Information loss in riffle shuffling

Dudley Stark[*], A. Ganesh[†], Neil O'Connell[‡]

May 8, 2000
Submitted to *Combinatorics, Probability & Computing*

Abstract

We study the asymptotic behavior of the relative entropy (to stationarity) for a commonly used model for riffle shuffling a deck of n cards m times. Our results establish and were motivated by a prediction in a recent numerical study of Trefethen and Trefethen. Loosely speaking, the relative entropy decays approximately *linearly* (in m) for $m < \log_2 n$, and approximately *exponentially* for $m > \log_2 n$. The deck becomes random in this information theoretic sense after $m = \frac{3}{2}\log_2 n$ shuffles.

1 Introduction

The common riffle shuffle of a deck of n cards divides the deck into two packets and then interleaves them. A model for the riffle shuffle was introduced by Gilbert and Shannon [6] and independently by Reeds [7]. The Gilbert-Shannon-Reeds (GSR) model is described in Section 2. Aldous [1] and Bayer and Diaconis [2] have shown that, under the GSR model, it takes $m = \frac{3}{2}\log_2 n$ shuffles to randomize a deck of n cards. This is to say, they show that this is the point at which there is a sharp cutoff (from 1 to 0) in the total variation distance between the law of the deck configuration and the uniform

[*]Queen Mary, University of London, Mile End Road, London, E1 4NS, UK
[†]Microsoft Research, St. George House, 1 Guildhall Street, Cambridge CB2 3NH, UK
[‡]BRIMS, Hewlett-Packard Laboratories, Filton Road, Bristol BS34 8QZ, UK

distribution, assuming n is large. Bayer and Diaconis obtain more precise asymptotics in the region of the cutoff. The use of total variation distance to measure rates of convergence to equilibrium for Markov chains (and random walks on groups) is largely a matter of convention, although often bounds are obtained on the l_2 distance which in turn provide bounds on the total variation distance. The relative entropy (which typically lies somewhere in between the two) has not been considered to the same extent, particularly in the context of cutoff phenomena. There is a chapter in the thesis of Su [9] which is devoted to this topic, where some elementary inequalities relating the relative entropy to total variation distance and l_2 distance, as well as detailed asymptotics for the random walk on a hypercube, are presented. Similar inequalities are discussed by Diaconis and Saloff-Coste [5]. The relative entropy decreases monotonically for Markov chains; see [3], Section 2.9.

In a recent numerical study, Trefethen and Trefethen [10] observed that for the GSR model:

(a) the threshold at $m = \frac{3}{2}\log_2 n$ shuffles is also observed in relative entropy;

(b) after exactly $m = \frac{3}{2}\log_2 n$ shuffles, 0.0601 bits of information remain;

(c) the relative entropy decays *linearly* for $m < \log_2 n$, and *geometrically* for $m > \log_2 n$.

The purpose of this paper is to make these observations precise, by careful analysis of the GSR model. Our relative entropy asymptotics complement the total variation distance asymptotics of [2].

The observation (a) is not surprising, and can be expected to hold quite generally. However, the observation (c) is rather puzzling and it is not clear in what sense it might be a general phenomenon. Consider, for example, the (continuous time) random walk on a hypercube, where the relative entropy between the position of the walker after time t and the uniform distribution on an n-dimensional cube is given by

$$\frac{n}{2}\left[(1-e^{-2t/n})\log_2(1-e^{-2t/n})+(1+e^{-2t/n})\log_2(1+e^{-2t/n})\right].$$

For this model, there is a sharp threshold in total variation distance at $t = \frac{1}{4}n\ln n$, and the same threshold is observed in relative entropy. There is also a similar kind of secondary phase transition at $t = n$: for $t << n$, the relative entropy is

$$n + t \log_2 \frac{2t}{n} + O(t)$$

and for $t >> n$, the relative entropy is

$$\frac{n}{2} e^{-4t/n} + O\left(ne^{-6t/n}\right).$$

The rate of decay before the transition is not linear, but we do see a sudden switch to geometric decay. If one were to speculate on what might be a general rule, perhaps it would be that there is a kind of 'pre-threshold', at some point earlier than the absolute cutoff, at which point the relative entropy begins to decay geometrically.

In the next two sections we state and prove the main results of the paper, on the relative entropy to equilibrium in the GSR model.

2 The main results

The Gilbert-Shannon-Reeds model [6, 7] is a model for riffle shuffling a deck of cards. Statistical experiments presented in [4] show that it is a realistic description of the way people shuffle cards.

Several equivalent descriptions of the GSR model are given in [2]. Here is one of them. A deck of n cards is cut into two packets according to a binomial distribution. The probability that k cards are cut off is $\binom{n}{k}/2^n$. The two packets are then riffled together in such a way that cards drop from the left or right heaps with probability proportional to the number of cards in each heap. Thus, if there are A and B remaining in the left and right heaps, then the chance that the next card will drop from the left heap is $A/(A+B)$.

For example, suppose that the number of cards in a deck is $n = 8$ and that the cards are marked successively with the numbers $1, 2, 3, 4, 5, 6, 7, 8$. The first packet could, for example, be $1, 2, 3, 4, 5$ with probability $\binom{8}{5}/2^8 = 56/1024$, in which case the second packet is $6, 7, 8$. The probability that 5 drops first is $\frac{5}{8}$, the probability that 8 drops next is $\frac{3}{7}$, the probability that 4 drops next is $\frac{4}{6}$, and so on, so that the probability the packets are riffled together in such a way that the new deck has the order $6, 1, 2, 3, 7, 4, 8, 5$ is

$$\frac{5}{8} \times \frac{3}{7} \times \frac{4}{6} \times \frac{2}{5} \times \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} \times \frac{1}{1} = \frac{5!3!}{8!} = \frac{720}{40320}.$$

Each shuffle of a deck of n cards corresponds to a permutation in S_n. In our example the corresponding permutation is given by

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 1 & 2 & 3 & 7 & 4 & 8 & 5 \end{pmatrix}. \tag{1}$$

