CLUMPY RIFFLE SHUFFLES

STEFAN WAGER
DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY

JUNE 7, 2011

ABSTRACT. Sticky or clumpy riffle shuffles appear quite naturally in applications; however, the problem of precisely describing the convergence rate of such modified riffle shuffles has remained open for a long time. In this paper, we develop an alternative family of clumpy shuffles, and analyze their convergence rate. We then provide empirical evidence that our alternative shuffles converge at a similar rate as the clumpy riffle shuffle in cases that apply to real shuffling problems.

1. INTRODUCTION

The riffle shuffle – which consists of cutting a deck of cards roughly in half and then riffling the two halves together – is one of the most commonly used card shuffling techniques. A natural model for this shuffle, developed independently by Gilbert and Shannon [10] and by Reeds [14], has led to general results and a fairly good understanding of the riffle shuffle. A Gilbert-Shannon-Reeds (GSR) shuffle consists of the following: Cut the deck roughly in half, where the location of the cut is binomially distributed, and then drop cards one by one from the two half decks A and B such that the probability that the next card comes from deck A is $|A|/(|A| + |B|)$. Bayer and Diaconis [3] give a

This paper was written as an undergraduate honors thesis. I wish to thank my advisor Persi Diaconis for his support, and for always giving me new and interesting questions to think about. I am also grateful to my parents Suzanne de Treville and Hannu Wager for their encouragement.

sharp mathematical analysis of a GSR riffle shuffle; a survey of various results relating to riffle shuffling is given in [7].

The GSR model has an algebraic structure that leads itself well to careful analysis, and allows us to answer fairly intricate questions about riffle shuffles. For example, Conger and Howald [4] describe the impact of dealing techinques after riffle shuffling, and Assaf & al. [2] describe the performance of a riffle shuffle when only certain features (such as suits) matter. Meanwhile, Conger and Viswanath [5], [6] develop a Monte-Carlo method for numerically answering less tractable questions about riffle shuffles (such as what kinds of bridge hands are induced by riffle shuffles). The model also lends itself to developments that are not directly related to playing cards: for example, Diaconis & al. [8] give a description of the cycle structure induced by riffle shuffles. Trefethen and Trefethen use information theoretic methods to get quantitative results about cutoff phenomena in riffle shuffles [15].

We thus have a fairly good idea of how riffle shuffles behave, so long as they obey the GSR model. However, there are much fewer results about the behavior of riffle shuffling outside of the GSR model. Fulman [9] develops a model for a biased riffle shuffle, which differs from a GSR shuffle in that the distribution of the cut to the deck is skewed; and Uyemura-Reyes [16] analyzes "perfect" riffle shuffles, where the only degree of freedom is which half of the deck we start dropping cards from first. The mixing time of the Thorp shuffle, which is another model for a dealer shuffle, was bounded by Morris [11]. To our knowledge, though, there are not many results available that would allow us to do error analysis on riffle shuffles, and explain how the approach to randomness of a riffle shuffle is impacted by a tweak to the shuffling technique.

One question in particular that could be interesting to people shuffling cards – but has yet to be answered precisely – is the following: if a riffle shuffle is clumpy (i.e. more pairs of consecutive cards are preserved than should be according to the GSR model), how will approach to randomness be affected? A related question is the behavior of dealer shuffles, which preserves fewer pairs of consecutive cards than a GSR shuffle. Both of these questions are relevant to "real life" card shuffling situations. Casino dealers tend to have shuffles that are significantly cleaner than the GSR shuffle, which motivates the study of dealer shuffle; in section 1.1, we show evidence that sub-routines of a wash-shuffle can be modeled as clumpy GSR shuffles.

In this paper, we attempt to shed some light on the convergence rate of clumpy riffle shuffles. This one-parameter family of deformations of the GSR shuffle was, to our knowledge, first suggested by Aldous in

1983 [1]. More recently, this model has been mentioned in [6] and [7]. A sharp analysis of the convergence rate of such modified GSR shuffles, however, remains a largely open problem. For the purpose of this paper, we use the following definition of a clumpy riffle shuffle, which is equivalent to the *Markovian Model* suggested in [7] with $p_{ii} \geq 1/2$.

Definition 1.1. A q-clumpy riffle shuffle of a deck of cards, with $0 \leq q < 1$, consists of the following:

- With probability $2^q - 1$, glue consecutive cards together. (Each pair of consecutive cards is considered independently.)
- Perform a GSR riffle shuffle on the resulting deck, where each chunk of glued cards is treated as a single card.
- Un-glue all cards.

We do not attempt to analyze this shuffle directly, but rather give a sharp analysis of a related shuffle: the q-clumpy a-shuffle. We then present empirical evidence that these two shuffles approach randomness at similar rates (at least in the case $n = 52$). If one of our motivations for studying clumpy riffle shuffles is to use such shuffles as a model for real-world processes, it does not seem too unreasonable to use another model which is easier to understand as a tool to approximate clumpy riffle shuffles.

1.1. **Modeling a Wash Shuffle.** The motivation for our study of clumpy riffle shuffles came from an attempt to model the wash shuffle. One promising model of the wash shuffle consists of dividing the deck into multiple small piles, and then randomly shuffling or smushing adjacent piles together and cutting existing piles apart. This model corresponds fairly well qualitatively at least to the way I wash-shuffle cards. When I shuffle a deck of 52 cards, the cards are typically spread out between around 10 piles, with a limited number of overlaps between the piles.

The question then becomes: how can we model the way in which these small piles are smushed together during a wash shuffle? One strategy would be to combine these piles just using a regular GSR riffle shuffle. However, our experiments suggest that the smushing subroutine of a wash shuffle is significantly clumpier than a GSR shuffle.

