Séminaire Lotharingien de Combinatoire **93B** (2025)
Article #6, 11 pp.

Proceedings of the 37th Conference on Formal Power Series and Algebraic Combinatorics (Sapporo)

Cutoff for the Biased Random Transposition Shuffle

Evita Nestoridi*[1] and Alan Yan†[2]

[1] *Mathematics Department, Stony Brook University, Stony Brook NY, 11794-3651, USA*
[2] *Department of Mathematics, Harvard University, Cambridge MA, 02138, USA*

Abstract. We study the biased random transposition shuffle, a natural generalization of the classical random transposition shuffle studied by Diaconis and Shahshahani. Using the representation theory of the symmetric group, we diagonalize the transition matrix of the shuffle. We use these eigenvalues to prove that the shuffle exhibits total variation cutoff at time $t_N = \frac{1}{2b} N \log N$ with window N. We also prove that the limiting distribution of the number of fixed cards near the cutoff time is Poisson.

Keywords: symmetric group, partitions, Littlewood–Richardson coefficients, cutoff phenomenon, hive models, card shuffling

1 Introduction

The study of card shuffling with the help of representation theory was initially introduced by Diaconis and Shahshahani [12] for the random transpositions shuffle. In this shuffle, consider a deck of N distinct cards. Pick two cards uniformly at random with repetition and swap them. Diaconis and Shahshahani proved that it takes $\frac{1}{2} N \log N$ repetitions until the deck is shuffled sufficiently well. This model is now famously known as the *random transpositions shuffle*, and to this day, it is considered the best approximation of the Markov chain that genes in DNA sequences follow [5]. Following Diaconis and Shahshahani, there has been a variety of other card shuffles that have been studied with the use of representation theory (see [15, 17, 13, 8, 2, 23, 9, 10]).

In this extended abstract, we consider a generalization of random transpositions, where we partition the cards into two sets and pick cards with probabilities depending on which half of the deck they belong. More precisely, consider a deck of $N = 2n$ distinct cards and denote $[N] := \{1, \ldots, N\}$. We partition the cards into two sets $[N] = A \sqcup B$. Fix numbers $0 < b \leq a$ satisfying $a|A| + b|B| = N$. Let $\mu_{a,b}$ be the probability measure on $[N]$ given by

$$\mu_{a,b}(x) = \begin{cases} \frac{a}{N} & \text{if } x \in A \\ \frac{b}{N} & \text{if } x \in B. \end{cases}$$

*evrydiki.nestoridi@stonybrook.edu. Partially supported by NSF grant DMS-2346986
†alanyan@math.harvard.edu

We pick two cards uniformly at random with repetition according to $\mu_{a,b}$ and swap them. By identifying all of the configurations of our deck with elements in $\mathfrak{S}_N$, we can think of the shuffle as a random walk on $\mathfrak{S}_N$. We refer to it as the **biased random transpositions shuffle** to emphasize its connection to the classical case of random transpositions considered in [12]. Indeed, in the case $a = b = 1$, we recover the original random transposition shuffle.

For $x \in \mathfrak{S}_N$, let $P^t(x, \bullet)$ be the distribution of our deck configuration after t shuffles starting from the initial configuration x. For two probability measures μ, ν on $\mathfrak{S}_N$, let $d_{\mathsf{TV}}(\mu, \nu)$ be the **total variation distance** between μ and ν. Explicitly, the total variation distance is equal to

$$d_{\mathsf{TV}}(\mu, \nu) := \frac{1}{2} \sum_{x \in \mathfrak{S}_N} |\mu(x) - \nu(x)|.$$

We prove that the biased random transposition shuffle exhibits **total variation distance cutoff** at time $\frac{1}{2b} N \log N$ by diagonalizing the transition matrix $P = (P(x, y))_{x,y \in \mathfrak{S}_N}$ and using this spectral information to produce lower and upper bounds for the mixing time. We are ready to state our main result, where for simplicity we assume $N = 2n$ is even.

