Knowing When to Fold'em: A Monte Carlo Exploration of Card Shuffling and How Poker Players Can Gain an Advantage

David W. Merz
Peter B. Chi

Abstract

This work demonstrates that the card shuffling procedure commonly performed in casino poker rooms is insufficient for randomizing a deck of cards. We explore this in the context of Texas Hold'em, which has established itself as the most popular form of poker worldwide over the past few decades. We show the degree to which any given card may be more (or less) likely to show up as an important card in the subsequent hand. Additionally, we find that the shuffling procedure does not sufficiently separate cards from their starting point; that is, cards are more likely to stay close together after shuffling than they should by chance. Thus, in this work, we demonstrate that Texas Hold'em players can gain an advantage over their opponents by recognizing these deficiencies in the shuffling procedure.

Keywords: card shuffling; poker; Texas Hold'em; applied probability

David W. Merz
Undergraduate student
University of Notre Dame
dmerz2@nd.edu

Peter B. Chi
Assistant
Professor of Statistics
Villanova University
peter.chi@villanova.edu

Introduction

No Limit Texas Hold'em (NLHE) is a poker game utilizing a single standard 52-card deck. Before each new round is dealt, the deck is shuffled to randomize the distribution of cards. This work investigates the extent to which standard manual shuffling procedures sufficiently randomize a deck of cards for usage in NLHE. Theoretically, an ideal shuffling procedure of an ordinary deck of 52 cards would leave each possible ordering of the 52 cards equally likely, with a 1/52! probability of each ordering. As a result, such a shuffling procedure would give each individual card the same probability of ending in each position of the deck after the deck has been shuffled. For example, the top card of the deck before any shuffling is performed should have exactly a 1/52 probability of being any position in the deck after the shuffle is performed; in other words, the location of the card after shuffling should follow a discrete uniform distribution. Here, we investigate the degree to which card shuffling in casino poker rooms may not produce such an outcome, and how this imperfection can be exploited by players to gain an advantage.

There exists a well-known rule of thumb that seven riffle shuffles are recommended in order to render a deck of cards sufficiently random (according to the total variation distance metric), where a riffle shuffle is the commonly used technique of splitting the deck into two roughly equal portions and then interleaving the cards together (Aldous, 1983; Aldous & Diaconis, 1986; Bayer & Diaconis, 1992). However, casino poker rooms across the world do not generally perform seven riffle shuffles as their shuffling procedure. Instead, the following two approaches are common:

1. Automatic card shuffler: these machines are meant to randomize a deck of cards with minimal human effort for usage in casino games. A careful treatment of such machines has already been performed (Diaconis et al., 2013); additionally, many casino poker tables are still not equipped with these machines. Thus, they will not be the focus of our research.

2. The Casino Shuffle: What we will refer to as the CS for the remainder of this manuscript is a standard shuffling procedure in casinos that is performed by dealers when an automatic card shuffler is not used. The CS procedure is the following: two riffle shuffles, one strip shuffle, one riffle shuffle, and one cut (in that order). A cut is simply the process of splitting the deck in half and placing what was previously the bottom half on top of what was previously the top half; a strip shuffle is a similar process but instead of rearranging just two halves, some greater number of portions (often four) are rearranged; and a riffle shuffle is as described above. In the Methodology section below, we will go into greater detail on how we mathematically express these mechanics. The prevalence of the CS across poker rooms worldwide is addressed in the Discussion section.

Much of the existing work on card shuffling has largely focused on theoretical aspects (see Literature Review section below). Washburn (2018) introduces a metric of shuffling efficacy that is motivated by the subsequent usage of the cards in a variety of games; however, he does not specifically investigate its actual impact on the play of any of these games. Hannum (2000) analyzed data arising from repeated manual shuffling according to procedures known to be utilized at Las Vegas Strip casinos in early 1997 for single-deck, double-deck and six-deck games, clearly with the resulting impact on these games in mind. However, while the analyses performed all give various statistical measures describing the degree of non-randomness resulting from the shuffling procedures, none of them address what tends to happen to the cards as they are dealt in an actual card game, nor specifically how these deficiencies in randomness can be exploited by a player in any given game. The only existing published research, to our knowledge, investigating the question of how deficiencies in the randomness of shuffling procedures can be exploited by a card player is Thorp (1973), within a currently obscure game called Faro and with a brief description of how the concepts therein may apply to Blackjack.

Thus, the literature remains quite sparse with regard to the impact of a card shuffling procedure within the context of an actual game in which the cards are to be used. The question that we set out to investigate was whether the CS produces a deck that is sufficiently random for usage in NLHE, and if not, how much of an advantage might a NLHE player be able to gain over their opponents by analyzing the CS shuffle. All code for Monte Carlo simulation experiments associated with this manuscript is available at https://github.com/d-merz/casino-shuffle-research-code.

Literature Review

Card shuffling has been well-studied in the mathematics and computer science literatures, in part because various shuffling methodologies have direct relationships with other real-world processes outside of their usage on decks of cards. For example, the *top-swap*, first described by Durrett (2005), is a card shuffling method that was motivated by the need for a model of genomic rearrangements in biological studies. Because genetic material (e.g., DNA or RNA molecules) consists of contiguous nucleotides in the same way that decks of cards consist of contiguous individual cards, models created for one of these processes are also applicable to the other. Genomic rearrangements include translocations, in which portions of genetic material from two different chromosomes switch places with each other, and inversions, in which portions of genetic material reverse their order within their original location.

In the top-swap card shuffling model, we assume that there are k decks of cards, and the i^{th} deck has n_i cards. To perform the shuffle, two positions across the k decks are chosen at random. If the chosen positions are in two different decks, each deck is cut at these respective positions, and the resulting top portions of each deck are swapped (analogous to a genomic translocation as described above). If the positions are in the same deck, the order of the cards within the interval contained between the two positions is reversed (analogous to a genomic inversion as described above). Durrett (2005) formulated a Markov chain model for this process, and estimated rates of convergence to stationarity of the chain. Further study of this model includes characterizing the relaxation time of the chain, which is a central quantity with respect to the rate of convergence to stationarity (Bhatnagar et al., 2007).

