ANALYSIS OF CASINO SHELF SHUFFLING MACHINES

By Persi Diaconis[*] , Jason Fulman [†] and Susan Holmes[‡]

Many casinos routinely use mechanical card shuffling machines. We were asked to evaluate a new product, a shelf shuffler. This leads to new probability, new combinatorics, and to some practical advice which was adopted by the manufacturer. The interplay between theory, computing, and real-world application is developed.

1. Introduction. We were contacted by a manufacturer of casino equipment to evaluate a new design for a casino card-shuffling machine. The machine, already built, was a sophisticated "shelf shuffler" consisting of an opaque box containing ten shelves. A deck of cards is dropped into the top of the box. An internal elevator moves the deck up and down within the box. Cards are sequentially dealt from the bottom of the deck onto the shelves; shelves are chosen uniformly at random at the command of a random number generator. Each card is randomly placed above or below previous cards on the shelf with probability 1/2. At the end, each shelf contains about 1/10 of the deck. The ten piles are now assembled into one pile, in random order. The manufacturer wanted to know if one pass through the machine would yield a well-shuffled deck.

Testing for randomness is a basic task of statistics. A standard approach is to design some ad hoc tests such as: Where do the original top and bottom cards wind up? What is the distribution of cards that started out together? What is the distribution, after one shuffle, of the relative order of groups of consecutive cards? Such tests *had* been carried out by the engineers who designed the machine, and seemed satisfactory.

We find closed-form expressions for the probability of being at a given permutation after the shuffle. This gives exact expressions for various global distances to uniformity, e.g., total variation. These suggest that the machine has flaws. The engineers (and their bosses) needed further convincing; using our theory, we were able to show that a knowledgeable player could guess about 9 1/2 cards correctly in a single run through a 52-card deck. For a well-shuffled deck, the optimal strategy gets about 4 1/2 cards correct. This data *did* convince the company. The theory also suggested a useful remedy. Journalist accounts of our shuffling adventures can be found in Klarreich (2002, 2003); Mackenzie (2002).

Section 2 gives background on casino shufflers, needed probability, and the literature of shuffling. Section 3 gives an analysis of a single shuffle; we give a closed formula for the chance that a deck of n cards passed through a machine with m shelves is in final order w. This is used to compute several classical distances to randomness. In particular it is shown that, for n cards, the $l(\infty)$ distance is asymptotic to $e^{1/12c^2} - 1$ if the number of shelves $m = cn^{3/2}$ and n is large. The combinatorics of shelf shufflers turns out to have connections to the "peak algebra" of algebraic combinatorics. This allows nice formulae for the distribution of several classical test statistics: the cycle structure (e.g., the number of fixed points), the descent structure, and the length of the longest increasing subsequence.

Section 4 develops tools for analyzing repeated shelf shuffling. Section 5 develops our "how many can be correctly guessed" tests. This section also contains our final conclusions.

2. Background. This section gives background and a literature review. Section 2.1 treats shuffling machines; Section 2.2 gives probability background; Section 2.3 gives an overview of related literature and results on the mathematics of shuffling cards.

2.1. *Card shuffling machines.* Casinos worldwide routinely employ mechanical card-shuffling machines for games such as blackjack and poker. For example, for a single deck game, two decks are used. While the dealer is using the first deck in the usual way, the shuffling machine mixes the second deck. When the first deck is used up (or perhaps half-used), the second deck is brought into play and the first deck is inserted into the machine. Two-, four-, and six-deck machines of various designs are also in active use.

The primary rationale seems to be that dealer shuffling takes time and use of a machine results in approximately 20% more hands per hour. The machines may also limit dealer cheating.

The machines in use are sophisticated, precision devices, rented to the casino (with service contracts) for approximately $500 per month per machine. One company told us they had about 8,000 such machines in active use; this amounts to millions of dollars per year. The companies involved are substantial businesses, listed on the New York Stock Exchange.

One widely used machine simulates an ordinary riffle shuffle by pushing two halves of a single deck together using mechanical pressure to make the halves interlace. The randomness comes from slight physical differences in alignment and pressure. In contrast, the shelf shufflers we analyze here use computer-generated pseudo-random numbers as a source of their randomness.

The pressure shufflers require multiple passes (perhaps seven to ten) to adequately mix 52 cards. Our manufacturer was keen to have a single pass through suffice.

2.2. *Probability background.* Let S_n denote the group of permutations of n objects. Let $U(\sigma) = 1/n!$ denote the uniform distribution on S_n. If P is a probability on S_n, the total variation, separation, and l_∞ distances to uniformity are

(2.1)
$$\|P - U\|_{\mathrm{TV}} = \frac{1}{2} \sum_w |P(w) - U(w)| = \max_{A \subseteq S_n} |P(A) - U(A)| = \frac{1}{2} \max_{\|f\|_\infty \leq 1} |P(f) - U(f)|,$$

$$\mathrm{sep}(P) = \max_w \left(1 - \frac{P(w)}{U(w)} \right), \qquad \|P - U\|_\infty = \max_w \left| 1 - \frac{P(w)}{U(w)} \right|.$$

Note that $\|P - U\|_{\mathrm{TV}} \leq \mathrm{sep}(P) \leq \|P - U\|_\infty$. The first two distances are less than 1; the $\| \ \|_\infty$ norm can be as large as $n! - 1$.

If one of these distances is suitably small then many test statistics evaluate to approximately the same thing under P and U. This gives an alternative to ad hoc tests. The methods developed below allow exact evaluation of these and many further distances (e.g., chi-square or entropy).

Repeated shuffling is modeled by convolution:

$$P * P(w) = \sum_v P(v)P(wv^{-1}), \qquad P^{*k}(w) = P * P^{*(k-1)}(w).$$

All of the shelf shufflers generate ergodic Markov chains (even if only one shelf is involved), and so $P^{*k}(w) \to U(w)$ as $k \to \infty$. One question of interest is the quantitative measurement of this convergence using one of the metrics above.

2.3. *Previous work on shuffling.*

Early work.   The careful analysis of repeated shuffles of a deck of cards has challenged probabilists for over a century. The first efforts were made by Markov (1906) in his papers on Markov chains. Later, Poincaré (1912) studied the problem. These great mathematicians proved that *in principle* repeated shuffling *would* mix cards at an exponential rate but gave no examples or quantitative methods to get useful numbers in practical problems.

Borel and Chéron (1955) studied riffle shuffling and concluded heuristically that about seven shuffles would be required to mix 52 cards. Emile Borel also reported joint work with Paul Levy, one of the great probabilists of the twentieth century; they posed some problems but were unable to make real progress.

Isolated but serious work on shuffling was reported in a 1955 Bell Laboratories report by Edgar Gilbert. He used information theory to attack the problems and gave some tools for riffle shuffling developed jointly with Claude Shannon.

They proposed what has come to be called the Gilbert–Shannon–Reeds model for riffle shuffling; this presaged much later work. Thorp (1973) proposed a less realistic model and showed how poor shuffling could be exploited in casino games. Thorp's model is analyzed in Morris (2009). Epstein (1977) reports practical studies of how casino dealers shuffle with data gathered with a very precise microphone! The upshot of this work was a well-posed mathematics problem and some heuristics; further early history appears in Chapter 4 of Diaconis (1988).

The modern era.   The modern era in quantitative analysis of shuffling begins with papers of Diaconis and Shahshahani (1981) and Aldous (1983). They introduced rigorous methods, Fourier analysis on groups, and coupling. These gave sharp upper and lower bounds, suitably close, for real problems. In particular, Aldous sketched out a proof that $\frac{3}{2}\log_2 n$ riffle shuffles mixed n cards. A more careful argument for riffle shuffling was presented by Aldous and Diaconis (1986). This introduced "strong stationary times," a powerful method of proof which has seen wide application. It is applied here in Section 4.

A definitive analysis of riffle shuffling was finally carried out in Bayer and Diaconis (1992) and Diaconis et al. (1995). They were able to derive simple closed-form expressions for all quantities involved and do exact computations for $n = 52$ (or 32 or 104 or ...). This results in the "seven shuffles theorem" explained below. A clear elementary account of these ideas is in Mann (1994, 1995) reprinted in Grinstead and Snell (1997). See Ethier (2010) for an informative textbook account.

The successful analysis of shuffling led to a host of developments, the techniques refined and extended. For example, it is natural to want not only the order of the cards, but also the "up-down pattern" of one-way backs to be randomized. Highlights include work of Bidigare et al. (1999) and Brown and Diaconis (1998) who gave a geometric interpretation of shuffling which had many extensions to which the same analysis applied. Lalley (1996, 1999) studied less random methods of riffle shuffling. Fulman (2000a,b, 2001) showed that interspersing cuts doesn't materially effect things and gave high level explanations for miraculous accidents connecting shuffling and Lie theory. The work is active and ongoing. Recent surveys are given by Diaconis (1996, 2003); Fulman (1998); Stark et al. (2002).

In recent work, Conger and Viswanath (2006); Diaconis et al. (1995) and Assaf et al. (2011) have studied the number of shuffles required to have selected features randomized (e.g., the original top card, or the values but not the suits). Here, fewer shuffles suffice. Conger and Howald (2010) shows that the way the cards are dealt out after shuffling affects

things. The mathematics of shuffling is closely connected to modern algebraic combinatorics through quasi-symmetric functions Stanley (2001). The descent theory underlying shuffling makes equivalent appearances in the basic task of carries when adding integers Diaconis and Fulman (2009a,b); Diaconis and Fulman (2011).

3. Analysis of one pass through a shelf shuffler.   This section gives a fairly complete analysis of a single pass through a shelf shuffler. Section 3.1 gives several equivalent descriptions of the shuffle. In Section 3.2, a closed-form formula for the chance of any permutation w is given. This in turn depends only on the number of "valleys" in w. The number of permutations with j valleys is easily calculated and so exact computations for any of the distances above are available. Section 3.3 uses the exact formulae to get asymptotic rates of convergence for $l(\infty)$ and separation distances. Section 3.4 gives the distribution of such permutations by cycle type. Section 3.5 gives the distribution of the "shape" of such a permutation under the Robinson–Schensted–Knuth map. Section 3.6 gives the distribution of the number of descents. We find it surprising that a real-world applied problem makes novel contact with elegant combinatorics. In Section 4, iterations of a shelf shuffler are shown to be equivalent to shelf shuffling with more shelves. Thus all of the formulae of this section apply.