Theorem 1.1. *Let $|A| = |B| = n$. Let $c > 0$ be a positive real number. For sufficiently large N and some universal constant $C > 0$, we have*

$$d_{\mathsf{TV}}\left(P^{\frac{N}{2b}(\log N + c)}(id, \cdot), U\right) \le C \cdot e^{-c} \quad \textit{and}$$
$$d_{\mathsf{TV}}\left(P^{\frac{N}{2b}(\log N - c)}(id, \cdot), U\right) \ge 1 - e^{-\frac{1}{2}\left(\sqrt{1+\frac{1}{2}e^c} - 1\right)^2} + o(1),$$

where U is the uniform measure on $\mathfrak{S}_N$.

Theorem 1.1 describes the occurrence of the **cutoff phenomenon**. A family of Markov chains on $\mathfrak{S}_N$ is said to have (total variation) cutoff at t_N with window $w_N = o(t_N)$ if

$$\lim_{c\to\infty} \lim_{N\to\infty} d(t_N - cw_N) = 1 \quad \text{and} \quad \lim_{c\to\infty} \lim_{N\to\infty} d(t_N + cw_N) = 0.$$

Figure 1 is a schematic picture of the cutoff phenomenon. The curve shown is the graph of the function $d(t)$ representing the total variation distance from our deck of cards at time t to the uniform distribution. The cutoff phenomenon concerns the curve outside of the window $[t_N - cw_N, t_N + cw_N]$, specifically its convergence to a step function. The question about the limiting curve of $d(t)$ within the window $[t_N - cw_N, t_N + cw_N]$ concerns the **limit profile**. In this extended abstract, we do not discuss the limit profile and direct the interested reader to [24].

The occurrence of cutoff is a central question in Markov chain mixing, see more in [25]. The first instances of cutoff were in card shuffling in works by Aldous, Diaconis and Shahshahani [1, 12], and since then cutoff has been studied in terms of many other card shuffles [3, 6, 8, 17, 19].

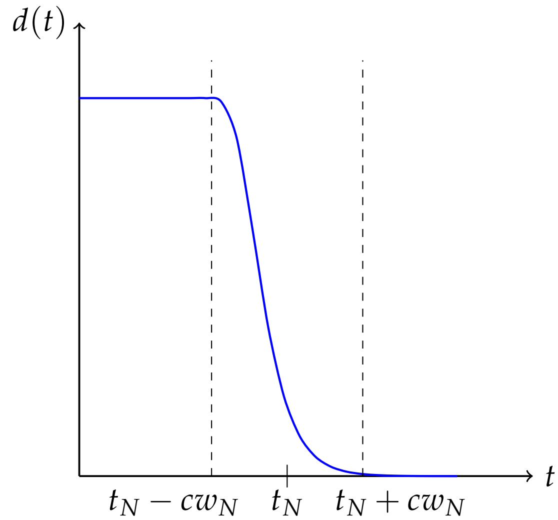

Figure 1: The cutoff phenomenon

The biased random transposition shuffle was also studied in [7], where the authors focused on separation distance. Their proof relied on analyzing the behavior of a stopping time, that unfortunately turned out not to be a strong stationary time as explained in Section 5.3 of [26]. In this paper, we follow a spectral approach to prove the behavior that was predicted in [7].

The upper bound of Theorem 1.1 is proven by an ℓ_2-bound, which makes use of the eigenvalues of the transition matrix P. Our second theorem explicitly describes the eigenvalue spectrum of P. The statement involves the **diagonal index** of a partition λ, defined to be

$$\mathrm{Diag}(\lambda) := \sum_{i \geq 0} \binom{\lambda_i}{2} - \sum_{i \geq 0} \binom{\lambda_i^*}{2}.$$

Here, λ^* is the partition conjugate to λ. The eigenvalue spectrum of P is given by the following theorem.

