CARD SHUFFLING

EKAM KAUR

Abstract. We look at well known shuffling methods such as "Top to Random" Shuffle and Riffle Shuffle. We are mainly interested in a simple form of the the number of shuffles it takes of a deck of n cards to be uniformly random.

1. Introduction

A Riffle Shuffle is the most common type of shuffling method where a deck of cards is cut into about half and then riffles together. A precise model of doing this was introduced Gilbert and Shannon where a deck of n cards in cut into two portions according to a binomial distribution. Thus, the chance that k cards are cut-off is $\frac{\binom{n}{k}}{2^n}$. The two heaps are then riffled such that the probability that a card drops from the left or right heap is proportional to the number of cards in each heap.

2. Mixing Times

We start with the following lemma:

Lemma 2.1. *Let P be the transition matrix of a markov chain on a state space S where x is the starting state and t is the time. Then*

$$\|P^t(x,\cdot) - \pi\|_{TV} \leq s_x(t).$$

Proof.

$$\begin{aligned} \left\|P^t(x,.) - \pi\right\|_{TV} &= \sum_{y \in S} \left[\pi(y) - P^t(x,y)\right] \\ &= \sum_{y \in S} \quad \pi(y) \left[1 - \frac{P^t(x,y)}{\pi(y)}\right] \\ &=\leq \max_y \left[1 - \frac{P^t(x,y)}{\pi(y)}\right] = s_x(t). \end{aligned}$$

■

Definition 2.2. A *stopping time* is a map $\Gamma : W* \to [0,1]$ which shows the probability of continuing a walk $w \in (w_1, w_2, \dots, w_t) \in W*$. We define the *exact mixing time* of a chain to be the expected stopping time for an optimal stopping rule.

We also define the *halting time* as a state that given that we already stopped, we know it was never exited.

Combining the two definitions, we present the following theorem which we won't prove.

Theorem 2.3. *A stopping rule is optimal if and only if it has a halting state.*

Date: November 2020.

We don't prove this theorem in this paper, however it can be found in [3].

Lemma 2.4. *Let x be the starting state of a random walk (X_t) and τ be a strong stationary time. Then $s_x \leq Pr_x(\tau > t)$.*

Proof. Let y be the state for which the maximum of $1 - \Pr_x(X_t = y)/\pi(y)$ is achieved. Then,

$$s_x(t) = 1 - \frac{\Pr_x(X_t = y)}{\pi(y)} \tag{2.1}$$

$$\leq 1 - \frac{\Pr_x(X_t = y, \tau \leq t)}{\pi(y)} \tag{2.2}$$

$$= 1 - \frac{\Pr_x(\tau \leq t)\pi(y)}{\pi(y)} \tag{2.3}$$

$$= \Pr_x(\tau > t). \tag{2.4}$$

■

3. Top to Random Shuffles

We start by defining a "Top to Random Shuffle" as the following:

Take a deck with n cards. At each step, take the first card and insert it uniformally in any of the n places left in the deck.

The following two theorems give useful bounds on the mixing time on Top to Random Shuffles.

Theorem 3.1. *For a deck of n cards, the exact mixing time for a top to random shuffle is $n(H_{n-1} - 1) + 1$. As $n \to \infty$, the exact mixing time converges to $n \log(n-1) - n + 1$.*

Proof. Following is a an optimal stopping rule for top to random shuffle. Let's say we start from $\sigma = (\sigma_1, \ldots, \sigma_n)$. Mark the card σ_{n-1}. Shuffle until the marked card gets to the top. Do one more shuffle. Stop. It is easy to see that when we stop we are in uniform distribution. Moreover, this rule is optimal since any state in which the deck has σ_n on top is a halting state. Now, we need to calculate the expected time that takes the rule to stop.