Another card shuffling method, dubbed the *swap-or-not*, has connections to cryptology (Hoang et al., 2012). The swap-or-not shuffle proceeds by choosing pairs of cards from the deck at random and then flipping a fair coin to determine whether the two chosen cards should switch positions or not. Hoang et al. (2012) demonstrate that this procedure has applications to the encryption of credit card numbers and its optimality therein. An adaptation of the swap-or-not, called the *mix-and-cut*, was developed by Ristenpart and Yilek (2013), with the aim of improving the security of encrypted objects to a greater number of queries. The mix-and-cut begins by using the methodology of the swap-or-not, but then performs a deck cut (as described above) after which each portion of the deck then undergoes a round of swap-or-not. Each portion is then cut again, and then a round of swap-or-not is performed on each of the four resulting portions. This continues until no more deck cuts can be performed. Further work on the mix-and-cut method includes a reformulation that enabled it to be performed much faster (Morris & Rogaway, 2014). Clearly, the swap-or-not and mix-and-cut shuffling methods could not reasonably be performed manually by a dealer at a live poker table, but they could conceivably be adapted by automatic shuffling machines or for the shuffling of virtual cards in online poker.

Other shuffling methods have less direct connection to other processes outside of decks of cards but have still broad usefulness as toy models on which to develop new tools for mathematical problems. For example, consider the *top-to-random* shuffle, in which one shuffle consists of moving the top card of the deck to any position of the deck at random (including remaining as the top card). Aldous and Diaconis (1986) use this as their first setting on which to develop their mathematical results regarding an upper bound and asymptotic behavior of the total variation distance between the distribution of cards in the resulting deck after any number of shuffles and the distribution of cards in a perfectly shuffled deck (as defined above). The top-to-random shuffle was further studied to derive a closed-form expression for the distribution of cards after any number of iterations of the top-to-random shuffle, enabling further characterization of the asymptotic behavior of the approach to stationarity of the Markov chain associated with the shuffling procedure (Diaconis et al., 1992).

Following their analysis of the top-to-random shuffle, Aldous and Diaconis (1986) turned their attention to the *riffle shuffle* as briefly mentioned above. The Gilbert-Shannon-Reeds model, described in further detail below, is a mathematical model that had been previously developed to mimic the mechanics of the riffle shuffle (Gilbert, 1955; Reeds, 1981). Utilizing it, Aldous and Diaconis (1986) derive an upper bound to the total variation distance between the resulting deck after any number of riffle shuffles and the distribution of cards in a perfectly shuffled deck. Their result corroborates a previous estimate that approximately seven riffle shuffles are recommended to randomize 52 cards (Aldous, 1983). A given deck of cards' movement towards the same distribution as a perfectly shuffled deck via riffle shuffling according to the Gilbert-Shannon-Reeds model is further characterized by Aldous and Diaconis (1986), who demonstrate that $(3/2)\log_2(n)$ shuffles are necessary and sufficient to randomize a deck of n cards; for $n = 52$, this is approximately 8.55 shuffles.

While the Gilbert-Shannon-Reeds model has been demonstrated to be a very close approximation of an actual dealer's manual riffle shuffling of a deck (Diaconis, 1988), the exact nuances of any given dealer's technique at riffle shuffling will clearly have an impact on the resulting distribution of the cards. However, the extent of this impact and whether the number of riffle shuffles required to randomize a deck of cards may significantly vary as a result is unknown. Diaconis et al. (1983) studied the case in which a deck is deterministically split into two equally sized portions and riffled in such a manner that exactly one card falls from each portion in turn. Of note is that when this deterministic procedure is performed on a deck of 52 cards, eight such shuffles will return the deck to its original order. Thus, some randomness as to the size of the portions and the number of cards that drop from each portion is crucial to the mixing of the deck. As the Gilbert-Shannon-Reeds model remains the best-known model of the riffle shuffle to date, we hinge our analyses on this model to investigate the performance of the CS and explore the extent to which the CS has deficiencies that are exploitable in the context of NLHE.

Methodology

NLHE Shuffling

The shuffling techniques that make up the CS procedure were simulated according to the following assumptions:

Riffle shuffle

We follow the Gilbert-Shannon-Reeds model (Gilbert, 1955; Reeds, 1981):

1. The deck of n cards is first split into two portions, with the probability of each possible number of cards in each portion following a binomial distribution with $p = 0.5$; that is, $P(K = k) = \binom{n}{k}/2^n$, where k represents the number of cards that end up in the first portion.

2. The two portions are then riffled together, with cards dropping into place from either portion randomly and with probabilities proportional to the number of cards remaining in each portion; that is, if A and B represent the number of cards remaining in the first and second portions respectively, the probability that the next card falls from the first portion is $A/(A+B)$.

Strip shuffle

1. A deck of n cards is cut into two portions according to a binomial distribution with probability equal to 1/4. The resulting top portion is approximately the first quarter of the deck; call it A.

2. The remainder of the deck of n cards that contains approximately 3/4 of the deck is cut into two portions according to a binomial distribution with probability equal to 1/3. The resulting top portion here is approximately the second quarter of the original deck; call it B.

3. With the remaining approximately half of the deck, cut it into two portions according to a binomial distribution with probability equal to 1/2. Call the two resulting portions C and D. The result at this point is then four roughly equal heaps: A, B, C, and the remainder of the deck D.

4. Place heap B on top of heap A, then place heap C on top of heap BA. Finally, place heap D on top of heap CBA. The final order of the deck is $DCBA$.

Deck Cut

1. A deck of n cards is cut into two portions according to a binomial distribution (in the same manner as the first step of a riffle shuffle); thus, the probability that the first portion contains k cards is again $\binom{n}{k}/2^n$.

2. The second portion is placed on top of the first portion.

NLHE Dealing

A round of NLHE proceeds as such: after the cards are shuffled using the CS, they remain in the resulting order for the entirety of the round. In a hand with 6 players, Player 1, the first player to receive a card, receives cards 1 and 7 from the top of the deck from the dealer. Next, Player 2 receives cards 2 and 8 from the dealer, Player 3 receives cards 3 and 9, and so on for each player.

