1637	0.2747	0.2931	0.7293	0.9482

TABLE 4. Probability distribution of each permutation after an inverse 2-shuffle from different starting permutation.

Starting permutation	**Probability distribution**					
	1122	1212	1221	2112	2121	2211
1122	$\frac{7}{16}$	$\frac{4}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{0}{16}$	$\frac{1}{16}$
1212	$\frac{3}{16}$	$\frac{6}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{1}{16}$
1221	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{6}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{2}{16}$

Other results for 2-decks are much harder to generalize, which we may explore in the future.
Define a set of $\pi-$*inducing packets* to be a set $P = \{\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_a\}$ such that the concatenation of $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_a$ equals π, and that P is achievable through the a-shuffle. Denote P to be the set of all $\pi-$*inducing packets* for a given permutation π. Define $x_{i,j}$ to be the number of occurrences of the number j in $\mathcal{P}_i$.

Theorem 2.9. The number of ways to achieve π after an inverse a-shuffle can be expressed as $\sum_{P \in \mathcal{P}(\pi)} \prod_{j=1}^{n} \binom{m}{x_{1,j}, x_{2,j}, \ldots, x_{a,j}}$ where $x_{i,j}$ varies depending on P.

Proof. For each $1 \leq j \leq n$, there are m copies of j distributed in $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_a$. We treat the m copies of j to be distinguishable, hence the number of ways to allocate the m unique j into $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_a$ is $\binom{m}{x_{1,j}, x_{2,j}, \ldots, x_{a,j}}$.

Taking for all j, we arrive at $\prod_{j=1}^{n} \binom{m}{x_{1,j}, x_{2,j}, \ldots, x_{a,j}}$ for a given set of $P = \{\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_a\}$. Hence the number of ways to achieve π after an inverse a-shuffle is $\sum_{P \in P(\pi)} \prod_{j=1}^{n} \binom{m}{x_{1,j}, x_{2,j}, \ldots, x_{a,j}}$ where $x_{i,j}$ varies depending on P. ■

3. Randomness of a Player's Hand After Riffle Shuffling

It is important to realize that in card games, the randomization of player hands is of greater significance than the randomization of the entire deck, as after all, a players hand is what matters in a card game. Despite this, the probabilistic analysis of hand distributions following a shuffle has not been studied at all. For the purposes of this analysis, a hand is defined as the collection of cards held by a player. In this paper, we analyze the randomness of players' hands consisting of the first m cards in a shuffled deck. To our knowledge, this is the first time such a problem has been studied.

Let $\mathbf{F}_{n,m}$ be the set of all subsets of {1,2,...,n} with cardinality m. This section computes the probability of obtaining any specific hand $\mathcal{E}$ from a deck of n cards, where $\mathcal{E} \in \mathbf{F}_{n,m}$, after an inverse a-shuffle.

Using this analysis, we compute the separation distance of the hands after an inverse a-shuffle.

Theorem 3.1. Starting with the initial ordering of (1,2,...,n), the probability of obtaining the hand $\mathcal{E} = \{e_1, e_2, \ldots, e_m\} \in \mathbf{F}_{n,m}$ with $e_1 < e_2 < e_3 < \cdots < e_m$ as the first m cards of the shuffled deck after an inverse a-shuffle is

$$H_{a,n}(\mathcal{E}) = \frac{1}{a^n} \sum_{L=1}^{a} \sum_{k=1}^{m} \sum_{t=1}^{m} \binom{t-1}{k-1} (L-1)^{m-k} (a-L+1)^{n-m-e_t+t} (a-L)^{e_t-t}$$

Proof. We adopt the second definition for the inverse a-shuffle in this proof. We assign one of a colors $C_1, C_2, \ldots, C_a$ to each element in the initial ordering, with equal probability.

Let L be the largest color index such that at least one of the elements of $\mathcal{E}$ is assigned to the color C_L. Let k be the number of elements within $e_1, e_2, \ldots, e_m$ assigned to the color C_L. Let $e_t \in \mathcal{E}$ be the largest element in $\mathcal{E}$ which is assigned the color C_L.

To compute the probability of obtaining the hand $\mathcal{E}$ for a fixed L, k and t, we must separately calculate the number of ways to assign the color C_L to elements within $\mathcal{E}$, the number of ways to assign the colors $C_1, C_2, \ldots, C_{L-1}$ to elements within $\mathcal{E}$, the number of ways to assign $C_{L+1}, C_{L+2}, \ldots, C_a$ to elements outside $\mathcal{E}$ that hold a smaller value than e_t, and the number of ways to assign the colors $C_L, C_{L+1}, \ldots, C_a$ to elements outside $\mathcal{E}$ that hold a larger value than e_t.