In the GSR model all interleavings after the deck has been cut into two packets have equal probability. The probability of a given cut followed by a given interleaving is $\binom{n}{k} 2^{-n} \times \binom{n}{k}^{-1} = 2^{-n}$. Another way of describing the GSR model from [1, 2] is to let B_i, $i = 1, \dots, n$ be indicators such that $B_i = 0$ when the ith card of the shuffled deck comes from the first packet and such that $B_i = 1$ when the ith card of the shuffled deck comes from the second packet. In the GSR model, the B_i are i.i.d. Bernoulli variables with $\mathbb{P}(B_i = 0) = \mathbb{P}(B_i = 1) = 1/2$. Suppose that $n = 6$ and $B_1 = 1$, $B_2 = 1$, $B_3 = 0$, $B_4 = 1$, $B_5 = 1$, and $B_6 = 0$. Since only two B_i's take on the value 0 there were only two cards in the first packet. These two cards must have marks 1 and 2. There were four cards in the second packet and they had marks $3, 4, 5, 6$. Since the cards marked 1 and 2 must retain their original order in the shuffled deck, we deduce that after the riffle shuffle the card marked 1 is now the third card in the deck and the card marked 2 is now the sixth card in the deck. Similarly, the card marked 3 is now the first card, the card marked 4 is now the second card, the card marked 5 is now the fourth card, and the card marked 6 is now the fifth card. The corresponding permutation is

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 1 & 5 & 6 & 2 \end{pmatrix}. \tag{2}$$

Consider the process of performing m riffle shuffles. Each card $k = 1, \dots n$ will be assigned m independent Bernoulli variables $B_1(k), B_2(k), \dots, B_m(k)$. The variable $B_j(k)$ indicates whether or not the kth card in the deck after the m shuffles have been performed was in the first packet packet of the jth shuffle. If $B_1(k) < B_1(k')$ for $k \neq k'$, then for the first shuffle the card now in the kth position was in the first packet and the card now in the k'th position was in the second packet. Therefore the card now in the kth position has a smaller mark than the card now in the k'th position. Suppose that $B_1(k) = B_1(k')$ but that $B_2(k) < B_2(k')$. Then the cards in the kth and k'th positions in the shuffled decks were in the same packets in the first shuffle and hence retained their original order relative to one another after the first shuffle. For the second shuffle the card now in the kth position was in the first packet and the card now in the k'th position was in the second packet. Therefore, the card now in the kth position has a smaller mark than the card now in the k'th position.

Let $D_k = \sum_{j=1}^{m} B_j(k)2^{m-j}$. Arguing as in the last paragraph, we see that if $D_k < D_{k'}$, $k \neq k'$, then the card in the kth position of the shuffled deck has a smaller mark than the card in the k'th position. Cards in the shuffled deck in positions with identical D_k retain their original orders relative to one another. Moreover, the D_k, $k = 1, 2, \ldots, n$ are independent uniformly distributed variables on $\{0, 1, 2, \ldots, 2^m - 1\}$.

Let us illustrate a deck shuffled m times under the GSR model with an example. Suppose $n = 8$, $m = 2$ and the values of the D_i are $D_1 = 3$, $D_2 = 1$, $D_3 = 0$, $D_4 = 0$, $D_5 = 2$, $D_6 = 3$, $D_7 = 0$, $D_8 = 2$. The cards in the shuffled deck with $D_k = 0$ have smaller marks than the other cards and they remain in their original order relative to one another. Therefore the card marked 1 is now in the third position, the card marked 2 is in the fourth position, and the card marked 3 is in the seventh position. The card with $D_k = 1$ has the next smallest mark, so the card marked 4 is now in the second position. The cards with $D_k = 2$ have the next smallest marks, so the card marked 5 is in the fifth position and the card marked 6 is in the eighth position. The cards with $D_k = 3$ have the largest marks and the card marked 7 is now in the first position and the card marked 8 is now in the sixth position. The corresponding permutation is

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 7 & 4 & 1 & 2 & 5 & 8 & 3 & 6 \end{pmatrix}. \tag{3}$$

There are some card orders, such as $8, 7, 6, 5, 4, 3, 2, 1$, which are impossible to achieve after one riffle shuffle because all of the cards in a packet must appear in increasing order in the shuffled deck. The permutations corresponding to possible orderings of the deck after one riffle shuffle can have at most two *rising sequences*. A rising sequence is a maximal subsequence of an arrangement of cards, consisting of successive values displayed in any order. In example (1) there are two rising sequences: $1, 2, 3, 4, 5$ and $6, 7, 8$.

In example (2) there are two rising sequences: $1, 2$ and $3, 4, 5, 6$. In example (3) there are three rising sequences: $1, 2, 3$ and $4, 5, 6$ and $7, 8$.

We denote the distribution on S_n induced by the GSR model with m shuffles by $P_{n,m}$. It is shown in [2] that the probability density $P_{n,m}(\pi)$, $\pi \in S_n$ depends only on the number of rising sequences of π. If $r = r(\pi)$ is the number of rising sequences in $\pi \in S_n$ then

$$P_{n,m}(\pi) = \binom{2^m + n - r}{n} 2^{-mn}. \tag{4}$$

Note that a deck of cards shuffled m times can have no more than $2^m \wedge n$ rising sequences.

Let U be the uniform distribution on S_n, so that $U(\pi) = 1/n!$ for all $\pi \in S_n$. It is well known that the GSR model has U as a stationary distribution. The relative entropy to stationarity $H(P_{n,m}|U)$ is

$$H(P_{n,m}|U) = \sum_{\pi \in S_n} P_{n,m}(\pi) \log_2 \left(n! P_{n,m}(\pi)\right). \tag{5}$$

We derive expressions for $H(P_{n,m}|U)$, which demonstrate a linear/geometric transition in the decay of $H(P_{n,m}|U)$ at about $m = \log_2 n$.

Theorem 1 *Let $\omega(n)$ be any sequence such that $\omega(n) \geq 2$ for all n and such that $\omega(n) \to \infty$. Uniformly for all m such that $m \leq \log_2 n - \omega(n)$ we have*

$$H(P_{n,m}|U) \geq \log_2 n! - mn$$

and

$$H(P_{n,m}|U) \leq \log_2 n! - mn + n[\omega(n) + 3]2^{-\omega(n)}.$$

Define the constant $\alpha := \log_2 e \doteq 1.44270$. and let $k > 2$ be any constant. Uniformly for m such that

$$m \in [\log_2 n - k, \log_2 n - 2],$$

the relative entropy is bounded below by

$$H(P_{n,m}|U) \geq (2 - \alpha)n$$

and bounded above by

$$H(P_{n,m}|U) \leq (k - \alpha + 2)n + \frac{1}{2} + \frac{\alpha}{8n}.$$

Let $\omega(n)$ be any sequence such that $\omega(n) \geq 0$ for all n and such that $\omega(n) \to \infty$. Uniformly for all m such that $\log_2 n + \omega(n) \leq m \leq 2\log_2 n$ we have

$$H(P_{n,m}|U) = \frac{\alpha}{24} n^3 2^{-2m} + O(n^4 2^{-3m}). \tag{6}$$

Uniformly for all m such that $m > 2\log_2 n$ we have

$$H(P_{n,m}|U) = \frac{\alpha}{24} n^3 2^{-2m} + O(n^2 2^{-2m}). \tag{7}$$

We display the linear/geometric transition more explicitly in our first corollary.