Theorem 1.2. *The transition matrix P has eigenvalues*

$$\frac{a^2|A| + b^2|B|}{N^2} + \frac{2(a^2 - ab)}{N^2} \mathrm{Diag}(\mu) + \frac{2(b^2 - ab)}{N^2} \mathrm{Diag}(\nu) + \frac{2ab}{N^2} \mathrm{Diag}(\lambda),$$

with multiplicities $f_\lambda f_\mu f_\nu c^\lambda_{\mu,\nu}$ for all partitions $\lambda \vdash N$, $\mu \vdash |A|$, and $\nu \vdash |B|$, where f_λ is the number of standard Young tableaux of shape λ and $c^\lambda_{\mu,\nu}$ is the Littlewood–Richardson coefficient.

The lower bound of Theorem 1.1 is proven by studying the number of fixed cards at time t. Let Fix_c be the number of fixed points after shuffling our deck of cards using the biased random transposition shuffle $\frac{1}{2b}N(\log N - c)$ times. Let Fix be the number of fixed points of a uniformly chosen permutation. The random variable Fix_c has the following limiting distribution.

Theorem 1.3. *Suppose that $b < 1$. Then, we have*

$$\mathrm{Fix}_c \xrightarrow{dist} \mathrm{Poiss}\left(1 + \frac{1}{2}e^c\right), \quad \text{as } N \to \infty. \tag{1.4}$$

where the convergence is convergence in distribution. Also, we have

$$d_{\mathsf{TV}}\left(\mathrm{Fix}_c, \mathrm{Fix}\right) \longrightarrow d_{\mathsf{TV}}\left(\mathrm{Poiss}(1), \mathrm{Poiss}\left(1 + \frac{1}{2}e^c\right)\right), \quad \text{as } N \to \infty. \tag{1.5}$$

Note that Equation 1.5 follows immediately from Equation 1.4 by definition of convergence in distribution. Theorem 1.3 is analogous to [22, Equation 1.6] which was proven for random transpositions.

Remark 1.6. Our proof is motivated by the proof in [12] for random transpositions. However, in the biased case we run into several difficulties which require new ideas. Since our shuffle is not conjugacy class invariant, we need another way to diagonalize our shuffle. In terms of the analysis, we need an understanding of the positivity of the Littlewood–Richardson coefficients, which plays no role in random transpositions. Finally, we provide an algebraic proof of the fixed point result of [22] which also generalizes to the biased case.

2 Spectral Analysis

In this section, we prove Theorem 1.2, which gives an explicit description of the eigenvalue spectrum of the biased random transposition shuffle. We will do this by relating the card shuffle with the representation theory of the symmetric group. We first define some relevant elements in the group algebra which encode the transition matrix.

Definition 2.1. For disjoint subsets $S, T \subseteq [N]$, let $\mathrm{Flip}(S)$ be the set of transpositions swapping two elements in S and let $\mathrm{Flip}(S, T)$ be the set of transpositions swapping one element of S with one element in T. Define the group algebra elements

$$\mathscr{T}_S := \sum_{x \in \mathrm{Flip}(S)} x, \quad \text{and} \quad \mathscr{T}_{S,T} := \sum_{x \in \mathrm{Flip}(S,T)} x. \tag{2.2}$$

We also define the group algebra element

$$\mathscr{A} := \left(\frac{a^2|A| + b^2|B|}{N^2}\right) \cdot \mathrm{id} + \frac{2a^2}{N^2}\mathscr{T}_A + \frac{2b^2}{N^2}\mathscr{T}_B + \frac{2ab}{N^2}\mathscr{T}_{A,B}. \tag{2.3}$$