To gain a better understanding of how cards are smushed together during a wash-shuffle, we ran the following experiment: we smushed two decks of 13 cards together, and then counted the number of clumps of size k preserved by the shuffle. In order to smush the two decks of cards together, we first spread each deck out, and then combined the decks by pushing the cards inward. See figure 2 for an illustration. We

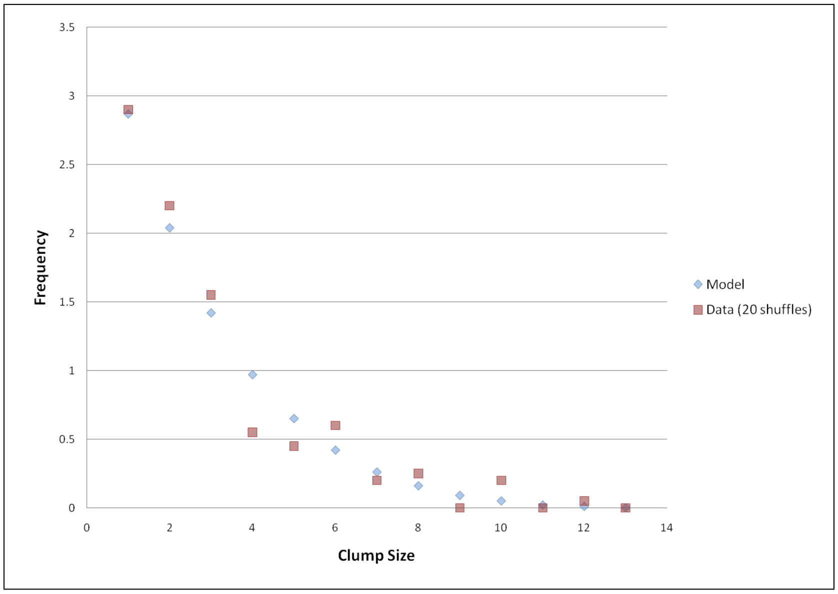

FIGURE 1. Distribution of clump sizes with a wash shuffle

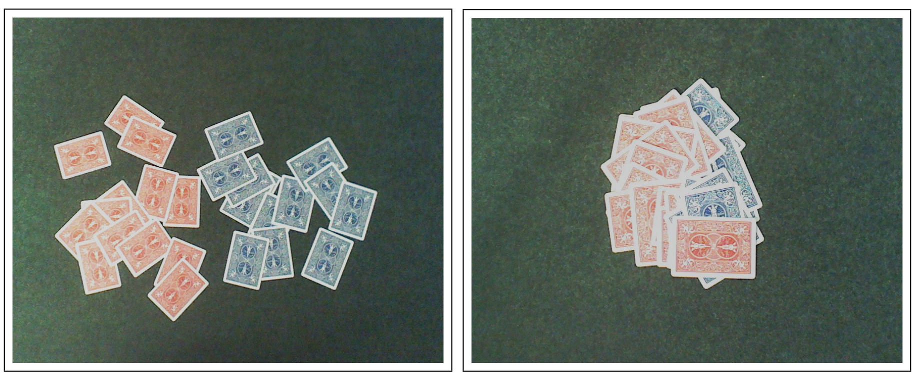

FIGURE 2. Smushing two piles of cards together (before and after)

ran this experiment 20 times. Our results are summarized in figure 1. For each possible clump size $k = 1, ..., 13$ we recorded the average number of times clumps of this size were produced by a wash shuffle.

For comparison, we present the expected number of clumps of size k generated by a clumpy riffle shuffle, where consecutive cards are glued together with probability 0.4 (this corresponds to a clumpiness parameter $q = 0.5$). These values were obtained by simulation. This data seems to present some evidence that the smushing in a wash-shuffle behaves like a clumpy riffle shuffle. As we shall see, a clumpy riffle shuffle with parameter $q = 0.5$ is approximately 4/3 times slower than a regular riffle shuffle, which means we cannot just ignore this clumpiness.

2. The Clumpy a-Shuffle

The Gilbert-Shannon-Reeds (GSR) model for card shuffling has an algebraic structure that lends itself to very natural mathematical abstractions; in fact, it leads to an independent development of Solomon's descent algebra [12]. One of the key properties of the GSR riffle shuffle is that we can generalize it to an a-shuffle, where we obtain an a-shuffle by cutting a deck of cards into a packets according to a multinomial distribution, and then riffling these a packets together. Bayer and Diaconis prove in [3] that an a-shuffle followed by a b-shuffle is equivalent to an ab-shuffle. This result enables them to treat the composition of k riffle shuffles as a single 2^k-shuffle, which opens the door to a whole new suite of combinatorial tools. Unfortunately, clumpy riffle shuffles do not naturally lend themselves to a similar generalization.

The main difficulty with clumpy riffle shuffles is that the distributions generated by clumpy inverse shuffles do not fall into Stanley's theory of QS-distributions, which is the most general form of the distributions commonly used to analyze riffle shuffles [13]. Inverse shuffling a deck k times is equivalent to labeling each a card with a random binary string of length k, and then sorting the cards by label (ties are kept in the initial order). With a regular GSR riffle shuffle, these labels are drawn from a uniform distribution; with Fulman's biased shuffles, these labels are no longer uniform, but all permutations of a given set of labels are equiprobable. This exchangeability condition is necessary for the inverse shuffle to generate a QS-distribution.

Clumpy shuffles, meanwhile, do not satisfy this exchangeability condition: Consider shuffling a 3-card deck twice with a clumpy inverse riffle shuffle, and let p be the probability that two consecutive cards get the same random bit in a single shuffle. Then, the probability of obtaining the labeling $[11, 10, 01]$ is $p^2(1-p)^2/4$, but the probability of obtaining $[11, 01, 10]$ is $p(1-p)^3/4$; thus we do not have the kind of exchangeability property we might have wanted.