After each player receives two cards, card 13 from the top of the deck is "burned" by being placed face down on the table, and then the next three cards, cards 14, 15 and 16, are placed face up (known as the flop). Next, card 17 is burned, and card 18 (the turn) is placed face up next to the flop. Finally, card 19 is burned, and card 20 (the river) is placed face up next to the flop and turn. After the round is finished, the dealer collects the cards, often placing the cards from the round on the top or bottom of the deck.

Monte Carlo simulations were devised to investigate the performance of the CS in the context of NLHE. In order to evaluate the effectiveness of the CS, the following measures were investigated: the probability of cards being dealt to certain locations, the likelihood of neighbor cards staying next to each other, and the number of cards that can be guessed correctly in a deck using an optimized algorithm. We also propose a set of alternative shuffling procedures for casinos to consider, utilizing various permutations of the techniques present in the CS described above. Code to implement the simulation of these shuffling procedures and Monte Carlo simulations was written in the R programming language (R Core Team, 2021), with graphical summaries produced using the ggplot2 (Wickham, 2016) and ggpubr (Kassambara, 2020) packages.

Results

Empirical Probabilities of Dealt Locations

First, we investigate potential deficiencies in the CS by examining simulated empirical probabilities of certain cards being dealt to certain players or board spots after the CS has been performed and comparing them to what they should be with a perfectly randomized deck. Here, we specifically track the first five and last five cards in the deck before any shuffling is performed, as these are likely locations for the community cards from the previous hand to be placed prior to shuffling. While this is of course not always the case, we use this as a starting point for our demonstrations (see Discussion section for other considerations). We evaluate the performance of the CS by comparing it to the performance of seven riffle shuffles, and to a theoretically perfect shuffling procedure (that is, one that would result in each card's location following a discrete uniform distribution).

Via Monte Carlo simulations, we:

1. Perform the necessary shuffle on a deck of cards.
2. Deal the shuffled deck to a table of six players.
3. Track the occurrences of each of the first five and last five cards in the starting deck being dealt to each player and as the community cards.
4. Determine the relative frequency (empirical probability) of each card going to each player or being dealt as a community card.

For each empirical probability, the 95% confidence interval is shown as the upper and lower lines that bound the point estimate obtained, which is the calculated relative frequency from the Monte Carlo simulations. The confidence interval bounds were obtained via the standard Wald interval assuming approximate normality of the sample proportion.

As shown in Figure 1, the confidence interval bounds corresponding to the seven riffle shuffles nearly always include the black dotted line which represents the expected probabilities under perfect randomness. In contrast, the bounds corresponding to the CS rarely encompass this line. Thus, depending on the position of a card before any shuffling, it can be far more likely or far less likely to be in the hand than it should be.

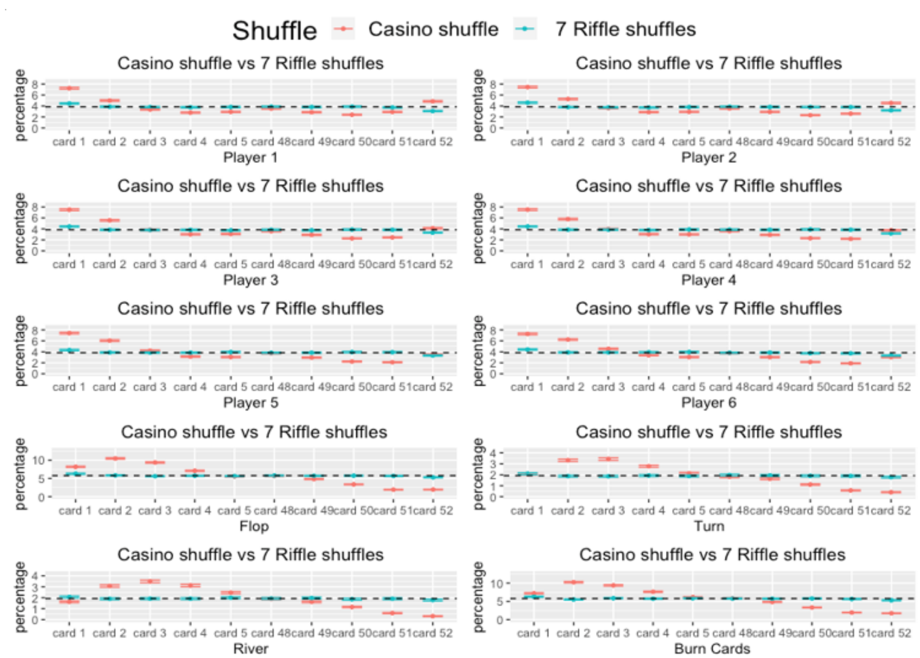

Figure 1
Empirical probabilities of cards being dealt to a player or as a community card (100,000 iterations).

For example, consider the second and third cards in the deck prior to any shuffling. As shown in the Flop panel in Figure 1, the CS favors these cards being dealt on the flop.

Card #2 has more than a 10% chance and card #3 has approximately a 10% chance, indicating that these two cards are about twice as likely to come on the flop than they should be. Therefore, if a player notices that the dealer is placing the community cards from the previous hand on top of the deck before shuffling, they would know that these cards would be much more likely to come on the flop than they should be. Overall, a simple trend that can be easily recognized in Figure 1 is that when using the CS, cards at the front of the deck tend to stay at the front of the deck, and cards at the end of the deck tend to stay at the end.