To calculate the number of ways to assign the color C_L to elements within $\mathcal{E}$, note that there are k elements in $\mathcal{E}$ assigned to the color C_L, yet each element cannot hold a value that exceeds e_t. Hence the remaining $k-1$ elements assigned the color C_L must hold indexes strictly less than t, giving a total of $\binom{t-1}{k-1}$ ways to do so.
To calculate the number of ways to assign the colors $C_1, C_2, \ldots, C_{L-1}$ to elements within $\mathcal{E}$, note that $m-k$ elements within $e_1, e_2, \ldots, e_m$ are not assigned the color C_L, giving a total of $(L-1)^{m-k}$ ways to assign the colors.
To calculate the number of ways to assign the colors $C_{L+1}, C_{L+2}, \ldots, C_a$ to elements outside $\mathcal{E}$ that hold a smaller value than e_t, note that the number of elements outside $\mathcal{E}$ smaller than e_t is $e_t - t$. Thus the total number of possibilities to assign the $a-L$ colors to $e_t - t$ elements is $(a-L)^{e_t - t}$.
To calculate the number of ways to assign the colors $C_L, C_{L+1}, \ldots, C_a$ to elements outside $\mathcal{E}$ that hold a larger value than e_t, note that there are $n-m$ total elements outside $\mathcal{E}$. To calculate the number of elements larger than e_t, we take away the elements less than or equal to e_t as a whole, and add back the elements less than or equal to e_t belonging to $\mathcal{E}$. Thus, in total there are $n-m-e_t+t$ elements satisfying the condition. As all of these elements can be assigned one of the $a-L+1$ colors of $C_L, C_{L+1}, \ldots, C_a$, the total number of colorings for this group of elements will be $(a-L+1)^{n-m-e_t+t}$.
Hence for a fixed L, k and t, the total number of colorings is

$$\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m-e_t+t}(a-L)^{e_t-t}.$$

Taking the sum across all L, k and t, we arrive at our result. ∎

Lemma 3.2. Starting with the initial ordering of $(1, 2, \ldots, n)$, the probability of obtaining the hand $\mathcal{E} = \{1, 2, \ldots, m\}$ as the first m cards of the shuffled deck after an inverse a-shuffle is

$$H_{a,n}(\mathcal{E}) = \frac{1}{a^n}\sum_{L=1}^{a}\big(L^m - (L-1)^m\big)(a-L+1)^{n-m}.$$

Proof. Note that for every element in $\mathcal{E}$, $e_t = t$. Thus by Theorem 3.1,

$$\begin{aligned}
&H_{a,n}(\mathcal{E})\\
&\quad= \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m-t+t}(a-L)^{t-t}\\
&\quad= \frac{1}{a^n}\sum_{t=1}^{m}\binom{t-1}{k-1}\sum_{L=1}^{a}\sum_{k=1}^{m}(L-1)^{m-k}(a-L+1)^{n-m}.
\end{aligned}$$

By the Hockey-stick identity,

$$\sum_{t=1}^{m}\binom{t-1}{k-1} = \binom{m}{k},$$

we have

$$H_{a,n}(\mathcal{E}) = \sum_{L=1}^{a}\sum_{k=1}^{m}\binom{m}{k}(L-1)^{m-k}(a-L+1)^{n-m}.$$

Because

$$\sum_{k=0}^{m}\binom{m}{k}(L-1)^{m-k} = (L-1+1)^m = L^m$$

we have

$$\sum_{k=1}^{m}\binom{m}{k}(L-1)^{m-k} = L^m - (L-1)^m.$$

Hence,

$$H_{a,n}(\mathcal{E}) = \sum_{L=1}^{a}\big(L^m - (L-1)^m\big)(a-L+1)^{n-m}.$$

■

Lemma 3.3. Starting with the initial ordering of $(1,2,\dots,n)$, the probability of obtaining cards $\mathcal{E} = \{n-m+1, n-m+2, \dots, n\}$ as the first m cards of the shuffled deck after an inverse a-shuffle is

$$H_{a,n}(\mathcal{E}) = \frac{1}{a^n}\sum_{L=1}^{a}(a-L)^{n-m}\big(L^m - (L-1)^m\big).$$