This lack of exchangeability also makes it difficult to find stopping times at which the deck is completely random. Under an inverse shuffle with a QS-distribution, once all cards have different labels, we know that the deck has been randomized, since any permutation of the labels is equally likely. With clumpy shuffles this, again, is not true. Suppose we have inverse shuffled the deck abc twice. If we know only that all cards now have distinct labels, we have the following probability distribution on the current state of the deck.

$$P(abc) = \frac{1 - 2p + 2p^2}{2(1 + p)}$$
$$P(bca, cab) = \frac{1 - 2p + 4p^2}{4(1 + p)}$$
$$P(acb, bac, cba) = \frac{p - p^2}{1 + p}$$

Noting these difficulties, in this paper we develop an alternative family of shuffles: the q-clumpy a-shuffles; these shuffles are analoguous to the a-shuffles in [3]. They still do not generate an inverse shuffle with exchangeable labels, but they are simple enough that we can get around this. Clumpy a-shuffles in themselves are not of much interest, since, unlike non-clumpy a-shuffles, the distributions generated by such clumpy a-shuffles do not form a semigroup. However, in section 3, we present empirical evidence that clumpy 2^m-shuffles have similar distributions to the composition of m clumpy riffle shuffles (at least in the case $n = 52$), which suggests that understanding the former can help us understand the latter.

Definition 2.1. A q-clumpy a-shuffle of a deck of cards, with $0 \leq q < 1 \in \mathbb{R}$, $a \geq 2 \in \mathbb{N}$, consists of the following:

- With probability $\frac{a^q - 1}{a - 1}$, glue consecutive cards together. (Each pair of consecutive cards is considered independently.)
- Perform an a-shuffle on the resulting deck, where each chunk of glued cards is treated as a single card.
- Un-glue all cards.

Above, we characterized inverse shuffles in terms of labelings. A more shuffling-oriented way of describing a regular inverse a-shuffle is as follows: Take cards sequentially from the top of the deck, and place them face up on one of a packets (where the packet the card is placed on is selected uniformly at random); then stack the a packets. Our definition of a clumpy a-shuffle also admits a similar inverse description.

Proposition 2.2. *The inverse of a q-clumpy a-shuffle admits the following description: Take a deck of cards, and distribute the cards one by one (face up) into a separate packets, such that a card goes into the same pile as the card before it with probability a^{q-1}. If a card does not follow the card above, the card is equally likely to end up in any of the other piles.*

Proof. The inverse shuffle we just described is equivalent to first gluing consecutive cards together with probability $\frac{a^q-1}{a-1}$, and then performing a regular inverse a-shuffle. To see this, we notice that if consecutive cards are glued together with probability $\frac{a^q-1}{a-1}$, then the probability of one card following the card above it is:

$$\frac{a^q-1}{a-1} + \frac{1}{a} \cdot (1 - \frac{a^q-1}{a-1}) = \frac{1 + (a-1) \cdot \frac{a^q-1}{a-1}}{a} = a^{q-1}$$

We get the desired result by noting the equivalence of the sequential and inverse descriptions of a regular a-shuffle.

□

This inverse description gives some motivation for our q-clumpy a-shuffle. Indeed, although the composition of a q-clumpy a-shuffle and b-shuffle does not give an ab-shuffle, the probability that two originally consecutive cards remain stuck together through both an a-shuffle and then a b-shuffle is the same as the probability that they remain stuck together in an ab-shuffle. We might hope that the relationship between two formerly consecutive cards becomes essentially random as soon as they get separated; in this case, the composition of multiple a-shuffles might have a similar approach to randomness as a single a^k-shuffle.

2.1. **Convergence of the clumpy a-shuffle.** In this section, our goal is to give an estimate of the rate of convergence in total variation distance of the probability measures P_q^a generated by a q-clumpy a-shuffles to the uniform distribution U on S_n. We first obtain bounds on $||P_q^a - P_0^a||_{TV}$, and then use Bayer and Diaconis' result from [3], given below as theorem 2.3, to get estimates for $||P_0^a - U||_{TV}$. Since the transition probabilities of non-clumpy a-shuffles form a semi-group, the GSR distribution Q^m used below is equal to the a-shuffle distribution $P_0^{2^m}$.

Theorem 2.3. *(Bayer, Diaconis) Let Q^m be the Gilbert-Shannon-Reeds distribution on the symmetric group S_n. Let U be the uniform distribution. For $m = log_2[n^{\frac{3}{2}}c]$, with $0 < c < \infty$ fixed, as n tends to ∞,*

$$||Q^m - U||_{TV} = 1 - 2\Phi\left(\frac{-1}{4c\sqrt{3}}\right) + O_c\left(\frac{1}{n^{1/4}}\right)$$

with $\Phi = \int_{-\infty}^{x} e^{-t^2/2}\, dt/\sqrt{2\pi}$.

In order to proceed with our analysis of q-clumpy a-shuffles, we first need to obtain an expression for their transition probabilities. When

$q = 0$, these transition probabilities only depend on the number of descents in the shuffled deck. Once we make the shuffle clumpy, we also have to take into account the number of pairs of consecutive cards preserved by the shuffle.

