Electronic Journal of Probability

Electron. J. Probab. **18** (2012), no. 20, 1–20.
ISSN: 1083-6489 DOI: 10.1214/EJP.v18-1950

A lower bound for the mixing time of the random-to-random insertions shuffle*

Eliran Subag†

Abstract

The best known lower and upper bounds on the total variation mixing time for the random-to-random insertions shuffle are $\left(\frac{1}{2}-o\left(1\right)\right)n\log n$ and $\left(2+o\left(1\right)\right)n\log n$. A long standing open problem is to prove that the mixing time exhibits a cutoff. In particular, Diaconis conjectured that the cutoff occurs at $\frac{3}{4}n\log n$. Our main result is a lower bound of $t_n=\left(\frac{3}{4}-o\left(1\right)\right)n\log n$, corresponding to this conjecture.

Our method is based on analysis of the positions of cards yet-to-be-removed. We show that for large n and t_n as above, there exists $f(n)=\Theta(\sqrt{n\log n})$ such that, with high probability, under both the measure induced by the shuffle and the stationary measure, the number of cards within a certain distance from their initial position is $f(n)$ plus a lower order term. However, under the induced measure, this lower order term is strongly influenced by the number of cards yet-to-be-removed, and is of higher order than for the stationary measure.

Keywords: Mixing-time ; card shuffling ; random insertions ; cutoff phenomenon.
AMS MSC 2010: 60J10.
Submitted to EJP on April 15, 2012, final version accepted on December 29, 2012.
Supersedes arXiv:1112.5847v2.

1 Introduction

In the random-to-random insertions shuffle a card is chosen at random, removed from the deck and reinserted in a random position. Assuming the cards are numbered from 1 to n, let us identify an ordered deck with the permutation $\sigma\in S_n$ such that $\sigma\left(j\right)$ is the position of the card numbered j. The shuffling process induces a random walk Π_t, $t=0,1,\ldots$, on S_n. Let $\mathbb{P}_\sigma^n$ be the probability measure corresponding to the random walk starting from $\sigma\in S_n$.

Clearly, Π_t is an irreducible and aperiodic Markov chain. Therefore $\mathbb{P}_\sigma^n\left(\Pi_t\in\cdot\right)$ converges, as $t\to\infty$, to the stationary measure $\mathbb{U}^n$, which, since the transition matrix is symmetric, is the uniform measure on S_n. To quantify the distance from stationarity, one usually uses the total variation (TV) distance

$$d_n\left(t\right)\triangleq\max_{\sigma\in S_n}\left\|\mathbb{P}_\sigma^n\left(\Pi_t\in\cdot\right)-\mathbb{U}^n\right\|_{TV}=\left\|\mathbb{P}_{id}^n\left(\Pi_t\in\cdot\right)-\mathbb{U}^n\right\|_{TV},$$

*Research supported in part by Israel Science Foundation 853/10 and USAFOSR FA8655-11-1-3039.
†Electrical Engineering, Technion, Israel. E-mail: elirans@techunix.technion.ac.il

where equality follows since the chain is transitive. The mixing time is then defined by

$$t_{mix}^{(n)}(\varepsilon) \triangleq \min\{t : d_n(t) \le \varepsilon\}.$$

In order to study the rate of convergence to stationarity for large n, one studies how the mixing time grows as $n \to \infty$. In particular, one is interested in finding conditions on $(t_n)_{n=1}^{\infty}$ such that $\lim_{n\to\infty} d_n(t_n)$ equals 0 or 1.

The random-to-random insertions shuffle is known to have a pre-cutoff of order $O(n \log n)$. Namely, for $c_1 = \frac{1}{2}$, $c_2 = 2$:

(i) for any sequence of the form $t_n = c_1 n \log n - k_n n$ with $\lim_{n\to\infty} k_n = \infty$, $\lim_{n\to\infty} d_n(t_n) = 1$; and

(ii) for any sequence of the form $t_n = c_2 n \log n + k_n n$ with $\lim_{n\to\infty} k_n = \infty$, $\lim_{n\to\infty} d_n(t_n) = 0$.

Diaconis and Saloff-Coste [4] showed that the mixing time is of order $O(n \log n)$. Uyemura-Reyes [6] used a comparison technique from [4] to show that the upper bound above holds with $c_2 = 4$ and proved the lower bound with $c_1 = \frac{1}{2}$ by studying the longest increasing subsequence. In [7] the upper bound is improved by Saloff-Coste and Zúñiga, also by applying a comparison technique, and shown to hold with $c_2 = 2$. An alternative proof to the lower bound with $c_1 = \frac{1}{2}$ is also given there.

A long standing open problem is to prove the existence of a cutoff in TV (see [3, 2]); that is, a value c such that for any $\varepsilon > 0$:

(i) for any sequence $t_n \le (c-\varepsilon) n \log n$, $\lim_{n\to\infty} d_n(t_n) = 1$; and

(ii) for any sequence $t_n \ge (c+\varepsilon) n \log n$, $\lim_{n\to\infty} d_n(t_n) = 0$.

In particular, in [3] Diaconis conjectured that there is a cutoff at $\frac{3}{4} n \log n$.

Our main result is a lower bound on the mixing time with this rate.

Theorem 1.1. *Let $t_n = \frac{3}{4} n \log n - \frac{1}{4} n \log\log n - c_n n$ be a sequence of natural numbers with $\lim_{n\to\infty} c_n = \infty$. Then $\lim_{n\to\infty} d_n(t_n) = 1$.*

The proof is based on analysis of the distribution of the positions of cards yet-to-be-removed. Let $[n] = \{1, \dots, n\}$, and denote the set of cards that have not been chosen for removal and reinsertion up to time t by $A^t = A^{n,t}$. The following result describes the limiting distribution for a card in A^t as the size of the deck grows (in the sense below).

Recall that for a permutation $\sigma \in S^n$, the image of j under σ, $\sigma(j)$, is the position of card j in the deck with corresponding ordering. Hence, for the random walk Π_t, $\Pi_t(j)$ corresponds to the position of card j after t random-to-random insertion shuffles. Let $\Rightarrow$ denote weak convergence and $N(0,1)$ denote the standard normal distribution.

Theorem 1.2. *Let $j_n \in [n]$ and $t_n \in \mathbb{N}$ be sequences. Assume that $\gamma \triangleq \lim_{n\to\infty} \frac{j_n}{n}$ exists, and that*

$$\lim_{n\to\infty} \frac{n^2}{t_n j_n (n - j_n)} = \lim_{n\to\infty} \frac{t_n}{j_n (n - j_n)} = 0.$$

Then

$$\mathbb{P}_{id}^n\left(\frac{\Pi_{t_n}(j_n) - j_n}{\sqrt{2 t_n \lambda_n}} \in \cdot \,\middle|\, j_n \in A^{t_n}\right) \Longrightarrow \mathbb{P}(N(0,1) \in \cdot),$$

where

$$\lambda_n = \begin{cases} \frac{j_n}{n} & \text{if } \gamma = 0, \\ \frac{n - j_n}{n} & \text{if } \gamma = 1, \\ \gamma(1-\gamma) & \text{if } \gamma \in (0,1). \end{cases}$$

This can be explained by the following heuristic. Conditioned on $j \in A^t$, $\Pi_m(j) - j$, $m = 0, 1, \ldots, t$, is a Markov chain starting at 0 with increments in $\{0, \pm 1\}$. If the increments were independent and identically distributed as the first increment, Theorem 1.2 would have readily followed from Lindeberg's central limit theorem for triangular arrays ([1], Theorem 27.2). While this is not the case, if with high probability the conditional increment distributions (given in (2.2) below),

$$\mathbb{P}_{id}^n \left(\Pi_{m+1}(j) = i + k | \, \Pi_m(j) = i, \, j \in A^t \right), \quad k = 0, \pm 1,$$

are 'close enough' to be identical for all the states $\Pi_m(j)$ visits in times $m = 0, 1, \ldots, t$, one should expect a similar result. This, however, follows under mild conditions on t and j, since the conditional transition probabilities above are very close to being symmetric, and so, with high probability, $\Pi_m(j)$ remains up to time t in a small neighborhood of j, where the transition probabilities hardly vary.

To prove the lower bound on the TV distance of $\mathbb{P}_{id}^n(\Pi_{t_n} \in \cdot)$ and $\mathbb{U}^n$, we study the size of sets of the form

$$\triangle_\alpha(\sigma) \triangleq \left\{ j \in D^n : \, |\sigma(j) - j| \le \alpha \sqrt{n \log n} \right\}, \; \sigma \in S_n,$$

where $D^n = [n] \cap [n(1-\varepsilon)/2, n(1+\varepsilon)/2]$, for fixed $\varepsilon \in (0,1)$ and a parameter $\alpha > 0$. We shall see that for t_n as in Theorem 1.1, as long as $\limsup c_n / \log n < 1/4$,

$$|\triangle_\alpha| / \left(2\varepsilon\alpha\sqrt{n \log n} \right) \Longrightarrow 1,$$

under both measures. However, the deviation $|\triangle_\alpha| - 2\varepsilon\alpha\sqrt{n\log n}$, which for $\mathbb{P}_{id}^n(\Pi_{t_n} \in \cdot)$ is strongly influenced by $|\triangle_\alpha(\Pi_{t_n}) \cap A^{t_n}|$, i.e. by the cards yet-to-be-removed, is of different order for the two measures.

In Section 2 we prove Theorem 1.2 and other related results. We analyze the distribution of $|\triangle_\alpha(\sigma)|$ under $\mathbb{U}^n$, and the distributions of $|\triangle_\alpha(\Pi_{t_n}) \cap A^{t_n}|$ and $|\triangle_\alpha(\Pi_{t_n}) \setminus A^{t_n}|$ under $\mathbb{P}_{id}^n$ in Section 3. The proof of Theorem 1.1, given in Section 4, then easily follows. Lastly, in Section 5 we prove a result which is used in the previous sections.

2 The Position of Cards Yet-to-be-Removed

In this section we prove Theorem 1.2 and other related results.