3.1. *Alternative descriptions.*   Consider two basic shelf shufflers: for the first, a deck of n cards is sequentially distributed on one of m shelves. (Here, $n = 52$, $m = 10$, are possible choices.) Each time, the cards are taken from the *bottom* of the deck, a shelf is chosen at random from one to m, and the bottom card is placed on top of any previous cards on the shelf. At the end, the packets on the shelves are unloaded into a final deck of n. This may be done in order or at random; it turns out not to matter. Bayer and Diaconis (1992) called this an *inverse m-shuffle*.

The second shuffling scheme, that is the main object of the present study, is based on m shelves. At each stage that a card is placed on a shelf, the choice of whether to put it on the top or the bottom of the existing pile on that shelf is made at random (1/2 each side). This will be called a *shelf shuffle*. There are several equivalent descriptions of shelf shuffles:

Description 1 (Shelf shuffles)**.** A deck of cards is initially in order $1, 2, 3, \ldots, n$. Label the back of each card with n random numbers chosen at random between one and $2m$. Remove all cards labeled 1 and place them on top, keeping them in the same relative order. Then remove all cards labeled 2 and place them under the cards labeled 1, reversing their relative order. This continues with the cards labeled 3, labeled 4, and so on, reversing the order in each even labeled packet. If at any stage there are no cards with a given label, this empty packet still counts in the alternating pattern.

For example, a twelve-card deck with $2m = 4$,

Label	2	1	1	4	3	3	1	2	4	3	4	1
Card	1	2	3	4	5	6	7	8	9	10	11	12

is reordered as

$$2 \quad 3 \quad 7 \quad 12 \quad 8 \quad 1 \quad 5 \quad 6 \quad 10 \quad 11 \quad 9 \quad 4.$$

Description 2 (Inverse shelf shuffles)**.** Cut a deck of n cards into $2m$ piles according to a multinomial distribution; thus the number of cards cut off in pile i has the same distribution as the number of balls in the ith box if n balls are dropped randomly into $2m$ boxes. Reverse the order of the even-numbered packets. Finally, riffle shuffle the $2m$

packets together by the Gilbert–Shannon–Reeds (GSR) distribution Bayer and Diaconis (1992) dropping each card sequentially with probability proportional to packet size. This makes all possible interleavings equally likely.
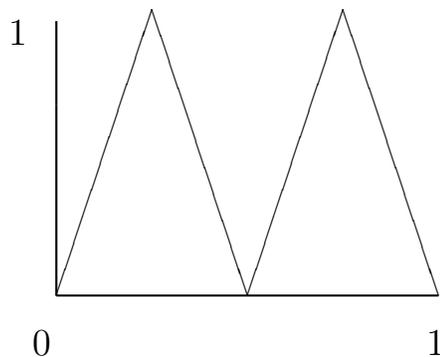

Fig 1: Two shelves in shelf shuffle.

Description 3 (Geometric description). Consider the function $f_m(x)$ from $[0,1]$ to $[0,1]$ which has "tents," each of slope $\pm 2m$ centered at $\frac{1}{2m}, \frac{3}{2m}, \frac{5}{2m}, \ldots, \frac{2m-1}{2m}$. Figure 1 illustrates an example with $m = 2$. Place n labeled points uniformly at random into the unit interval. Label them, from left to right, $x_1, x_2, \ldots, x_n$. Applying f_m gives $y_i = f_m(x_i)$. This gives the permutation

$$
\begin{array}{cccc}
1 & 2 & \cdots & n \\
\pi_1 & \pi_2 & \cdots & \pi_n
\end{array}
$$

with π_1 the relative position from the bottom of $y_1, \ldots, \pi_i$ the relative position from the bottom of y_i among the other y_j. This permutation has the distribution of an inverse shelf shuffle. It is important to note that the natural distances to uniformity (total variation, separation, l_∞) are the same for inverse shuffles and forward shuffles. In Section 4, this description is used to show that repeated shelf shuffling results in a shelf shuffle with more shelves.

3.2. *Formula for the chance of a permutation produced by a shelf shuffler.* To describe the main result, we call i a *valley* of the permutation $w \in S_n$ if $1 < i < n$ and $w(i-1) > w(i) < w(i+1)$. Thus $w = 5136724$ has two valleys. The number of valleys is classically used as a test of randomness for time series. See Warren and Seneta (1996) and their references. If $v(n, k)$ denotes the number of permutations on n symbols with k valleys, then Warren and Seneta (1996) $v(1, 0) = 1$, $v(n, k) = (2k + 2)v(n - 1, k) + (n - 2k)v(n - 1, k - 1)$. So $v(n, k)$ is easy to compute for numbers of practical interest. Asymptotics are in Warren and Seneta (1996) which also shows the close connections between valleys and descents.

Theorem 3.1. *The chance that a shelf shuffler with m shelves and n cards outputs a permutation w is*

$$
\frac{4^{v(w)+1}}{2(2m)^n} \sum_{a=0}^{m-1} \binom{n+m-a-1}{n} \binom{n-1-2v(w)}{a-v(w)}
$$

where $v(w)$ is the number of valleys of w. This can be seen to be the coefficient of t^m in

$$
\frac{1}{2(2m)^n} \frac{(1+t)^{n+1}}{(1-t)^{n+1}} \left(\frac{4t}{(1+t)^2} \right)^{v(w)+1} .
$$

Example. Suppose that $m = 1$. Then the theorem yields the uniform distribution on the 2^{n-1} permutations with no valleys; a permutation with one or more valleys occurs with probability 0. Permutations with no valleys are also sometimes called unimodal permutations. These arise in social choice theory through Coombs' "unfolding" hypothesis (Diaconis, 1988, Chap. 6). They also appear in dynamical systems and magic tricks see (Diaconis and Graham, 2011, Chap. 5).

Remark. By considering the cases $m \geq n$ and $n \geq m$ we see that, in the formula of Theorem 3.1, the range of summation can be taken up to $n - 1$ instead of $m - 1$. This will be useful later.

Theorem 3.1 makes it easy to compute the distance to stationarity for any of the metrics in Section 2.2. Indeed, the separation and $l(\infty)$ distance is attained at either permutations with a maximum number of valleys (when $n = 52$, this maximum is 25) or for permutations with 0 valleys. For the total variation distance, with $P_m(v)$ denoting the probability in Theorem 3.1,

$$\|P_m - U\|_{\mathrm{TV}} = \frac{1}{2} \sum_{a=0}^{\lfloor \frac{n-1}{2} \rfloor} v(n, a) \left| P_m(a) - \frac{1}{n!} \right|.$$

TABLE 1
Distances for various numbers of shelves m.

m	10	15	20	25	30	35	50	100	150	200	250	300
$\|P_m - U\|_{\mathrm{TV}}$	1	.943	.720	.544	.391	.299	.159	.041	.018	.010	.007	.005
$\mathrm{sep}(P_m)$	1	1	1	1	1	.996	.910	.431	.219	.130	.085	.060
$\|P_m - U\|_\infty$	∞	∞	∞	45118	3961	716	39	1.9	.615	.313	.192	.130

Table 1 gives these distances when $n = 52$ for various numbers of shelves m. Larger values of m are of interest because of the convolution results explained in Section 4. These numbers show that ten shelves are woefully insufficient. Indeed, 50 shelves are hardly sufficient.

To prove Theorem 3.1, we will relate it to the following $2m$-shuffle on the hyperoctahedral group B_n: cut the deck multinomially into $2m$ piles. Then flip over the odd numbered stacks, and riffle the piles together, by dropping one card at a time from one of the stacks (at each stage with probability proportional to stack size). When $m = 1$ this shuffle was studied in Bayer and Diaconis (1992), and for larger m it was studied in Fulman (2001).

It will be helpful to have a description of the inverse of this $2m$-shuffle. To each of the numbers $\{1, \ldots, n\}$ is assigned independently and uniformly at random one of $-1, 1, -2, 2, \ldots, -m, m$. Then a signed permutation is formed by starting with numbers mapped to -1 (in decreasing order and with negative signs), continuing with the numbers mapped to 1 (in increasing order and with positive signs), then continuing to the numbers mapped to -2 (in decreasing order and with negative signs), and so on. For example the assignment

$$\{1, 3, 8\} \mapsto -1, \{5\} \mapsto 1, \{2, 7\} \mapsto 2, \{6\} \mapsto -3, \{4\} \mapsto 3$$

leads to the signed permutation

$$-8 \ -3 \ -1 \ 5 \ 2 \ 7 \ -6 \ 4.$$

The proof of Theorem 3.1 depends on an interesting relation with shuffles for signed permutations (hyperoctahedral group). This is given next followed by the proof of Theorem 3.1.

Theorem 3.2 gives a formula for the probability for w after a hyperoctahedral $2m$-shuffle, when one forgets signs. Here $p(w)$ is the number of peaks of w, where i is said to be a peak of w if $1 < i < n$ and $w(i-1) < w(i) > w(i+1)$. Also $\Lambda(w)$ denotes the peak set of w and $D(w)$ denotes the descent set of w (i.e., the set of points i such that $w(i) > w(i+1)$). Finally, let $[n] = \{1, \ldots, n\}$.

Theorem 3.2. *The chance of a permutation w obtained by performing a $2m$ shuffle on the hyperoctahedral group and then forgetting signs is*

$$\frac{4^{p(w^{-1})+1}}{2(2m)^n} \sum_{a=0}^{m-1} \binom{n+m-a-1}{n}\binom{n-1-2p(w^{-1})}{a-p(w^{-1})}$$

where $p(w^{-1})$ is the number of peaks of w^{-1}.

Proof. Let $P'(m)$ denote the set of nonzero integers of absolute value at most m, totally ordered so that

$$-1 \prec 1 \prec -2 \prec 2 \prec \cdots \prec -m \prec m.$$

Then given a permutation $w = (w_1, \cdots, w_n)$, page 768 of Stembridge (1997) defines a quantity $\Delta(w)$. (Stembridge calls it $\Delta(w, \gamma)$, but throughout we always choose γ to be the identity map on $[n]$, and so suppress the symbol γ whenever he uses it). By definition, $\Delta(w)$ enumerates the number of maps $f : [n] \mapsto P'(m)$ such that

- $f(w_1) \preceq \cdots \preceq f(w_n)$
- $f(w_i) = f(w_{i+1}) > 0 \;\Rightarrow i \notin D(w)$
- $f(w_i) = f(w_{i+1}) < 0 \;\Rightarrow i \in D(w)$

We claim that the number of maps $f : [n] \mapsto P'(m)$ with the above three properties is equal to $(2m)^n$ multiplied by the chance that a hyperoctahedral 2m-shuffle results in the permutation w^{-1}. This is most clearly explained by example:

$$\begin{array}{ccccccccc}
w & = & 8 & 3 & 1 & 5 & 2 & 7 & 6 & 4 \\
f & = & -1 & -1 & -1 & 1 & 2 & 2 & -3 & 3
\end{array}$$

From the description of the inverse hyperoctahedral $2m$-shuffle stated before the proof, the assignment yields w. This proves the claim.