Theorem 2.4. *Consider a permutation π of $[n]$, such that π^{-1} has d descents (i.e. π has $d+1$ rising sequences), and π preserves f pairs of consecutive cards. Then the probability of obtaining π with a q-clumpy a-shuffle is given by:*

$$P_q^a(\pi) = \frac{1}{a} \sum_{g=0}^{f} a^{g(q-1)} \left(\frac{1-a^{q-1}}{a-1}\right)^{n-g-1} \binom{f}{g} \binom{n+a-1-d-f}{n-g}$$

Proof. We start by recalling the proof given in [3] for the result:

$$P_0^a(\pi) = \frac{1}{a^n} \binom{n+a-1-d}{n}$$

The proof relies on the following idea. An inverse riffle shuffle is equivalent to assigning each card a random label between 1 and a, and then sorting the cards first by label, and then, in case of tie, by their original order. Notice that if $\pi(k) > \pi(k+1)$, i.e. π^{-1} has a descent at k, then the cards k and $k+1$ must have distinct labels. The result then follows by a classical bars and stars argument: We need to divide n cards (in the order given by π^{-1}) into a packets, with the additional constraint that cards k_i and k_{i+1} not be in the same packet for $1 \leq i \leq d$. We thus have $a-1-d$ remaining separators that we can freely place between the cards, which we can do in $\binom{n+a-1-d}{a-1-d}$ ways. To complete the argument, we note that there are a^n equiprobable ways of performing an a-shuffle.

In the case of a q-clumpy a-shuffle, we no longer can assume that all ways of performing an a-shuffle are equiprobable. In fact, if a given shuffle has g pairs of consecutive cards that followed each other into the same packet during the shuffle, then, by the inverse description of the q-clumpy a-shuffle in proposition 2.2, such a shuffle must have probability

$$p_g = \frac{1}{a} \cdot a^{g(q-1)} \cdot \left(\frac{1-a^{q-1}}{a-1}\right)^{n-g-1}$$

Here, the a^{-1} factor up front indicates that the first card is equally likely to end up in any pile. We now choose g pairs of consecutive cards that are preserved by π, and add the constraint that these g card pairs must be kept together by the shuffle, while the other $f-g$ pairs that were preserved by π must be separated by the shuffle (i.e. there must be a packet separation placed between them). We can emulate the bars and stars argument from above to find the number of shuffles that satisfy this constraint: Having g pairs of cards that must follow each other throughout the shuffle is equivalent to gluing these cards together, and mapping $n \mapsto n - g$ in the above argument. Meanwhile, the $f - g$ pairs of cards that must be separated fix the location of $f - g$ additional separators (in exactly the same way as descents); this can be described by sending $d \mapsto d + f - g$ in the same arguement. Thus, given that we chose exactly which g pairs of cards should follow each other through the shuffle, we find that we can perform an a-shuffle satisfying all our constraints in $N_g^a(\pi)$ ways:

$$N_g^a(\pi) = \binom{n + a - 1 - d - f}{n - g}$$

We notice in addition that all the $N_g(\pi)$ shuffles described above have probability p_g. Finally, summing over all choices of g pairs we could glue together, we find that:

$$P_q^a(\pi) = \sum_{g=0}^{f} p_g \binom{f}{g} N_g^a(\pi)$$

□

We can conveniently formulate the above result as follows:

Corollary 2.5. *For a permutation π of $[n]$ such that π^{-1} has d descents and π preserves f pairs of consecutive cards:*

$$\frac{P_q^a(\pi)}{P_0^a(\pi)} = \sum_{g=0}^{f} a^{qg} \left(\frac{1 - a^{q-1}}{1 - a^{-1}}\right)^{n-g-1} \binom{f}{g} \frac{\binom{n+a-1-d-f}{n-g}}{\binom{n+a-1-d}{n}}$$

Under the uniform distribution, the vast majority of permutations have only very few consecutive elements. To see this, consider the natural bijection $\sigma : S_n \to Aut([n])$ given by $\sigma(\pi)(k) = \pi(k+1) - 1 \pmod{n}$. Notice that this bijection associates pairs of consecutive cards in π to fixed points of $\sigma(\pi)$. The distribution of the number of pairs of consecutive cards in random permutations is thus the same as that of the number of fixed points in a random automorphism over a finite set, which is Poisson(1) for n large.

Using this remark, we can give fairly tight bounds for the convergence rate of $||P_q^a - P_0^a||_{TV}$.

Lemma 2.6. *Let $q \in [0,1]$ be fixed, and let P_*^a be clumpy a-shuffle measures on S_n such that $a(n) = (n/\phi(n))^{\frac{1}{1-q}}$ where ϕ is a bounded function of n. Then, for n large,*

$$||P_q^a - P_0^a||_{TV} = \left(\frac{1}{e} \pm ||P_0^a - U||_{TV}\right) \cdot \left(1 - e^{-\phi(n)} + \frac{e^{\phi(n)} - \phi(n) - 1}{e^{\phi(n)}}\right) + O\left(\frac{1}{n},\ \left(\frac{\phi(n)}{n^q}\right)^{\frac{1}{1-q}}\right)$$

with $\pm\zeta$ indicating an error term bounded by ζ.

Proof. We begin by approximating our expression in corollary 2.5. To use this equation, however, we must choose permutations $\pi(n)$ in S_n with a number of descents $d(n)$ and of consecutive pairs $f(n)$. For this purpose, we select a sequence $d(n) \leq n$ (as we shall see the $d(n)$ disappear into the error terms), and hold $f(n)$ to a fixed value (recall that the distribution of f does not depend on n for n large). We start by observing that

$$\log\left[\frac{\binom{n+a-d-1-f}{n-g}}{\binom{n+a-d-1}{n}}\right] = \log\left[\frac{n_g}{(a-d-1+g)_g}\right] + \log\left[\frac{(a-d-1+g)_f}{(n+a-d-1)_f}\right] = g\log\left[\frac{n}{a}\right] + O_f\left(\frac{1}{n},\ \frac{n}{a}\right)$$

We also notice that as a gets large

$$\log\left[\frac{1-a^{q-1}}{1-a^{-1}}\right] = \frac{1-a^q}{a} + O(a^{2(q-1)})$$