The proof of Theorem 1.2 involves diagonalizing the linear operator on $\mathbb{C}[\mathfrak{S}_N]$ given by left multiplication by $\mathscr{A}$. The definition of $\mathscr{A}$ in Equation 2.3 exactly encodes the transition probabilities of our card shuffle. Specifically, if we begin with a probability distribution π on $\mathfrak{S}_N$, the distribution after one shuffle will be $\mathscr{A}\pi$. Thus, the eigenvalue spectrum of the transition matrix will be equal to the eigenvalue spectrum of left multiplication by $\mathscr{A}$ on $\mathbb{C}[\mathfrak{S}_N]$. To describe the eigenvalues of our transition matrix, we define $\mathrm{Eig}^{\lambda}_{\mu,\nu}$ as

$$\mathrm{Eig}^{\lambda}_{\mu,\nu} := \frac{a^2|A| + b^2|B|}{N^2} + \frac{2(a^2-ab)}{N^2}\,\mathrm{Diag}(\mu) + \frac{2(b^2-ab)}{N^2}\,\mathrm{Diag}(\nu) + \frac{2ab}{N^2}\,\mathrm{Diag}(\lambda),$$

for all partitions $\lambda \vdash N$, $\mu \vdash |A|$, and $\nu \vdash |B|$. The following result states that the eigenvalues of $\mathscr{A}$ are exactly $\mathrm{Eig}^{\lambda}_{\mu,\nu}$ for various triples of partitions (λ, μ, ν).

Theorem 2.4. *Left multiplication by $\mathscr{A}$ on $\mathbb{C}[\mathfrak{S}_N]$ has eigenvalue spectrum $\mathrm{Eig}^{\lambda}_{\mu,\nu}$ with multiplicities $f_\lambda f_\mu f_\nu c^{\lambda}_{\mu,\nu}$ for all $\lambda \vdash N$, $\mu \vdash |A|$, and $\nu \vdash |B|$.*

Proof. From Maschke's theorem, we can decompose the group algebra $\mathbb{C}[\mathfrak{S}_N]$ as $\mathfrak{S}_N$-representations as

$$\mathbb{C}[\mathfrak{S}_N] = \bigoplus_{\lambda \vdash N} (S^{\lambda})^{\oplus f_\lambda}, \tag{2.5}$$

where S^{λ} is the Specht module corresponding to the partition λ. Since subrepresentations are invariant under left multiplication by $\mathscr{A}$, it is enough to diagonalize $\mathscr{A}$ on S^{λ}. We can rewrite Equation 2.3 as

$$\mathscr{A} = \left(\frac{a^2|A| + b^2|B|}{N^2}\right)\cdot \mathrm{id} + \frac{2(a^2-ab)}{N^2}\mathscr{T}_A + \frac{2(b^2-ab)}{N^2}\mathscr{T}_B + \frac{2ab}{N^2}\mathscr{T}_{A\cup B}. \tag{2.6}$$

It is well known (see [14, Exercise 27.9]) that $\mathscr{T}_{[N]}$ acts on S^{λ} for $\lambda \vdash N$ by scalar multiplication by $\mathrm{Diag}(\lambda)$. Since the outer two terms in Equation 2.6 act by scalar multiplication on S^{λ}, it is enough to diagonalize the sum of the inner two terms. To this end, we view S^{λ} as a $\mathfrak{S}_A \times \mathfrak{S}_B$ representation. It follows from the Littlewood–Richardson rule [27, Proposition 1] that

$$\mathrm{Res}^{\mathfrak{S}_N}_{\mathfrak{S}_A\times\mathfrak{S}_B} S^{\lambda} = \bigoplus_{\substack{\mu\vdash|A|\\ \nu\vdash|B|}} (S^{\mu}\boxtimes S^{\nu})^{\oplus c^{\lambda}_{\mu,\nu}} \tag{2.7}$$

as $\mathfrak{S}_A \times \mathfrak{S}_B$ modules. The sum of the inner two terms in Equation 2.6 act on each copy of $S^{\mu}\boxtimes S^{\nu}$ by scalar multiplication by

$$\frac{2(a^2-ab)}{N^2}\,\mathrm{Diag}(\mu) + \frac{2(b^2-ab)}{N^2}\,\mathrm{Diag}(\nu). \tag{2.8}$$