The increment distribution of Π_t is given by

$$\mu(\tau) = \begin{cases} 1/n & \text{if } \tau = id, \\ 2/n^2 & \text{if } \tau = (i,j) \text{ with } 1 \le i,j \le n \text{ and } |i-j| = 1, \\ 1/n^2 & \text{if } \tau = c_{i,j} \text{ with } 1 \le i,j \le n \text{ and } |i-j| > 1, \\ 0 & \text{otherwise,} \end{cases} \tag{2.1}$$

where $c_{i,j}$ is the cycle corresponding to removing the card in position i and reinserting it in position j, that is

$$c_{i,j} = \begin{cases} id & \text{if } i = j, \\ (j, j-1, \ldots, i+1, i) & \text{if } i < j, \\ (j, j+1, \ldots, i-1, i) & \text{if } i > j. \end{cases}$$

Let $2 \le n \in \mathbb{N}$ and $j \in [n]$. Under conditioning on $\{j \in A^t\}$, $\Pi_m(j)$, $m = 0, \ldots, t$ is a time homogeneous Markov chain with transition probabilities

$$p^{\Pi(j)}_{i,i+k} \triangleq \mathbb{P}^n_{id}\left(\Pi_{m+1}(j) = i+k \middle| \Pi_m(j) = i,\, j \in A^t\right)$$
$$= \begin{cases} \frac{i(n-i)}{n(n-1)} & \text{if } k = +1, \\ \frac{(i-1)(n-i+1)}{n(n-1)} & \text{if } k = -1, \\ \frac{(i-1)^2+(n-i)^2}{n(n-1)} & \text{if } k = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

One of the difficulties in analyzing the chain is the fact that the transition probabilities $p^{\Pi(j)}_{i,i+k}$ are inhomogeneous in i. To overcome this, we consider a modification of the process for which inhomogeneity is 'truncated' by setting transition probabilities far from the initial state to be identical to these in the initial state. As we shall see, a bound on the TV distance of the marginal distributions of the modified and original processes is easily established.

For $j \in [n]$ and $M > 0$, let $\overline{j \pm M} \triangleq [n] \cap [j-M, j+M]$, and let $\zeta_m = \zeta^{n,j,M}_m$, $m = 0, 1, \ldots$, be a Markov process starting at $\zeta_0 = j$ with transition probabilities $p^{\zeta,j,M}_{i,i+k} \triangleq \mathbb{P}\left(\zeta_{m+1} = i+k \middle| \zeta_m = i\right)$ such that

$$\forall i \in \overline{j \pm M} : p^{\zeta,j,M}_{i,i+k} = p^{\Pi(j)}_{i,i+k},$$
$$\forall i \in \mathbb{Z} \setminus \overline{j \pm M} : p^{\zeta,j,M}_{i,i+k} = p^{\Pi(j)}_{j,j+k}.$$

Clearly, for any sequence $(k_m)^t_{m=0} \in \mathbb{Z}^{t+1}$ if $\max_{0 \le m \le t} |k_m - j| \le M$ then

$$\mathbb{P}\left((\zeta_m)^t_{m=0} = (k_m)^t_{m=0}\right) = \mathbb{P}^n_{id}\left((\Pi_m(j))^t_{m=0} = (k_m)^t_{m=0} \middle| j \in A^t\right). \tag{2.3}$$

Therefore, by taking complements, for any $u \le M$

$$\mathbb{P}^n_{id}\left(\max_{0 \le m \le t} |\Pi_m(j) - j| > u \middle| j \in A^t\right) = \mathbb{P}\left(\max_{0 \le m \le t} |\zeta_m - j| > u\right). \tag{2.4}$$

Moreover, (2.3) implies that for any $B \subset \mathbb{Z}^{t+1}$

$$\mathbb{P}^n_{id}\left((\Pi_m(j))^t_{m=0} \in B \middle| j \in A^t\right) - \mathbb{P}\left((\zeta_m)^t_{m=0} \in B\right)$$
$$= \mathbb{P}^n_{id}\left((\Pi_m(j))^t_{m=0} \in B, \max_{0 \le m \le t} |\Pi_m(j) - j| > M \middle| j \in A^t\right)$$
$$- \mathbb{P}\left((\zeta_m)^t_{m=0} \in B, \max_{0 \le m \le t} |\zeta_m - j| > M\right).$$

Since both terms in the last equality are bounded from above by the equal expressions of (2.4) (and from below by zero), it follows that

$$\begin{aligned} &\left\|\mathbb{P}^n_{id}\left((\Pi_m(j))^t_{m=0} \in \cdot \middle| j \in A^t\right) - \mathbb{P}\left((\zeta_m)^t_{m=0} \in \cdot\right)\right\|_{TV} \\ &\quad \le \mathbb{P}\left(\max_{0 \le m \le t} |\zeta_m - j| > M\right). \end{aligned} \tag{2.5}$$

A simple computation shows that $\left|p^{\Pi(j)}_{i,i+1} - p^{\Pi(j)}_{i,i-1}\right|$ is bounded by $\frac{1}{n}$ for any i. On the other hand, $p^{\Pi(j)}_{i,i\pm 1}$ is roughly equal to $i(n-i)/n^2$. Thus if j is large enough and M, and thus $\left|\overline{j \pm M}\right|$, is small compared to j, we can think of $\zeta^{n,j,M}_m$ as a perturbation of a random walk with a very small bias. In order to make this precise, we decompose $\zeta^{n,j,M}_m$ as a sum of a random walk determined by the increment distribution in state j and two additional random processes related to the 'defects' in symmetry and homogeneity in state.

Consider the vector-valued Markov process

$$(S_m, X_m, Y_m) = \left(S_m^{n,j,M}, X_m^{n,j,M}, Y_m^{n,j,M}\right)$$

starting at $(S_0, X_0, Y_0) = (0,0,0)$ with transition probabilities as follows. For each $k \in \mathbb{Z}$ define

$$\begin{aligned} q_k &= \min\left\{p_{k,k+1}^{\zeta,j,M}, p_{k,k-1}^{\zeta,j,M}\right\}, \\ r_k &= \max\left\{p_{k,k+1}^{\zeta,j,M}, p_{k,k-1}^{\zeta,j,M}\right\}. \end{aligned} \tag{2.6}$$

For a state (i_1, i_2, i_3) set $i = i_1 + i_2 + i_3$ and define

$$\begin{aligned} w_i &= \arg\max_{k=\pm1}\left(p_{j+i,j+i+k}^{\zeta,j,M}\right), \\ z_i &= \operatorname{sgn}\left(q_j - q_{j+i}\right), \end{aligned}$$

where sgn is the sign function (the definition of sgn at zero will not matter to us). Define the transition probabilities by

$$\begin{aligned} &\mathbb{P}\left(\left(S_{m+1}, X_{m+1}, Y_{m+1}\right) = (i_1+k_1, i_2+k_2, i_3+k_3) \middle| \left(S_m, X_m, Y_m\right) = (i_1, i_2, i_3)\right) \\ &= \begin{cases} \min\{q_{j+i}, q_j\} & \text{if } (k_1,k_2,k_3) = (+1,0,0), \\ \min\{q_{j+i}, q_j\} & \text{if } (k_1,k_2,k_3) = (-1,0,0), \\ |q_j - q_{j+i}| & \text{if } (k_1,k_2,k_3) = \left(+\frac{1+z_i}{2}, -1, 0\right), \\ |q_j - q_{j+i}| & \text{if } (k_1,k_2,k_3) = \left(-\frac{1+z_i}{2}, +1, 0\right), \\ r_{j+i} - q_{j+i}, & \text{if } (k_1,k_2,k_3) = (0,0,w_i), \\ c_i, & \text{if } (k_1,k_2,k_3) = (0,0,0). \end{cases} \end{aligned}$$

where c_i is chosen such that the sum of probabilities is 1.

It is easy to verify that $(S_m + X_m + Y_m)_{m=0}^{\infty}$ is a Markov process with transition probabilities identical to those of $(\zeta_m - j)_{m=0}^{\infty}$. Therefore the two processes have the same law. It is also easy to check that S_n is a random walk with increment distribution

$$\mu(+1) = \mu(-1) = q_j\ ,\ \mu(0) = 1 - 2q_j.$$

In order to study X_m and Y_m we need the following proposition.

Proposition 2.1. *Let $\{A_m\}_{m=0}^{\infty}$ and $\{B_m\}_{m=0}^{\infty}$ be integer-valued random processes starting at the same point $A_0 = B_0$. Suppose that there exist $p_{ik}^A \in [0,1]$ such that for any $m \geq 0$ and $k, i, i_0, \ldots, i_{m-1} \in \mathbb{Z}$ (such that the conditional probabilities are defined)*

$$\begin{aligned} p_{ik}^A &= \mathbb{P}\left(A_{m+1} = k \middle| A_{m+1} \neq i, A_m = i\right) \\ &= \mathbb{P}\left(A_{m+1} = k \middle| A_{m+1} \neq i, A_m = i, A_{m-1} = i_{m-1}, \ldots, A_0 = i_0\right) \end{aligned}$$

and similarly for B_m with p_{ik}^B. Assume that for any $i, k \in \mathbb{Z}$, $p_{ik}^A = p_{ik}^B$. Finally, suppose that for any $m \geq 0$ and $k, i, i_0, \ldots, i_{m-1}, j_0, \ldots, j_{m-1} \in \mathbb{Z}$, (whenever defined)

$$\begin{aligned} &\mathbb{P}\left(A_{m+1} \neq i \middle| A_m = i, A_{m-1} = i_{m-1}, \ldots, A_0 = i_0\right) \\ &\quad \geq \mathbb{P}\left(B_{m+1} \neq i \middle| B_m = i, B_{m-1} = j_{m-1}, \ldots, B_0 = j_0\right). \end{aligned}$$

Then for any $t \in \mathbb{N}$ and $\delta > 0$

$$\mathbb{P}\left(\max_{0 \leq m \leq t} |A_m| \geq \delta\right) \geq \mathbb{P}\left(\max_{0 \leq m \leq t} |B_m| \geq \delta\right).$$

Proof. The processes $\{A_m\}$ and $\{B_m\}$ can be coupled so that they jump from a given state to a new state according to the same order of states, say according to the order $\{k_m\}_{m=0}^{\infty}$, and such that the amount of time that $\{B_m\}$ spends in any given state k_m before jumping to state k_{m+1} is at least as much as $\{A_m\}$ spends there. The proposition follows easily from this. $\square$

The only nonzero increments of X_m are ± 1. Note that

$$\begin{aligned}
&\mathbb{P}\left(X_{m+1}=i_m+1\middle|X_{m+1}\neq i_m,\{X_p\}_{p=0}^m=\{i_p\}_{p=0}^m\right)\\
&=\sum_{(k_1,k_2)\in\mathbb{Z}^2}\mathbb{P}\left(X_{m+1}=i_m+1\middle|X_{m+1}\neq i_m,\{X_p\}_{p=0}^m=\{i_p\}_{p=0}^m,\dots\right.\\
&\left.S_m=k_1,Y_m=k_2\right)\mathbb{P}\left(S_m=k_1,Y_m=k_2\middle|X_{m+1}\neq i_m,\{X_p\}_{p=0}^m=\{i_p\}_{p=0}^m\right)\\
&=\frac{1}{2}.
\end{aligned}$$

The last equality follows from Markov property of (S_m, X_m, Y_m). The same, of course, holds for the negative increment. In addition, again by Markov property,

$$\begin{aligned}
&\mathbb{P}\left(X_{m+1}\neq i_m\middle|\{X_p\}_{p=0}^m=\{i_p\}_{p=0}^m\right)\\
&=\sum_{(k_1,k_2)\in\mathbb{Z}^2}\mathbb{P}\left(X_{m+1}\neq i_m\middle|\{X_p\}_{p=0}^m=\{i_p\}_{p=0}^m,S_m=k_1,Y_m=k_2\right)\times\\
&\mathbb{P}\left(S_m=k_1,Y_m=k_2\middle|\{X_p\}_{p=0}^m=\{i_p\}_{p=0}^m\right)\\
&\leq\max_{k_1,k_2}\mathbb{P}\left(X_{m+1}\neq i_m\middle|X_m=i_m,S_m=k_1,Y_m=k_2\right)\leq 2\max_{i\in\overline{j\pm M}}|q_i-q_j|\\
&\leq 2M\max_{x\in[1,n]}\left(\max\left\{\left|\frac{d}{dx}\frac{x(n-x)}{n(n-1)}\right|,\left|\frac{d}{dx}\frac{(x-1)(n-x+1)}{n(n-1)}\right|\right\}\right)\\
&\leq\frac{2M}{n-1},
\end{aligned}$$

where the maximum in the first inequality is over all k_1, k_2 such that the conditional probability is defined.