Let $\Lambda(w)$ denote the set of peaks of w. From Proposition 3.5 of Stembridge (1997),

$$\Delta(w) = 2^{p(w)+1} \sum_{E \subseteq [n-1]: \Lambda(w) \subseteq E \triangle (E+1)} L_E.$$

Here $E + 1$ means the elements of E with one added to each,

$$L_E = \sum_{\substack{1 \le i_1 \le \cdots \le i_n \le m \\ k \in E \Rightarrow i_k < i_{k+1}}} 1.$$

Here L_E is the number of ordered n-tuples $(i_1, i_2, \ldots, i_n)$ of integers between 1 and m which are non-decreasing, and strictly increasing at positions k in E. $\triangle$ denotes symmetric difference, i.e., $A \triangle B = (A - B) \cup (B - A)$. Now a simple combinatorial argument shows

that $L_E = \binom{n+m-|E|-1}{n}$. Indeed, L_E is equal to the number of integral $i_1, \ldots, i_n$ with $1 \leq i_1 \cdots \leq i_n \leq m - |E|$, which by a "stars and bars" argument is $\binom{n+m-|E|-1}{n}$. Thus

$$\Delta(w) = 2^{p(w)+1} \sum_{E \subseteq [n-1]: \Lambda(w) \subseteq E \triangle (E+1)} \binom{n+m-|E|-1}{n}.$$

Now let us count the number of E of size a appearing in this sum. For each $j \in \Lambda(w)$, exactly one of j or $j-1$ must belong to E, and the remaining $n - 1 - 2p(w)$ elements of $[n-1]$ can be independently and arbitrarily included in E. Thus the number of sets E of size a appearing in the sum is $2^{p(w)} \binom{n-1-2p(w)}{a-p(w)}$. Hence

$$\Delta(w) = \frac{4^{p(w)+1}}{2} \sum_{a=0}^{n-1} \binom{n+m-a-1}{n} \binom{n-1-2p(w)}{a-p(w)},$$

which completes the proof. □

Proof of Theorem 3.1. To deduce Theorem 3.1 from Theorem 3.2, we see that a shelf shuffle with m shelves is equivalent to taking w' to be the inverse of a permutation after a hyperoctahedral $2m$-shuffle (forgetting about signs), then taking a permutation w defined by $w(i) = n - w'(i) + 1$. Thus the shelf shuffle formula is obtained from the hyperoctahedral $2m$-shuffle formula by replacing peaks by valleys. □

Remarks.

- The paper Fulman (2001) gives an explicit formula for the chance of a signed permutation after a $2m$-shuffle on B_n in terms of cyclic descents. Namely it shows this probability to be

$$\frac{\binom{m+n-cd(w^{-1})}{n}}{(2m)^n}$$

where $cd(w)$ is the number of cyclic descents of w, defined as follows: Ordering the integers $1 < 2 < 3 < \cdots < \cdots < -3 < -2 < -1$,

  - w has a cyclic descent at position i for $1 \leq i \leq n - 1$ if $w(i) > w(i+1)$.
  - w has a cyclic descent at position n if $w(n) < 0$.
  - w has a cyclic descent at position 1 if $w(1) > 0$.

For example the permutation $3\ 1\ -2\ 4\ 5$ has two cyclic descents at position 1 and a cyclic descent at position 3, so $cd(w) = 3$.
This allows one to study aspects of shelf shufflers by lifting the problem to B_n, using cyclic descents (where calculations are often easier), then forgetting about signs. This idea was used in Fulman (2001) to study the cycle structure of unimodal permutations, and in Aguiar et al. (2004) to study peak algebras of types B and D.
The idea of lifting the problem to type B_n leads to a total variation upper bound. Indeed, from the proof of Theorem 3.1 the total variation distance after a shelf-shuffler with m shelves to uniform is equal to the total variation distance of a hyperoctahedral $2m$ shuffle to uniform, after one forgets about signs. Now from Bayer and Diaconis (1992) or Fulman (2001), the total variation distance of a hyperoctahedral $2m$ shuffle to uniform, when one doesn't forget about signs, is equal to the total variation distance of an ordinary m riffle shuffle to uniform on the symmetric group - a quantity

thoroughly studied in Bayer and Diaconis (1992). Thus the total variation distance after a shelf-shuffler with m shelves to uniform is at most the total variation distance of an ordinary m riffle shuffle to uniform on the symmetric group.

- The appearance of peaks in the study of shelf shufflers is interesting, as peak algebras have appeared in various parts of mathematics. Nyman (2003) proves that the peak algebra is a subalgebra of the symmetric group algebra, and connections with geometry of polytopes can be found in Aguiar et al. (2006) and Billera et al. (2003). There are also close connections with the theory of P-partitions Petersen (2005, 2007); Stembridge (1997).

The following corollary shows that for a shelf shuffler of n cards with m shelves, the chance of a permutation w with v valleys is monotone decreasing in v. Thus, the identity (or any other unimodal permutation) is most likely and an alternating permutation (down, up, down, up, ...) is least likely. From Theorem 3.1, the chance of a fixed permutation with v valleys is

$$(3.1) \qquad P(v) = \frac{4^{v+1}}{2(2m)^n} \sum_{a=0}^{n-1} \binom{n+m-1-a}{n}\binom{n-1-2v}{a-v}.$$

Corollary 3.3. *For $P(v)$ defined at (3.1), $P(v) \geq P(v+1)$, $0 \leq v \leq (n-1)/2$.*

Proof. Canceling common terms, and setting $a - v = j$ (so $a = j + v$) in (3.1), we have $\frac{2(2m)^n}{4^{v+1}}P(v) = \sum_{j=0}^{n-1-2v} f(j+v)\binom{n-1-2v}{j} = 2^{n-1-2v}E(f(S_{n-1-2v}+v))$ with $f(a) = \binom{n+m-1-a}{n}$ and S_{n-1-2v} distributed as binomial $(n-1-2v, \frac{1}{2})$. The proposed inequality is equivalent to

$$(3.2) \qquad E\left(f\left(S_{n-1-2v}+v\right)\right) \geq E\left(f\left(S_{n-1-2v-2}+v+1\right)\right).$$

To prove this, represent $S_{n-1-2v} = S_{n-1-2v-2} + Y_1 + Y_2$, with Y_i independent taking values in $\{0,1\}$, with probability $1/2$. Then (3.2) is equivalent to
$$(3.3)$$
$$\sum_j \left[\frac{1}{4}f(j+v) + \frac{1}{2}f(j+v+1) + \frac{1}{4}f(j+v+2) - f(j+v+1)\right] P\{S_{n-1-2v-2} = j\} \geq 0.$$

Thus if $\frac{1}{2}f(j+v) + \frac{1}{2}f(j+v+2) \geq f(j+v+1)$, e.g., $f(a)$ is convex, we are done. Writing out the expression $f(a) + f(a+2) \geq 2f(a+1)$ and canceling common terms, it must be shown that

$$(3.4) \quad (m+n-1-a)(m+n-2-a) + (m-1-a)(m-2-a) \geq 2(m+n-2-a)(m-1-a)$$

for all $0 \leq a \leq n-1$. Subtracting the right side from the left, the coefficients of a^2 and a cancel, leaving $n(n-1) \geq 0$. □

3.3. *Asymptotics for the $\|P-U\|_\infty$ and separation distances.* Recall the distances $\|P - U\|_\infty = \max_w \left|1 - \frac{P(w)}{U(w)}\right|$ and $\text{sep}(P) = \max_w \left(1 - \frac{P(w)}{U(w)}\right)$.

Theorem 3.4. *Consider the shelf shuffling measure P_m with n cards and m shelves. Suppose that $m = cn^{3/2}$. Then, as n tends to infinity with $0 < c < \infty$ fixed,*

$$\|P_m - U\|_\infty \sim e^{1/(12c^2)} - 1,$$
$$\text{sep}(P_m) \sim 1 - e^{-1/(24c^2)}.$$

Remark. We find it surprising that this many shelves are needed. For example, when $n = 52$, to make the distance less than $1/100$, $m \doteq 1,085$ shelves are required for $\|P_m - U\|_\infty$ and $m \doteq 764$ are required for $\mathrm{sep}(P_m)$. The order $m \doteq n^{\frac{3}{2}}$ in Theorem 3.4 can be understood as follows: Bayer and Diaconis (1992) show that it takes $k \doteq \frac{3}{2}\log_2 n$ riffle shuffles to mix n cards in total variation. Now k riffle shuffles correspond to $2^k = n^{\frac{3}{2}}$ shelves. Of course, the $\frac{3}{2}$ in riffle shuffling is the result of a careful computation.

Proof. Using Corollary 3.3, the distance is achieved at the identity permutation or a permutation with $\lfloor (n-1)/2 \rfloor$ valleys. For the identity, consider $n!P_m(\mathrm{id})$. Using Theorem 3.1,

$$(3.5) \qquad n!P_m(\mathrm{id}) = \frac{2(n!)}{(2m)^n} \sum_{a=0}^{n-1} \binom{m+n-a-1}{n}\binom{n-1}{a}.$$

To bound this sum, observe that $\binom{n-1}{a}/2^{n-1}$ is the binomial probability density. To keep the bookkeeping simple, assume throughout that n is odd. The argument for even n is similar.

For $a = \frac{n-1}{2} + j$, the local central limit theorem as in Feller (1968, Chap. VII.2), shows

$$(3.6) \qquad \frac{\binom{n-1}{\frac{n-1}{2}+j}}{2^{n-1}} \sim \frac{e^{-2j^2/n}}{\sqrt{\pi n/2}} \quad \text{for } j = o(n^{2/3}).$$

In the following, we show further that

$$(3.7) \qquad \frac{n!}{m^n}\binom{m+\frac{n-1}{2}-j}{n} \sim e^{-\frac{1}{24c^2}+\frac{j}{c\sqrt{n}}} \quad \text{uniformly for } j = o(n).$$

Combining (3.6), (3.7), gives a Riemann sum for the integral

$$\frac{e^{-\frac{1}{24c^2}}}{\sqrt{\pi/2}} \int_{-\infty}^{\infty} e^{-2x^2+x/c}\, dx = e^{1/(12c^2)},$$

the claimed result. This part of the argument follows Feller (1968, Chap. VII.2) and we suppress further details. To complete the argument the tails of the sum in (3.5) must be bounded.