Adding the scalar which the sum of the outer two terms in Equation 2.6 act on S^{λ}, we get the desired eigenvalue. The multiplicity is given by the fact that $\dim(S^{\mu}\boxtimes S^{\nu}) = f_\mu f_\nu$ and $S^{\mu}\boxtimes S^{\nu}$ appears $f_\lambda c^{\lambda}_{\mu,\nu}$ times in $\mathbb{C}[\mathfrak{S}_N]$. This completes the proof to the theorem. $\square$

Proof of Theorem 1.2. As mentioned in the paragraph after Equation 2.3, the eigenvalue spectrum of the transition matrix is exactly the eigenvalue spectrum of left multiplication of $\mathscr{A}$ on $\mathbb{C}[\mathfrak{S}_N]$. Theorem 1.2 follows immediately from Theorem 2.4. □

3 Lower Bound

In this section, we prove the lower bound of Theorem 1.1. This will follow from the limiting behavior of the number of fixed points described by Theorem 1.3.

3.1 Asymptotics of fixed points

For the rest of this section, let V be the defining representation of $\mathfrak{S}_N$. For any permutation $\pi \in \mathfrak{S}_N$, let $\mathrm{Fix}(\pi)$ be the number of fixed points of π. The following lemma gives a representation-theoretic interpretation of the moments of $\mathrm{Fix}(\pi)$ for a random permutation π.

Lemma 3.1. *Let $\theta : \mathfrak{S}_N \to [0,1]$ be a probability distribution on $\mathfrak{S}_N$ and let $\vartheta \in \mathbb{C}[\mathfrak{S}_N]$ the corresponding element in the group algebra.*

(a) The expected number of fixed points of a permutation picked according to θ is the trace of ϑ on V.

(b) The p^{th} moment of the number of fixed points of a permutation picked according to θ is the trace of ϑ on $V^{\otimes p}$.

Proof. The group algebra element ϑ is given by $\sum \theta(\pi)\pi \in \mathbb{C}[\mathfrak{S}_N]$ where the sum ranges over all permutations. Viewing π as an element of $\mathrm{End}(V)$, the trace $\mathrm{tr}(\pi)$ is equal to the number of fixed points of π. Thus, we have

$$\mathrm{tr}(\vartheta) = \sum_{\pi \in \mathfrak{S}_N} \theta(\pi)\,\mathrm{tr}(\pi) = \sum_{\pi \in \mathfrak{S}_N} \theta(\pi)\,\mathrm{Fix}(\pi) = \mathbb{E}_\theta[\mathrm{Fix}(\pi)]. \tag{3.2}$$

This proves (1). To prove (2), observe that $\mathrm{Fix}(\pi) = \mathrm{tr}(\pi) = \chi_V(\pi)$ where χ_V is the character of V. The character of the tensor product $V^{\otimes p}$ is equal to χ_V^p. Thus, viewing each π as an element of $\mathrm{End}(V^{\otimes p})$, we have

$$\mathrm{tr}(\vartheta) = \sum_{\pi \in \mathfrak{S}_N} \theta(\pi)\,\mathrm{tr}(\pi) = \sum_{\pi \in \mathfrak{S}_N} \theta(\pi)\,\mathrm{Fix}(\pi)^p = \mathbb{E}_\theta[\mathrm{Fix}(\pi)^p]. \tag{3.3}$$

□

To prove Theorem 1.3 we also need the following lemma from [4, Equation 5.5] describing the irreducible subrepresentations of tensor powers of the defining representation. Let $S(p,t)$ denote the **Stirling number of the second kind**, and let $K_{\lambda,\mu}$ denote the **Kostka number**.

Lemma 3.4. *As $\mathfrak{S}_N$-representations, we have the decomposition*

$$V^{\otimes p} = \bigoplus_{\lambda} m_p^{\lambda} S^{\lambda}, \quad \textit{where} \quad m_p^{\lambda} := \sum_{t=0}^{N} S(p,t) K_{\lambda,(N-t,1^t)},$$

We are ready for the proof of Theorem 1.3.