Thus, according to Proposition 2.1, for $\delta > 0$,

$$\mathbb{P}\left(\max_{0\leq m\leq t}\left|X_m^{n,j,M}\right|\geq\delta\right)\leq\mathbb{P}\left(\max_{0\leq m\leq t}\left|W_m^{n,M}\right|\geq\delta\right),\tag{2.7}$$

where $W_m = W_m^{n,M}$ is a random walk starting at 0 with increment distribution

$$\nu(+1)=\nu(-1)=\frac{M}{n-1}\ ,\ \nu(0)=1-2\frac{M}{n-1}.$$

Similarly, for the process $\widetilde{Y}_t=\sum_{m=1}^t|Y_m-Y_{m-1}|$, whose increments are 0 and 1, we have

$$\begin{aligned}
&\mathbb{P}\left(\widetilde{Y}_{m+1}=i_m+1\middle|\left\{\widetilde{Y}_p\right\}_{p=0}^m=\{i_p\}_{p=0}^m\right)\\
&\leq\max_{k_1,k_2,k_3}\mathbb{P}\left(Y_{m+1}\neq k_1|\,Y_m=k_1,S_m=k_2,X_m=k_3\right)\\
&\leq\max_{i\in\mathbb{Z}}\left(r_{j+i}-q_{j+i}\right)=\max_{i\in\overline{j\pm M}}\left|\frac{i(n-i)}{n(n-1)}-\frac{(i-1)(n-i+1)}{n(n-1)}\right|\\
&=\max_{i\in\overline{j\pm M}}\left|\frac{n-2i+1}{n(n-1)}\right|\leq\frac{1}{n}.
\end{aligned}$$

Therefore, for $\delta > 0$,

$$\mathbb{P}\left(\max_{0\le m\le t}\left|Y_m^{n,j,M}\right| \ge \delta\right) \le \mathbb{P}\left(\widetilde{Y}_t \ge \delta\right) \le \mathbb{P}\left(N_t^n \ge \delta\right), \tag{2.8}$$

where $N_t = N_t^n \sim \mathrm{Bin}\left(t, \frac{1}{n}\right)$.

Since the increment distributions of W_m and S_m are symmetric, the classical Lévy inequality ([5], Theorem 2.2) yields, for any $\delta > 0$,

$$\mathbb{P}\left(\max_{0\le m\le t}|W_m| \ge \delta\right) \le 4\mathbb{P}\left(W_t \ge \delta\right). \tag{2.9}$$

and

$$\mathbb{P}\left(\max_{0\le m\le t}|S_m| \ge \delta\right) \le 4\mathbb{P}\left(S_t \ge \delta\right). \tag{2.10}$$

Having established the connections between the different processes, we are now ready to prove Theorem 1.2.

Proof of Theorem 1.2. The case where $\gamma = 1$ follows by symmetry from the case with $\gamma = 0$. Assume $\gamma \in [0, 1)$. In this case, the hypothesis in the theorem are equivalent to

$$\lim_{n\to\infty}\frac{n}{t_n j_n} = \lim_{n\to\infty}\frac{t_n}{n j_n} = 0.$$

Let $n \in \mathbb{N}$, $j \in [n]$ and $M > 0$. Based on (2.7)-(2.9) and a union bound, for $u \in \mathbb{R}$, $\delta > 0$, we have

$$\begin{aligned}
&\mathbb{P}\left(\zeta_t - j \ge u\right)\\
&\quad\le \mathbb{P}\left(S_t \ge u - \delta\right) + \mathbb{P}\left(\max_{0\le m\le t}|X_m| \ge \frac{\delta}{2}\right) + \mathbb{P}\left(\max_{0\le m\le t}|Y_m| \ge \frac{\delta}{2}\right)\\
&\quad\le \mathbb{P}\left(S_t \ge u - \delta\right) + 4\mathbb{P}\left(W_t \ge \frac{\delta}{2}\right) + \mathbb{P}\left(N_t \ge \frac{\delta}{2}\right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\mathbb{P}\left(\zeta_t - j \ge u\right)\\
&\quad\ge \mathbb{P}\left(S_t \ge u + \delta\right) - \mathbb{P}\left(\max_{0\le m\le t}|X_m| \ge \frac{\delta}{2}\right) - \mathbb{P}\left(\max_{0\le m\le t}|Y_m| \ge \frac{\delta}{2}\right)\\
&\quad\ge \mathbb{P}\left(S_t \ge u + \delta\right) - 4\mathbb{P}\left(W_t \ge \frac{\delta}{2}\right) - \mathbb{P}\left(N_t \ge \frac{\delta}{2}\right).
\end{aligned}$$

Assume $\frac{\delta}{2} - \frac{t}{n} > 0$. By computing moments and applying the Berry-Esseen theorem to approximate the tail probability function of S_t, and applying Chebyshev's inequality to bound the tail probability functions of W_t and N_t, we arrive at

$$\mathbb{P}\left(\zeta_t^{n,j,M} - j \ge u\right) \le \Psi\left(\frac{u-\delta}{\sqrt{2tq_j}}\right) + \frac{C}{\sqrt{2tq_j}} + \frac{32Mt}{\delta^2(n-1)} + \frac{\frac{t}{n}\frac{n-1}{n}}{\left(\frac{\delta}{2} - \frac{t}{n}\right)^2} \tag{2.11}$$

and

$$\mathbb{P}\left(\zeta_t^{n,j,M} - j \ge u\right) \ge \Psi\left(\frac{u+\delta}{\sqrt{2tq_j}}\right) - \frac{C}{\sqrt{2tq_j}} - \frac{32Mt}{\delta^2(n-1)} - \frac{\frac{t}{n}\frac{n-1}{n}}{\left(\frac{\delta}{2} - \frac{t}{n}\right)^2}, \tag{2.12}$$

where q_j is defined in (2.6), C is the constant from the Berry-Esseen theorem and Ψ is the tail probability function of a standard normal variable.

For two sequences of positive numbers v_n, v'_n let us write $v_n \ll v'_n$ if and only if $\lim_{n\to\infty} v_n/v'_n = 0$. By assumption, $\sqrt{\frac{t_n j_n}{n}} \ll j_n$, therefore we can choose a sequence M_n such that $\frac{t_n}{n}, 1 \ll \sqrt{\frac{t_n j_n}{n}} \ll M_n \ll j_n$. Similarly, since $M_n \ll j_n$ we can set δ_n with $\sqrt{\frac{t_n M_n}{n}} \ll \delta_n \ll \sqrt{\frac{t_n j_n}{n}}$, which also implies that $\sqrt{\frac{t_n}{n}}, \frac{t_n}{n} \ll \delta_n$.

Now, let $x \in \mathbb{R}$ and set $u_n = x\sqrt{2t_n\lambda_n}$. Let us consider the inequalities derived from (2.11) and (2.12) by replacing each of the parameters by a corresponding element from the sequences above. Based on the relations established for the sequences and the assumptions on t_n and j_n it can be easily verified that, upon letting $n \to \infty$, all terms but those involving Ψ go to zero. Relying, in addition, on the fact that Ψ is continuous, it can be easily verified that

$$\lim_{n\to\infty} \Psi\left(\frac{u_n \pm \delta_n}{\sqrt{2t_n q_{j_n}}}\right) = \Psi(x).$$

Hence we conclude that

$$\lim_{n\to\infty} \mathbb{P}\left(\zeta_{t_n}^{n,j_n,M_n} - j_n \geq u_n\right) = \Psi(x). \tag{2.13}$$

Based on (2.7)-(2.10),

$$\begin{aligned}
\mathbb{P}\left(\max_{0\le m\le t}\left|\zeta_m^{n,j,M} - j\right| \geq M\right) &\leq \mathbb{P}\left(\max_{0\le m\le t}|S_m| \geq M-\delta\right) \\
&\quad + \mathbb{P}\left(\max_{0\le m\le t}|Y_m| \geq \frac{\delta}{2}\right) + \mathbb{P}\left(\max_{0\le m\le t}|X_m| \geq \frac{\delta}{2}\right) \\
&\leq 4\mathbb{P}\left(S_t \geq M-\delta\right) + 4\mathbb{P}\left(W_t \geq \frac{\delta}{2}\right) + \mathbb{P}\left(N_t \geq \frac{\delta}{2}\right) \\
&\leq \frac{8tq_j}{(M-\delta)^2} + \frac{32Mt}{\delta^2(n-1)} + \frac{\frac{t}{n}\frac{n-1}{n}}{\left(\frac{\delta}{2} - \frac{t}{n}\right)^2},
\end{aligned} \tag{2.14}$$

where the last inequality follows from Chebyshev's inequality.

As before, consider the inequality derived from (2.14) by replacing each of the parameters by a corresponding element from the sequences above. The middle and right-hand side summands of (2.14) were already shown to go to zero as $n \to \infty$. Since $\delta_n \ll \sqrt{\frac{t_n j_n}{n}} \ll M_n$ the additional term also goes to zero. Combined with (2.5) and (2.13) this gives

$$\lim_{n\to\infty} \mathbb{P}_{id}^n\left(\Pi_{t_n}(j_n) - j_n \geq u_n \middle| j_n \in A^{t_n}\right) = \Psi(x),$$

which completes the proof. □

In Theorem 1.2 for each n only a single card j_n of the deck of size n is involved. The following gives a uniform bound (in initial position and in time) for the tail distributions of the difference from the initial position.