We first prove (3.7). From the definitions

$$\frac{n!}{m^n}\binom{m-j+\frac{n-1}{2}}{n} = \prod_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} \left(1 - \frac{j}{m} + \frac{i}{m}\right)$$

using $\log(1-x) = -x - \frac{x^2}{2} + O(x^3)$,

(3.8)

$$\sum_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} \log\left(1 - \frac{j}{m} + \frac{i}{m}\right) = -\sum_i \left(-\frac{j}{m} + \frac{i}{m}\right) - \frac{1}{2}\sum_i \left(-\frac{j}{m} + \frac{i}{m}\right)^2 + nO\left(\left(\frac{n}{m}\right)^3\right)$$

$$= \frac{nj}{m} - \frac{1}{2}\left(\frac{nj^2}{m^2} + \frac{1}{12}\frac{n(n^2-1)}{m^2}\right) + O\left(\frac{1}{\sqrt{n}}\right)$$

$$= \frac{j}{c\sqrt{n}} - \frac{j^2}{2c^2n^2} - \frac{1}{24c^2} + O\left(\frac{1}{\sqrt{n}}\right).$$

The error term in (3.8) is uniform in j. For $j = o(n)$, $j^2/n^2 = o(1)$ and (3.7) follows.

To bound the tails of the sum, first observe that (3.8) implies that $\frac{n!}{m^n}\binom{m-j+\frac{n-1}{2}}{n} = e^{O(\sqrt{n})}$ for all j. From Bernstein's inequality, if $X_i = \pm 1$ with probability $1/2$, $P(|X_1 + \cdots + X_{n-1}| > a) \leq 2e^{-a^2/(n-1)}$. Using this, the sum over $|j| \geq An^{3/4}$ is negligible for A sufficiently large.

The Gaussian approximation to the binomial works for $j \ll n^{2/3}$. To bound the sum for $|j|$ between $n^{2/3}$ and $n^{3/4}$, observe from (3.8) that in this range, $\frac{n!}{m^n}\binom{m-j+\frac{n-1}{2}}{n} = O\left(e^{n^{1/4}}\right)$. Then, Feller (1968, p. 195) shows

$$\frac{\binom{n-1}{\frac{n-1}{2}+j}}{2^{n-1}} \sim \frac{1}{\sqrt{\pi n/2}} e^{-\frac{1}{2}(j)^2/(n/4)-f\left(j/\sqrt{n/4}\right)}$$

with $f(x) = \sum_{a=3}^{\infty} \frac{\left(\frac{1}{2}\right)^{a-1}+\left(-\frac{1}{2}\right)^{a-1}}{a(a-1)}\left(\frac{1}{\sqrt{n/4}}\right)^{a-2} x^a = c_1\frac{x^4}{n} + c_2\frac{x^6}{n^2} + \ldots$ for explicit constants $c_1, c_2, \ldots$. For $\theta_1 n^{2/3} \leq |j| \leq \theta_2 n^{3/4}$, the sum under study is dominated by $A\sum_{j \geq n^{2/3}} e^{-Bj^{1/6}}$ which tends to zero.

The separation distance is achieved at permutations with $\frac{n-1}{2}$ valleys (recall we are assuming that n is odd). From (3.1),

$$1 - n!P_m\left(\frac{n-1}{2}\right) = 1 - \frac{n!}{m^n}\binom{m+\frac{n-1}{2}}{n}.$$

The result now follows from (3.7) with $j = 0$. □

Remark. A similar argument allows asymptotic evaluation of total variation. We have not carried out the details.

3.4. *Distribution of cycle type.* The number of fixed points and the number of cycles are classic descriptive statistics of a permutation. More generally, the number of i-cycles for $1 \leq i \leq n$ has been intensively studied (Diaconis et al., 1995; Shepp and Lloyd, 1966). This section investigates the distribution of cycle type of a permutation w produced from a shelf shuffler with m shelves and n cards. Similar results for ordinary riffle shuffles appeared in Diaconis et al. (1995), and closely related results in the type B case (not in the language of shelf-shuffling) appear in Fulman (2001, 2002). Recall also that in the case of one shelf, the shelf shuffler generates one of the 2^{n-1} unimodal permutations uniformly at random. The cycle structure of unimodal permutations has been studied in several papers in the literature: see Fulman (2001, 2002); Thibon (2001) for algebraic/combinatorial approaches and Gannon (2001); Rogers (1981) for approaches using dynamical systems.

For what follows, we define

$$f_{i,m} = \frac{1}{2i}\sum_{\substack{d|i \\ d \text{ odd}}} \mu(d)(2m)^{i/d}$$

where μ is the Möebius function of elementary number theory: $\mu(d) = (-1)^k$ if d is a square free number with k prime factors, $\mu(1) = 1$ and $\mu(d) = 0$ otherwise.

Theorem 3.5. *Let $P_m(w)$ denote the probability that a shelf shuffler with m shelves produces a permutation w. Let $N_i(w)$ denote the number of i-cycles of a permutation w in S_n. Then*

(3.9) $$1 + \sum_{n \geq 1} u^n \sum_{w \in S_n} P_m(w) \prod_{i \geq 1} x_i^{N_i(w)} = \prod_{i \geq 1}\left(\frac{1 + x_i(u/2m)^i}{1 - x_i(u/2m)^i}\right)^{f_{i,m}}.$$

Proof. By the proof of Theorem 3.1, a permutation produced by a shelf shuffler with m shelves is equivalent to forgetting signs after the inverse of a type B riffle shuffle with $2m$ piles, then conjugating by the longest element $n, n-1, \ldots, 1$. Since a permutation and its inverse have the same cycle type and conjugation leaves cycle type invariant, the result follows from either (Fulman, 2001, Thm. 7) or (Fulman, 2002, Thm. 9), both of which derived the generating function for cycle type after type B shuffles. $\square$

Theorem 3.5 leads to several corollaries. We say that a random variable X is binomial (n,p) if $\mathbb{P}(X = j) = \binom{n}{j}p^j(1-p)^{n-j}$, $0 \le j \le n$, and that X is negative binomial with parameters (f, p) if $\mathbb{P}(X = j) = \binom{f+j-1}{j}p^j(1-p)^f$, $0 \le j < \infty$. As usual, the products in the generating function on the right of (3.9) correspond to the convolution of the corresponding measures:

Corollary 3.6. *Let $N_i(w)$ be the number of i-cycles of a permutation w.*

1. *Fix u such that $0 < u < 1$. Then choose a random number N of cards so that $\mathbb{P}(N = n) = (1 - u)u^n$. Let w be produced by a shelf shuffler with m shelves and N cards. Then any finite number of the random variables $\{N_i\}$ are independent, and N_i is distributed as the convolution of a binomial $\left(f_{i,m}, \frac{(u/2m)^i}{1+(u/2m)^i}\right)$ and a negative binomial with parameters $(f_{i,m}, (u/2m)^i)$.*
2. *Let w be produced by a shelf shuffler with m shelves and n cards. Then in the $n \to \infty$ limit, any finite number of the random variables $\{N_i\}$ are independent. The N_i are distributed as the convolution of a binomial $\left(f_{i,m}, \frac{1}{(2m)^i+1}\right)$ and a negative binomial with parameters $(f_{i,m}, (1/2m)^i)$.*

Proof. Setting all $x_i = 1$ in equation (3.9) yields the equation

$$(3.10) \qquad (1-u)^{-1} = \prod_{i \ge 1} \left(\frac{1 + (u/2m)^i}{1 - (u/2m)^i}\right)^{f_{i,m}}.$$

Taking reciprocals of equation (3.10) and multiplying by equation (3.9) gives the equality

$$(3.11) \quad (1-u) + \sum_{n \ge 1}(1-u)u^n \sum_{w \in S_n} P_m(w) \prod_{i \ge 1} x_i^{n_i(w)}$$
$$= \prod_{i \ge 1}\left(\frac{1 + x_i(u/2m)^i}{1 + (u/2m)^i}\right)^{f_{i,m}} \cdot \left(\frac{1 - (u/2m)^i}{1 - x_i(u/2m)^i}\right)^{f_{i,m}}.$$

This proves part 1 of the theorem, the first term on the right corresponding to the convolution of binomials, and the second term to the convolution of negative binomials.

The second part follows from the claim that if a generating function $f(u)$ has a Taylor series which converges at $u = 1$, then the $n \to \infty$ limit of the coefficient of u^n in $f(u)/(1 - u)$ is $f(1)$. Indeed, write the Taylor expansion $f(u) = \sum_{n=0}^\infty a_n u^n$ and observe that the coefficient of u^n in $f(u)/(1-u)$ is $\sum_{i=0}^n a_i$. Now apply the claim to equation (3.11) with all but finitely many x_i equal to 1. $\square$

Remark. For example, when $i = 1$, $f_{i,m} = m$; the number of fixed points are distributed as a sum of binomial $\left(m, \frac{1}{2m+1}\right)$ and negative binomial$\left(m, \frac{1}{2m}\right)$. Each of these converges to Poisson(1/2) and so the number of fixed points is approximately Poisson(1). A similar

analysis holds for the other cycle counts. Corollary 3.6 could also be proved by the method of moments, along the lines of the arguments of Diaconis et al. (1995) for the case of ordinary riffle shuffles.

For the next result, recall that the limiting distribution of the large cycles of a uniformly chosen permutation in S_n has been determined by Goncharov (1944); Gontcharoff (1942), Shepp and Lloyd (1966), Vershik and Shmidt (1977, 1978), and others. For instance the average length of the longest cycle L_1 is approximately $.63n$ and L_1/n has a known limiting distribution. The next result shows that even with a fixed number of shelves, the distribution of the large cycles approaches that of a uniform random permutation, as long as the number of cards is growing. We omit the proof, which goes exactly along the lines of the corresponding result for riffle shuffles in Diaconis et al. (1995).

Corollary 3.7. *Fix k and let $L_1(w), L_2(w), \ldots, L_k(w)$ be the lengths of the k longest cycles of $w \in S_n$ produced by a shelf shuffler with m shelves. Then for m fixed, or growing with n, as $n \to \infty$,*

$$|P_m\{L_1/n \le t_1, \ldots, L_k/n \le t_n\} - P_\infty\{L_1/n \le t_1, \ldots, L_k/n \le t_n\}| \to 0$$

uniformly in $t_1, t_2, \ldots, t_k$.