Proof. Note that for every element in $\mathcal{E}$, $e_t = n-m+t$. Thus by Theorem 3.1,

$$\begin{aligned}
&H_{a,n}(\mathcal{E})\\
&= \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m-(n-m+t)+t}(a-L)^{n-m+t-t}\\
&= \frac{1}{a^n}\sum_{t=1}^{m}\binom{t-1}{k-1}\sum_{L=1}^{a}\sum_{k=1}^{m}(L-1)^{m-k}(a-L)^{n-m}.
\end{aligned}$$

By the Hockey-stick identity,

$$\sum_{t=1}^{m}\binom{t-1}{k-1} = \binom{m}{k},$$

we have

$$H_{a,n}(\mathcal{E}) = \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\binom{m}{k}(L-1)^{m-k}(a-L)^{n-m}.$$

Because

$$\sum_{k=0}^{m}\binom{m}{k}(L-1)^{m-k} = (L-1+1)^m = L^m,$$

we have

$$\sum_{k=1}^{m}\binom{m}{k}(L-1)^{m-k} = L^m - (L-1)^m.$$

Hence,

$$H_{a,n}(\mathcal{E}) = \frac{1}{a^n}\sum_{L=1}^{a}\big(L^m - (L-1)^m\big)(a-L)^{n-m}.$$

■

Theorem 3.4. Assuming the initial ordering of an n card deck is $(1, 2, \ldots, n)$, after an inverse a-shuffle, the most likely hand formed by the first m cards of the shuffled deck is $\mathcal{E}_1 = \{1, 2, \ldots, m\}$, and the least likely hand formed by the first m cards of the shuffled deck is $\mathcal{E}_2 = \{n-m+1, n-m+2, \ldots, n\}$.

Proof. It is easy to see that

$$(a-L+1)^{n-m-e_t+t}(a-L)^{e_t-t} \le (a-L+1)^{n-m},$$
$$(a-L+1)^{n-m-e_t+t}(a-L)^{e_t-t} \ge (a-L)^{n-m}.$$

Hence

$$\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m-e_t+t}(a-L)^{e_t-t}$$
$$\le \sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m},$$

and

$$\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m-e_t+t}(a-L)^{e_t-t}$$
$$\ge \sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L)^{n-m}.$$

On the other hand, by Theorem 3.1, the probability of the hand $\{1, 2, \ldots, m\}$ occurring is

$$H_{a,n}(\mathcal{E}_1) = \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m}(a-L)^0$$

as $e_t = t$. Similarly, the probability of the hand $\{n-m+1, n-m+2, \ldots, n\}$ occurring is

$$H_{a,n}(\mathcal{E}_2) = \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^0(a-L)^{n-m}$$

as $e_t = n-m+t$. Hence it is easy to see that for all L, t and k,

$$\frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m-e_t+t}(a-L)^{e_t-t} \le H_{a,n}(\mathcal{E}_1),$$

and

$$\frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m-a_t+t}(a-L)^{e_t-t} \ge H_{a,n}(\mathcal{E}_2).$$

Hence our result stands. ■

Corollary 3.5. Starting from any initial ordering of an n card deck, the separation distance after an a-shuffle or an inverse a-shuffle for the probability distribution of a hand of m cards is

$$\mathrm{SEP}(a,n,m) = 1 - \frac{\binom{n}{m}}{a^n}\sum_{L=1}^{a}(a-L)^{n-m}\big(L^m - (L-1)^m\big).$$

Proof. In the single deck setting, an a-shuffle and an inverse a-shuffle yield the same probability distribution, meaning we only need to prove this statement in the case of an inverse a-shuffle. For a single deck, Separation Distance is not dependent on the initial ordering of the deck, thus without loss of generality, we assume that the initial ordering is $(1,2,\ldots,n)$. Hence the Separation Distance is

$$\begin{aligned} 1 - &\frac{\frac{1}{a^n}\sum_{L=1}^{a}(a-L)^{n-m}\big(L^m-(L-1)^m\big)}{\frac{1}{\binom{n}{m}}} \\ &= 1 - \frac{\binom{n}{m}}{a^n}\sum_{L=1}^{a}(a-L)^{n-m}\big(L^m-(L-1)^m\big). \end{aligned}$$

■

With this, we are able to compute the separation distance of the first m cards of an n-card deck after an a-shuffle for any m, n and $a = 2^s$, where s is the number of GSR riffle shuffles. Table 7 lists the separation distance in the 2-deck $Z_{n,2}$ setting, where m is fixed to be $m = n/2$. It is interesting to point out that the separation distance doesn't seem to be very sensitive to m as long as m is not too small compared to n. For example, if we pick $m = n$ instead of $n/2$, the numbers in the table would look virtually identical.