Corollary 1 *Let $\omega(n)$ be any sequence such that $\omega(n) \geq 0$ for all n and such that $\omega(n) \to \infty$. Uniformly for all m such that $m \leq \log_2 n - \omega(n)$, $\omega(n) \to \infty$ we have*

$$H(P_{n,m+1}|U) - H(P_{n,m}|U) = -n + o(n).$$

Uniformly for all m such that $m \geq \log_2 n + \omega(n)$, $\omega(n) \to \infty$ we have

$$\frac{H(P_{n,m+1}|U)}{H(P_{n,m}|U)} = \frac{1}{4} + o(1).$$

Proof The first statement follows automatically from Theorem 1. The second statement follows from Theorem 1 because whenever $m/\log_2 n \to \infty$

$$n^4 2^{-3m}/(n^3 2^{-2m}) = n 2^{-m} = o(1)$$

and

$$n^2 2^{-2m}/(n^3 2^{-2m}) = n^{-1} = o(1).$$

■

Our second corollary demonstrates that $m = \frac{3}{2}\log_2 n$ is the threshold to mix a deck of cards in relative entropy.

Corollary 2 *If $m = \frac{3}{2}\log_2 n + c$, c a real constant, then*

$$H(P_{n,m}|U) = \frac{\alpha}{24} 4^{-c} + O_c(n^{-1/2}).$$

As we remarked in the Introduction, it is observed in [10] that $H(P_{n,m}|U) \doteq 0.06011$ when $m = \frac{3}{2}\log_2 n$, agreeing with our limit $H(P_{n,m}|U) \to \alpha/24 \doteq 0.060112293$. Corollary 2 may also be shown by using the normality of the Eulerian numbers in the way that [2] proves limits for total variation distance when $m = \frac{3}{2}\log_2 n + c$.

3 Proofs

Lemma 1, Lemma 2, and, following some technical lemmas, Lemma 11, handle the cases $m \leq \log_2 n - \omega(n)$, $m = \log_2 n + O(1)$, and $m \geq \log_2 n + \omega(n)$, respectively.

Lemma 1 *Let $\omega(n)$ be any integer valued function such that $\omega(n) \geq 2$ for all n and such that $\omega(n) \to \infty$. Uniformly for all m such that $m \leq \log_2 n - \omega(n)$ we have*

$$H(P_{n,m}|U) \geq \log_2 n! - mn \tag{8}$$

and

$$H(P_{n,m}|U) \leq \log_2 n! - mn + n[\omega(n) + 3]2^{-\omega(n)}. \tag{9}$$

Proof For $n \geq 1$, $m \geq 1$, and $r \in [1, 2^m \wedge n]$, define $A_{n,m}(r)$ to be

$$A_{n,m}(r) := \binom{n + 2^m - r}{n}.$$

Clearly

$$A_{n,m}(r) \geq 1 \tag{10}$$

and

$$A_{n,m}(r) \leq \binom{n + 2^m}{n} \tag{11}$$

uniformly for r restricted to its possible values $r \in [1, 2^m \wedge n]$. The probability density on S_n of m riffle shuffles on n cards (4) can be written for $\pi \in S_n$ and $r = r(\pi)$ as

$$P_{n,m}(\pi) = 2^{-mn} A_{n,m}(r). \tag{12}$$

Use of (10) and (12) in (5) and the monotonicity of the logarithm function now produces

$$\begin{aligned} H(P_{n,m}|U) &= \sum_{\pi \in S_n} P_{n,m}(\pi) \log_2 (n! P_{n,m}(\pi))) \\ &\geq \sum_{\pi \in S_n} P_{n,m}(\pi) \log_2 \left(n! 2^{-mn}\right) \\ &= \log_2 \left(n! 2^{-mn}\right) \sum_{\pi \in S_n} P_{n,m}(\pi) \\ &= \log_2 n! - mn, \end{aligned} \tag{13}$$

which is (8).

As for (9), because (11) holds uniformly for every $r \in [1, 2^m \wedge n]$, we have

$$\begin{aligned}
\log_2(n! P_{n,m}(\pi)) &= \log_2(n! 2^{-mn} A_{n,m}(r)) \\
&\leq \log_2 \left(n! 2^{-mn} \binom{n+2^m}{n} \right)
\end{aligned}$$

and so

$$\begin{aligned}
H(P_{n,m}|U) &= \sum_{\pi \in S_n} P_{n,m}(\pi) \log_2 \left(n! P_{n,m}(\pi) \right) \\
&\leq \sum_{\pi \in S_n} P_{n,m}(\pi) \log_2 \left(n! 2^{-mn} \binom{n+2^m}{n} \right) \\
&= \log_2 \left(n! 2^{-mn} \binom{n+2^m}{n} \right) \sum_{\pi \in S_n} P_{n,m}(\pi) \\
&= \log_2 n! - mn + \log_2 \binom{n+2^m}{n}. \qquad (14)
\end{aligned}$$

We observe that

$$\binom{n+2^m}{n} p^n (1-p)^{2^m} \leq 1 \quad \text{for all} \;\; p \in [0,1],$$

and so,

$$\log_2 \binom{n+2^m}{n} \leq -n \log_2 p - 2^m \log_2(1-p) \quad \text{for all} \;\; p \in [0,1].$$

Take $p = (1 + 2^{-\omega(n)})^{-1}$, so that $1 - p = (1 + 2^{\omega(n)})^{-1}$. Substituting in the above, and using the fact that $-\log_2(1-p)$ is positive and $2^m \leq n 2^{-\omega(n)}$ by assumption, we get

$$\log_2 \binom{n+2^m}{n} \leq n \log_2(1 + 2^{-\omega(n)}) + n 2^{-\omega(n)} \log_2(1 + 2^{\omega(n)}). \qquad (15)$$

Now $\log_2(1+x) = \log_2 e \cdot \log_e(1+x) \leq x \log_2 e$, while $1 + 2^{\omega(n)} \leq 2 \cdot 2^{\omega(n)}$. Using these inequalities in (15), we obtain,

$$\begin{aligned}
\log_2 \binom{n+2^m}{n} &\leq n 2^{-\omega(n)} \log_2 e + n 2^{-\omega(n)} (1 + \omega(n)) \\
&\leq n 2^{-\omega(n)} (\omega(n) + 3).
\end{aligned}$$