As a final corollary, we note that Theorems 3.1 and 3.5 give the following generating function for the joint distribution of permutations by valleys and cycle type. Note that this gives the joint generating function for the distribution of permutations by peaks and cycle type, since conjugating by the permutation $n, n-1, \ldots, 1$ preserves the cycle type and swaps valleys and peaks.

Corollary 3.8. *Let $v(w)$ denote the number of valleys of a permutation w. Then*

$$\frac{t}{1-t} + \sum_{n \ge 1} u^n \sum_{w \in S_n} \frac{1}{2} \frac{(1+t)^{n+1}}{(1-t)^{n+1}} \left(\frac{4t}{(1+t)^2} \right)^{v(w)+1} \prod_{i \ge 1} x_i^{N_i(w)} = \sum_{m \ge 1} t^m \prod_{i \ge 1} (\frac{1 + x_i u^i}{1 - x_i u^i})^{f_{i,m}}.$$

The same result holds with $v(w)$ replaced by $p(w)$, the number of peaks of w.

Remark. There is a large literature on the joint distribution of permutations by cycles and descents (Blessenohl et al., 2005; Diaconis et al., 1995; Fulman, 2000b; Gessel and Reutenauer, 1993; Poirier, 1998; Reiner, 1993) and by cycles and cyclic descents Fulman (2000a, 2001, 2002), but Corollary 3.8 seems to be the first result on the joint distribution by cycles and peaks.

3.5. *Distribution of RSK shape.* In this section we obtain the distribution of the Robinson–Schensted–Knuth (RSK) shape of a permutation w produced from a shelf shuffler with m shelves and n cards. For background on the RSK algorithm, see Stanley (1999). The RSK bijection associates to a permutation $w \in S_n$ a pair of standard Young tableaux $(P(w), Q(w))$ of the same shape and size n. $Q(w)$ is called the recording tableau of w.

To state our main result, we use a symmetric function S_λ studied in Stembridge (1997) (a special case of the extended Schur functions in Kerov and Vershik (1986)). One definition of the S_λ is as the determinant

$$S_\lambda(y) = \det(q_{\lambda_i - i + j})$$

where $q_{-r} = 0$ for $r > 0$ and for $r \geq 0$, q_r is defined by setting

$$\sum_{n \geq 0} q_n t^n = \prod_{i \geq 1} \frac{1 + y_i t}{1 - y_i t}.$$

We also let f_λ denote the number of standard Young tableaux of shape λ.

Theorem 3.9. *The probability that a shelf shuffler with m shelves and n cards produces a permutation with recording tableau T is equal to*

$$\frac{1}{2^n} S_\lambda \left(\frac{1}{m}, \ldots, \frac{1}{m} \right)$$

for any T of shape λ, where S_λ has m variables. Thus the probability that w has RSK shape λ is equal to

$$\frac{f_\lambda}{2^n} S_\lambda \left(\frac{1}{m}, \ldots, \frac{1}{m} \right).$$

Proof. By the proof of Theorem 3.1, a permutation produced by a shelf shuffler with m shelves is equivalent to forgetting signs after the inverse of a type B $2m$-shuffle, and then conjugating by the permutation n $n-1 \ldots 1$. Since a permutation and its inverse have the same RSK shape (Stanley, 1999, Sect. 7.13), and conjugation by $n, n-1, \ldots, 1$ leaves the RSK shape unchanged (Stanley, 1999, Thm. A1.2.10), the result follows from Theorem 8 of Fulman (2002), which studied RSK shape after type B riffle shuffles. $\qquad\square$

3.6. *Distribution of descents.* A permutation w is said to have a descent at position i ($1 \leq i \leq n - 1$) if $w(i) > w(i + 1)$. We let $d(w)$ denote the total number of descents of w. For example the permutation 3 1 5 4 2 has $d(w) = 3$ and descent set $1, 3, 4$. The purpose of this section is to derive a generating function for the number of descents in a permutation w produced by a shelf shuffler with m shelves and n cards. More precisely, we prove the following result.

Theorem 3.10. *Let $P_m(w)$ denote the probability that a shelf shuffler with m shelves and n cards produces a permutation w. Letting $[u^n] f(u)$ denote the coefficient of u^n in a power series $f(u)$, one has that*

$$(3.12) \qquad \sum_{w \in S_n} P_m(w) t^{d(w)+1} = \frac{(1-t)^{n+1}}{2^n} \sum_{k \geq 1} t^k [u^n] \frac{(1 + u/m)^{km}}{(1 - u/m)^{km}}.$$

The proof uses the result about RSK shape mentioned in Section 3.5, and symmetric function theory; background on these topics can be found in the texts by Stanley (1999) and Macdonald (1995) respectively.

Proof. Let w be a permutation produced by a shelf shuffler with m shelves and n cards. The RSK correspondence associates to w a pair of standard Young tableaux $(P(w), Q(w))$ of the same shape. Moreover, there is a notion of descent set for standard Young tableaux, and by Lemma 7.23.1 of Stanley (1999), the descent set of w is equal to the descent set of $Q(w)$. Let $f_\lambda(r)$ denote the number of standard Young tableaux of shape λ with r descents. Then Theorem 3.9 implies that

$$\mathbb{P}(d(w) = r) = \sum_{|\lambda|=n} \frac{f_\lambda(r)}{2^n} S_\lambda \left(\frac{1}{m}, \ldots, \frac{1}{m} \right).$$

By equation 7.96 of Stanley (1999), one has that

$$\sum_{r \geq 0} f_\lambda(r) t^{r+1} = (1-t)^{n+1} \sum_{k \geq 1} s_\lambda(1, \ldots, 1) t^k$$

where in the kth summand, $s_\lambda(1, \ldots, 1)$ denotes the Schur function with k variables specialized to 1. Thus

$$\sum_{r \geq 0} \mathbb{P}(d(w) = r) \cdot t^{r+1} = \sum_{r \geq 0} \sum_{|\lambda|=n} \frac{f_\lambda(r)}{2^n} S_\lambda \left(\frac{1}{m}, \ldots, \frac{1}{m} \right) \cdot t^{r+1}$$

$$= \frac{(1-t)^{n+1}}{2^n} \sum_{k \geq 1} t^k \sum_{|\lambda|=n} S_\lambda \left(\frac{1}{m}, \ldots, \frac{1}{m} \right) s_\lambda(1, \ldots, 1)$$

$$= \frac{(1-t)^{n+1}}{2^n} \sum_{k \geq 1} t^k [u^n] \sum_{n \geq 0} \sum_{|\lambda|=n} S_\lambda \left(\frac{1}{m}, \ldots, \frac{1}{m} \right) s_\lambda(1, \ldots, 1) \cdot u^n.$$

From Appendix A.4 of Stembridge (1997), if λ ranges over all partitions of all natural numbers, then

$$\sum_\lambda s_\lambda(x) S_\lambda(y) = \prod_{i,j \geq 1} \frac{1 + x_i y_j}{1 - x_i y_j}.$$

Setting $x_1 = \cdots = x_k = u$ and $y_1 = \cdots = y_m = \frac{1}{m}$ completes the proof of the theorem. $\square$

For what follows we let $A_n(t) = \sum_{w \in S_n} t^{d(w)+1}$ be the generating function of elements in S_n by descents. This is known as the Eulerian polynomial and from page 245 of Comtet (1974), one has that

$$(3.13) \qquad\qquad A_n(t) = (1-t)^{n+1} \sum_{k \geq 1} t^k k^n.$$

This also follows by letting $m \to \infty$ in equation (3.12).

The following corollary derives the mean and variance of the number of descents of a permutation produced by a shelf shuffler.

Corollary 3.11. *Let w be a permutation produced by a shelf shuffler with m shelves and $n \geq 2$ cards.*

1. *The expected value of $d(w)$ is $\frac{n-1}{2}$.*
2. *The variance of $d(w)$ is $\frac{n+1}{12} + \frac{n-2}{6m^2}$.*

Proof. The first step is to expand $[u^n] \frac{(1+u/m)^{km}}{(1-u/m)^{km}}$ as a series in k. One calculates that

$$[u^n] \frac{(1+u/m)^{km}}{(1-u/m)^{km}} = \frac{1}{m^n} \sum_{a \geq 0} \binom{km}{a} \binom{km+n-a-1}{n-a}$$

$$= \frac{1}{m^n} \sum_{a \geq 0} \left[\frac{(km) \ldots (km-a+1)}{a!} \right] \left[\frac{(km+n-a-1) \ldots (km)}{(n-a)!} \right]$$

$$= \frac{1}{n!} \left[2^n k^n + \frac{2^n n(n-1)(n-2)}{12m^2} k^{n-2} + \ldots \right]$$

where the ... in the last equation denote terms of lower order in k. Thus Theorem 3.10 gives

$$\sum_w P_m(w)t^{d(w)+1} = \left[\frac{(1-t)^{n+1}}{n!}\sum_{k\geq 1}t^k k^n\right] + \frac{n-2}{12m^2}(1-t)^2\left[\frac{(1-t)^{n-1}}{(n-2)!}\sum_{k\geq 1}t^k k^{n-2}\right] + (1-t)^3 C(t)$$

where $C(t)$ is a polynomial in t. By equation (3.13), it follows that

$$\sum_w P_m(w)t^{d(w)+1} = \frac{A_n(t)}{n!} + (1-t)^2\frac{n-2}{12m^2}\frac{A_{n-2}(t)}{(n-2)!} + (1-t)^3 C(t).$$

Since the number of descents of a random permutation has mean $(n-1)/2$ and variance $(n+1)/12$ for $n \geq 2$, it follows that $\frac{A'_n(1)}{n!} = \frac{(n+1)}{2}$ and also that $\frac{A''_n(1)}{n!} = (3n^2 + n - 2)/12$. Thus

$$\sum_w P_m(w)d(w) = \frac{n-1}{2}$$

and

$$\sum_w P_m(w)d(w)[d(w)+1] = \frac{3n^2+n-2}{12} + \frac{n-2}{6m^2}$$

and the result follows. □

Remarks.

- Part 1 of Corollary 3.11 can be proved without generating functions simply by noting that by the way the shelf shuffler works, w and its reversal are equally likely to be produced.
- Theorem 3.10 has an analog for ordinary riffle shuffles which is useful in the study of carries in addition. See Diaconis and Fulman (2009a) for details.