TABLE 5. Separation distance after s inverse riffle shuffles on a single deck of n cards for n = 16, 32, 52, 64, 104, 108, 128 and 208.

$s \setminus n$	16	32	52	64	104	108	128	208
4	0.373	0.573	0.708	0.758	0.848	0.853	0.876	0.922
5	0.215	0.370	0.508	0.570	0.706	0.716	0.756	0.847
6	0.116	0.214	0.317	0.369	0.507	0.518	0.569	0.705
7	0.060	0.115	0.179	0.214	0.316	0.325	0.369	0.506
8	0.031	0.060	0.095	0.115	0.178	0.184	0.213	0.316
9	0.015	0.031	0.049	0.060	0.095	0.098	0.115	0.178
10	0.008	0.015	0.025	0.031	0.049	0.051	0.060	0.095

Table 6 shows the separation distance for three popular games: Bridge, Big Two and Dou Di Zhu. Interestingly while we have no good ways to compute the total variation distance of hands after shuffling, we may speculate what it could be using the separation distance data. For example, for the standard 52 card deck, after

7 shuffles the separation distance is still essentially 1.0, and it is still larger than 0.99 after 8 shuffles. However, for a bridge hand, even after just 4 shuffles we have lowered the separation distance to 0.7080. Could this mean that ***4 shuffles suffice*** for a bridge game?

TABLE 6. Separation distance of various single deck card games

m/n s	Game Type Bridge (13/52)	 Big Two (27/54)	 Dou Di Zhu (17/54)
4	0.7080	0.7175	0.7175
5	0.5085	0.5197	0.5197
6	0.3166	0.3257	0.3257
7	0.1786	0.1846	0.1846
8	0.0951	0.0985	0.0985
9	0.0491	0.0510	0.0510
10	0.0250	0.0259	0.0259

Lemma 3.6. Assuming the initial ordering $(1, 2, \ldots, n)$, after an inverse a-shuffle on a deck of n cards, the second least likely hand formed by the first m cards of the shuffled deck is $\{n-m, n-m+2, \ldots, n\}$.

Proof. By Theorem 3.4, we know that the minimum value of $H_{a,,n}(\mathcal{E})$ is achieved at $\mathcal{E}_2 = \{n-m+1, \ldots, n\}$. Let $\mathcal{E}_3 = \{n-m, n-m+2, \ldots, n\}$, we prove that for all $\mathcal{E} \in \mathbf{F}_{n,m}$ such that $\mathcal{E} \neq \mathcal{E}_2$, $H_{a,n}(\mathcal{E}) \geq H_{a,n}(\mathcal{E}_3)$. We do this by comparing the probability between the two.

We first note that for every element in $\mathcal{E}_3$, $e_t = n-m+t$ except for e_1, which equals to $n-m$. Thus by Theorem 3.1,

$$\begin{aligned}
&H_{a,n}(\mathcal{E}_3)\\
&= \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=2}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m-(n-m+t)+t}(a-L)^{n-m+t-t}\\
&+ \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\binom{1-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m-(n-m)+1}(a-L)^{n-m-1}\\
&= \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=2}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L)^{n-m}\\
&+ \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\binom{0}{k-1}(L-1)^{m-k}(a-L+1)(a-L)^{n-m-1}.
\end{aligned}$$

Let $\mathcal{E} = \{e_1, e_2, \dots, e_m\}$ with $e_1 < e_2 < \cdots < e_m$, we know that for a fixed m, k, a, L and any $t \leq 2$,

$$\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m-e_t+t}(a-L)^{e_t-t} \geq \binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{0}(a-L)^{n-m}.$$

Note that for any $\mathcal{E}$, e_1 in $\mathcal{E}$ must be smaller or equal to $n-m$, as if $e_1 = n-m+1$, then $\mathcal{E}$ must equal to $\mathcal{E}_2$, a contradictory statement. Thus,

$$\begin{aligned}
H_{a,n}(\mathcal{E}) &= \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{n-m-e_t+t}(a-L)^{e_t-t} \\
&\geq \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=2}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{0}(a-L)^{n-m} \\
&+ \frac{1}{a^n}\sum_{L=1}^{a}\sum_{k=1}^{m}\binom{1-1}{k-1}(L-1)^{m-k}(a-L+1)(a-L)^{n-m-1}. \\
&= H_{a,n}(\mathcal{E}_3).
\end{aligned}$$