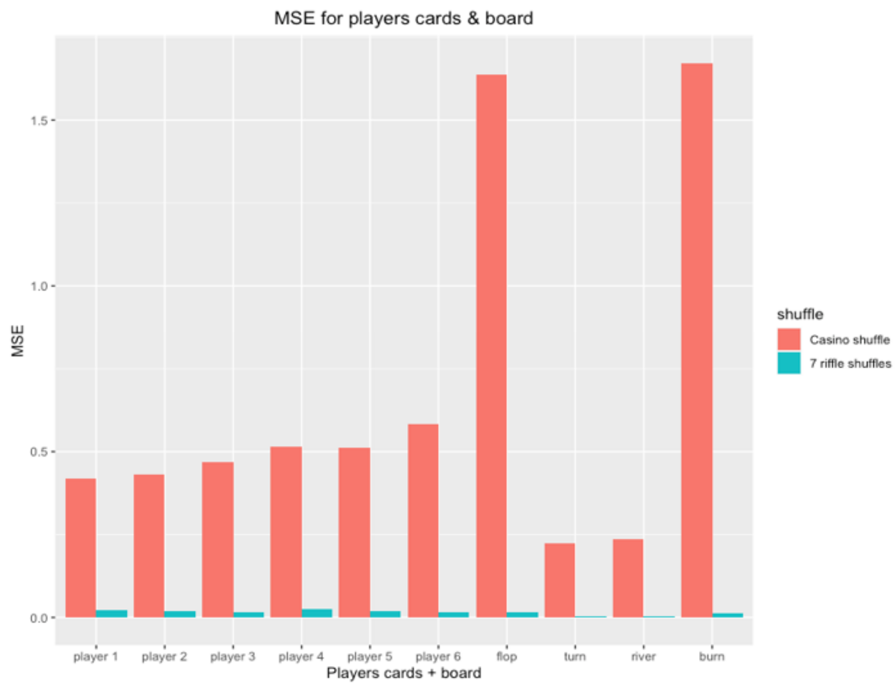

Figure 2
MSE of the set of probabilities for a card to be dealt to a player or as a community card (100,000 iterations).

While Figure 1 focuses solely on the first five and last five cards in the starting deck, in Figure 2 we consider the entire deck and the relative frequency for each card to be in each location of interest. We utilize the mean squared error (MSE), which combines both the bias and the variance into one summary measure. If the deck is perfectly randomized, then all cards should have the correct probability of landing into any particular spot, giving an MSE of 0; on the other hand, large values of the MSE indicate that the empirical probabilities for the cards are either far from what they should be, highly variable, or both. In Figure 1, we observe that the CS gives a wide range of empirical probabilities for the ten highlighted cards to end up in any given dealt location; this is further summarized here in Figure 2 for the entire deck as we plot the MSE of the percentages for each dealt location, for the CS and seven riffle shuffles. For each location, the MSE for the CS is significantly larger than the MSE for the seven riffle shuffles. This emphasizes the extent to which, when the CS is used, there are a significant number of cards that are dealt to a player or on the board at a different probability than they should be.

Card Neighbors

Perhaps the most important test of the CS is how well it breaks up "clumps" of cards, or cards in consecutive locations in the deck. This is important to a game like NLHE which is dictated at each step by what the next card in the deck may be, and the board cards are dealt face-up and sequentially. Consider two cards, x and y, that are located consecutively in the deck prior to any shuffling, with x coming before y. We call these cards "neighbors," and say that their "cards apart" value is equal to 1. Our interest is in the probability of each possible cards apart value after shuffling, and how far they deviate from what would be expected from a perfectly shuffled deck. Additionally, we focus our interest on positive values of cards apart whereby card x still comes before card y, because it is in this direction that the CS suffers more in a way that can be taken advantage of. This is because, due to the mechanics of the CS, it is disproportionately unlikely that the CS will move card y to be positioned anywhere before card x; thus, observing card y being dealt would suggest that card x will have a lower probability of coming shortly thereafter than it should. This reduction in probability, however, is not nearly as great as the magnitude of the increase in the probability of card y coming within a few cards after card x that results from the CS, as shown in Figure 3. Here, we present Monte Carlo estimates of the probabilities of each possible positive value of cards apart after shuffling, for any two cards that were neighbors prior to shuffling.

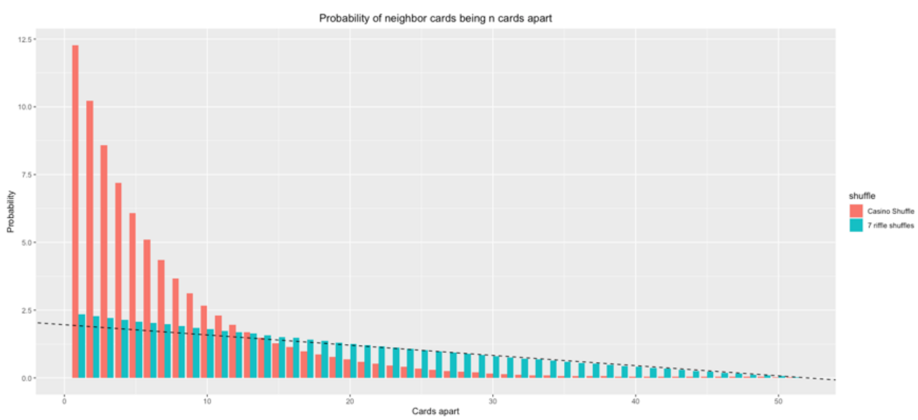

Figure 3
Bar plot showing the probability of neighbor cards ending n cards apart (100,000 iterations).

The black dotted line represents the theoretical probability of cards that are neighbors prior to any shuffling ending up with a cards apart value of n after a perfect shuffling procedure is performed, derived via elementary combinatorics. The CS performs poorly compared to this line. For example, as indicated by the left-most bar, the result from the CS is almost 6x more likely than it should be with respect to neighboring cards remaining together as neighbors after the shuffle. Compared to seven riffle shuffles which is only slightly worse than a perfectly random shuffle, the CS clearly offers very poor randomization according to this metric.

The probability of neighbor cards ending up n cards apart for $n \in \{1,2,3,4,5,6\}$ is shown in more detail in Figure 4 below, with corresponding numeric values shown above each bar. For two cards that are neighbors, the probability of those cards staying in that exact order after the CS is about 12.24%, which is again almost six times greater than it should be. In fact, a heuristic argument demonstrates that this result is not surprising based on the process that comprises the CS. First, consider that both the strip shuffle and cut are largely unlikely to put any separation between any given two neighbor cards (that is, out of all 51 pairs of neighbors in the deck, exactly three pairs of neighbors will be separated by the strip shuffle, and one pair of neighbors will be separated by the cut, with varying probabilities as described previously). Therefore, any separation of two neighbors after the

CS, if it were to happen, would most likely come from the three riffle shuffles. On each riffle, the quantity $A/(A+B)$ is 0.5 in expected value and should rarely deviate very far from that. Thus, the probability that two neighbors remain as neighbors after one riffle is 0.5. Therefore, after three riffles, the probability that they are still neighbors should be approximately $(0.5)^3 = 0.125$ which matches quite nicely with our Monte Carlo estimate.