Let T_i be time that it takes until i cards get under card σ_{n-1}. We know that $E[T_1] = 0$. Now consider the time $T_{i+1} - T_i$. This is the time that is needed for another card to get under the card σ_{n-1} given that there are already i cards below σ_{n-1}. Note that $T_{i+1} - T_i$ has geometric distribution with parameter $(i+1)/n$. Therefore, we have $E[T_{i+1} - T_i] = n/(i+1)$ Let τ be the stopping time for top to random shuffle. We have, $E[\tau] = E[T_{n-1}] + 1 = E[T_1] + E[T_2 - T_1] + \ldots + E[T_{n-1} - T_{n-2}] + 1 = n\sum_{i=2}^{n-1}(1/i) + 1 = n(H_{n-1} - 1) + 1$ ■

Lemma 3.2. *Given a deck on n cards, the mixing time of the top to random shuffle is less than $n \log n$.*

Proof. The stopping time τ that we gave in proof of Lemma 2.1 is in fact a strong stationary time. Let P be the transition matrix of the chain and U the uniform distribution. We have $\|P^t(\sigma, .) - U\| \leq s_\sigma(t) \leq \Pr(\tau > t)$ Claim. $\Pr(\tau > n \log n + cn) \leq e^{-c}$. Consider the coupon collector problem [4]. Notice that for T_i s in proof of Lemma 3.1 we have $\Pr(T_i - T_{i-1} = j) = \frac{i}{n}\left(1 - \frac{i}{n}\right)^{j-1}$ which is the same probability of how long it takes for the coupon collector to collect the $n - i + 1$ st coupon after collecting the $n - i$ th one. The stopping time τ in the above proof is in fact equal to the time it takes for the coupon collector to collect the last

$n-1$ st cards ($2\text{nd}, 3\text{rd}, \dots n$ th card) plus one extra step which is equal to the time needed to collect all $1, \dots, n$ coupons. Now, let's try to upper bound $\Pr(\tau > n \log n + cn)$. Let A_i be the event that the collector does not collect the coupon number i till time $n \log n + cn$. We have

$$\Pr(\tau > n \log n + cn) \leq \sum_{i=1}^{n} \Pr\left(A_i\right) = \sum_{i=1}^{n} \left(1 - \frac{1}{n}\right)^{n \log n + cn} \leq n \exp\left(-\frac{n \log n + cn}{n}\right) = e^{-c}$$

■

4. Riffle Shuffles

Definition 4.1. A Riffle Shuffle is a shuffle in which n cards are split in half and then alternatively interleaved from left to right. For example, 8 cards arranged as 1 2 3 4 5 6 7 8 becomes 5 1 6 2 7 3 8 4.

Here, we define an "unshuffle" which will help us understand the next lemma.

Definition 4.2. (Unshuffle) To each card in the deck assign a uniformly random bit ($0 or 1$). Pull the cards with label 0 to top of the deck preserving their relative order. The cards with label 1 will stay at the bottom preserving their relative order.

Lemma 4.3. *Let τ be an optimal stopping time for Riffle shuffle and $\tilde{\tau}$ to be an optimal stopping time for unshuffle. Then*

$$E\left[\tau\right] = E\left[\tilde{\tau}\right] \leq 2 \log n$$

.

Proof. The following is an optimal stopping rule for unshuffle. Unshuffle the cards and at each step, keep track of the bits that are assigned to each card. After t steps any card will be associated with a length t binary number. Stop when all n numbers are different. It is easy to check that this stopping rule generates the uniform distribution and the inverse of starting permutation is the halting state. Now, we should calculate the expectation of stopping time. Notice that we stop after t steps if we got n different numbers when we are allowed to choose from $\{0, 1, \dots, 2^t - 1\}$. Therefore, we have an instance of the Birthday problem. We use the results from Birthday problem to bound the expected stopping time. $\Pr(\tau \leq t) = \Pr($ we have n distinct numbers in range $\{1, \dots 2^t\}) = \left(1 - \frac{1}{2^t}\right)\left(1 - \frac{2}{2^t}\right) \dots (1 - \frac{n}{2^t}) \simeq \Pi_{i=1}^{n} \left(e^{-i/2^t}\right) = e^{\left(-\frac{1}{2^t}\right)\sum_{i=1}^{n} i} \simeq e^{\left(n^2/2^t\right)}$ Therefore, we have $E(\tau) = \sum_{t=1}^{\infty} \Pr(\tau \geq t) \simeq \sum_{t=1}^{\infty} \left(1 - e^{\left(n^2/2^t\right)}\right) = \sum_{t=1}^{\log n^2} \left(1 - e^{\left(n^2/2^t\right)}\right) + \sum_{t=\log n^2+1}^{\infty} \left(1 - e^{\left(n^2/2^t\right)}\right) \leq \log n^2 - \sum_{\log n+1}^{\infty} \left(n^2/2^t\right) \simeq 2 \log(n)$ Using the result from Lovász and Winkler ([8])^6, we know that the above exact mixing time for unshuffle is also the exact mixing time of shuffle. ■

5. Main Result

We first define a *rising sequence.*

Definition 5.1. A rising sequence in a permutation is the maximal set of consecutive numbers that occur in the correct order.