Hence, $\mathcal{E}_3$ is the second least likely m card hand achieved under an inverse a-shuffle on the ordered deck $(1, 2, \dots, n)$.
We can simplify $H_{a,n}(\mathcal{E}_3)$. By Lemma 3.3,

$$\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{0}(a-L)^{n-m} = \sum_{L=1}^{a}(L^m-(L-1)^m)(a-L)^{n-m}.$$

So

$$\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=2}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{0}(a-L)^{n-m} = \sum_{L=1}^{a}(L^m-(L-1)^m)(a-L)^{n-m} - \sum_{L=1}^{a}(L-1)^{m-1}(a-L)^{n-m}.$$

Also, in the second summation, the only nonzero term is when $k = 1$, thus

$$\begin{aligned}
&\sum_{L=1}^{a}\sum_{k=1}^{m}\binom{1-1}{k-1}(L-1)^{m-k}(a-L+1)(a-L)^{n-m-1} \\
&= \sum_{L=1}^{a}(L-1)^{m-1}(a-L+1)(a-L)^{n-m-1} \\
&= \sum_{L=1}^{a}(L-1)^{m-1}(a-L)^{n-m} + \sum_{L=1}^{a}(L-1)^{m-1}(a-L)^{n-m-1}.
\end{aligned}$$

Therefore, adding the two and multiplying by $\frac{1}{a^{2n}}$ achieves

(3.3)

$$H_{a,n}(\mathcal{E}_3) = \frac{1}{a^n}\sum_{L=1}^{a}(L^m - (L-1)^m)(a-L)^{n-m} + \frac{1}{a^n}\sum_{L=1}^{a}(L-1)^{m-1}(a-L)^{n-m-1}.$$

■

We now turn our attention to the multi-deck $Z_{n,s}$. We consider the probability distribution of the m card hands from the initial ordering of $\pi_0^* = 1\cdots 12\cdots 2\cdots s\cdots s$. We shall lift the elements of $Z_{n,s}$ to $Z_{sn,1}$, which is a single deck with sn cards. We then project the elements of $Z_{sn,1}$ to $Z_{n,s}$ through the projection map $f : Z_{sn,1} \mapsto Z_{n,s}$, where

$$f(k) = \left\lceil \frac{k}{s} \right\rceil.$$

Thus, the elements $1, 2, \ldots, s$ in $Z_{sn,1}$ correspond to $1, 1, \ldots, 1$ in $Z_{n,s}$, the elements $s+1, s+2, \ldots, 2s$ in $Z_{sn,1}$ correspond to $2, 2, \ldots, 2$ in $Z_{n,s}$ and so on. The idea is that we know the probability distribution for m card hands in a single deck with sn cards, thus we can compute the probability distribution of m card hands in the multi-deck $Z_{n,s}$ through the projection from a single deck to a multi-deck. Note that this is essentially the same idea as adding 0.5 to one of the repeated elements of $Z_{n,2}$ in the 2-deck case as we have used in Section 2, albeit a more elegant way.

Theorem 3.7. Let the initial ordering of the multi-deck $Z_{n,s}$ be

$$\pi_0^*, \mathcal{E} = \{e_1, e_2, \ldots, e_m\}$$

be a hand from $Z_{n,s}$, i.e. $\mathcal{E} \subseteq Z_{n,s}$, where $e_1 \le e_2 \le e_3 \le \cdots \le e_m$.
Let $W(\mathcal{E})$ be the set of all sequences $\Gamma = (w_1, w_2, \ldots, w_m)$
with $1 \le w_1 < w_2 < \cdots < w_m \le sn$ such that $\left\lceil \frac{w_t}{s} \right\rceil = e_t$ for all $t \in \{1, 2, \ldots, m\}$.
Then the probability $\widehat{H}_{a,n,s}(\mathcal{E})$ of obtaining the hand $\mathcal{E}$ after an inverse a-shuffle on the multi-deck $Z_{n,s}$ is

$$\frac{1}{a^{sn}}\sum_{\Gamma=(w_i)_{i=1}^m \in W(\mathcal{E})}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{sn-m-w_t+t}(a-L)^{w_t-t}.$$