Substituting the last inequality in (14) yields (9). ■

Lemma 2 *Let $k > 2$ be any constant. Let α be defined by $\alpha := \log_2 e \doteq 1.44270 < 2$. Uniformly for m such that*

$$m \in [\log_2 n - k, \log_2 n - 2],$$

the relative entropy is bounded below by

$$H(P_{n,m}|U) \geq (2-\alpha)n$$

and bounded above by

$$H(P_{n,m}|U) \leq (k-\alpha+2)n + \frac{1}{2} + \frac{\alpha}{8n}.$$

Proof We will use a version of Stirling's formula from [7]:

$$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n} e^{c_n}, \tag{16}$$

where

$$\frac{1}{12n+1} < c_n < \frac{1}{12n}.$$

The lower bound from (13) is valid for all $m \geq 1$. Thus, by (16), we get

$$\begin{aligned} H(P_{n,m}|U) &\geq \log_2(n!) - mn \\ &\geq \log_2\left(\left(\frac{n}{e}\right)^n\right) - mn \\ &\geq (n\log_2 n - \alpha n) - (\log_2 n - 2)n \\ &= (2-\alpha)n. \end{aligned}$$

For the upper bound, we will use (14). Note that

$$\begin{aligned} \log_2\binom{n+2^m}{n} &\leq \log_2\binom{2n}{n} \\ &= \log_2(2n)! - 2\log_2(n!) \\ &= 2n\log_2(2n) - 2(n\log_2 n) + \frac{1}{2}\log_2\left(\frac{2n}{2\pi n^2}\right) + \log_2\left(e^{c_{2n}-2c_n}\right) \\ &\leq 2n\log_2(2n) - 2(n\log_2 n) - \frac{1}{2}\log_2(\pi n) + \log_2 e^{c_{2n}} \\ &= 2n - \frac{1}{2}\log_2(\pi n) + \alpha c_{2n} \\ &\leq 2n - \frac{1}{2}\log_2(\pi n) + \frac{\alpha}{24n}, \end{aligned}$$

uniformly for all $m \le \log_2 n$. From (14), we now have

$$\begin{aligned}
H(P_{n,m}|U) &\le \log_2 n! - mn + 2n - \frac{1}{2}\log_2(\pi n) + \frac{\alpha}{24n} \\
&\le \log_2 n! - (\log_2 n - k)n + 2n - \frac{1}{2}\log_2(\pi n) + \frac{\alpha}{24n} \\
&= n\log_2 n - n\alpha + \frac{1}{2}\log_2(2\pi n) + c_n\alpha - (\log_2 n - k)n + 2n \\
&\quad -\frac{1}{2}\log_2(\pi n) + \frac{\alpha}{24n} \\
&\le (k - \alpha + 2)n + \frac{1}{2} + \frac{\alpha}{12n} + \frac{\alpha}{24n} \\
&< (k - \alpha + 2)n + \frac{1}{2} + \frac{\alpha}{8n}.
\end{aligned}$$

■

Let R be the random variable defined on the measure space of permutations that counts the number of rising sequences. In bounding the relative entropy for $m \ge \log_2 n + \omega(n)$, $\omega(n) \to \infty$, the major effort goes into estimating $\mathbb{E}(R)$ and $\mathrm{Var}(R)$ under the $P_{n,m}$ measure. Lemma 3 gives a formula for $\mathbb{E}(R)$ and an asymptotic expression for $\mathbb{E}(R)$ is found in Lemma 9. The intervening lemmas are technical in nature. A bound on $\mathrm{Var}(R)$ is given in Lemma 10. We then derive an expression for the relative entropy in Lemma 11.

We begin with Lemma 3, which gives formulas for R and $\mathbb{E}(R)$.

Lemma 3 *Suppose that a realization of the D_i takes on L distinct values between 0 and $2^m - 1$, so that L is a random variable, and list the values in order $0 \le c_1 < c_2 < \cdots < c_{L-1} < c_L \le 2^m - 1$. Define the indicator variables I_j for $j \in [1, L-1]$ by*

$$I_j = I\left[\inf\{i : D_i = c_{j+1}\} > \sup\{i : D_i = c_j\}\right]. \tag{17}$$

The number of rising sequences is given by

$$R = L - \sum_{j=1}^{L-1} I_j.$$

Let $N_j = \sum_{k=1}^{n} I[D_k = c_j]$, $j \in [1, L]$, count the number of times the value c_j is taken on. Let J_{l_1,l_2}, $1 \le l_1, l_2 \le n$, be the random variables defined *by*

$$J_{l_1,l_2} = \sum_{j=1}^{L-1} I\left[N_j = l_1, N_{j+1} = l_2\right]. \tag{18}$$

The expectation of R under the $P_{n,m}$ measure is

$$\mathbb{E}(R) = \mathbb{E}(L) - \sum_{1\le l_1,l_2\le n} \mathbb{E}(J_{l_1,l_2}) \binom{l_1+l_2}{l_2}^{-1}. \tag{19}$$

Proof For $k \in [1, L]$, let $1 \le i_{1,k} < i_{2,k} < \ldots < i_{N_k,k} \le n$ denote the indices such that $D_{i_{j,k}} = c_k$ for all $j \in [1, N_k]$. Let $\pi_{\mathbf{D}}$ denote the permutation corresponding to $\mathbf{D} = (D_1, D_2, \ldots, D_n)$. By definition $\pi_{\mathbf{D}}(i_{j,1}) = j$ for $j \in [1, N_1]$. Similarly, $\pi_{\mathbf{D}}(i_{j,2}) = N_1 + j$ for all $j \in [1, N_2]$ and $\pi_{\mathbf{D}}(i_{j,k}) = N_1 + \cdots + N_{k-1} + j$ for all $k \in [1, L]$ and $j \in [1, N_k]$.