Now, since all the terms in our sum for $P_q^a/P_0^a(\pi)$ are strictly positive, and there are only f terms in this sum for each n, we can bound all these log-space errors ζ_g using a single error term $\zeta_f = O_f(1/n,\ n/a)$.

$$
\begin{aligned}
\log\left[\frac{P_q^a}{P_0^a}(\pi)\right] &= \log\left[\sum_{g=0}^{f}\binom{f}{g}\left(\frac{na^q}{a}\right)^g\left(\frac{1-a^{q-1}}{1-a^{-1}}\right)^{n-g-1}e^{\zeta_g}\right] \\
&= \log\left[\left(\frac{1-a^{q-1}}{1-a^{-1}}\right)^{n-f-1}\left(na^{q-1}+\frac{1-a^{q-1}}{1-a^{-1}}\right)^f\right]+\zeta_f \\
&= f\log\left[1+na^{q-1}\frac{1-a^{-1}}{1-a^{q-1}}\right]+(n-1)\left[\frac{1-a^q}{a}+O(a^{2(q-1)})\right]+\zeta_f \\
&= f\log\left[1+na^{q-1}(1+O(a^{q-1}))\right]-na^{q-1} \\
&\quad + O_f\left(\frac{1}{n},\ \frac{n}{a},\ na^{2(q-1)},\ a^{q-1}\right) \\
&= f\log[1+na^{q-1}]-na^{q-1}+O_f\left(\frac{1}{n},\ \frac{n}{a}\right)
\end{aligned}
$$

On the last line, we used the assumption that na^{q-1} is bounded to simplify the error term. Now for large enough n, the ratio $\frac{P_q^a}{P_0^a}(\pi)$ is greater than 1 for $f \geq 2$. Thus, in order to study the asymptotic behavior of $||P_q^{a_k} - P_0^{a_k}||_{TV}$, we only need to consider the subset of S_n with $f = 0, 1$. Using a slight abuse of notation (P_0^a and P_q^a clearly have densities with respect to U, and share the same null sets so have densities with respect to each other), and error terms $\zeta_f = O_f(n^{-1},\ n/a)$, we see that for a large enough,

$$
\begin{aligned}
||P_q^a - P_0^a||_{TV} =& \int_{\pi \in S_n : (P_0^a - P_q^a)(\pi) > 0} (P_0^a - P_q^a)(\pi) \; dU \\
=& \int_{\pi \in S_n : f = 0,1} 1 - \frac{P_q^a}{P_0^a}(\pi) \; dP_0^q \\
=& \int_{\pi \in S_n : f = 0} 1 - e^{-na^{q-1} + \zeta_0} \; dP_0^q \\
& + \int_{\pi \in S_n : f = 1} 1 - (1 + na^{q-1}) e^{-na^{q-1} + \zeta_1} \; dP_0^q \\
=& P_0^a(\{f = 0\}) \cdot (1 - e^{-na^{q-1}}) \\
& + P_0^a(\{f = 1\}) \cdot (e^{na^{q-1}} - na^{q-1} - 1) e^{-na^{q-1}} + \zeta \\
=& (P_U(\{f = 0\} \pm ||P_0^a - U||_{TV}) \cdot (1 - e^{-na^{q-1}}) \\
& + (P_U(\{f = 1\} \pm ||P_0^a - U||_{TV}) \cdot \frac{e^{na^{q-1}} - na^{q-1} - 1}{e^{na^{q-1}}} + \zeta \\
=& \left(\frac{1}{e} \pm ||P_0^a - U||_{TV} \right) \cdot \left(1 - e^{-na^{q-1}} + \frac{e^{na^{q-1}} - na^{q-1} - 1}{e^{na^{q-1}}} \right) \\
& + O\left(\frac{1}{n}, \; \frac{n}{a} \right)
\end{aligned}
$$

On the last line, we use the fact that f is distributed according to Poisson(1) under U for large n, and $P_U(\{f = 0\}) = P_U(\{f = 1\}) = 1/e + O(1/n!)$. Finally, writing a in terms of $\phi(n)$ gives us the desired result.

□

It is not surprising that the clumpy a-shuffle converges to the regular GSR shuffle as na^{q-1} becomes small. As we saw in proposition 2.2, the probability that any two cards are stuck together before being riffled in an q-clumpy a-shuffle is a^{q-1}; thus, when na^{q-1} is small, the expected number of pairs of consecutive cards preserved by the clumpy shuffle is approximately $E_{n,a} + (n-1)a^{q-1}$, where $E_{n,a}$ is the expected number of clumps under a non-sticky a-shuffle.

Combining the above result with the bound for $||P_0^a - U||_{TV}$ given in theorem 2.3, we can get our desired bound for $||P_q^{a_k} - U||_{TV}$. If $q < \frac{1}{3}$, the clumpiness of the shuffle disappears before P_0^a approaches randomness; on the other hand, if $q > \frac{1}{3}$, clumpiness becomes the dominating source of non-randomness.

Theorem 2.7. *Let P_q^a be the q-clumpy a-shuffle distribution on S_n for some fixed $q \in [0,1]$; let U be the uniform distribution.*

If $q < \frac{1}{3}$, let $a = cn^{\frac{3}{2}}$, where $0 < c < \infty$. Then, as n goes to ∞,

$$||P_q^a - U||_{TV} = 1 - 2\Phi\left(\frac{-1}{4c\sqrt{3}}\right) + O_c\left(n^{-\min[\frac{1}{4}, \frac{1-3q}{2}]}\right)$$

with $\Phi = \int_{-\infty}^{x} e^{-t^2/2}\, dt/\sqrt{2\pi}$.