4. Iterated shuffling. This section shows how to analyze repeated shuffles. Section 4.1 shows how to combine shuffles. Section 4.2 gives a clean bound for the separation distance.

4.1. *Combining shuffles.* To describe what happens to various combinations of shuffles, we need the notion of a signed $m-$shuffle. This has the following geometric description: divide the unit interval into sub-intervals of length $\frac{1}{m}$; each sub-interval contains the graph of a straight line of slope $\pm m$. The left-to-right pattern of signs $\pm s$ is indicated by a vector x of length m. Thus if $m = 4$ and $x = ++++$, an $x-$shuffle is generated as shown on the left side of Figure 2. If $m = 4$ and $x = +--+$, the graph becomes that of the right side of Figure 2. Call this function f_x.

The shuffle proceeds as in the figure with n points dropped at random into the unit interval, labeled left to right, $y_1, y_2, \ldots, y_n$ and then permuted by f_x. In each case there is a simple forward description: the deck is cut into m piles by a multinomial distribution and piles corresponding to negative coordinates are reversed. Finally, all packets are shuffled together by the GSR procedure in which one drops each card sequentially with probability proportional to packet size. Call the associated measure on permutations P_x.
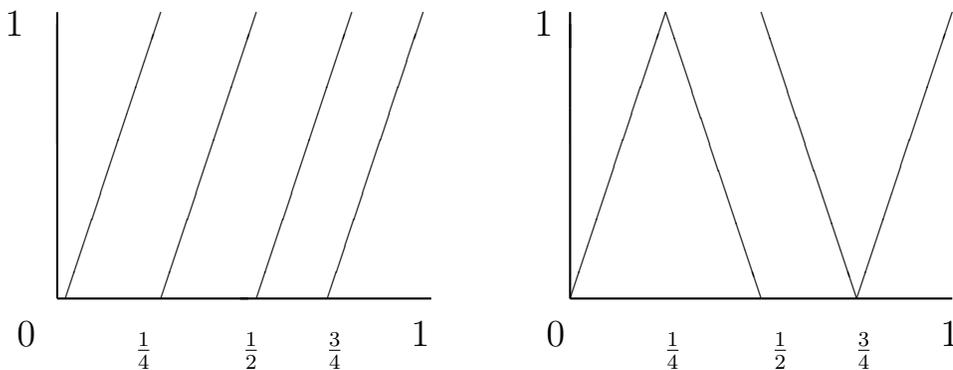
Fig 2: *Left*: $m = 4$, $x = + + ++$; *right*: $m = 4$, $x = + - -+$.

Remark. Thus, ordinary riffle shuffles are $++$ shuffles. The shelf shuffle with 10 shelves is an inverse $+ - + - \cdots + -$ (length 20) shuffle in this notation.

The following theorem reduces repeated shuffles to a single shuffle. To state it, one piece of notation is needed. Let $x = (x_1, x_2, \ldots, x_a)$ and $y = (y_1, y_2, \ldots, y_b)$ be two sequences of $\pm$ signs. Define a sequence of length ab as $x * y = y^{x_1}, y^{x_2}, \ldots, y^{x_a}$ with $(y_1, \ldots, y_b)^1 = (y_1, \ldots, y_b)$ and $(y_1, \ldots, y_b)^{-1} = (-y_b, -y_{b-1}, \ldots, -y_1)$. This is an associative product on strings; it is not commutative. Let P_x be the measure induced on S_n (forward shuffles).

Example.

$$(+ + +) * (++) = + + + + + ++$$
$$(+-) * (+-) = + - +-$$
$$(+-) * (+ + -+) = + + - + - + --$$

Theorem 4.1. *If x and y are ± 1 sequences of length a and b, respectively, then*

$$P_x * P_y = P_{x*y}.$$

Proof. In outline, this follows most easily from the geometric description underlying Figure 1 and Figure 2. If a uniformly chosen point in $[0, 1]$ is expressed base a, the "digits" are uniform and independently distributed in $\{0, 1, \ldots, a - 1\}$. Because of this, iterating the maps on the same uniform points gives the convolution. The iterated maps have the claimed pattern of slopes by a simple geometric argument.

In more detail, consider a $\pm$ string x of length m. The function f_x sends $[0, 1]$ to itself by mapping $\eta \in [\frac{i-1}{m}, \frac{i}{m}]$ to $x_i m \eta \pmod 1$. If x_i is positive, all points in $[\frac{i-1}{m}, \frac{i}{m}]$ are sent to $[0, 1]$ in an order preserving way. If x_i is negative, the order is reversed. In either case, the map f_x is m to 1 and measure preserving on $[0, 1]$ (that is, f_x^{-1} preserves Lebesgue measure). Now consider $f_y \circ f_x$ (apply f_x first, then f_y), where x is of length m and y is of length n. The composition sends all elements $\eta \in [\frac{i-1}{mn}, \frac{i}{mn}]$ to sgn(i) $mn\eta \pmod 1$; where sgn(i) is ± 1. A simple argument shows sgn(i) is given by the $x * y$ rule.

There is a subtle point : the convolution $P_x * P_y$ (first apply P_x then independently P_y) involves independent shuffles while $f_y \circ f_x$ and f_{x*y} are applied to a single random uniform set of points. The induced measures are the same because the digits of $\eta \in [0, 1]$ base m are independent and uniformly distributed in $\{0, 1, \ldots, m\}$ and f_x preserves measure. It

18

follows that the image of independent uniform points in $[0, 1]$ under f_x are independent and uniform even conditional on the induced permutation which is determined by how many points fall in each $[\frac{i-1}{m}, \frac{i}{m}]$. □

Corollary 4.2. *The convolution of k $+-$ shuffles is a $+-+-\cdots+-$ (2^k terms) shuffle. Further, the convolution of a shelf shuffler with m_1 and then m_2 shelves is the same as a shelf shuffler with $2m_1 m_2$ shelves.*

4.2. *Bounds for separation distance.* The following theorem gives a bound for separation (and so for total variation) for a general P_x shuffle on S_n.

Theorem 4.3. *For any ± 1 sequence x of length a, with P_x the associated measure on S_n, and $\text{sep}(P_x)$ from (2.1),*

$$(4.1) \qquad \text{sep}(P_x) \leq 1 - \prod_{i=1}^{n-1} \left(1 - \frac{i}{a} \right).$$

Proof. It is easiest to argue using shuffles as in Description 1. There, the backs of cards are labeled, independently and uniformly, with symbols $1, 2, \ldots, a$. For the inverse shuffle, all cards labeled 1 are removed, keeping them in their same relative order, and placed on top followed by the cards labeled 2 (placed under the 1s) and so on, with the following proviso: if the ith coordinate of x is -1, the cards labeled i have their order reversed; so if they are 1, 5, 17 from top down, they are placed in order 17, 5, 1. All of this results in a single permutation drawn from P_x. Repeated shuffles are modeled by labeling each card with a vector of symbols. The kth shuffle is determined by the kth coordinate of this vector. The first time t that the first k coordinates of those n vectors are all distinct forms a strong stationary time. See Aldous and Diaconis (1986) or Fulman (1998) for further details. The usual bound for separation yields

$$\text{sep}(P_x) \leq P\{\text{all } n \text{ labels are distinct}\}.$$

The bound (4.1) now follows from the classical birthday problem. □

Remarks.

- For a large with respect to n, the right side is well-approximated by $1 - e^{-\frac{n(n-1)}{2a}}$. This is small when $n^2 \ll a$.
- The theorem gives a clean upper bound on the distance to uniformity. For example, when $n = 52$, after 8 ordinary riffle shuffles (so $x = ++\cdots++$, length 256), the bound (4.1) is $\text{sep}(P_x) \leq 0.997$, in agreement with Table 1 of Assaf et al. (2011). For the actual shelf shuffle with $x = +-+-\cdots+-$ (length 20), the bound gives $\text{sep}(P_x) = 1$ but $\text{sep}(P_x * P_x) \leq 0.969$ and $\text{sep}(P_x * P_x * P_x) \leq 0.153$.
- The bound in Theorem 4.3 is simple and general. However, it is not sharp for the original shelf shuffler. The results of Section 3.3 show that $m = cn^{3/2}$ shelves suffice to make $\text{sep}(P_m)$ small when c is large. Theorem 4.3 shows that $m = cn^2$ steps suffice.
- The upper bound in (4.1) is achieved. If the length a sequence consists only of $+$ signs, we have the ordinary "riffle shuffle". Then the formula in Bayer and Diaconis (1992) for the chance of a permutation after an a shuffle implies that the separation distance is attained for the permutation $n, n-1,, \ldots, 1$, and is equal to the right hand side of (4.1).

5. Practical tests and conclusions. The engineers and executives who consulted us found it hard to understand the total variation distance. They asked for more down-to-earth notions of discrepancy. This section reports some ad hoc tests which convinced them that the machine had to be used differently. Section 5.1 describes the number of cards guessed correctly. Section 5.2 briefly describes three other tests. Section 5.3 describes conclusions and recommendations.

5.1. *Card guessing with feedback.* Suppose, after a shuffle, cards are dealt face-up, one at a time, onto the table. Before each card is shown, a guess is made at the value of the card. Let X_i, $1 \le i \le n$, be one or zero as the ith guess is correct and $T_n = X_1 + \cdots + X_n$ the total number of correct guesses. If the cards were perfectly mixed, the chance that $X_1 = 1$ is $1/n$, the chance that $X_2 = 1$ is $1/(n-1)$, ..., that $X_i = 1$ is $1/(n-i+1)$. Further, the X_i are independent. Thus elementary arguments give the following.

Proposition 5.1. *Under the uniform distribution, the number of cards guessed correctly T_n satisfies*

- $E(T_n) = \frac{1}{n} + \frac{1}{n-1} + \cdots + 1 \sim \log n + \gamma + O\left(\frac{1}{n}\right)$ *with $\gamma \doteq 0.577$ Euler's constant.*
- $\text{var}(T_n) = \frac{1}{n}\left(1 - \frac{1}{n}\right) + \frac{1}{n-1}\left(1 - \frac{1}{n-1}\right) + \cdots + \frac{1}{2}\left(1 - \frac{1}{2}\right) \sim \log n + \gamma - \frac{\pi^2}{6} + O\left(\frac{1}{n}\right).$
- *Normalized by its mean and variance, T_n has an approximate normal distribution.*

When $n = 52$, T_n has mean approximately 4.5, standard deviation approximately $\sqrt{2.9}$, and the number of correct guesses is between 2.7 and 6.3, 70% of the time.