The rising sequence starting with 1 begins with $1, 2, \ldots, N_1$. A new rising sequence begins with $N_1+1, N_1+2, \ldots, N_1+N_2$ if and only if $i_{1,2} < i_{N_1,1}$. In general, a new rising sequence begins with $N_1+\cdots+N_{k-1}+1, \ldots, N_1+\cdots+N_k$ if and only if $i_{1,k} < i_{N_{k-1},k-1}$. The number of rising sequences is therefore

$$R = L - \sum_{j=1}^{L-1} I_j.$$

Define $\mathcal{F}$ to be the σ-field generated by L, $(c_1, c_2, \ldots, c_L)$, and $(N_1, N_2, \ldots, N_L)$. The conditional expectation of an indicator I_j given $\mathcal{F}$ is just the number of ways linearly arranging N_j indistinguishable blue balls and N_{j+1} indistinguishable red balls so that the blue balls are all to the left of the red balls, which is 1, divided by the total number of ways of arranging the balls, which is $\binom{N_j+N_{j+1}}{N_j}$. Therefore

$$\mathbb{E}(I_j|\mathcal{F}) = \binom{N_j+N_{j+1}}{N_j}^{-1}. \tag{20}$$

The expectation of R is

$$\begin{aligned} \mathbb{E}(R) &= \mathbb{E}(\mathbb{E}(R|\mathcal{F})) \qquad (21)\\ &= \mathbb{E}\left(L - \sum_{j=1}^{L-1}\binom{N_j+N_{j+1}}{N_j}^{-1}\right) \end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}(L) - \mathbb{E}\left(\sum_{j=1}^{L-1} \sum_{1\le l_1,l_2\le n} I\left[N_j = l_1, N_{j+1} = l_2\right] \binom{l_1+l_2}{l_2}^{-1}\right) \\
&= \mathbb{E}(L) - \mathbb{E}\left(\sum_{1\le l_1,l_2\le n} \sum_{j=1}^{L-1} I\left[N_j = l_1, N_{j+1} = l_2\right] \binom{l_1+l_2}{l_2}^{-1}\right) \\
&= \mathbb{E}(L) - \sum_{1\le l_1,l_2\le n} \mathbb{E}\left(\sum_{j=1}^{L-1} I\left[N_j = l_1, N_{j+1} = l_2\right]\right) \binom{l_1+l_2}{l_2}^{-1} \\
&= \mathbb{E}(L) - \sum_{1\le l_1,l_2\le n} \mathbb{E}(J_{l_1,l_2}) \binom{l_1+l_2}{l_2}^{-1}.
\end{aligned}
$$

■

Lemma 4 *Let $\omega(n) \ge 0$ be any sequence such that $\omega(n) \to \infty$. If $m \ge \log_2 n + \omega(n)$, then the expectation of the variable L defined in Lemma 3 has asymptotics*

$$\mathbb{E}(L) = n - \binom{n}{2} 2^{-m} + O\left(n^3 2^{-2m}\right).$$

Proof Recall that $D_i, i = 1, \ldots, n$ are independent, and uniformly distributed on $\{0, \ldots, 2^m - 1\}$. The probability that any given value is taken on by at least one D_i is $1 - (1 - 2^{-m})^n$. Therefore,

$$
\begin{aligned}
\mathbb{E}(L) &= 2^m \left(1 - \left(1 - 2^{-m}\right)^n\right) \\
&= 2^m \left(n 2^{-m} - \binom{n}{2} 2^{-2m} + O\left(n^3 2^{-3m}\right)\right) \\
&= n - \binom{n}{2} 2^{-m} + O\left(n^3 2^{-2m}\right).
\end{aligned}
$$

■

Only $\mathbb{E}(J_{1,1})$, $\mathbb{E}(J_{1,2})$, and $\mathbb{E}(J_{2,1})$ contribute significantly to (19). Estimating the $\mathbb{E}(J_{l_1,l_2})$ takes the next several lemmas.

Lemma 5 *Let $\omega(n) > 0$ be any sequence such that $\omega(n) \to \infty$. For $m \ge \log_2 n + \omega(n)$, the expected value of $J_{2,2}$ defined by (18) is bounded by*

$$\mathbb{E}(J_{2,2}) = O(n^3 2^{-2m}).$$

Proof Let s and t, $0 \leq s < t \leq 2^m - 1$ be such that $N_s = N_t = 2$ and such that none of the D_i takes any of the values between s and t. We can use estimates of the probability of this event for all s and t to obtain estimates of $\mathbb{E}(J_{2,2})$. Thus,

$$\begin{aligned}
\mathbb{E}(J_{2,2}) &= \sum_{1 \leq s < t \leq 2^m} \binom{n}{2}\binom{n-2}{2} 2^{-4m}\left(1 - \frac{1+t-s}{2^m}\right)^{n-4} \\
&\leq n^4 2^{-4m} \sum_{1 \leq s < t \leq 2^m} \left(1 - \frac{1+t-s}{2^m}\right)^{n-4},
\end{aligned}$$

where the binomial coefficients are the number of ways of choosing two i such that $D_i = s$ and another two i such that $D_i = t$, the 2^{-4m} is the probability that the chosen D_i are s and t, and the factor $\left(1 - \frac{1+t-s}{2^m}\right)^{n-4}$ is the probability that none of the other D_i's equal s or t or a value in between them. The sum is bounded by

$$\begin{aligned}
\sum_{1 \leq s < t \leq 2^m} \left(1 - \frac{1+t-s}{2^m}\right)^{n-4} &= \sum_{k=2}^{2^m} (2^m - k + 1)\left(1 - \frac{k}{2^m}\right)^{n-4} \\
&\leq 2^m \sum_{k=2}^{\infty} \exp\left(-\frac{k(n-4)}{2^m}\right) \\
&= 2^m \frac{\exp\left(-\frac{2(n-4)}{2^m}\right)}{1 - \exp\left(-\frac{(n-4)}{2^m}\right)} \\
&= 2^m\, O\left(\frac{2^m}{n}\right) \\
&= O\left(\frac{2^{2m}}{n}\right),
\end{aligned}$$

hence

$$\mathbb{E}(J_{2,2}) = O(n^3 2^{-2m}). \tag{22}$$

■

In estimating the remaining $\mathbb{E}(J_{l_1,l_2})$ it is helpful to define random variables K_l, $l \in [1, n]$ by

$$K_l = \sum_{j=1}^{L} I[N_j = l]. \tag{23}$$

The observation that for each c_j, $j \in [1, L-1]$ such that $N_j = l$ there is one any only one $l_2 \in [1, n]$ such that $c_{j+1} = l_2$ gives the equation for $l \in [1, n]$

$$K_l = \sum_{l_2=1}^{n} J_{l,l_2} + I[N_L = l]. \tag{24}$$

Similarly, we have for $l \in [1, n]$

$$K_l = \sum_{l_1=1}^{n} J_{l_1,l} + I[N_1 = l]. \tag{25}$$

It follows from (24) that

$$J_{l,l_2} \leq K_l \tag{26}$$

for all $l \in [1, n]$ and $l_2 \in [1, n]$.