Proof of Theorem 1.3. Since the Poisson distribution is determined by its moments [11, Example 6.3], it is enough to show that the p^{th} moment of Fix_c converges to the p^{th} moment of a $\text{Poiss}(1 + e^c/2)$ random variable. From Lemma 3.1, the pth moment of Fix_c is exactly the trace of $\mathscr{A}^K$ on $V^{\otimes p}$. From Lemma 3.4, Equation 2.7, and Theorem 2.4, we have

$$\operatorname{tr}\left(\mathscr{A}^K\right) = \sum_{\lambda \vdash N} \sum_{\mu,\nu \vdash n} m_p^{\lambda} c_{\mu,\nu}^{\lambda} f^{\mu} f^{\nu} (\text{Eig}_{\mu,\nu}^{\lambda})^K \tag{3.5}$$

$$= \sum_{j=0}^{p} \sum_{\substack{\lambda \vdash N \\ \lambda_1 = N-j}} m_p^{\lambda} \left\{ \sum_{\mu,\nu \vdash n} c_{\mu,\nu}^{\lambda} f^{\mu} f^{\nu} (\text{Eig}_{\mu,\nu}^{\lambda})^K \right\}. \tag{3.6}$$

Since we are fixing p and letting N tend to infinity, there are only finitely many partitions (depending on p) which appear in the sum. These partitions are of the form $\lambda = (N - j, T)$ where $0 \leq j \leq p$, and T is a partition of j. Take any partition $\lambda = (N - j, T)$ of this form. We first compute the asymptotics of

$$\sum_{\mu,\nu \vdash n} c_{\mu,\nu}^{\lambda} f^{\mu} f^{\nu} (\text{Eig}_{\mu,\nu}^{\lambda})^K. \tag{3.7}$$

Suppose that $\mu, \nu \vdash n$ satisfy $c_{\mu,\nu}^{\lambda} > 0$. Then, [24, Lemma 2.52] implies that they must be of the form

$$\mu = (n - i_1, T_1), \quad \nu = (n - i_2, T_2), \quad i_1 + i_2 \leq j, \quad T_1 \vdash i_1, T_2 \vdash i_2.$$

In particular, the number of non-zero summands in Equation 3.7 is bounded above by a constant depending only on p. From [24, Lemma 2.25], we have the asymptotics

$$\text{Eig}_{\mu,\nu}^{\lambda} = 1 - \frac{B_{i_1,i_2,j}}{N} + O\left(\frac{1}{N^2}\right)$$

where $B_{i_1,i_2,j} := i_1 a^2 + i_2 b^2 + (2j - i_1 - i_2)ab$. Using the approximation $1 - x = e^{-x} + O(x^2)$, the contribution of all copies of $S^{\mu} \boxtimes S^{\nu}$ in S^{λ} to the trace is

$$c_{\mu,\nu}^{\lambda} f_{\mu} f_{\nu} (\text{Eig}_{\mu,\nu}^{\lambda})^K = \frac{N^{i_1+i_2}}{2^{i_1+i_2}(i_1)!(i_2)!} f_{T_1} f_{T_2} c_{\mu,\nu}^{\lambda} \cdot \frac{e^{\frac{B_{i_1,i_2,j}}{2b} c}}{N^{\frac{B_{i_1,i_2,j}}{2b}}} + o(1) \tag{3.8}$$

$$= C \cdot N^{\frac{2b(i_1+i_2) - B_{i_1,i_2,j}}{2b}} + o(1) \tag{3.9}$$

$$= C \cdot N^{E_{i_1,i_2,j}} + o(1), \tag{3.10}$$

where we define $E_{i_1,i_2,j} := \frac{2b(i_1+i_2)-B_{i_1,i_2,j}}{2b}$. In Equation 3.9 we consolidated all constants which are of constant order with respect to N in the constant C.

We study the maximum value of $E_{i_1,i_2,j}$ depending on the values of i_1, i_2, j. The following lemma implies that $E_{i_1,i_2,j}$ is will less than or equal to 0, and it will be equal to zero for a single choice of the pair (i_1, i_2).