However, in NLHE it is often not the next card in the deck that matters, but rather the card following that one (that is, with a value of cards apart equal to 2). This is because of the NLHE built-in "bad shuffle catcher" whereby a card is burned between rounds of betting, as described in the dealing procedure above. However, while the intention of the bad shuffle catcher is to mitigate the chance for a player to gain an advantage, we observe that the CS produces a probability that is approximately five times more likely than it should be for neighboring cards ending up two cards apart after the CS. Thus, during a hand after the CS is performed, if for example card x is observed on the turn, then a player can assume that card y is much more likely to appear as the river card than it should be by chance.

An example of how a player could gain an advantage with this information:

Suppose the following community cards are dealt: A♠ 6♢ 3♡ 9♠ 4♣

Once the round is finished, one possible occurrence is that the dealer will then collect the cards from their right to left such that the 4♣ is on the bottom and the A♠ is on top of the 5-card pile, which is face up. The cards will then be flipped over and placed on the burn pile and then on top or bottom of the existing deck. The order of these five cards in the deck is now in reverse order: 4♣ 9♠ 3♡ 6♢ A♠

The dealer then shuffles the cards using the CS and deals out the next hand. A player is dealt and sees the flop with the following hand: A♣ 6♣

The flop is: K♡ Q♣ 6♢

The flop has given this player pair of sixes, a hand that is not extremely powerful, especially with a king and queen on the flop. However, if the player were to gain an Ace by the river to make two pairs, this would greatly increase their chance of having the best hand at the table. A simple consult of the effectiveness of the CS shows that there is a 10.26% chance that the next card dealt (i.e., the second card in the deck, as the next card will be burned) is the A♠. Going even further, there is a 7.22% chance that the river is the A♠. That means that there is approximately a 17.48% chance that the A♠ is dealt on the next two cards. On top of that, from the two other Aces in the deck for which there is a combined 8.42% chance of hitting by the river, the total probability of getting a two pair with aces is then approximately 25.9%. In a random deck, there should only be a 12.5% chance of being dealt an ace on the turn or river here. Therefore, with the CS, the probability of this player's hand improving is far more likely than it should be with a properly randomized deck. This can then be weighed into their decisions, e.g., to play their hand more aggressively than they might otherwise.

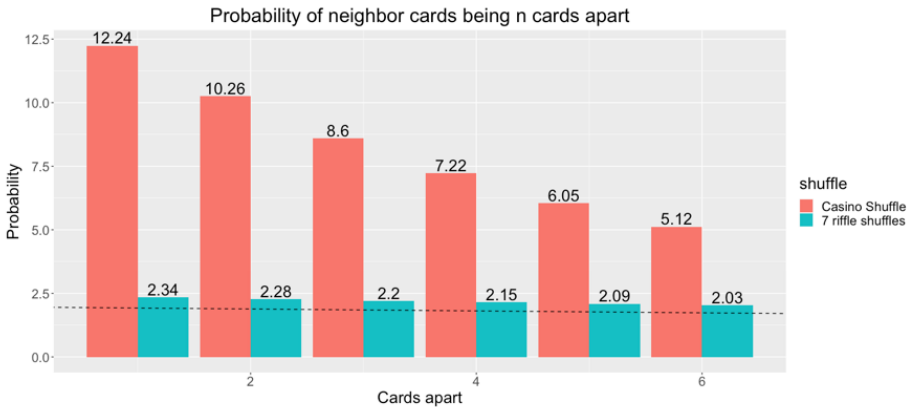

Figure 4
Bar plot showing the probability of neighbor cards ending within 6 cards apart (100,000 iterations).

A Different Distance

In section 5.1 of Bayer and Diaconis (1992), one method used to test the effectiveness of a shuffle is based on the following problem: a deck of cards is face down on the table. The deck has been shuffled, but its original order prior to shuffling is known. A guesser tries to guess what the next card will be one at a time. After each guess, the current top card is turned over to reveal its value and is then discarded. If the guesser believes that the deck is perfectly random, the optimal strategy is to simply guess any card first (with a probability equal to 1/52 of being correct) and thereafter guess any card known to be in the remaining deck. The expected number of correct guesses is:

$$\frac{1}{52}+\frac{1}{51}+\frac{1}{50}+\frac{1}{49}+\cdots+1=\sum_{n=0}^{51}\frac{1}{52-n}\approx 4.54$$

However, when the deck is not perfectly shuffled, there is another method that can be used to guess more than an average of 4.54 cards correctly. The method proceeds as follows:

1. Guess the original top card as the first guess.
2. After each guess is made, check off the card revealed from a list of the deck in its original order. The list will show possible cards unmarked and cards already revealed as crossed out.
3. From the longest block of consecutive possible cards, guess the top card.

A comparison of the number of cards guessed correctly with this strategy is shown in the following figure. The dotted black line represents the expected number of correct guesses if the deck was perfectly shuffled (approximately 4.54 as shown above). The plot shows that more than twice as many cards are expected to be guessed correctly when using the casino shuffle as compared to seven riffle shuffles and to complete randomness. While this method would be practically impossible to employ at an actual poker table, this result further illustrates the deficiencies in the CS by highlighting that far more cards can be guessed correctly after the CS than should be expected by chance.

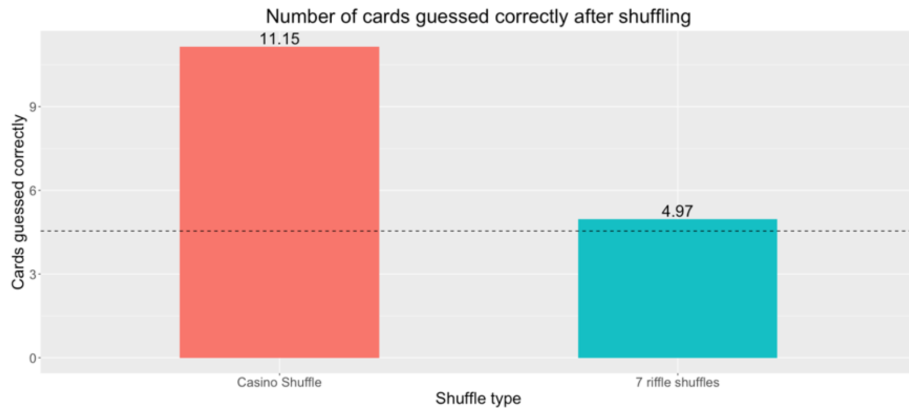

Figure 5
A bar plot showing the number of cards guessed correctly after a shuffling method (100,000 iterations).