Conversely, if $q > \frac{1}{3}$, let $a = (n/c)^{\frac{1}{1-q}}$, with $0 < c < \infty$. Then, as n gets large,

$$||P_q^a - U||_{TV} = \frac{1}{e}\left(1 - e^{-c} + \frac{e^c - c - 1}{e^c}\right) + O_c\left(n^{-\min[\frac{1}{4}, \frac{3q-1}{2(1-q)}]}\right)$$

Proof. We start with the case $q < \frac{1}{3}$. By the triangle inequality, we know that

$$||P_q^a - U||_{TV} = ||P_0^a - U||_{TV} \pm ||P_q^a - P_0^a||_{TV}$$

with $\pm\zeta$ indicating an error term bounded by ζ. If we set

$$\phi(n) = \frac{n}{a^{1-q}} = \frac{c^{q-1}}{n^{\frac{1-3q}{2}}}$$

We find, as in lemma 2.6, that

$$\begin{aligned} ||P_q^a - P_0^a||_{TV} &= \left(\frac{1}{e} \pm ||P_0^a - U||_{TV}\right) \cdot \left(1 - e^{-\phi(n)} + \frac{e^{\phi(n)} - \phi(n) - 1}{e^{\phi(n)}}\right) \\ &\quad + O\left(\frac{1}{n}, \left(\frac{\phi(n)}{n^q}\right)^{\frac{1}{1-q}}\right) \\ &= O_c\left(n^{-\frac{1-3q}{2}}\right) + O_c\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

Thus, using theorem 2.3, we get that

$$||P_q^a - U||_{TV} = 1 - 2\Phi\left(\frac{-1}{4c\sqrt{3}}\right) + O_c\left(n^{-\min[\frac{1}{4}, \frac{1-3q}{2}]}\right)$$

Conversely, in the case $q > 1/3$, we know from theorem 2.3 that, with $m = \gamma n^{\frac{3}{2}}$,

$$||P_0^m - U||_{TV} = 1 - 2\Phi\left(\frac{-1}{4\gamma\sqrt{3}}\right) + O_\gamma\left(\frac{1}{\sqrt[4]{n}}\right)$$

From [3] it is clear that the error term is bounded in γ as γ gets large. Moreover, for large γ, we have the asymptotic behavior

$$1 - 2\Phi\left(\frac{-1}{4\gamma\sqrt{3}}\right) \sim \frac{1}{2\gamma\sqrt{6\pi}}$$

Thus, with $a = \kappa n^{\frac{1}{1-q}}$ and $q > 1/3$, we find that

$$||P_0^a - U||_{TV} = O_\kappa\left(n^{-\min[\frac{3q-1}{2(1-q)},\frac{1}{4}]}\right)$$

Finally, all the other errors in our expression for $||P_q^a - P_0^a||_{TV}$ in lemma 2.6 decay faster than $1/\sqrt{n}$ since $q > 1/3$, so an application of the triangle inequality gives us our conclusion.

□

Clumpiness thus disappears faster than the rising-sequence signature with $q < 1/3$, but becomes the asymptotically dominant source of non-randomness in a q-clumpy a-shuffle once $q > 1/3$. However, when q is near 1/3, both sources of non-randomness decay at fairly similar rates, which makes the error terms decay quite slowly.

3. Clumpy Riffle Shuffles

In the previous section, we developed a fairly precise description of the clumpy a-shuffle. Of course, this process is not really a shuffle in the regular sense. The non-clumpy a-shuffle distribution $P_0^{2^m}$ is equal to the GSR distribution Q_0^m; however, there is no such exact correspondance between the clumpy a-shuffle and clumpy riffle shuffle distributions $P_q^{2^m}$ and Q_q^m. It appears, however, that $P_q^{2^m}$ and Q_q^m converge to the uniform distribution at similar rates; if this were true, then we could use our above results on the asymptotic behavior of $||P_q^{2^m} - U||_{TV}$ to approximate the behavior of $||Q_q^m - U||_{TV}$ for large m.

In this section, we present some empirical evidence that $P_q^{2^m}$ and Q_q^m converge at similar rates, at least in the case where $n = 52$ and q is not too close to 1. As a proxy for the quantity $||Q_q^m - U||_{TV}$, we used the probability that a shuffle preserves no pairs of consecutive cards under Q_q^m. Indeed, as we saw in the previous section, it is this probability under $P_q^{2^m}$ that explains most of the separation between $P_q^{2^m}$ and U for q large enough.

In figure 3, we see how the probability of obtaining no consecutive cards, which should be $1/e$ under the uniform distribution, approaches this uniform probability as we increase the number of shuffles m. We notice that although Q_q^m and $P_q^{2^m}$ don't exactly match, they converge at very similar rates (especially in comparison to Q_0^m).