Proof. Note that our lift has resulted in

$$\begin{array}{ccc} (1,1,\ldots,1) & \mapsto & (1,2,\ldots,s) \\ (2,2,\ldots,2) & \mapsto & (s+1,s+2,\ldots,2s) \\ & \vdots & \\ (n,n,\ldots,n) & \mapsto & ((n-1)s+1,(n-1)s+2,\ldots,sn). \end{array}$$

For any $\Gamma = (w_1, w_2, \ldots, w_m) \in W(\mathcal{E})$, it is easy to see that any inverse a-shuffle on the initial ordering $(1, 2, \ldots, sn)$ of $Z_{sn,1}$ that achieves Γ will also achieve $\mathcal{E}$ on $Z_{n,s}$ with initial ordering π_0^* via the projection. By Theorem 3.1, the probability of achieving Γ from the initial ordering $(1, 2, \ldots, sn)$ is

$$\frac{1}{a^{sn}}\sum_{L=1}^{a}\sum_{k=1}^{m}\sum_{t=1}^{m}\binom{t-1}{k-1}(L-1)^{m-k}(a-L+1)^{sn-m-w_t+t}(a-L)^{w_t-t}.$$

Hence, the desired probability is the sum of probabilities of achieving all $\Gamma \in W(\mathcal{E})$ from the initial ordering of $(1, 2, \ldots, sn)$, namely

$$\frac{1}{a^{sn}} \sum_{\Gamma=(w_i)_{i=1}^m \in W(\mathcal{E})} \sum_{L=1}^{a} \sum_{k=1}^{m} \sum_{t=1}^{m} \binom{t-1}{k-1} (L-1)^{m-k} (a-L+1)^{sn-m-w_t+t} (a-L)^{w_t-t}.$$

■

Theorem 3.8. Let the initial ordering of the 2-deck $Z_{n,2}$ be π_0^*. After an inverse a-shuffle, the least likely hand formed by the first m cards of the shuffled deck is

$$\left\{n - \frac{m}{2} + 1, n - \frac{m}{2} + 1, \ldots, n, n\right\} \text{ when } m \text{ is even}$$

$$\left\{n - \frac{m-1}{2}, n - \frac{m-1}{2} + 1, n - \frac{m-1}{2} + 1, \ldots, n, n\right\} \text{ when } m \text{ is odd}$$

Proof. Let $\mathcal{E} = \{e_1, e_2, \ldots, e_m\} \subseteq Z_{n,2}$ be an arbitrary hand of m cards. Consider the orphan elements in $\mathcal{E}$, i.e. elements that only appear once, which we note as $g_1, g_2, \ldots, g_d$. Following the lift of $(1, 1, 2, 2, \ldots, n, n)$ to $(1, 2, \ldots, 2n)$ as described in Theorem 3.7, we notice that for any g_i such that $i \in \{1, 2, \ldots, d\}$, g_i is lifted to either $2g_i - 1$ or $2g_i$, which gives us two choices. For any pair of elements (e_i, e_{i+1}) with $e_i = e_{i+1}$, (e_i, e_{i+1}) is lifted to $(2e_i - 1, 2e_i)$, leaving only one choice. Thus $|W(\mathcal{E})| = 2^d$, where d is the number of orphan elements in $\mathcal{E}$.
Now we find the least likely m card hand for both odd and even cases.
Case 1: m is even. As $|W(\mathcal{E})| = 2^d$, it is easy to see that

$$\widehat{H}_{a,n,2}(\mathcal{E}) = \sum_{\Gamma \in W(\mathcal{E})} H_{a,2n}(\Gamma) \geq 2^d \min_{\Gamma}(H_{a,2n}(\Gamma)).$$

Since $d \geq 0$, $2^d \geq 1$, by Theorem 3.4

$$\begin{aligned} \widehat{H}_{a,n,2}(\mathcal{E}) &= \sum_{\Gamma \in W(\mathcal{E})} H_{a,2n}(\Gamma) \geq \min_{\Gamma}(H_{a,2n}(\Gamma)) \\ &= \frac{1}{a^{2n}} \sum_{L=1}^{a} (a-L)^{2n-m} \big(L^m - (L-1)^m\big). \end{aligned} \tag{3.4}$$

It is easy to check that equality is achieved when

$$\mathcal{E} = \left\{n - \frac{m}{2} + 1, n - \frac{m}{2} + 1, \ldots, n, n\right\}$$