Now we introduce the following theorem:

Theorem 5.2. *If n cards are shuffled m times, then the chance that the deck is in arrangement π is $\binom{2^m+n-r}{n}/2^{mn}$, where r is the number of rising sequences in π.*

Sketch of Proof. First consider a generalization of Riffle shuffle to a -shuffle where the deck is cut to a piles and then the piles will be interleaving into each other. Definition 4.1 will have the following formulation in general case. ■

Theorem 5.3. *If n cards are shuffled m times with $m = 3\log_2 n + \theta$, then for large n,*

$$||Q^m - U|| = 1 - 2\Phi(-2^{-\theta}) + O(\frac{1}{n^4}),$$

with

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt.$$

Suppose that a deck consists of n cards, arranged $1, 2, ..., n$. If a riffle shuffle divides the deck into packets of k and $n - k$ cards, then riffling together these packets interleaves cards $1, 2, ..., k$ with cards $k + 1, ..., n$. This creates two rising sequences: Cards $1, 2, ..., k$ remain in relative order within the deck, as do cards $k + 1, k + 2, ..., n$. Successive shuffles tend to double the number of rising sequences (until the capacity of the deck is approached), so shuffling a 52 card deck three times usually creates eight rising sequences. From these eight rising sequences, one can reconstruct exactly how the deck was shuffled.

Now the performer takes back the pack, spreads it in a wide arc on the table, and, after staring intensely, names the selected card. To explain, consider what happens instead if the deck is never cut and the card is moved after the final shuffle. After three shuffles, the deck will usually have eight rising sequences, each consisting of an average of six and-a-half cards. Moving a card from the top to the middle of the deck usually creates a ninth rising sequence consisting of only the moved card, which is easily spotted.

If we view both the positions and face values of cards as having a cyclic order, then we can graph arrangements of cards on a torus, viewed as the product of two cycles. An unshuffled deck embeds as a $(1, l)$-cycle and a once shuffled deck embeds as a $(2, l)$-cycle. One sees that rising sequences are an artifact of where the torus is cut to make a square. Define the winding number of a deck to be the number of laps required to cycle through the deck by successive face values. A deck begins with winding number 1; each of the first few shuffles of a deck doubles its winding number. Moving a card usually increases the winding number by 1.We can identify the moved card by associating a count with each card, giving the total number of cards between its predecessor and successor, as we follow the winding sequence through the deck: Let $u(i)$ give the position of card i, and let $d(i, j)$ be the least positive integer so $d(i, j) = u(j) - u(i) \pmod n$. Then we associate with each card i the count $d(i - 1, i) + d(i, i + 1) - 1$. Ideally, the moved card will sit on its own lap of the winding sequence and its count will be the only count greater than n.

With this in mind, we present the following lemms.

Lemma 5.4. *The four descriptions generate the same permutation distribution. Moreover, in each model an a-shuffle followed by a b-shuffle is equivalent to an ab-shuffle.*

Proof. Each description results in a multinomial number of cards in each packet. This holds by decree for the sequential description and is clear for the inverse description. For the geometric description, the packet sizes are determined by how many points are chosen in each interval $[(i - 1)/a, (i/a)]$, which is also multinomial. For the maximum entropy description, the number of possible interleavings starting from a given cut is multinomial; they are in 1: 1 correspondence with the ways of dividing a deck into a subsets with the corresponding packet sizes.