Theorem 2.2. *Let $\alpha > 0$ and let t_n be a sequence of natural numbers such that $\lim_{n\to\infty} t_n = \lim_{n\to\infty} n^2/t_n = \infty$. Then*

$$\limsup_{n\to\infty} \max_{j\in[n]} \mathbb{P}_{id}^n\left(\max_{0\le m\le t_n}|\Pi_m(j) - j| > \alpha\sqrt{\frac{t_n}{2}} \middle| j \in A^{t_n}\right) \leq 4\Psi(\alpha).$$

Proof. Set $u_n = \alpha\sqrt{\frac{t_n}{2}}$ and $j_n = \left\lfloor \frac{n}{2} \right\rfloor$. Let $M_n \geq u_n$ and $\delta_n > 0$ be sequences to be determined below. From (2.4) it follows that

$$\begin{aligned}
&\max_{j\in[n]} \mathbb{P}^n_{id}\left(\max_{0\leq m\leq t_n}|\Pi_m(j) - j| > u_n \middle| j \in A^{t_n}\right) \\
&\quad = \max_{j\in[n]} \mathbb{P}\left(\max_{0\leq m\leq t_n}\left|\zeta^{n,j,M_n}_m - j\right| > u_n\right) \\
&\quad \leq \max_{j\in[n]}\left\{\mathbb{P}\left(\max_{0\leq m\leq t_n}\left|S^{n,j,M_n}_m\right| \geq u_n - \delta_n\right) + \right. \\
&\qquad \left. \mathbb{P}\left(\max_{0\leq m\leq t_n}\left|W^{n,M_n}_m\right| \geq \frac{\delta_n}{2}\right) + \mathbb{P}\left(N^n_{t_n} \geq \frac{\delta_n}{2}\right)\right\}.
\end{aligned}$$

It is easy to check that for the random walks S^{n,j,M_n}_m, $j \in [n]$, the probabilities of the nonzero increments, ± 1, are maximal when $j = j_n$. Therefore according to Proposition 2.1 and equations (2.9) and (2.10),

$$\begin{aligned}
&\max_{j\in[n]} \mathbb{P}^n_{id}\left(\max_{0\leq m\leq t_n}|\Pi_m(j) - j| > u_n \middle| j \in A^{t_n}\right) \\
&\quad \leq \mathbb{P}\left(\max_{0\leq m\leq t_n}\left|S^{n,j_n,M_n}_m\right| \geq u_n - \delta_n\right) + \\
&\qquad \mathbb{P}\left(\max_{0\leq m\leq t_n}\left|W^{n,M_n}_m\right| \geq \frac{\delta_n}{2}\right) + \mathbb{P}\left(N^n_{t_n} \geq \frac{\delta_n}{2}\right) \\
&\quad \leq 4\left\{\mathbb{P}\left(S^{n,j_n,M_n}_{t_n} \geq u_n - \delta_n\right) + \mathbb{P}\left(W^{n,M_n}_{t_n} \geq \frac{\delta_n}{2}\right) + \mathbb{P}\left(N^n_{t_n} \geq \frac{\delta_n}{2}\right)\right\}.
\end{aligned}$$

Finally, note that j_n and t_n meet the conditions of Theorem 1.2 with $\gamma = \frac{1}{2}$. Therefore, defining M_n and δ_n as in the proof of the theorem (which also implies $M_n \geq u_n$) and following the same arguments therein, as $n \to \infty$ the expression in the last line of the inequality above converges to

$$4\Psi\left(\lim_{n\to\infty}\frac{u_n - \delta_n}{\sqrt{t_n/2}}\right) = 4\Psi(\alpha).$$

This completes the proof. □

3 Cards of Distance $O(\sqrt{n\log n})$ from their Initial Position

The results of Section 2 show that the position of a card that has not been removed is fairly concentrated around the initial position. This, of course, is a rare event for each card under the uniform measure $\mathbb{U}^n$. In this section we shall develop the tools to exploit this to derive a lower bound for the TV distance between $\mathbb{U}^n$ and $\mathbb{P}^n_{id}(\Pi_t \in \cdot)$ whenever sufficiently many (in expectation) of the cards have not been removed. Here, 'sufficiently many' means, of course, that t is not too large.

More precisely, we shall consider the size of sets of the form

$$\triangle_\alpha(\sigma) \triangleq \left\{j \in D^n : |\sigma(j) - j| \leq \alpha\sqrt{n\log n}\right\}, \ \sigma \in S_n,$$

where $D^n = [n] \cap [n(1-\varepsilon)/2, n(1+\varepsilon)/2]$, and $\varepsilon \in (0,1)$ is arbitrary and will be fixed throughout the proofs. Under $\mathbb{U}^n$, for $i \neq j$, the events $\{i \in \triangle_\alpha\}$ and $\{j \in \triangle_\alpha\}$ are 'almost' independent, as $n \to \infty$. Therefore one should expect $|\triangle_\alpha| - \mathbb{E}^{\mathbb{U}^n}\{|\triangle_\alpha|\}$ to be of order $\left(\mathbb{E}^{\mathbb{U}^n}\{|\triangle_\alpha|\}\right)^{1/2}$. Under $\mathbb{P}^n_{id}$, if $|A^{t_n}|$ is relatively small, it seems natural that the positions of the cards that have been removed are distributed approximately as they would under $\mathbb{U}^n$. Thus, $|\triangle_\alpha(\Pi_{t_n}) \setminus A^{t_n}|$ under $\mathbb{P}^n_{id}$ should be distributed roughly as $|\triangle_\alpha|$ is under $\mathbb{U}^n$. By this logic, we need to choose t_n so that $|\triangle_\alpha(\Pi_{t_n}) \cap A^{t_n}|$ is larger than $\left(\mathbb{E}^{\mathbb{U}^n}\left\{|\triangle_\alpha|\right\}\right)^{1/2}$ with high probability, which leads us to set t_n to be as in Theorem 1.1.

The three subsections below are devoted to separately study the distribution of $|\triangle_\alpha|$ under $\mathbb{U}^n$ and the distributions of $|\triangle_\alpha(\Pi_{t_n}) \cap A^{t_n}|$ and $|\triangle_\alpha(\Pi_{t_n}) \setminus A^{t_n}|$ under $\mathbb{P}^n_{id}$.

3.1 The distribution of $|\triangle_\alpha|$ under $\mathbb{U}^n$

In this case, the first and second moments of $|\triangle_\alpha|$ can be easily computed in order to apply Chebyshev's inequality. In what follows, let R_j denote the event $\{j \in \triangle_\alpha\}$.

Lemma 3.1. *For any α, $k > 0$,*

$$\limsup_{n\to\infty} \mathbb{U}^n \left(\left| |\triangle_\alpha(\sigma)| - 2\varepsilon\alpha\sqrt{n\log n} \right| \geq k\sqrt{2\varepsilon\alpha}\,(n\log n)^{\frac{1}{4}} \right) \leq \frac{1}{k^2}.$$

Proof. Suppose n is large enough so that $n(1-\varepsilon)/2 \geq \alpha\sqrt{n\log n}$. Then

$$\begin{aligned}\mathbb{E}^{\mathbb{U}^n}\left\{|\triangle_\alpha(\sigma)|\right\} &= \sum_{j\in D^n} \mathbb{U}^n(R_j) = |D^n| \frac{1+2\left\lfloor \alpha\sqrt{n\log n}\right\rfloor}{n} \\ &= 2\varepsilon\alpha\sqrt{n\log n} + O(1).\end{aligned}$$

The second moment satisfies the bound

$$\begin{aligned}\mathbb{E}^{\mathbb{U}^n}\left\{|\triangle_\alpha(\sigma)|^2\right\} &= \sum_{j\in D^n}\mathbb{U}^n(R_j) + \sum_{i,j\in D^n: i\neq j}\mathbb{U}^n(R_i\cap R_j) \\ &\leq \mathbb{E}^{\mathbb{U}^n}\left\{|\triangle_\alpha(\sigma)|\right\} + |D^n|^2 \frac{\left(1+2\left\lfloor\alpha\sqrt{n\log n}\right\rfloor\right)^2}{n(n-1)} \\ &= \mathbb{E}^{\mathbb{U}^n}\left\{|\triangle_\alpha(\sigma)|\right\} + \frac{n}{n-1}\left(\mathbb{E}^{\mathbb{U}^n}\left\{|\triangle_\alpha(\sigma)|\right\}\right)^2,\end{aligned}$$

which implies

$$\mathbf{Var}^{\mathbb{U}^n}\left\{|\triangle_\alpha(\sigma)|\right\} \leq 2\varepsilon\alpha\sqrt{n\log n} + 4\varepsilon^2\alpha^2\log n + O(1).$$

Applying Chebyshev's inequality and letting $n\to\infty$ yields the required result. □

3.2 The distribution of $|\triangle_\alpha(\Pi_{t_n}) \cap A^{t_n}|$ under $\mathbb{P}^n_{id}$

We begin with the following lemma which, when combined with Theorem 2.2, yields a bound on the probability that $|\triangle_\alpha(\Pi_{t_n}) \cap A^{t_n}|$ is less than a fraction of its expectation. This bound is the content of Lemma 3.3 which takes up the rest of the subsection.

Lemma 3.2. *Let $n, t \in \mathbb{N}$, let $B \subset [n]$ be a random set, and let $D \subset [n]$ be a deterministic set. Suppose that for some $c > 0$*

$$\min_{j\in D} \mathbb{P}^n_{id}\left(j\in B \middle| j\in A^t\right) \geq c.$$

Then, denoting $K = \mathbb{E}^n_{id}|D\cap A^t|$, for any $r\in(0,1)$,

$$\mathbb{P}^n_{id}\left(\left|B\cap D\cap A^t\right| \leq r\cdot\mathbb{E}^n_{id}\left\{\left|B\cap D\cap A^t\right|\right\}\right) \leq \frac{K + \left(1-c^2\right)K^2}{(1-r)^2 c^2K^2}.$$

Proof. By our assumption,

$$\mathbb{E}_{id}^{n}\left\{\left|B\cap D\cap A^{t}\right|\right\}=\sum_{j\in D}\mathbb{P}_{id}^{n}\left(j\in B\middle|\,j\in A^{t}\right)\mathbb{P}_{id}^{n}\left(j\in A^{t}\right)\geq cK.$$

Write

$$\begin{aligned}\mathbb{E}_{id}^{n}\left\{\left|B\cap D\cap A^{t}\right|^{2}\right\}&\leq\mathbb{E}_{id}^{n}\left\{\left|D\cap A^{t}\right|^{2}\right\}\\&=\sum_{j\in D}\mathbb{P}_{id}^{n}\left(j\in A^{t}\right)+\sum_{i,j\in D:i\neq j}\mathbb{P}_{id}^{n}\left(i,j\in A^{t}\right)\\&=K+|D|\left(|D|-1\right)\left(\frac{n-2}{n}\right)^{t}.\end{aligned}$$

Since $K=|D|((n-1)/n)^{t}$ it follows that

$$\mathbb{E}_{id}^{n}\left\{\left|B\cap D\cap A^{t}\right|^{2}\right\}\leq K+K^{2},$$

therefore

$$\mathbf{Var}_{id}^{n}\left\{\left|B\cap D\cap A^{t}\right|\right\}\leq K+\left(1-c^{2}\right)K^{2}.$$

Applying Chebyshev's inequality completes the proof. □

Now, let $R_{j,t}$ and $R_{j,t}^{A^c}$ denote the events $\{j\in\triangle_{\alpha}\left(\Pi_{t}\right)\}$ and $\{j\in\triangle_{\alpha}\left(\Pi_{t}\right)\}\cap\{j\notin A^{t}\}$, respectively. Let $p_{t,n}\triangleq\mathbb{P}_{id}^{n}\left(j\in A^{t}\right)$ (which is, of course, independent of j).