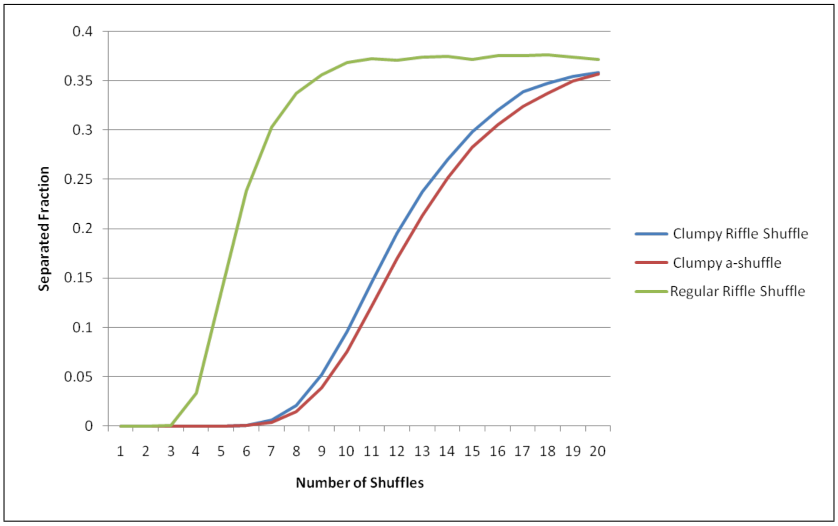

FIGURE 3. Probability of obtaining no consecutive cards; with n = 52 and q = 0.5

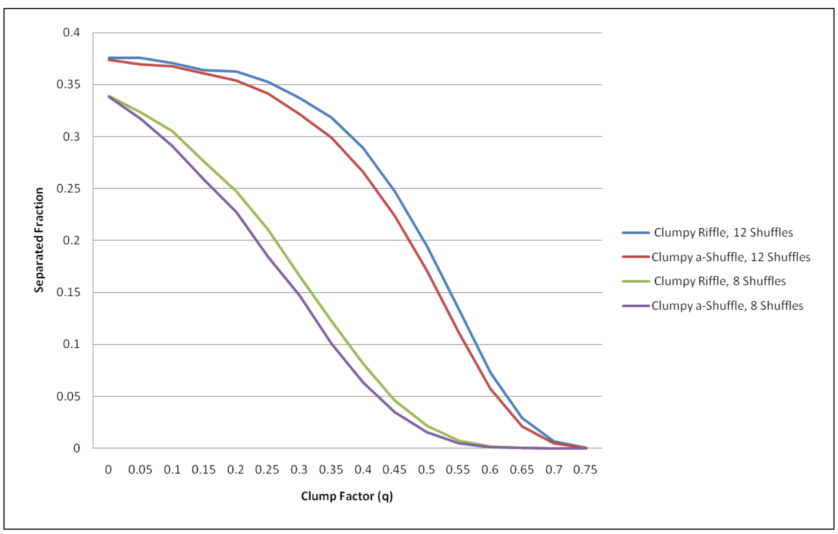

FIGURE 4. Probability of obtaining no consecutive cards; with n = 52

In figure 4 we vary the clumpiness parameter q instead of the number of shuffles, and see how the clumpy shuffles get further from the uniform distribution as q increases.

Although these results are far from providing any proof that P_q^{2m} and Q_q^m have similar rates of convergence in general, they do provide some evidence that, at least in the cases that we might care about most (e.g. the smush shuffle example with $n = 52$, $q = 0.5$), P_q^{2m} can be a useful device in understanding the approach to randomness of Q_q^m.

3.1. **How Clumpy is Too Clumpy?** Armed with our results from the previous sections, we might consider the following conjecture as a corollary to theorem 2.7:

Conjecture 3.1. *When shuffling a deck of n cards using a clumpy riffle shuffle of parameter q, m shuffles are both necessary and sufficient to randomize the deck of cards, where*

$$m = \max\left[\frac{3}{2}, \frac{1}{1-q}\right] \log_2(n) + O_{q,\epsilon}(1)$$

Here, the deck is considered to be random once $||Q_q^m - U||_{TV} < \epsilon$.

The lower bound in conjecture 3.1 is fairly easy to prove directly using techniques similar to those used in theorem 2.4.

Lemma 3.2. *When shuffling a deck of n cards using a clumpy riffle shuffle of parameter q, at least*

$$m = \frac{1}{1-q} \log_2(n) + O_{q,\epsilon}(1)$$

shuffles are necessary to get $||Q_q^m - U||_{TV} < \epsilon$.

Proof. Any permutation $\pi \in N_0$ that preserves no pairs of consecutive cards can only be generated by a separating shuffle $\sigma \in S_0$, where a separating shuffle is a shuffle that separates each pair of initially consecutive cards at some point during the shuffle (but the pair can later be reunited).

The probability of never separating a given pair of adjacent cards over the course of k q-clumpy riffle shuffles is $2^{k(q-1)}$. Moreover, from definition 1.1, it is clear that all $n-1$ initial pairs of consecutive cards are either separated at some point or preserved through all k riffles independently. Thus, the probability of obtaining a separating shuffle is $(1 - 2^{k(q-1)})^{n-1}$. It follows that

$$P(\pi \in N_0) \leq P(\sigma \in S_0) = (1 - 2^{k(q-1)})^{n-1} \leq e^{-(n-1)2^{k(q-1)}}$$

Now, since the number of consecutive cards preserved by π is very nearly Poisson distributed for large n, we know that $P(\pi \in N_0) = e^{-1} + O(1/n!)$, and so, for $(n-1)2^{k(q-1)} > 1$

$$||Q_q^k - U||_{TV} \geq e^{-1} - e^{-(n-1)2^{k(q-1)}}$$

Thus, in order to have $||Q_q^m - U||_{TV} < \epsilon$, we need

$$(n-1)2^{k(q-1)} < 1 - \log(1 - e \cdot \epsilon)$$

which leads directly to our conclusion.

□

This lemma, combined with 2.3, gives us the desired lower bound in conjecture 3.1.

This result could be useful, for example, when we're playing cards and need to decide whether to make our friends who have very clumpy shuffling technique shuffle the deck more times than everyone else. Our first task, though, is to translate the clumpiness parameter q into something that can be easily measured.

Definition 3.3. The packetification ψ of a shuffle is the expected number of packets of consecutive cards formed by the shuffle, divided by the total number of cards shuffled.

The relationship between q and ψ in a clumpy riffle shuffle is quite neat.