Given the packet sizes, the maximum entropy description asserts that all possible interleavings are equally likely. This also clearly holds for the inverse description. For the sequential description, observe that when the first two piles of size j_1, j_2 are shuffled, the chance of any specific sequence of left-right drops is

$$\frac{j_1\,(j_1-1)\cdots 1\cdot j_2\,(j_2-1)\cdots 1}{(j_1+j_2)\,(j_1+j_2-1)\cdots 1}=\left(\begin{array}{c} j_1+j_2 \\ j_1 \end{array}\right)^{-1}$$

When these cards are shuffled into the third packet of size j_3, all $\left(\begin{array}{c} j_1+j_2+j_3 \\ j_3 \end{array}\right)$ positions for its cards are equally likely. This continues to hold for each successive packet.

No state information is retained between shuffles in these three models, so the product rule for a sequence of shuffles holds in each model once it is established for one. This easily follows from the inverse description: Lexicographically combining the pile assignments from an inverse a -shuffle and an inverse b -shuffle yields uniform and independent pile assignments for an inverse ab -shuffle.

For the geometric description, the lemma follows from the independence of base a digits of points picked uniformly in $[0,1]$: Choosing n points in $[0,1]$

labeling them with their leading digits and applying the map $x \to ax(\text{mod}1)$ is the same as choosing n points in $[0,1]$ and labeling them arbitrarily with integers from $\{0,\ldots,a-1\}$. Thus, all interleavings are equally likely for a given set of packet sizes. Moreover, the sets of n points which yield a given shuffle map to sets of n points distributed uniformly in [0,1] . Thus, successive shuffles can reuse the points x_i without first having to reposition them at random in $[0,1]$, so the product rule follows from the identity

$$b(ax(\text{mod}1))\,(\text{mod}1)=abx(\text{mod}1).$$

■

Theorem 5.5. *The probability that an a-shuffle will result in the permutation π is*

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

Proof. Using the maximum entropy description, this probability is determined by the number of ways of cutting an ordered deck into a packets, so π is a possible interleaving. Because each packet stays in order as the cards are riffled together, each rising sequence in the shuffled deck is a union of packets. Thus, we want to count the number of ways of refining r rising sequences into a packets.

We emulate the classical stars and bars argument, counting arrangements of cuts on the ordered deck before shuffling: At least one cut must fall between each successive pair of rising sequences of π, but the remaining cuts can be located arbitrarily. Thus, the n cards form dividers creating $n+1$ bins, into which the $a-r$ spare cuts are allocated. There are $\binom{a+n-r}{n}$ ways of doing this. There are a^n possible a -shuffles in all, giving the stated probability. ■

Summing the formula of the previous theorem over all permutations π gives 1. There are $A_{n,r}$ permutations with r rising sequences, where the $A_{n,r}$ are the Eulerian numbers. Thus, multiplying this sum by an gives

$$a^n = \sum_{r=1}^{n} A_{n,r} \binom{a+n-r}{n},$$

which is Worpitzky's identity.

Corollary 5.6. *If a deck of cards is given a sequence of m shuffles of types $a_1, a_2, \dots, a_m$, then the chance that the deck is in arrangement π is given by*

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

where $a = a_1, a_2, \dots, a_k$ and r is the number of rising sequences in π.

Proof. Combine Lemma 4.4 and Theorem 4.5. ■

Corollary 5.7. *Let a Markov Chain on the symmetric group begin at the identity and proceed by successive independent a-shuffles chosen from Gilbert-Shannon-Reeds measure. Then $R(\pi)$, the number of rising sequences, forms a Markov Chains.*

Proof. From Theorem 4.5 , the conditional law of π given $R(\pi)$ is uniform. Rogers and Pitman [(1981), Lemma 1] show that this, coupled with a completeness condition on the induced family of distributions for the process of rising sequences, is sufficient. In the present setting, completeness amounts to showing

$$\sum_{i=1}^{m} \left(a^m + n - r\right) f(r) = \textstyle\sum_{i=1}^{n} \binom{a^m+n-r}{n} g(r) \qquad \text{for } m = 0, 1, 2, \dots, \ .$$

implies $f = g$. The left side is the polynomial

$$\begin{aligned} &\tfrac{1}{n!}[(x+n-1)(x+n-2)\cdots x f(1) \\ &\quad +(x+n-2)(x+n-3)\cdots(x-1)f(2) \\ &\quad +\cdots + x(x-1)\cdots(x-(n-1))f(n)] \end{aligned}$$