Based on the theory developed in Section 3 we constructed a guessing strategy — conjectured to be optimal — for use after a shelf shuffle.

Strategy

- To begin, guess card 1.
- If guess is correct, remove card 1 from the list of available cards. Then guess card 2, card 3, ...
- If guess is incorrect and card i is shown, remove card i from the list of available cards and guess card $i + 1$, card $i + 2$, ...
- Continue until a descent is observed (order reversal with the value of the current card smaller than the value of the previously seen card). Then change the guessing strategy to guess the next-smallest available card.
- Continue until an ascent is observed, then guess the next-largest available card, and so on.

TABLE 2
Mean and variance for $n = 52$ after a shelf shuffle with m shelves under the conjectured optimal strategy.

m	1	2	4	10	20	64
mean	39	27	17.6	9.3	6.2	4.7
variance	3.2	5.6	6.0	4.7	3.8	3.1

A Monte Carlo experiment was run to determine the distribution of T_n for $n = 52$ with various values of m (10,000 runs for each value). Table 2 shows the mean and variance for various numbers of shelves. Thus for the actual shuffler, $m = 10$ gives about 9.3 correct guesses versus 4.5 for a well-shuffled deck. A closely related study of optimal strategy for the GSR measure (without feedback) is carried out by Ciucu (1998).

5.2. *Three other tests.* For the shelf shuffler with m shelves, an easy argument shows that the chance that the original top card is still on top is at least $1/2m$ instead of $1/n$. When $n = 52$, this is $1/20$ versus $1/52$. The chance that card 2 is on top is approximately $\frac{1}{2m}\left(1 - \frac{1}{2m}\right)$ while the chance that card 2 is second from the top is roughly $\frac{1}{(2m)^2}$. The same probabilities hold for the bottom cards. While not as striking as the guessing test of Section 5.1, this still suggests that the machine is "off."

Our second test supposed that the deck was originally arranged with all the red cards on top and all the black cards at the bottom. The test statistic is the number of changes of color going through the shuffled deck. Under uniformity, simulations show this has mean 26 and standard deviation 3.6. With a 10-shelf machine, simulations showed 17 ± 1.83, a noticeable deviation. The third test is based on the spacings between cards originally near
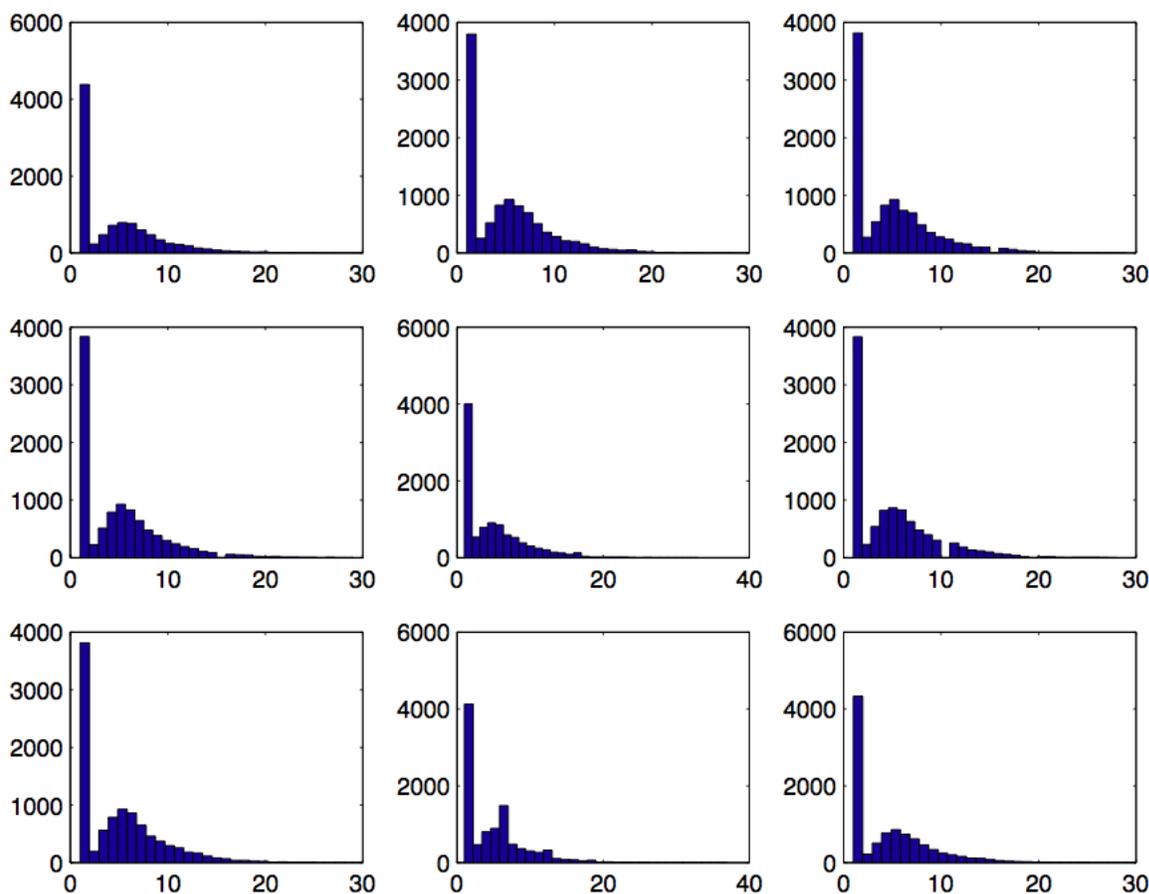

Fig 3: 9 spacings from a 10-shelf shuffle; j varies from top left to bottom right, $1 \le j \le 9$.

the top of the deck. Let w_j denote the position of the card originally at position j from the top. Let $D_j = |w_j - w_{j+1}|$. Figure 3 shows a histogram of D_j for $1 \le j \le 9$, from a simulation with $n = 52$ based on a 10-shelf shuffler. Figure 4 shows histograms for the same statistics for a well-shuffled deck; there are striking discrepancies.

5.3. *Conclusions and recommendations.* The study above shows that a single iteration of a 10-shelf shuffler is not sufficiently random. The president of the company responded "We are not pleased with your conclusions, but we believe them and that's what we hired
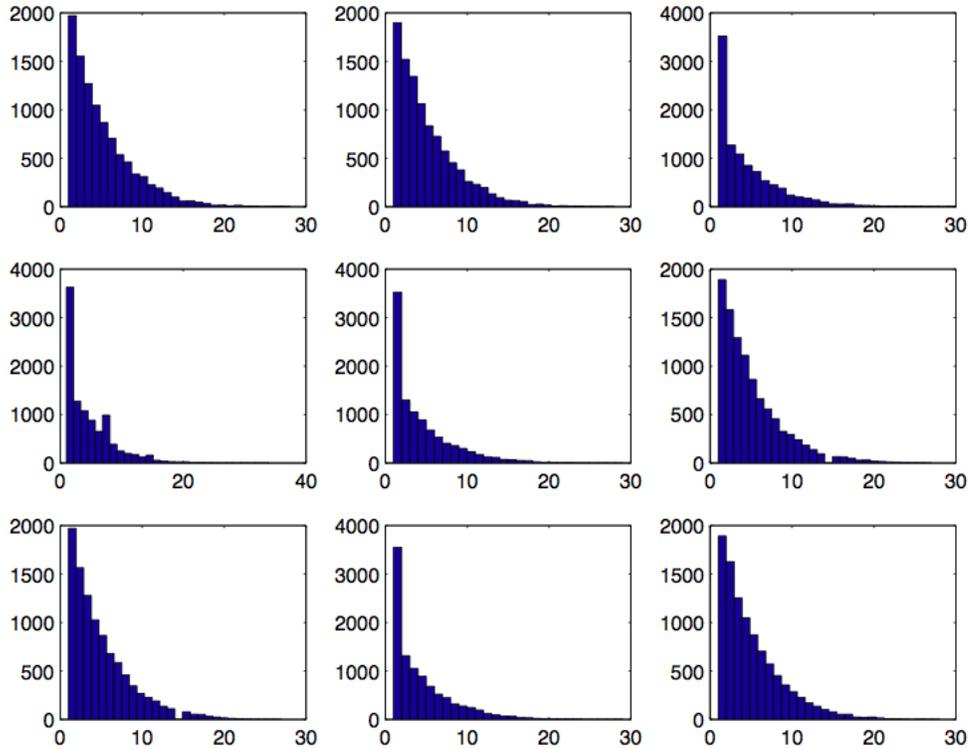
Fig 4: 9 spacings from a uniform shuffle; j varies from top left to bottom right, $1 \le j \le 9$.

you for."

We suggested a simple alternative: use the machine twice. This results in a shuffle equivalent to a 200-shelf machine. Our mathematical analysis and further tests, not reported here, show that this *is* adequately random. Indeed, Table 1 shows, for total variation, this is equivalent to 8-to-9 ordinary riffle shuffles.

22

References.

AGUIAR, M., BERGERON, N. and NYMAN, K. (2004). The peak algebra and the descent algebras of types B and D. *Trans. Amer. Math. Soc.*, **356** 2781–2824. URL `http://dx.doi.org/10.1090/S0002-9947-04-03541-X`.

AGUIAR, M., BERGERON, N. and SOTTILE, F. (2006). Combinatorial Hopf algebras and generalized Dehn-Sommerville relations. *Compos. Math.*, **142** 1–30. URL `http://dx.doi.org/10.1112/S0010437X0500165X`.

ALDOUS, D. (1983). Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on probability, XVII*, vol. 986 of *Lecture Notes in Math.* Springer, Berlin, 243–297.

ALDOUS, D. and DIACONIS, P. (1986). Shuffling cards and stopping times. *Amer. Math. Monthly*, **93** 333–348.

ASSAF, S., DIACONIS, P. and SOUNDARARAJAN, K. (2011). A rule of thumb for riffle shuffling. *Ann. Appl. Probab.*, **21** 843–875.

BAYER, D. and DIACONIS, P. (1992). Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.*, **2** 294–313.

BIDIGARE, P., HANLON, P. and ROCKMORE, D. (1999). A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements. *Duke Math. J.*, **99** 135–174.

BILLERA, L. J., HSIAO, S. K. and VAN WILLIGENBURG, S. (2003). Peak quasisymmetric functions and Eulerian enumeration. *Adv. Math.*, **176** 248–276. URL `http://dx.doi.org/10.1016/S0001-8708(02)00067-1`.