Lemma 3.3. *Let $v\left(\alpha\right)=1-4\Psi\left(\alpha\sqrt{\frac{8}{3}}\right)$. Let t_n be a sequence of natural numbers such that $t_n\leq\frac{3}{4}n\log n$ and suppose α satisfies $v\left(\alpha\right)>0$. Then, for any $r\in(0,1)$,*

$$\begin{aligned}&\limsup_{n\to\infty}\mathbb{P}_{id}^{n}\left(\left|\triangle_{\alpha}\left(\Pi_{t_n}\right)\cap A^{t_n}\right|\leq rv\left(\alpha\right)\varepsilon np_{t_n,n}\right)\\&\qquad\leq\left(1-r\right)^{-2}\left(v^{-2}\left(\alpha\right)-1\right).\end{aligned}$$

Proof. With $\mathcal{S}\left(n,\alpha\right)$ defined by

$$\begin{aligned}&\mathcal{S}\left(n,\alpha\right)\\&\triangleq\min_{j\in D^{n}}\mathbb{P}_{id}^{n}\left(\max_{0\leq m\leq\frac{3}{4}n\log n}\left|\Pi_{m}\left(j\right)-j\right|\leq\alpha\sqrt{n\log n}\,\middle|\,j\in A^{\left\lfloor\frac{3}{4}n\log n\right\rfloor}\right)\\&\leq\min_{j\in D^{n}}\mathbb{P}_{id}^{n}\left(R_{j,t_n}\middle|\,j\in A^{\left\lfloor\frac{3}{4}n\log n\right\rfloor}\right)\\&=\min_{j\in D^{n}}\mathbb{P}_{id}^{n}\left(R_{j,t_n}\middle|\,j\in A^{t_n}\right),\end{aligned}$$

Lemma 3.2 yields

$$\begin{aligned}&\mathbb{P}_{id}^{n}\left(\left|\triangle_{\alpha}\left(\Pi_{t_n}\right)\cap A^{t_n}\right|\leq r\cdot\mathbb{E}_{id}^{n}\left\{\left|\triangle_{\alpha}\left(\Pi_{t_n}\right)\cap A^{t_n}\right|\right\}\right)\\&\qquad\leq\frac{K_{t_n}+\left(1-\mathcal{S}^{2}\left(n,\alpha\right)\right)K_{t_n}^{2}}{\left(1-r\right)^{2}\mathcal{S}^{2}\left(n,\alpha\right)K_{t_n}^{2}},\end{aligned}$$

where $K_{t_n}\triangleq\mathbb{E}_{id}^{n}\left\{\left|D^{n}\cap A^{t_n}\right|\right\}$.

A simple calculation shows that $\lim_{n\to\infty}K_{t_n}=\infty$. Theorem 2.2 (with $t_n=\left\lfloor\frac{3}{4}n\log n\right\rfloor$) implies that

$$\liminf_{n\to\infty}\mathcal{S}\left(n,\alpha\right)\geq1-4\Psi\left(\alpha\sqrt{\frac{8}{3}}\right)=v\left(\alpha\right)>0. \tag{3.1}$$

Therefore

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbb{P}^n_{id}\left(\left|\triangle_\alpha\left(\Pi_{t_n}\right)\cap A^{t_n}\right| \le r\cdot \mathbb{E}^n_{id}\left\{\left|\triangle_\alpha\left(\Pi_{t_n}\right)\cap A^{t_n}\right|\right\}\right)\\
&\quad\le \frac{1}{\left(1-r\right)^2}\limsup_{n\to\infty}\frac{\left(1-\mathcal{S}^2\left(n,\alpha\right)\right)}{\mathcal{S}^2\left(n,\alpha\right)}\\
&\quad\le \left(1-r\right)^{-2}\left(v^{-2}\left(\alpha\right)-1\right).
\end{aligned} \tag{3.2}$$

Note that by (3.1), for any $\delta\in(0,1)$ and sufficiently large n,

$$\begin{aligned}
\mathbb{E}^n_{id}\left\{\left|\triangle_\alpha\left(\Pi_{t_n}\right)\cap A^{t_n}\right|\right\} &= \sum_{j\in D^n}\mathbb{P}^n_{id}\left(R_{j,t_n}\middle|\, j\in A^{t_n}\right)\mathbb{P}^n_{id}\left(j\in A^{t_n}\right)\\
&\ge \left|D^n\right| p_{t,n}\mathcal{S}\left(n,\alpha\right) \ge \delta\varepsilon n p_{t_n,n} v\left(\alpha\right).
\end{aligned}$$

Together with (3.2), this implies

$$\begin{aligned}
&\limsup_{n\to\infty}\mathbb{P}^n_{id}\left(\left|\triangle_\alpha\left(\Pi_{t_n}\right)\cap A^{t_n}\right| \le rv\left(\alpha\right)\varepsilon n p_{t_n,n}\right)\\
&\quad\le \left(1-r/\delta\right)^{-2}\left(v^{-2}\left(\alpha\right)-1\right).
\end{aligned}$$

By letting $\delta\to 1$, the lemma follows. □

3.3 The distribution of $|\triangle_\alpha(\Pi_{t_n})\setminus A^{t_n}|$ under $\mathbb{P}^n_{id}$

As in Subsection 3.1, we shall use Chebyshev's inequality to bound the deviation of $|\triangle_\alpha(\Pi_{t_n})\setminus A^{t_n}|$ from its expectation with high probability. Here, however, the computations are much more involved, and the main difficulty is to compute probabilities that depend to joint distributions of $\Pi_t(i)$ and $\Pi_t(j)$, for general $i\neq j$ in D^n, conditioned on i and j not being in A^t. This is treated in the lemma below, in which we denote by τ^t_m the last time up to time t at which the card numbered m is chosen for removal, and set $\tau^t_m=\infty$ if it is not chosen up to that time.

Lemma 3.4. *Let $i,j\in[n]$ such that $i\neq j$, let $\delta>0$, and let $1\le t_1<t_2\le t$ be natural numbers with $t\le n\log n$. Then*

$$\begin{aligned}
&\mathbb{P}^n_{id}\left(\left(\Pi_t\left(i\right),\Pi_t\left(j\right)\right)\in\overline{i\pm\delta}\times\overline{j\pm\delta}\,\middle|\,\tau^t_i=t_1,\tau^t_j=t_2\right)\\
&\qquad\le\frac{1}{n^2}\left(\left(1+2\left\lfloor\delta\right\rfloor\right)^2+4\delta+g\left(n\right)\right),
\end{aligned}$$

where $g\left(n\right)=\Theta\left(\log^2 n\right)$ is a function independent of all the parameters above.

The proof of Lemma 3.4 is given in Section 5. Now, let us see how it is used to prove the following.

Lemma 3.5. *Let t_n be a sequence of integers such that $0\le t_n\le\left\lfloor\frac{3}{4}n\log n\right\rfloor$ and let k, $\alpha>0$ be real numbers. Then,*

$$\begin{aligned}
&\limsup_{n\to\infty}\mathbb{P}^n_{id}\left(\left|\left|\triangle_\alpha\left(\Pi_{t_n}\right)\setminus A^{t_n}\right|-2\varepsilon\alpha\left(1-p_{t_n,n}\right)\sqrt{n\log n}\right|\dots\right.\\
&\qquad\left.\dots\ge k\sqrt{6\varepsilon\alpha}(n\log n)^{1/4}\right)\le\frac{1}{k^2}.
\end{aligned}$$

Proof. For some $t_1\le t$, consider the Markov chain $\Pi_{t'}\left(j\right)$, $t'=t_1,\dots,t$, conditioned on $\tau^t_j=t_1$. By definition, its initial distribution is the uniform measure on $[n]$. One can easily check that the transition matrix of this chain is symmetric. Therefore its stationary measure, and thus its distribution at time t, is also the uniform measure.

Thus, assuming $n(1-\varepsilon)/2 \geq \alpha\sqrt{n\log n}$,

$$\begin{aligned}\mathbb{E}_{id}^{n}\left\{\left|\triangle_{\alpha}\left(\Pi_{t}\right)\setminus A^{t}\right|\right\} &= \sum_{j\in D^{n}} \mathbb{P}_{id}^{n}\left(R_{j,t}^{A^{c}}\middle| j\notin A^{t}\right)\mathbb{P}_{id}^{n}\left(j\notin A^{t}\right)\\ &= \left|D^{n}\right| \frac{1+2\left\lfloor \alpha\sqrt{n\log n}\right\rfloor}{n}\left(1-p_{t,n}\right).\end{aligned}$$

For the second moment write

$$\mathbb{E}_{id}^{n}\left\{\left|\triangle_{\alpha}\left(\Pi_{t}\right)\setminus A^{t}\right|^{2}\right\} = \mathbb{E}_{id}^{n}\left\{\left|\triangle_{\alpha}\left(\Pi_{t}\right)\setminus A^{t}\right|\right\} + \sum_{i,j\in D: i\neq j} \mathbb{P}_{id}^{n}\left(R_{i,t}^{A^{c}}\cap R_{j,t}^{A^{c}}\right).$$

From Lemma 3.4,

$$\begin{aligned}\sum_{i,j\in D: i\neq j} \mathbb{P}_{id}^{n}\left(R_{i,t}^{A^{c}}\cap R_{j,t}^{A^{c}}\right) &\leq \frac{\left|D^{n}\right|^{2}}{n^{2}}\left(1-p_{t,n}\right)^{2}\times\\ &\left\{\left(1+2\left\lfloor\alpha\sqrt{n\log n}\right\rfloor\right)^{2} + 4\alpha\sqrt{n\log n} + \Theta\left(\log^{2} n\right)\right\}.\end{aligned}$$