evaluated at $x = a^m$. Evaluating at $x = i$ gives $f(i)$ ■

The final corollary connects shuffling to results in algebra. To describe things, define the group algebra $L(S_n)$ as the set of all functions from S, into the rational numbers $\mathbb{Q}$. Elements of L may be thought of as formal linear combinations of permutations with rational coefficients. Multiplication is given by formally multiplying the linear expressions using the multiplication on S_n. This is the same as convolving together the associated functions. In $L(S_n)$ let

$$A_i = \sum_{R(\pi)=i} \pi, \qquad\qquad i = 1, 2, \dots, n.$$

Corollary 5.8. *Let $\mathscr{A}$ be the subalgebra of $L(S_n)$ generated by $A_1, A_2, \dots, A_n$. Then $\mathscr{A}$ is a commutative, semisimple algebra of dimension n. A basis of primitive endempotents is given by $e_n(l) = \sum_{r=1}^{n} \sigma_i(n-r, \dots, 1-r)A_r$ with σ_l the l_{th} elementary symmetric function.*

Proof. Using the theorem, an a -shuffle can be represented in $\mathscr{A}$ as

$$B_a = \frac{1}{a^n} \sum_{r=1}^{n} \left(\begin{array}{c} a+n-r \\ n \end{array} \right) A_r$$

Now B_2^2 is a positive linear combination of $A_1 = Id, A_2$ and A_2^2. Theorem 3 gives $\bar{B}_2^2 = B_4 \in \mathscr{A}$. Thus A_2^2 is in $\mathscr{A}$. Next, B_2B_3 is a positive linear combination of A_2A_3, A_2^2, A_2, A_3 and A_1. It follows that $A_2A_3 \in \mathscr{A}$ and, from $B_2B_3 = B_3B_2 = B_6, A_2A_3 = A_3A_2$. From here A_3^2 and then $A_iA_3 = A_3A_i$ are in $\mathscr{A}$. Continuing inductively proves that $\mathscr{A}$ is a commutative algebra. To complete the proof, consider successive powers of B_2 :

$$\begin{aligned} B_2^m = B_{2^m} &= \frac{1}{2^{mn}} \sum_{r=1}^{n} \tbinom{2^m+n-r}{n} A_r \\ &= \frac{1}{n!} \sum_{l=0}^{n-1} \frac{1}{2^{lm}} \sum_{r=1}^{n} \sigma_l(n-r, \ldots, 1-r) A_r \end{aligned}$$

From this it follows that the linear map $\mathscr{A} \to \mathscr{A}$ given by multiplying by B_2 has distinct eigenvalues $1, 1/2, 1/2^2, \ldots, 1/2^{n-1}$. It is thus diagonalizable with the $e(l)$ as eigenvectors.

Left multiplication on itself gives a faithful representation of $\mathscr{A}$ as a commutative matrix algebra which we have shown contains an element B_2 with distinct eigenvalues. It follows that the set of matrices that commute with B_2 is all polynomials in B_2. Thus B_2 generates $\mathscr{A}$ (since elements of $\mathscr{A}$ commute with B_2). Thus, the $e(l)$ simultaneously diagonalize $\mathscr{A}$ which is therefore semisimple. ■

Proposition 5.9. *Let* $Q^m(r) = \frac{\binom{2^m+n-r}{n}}{2^{mn}}$ *be the probability of a permutation with* r *rising sequences after* m *shuffles from the GSR distribution. Let* $r = \frac{n}{2} + h$ *with* $-\frac{n}{2} + 1 \leq h \leq \frac{n}{2}$*. Let* $m = \log_2(n^{\frac{3}{2}}c)$ *with* $0 < c < \infty$ *fixed. Then*

$$Q^m(r) = \frac{1}{n!} exp \left\{ \frac{1}{c\sqrt{n}} \left(-h + \frac{1}{2} + O_c\left(\frac{n}{m}\right) \right) - \frac{1}{24c^2} - \frac{1}{2}\left(\frac{h}{cn}\right)^2 + O_c(\frac{1}{n}) \right\}.$$

Proof.