Case 2: m is odd. As m is odd, $d \geq 1$, thus $|W(\mathcal{E})| \geq 2$.
Therefore, $\widehat{H}_{a,n,2}(\mathcal{E})$

$$= \sum_{\Gamma \in W(\mathcal{E})} H_{a,2n}(\Gamma) \geq \min_{\Gamma}(H_{a,2n}(\Gamma)) + \min_{\Gamma \neq \{2n-m+1, 2n-m+2, \ldots, 2n\}} (H_{a,2n}(\Gamma)).$$

By Lemma 3.6 and Theorem 3.4, the above is equal to

$$\begin{aligned}\widehat{H}_{a,n,2}(\mathcal{E}) &= \frac{2}{a^{2n}}\sum_{L=1}^{a}(a-L)^{2n-m}\big(L^m-(L-1)^m\big) \\ &\quad + \frac{1}{a^{2n}}\sum_{L=1}^{a}(L-1)^{m-1}(a-L)^{2n-m-1}.\end{aligned} \tag{3.5}$$

It is easy to check that equality is achieved when

$$\mathcal{E}=\left\{n-\frac{m-1}{2}, n-\frac{m-1}{2}+1, n-\frac{m-1}{2}+1,\ldots,n,n\right\}$$

Thus proving our result. ■

We would like to analyze the separation distance for hands in a 2-deck. However there is a problem. Unlike in a single deck where all hands have the same probability if the permutation has uniform distribution, in a multi-deck setting this is not true. For example, in a perfectly shuffled 2-deck $Z_{2,2}$, the probability of getting $\{1,1\}$ and $\{2,2\}$ are both $1/6$ while the probability of getting the hand $\{1,2\}$ is $2/3$. Thus we need to be able to define the separation distance when the distribution we want to compare to is not uniform. Here we introduce a *generalized separation distance*: we compare the probability of the least likely hand after an inverse a-shuffle with the probability of the least likely hand if we randomly draw m elements from $Z_{n,2}$, with each element having the same probability being selected.

Definition 3.1. Starting from the initial ordering π_0^* of the 2-deck $Z_{n,2}$, the generalized separation distance after an inverse a-shuffle is defined as

$$\mathrm{SEP}_2(a,n,m)=1-\frac{\min_{\mathcal{E}\subseteq Z_{n,2},|\mathcal{E}|=m}\widehat{H}_{a,n,2}(\mathcal{E})}{U_{n,m}},$$

where $\widehat{H}_{a,n,2}(\mathcal{E})$ denotes the probability of obtaining the hand $\mathcal{E}$ of $Z_{n,2}$ after an inverse a-shuffle, and $U_{n,m}$ is the probability of the least likely hand of m elements if we randomly draw m elements from $Z_{m,2}$ with equal probability.

To compute the generalized separation distance after an inverse a-shuffle, we have

Corollary 3.9. Let the initial ordering of the 2-deck $Z_{n,2}$ be π_0^*. Then for even m we have $U_{n,m}=\dfrac{1}{\binom{2n}{m}}$ and

$$\min_{\mathcal{E}\subseteq Z_{n,2},|\mathcal{E}|=m}(H_{a,2n}(\mathcal{E}))=\frac{1}{a^{2n}}\sum_{L=1}^{a}(a-L)^{2n-m}\big(L^m-(L-1)^m$$

For odd m we have $U_{n,m} = 2/\binom{2n}{m}$ and

$$\min_{\mathcal{E}\subseteq Z_{n,2},|\mathcal{E}|=m} H_{a,2n}(\mathcal{E})) = \frac{2}{a^{2n}}\sum_{L=1}^{a}(a-L)^{2n-m}\big(L^m-(L-1)^m\big) + \frac{1}{a^{2n}}\sum_{L=1}^{a}(L-1)^{m-1}(a-L)^{2n-m-1}.$$

Proof. We have already computed $\min_{\mathcal{E}\subseteq Z_{n,2},|\mathcal{E}|=m}(H_{a,2n}(\mathcal{E}))$ in Theorem 3.8 by the equations (3.4) and (3.5). For $U_{n,m}$ this is also rather straightforward. For even m, the least likely set of m elements are sets that contain only pairs, and it is easy to see the probability is exactly $1/\binom{2n}{m}$ as there is only one way to select the it from $2n$ elements. For odd m there is one orphan element, which can be selected in two ways from $2n$ elements, and hence the probability is $\frac{2}{\binom{2n}{m}}$. ■