Lemma 6 *The expectation of K_l defined by (23) for $l \in [1, n]$ is*

$$\mathbb{E}(K_l) = 2^m \binom{n}{l} 2^{-lm}\left(1 - 2^{-m}\right)^{n-l} \leq 2^m \frac{(n2^{-m})^l}{l!}. \tag{27}$$

Proof In (27) 2^m is the number of possible values for the D_i, $\binom{n}{l}$ is the number of ways of choosing l D_i's such that D_i is a particular value, 2^{-lm} is the probability that the chosen D_i's actually equal that value and the factor $(1 - 2^{-m})^{n-l}$ is the probability that none of the other D_i's equal that value. ■

Lemma 7 *Let $\omega(n) > 0$ be any sequence such that $\omega(n) \to \infty$. For $m \geq \log_2 n + \omega(n)$, the expected value of $J_{2,1}$ defined by (18) equals*

$$\mathbb{E}(J_{2,1}) = \mathbb{E}(K_2) + O\left(n^3 2^{-2m}\right) + O\left(n2^{-m}\right). \tag{28}$$

and the expected value of $J_{1,2}$ equals

$$\mathbb{E}(J_{1,2}) = \mathbb{E}(K_2) + O\left(n^3 2^{-2m}\right) + O\left(n2^{-m}\right). \tag{29}$$

Note that $n^3 2^{-2m}/(n2^{-m}) = n^2 2^{-m}$, which converges to ∞ if $m/(2\log_2 n) \to 0$ and converges to 0 if $m/(2\log_2 n) \to \infty$. Thus, neither of the $O(\cdot)$ terms dominates the other in the range we are interested in. Similar pairs of $O(\cdot)$

terms will appear in other lemmas.

Proof Taking $l = 2$ gives in (24) gives

$$K_2 = J_{2,1} + J_{2,2} + \cdots + J_{2,n} + I[N_L = 2]. \tag{30}$$

Because $J_{l,l_2} \leq K_l$ (see (26)) and because of Lemma 5 and Lemma 6 we have

$$\begin{aligned} \mathbb{E}(J_{2,2}) + \mathbb{E}(J_{2,3}) + \mathbb{E}(J_{2,4}) + \cdots + \mathbb{E}(J_{2,n}) &\leq \mathbb{E}(J_{2,2}) + \mathbb{E}(K_3) + \cdots + \mathbb{E}(K_n) \\ &\leq \mathbb{E}(J_{2,2}) + \sum_{l=3}^{\infty} 2^m \frac{(n2^{-m})^l}{l!} \\ &= O(n^3 2^{-2m}). \end{aligned} \tag{31}$$

The expectation of the indicator in (30) is

$$\mathbb{P}(N_L = 2) = \sum_{k=1}^{2^m} \binom{n}{2} 2^{-2m} \left(1 - \frac{k}{2^m}\right)^{n-2}, \tag{32}$$

where $n-k+1$ represents c_L, $\binom{n}{2}$ is the number of ways of choosing two D_i's such that $D_i = c_L$, 2^{-2m} is the probability that the two chosen D_i's actually equal c_L, and $\left(1 - \frac{k}{2^m}\right)^{n-2}$ is the probability that none of the other D_i's are in the range $[c_L, n]$. We have

$$\begin{aligned} \mathbb{P}(N_L = 2) &\leq \sum_{k=1}^{2^m} \frac{n^2}{2} 2^{-2m} \exp\left(-\frac{k(n-2)}{2^m}\right) \\ &\leq \frac{n^2}{2} 2^{-2m} \frac{\exp\left(-\frac{n-2}{2^m}\right)}{1 - \exp\left(-\frac{n-2}{2^m}\right)} \\ &= O(n^2 2^{-2m} \cdot 2^m n^{-1}) \\ &= O(n 2^{-m}). \end{aligned} \tag{33}$$

Using (31) and (33) in the expected value of both sides of (30) gives (28). A similar argument using (25) produces (29). ∎

Lemma 8 *Let $\omega(n) > 0$ be any sequence such that $\omega(n) \to \infty$. For $m \geq \log_2 n + \omega(n)$, the expected value of $J_{1,1}$ defined by (18) equals*

$$\mathbb{E}(J_{1,1}) = n - 3\mathbb{E}(K_2) - 1 + O(n^3 2^{-2m}) + O(n 2^{-m}). \tag{34}$$

Letting $l = 1$ in (25) gives

$$K_1 = J_{1,1} + J_{2,1} + \cdots + J_{n,1} + I[N_1 = 1]. \tag{35}$$

If one takes expectations on both sides of (35), then by Lemma 7 and the argument showing (31) one has

$$\mathbb{E}(K_1) = \mathbb{E}(J_{1,1}) + \mathbb{E}(K_2) + \mathbb{P}(N_1 = 1) + O(n^3 2^{-2m}) + O\left(n2^{-m}\right). \tag{36}$$

By definition,

$$n = K_1 + 2K_2 + 3K_3 + \cdots + nK_n.$$

An argument like the one showing (31) implies

$$3\mathbb{E}(K_3) + \cdots + n\mathbb{E}(K_n) = O(n^3 2^{-2m}),$$

so that

$$\mathbb{E}(K_1) = n - 2\mathbb{E}(K_2) + O(n^3 2^{-2m}). \tag{37}$$

By arguing as we did at (32), we have

$$\begin{aligned}
\mathbb{P}(N_1 > 1) &= \sum_{k=2}^{n} \mathbb{P}(N_1 = k) \\
&= \sum_{k=2}^{n} \sum_{i=1}^{2^m} \binom{n}{k} 2^{-km} \left(1 - \frac{i}{2^m}\right)^{n-k} \\
&= \sum_{k=2}^{\lceil n/2 \rceil} \sum_{i=1}^{2^m} \binom{n}{k} 2^{-km} \left(1 - \frac{i}{2^m}\right)^{n-k} + \sum_{k=\lceil n/2 \rceil + 1}^{n} \sum_{i=1}^{2^m} \binom{n}{k} 2^{-km} \left(1 - \frac{i}{2^m}\right)^{n-k}.
\end{aligned}$$

The first sum is bounded by

$$\begin{aligned}
\sum_{k=2}^{\lceil n/2 \rceil} \sum_{i=1}^{2^m} \binom{n}{k} 2^{-km} \left(1 - \frac{i}{2^m}\right)^{n-k} &\le \sum_{k=2}^{\lceil n/2 \rceil} \sum_{i=1}^{2^m} \frac{n^k}{k!} 2^{-km} \exp\left(-\frac{i(n-k)}{2^m}\right) \\
&\le \sum_{k=2}^{\lceil n/2 \rceil} \sum_{i=1}^{2^m} \frac{n^k}{k!} 2^{-km} \exp\left(-\frac{in}{2^{m+1}}\right) \\
&\le \sum_{k=2}^{\lceil n/2 \rceil} \frac{n^k}{k!} 2^{-km} \frac{\exp\left(n2^{-m-1}\right)}{1 - \exp\left(n2^{-m-1}\right)} \\
&= O\left(n^2 2^{-2m} \cdot 2^m n^{-1}\right) \\
&= O(n2^{-m}).
\end{aligned}$$