Therefore,

$$\limsup_{n\to\infty} \frac{\mathbf{Var}_{id}^{n}\left\{\left|\triangle_{\alpha}\left(\Pi_{t_n}\right)\setminus A^{t_n}\right|\right\}}{\sqrt{n\log n}} \leq 4\varepsilon^{2}\alpha + 2\varepsilon\alpha \leq 6\varepsilon\alpha.$$

By Chebyshev's inequality, the lemma follows. □

Remark 3.6. *Assume t_n is of the form in Theorem 1.1 with c_n satisfying $\limsup c_n/\log n < 1/4$. From Lemma 3.5,*

$$\left|\triangle_{\alpha}(\Pi_{t_n})\setminus A^{t_n}\right| / \left(2\varepsilon\alpha\sqrt{n\log n}\right) \Longrightarrow 1.$$

By a simple computation, taking into account our restriction on c_n, it is seen that $\mathbb{E}_{id}^{n}\left|A^{t_n}\right| = o(\sqrt{n})$. Therefore

$$\left|\triangle_{\alpha}\right| / \left(2\varepsilon\alpha\sqrt{n\log n}\right) \Longrightarrow 1, \tag{3.3}$$

under $\mathbb{P}_{id}^{n}(\Pi_{t_n}\in\cdot)$. From Lemma 3.1, the convergence in (3.3) *clearly holds under the stationary measure $\mathbb{U}^n$ as well.*

4 Proof of Theorem 1.1

In order to prove the lower bound on TV distance we consider the deviation of $|\triangle_\alpha|$ from $2\varepsilon\alpha\sqrt{n\log n}$. Assume t_n is as in the theorem. Let $k>0$ and α be real numbers such that $v(\alpha)>0$ (where $v(\alpha)$ was defined in Lemma 3.3). The parameters k and α will be fixed until (4.5), where we derive a lower bound on the TV distance which depends on them. Then, maximizing over the two parameters, we shall obtain the required bound on TV distance.

Suppose that for some n

$$\left|\left|\triangle_{\alpha}\left(\Pi_{t}\right)\setminus A^{t_n}\right| - 2\varepsilon\alpha\left(1-p_{t_n,n}\right)\sqrt{n\log n}\right| < k\sqrt{6\varepsilon\alpha}(n\log n)^{1/4}, \tag{4.1}$$

and

$$\left|\triangle_{\alpha}\left(\Pi_{t_n}\right)\cap A^{t_n}\right| > \frac{1}{2}v\left(\alpha\right)\varepsilon n p_{t_n,n}. \tag{4.2}$$

Then, if n is sufficiently large,

$$\begin{aligned}
&|\triangle_\alpha(\Pi_{t_n})| - 2\varepsilon\alpha\sqrt{n\log n}\\
&\geq \varepsilon n p_{t_n,n}\left(\frac{1}{2}v(\alpha) - 2\alpha\sqrt{n\log n}/n\right) - k\sqrt{6\varepsilon\alpha}(n\log n)^{1/4}\\
&\geq k\sqrt{2\varepsilon\alpha}(n\log n)^{1/4},
\end{aligned}\tag{4.3}$$

where the last inequality follows from the following calculation: writing

$$\log\frac{np_{t_n,n}}{(n\log n)^{1/4}} = \frac{3}{4}\log n - \frac{1}{4}\log\log n + \log p_{t_n,n},$$

substituting $p_{t_n,n} = (1-1/n)^{t_n}$ and $t_n = \frac{3}{4}n\log n - \frac{1}{4}n\log\log n - c_n n$, and using the fact that $\log(1+x) = x + O(x^2)$ as $x\to 0$, we arrive at

$$\log\frac{np_{t_n,n}}{(n\log n)^{1/4}} = c_n + o(1) \to \infty. \tag{4.4}$$

Now, since for large n (4.1) and (4.2) imply (4.3), by a union bound, Lemma 3.3 and Lemma 3.5 imply

$$\begin{aligned}
&\liminf_{n\to\infty}\mathbb{P}_{id}^n\left(|\triangle_\alpha(\Pi_{t_n})| - 2\varepsilon\alpha\sqrt{n\log n} \geq k\sqrt{2\varepsilon\alpha}(n\log n)^{1/4}\right)\\
&\geq 1 - \frac{1}{k^2} - \left(1-\frac{1}{2}\right)^{-2}\left(v^{-2}(\alpha) - 1\right) \triangleq \phi(k,\alpha).
\end{aligned}$$

In addition, from Lemma 3.1,

$$\limsup_{n\to\infty}\mathbb{U}^n\left(|\triangle_\alpha(\sigma)| - 2\varepsilon\alpha\sqrt{n\log n} \geq k\sqrt{2\varepsilon\alpha}(n\log n)^{1/4}\right) \leq \frac{1}{k^2}.$$

Thus,

$$\liminf_{n\to\infty}\|\mathbb{P}_{id}^n(\Pi_{t_n}\in\cdot) - \mathbb{U}^n\|_{TV} \geq \phi(k,\alpha) - \frac{1}{k^2}. \tag{4.5}$$

Since k and α were arbitrary, and since as $k,\alpha\to\infty$, $\phi(k,\alpha)\to 1$ and $\frac{1}{k^2}\to 0$,

$$\lim_{n\to\infty}\|\mathbb{P}_{id}^n(\Pi_{t_n}\in\cdot) - \mathbb{U}^n\|_{TV} = 1.$$

□

Remark 4.1. *In Section 3, we have seen that the standard deviation of both $|\triangle_\alpha|$ under $\mathbb{U}^n$ and $|\triangle_\alpha(\Pi_{t_n})\setminus A^{t_n}|$ under $\mathbb{P}_{id}^n$ is of order $\Theta((n\log n)^{\frac{1}{4}})$. Since $\mathbb{E}_{id}^n\{|\triangle_\alpha(\Pi_{t_n})\cap A^{t_n}|\} = \Theta(np_{t_n,n})$, by (4.4), it is of higher order than $\Theta((n\log n)^{\frac{1}{4}})$, if and only if t_n is of the form in Theorem 1.1. In particular, this shows why the term $-\frac{1}{4}n\log\log n$ is essential to us in the choice of t_n.*

5 Proof of Lemma 3.4

In this section we prove Lemma 3.4 and additional results needed for the proof.

Proof of Lemma 3.4. For $m_1\in[n-1]$, $m_2\in[n]$ and $0\leq t'\in\mathbb{Z}$, let $\sigma\in S_n$ be some permutation such that $\sigma(j) = m_2$ and

$$\sigma(i) = \begin{cases} m_1+1 & \text{if } m_2\leq m_1,\\ m_1 & \text{if } m_2 > m_1,\end{cases}$$

and let $\mathcal{P}^{t'}_{m_1,m_2} = \mathcal{P}^{n,t'}_{m_1,m_2}$ be the probability measure on $[n] \times [n]$ defined by

$$\mathcal{P}^{n,t'}_{m_1,m_2}(\cdot) = \mathbb{P}^n_\sigma\left((\Pi_{t'}(i), \Pi_{t'}(j)) \in \cdot \mid i,j \in A^{t'} \right).$$

(Which, obviously, does not depend on the values $\sigma(k)$ for $k \notin \{i,j\}$.)

That is, starting with a deck whose ordering is obtained by inserting the card numbered j in position m_2, in a deck composed of the $n-1$ cards with numbers in $[n] \setminus \{j\}$ in which the position of card i is m_i, $\mathcal{P}^{t'}_{m_1,m_2}$ is the joint probability law of the positions of the cards numbered i and j after performing t' random-to-random insertion shuffles, conditioned on not choosing either of the cards i and j.

Now, let t, t_1 and t_2 be natural numbers as in the statement of the lemma, which will be fixed throughout the proof. Define the events

$$Q^+_m = \{\Pi_{t_2-1}(i) = m, \Pi_{t_2-1}(j) > m\},$$
$$Q^-_m = \{\Pi_{t_2-1}(i) = m, \Pi_{t_2-1}(j) < m\},$$

and define q^+_m and q^-_m by

$$q^\pm_m = \mathbb{P}^n_{id}\left(Q^\pm_m \;\middle|\; \tau^t_i = t_1, \tau^t_j = t_2\right).$$

Define the probability measure μ on $[n] \times [n]$ by

$$\begin{aligned}\mu(\cdot) &\triangleq \mathbb{P}^n_{id}\left((\Pi_t(i), \Pi_t(j)) \in \cdot \mid \tau^t_i = t_1, \tau^t_j = t_2\right) \\ &= \frac{1}{n}\sum_{m_2=1}^{n}\left\{\sum_{m_1=1}^{n-1} q^+_{m_1}\mathcal{P}^{t-t_2}_{m_1,m_2}(\cdot) + \sum_{m_1=2}^{n} q^-_{m_1}\mathcal{P}^{t-t_2}_{m_1-1,m_2}(\cdot)\right\}.\end{aligned}$$

Considering the Markov chain $\Pi_{t'}(i)$, $t' = t_1, \ldots, t_2 - 1$, conditioned on $\tau^t_i = t_1$ and $\tau^t_j = t_2$, by an argument similar to that given in the beginning of the proof of Lemma 3.5, $\Pi_{t_2-1}(i)$ is uniformly distributed on $[n]$. Thus, for any $m \in [n]$,

$$q^+_m + q^-_m = \frac{1}{n}. \tag{5.1}$$

Similarly, for any $s \in \mathbb{N}$, the transition matrix of the chain $(\Pi_{t'}(i), \Pi_{t'}(j))$, $t' = 0, \ldots, s$, conditioned on $i, j \in A^s$, is symmetric, and therefore the uniform measure on $\left\{(m_1, m_2) \in [n]^2 : m_1 \neq m_2\right\}$, which we denote by $\mathbb{U}^n_{(2)}$, is a stationary measure of the chain (the chain is reducible, thus the stationary measure is not unique). It therefore follows that for any $0 \le t' \in \mathbb{Z}$,

$$\frac{1}{n(n-1)}\sum_{m_1=1}^{n-1}\sum_{m_2=1}^{n}\mathcal{P}^{t'}_{m_1,m_2}(\cdot) = \mathbb{U}^n_{(2)}(\cdot). \tag{5.2}$$