We should remark that one drawback of this definition is that it gives the same result as the separation distance for hands of m cards in the single deck $Z_{2n,1}$ setting when m is even. For odd m, the two are different, but the difference is also numerically very small.
Table 7 shows the generalized separation distance for $Z_{n,2}$ after an inverse a-shuffle, with $m = n/2$ i.e. hands consist of 1/4 elements of the total deck. As one can see, it has exactly the same numbers as Table 5 as we mentioned. While for odd m the numbers are different, the numerical difference is smaller that the precision we use. Note that for $n = 54$ this is the case for the popular Chinese game Guandan. After just 4 shuffles the separation distance is already significantly lower than the separation distance of permutations for a single deck of 52 cards after 8 shuffles. Again, perhaps we can speculate that ***4 shuffles suffice*** for Guandan?

Table 7. Generalized separation distance after s inverse riffle shuffles on double deck $Z_{n,2}$ for n = 8, 16, 26, 32, 52, 54, 64 and 104.

$s \setminus 2n$	16	32	52	64	104	108	128	208
4	0.373	0.573	0.708	0.758	0.848	0.853	0.876	0.922
5	0.215	0.370	0.508	0.570	0.706	0.716	0.756	0.847
6	0.116	0.214	0.316	0.369	0.507	0.518	0.569	0.705
7	0.060	0.115	0.178	0.214	0.316	0.325	0.369	0.506
8	0.031	0.060	0.095	0.115	0.178	0.184	0.213	0.316
9	0.015	0.031	0.049	0.060	0.095	0.098	0.115	0.178
10	0.008	0.015	0.025	0.031	0.049	0.051	0.060	0.095

4. Conclusion

In this paper we study the question "How many shuffles are required to sufficiently randomize a deck of cards?" for multi-decks, i.e. decks with repeated cards. For a single deck, the most influential work on the subject remains the 1992 paper by Bayer and Diaconis, which provided an answer for a standard 52-card deck, showing that seven shuffles suffice. However, multiple decks of cards are commonly used in various card games, such as the popular Chinese game GuanDan (2 decks), making the randomness of a deck after shuffling highly relevant in both recreational and professional settings. Here we provide a mathematical analysis on the randomness of riffle shuffling multiple decks of cards, extending the result of Bayer and Diaconis [2] to the multi-deck setting. We obtain an explicit value of the total variation distance for a 2-deck setting after inverse riffle shuffles. A key contribution of ours is the investigation of the randomness of a player's hand after riffle shuffling in a deck, a question that has not been investigated at all previously, from which we are able to give explicit formulae for the separation distance for a player's hand after riffle shuffling for both the single deck and double deck setting.

References

[1] S. Assaf, P. Diaconis, and K. Soundararajan. Riffle shuffles of a deck with repeated cards. *Discrete Mathematics & Theoretical Computer Science*, (Proceedings), 2009.

[2] D. Bayer and P. Diaconis. Trailing the dovetail shuffle to its lair. *The Annals of Applied Probability*, pages 294–313, 1992.

[3] E. Gilbert. Theory of shuffling. Technical report, Bell Laboratories, 1955.

[4] G. Kolata. In shuffling cards, 7 is winning number, 1990.

[5] L. N. Trefethen and L. M. Trefethen. How many shuffles to randomize a deck of cards? *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 456(2002):2561–2568, 2000.

[6] A. Van Zuylen and F. Schalekamp. The achilles' heel of the gsr shuffle: A note on new age solitaire. *Probability in the Engineering and Informational Sciences*, 18(3):315–328, 2004.

EDITOR'S COMMENTS

This paper extends the well-known Bayer–Diaconis "seven shuffles" analysis to multiple decks, motivated by games such as GuanDan, and—its most original part—studies how random an individual player's *hand* is after riffle shuffling, giving explicit separation-distance formulae for both the single- and double-deck cases. The hand-randomness question appears not to have been studied before, and turning attention to what actually matters in a card game—the hand a player is dealt—is a smart and original idea.

The paper builds carefully on the established GSR and a-shuffle theory, extends it to the general m-deck case, and produces explicit closed-form separation-distance formulae. Handling repeated and indistinguishable cards correctly takes real care, and the results are concrete formulae rather than rough estimates. The hand-randomness part is the most original, while the multi-deck extension is a natural follow-on from known methods. The motivation is explained clearly and the command of the literature—Gilbert–Shannon–Reeds, Bayer–Diaconis, Trefethen, Van Zuylen–Schalekamp—is excellent as high school students. Overall, it is an interesting and well-focused paper about probability theory.