$$\begin{aligned} Q^m(r) &= \frac{1}{n!} \left(\frac{2^m+n-r}{2^m} \cdots \frac{2^m+1-r}{2^m} \right) \\ &= \frac{1}{n!} \exp \left\{ \sum_{i=0}^{n-1} \log\left(1 + \frac{(n/2)-h-i}{cn^{3/2}}\right) \right\} \end{aligned}$$

The logarithmic terms can be upper and lower bounded using

$$x - \frac{x^2}{2} + \frac{x^3}{3} - x^4 \leq \log(1+x) \leq x - \frac{x^2}{2} + \frac{x^3}{3}, \quad -\frac{1}{2} < x < 1$$

Standard summation formulas give

$$\begin{array}{l} \frac{1}{cn^{3/2}} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h - i\right) = \frac{-h+1/2}{c\sqrt{n}} \\ \frac{1}{2c^2n^3} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h - i\right)^2 = \frac{1}{24c^2} + \frac{1}{2}\left(\frac{h}{cn}\right)^2 + O_c\left(\frac{1}{n}\right) \\ \frac{1}{3c^3n^{9/2}} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h - i\right)^3 = O_c\left(\frac{h}{n^{3/2}}\right) \\ \frac{1}{c^4n^6} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h - i\right)^4 = O_c\left(\frac{1}{n}\right). \end{array}$$

■

Proposition 5.10. *With notation as in Proposition 4.9, let $h*$ be an integer such that $Q^m(\frac{n}{2}+h) \geq \frac{l}{n!} \iff h \leq h*$. Then, for any fixed c, as $n \to \infty$,*

$$h* = \frac{-\sqrt{n}}{24c} + \frac{1}{12c^3} + B + O_c\left(\frac{1}{\sqrt{n}}\right),$$

where $-1 \leq B \leq 1$.

Proof. Note that $Q^m(n/2+h) \geq 1/n!$ if and only if the exponent in Proposition 4.9 is nonnegative. Setting the exponent equal to 0 and collecting terms gives Proposition 4.10. □ ■

Theorem 5.11. *Let Q^m be the Gilbert-Shannon-Reeds distribution on the symmetric group S_n. Let U be the uniform distribution. For $m = \log_2\left(n^{3/2}c\right)$ with $0 < c < \infty$ fixed, as n tends to ∞,*

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

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

Proof. With notation as in Propositions 4.9 and 4.10 above, $\|Q^m - U\|$ equals

$$\sum_{-n/2<h\leq h^*} R_{nh}\left(Q^m\left(\frac{n}{2}+h\right) - \frac{1}{n!}\right) \tag{5.1}$$

where R_{nh} is the number of permutations with $n/2+h$ rising sequences. This uses the fact that the number of rising sequences is a sufficient statistic for both Q^m and U as explained in Section 3 and that total variation between two probabilities equals the total variation between the induced laws of any sufficient statistic [see, e.g., Diaconis and Zabell (1982), Lemma 6.1].

A permutation has r rising sequences if and only if π^{-1} has $r-1$ descents; see Section 3 for further discussion. The number of permutations with j descents is called the Eulerian number a_{nj}; see Tanny (1973), Stanley (1977) and other papers in the latter volume. Tanny and Stanley show that $a_{nj}/n!$ equals the chance that the sum of n random variables uniform on [0,1] is between j and $j+1$. Thus $a_{nj}/n!$ and $R_{nh}/n!$ obey the central limit theorem as in Tanny (1973). In particular if $x_n = h/\sqrt{n/12}$, the local limit theorem gives

$$\frac{R_{nh}}{n!} = \frac{e^{-(1/2)x_n^2}}{\sqrt{2\pi n/12}}\left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right) \quad \text{uniformly in } h \tag{5.2}$$

The usual form of the central limit theorem for the distribution function of $a_{nj}/n!$ and $R_{nh}/n!$ gives

$$\frac{1}{n!}\sum_{h=-n/2}^{h^*} R_{nh} = \Phi\left(\frac{-1}{4c\sqrt{3}}\right)\left(1 + O\left(\frac{1}{\sqrt{n}}\right)\right) \quad \text{uniformly}$$