Our next step is to define two additional Markov chains $\Pi^-_{t'}$ and $\Pi^+_{t'}$, $t' = t_2, t_2 + 1, \ldots, t$, with state space S_n, such that on $\left\{\tau^t_i = t_1, \tau^t_j = t_2\right\}$,

$$\Pi^-_{t'}(i) \le \Pi_{t'}(i) \le \Pi^+_{t'}(i) \quad \text{and} \quad \Pi^-_{t'}(j) = \Pi_{t'}(j) = \Pi^+_{t'}(j), \tag{5.3}$$

for any $t' = t_2, t_2 + 1, \ldots, t$. Once we have done so, defining μ^+, μ^- by

$$\mu^\pm(\cdot) \triangleq \mathbb{P}^n_{id}\left(\left(\Pi^\pm_t(i), \Pi^\pm_t(j)\right) \in \cdot \;\middle|\; \tau^t_i = t_1, \tau^t_j = t_2\right),$$

it will follow that

$$\begin{aligned}\mu\left(\overline{i \pm \delta} \times \overline{j \pm \delta}\right) \le \mu&\left([n] \times \overline{j \pm \delta}\right) \\ &- \mu^-\left(((i + \delta, n] \cap [n]) \times \overline{j \pm \delta}\right) \\ &- \mu^+\left(([1, i - \delta) \cap [n]) \times \overline{j \pm \delta}\right).\end{aligned} \tag{5.4}$$

For each $m_1 \in [n] \setminus \{1, n\}$ define the events $\widehat{Q}^+_{m_1}$, $\widehat{Q}^-_{m_1}$, $\widehat{Q}^+$ and $\widehat{Q}^-$ by

$$\begin{aligned} \widehat{Q}^{\pm}_{m_1} &= Q^{\pm}_{m_1} \bigcap \{\Pi_{t_2}(j) \neq m_1\} \bigcap \{\tau^t_i = t_1, \tau^t_j = t_2\}, \\ \widehat{Q}^{\pm} &= \bigcup_{m_1 = 2, \dots, n-1} Q^{\pm}_{m_1}. \end{aligned} \tag{5.5}$$

Let us define $\Pi^+_{t'}$ and $\Pi^-_{t'}$ by setting, for $t' = t_2, t_2 + 1, \dots, t$,

$$\begin{aligned} \Pi^+_{t'} &= \Pi^0_{t'} \text{ on } \widehat{Q}^-, & \Pi^+_{t'} &= \Pi_{t'} \text{ on } \left(\widehat{Q}^-\right)^c, \\ \Pi^-_{t'} &= \Pi^0_{t'} \text{ on } \widehat{Q}^+, & \Pi^-_{t'} &= \Pi_{t'} \text{ on } \left(\widehat{Q}^+\right)^c, \end{aligned}$$

where $\Pi^0_{t'}$ is an additional Markov chain defined on $\widehat{Q}^+ \cup \widehat{Q}^-$ as described below.

The random walk $\Pi_{t'}$ on S_n corresponds to the ordering of a deck of n cards as it is being shuffled by random-to-random insertion shuffles. Let us call this deck of cards deck A, and for each time t' let us denote by $c_{t'}$ and $d_{t'}$ the number of the card removed from the deck at that time and the position into which it is reinserted, respectively. (To avoid any confusion – we refer to the ordering of the deck after the removal the card numbered $c_{t'}$ and its reinsertion to position $d_{t'}$ as the state of the deck at time t', and not $t' + 1$.)

In order to define $\Pi^0_{t'}$, we describe a shuffling process on a deck of n cards, which we shall refer to as deck B, on the set of times $t' = t_2, t_2 + 1, \dots, t$, and set $\Pi^0_{t'}$ to be the permutation corresponding to the ordering of the deck at time t' (i.e., $\Pi^0_{t'}(k)$ is the position of the card numbered k).

We begin by defining the state of deck B at time $t' = t_2$ on $\widehat{Q}^+$ (respectively, $\widehat{Q}^-$) as the deck obtained by taking a deck of n cards ordered as deck A is ordered at the same time and transposing the card numbered i with the card which has position lower (receptively, higher) by 1 from the card numbered i.

For a given state of decks A and B, let us say that two cards with numbers in $[n] \setminus \{i, j\}$, one in each of the decks, 'match' each other, if after removing the cards numbered i and j from both decks they have the same position.

At each of the times $t' = t_2 + 1, \dots, t$, suppose deck B is shuffled based upon how deck A is as follows: when the card numbered $c_{t'}$ is removed from deck A, we also remove the matching card from deck B; then, we reinsert both cards to their decks in the same position, $d_{t'}$. This defines the state of deck B, and thus $\Pi^0_{t'}$, for times $t' = t_2 + 1, \dots, t$.

A concrete example of a simultaneous shuffle of both decks with $n = 8$, $i = 5$ and $j = 7$ is given in Figure 1. Cards i and j are colored gray. Card 1 is chosen for removal in deck A, and so the matching card, 4, is the one removed from deck B. Then, they are reinserted in the same position.

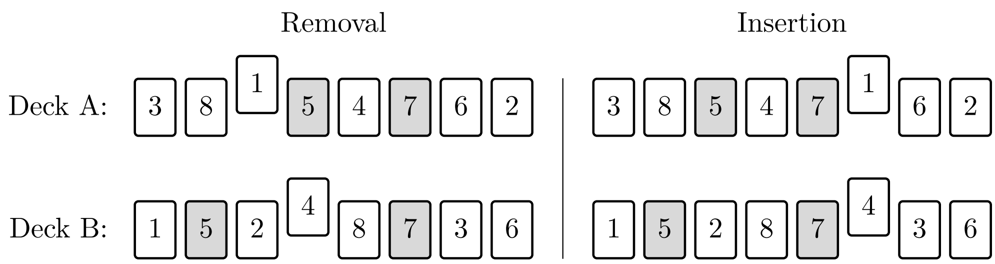

Figure 1: A shuffle of deck A and deck B.

Directly from definition, (5.3) holds for $t' = t_2$. It is also easy to verify that every single shuffle of decks A and B as described above preserves the relations in (5.3), which implies that, indeed, (5.3) holds for any $t' = t_2, t_2+1, \ldots, t$.

Note that, by definition,

$$\begin{aligned}\mu^+(\cdot) &= \sum_{m=1}^{n-1} q_m^+ \mathbb{P}_{id}^n \left(\left(\Pi_t^+(i), \Pi_t^+(j)\right) \in \cdot \,\middle|\, Q_m^+, \tau_i^t = t_1, \tau_j^t = t_2 \right) \\ &+ \sum_{m=2}^{n} q_m^- \mathbb{P}_{id}^n \left(\left(\Pi_t^+(i), \Pi_t^+(j)\right) \in \cdot \,\middle|\, Q_m^-, \tau_i^t = t_1, \tau_j^t = t_2 \right) \\ &= \frac{1}{n} \sum_{m_1=1}^{n-1} \sum_{m_2=1}^{n} q_{m_1}^+ \mathcal{P}_{m_1,m_2}^{t-t_2}(\cdot) + \frac{1}{n} \sum_{m_1=2}^{n} q_{m_1}^- \mathcal{P}_{m_1-1,m_1}^{t-t_2}(\cdot) \\ &+ \frac{1}{n} \sum_{m_1=2}^{n} \sum_{m_2 \in [n] \setminus \{m_1\}} q_{m_1}^- \mathcal{P}_{m_1,m_2}^{t-t_2}(\cdot).\end{aligned}$$

From this, together with (5.1) and (5.2), we obtain

$$\begin{aligned}\mu^+(\cdot) = \frac{n-1}{n} \mathbb{U}_{(2)}^n(\cdot) + \frac{1}{n^2} \sum_{m=1}^{n} \mathcal{P}_{n-1,m}^{t-t_2}(\cdot) \\ + \frac{1}{n} \sum_{m=2}^{n-1} \left\{ q_m^- \mathcal{P}_{m-1,m}^{t-t_2}(\cdot) - q_m^- \mathcal{P}_{m,m}^{t-t_2}(\cdot) \right\}.\end{aligned} \tag{5.6}$$

Similarly,

$$\begin{aligned}\mu^-(\cdot) = \frac{n-1}{n} \mathbb{U}_{(2)}^n(\cdot) + \frac{1}{n^2} \sum_{m=1}^{n} \mathcal{P}_{1,m}^{t-t_2}(\cdot) \\ + \frac{1}{n} \sum_{m=2}^{n-1} \left\{ q_m^+ \mathcal{P}_{m,m}^{t-t_2}(\cdot) - q_m^+ \mathcal{P}_{m-1,m}^{t-t_2}(\cdot) \right\}.\end{aligned} \tag{5.7}$$

According to (5.3), $\mu\left([n] \times \overline{j \pm \delta}\right) = \mu^+\left([n] \times \overline{j \pm \delta}\right)$. Hence, by substitution of (5.6) and (5.7) in (5.4), and using (5.1), it can be easily shown that

$$\begin{aligned}\mu\left(\overline{i \pm \delta} \times \overline{j \pm \delta}\right) \le \frac{n-1}{n} \mathbb{U}_{(2)}^n\left(\overline{i \pm \delta} \times \overline{j \pm \delta}\right) + \\ \frac{1}{n^2} \sum_{m=1}^{n} \mathcal{P}_{n-1,m}^{t-t_2}\left([n] \times \overline{j \pm \delta}\right) + \frac{1}{n^2} \sum_{m=2}^{n-1} \mathcal{P}_{m-1,m}^{t-t_2}\left([n] \times \overline{j \pm \delta}\right).\end{aligned} \tag{5.8}$$

The first summand is bounded by

$$\frac{n-1}{n} \mathbb{U}_{(2)}^n\left(\overline{i \pm \delta} \times \overline{j \pm \delta}\right) \le \frac{\left(1 + 2\lfloor \delta \rfloor\right)^2}{n^2}.$$

Note that, for fixed m_2, $\mathcal{P}_{m_1,m_2}^{t_2-t'}\left([n] \times \overline{j \pm \delta}\right)$ is identical for all m_1 such that $m_1 < m_2$, and for all m_1 such that $m_1 \ge m_2$. Thus,

$$\sum_{m=2}^{n-1} \mathcal{P}_{m-1,m}^{t-t_2}\left([n] \times \overline{j \pm \delta}\right) = \sum_{m=2}^{n-1} \mathcal{P}_{1,m}^{t-t_2}\left([n] \times \overline{j \pm \delta}\right).$$

Corollary 5.2 below provides an upper bound for this sum. Bounding the additional sum in (5.8) by the same bound can be done similarly, which completes the proof. □

Corollary 5.2, used in the previous proof, will follow from the following.

Lemma 5.1. *For any real number $\delta \geq 0$, integers $r,\, t \geq 0$, $i, j \in [n]$, and $m \in [n-1]$,*

$$\mathcal{P}^{n,t}_{1,m+1}\left([n] \times \overline{j \pm \delta}\right) \leq \mathbb{P}^{n-1}_{id}\left(\Pi_t\left(1\right) > r \mid 1 \in A^t\right) + \mathbb{P}^{n-1}_{id}\left(\Pi_t\left(m\right) \in \overline{j \pm (\delta + r)} \,\middle|\, m \in A^t\right).$$

Before we turn to proof of the lemma, let us state and prove the above mentioned corollary.