The second sum is bounded by

$$\begin{aligned}
\sum_{k=\lceil n/2\rceil+1}^{n} \sum_{i=1}^{2^m} \binom{n}{k} 2^{-km}\left(1-\frac{i}{2^m}\right)^{n-k} &\leq 2^{-(\lceil n/2\rceil+1)m} \sum_{k=\lceil n/2\rceil+1}^{n} \sum_{i=1}^{2^m} \binom{n}{k} \\
&\leq 2^{-\lceil n/2\rceil m} \sum_{k=\lceil n/2\rceil+1}^{n} \binom{n}{k} \\
&\leq 2^{-\lceil n/2\rceil m+n} \\
&= o(n2^{-m}).
\end{aligned}$$

Hence, $\mathbb{P}(N_1 > 1) = O(n2^{-m})$ and

$$\mathbb{P}(N_1 = 1) = 1 - O(n2^{-m}). \tag{38}$$

The asymptotic (34) now follows from (36), (37), and (38). ■

Lemma 9 is an estimate of $\mathbb{E}(R)$.

Lemma 9 *Let $\omega(n) \geq 0$ be any sequence such that $\omega(n) \to \infty$. If $m \geq \log_2 n + \omega(n)$, then under the $P_{n,m}$ measure,*

$$\mathbb{E}(R) = \frac{n}{2} - \frac{1}{12}n^2 2^{-m} + \frac{1}{2} + O(n^3 2^{-2m}) + O(n2^{-m})$$

Proof The sum in (19) equals

$$\sum_{1\leq l_1,l_2\leq n} \mathbb{E}(J_{l_1,l_2}) \binom{l_1+l_2}{l_2}^{-1} = \mathbb{E}(J_{1,1})\frac{1}{2} + \mathbb{E}(J_{1,2})\frac{1}{3} + \mathbb{E}(J_{2,1})\frac{1}{3} + \Delta, \tag{39}$$

where, by (24), Lemma 5, and Lemma 6,

$$\begin{aligned}
|\Delta| &\leq \mathbb{E}(J_{2,2}) + \mathbb{E}(J_{1,3}) + \mathbb{E}(J_{2,3}) + \mathbb{E}(J_{3,1}) + \mathbb{E}(J_{3,2}) \\
&\quad + \sum_{l_1\geq 3}\left(\sum_{l_2\geq 3} \mathbb{E}(J_{l_1,l_2})\right) \\
&\leq \mathbb{E}(J_{2,2}) + 4\mathbb{E}(K_3) + \sum_{l_1\geq 3} \mathbb{E}(K_{l_1}) \\
&= O(n^3 2^{-2m}).
\end{aligned} \tag{40}$$

Using Lemma 5, Lemma 7, Lemma 8, and (39) in the expression for $\mathbb{E}(R)$ in Lemma 3, we have

$$\begin{aligned}\mathbb{E}(R) &= \mathbb{E}(L) - \frac{1}{2}(n - 3\mathbb{E}(K_2) - 1) - \frac{2}{3}\mathbb{E}(K_2) + O(n^3 2^{-2m}) + O(n2^{-m}) \\ &= \mathbb{E}(L) - \frac{n}{2} + \frac{5}{6}\mathbb{E}(K_2) + \frac{1}{2} + O(n^3 2^{-2m}) + O(n2^{-m}). \qquad (41)\end{aligned}$$

If we let $l = 2$ in Lemma 6, we obtain

$$\mathbb{E}(K_2) = \binom{n}{2} 2^{-m} + O(n^3 2^{-2m}). \qquad (42)$$

Using Lemma 4 and (42) in (41) gives us

$$\begin{aligned}\mathbb{E}(R) &= \frac{n}{2} - \frac{1}{6}\binom{n}{2} 2^{-m} + \frac{1}{2} + O(n^3 2^{-2m}) + O(n2^{-m}) \\ &= \frac{n}{2} - \frac{1}{12} n^2 2^{-m} + \frac{1}{2} + O(n^3 2^{-2m}) + O(n2^{-m}).\end{aligned}$$

■

Lemma 10 *Let $\omega(n) \geq 0$ be any sequence such that $\omega(n) \to \infty$. If $m \geq \log_2 n + \omega(n)$, then under the $P_{n,m}$ measure,*

$$\mathrm{Var}(R) = O(n^3 2^{-m}) + O(n).$$

Proof Rewrite the expression for R in Lemma 3 as

$$R = 1 + \sum_{j=1}^{L-1} I'_j,$$

where

$$I'_j = I\left[\inf\{i : D_i = c_{j+1}\} < \sup\{i : D_i = c_j\}\right]. \qquad (43)$$

We will get an upper bound for $\mathbb{E}(R^2)$. Recall the definition of $\mathcal{F}$ from Lemma 3. Let

$$\mathcal{J} = \{j \in [1, L-1] : N_j = 1, N_{j+1} = 1\}.$$

$$
\begin{aligned}
\mathbb{E}(R^2) &= \mathbb{E}(\mathbb{E}(R^2|\mathcal{F})) \\
&= \mathbb{E}\left(\mathbb{E}\left(1 + 2\sum_{j=1}^{L-1} I'_j + \sum_{j=1}^{L-1}\sum_{k=1}^{L-1} I'_j I'_k \middle| \mathcal{F}\right)\right) \\
&\le \mathbb{E}\left(\mathbb{E}\left(2L + \sum_{j=1}^{L-1}\sum_{k=1}^{L-1} I'_j I'_k \middle| \mathcal{F}\right)\right) \\
&\le 2n + \mathbb{E}\left(\mathbb{E}\left(3(L-1) + \sum_{\substack{1\le j,k\le L-1 \\ |k-j|\ge 2}} I'_j I'_k \middle| \mathcal{F}\right)\right) \\
&\le 5n + \mathbb{E}\left(\mathbb{E}\left(2(L-1)(L-1-|\mathcal{J}|) + \sum_{\substack{j,k\in\mathcal{J} \\ |k-j|\ge 2}} I'_j I'_k \middle| \mathcal{F}\right)\right) \\
&\le 5n + 2n\mathbb{E}(L-1-|\mathcal{J}|) + \mathbb{E}\left(\mathbb{E}\left(\sum_{\substack{j,k\in\mathcal{J} \\ |k-j|\ge 2}} I'_j I'_k \middle| \mathcal{F}\right)\right).
\end{aligned}
$$