An Alternative Shuffling Procedure

As it has been shown throughout this paper, the current shuffling procedure used in casinos is not effective at sufficiently randomizing a deck of cards for NLHE. On the other hand, casinos may be resistant to implementing a new standard of seven riffle shuffles, due to the desire for speed. Therefore, we explore alternative procedures and compare their performance at separating neighbor cards (seen previously in Figure 4 above). Each alternative procedure consists of the same components as the CS (deck cut, riffle shuffle, and strip shuffle), but with different amounts of riffle shuffles and orderings of the components.

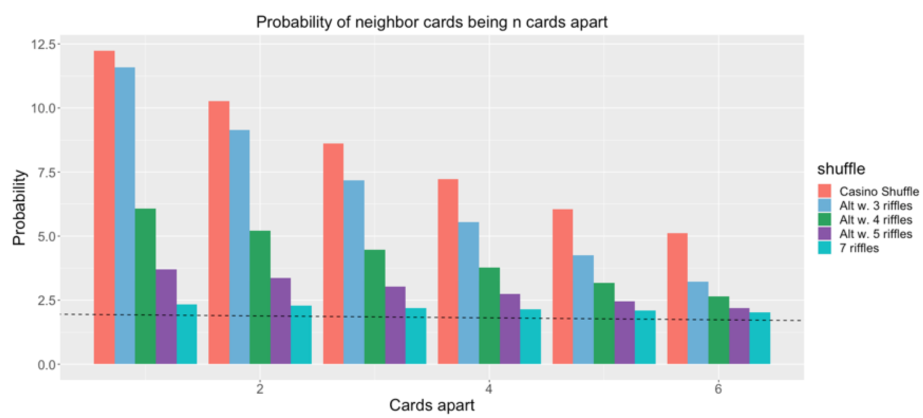

Figure 6
Bar plot comparison of cards apart probabilities from proposed alternative shuffling methods (100,000 iterations).

In Figure 6, we highlight three different proposed alternative shuffling procedures. Based upon an exploration of the impact of the order of each shuffling method within the procedure (see Appendix), in each alternative shuffling procedure all riffling is performed first, followed by a strip shuffle and a cut; thus, the only difference between the CS and the "Alt w. 3 riffles" is when the riffles occur.

While the proposed alternative shuffling procedure of five riffle shuffles, one strip shuffle, and one deck cut is not as effective as seven riffle shuffles, we observe that it is far more effective than the current CS. The proposed procedure only requires two more riffle shuffles, and the result is a shuffle that is greater than three times more effective at breaking up clumps of cards than the current CS.

Discussion

In this work, we have demonstrated that the standard methodology utilized by casinos for manually shuffling a deck of cards does not sufficiently randomize the deck, and the extent to which players can take advantage of this. As our work focused on manual shuffling, several questions as to both the validity and relevance of our findings may arise. The first is whether our method for simulating each riffle shuffle, the Gilbert-Shannon-Reeds model (Gilbert, 1955; Reeds, 1981) accurately emulates a human dealer's riffle shuffle. As mentioned above in the Literature Review section, an analysis by Diaconis (1988) showed that the Gilbert-Shannon-Reeds model does produce deck outcomes that are similar to what would be expected by a manual riffle shuffle with a real deck of cards. However, the extent to which individual dealers' deviations from the Gilbert-Shannon-Reeds model impacts the distribution of cards in the shuffled deck could be a subject of future study.

Secondly, one might ask whether our investigation is relevant since automatic card shufflers exist. The answer to this question is that our work here is indeed relevant to any poker table that is not equipped with an automatic shuffler, which is still relatively common due to how expensive these machines are. Notably, at the World Series of Poker (WSOP) in Las Vegas, NV, only a fraction of their tables are equipped with automatic shufflers, and only recently: in 2019, it was reported that, for the first time in their history, 100 of the tables at the WSOP would be equipped with automatic shufflers (Pitt, 2019); while a total count of tables required may vary by year and even day, it is safe to assume that it is far more than 100 on most days, due to the thousands of daily participants in the numerous events and side games at the WSOP. Additionally, it is worth noting that automatic shufflers have their own imperfections as highlighted by Diaconis, et al. (2013), although we did not focus on them here.

Finally, one may question whether, even if a poker room dealer is shuffling manually, is their procedure actually the CS as described here? To answer this question, we turned to a combination of personal experience, online footage, and any other information that could be gathered online. Firstly, as recently as March of 2022, one of this manuscript's authors (DWM) witnessed the CS in use at the Seminole Hard Rock Hotel & Casino (Hollywood, FL) when an automatic shuffler was not used. Secondly, we note that a WSOP Dealer Guide that can be found on a public-facing website prescribes exactly what we performed here for the CS (WSOP, 2013, page 3). Accordingly, WSOP footage from as recently as 2017 and 2021 showed at least one dealer from each year clearly performing our CS procedure (PokerGO, 2017; PokerGO, 2021). Thirdly, on recent footage from High Stakes Friday: Live at the Bike, while the camera angle did not allow us to see exactly how the dealers are shuffling, the audio while the dealer is shuffling is consistent with what is expected to be heard while the CS is being performed (Live at the Bike, 2021).