BLESSENOHL, D., HOHLWEG, C. and SCHOCKER, M. (2005). A symmetry of the descent algebra of a finite Coxeter group. *Adv. Math.*, **193** 416–437. URL `http://dx.doi.org/10.1016/j.aim.2004.05.007`.

BOREL, E. and CHÉRON, A. (1955). *Théorie mathématique du bridge à la portée de tous.* Gauthier-Villars, Paris. 2ème éd.

BROWN, K. S. and DIACONIS, P. (1998). Random walks and hyperplane arrangements. *Ann. Probab.*, **26** 1813–1854.

CIUCU, M. (1998). No-feedback card guessing for dovetail shuffles. *Ann. Appl. Probab.*, **8** 1251–1269. URL `http://dx.doi.org/10.1214/aoap/1028903379`.

COMTET, L. (1974). *Advanced combinatorics.* enlarged ed. D. Reidel Publishing Co., Dordrecht. The art of finite and infinite expansions.

CONGER, M. and VISWANATH, D. (2006). Riffle shuffles of decks with repeated cards. *Ann. Probab.*, **34** 804–819. URL `http://dx.doi.org/10.1214/009117905000000675`.

CONGER, M. A. and HOWALD, J. (2010). A better way to deal the cards. *Amer. Math. Monthly*, **117** 686–700. URL `http://dx.doi.org/10.4169/000298910X515758`.

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

DIACONIS, P. (1996). The cutoff phenomenon in finite Markov chains. *Proc. Nat. Acad. Sci. U.S.A.*, **93** 1659–1664.

DIACONIS, P. (2003). Mathematical developments from the analysis of riffle shuffling. In *Groups, Combinatorics & Geometry (Durham, 2001).* World Sci. Publ., River Edge, NJ, 73–97.

DIACONIS, P. and FULMAN, J. (2009a). Carries, shuffling, and an amazing matrix. *Amer. Math. Monthly*, **116** 788–803. URL http://dx.doi.org/10.4169/000298909X474864.

DIACONIS, P. and FULMAN, J. (2009b). Carries, shuffling, and symmetric functions. *Adv. in Appl. Math.*, **43** 176–196. URL http://dx.doi.org/10.1016/j.aam.2009.02.002.

DIACONIS, P. and FULMAN, J. (2011). Foulkes characters, Eulerian idempotents, and an amazing matrix. *Journ. Alg. Combin.* To appear.

DIACONIS, P. and GRAHAM, R. (2011). *Magical mathematics. The mathematical ideas that animate great magic tricks.* Princeton, NJ: Princeton University Press. xii, 244p. .

DIACONIS, P., MCGRATH, M. and PITMAN, J. (1995). Riffle shuffles, cycles, and descents. *Combinatorica*, **15** 11–29.

DIACONIS, P. and SHAHSHAHANI, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, **57** 159–179.

EPSTEIN, R. A. (1977). *The theory of gambling and statistical logic.* Revised ed. Academic Press [Harcourt Brace Jovanovich Publishers], New York.

ETHIER, S. N. (2010). *The Doctrine of Chances.* Probability and its Applications (New York), Springer-Verlag, Berlin. Probabilistic aspects of gambling, URL http://dx.doi.org/10.1007/978-3-540-78783-9.

FELLER, W. (1968). *An Introduction to Probability Theory and its Applications. Vol. I.* 3rd ed. John Wiley & Sons Inc., New York.

FULMAN, J. (1998). The combinatorics of biased riffle shuffles. *Combinatorica*, **18** 173–184.

FULMAN, J. (2000a). Affine shuffles, shuffles with cuts, the Whitehouse module, and patience sorting. *J. Algebra*, **231** 614–639.

FULMAN, J. (2000b). Semisimple orbits of Lie algebras and card-shuffling measures on Coxeter groups. *J. Algebra*, **224** 151–165. URL http://dx.doi.org/10.1006/jabr.1999.8157.

FULMAN, J. (2001). Applications of the Brauer complex: Card shuffling, permutation statistics, and dynamical systems. *J. Algebra*, **243** 96–122.

FULMAN, J. (2002). Applications of symmetric functions to cycle and increasing subsequence structure after shuffles. *J. Algebraic Combin.*, **16** 165–194.

Gannon, T. (2001). The cyclic structure of unimodal permutations. *Discrete Math.*, **237** 149–161. URL http://dx.doi.org/10.1016/S0012-365X(00)00368-X.

Gessel, I. M. and Reutenauer, C. (1993). Counting permutations with given cycle structure and descent set. *J. Combin. Theory Ser. A*, **64** 189–215. URL http://dx.doi.org/10.1016/0097-3165(93)90095-P.

Goncharov, V. (1944). Du domaine d'analyse combinatoire. *Bull. Acad. Sci. URSS Ser. Math* (Izv. Akad. Nauk SSSR), **8** 3–48. Amer. Math. Soc. Transl. (2) **19** (1962), 1–46.

Gontcharoff, W. (1942). Sur la distribution des cycles dans les permutations. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, **35** 267–269.

Grinstead, C. M. and Snell, J. L. (1997). *Introduction to Probability.* 2nd ed. American Mathematical Society, Providence, RI. URL http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/pdf.html.

Kerov, S. V. and Vershik, A. M. (1986). The characters of the infinite symmetric group and probability properties of the Robinson–Schensted–Knuth algorithm. *SIAM J. Algebraic Discrete Methods*, **7** 116–124. URL http://dx.doi.org/10.1137/0607014.

Klarreich, E. (2002). Coming up trumps. *New Scientist*, **175** 42–44.

Klarreich, E. (2003). Within every math problem, for this mathematician, lurks a card-shuffling problem. *SIAM News*, **36**. URL http://www.siam.org/pdf/news/295.pdf.

Lalley, S. P. (1996). Cycle structure of riffle shuffles. *Ann. Probab.*, **24** 49–73. URL http://dx.doi.org/10.1214/aop/1042644707.

Lalley, S. P. (1999). Riffle shuffles and their associated dynamical systems. *J. Theoret. Probab.*, **12** 903–932. URL http://dx.doi.org/10.1023/A:1021636902356.

Macdonald, I. G. (1995). *Symmetric Functions and Hall Polynomials.* 2nd ed. Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York. With contributions by A. Zelevinsky, Oxford Science Publications.

Mackenzie, D. (2002). The mathematics of . . . shuffling. *DISCOVER.* URL http://discovermagazine.com/2002/oct/featmath.

Mann, B. (1994). How many times should you shuffle a deck of cards? *UMAP J.*, **15** 303–332.

Mann, B. (1995). How many times should you shuffle a deck of cards? In *Topics in Contemporary Probability and its Applications.* Probab. Stochastics Ser., CRC, Boca Raton, FL, 261–289.

Morris, B. (2009). Improved mixing time bounds for the Thorp shuffle and L-reversal chain. *Ann. Probab.*, **37** 453–477. URL http://dx.doi.org/10.1214/08-AOP409.

Nyman, K. L. (2003). The peak algebra of the symmetric group. *J. Algebraic Combin.*, **17** 309–322. URL http://dx.doi.org/10.1023/A:1025000905826.

Petersen, T. K. (2005). Cyclic descents and P-partitions. *J. Algebraic Combin.*, **22** 343–375. URL http://dx.doi.org/10.1007/s10801-005-4532-5.

PETERSEN, T. K. (2007). Enriched P-partitions and peak algebras. *Adv. Math.*, **209** 561–610. URL http://dx.doi.org/10.1016/j.aim.2006.05.016.

POIRIER, S. (1998). Cycle type and descent set in wreath products. In *Proceedings of the 7th Conference on Formal Power Series and Algebraic Combinatorics (Noisy-le-Grand, 1995)*, vol. 180. 315–343. URL http://dx.doi.org/10.1016/S0012-365X(97)00123-4.

REINER, V. (1993). Signed permutation statistics and cycle type. *European J. Combin.*, **14** 569–579. URL http://dx.doi.org/10.1006/eujc.1993.1059.

ROGERS, T. D. (1981). Chaos in systems in population biology. In *Progress in Theoretical Biology, Vol. 6*. Academic Press, New York, 91–146.

SHEPP, L. A. and LLOYD, S. P. (1966). Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc.*, **121** 340–357.

STANLEY, R. P. (1999). *Enumerative Combinatorics. Vol. 2*, vol. 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

STANLEY, R. P. (2001). Generalized riffle shuffles and quasisymmetric functions. *Ann. Comb.*, **5** 479–491. Dedicated to the memory of Gian-Carlo Rota (Tianjin, 1999).

STARK, D., GANESH, A. and O'CONNELL, N. (2002). Information loss in riffle shuffling. *Combin. Probab. Comput.*, **11** 79–95. URL http://dx.doi.org/10.1017/S0963548301004990.

STEMBRIDGE, J. R. (1997). Enriched P-partitions. *Trans. Amer. Math. Soc.*, **349** 763–788. URL http://dx.doi.org/10.1090/S0002-9947-97-01804-7.

THIBON, J.-Y. (2001). The cycle enumerator of unimodal permutations. *Ann. Comb.*, **5** 493–500. Dedicated to the memory of Gian-Carlo Rota (Tianjin, 1999), URL http://dx.doi.org/10.1007/s00026-001-8024-6.

THORP, E. O. (1973). Nonrandom shuffling with applications to the game of Faro. *Journal of the American Statistical Association*, **68** 842–847. URL http://www.jstor.org/stable/2284510.

VERSHIK, A. M. and SHMIDT, A. A. (1977). Limit measures arising in the asymptotic theory of symmetric groups .1. *Theor. Probab. Appl.-Engl. Tr.*, **22** 70–85.

VERSHIK, A. M. and SHMIDT, A. A. (1978). Limit measures arising in the asymptotic theory of symmetric groups .2. *Theor. Probab. Appl.-Engl. Tr.*, **23** 36–49.

WARREN, D. and SENETA, E. (1996). Peaks and Eulerian numbers in a random sequence. *J. Appl. Probab.*, **33** 101–114.

DEPARTMENTS OF MATHEMATICS AND STATISTICS,
STANFORD UNIVERSITY
E-MAIL: diaconis@math.stanford.edu

DEPARTMENT OF MATHEMATICS,
UNIVERSITY OF SOUTHERN CALIFORNIA
E-MAIL: fulman@usc.edu

DEPARTMENT OF STATISTICS,
STANFORD UNIVERSITY
E-MAIL: susan@stat.stanford.edu