Corollary 5.2. *For any real number $\delta \geq 0$, integer $0 \leq t \leq n \log n$, and $j \in [n]$,*

$$\sum_{m=2}^{n-1} \mathcal{P}^{n,t}_{1,m}\left([n] \times \overline{j \pm \delta}\right) \leq 2\delta + \widehat{g}\left(n\right),$$

where $\widehat{g}\left(n\right) = \Theta\left(\log^2 n\right)$ is a function independent of the parameters above.

Proof. From Lemma 5.1, for any real $\delta \geq 0$, integers $r,\, t \geq 0$, and $j \in [n]$,

$$\sum_{m=2}^{n-1} \mathcal{P}^{n,t}_{1,m}\left([n] \times \overline{j \pm \delta}\right) \leq \left(n-2\right) \mathbb{P}^{n-1}_{id}\left(\Pi_t\left(1\right) > r \mid 1 \in A^t\right) + \sum_{m=1}^{n-1} \mathbb{P}^{n-1}_{id}\left(\Pi_t\left(m\right) \in \overline{j \pm (\delta + r)} \,\middle|\, m \in A^t\right). \tag{5.9}$$

Clearly, the transition probabilities of $\Pi_{t'}\left(m\right)$, $t' = 0, 1, \ldots, t$, conditioned on $m \in A^t$ do not depend on m. Thus, up to a factor of $n-1$, the sum on the right-hand side above is equal to the probability that at time t, the state of the Markov chain with those transition probabilities and with uniform initial distribution belongs to $\overline{j \pm (\delta + r)}$. Since the transition matrix of this chain is symmetric, the stationary measure for this chain is the uniform measure. Thus,

$$\sum_{m=1}^{n-1} \mathbb{P}^{n-1}_{id}\left(\Pi_t\left(m\right) \in \overline{j \pm (\delta + r)} \,\middle|\, m \in A^t\right) \leq 1 + 2\left(\delta + r\right). \tag{5.10}$$

By (2.4) and by the same argument as in (2.14), setting $t_n = \lfloor n \log n \rfloor$, for any sequence of integers $r_n \geq 0$,

$$\begin{aligned}\mathbb{P}^{n-1}_{id}\left(\Pi_t\left(1\right) > r_n \mid 1 \in A^{t_n}\right) &\leq \mathbb{P}^{n-1}_{id}\left(\max_{0 \leq t' \leq t_n} \left|\Pi_{t_n}\left(1\right) - 1\right| \geq r_n \,\middle|\, 1 \in A^{t_n}\right) \\ &= \mathbb{P}\left(\max_{0 \leq t' \leq t_n} \left|\zeta^{n-1,1,r_n}_{t'} - 1\right| \geq r_n\right) \leq 4\mathbb{P}\left(S^{n-1,1,r_n}_{t_n} \geq r_n/3\right) \\ &\quad + 4\mathbb{P}\left(W^{n-1,r_n}_{t_n} \geq r_n/3\right) + \mathbb{P}\left(N^{n-1}_{t_n} \geq r_n/3\right).\end{aligned}$$

Using Bernstein inequalities ([5], Theorem 2.8), it is easy to verify that one can choose a sequence $r_n = \Theta\left(\log^2 n\right)$ such that the last part of the inequality above is $o\left(\log^2 n/n\right)$. From this, together with (5.9) and (5.10), the corollary follows. □

We now turn the proof of Lemma 5.1.

Proof of Lemma 5.1. The proof is based on a coupling of the Markov chains corresponding to the shuffling of two decks of cards. The first of the two decks contains n cards, numbered from 1 to n, and at time 0 (the initial state) has card i at position 1 and card j at position $m+1$. The second deck contains $n-1$ cards, numbered from 1 to $n-1$, and at time 0 is ordered lexicographically, i.e., according to the numbers of the cards. Let us call the decks deck 1 and deck 2, respectively.

We want to define a procedure to simultaneously shuffle the decks such that:

1. At each step deck 1 is shuffled by choosing a random card, different from i and j, removing it from the deck, and inserting it back into the deck at a random position; with shuffles at different steps being independent.

2. At each step deck 2 is shuffled by choosing a random card, different from m, removing it from the deck, and inserting it back into the deck at a random position; with shuffles at different steps being independent.

3. For all $t \geq 0$,
$$J_t - I_t \leq M_t \leq J_t - 1, \tag{5.11}$$
where J_t (respectively, I_t) denotes the position of the card numbered j (respectively, i) in deck 1 after completing t shuffles, and M_t denotes the position of card m in deck 2 after completing t shuffles.

We shall also need the notation $\bar{J}_t$ (respectively, $\bar{I}_t$) for the position of the card numbered j (respectively, i) in deck 1, after completing $t-1$ shuffles and performing only the removal of the t-th shuffle. Note that since after the removal the deck contains only $n-1$ cards, these positions are values in $[n-1]$. Similarly, $\bar{M}_t$ shall denote the corresponding position of the card numbered m in deck 2.

The definition of the shuffling shall be done inductively, and so, let us begin by assuming that (5.11) holds for some time t'. Under the assumption, one can easily define a bijection from the set of cards in deck 1 that are different from i and j, to the set cards in deck 2 that are different from m, such that at time t':

1. any card with position between the cards numbered i and j in deck 1 is mapped to a card below m in deck 2; and

2. any card below m in deck 2 is the image of some card below j in deck 1.

See, for example, Figure 2.

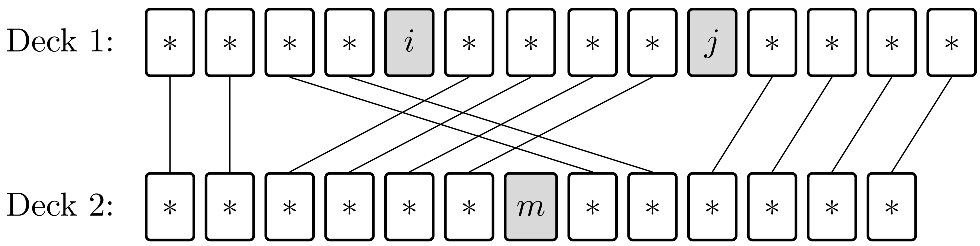

Figure 2: A bijection for decks 1 and 2.

Once the bijection is defined, one can perform the removal of step $t'+1$ from both decks by choosing a random card (different from i and j) from deck 1 and removing this card from deck 1 and its image under the bijection from deck 2. This ensures that

$$\bar{J}_{t'+1} - \bar{I}_{t'+1} \leq \bar{M}_{t'+1} \leq \bar{J}_{t'+1} - 1. \tag{5.12}$$

Denote

$$V_1 \triangleq (J_{t'+1} - I_{t'+1}) - \left(\bar{J}_{t'+1} - \bar{I}_{t'+1}\right),$$
$$V_2 \triangleq M_{t'+1} - \bar{M}_{t'+1}\ ,\ V_3 \triangleq J_{t'+1} - \bar{J}_{t'+1},$$

and note that $V_1,\, V_2,\, V_3 \in \{0,1\}$.

If we assume that the first two of the three conditions we need the shuffling to satisfy hold, then the conditional probabilities

$$p_i = p_i(\bar{I}_{t'+1}, \bar{J}_{t'+1}, \bar{M}_{t'+1}) \triangleq \mathbb{P}(V_i = 1|\bar{I}_{t'+1}, \bar{J}_{t'+1}, \bar{M}_{t'+1})\,, \quad i = 1,\, 2,\, 3,$$

satisfy

$$p_1 = \frac{\bar{J}_{t'+1} - \bar{I}_{t'+1}}{n} \le p_2 = \frac{\bar{M}_{t'+1}}{n-1} \le p_3 = \frac{\bar{J}_{t'+1}}{n}.$$

Therefore, since $\{V_1 = 1\} \subset \{V_3 = 1\}$, it is possible to couple the reinsertions of the cards back to their decks at step $t'+1$, so that the position of each of the cards after reinsertion is uniform in its deck, and so that (5.11) also holds for time $t'+1$.

By induction, this completes our definition of the shuffling of the two decks and implies that for any integers $t,\, r \ge 0$,

$$\mathcal{P}^{n,t}_{1,m+1}\left([n] \times \overline{j \pm \delta}\right) \le \mathcal{P}^{n,t}_{1,m+1}\left(([n] \setminus [1,r]) \times [n]\right) + \\ \mathbb{P}^{n-1}_{id}\left(\Pi_t\,(m) \in \overline{j \pm (\delta + r)}\,\middle|\, m \in A^t\right).$$

To finish the proof, note that by a coupling argument (remove cards as described in the proof of Lemma 3.4 for decks A and B, with the difference of removing card m instead of cards i and j from the smaller deck in order to compare positions for 'matchings', and define the random insertion appropriately),

$$\mathcal{P}^{n,t}_{1,m+1}\left(([n] \setminus [1,r]) \times [n]\right) \le \mathbb{P}^{n-1}_{id}\left(\Pi_t\,(1) > r \mid 1 \in A^t\right).$$

□

References

[1] P. Billingsley, *Probability and measure*, third ed., Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1995, A Wiley-Interscience Publication. MR-1324786

[2] P. Diaconis and L. Saloff-Coste, *Random walks on finite groups: a survey of analytic techniques*, Probability measures on groups and related structures, XI (Oberwolfach, 1994), World Sci. Publ., River Edge, NJ, 1995, pp. 44–75. MR-1414925

[3] P. Diaconis, *Mathematical developments from the analysis of riffle shuffling*, Groups, combinatorics & geometry (Durham, 2001), World Sci. Publ., River Edge, NJ, 2003, pp. 73–97. MR-1994961

[4] P. Diaconis and L. Saloff-Coste, *Comparison techniques for random walk on finite groups*, Ann. Probab. **21** (1993), no. 4, 2131–2156. MR-1245303

[5] V. V. Petrov, *Limit theorems of probability theory*, Oxford Studies in Probability, vol. 4, The Clarendon Press Oxford University Press, New York, 1995, Sequences of independent random variables, Oxford Science Publications. MR-1353441

[6] J.-C. Uyemura Reyes, *Random walk, semi-direct products, and card shuffling*, ProQuest LLC, Ann Arbor, MI, 2002, Thesis (Ph.D.)–Stanford University. MR-2703300

[7] L. Saloff-Coste and J. Zúñiga, *Refined estimates for some basic random walks on the symmetric and alternating groups*, ALEA Lat. Am. J. Probab. Math. Stat. **4** (2008), 359–392. MR-2461789

Acknowledgments. This work arose from a graduate course taught by Professor Ross G. Pinsky. I am grateful to him for introducing the problem and for valuable comments throughout the preparation of this work. I would also like to thank my advisor, Professor Robert J. Adler, and the anonymous referees for their helpful remarks.