Lemma 3.4. *The packetification ψ of a q-clumpy riffle shuffle is given by*

$$\psi(n) = 1 - 2^{q-1} + \zeta_n$$

where $\zeta_n \leq \frac{1}{n}$

Proof. The probability of obtaining a given set of j clumps with a q-clumpy riffle shuffle of n cards is clearly $2^{(n-j)\cdot(q-1)}(1-2^{q-1})^{j-1}$. Moreover, we can choose these j clumps in $\binom{n-1}{j-1}$ ways (by placing $j-1$ separators between clumps), so the expected number of clumps given by a q-clumpy riffle shuffle is given by:

$$\begin{aligned}\psi(n)\cdot n &= \sum_{j=1}^{n} j\binom{n-1}{j-1}2^{(n-j)\cdot(q-1)}(1-2^{q-1})^{j-1} \\ &= 1+\sum_{j=1}^{n}(j-1)\binom{n-1}{j-1}2^{(n-j)\cdot(q-1)}(1-2^{q-1})^{j-1} \\ &= 1+(n-1)\sum_{k=0}^{n-2}\binom{n-2}{k}2^{(n-2-k)\cdot(q-1)}(1-2^{q-1})^{k+1} \\ &= 1+(n-1)(1-2^{q-1})[(2^{q-1})+(1-2^{q-1})]^{n-2} \\ &= n\cdot(1-2^{q-1})+2^{q-1}\end{aligned}$$

For n large enough, we get that ψ is very nearly $1-2^{q-1}$.

□

We thus find that

$$q=\log_2(1-\psi(n))+1-\log_2(1-1/n)\approx\log_2(1-\psi)+1$$

This lets us re-write our conjecture in terms of ψ. For large enough n,

$$m=\min\left[\frac{3}{2},-\frac{1}{\log_2(1-\psi)}\right]\log_2(n)+\theta_\psi$$

shuffles are both necessary and sufficient to achieve randomness. This result implies that clumpiness becomes the asymptotically dominant source of non-randomness in riffle shuffle once $\psi>1-2^{-\frac{2}{3}}\approx 0.37$. An easy way to see if someone's shuffle is this clumpy is to have them shuffle a deck a 52 cards a few times and check whether, on average, their shuffle divides the deck into less than 20 packets.

We recall, though, that as q is near 1/3, the error terms decay quite slowly in n. This means that, even in the case $n=52$, clumpiness can have quite a significant impact on approach to randomness even when q is somewhat less than 1/3. In the case of a non-clumpy riffle shuffle, we know that 7 or 8 shuffles suffice to reach randomness. Meanwhile, from figure 4, we might argue that the deck is still unacceptably non-random after 8 shuffles once q is greater than, say, 0.2. Thus, if we expect a deck of 52 cards to be random after 8 shuffles, we should be wary of shufflers with a packetification factor $\psi\leq 0.43$, or who regularly obtain less than 23 packets when shuffling a standard deck of cards.

4. Conclusions

In this paper, we developed a class of q-clumpy a-shuffles, and proved results concerning the convergence of q-clumpy a-shuffle measures to the uniform measure. We also used empirical evidence to suggest that the convergence rate of clumpy a-shuffles can be used to approximate the convergence rate of a sequence of clumpy riffle shuffles, at least in some applications.

We gave a tight lower bound for the convergence rate of $||Q_q^m - U||_{TV}$ in lemma 3.2. It would be interesting to see whether there are conditions on q under which the corresponding upper bound given in conjecture 3.1 also holds. Another topic for further study would be to see whether similar techniques can be used to study the approach to randomness of dealer shuffles, which can be described as clumpy riffle shuffles with negative q.

References

[1] Aldous. D., Random walks on finite groups and rapidly mixing Markov chains. *Seminar on Probability XVII, Lecture Notes in Mathematics*, **986** (1983), 243-297.

[2] Assaf, S., Diaconis, P., and Soundararajan, K., A Rule of Thumb for Riffle Shuffling, To appear in: *Annals of Applied Probability*, (2008).

[3] Bayer, D. and Diaconis, P., Trailing the Dovetail Shuffle to its Lair, *Annals of Applied Probability*, **2.2** (1992), 294–313.

[4] Conger, M. and Howald, J., A Better Way to Deal the Cards, *The American Mathematical Monthly*, **117.8** (2010), 686–700.

[5] Conger, M. and Viswanath, D., Normal approximations for descents and inversions of permutations of multisets. *Journal of Theoretical Probabability*, **20.2** (2007), 309-325.

[6] Conger, M. and Viswanath, D., Shuffling Cards for Blackjack, Bridge, and Other Card Games, *arXiv* (2006), `http://arxiv.org/abs/math/0606031v1`.

[7] Diaconis, P., Mathematical Developments from the Analysis of Riffle-Shuffling. In Fuanou A. and Liebeck M. (eds.) *Groups Combinatorics and Geometry*, World Scientific, N.J. (2003), 73–97.

[8] Diaconis, P., McGrath, M. and Pitman J., Riffle Shuffles, Cycles and Descents, *Combinatorica*, **15.1** (1995), 11–29.

[9] Fulman, J., The Combinatorics of Biased Riffle Shuffles, *Combinatorica*, **18.2** (1998), 173–184.

[10] Gilbert, E., Theory of Shuffling, *Technical Report*, 1955, Bell Laboratories.

[11] Morris, B., The mixing time of the thorp shuffle, *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, Baltimore, ACM 2005 (2005), 403–412.

[12] Solomon, S., A Mackey formula in the group ring of a Coxeter group, *Journal of Algebra*, **41** (1976), 255–268.

[13] Stanley, R., Generalized Riffle Shuffles and Quasisymmetric Functions, *Annals of Combinatorics*, **5** (2001), 479–491.

[14] Reeds, J., Theory of shuffling, unpublished manuscript, 1981.
[15] Trefethen, L. and Trefethen, L., How many shuffles to randomize a deck of cards?, *Proceedings: Mathematical, Physical and Engineering Sciences*, **456.2002** (2000), 2561-2568.
[16] Uyemura-Reyes, J.C., Random walk, semi-direct products, and card shuffling, *Ph.D. Thesis*, 2002, Dept. of Mathematics, Stanford University.