Conversely, footage from the European Poker Tour (EPT) showed dealers possibly performing the CS, but with a wash preceding the first requisite riffle (PokerStars, 2018). A wash is a method for shuffling whereby the dealer leaves all cards face down on the table but moves them around in random fashion with both hands. The presence of a wash would clearly impact our analyses. However, for two reasons, our analyses remain useful even in the case of when a wash has been performed: firstly, after a wash, the dealer may hold the deck perpendicularly to the table while squaring them up (that is, preparing them) for the first riffle shuffle. This makes several cards visible to the players, especially the card at the very bottom of the deck, as shown in a dealer video tutorial (PokerListings, 2014); secondly, as the back side of all of the cards are generally visible during a wash, players can assess whether clumps have been broken up and roughly where in the deck certain cards will land, particularly if only a handful are being tracked (such as the board cards from the previous hand). This can still be used to a player's advantage even if the first point is mitigated by the dealer.

Of note, the WSOP Dealer Guide (WSOP, 2013) does not prescribe a wash to be done prior to each hand; a wash is only to be done when a fresh deck of cards is first opened. Indeed, in the aforementioned 2021 WSOP footage (PokerGO, 2021), it can be observed that no wash is performed after the cards are gathered. Moreover, with some dealers in this footage, we observe the cards to be held in a perpendicular manner when squaring them up for the first riffle shuffle, thereby making several cards visible prior to shuffling. Our analyses are thus very relevant to play at the WSOP.

Shuffle tracking, or the act of tracking cards through a series of shuffles, is indeed a well-known phenomenon among gamblers. While it appears anecdotally to be more prevalent in other casino card games outside of poker, it is nevertheless quite probable that poker players have already been taking advantage of the deficiencies of the CS described in this manuscript. Our contribution in this work is to characterize these deficiencies probabilistically, in order to demonstrate the magnitude of the need for casino poker rooms to make changes to their shuffling procedures.

References

Aldous, D. (1983). Random walks on finite groups and rapidly mixing Markov chains. In *Séminaire de Probabilités XVII 1981/82* (pp. 243–297). Springer, Berlin, Heidelberg.

Aldous, D., & Diaconis, P. (1986). Shuffling cards and stopping times. *The American Mathematical Monthly, 93*(5), 333. https://doi.org/10.2307/2323590

Bayer, D. & Diaconis, P. (1992). Trailing the dovetail shuffle to its lair. *The Annals of Applied Probability, 2*(2), 294-313.

Bhatnagar, N., Caputo, P., Tetali, P., & Vigoda, E. (2007). Analysis of top-swap shuffling for genome rearrangements. *The Annals of Applied Probability, 17*(4), 1424–1445.

Diaconis, P., Graham, R. L., & Kantor, W. M. (1983). The mathematics of perfect shuffles. *Advances in Applied Mathematics, 4*(2), 175–196.

Diaconis, P. (1988). *Group representations in probability and statistics.* Lecture Notes–Monograph Series, *11*. Hayward, CA: Institute of Mathematical Statistics.

Diaconis, P., Fill, J. A., & Pitman, J. (1992). Analysis of top to random shuffles. *Combinatorics, Probability and Computing, 1*(2), 135–155.

Diaconis, P., Fulman, J. & Holmes, S. (2013). Analysis of casino shelf shuffling machines. *The Annals of Applied Probability, 23*(4), 1692–1720.

Durrett, R. (2005). Genome Rearrangement. In: *Statistical Methods in Molecular Evolution*. Statistics for Biology and Health. New York, NY: Springer. https://doi.org/10.1007/0-387-27733-1_11.

Gilbert, E. (1955). Theory of shuffling. Technical memorandum, Bell Laboratories.

Hannum, R. (2000). Casino card shuffles: How random are they? *Finding the Edge: Mathematical Analysis of Casino Games*, 33–55.

Hoang, V. T., Morris, B., & Rogaway, P. (2012, August). An enciphering scheme based on a card shuffle. In *Annual Cryptology Conference* (pp. 1–13). Berlin, Heidelberg: Springer.

Kassambara, A. (2020). ggpubr: 'ggplot2' based publication ready plots. R package version 0.4.0. https://CRAN.R-project.org/package=ggpubr.

Live at the Bike. (2021). High Stakes Friday $50/$100 + $100 BB Ante NLH - Live at the Bike! Poker stream. Retrieved from https://www.youtube.com/watch?v=9Tiy7wx7qKY.

Morris, B., & Rogaway, P. (2014, May). Sometimes-recurse shuffle. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (pp. 311–326). Berlin, Heidelberg: Springer.

Pitt, M. (2019). Automatic shufflers, Main Event giveaways, and more at the 2019 WSOP. Retrieved from https://www.pokernews.com/news/2019/04/2019-wsop-automatic-shufflers-giveaways-34042.htm.

PokerGO. (2017). 2017 WSOP main event day 1. Retrieved from https://www.youtube.com/watch?v=zdTlTqqtcIk&t=420s.

PokerGO. (2021). 2021 WSOP main event day 1a. Retrieved from https://www.youtube.com/watch?v=vtbMIsH7Tls&t=921s.

PokerListings. (2014). How to shuffle cards Like a poker dealer - Live poker basics tutorials. Retrieved from https://www.youtube.com/watch?v=c6_8gia6eKU.

PokerStars. (2018). Main event day 2 - EPT Monte Carlo 2018. Retrieved from https://www.youtube.com/watch?v=WuNZEccZOsg.

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

Reeds, J. (1981). Unpublished manuscript.

Ristenpart, T., & Yilek, S. (2013, August). The mix-and-cut shuffle: small-domain encryption secure against *N* queries. In *Annual Cryptology Conference* (pp. 392–409). Berlin, Heidelberg: Springer.

Thorp, E. O. (1973). Nonrandom shuffling with applications to the game of Faro. *Journal of the American Statistical Association, 68*(344), 842–847.

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

WSOP. (2013). 2013 World Series of Poker official dealer guide. Retrieved from https://www.wsop.com/2013/2013_WSOP_Dealer_Guide.pdf.

Appendix

While searching for alternative shuffling procedures, we found that for a shuffling procedure such as the CS that consists of some number of riffle shuffles, a strip shuffle, and a deck cut, the placement of the strip shuffle within the procedure has an impact on the resulting randomness of the shuffled deck. In Figure 6 of the main text, each alternative shuffle has the best ordering of these methods among all possibilities with the same number of riffle shuffles; below, we show how this was determined.