The sum (4.1) can be broken into two zones. Recall from Proposition 2 that $h^* = -\sqrt{n}/24c + O(1)$

$$\text{zone 1: } \left\{\tfrac{-10n^{3/4}}{\sqrt{c}} \leq h \leq h^*\right\} = I_1$$
$$\text{zone 2 : } \left\{-\tfrac{n}{2} \leq h < \tfrac{-10n^{3/4}}{\sqrt{c}}\right\} = I_2$$

As will be shown, only zone 1 contributes. From (4.2) and Proposition 1, $\Sigma_{I_1} R_{nh} Q^m(n/2+h)$ equals

$$\begin{aligned}
&\tfrac{e^{-1/24c^2}}{\sqrt{2\pi n/12}} \textstyle\sum_{I_1} \exp\left\{-\tfrac{1}{2}\left(\tfrac{h}{\sqrt{n/12}}\right)^2 - \tfrac{h}{c\sqrt{n}} + O_c\left(\tfrac{1}{n^{1/4}}\right)\right\}\left\{1+o\left(\tfrac{1}{\sqrt{n}}\right)\right\}\\
&= \tfrac{e^{-1/24c^2}}{\sqrt{2\pi}} \textstyle\int_{-\infty}^{-(2c\sqrt{12})^{-1}} e^{-x^2/2-x/c\sqrt{3}} dx \left(1+O\left(\tfrac{1}{n^{1/4}}\right)\right)\\
&= \Phi\left(\tfrac{1}{4c\sqrt{3}}\right)\left(1+O\left(\tfrac{1}{n^{1/4}}\right)\right)
\end{aligned}$$

In zone $2, Q^m(n/2+h) \leq Q^m(1) \leq e^{\sqrt{n}/2c}/n!$ The standard large deviations bound as in Feller [(1971), Chapter 16] applied to the sum of n uniforms shows

$$\sum_{I_2} \frac{R_{nh}}{n!} \sim \frac{1}{10n^{1/4}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}}\right)^2\right]$$

Thus combining the bounds completes the proof. □ ■

The function $1-2\Phi(-1/4c\sqrt{3})$ has the following asymptotic behavior:

$$\begin{aligned}
&1-2\Phi\left(\tfrac{-1}{4c\sqrt{3}}\right) \sim \tfrac{1}{2c\sqrt{6\pi}} \text{ as } c \to \infty\\
&1-2\Phi\left(\tfrac{-1}{4c\sqrt{3}}\right) \sim 1-\tfrac{4c\sqrt{3}}{\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}\left(\tfrac{-1}{4c\sqrt{3}}\right)^2\right] \qquad \text{as } c \to 0.
\end{aligned}$$

Note that $m = \log_2\left(n^{3/2}c\right)$ has c inside the logarithm, so $c = 2^j$ where j is the number of shuffles beyond $\frac{3}{2}\log_2 n$ that have been performed. It follows that the variation distance tends to 0 exponentially in j for j positive. It tends to 1 doubly exponentially in j for j negative.

Thus it follows that about $\frac{3}{2}\log_2 n$ shuffles are needed to mix up n cards and that the deviation distance does not change much up until this point, and then drops a considerable amount (this is called the cut-off phenomenon).

We conclude by looking at the following table.

t	1	2	3	4	5	6	7	8	9	10
dv(t)	1.0	1.0	1.0	1.0	0.92	0.61	0.33	0.16	0.8	0.04

Note that the deviation does not change much until $t = 7$. However, at $t = 7$, the deviation distance is almost half of the previous amount, and continues to half each shuffle. Therefore 7 is considered as the ”cut-off” point.

References

[1] Bayer, Dave, Diaconis, Persi *Trailing the Dovetial Shuffle to its Lair* 1992. Columbia University. Harvard University. https://statweb.stanford.edu/~cgates/PERSI/papers/bayer92.pdf

[2] Haddadan, Shahrzad *Shuffling* March 2013. Dartmouth College. https://math.dartmouth.edu/~pw/math100w13/haddadan.pdf

[3] L. Lov´asz and P. Winkler *Mixing times, Microsurveys in Discrete Probability, D. Aldous and J. Propp, eds., DIMACS Series in Discrete Math. and Theoretical Computer Science 41, Amer. Math. Soc., Providence RI, pp. 85-134* 1998.