The expectation $\mathbb{E}(L-1-|\mathcal{J}|)$ is of order

$$
\mathbb{E}(L-1-|\mathcal{J}|) \le \sum_{j=2}^{n} \mathbb{E}(K_j) = O(n^2 2^{-m}),
$$

where we have used Lemma 6. If $j,k\in\mathcal{J}$ and $|k-j|\ge 2$, then the conditional expectation $\mathbb{E}(I_j I_k|\mathcal{F})$ is just the probability that a red ball, a blue ball, a yellow ball, and a green ball are linearly arranged at random and that the the blue ball is to the left of the red ball and the green ball is to the left of the yellow ball. (See the argument proving (20) in Lemma 3.) Thus, $\mathbb{E}(I_j I_k|\mathcal{F}) = 1/4$ for $j,k\in\mathcal{J}$ and

$$
\begin{aligned}
\mathbb{E}(R^2) &\le 5n + 2n\mathbb{E}(L-1-|\mathcal{J}|) + \frac{1}{4}\mathbb{E}(|\mathcal{J}|^2) \\
&\le 5n + 2n\mathbb{E}(L-1-|\mathcal{J}|) + \frac{n^2}{4} \\
&= \frac{n^2}{4} + O(n^3 2^{-m}) + O(n). \qquad (44)
\end{aligned}
$$

From Lemma 9 and (44) we have

$$\begin{aligned}
\mathrm{Var}(R) &= \mathbb{E}(R^2) - (\mathbb{E}(R))^2 \\
&\le \frac{n^2}{4} + O(n^3 2^{-m}) + O(n) - \left(\frac{n^2}{4} + O(n^3 2^{-m}) + O(n)\right) \\
&= O(n^3 2^{-m}) + O(n).
\end{aligned}$$

The statement of the Lemma now follows from $\mathrm{Var}(R) \ge 0$. ■

Lemma 11 *If $m = \log_2 n + \omega(n)$, $\omega(n) \to \infty$, then*

$$H(P_{n,m}|U) = \frac{\alpha}{24} n^3 2^{-2m} + O\left(n^4 2^{-3m}\right) + O\left(n^2 2^{-2m}\right).$$

Proof Recall that $\alpha = \log_2 e$ and that $H(P_{n,m}|U)$ is given by (5). The relative entropy is

$$H(P_{n,m}|U) = \sum_{\pi \in S_n} P_{n,m}(\pi) \log_2(n! P_{n,m}(\pi)) = \alpha \sum_{\pi \in S_n} P_{n,m}(\pi) \ln(n! P_{n,m}(\pi)). \tag{45}$$

From (4), we have

$$\begin{aligned}
\ln(n! P_{n,m}(\pi)) &= \ln\left((2^m + n - r)_n 2^{-mn}\right) \\
&= \ln\left(\prod_{k=0}^{n-1} \frac{2^m + n - k - r}{2^m}\right) \\
&= \sum_{k=0}^{n-1} \ln\left(1 + \frac{n-k-r}{2^m}\right) \\
&= \sum_{k=0}^{n-1} \frac{n-k-r}{2^m} - \frac{1}{2}\sum_{k=0}^{n-1}\left(\frac{n-k-r}{2^m}\right)^2 + O\left(n^4 2^{-3m}\right) \\
&= \left(\frac{n(n+1)}{2} - rn\right) 2^{-m} - \frac{1}{2}\left(\frac{n^3}{3} - rn(n+1) + r^2 n\right) 2^{-2m} \\
&\quad + O\left(n^4 2^{-3m}\right) + O\left(n^2 2^{-2m}\right),
\end{aligned}$$

where the $O(\cdot)$ terms are uniform for $r = r(\pi) \in [1, 2^m \wedge n]$. Therefore,

$$\sum_{\pi \in S_n} P_{n,m}(\pi) \ln(n! P_{n,m}(\pi)) = \left(\frac{n(n+1)}{2} - \mathbb{E}(R) n\right) 2^{-m}$$

$$
\begin{aligned}
&-\frac{1}{2}\left(\frac{n^3}{3}-\mathbb{E}(R)n(n+1)+\left(\mathrm{Var}(R)+(\mathbb{E}(R))^2\right)n\right)2^{-2m}\\
&+O\left(n^4 2^{-3m}\right)+O\left(n^2 2^{-2m}\right).
\end{aligned}
$$

Lemma 9, Lemma 10, and (45) give

$$
\begin{aligned}
H(P_{n,m}|U) &= \left(\frac{1}{12}-\frac{1}{6}+\frac{1}{4}-\frac{1}{8}\right)\alpha n^3 2^{-2m}+O\left(n^4 2^{-3m}\right)+O(n^2 2^{-2m})\\
&= \frac{\alpha}{24}n^3 2^{-2m}+O\left(n^4 2^{-3m}\right)+O\left(n^2 2^{-2m}\right).
\end{aligned}
$$

■

References

[1] Aldous, D. (1983) Random walks on finite groups and rapidly mixing Markov chains. *Séminaire de Probabilités XVII. Lecture Notes in Math.*, **986**, 243 –297. Springer, New York.

[2] Bayer, D., and Diaconis, P. (1992) Trailing the dovetail shuffle to its lair. *Ann. Appl. Prob.*, **2**, 294 –313.

[3] Cover, T. M. and Thomas J. A. (1991), *Elements of Information Theory*, John Wiley & Sons Inc., New York.

[4] Diaconis, P. (1988) *Group Representations in Probability and Statistics.* IMS, Hayward, Calif.

[5] Diaconis, P. and Saloff-Coste, L. (1996) Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Prob.*, **6**, 695 – 750.

[6] Gilbert, E. N. (1955) Theory of shuffling. Bell Telephone Laboratories memorandum.

[7] Reeds, J. (1981) Theory of Shuffling. Unpublished manuscript.

[8] Robbins, H. (1955) A remark on Stirling's formula. *Amer. Math. Monthly*, **62**, 26 –29.

[9] Su, F. E. (1995) Methods for Quantifying Rates of Convergence for Random Walks on Groups, PhD Thesis, Harvard, 1995.

[10] Trefethen, L. N. and Trefethen, L. M. (1998) How many shuffles to randomize a deck of cards? Oxford University Computing Laboratory, Technical Report Number 98/09.