In Figure S1, we show resulting cards apart values from shuffling procedures with a total of three riffle shuffles. From left to right, the first bar corresponds to the procedure with the strip shuffle occurring first, followed by all riffles; the second bar corresponds to a riffle shuffle occurring first, followed by the strip shuffle, and then the remainder of the riffle shuffles, and so on. Thus, the third bar corresponds to the CS, with two riffle shuffles occurring first, followed by a strip shuffle, and ending with a deck cut. The final bar, which is the closest to optimal (as indicated by the dotted line) has all three riffle shuffles occurring first, followed by the strip shuffle, and ending with the cut.

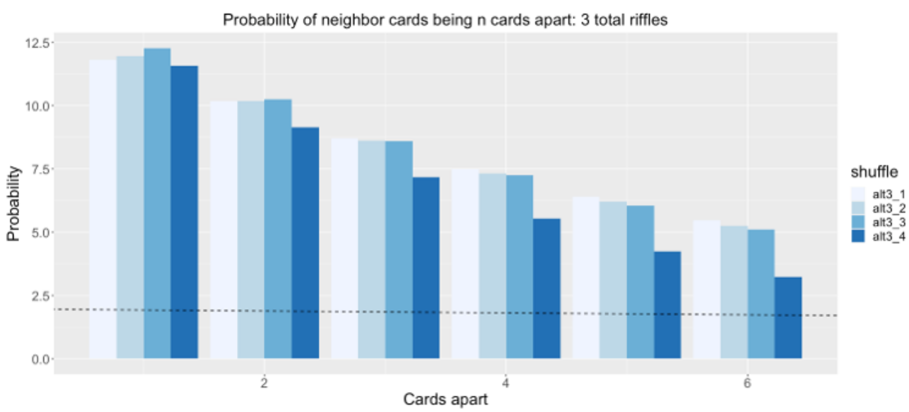

Figure S1

Neighbor bar plot comparison between shuffling methods with three riffle shuffles (100,000 iterations). Each bar represents a different position at which the strip shuffle occurs, in ascending order: the first bar corresponds to the strip shuffle occurring first, and so on.

We note that the relationship between strip shuffle position and probability is not necessarily monotonic (i.e., for cards apart values of 1 and 2, the probability at first increases as the strip shuffle moves later in the procedure, until finally decreasing), which is a consequence of the fact that each bar represents an average across all initial card positions prior to shuffling. Particularly, when examining only the first card in the deck prior to shuffling, the probabilities do monotonically decrease as the strip shuffle moves later in the procedure (results not shown). At other positions, however, this will not necessarily be the case; we posit that the reason is because other card positions have a greater probability of being separated from neighboring cards by the strip shuffle, as dictated by the Binomial distribution described in the Methods section for each type of shuffle. For all values of cards apart, however, the best result overall is achieved by having the strip shuffle occur after all riffle shuffles.

Figures S2 and S3 show the corresponding comparisons from shuffling procedures with a total of four and five riffle shuffles, respectively. In each case, the bar closest to optimal results from all riffle shuffles occurring first, followed by the strip shuffle, and ending with the deck cut. It is these bars that are then shown in Figure 6 of the main text.

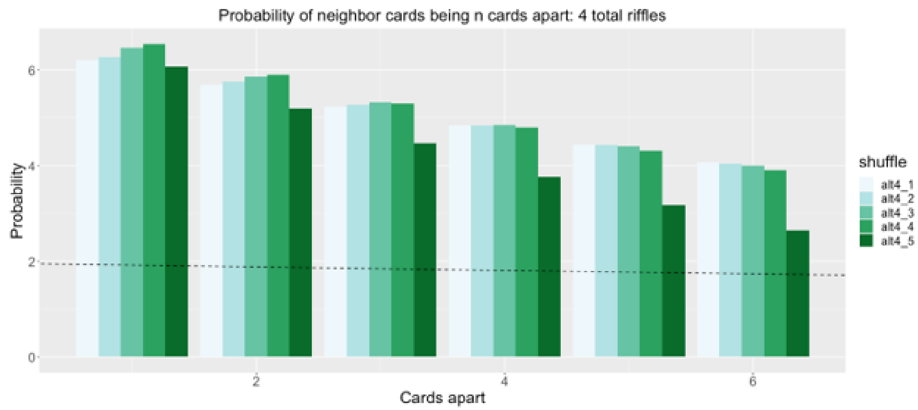

Figure S2
Neighbor bar plot comparison between shuffling methods with four riffle shuffles (100,000 iterations). Each bar represents a different position at which the strip shuffle occurs, in ascending order: the first bar corresponds to the strip shuffle occurring first, and so on, until the final bar which shows all riffles occurring before the strip shuffle.

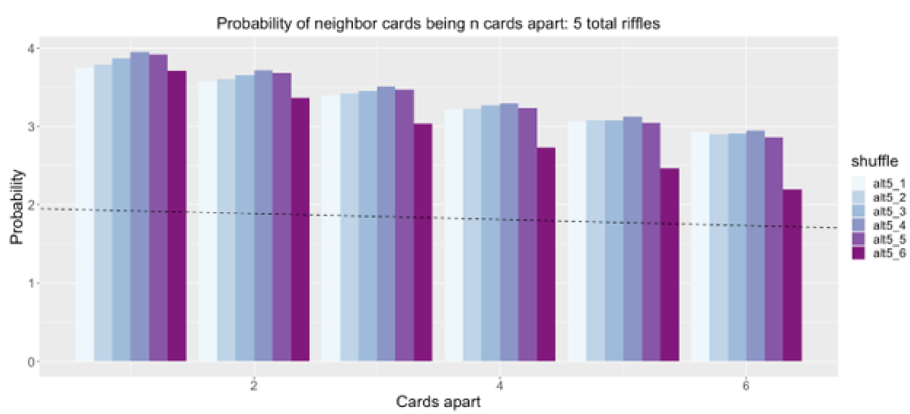

Figure S3
Neighbor bar plot comparison between shuffling methods with five riffle shuffles (100,000 iterations). Each bar represents a different position at which the strip shuffle occurs, in ascending order: the first bar corresponds to the strip shuffle occurring first, and so on, until the final bar which shows all riffles occurring before the strip shuffle.