Lemma 3.11. *Let $a > b$.*

(a) We have $E_{i_1,i_2,j} \leq 0$. If $E_{i_1,i_2,j} = 0$, then $i_1 = 0$, $i_2 = j$, and $B_{i_1,i_2,j} = 2jb$.

(b) Let $i_1 = 0$ and $i_2 = j$. Then $c^{\lambda}_{\mu,\nu} > 0$ if and only if $\mu = (n)$ and $\nu = (n-j, T)$. If this happens, then $c^{\lambda}_{\mu,\nu} = 1$.

From Lemma 3.11, the asymptotics of Equation 3.7 for any partition $\lambda = (N-j, T)$ are given by

$$\sum_{\mu,\nu\vdash n} c^{\lambda}_{\mu,\nu} f_\mu f_\nu \left(\mathrm{Eig}^{\lambda}_{\mu,\nu}\right)^K = \frac{f_T}{j!}\left(\frac{e^c}{2}\right)^j + o(1).$$

From Equation 3.5, we have

$$\mathrm{tr}\left(\mathscr{A}^K\right) = o(1) + \sum_{j=0}^{p}\sum_{T\vdash j}\frac{f_T}{j!}\left(\frac{e^c}{2}\right)^j\sum_{t=j}^{p} S(p,t)\binom{t}{j} f_T \tag{3.12}$$

$$= o(1) + \sum_{t=0}^{p} S(p,t)\sum_{j=0}^{t}\binom{t}{j}\left(\frac{e^c}{2}\right)^j\left[\frac{\sum_{T\vdash j} f_T^2}{j!}\right] \tag{3.13}$$

$$= o(1) + \sum_{t=0}^{p} S(p,t)\left(1+\frac{e^c}{2}\right)^t. \tag{3.14}$$

From [16, Equation 1.3-14], the right hand side of Equation 3.14 is exactly the p^{th} moment of a Poisson random variable of rate $1+\frac{e^c}{2}$. This suffices for the proof. □

Remark 3.15. In Theorem 1.3, it is important that $b < a$. Indeed, the condition that $b < a$ plays an important role in limiting the sources that contribute to the trace of $\mathscr{A}^K$. In contrast, in the case $a = b = 1$ corresponding to the random transposition shuffle, it is shown in [22, Equation 1.6] that

$$\mathrm{Fix}_c \xrightarrow{\text{dist}} \mathrm{Poiss}\left(1+e^c\right). \tag{3.16}$$

This illustrates the clear difference between the $a < b$ case and the $a = b = 1$ case.

3.2 Proof of lower bound

We can now prove the lower bound of Theorem 1.1 by comparing the distribution of our shuffle to the distribution of its fixed points.

Proof of lower bound of Theorem 1.1. For $0 \leq k \leq N$, let $\mathfrak{S}_N^{(k)}$ be the subset of permutations with exactly k fixed points. Then we have

$$\begin{aligned}
d_{\mathsf{TV}}(U, P^t(\mathrm{id}, \cdot)) &= \frac{1}{2}\sum_{k=0}^{N} \sum_{\pi \in \mathfrak{S}_N^{(k)}} |U(\pi) - P^t(\mathrm{id}, \pi)| \\
&\geq \frac{1}{2}\sum_{k=0}^{N} \left| U(\mathfrak{S}_N^{(k)}) - P^t(\mathrm{id}, \mathfrak{S}_N^{(k)}) \right| \\
&= d_{\mathsf{TV}}(\mathrm{Fix}, \mathrm{Fix}_c) \\
&= d_{\mathsf{TV}}\left(\mathrm{Poiss}(1), \mathrm{Poiss}\left(1 + \frac{e^c}{2}\right)\right) + o(1),
\end{aligned}$$

where the last equality follows from Theorem 1.3 and the classical *problème des rencontres*. The lower bound follows from [20, p. 44] giving

$$d_{\mathsf{TV}}\left(\mathrm{Poiss}(1), \mathrm{Poiss}\left(1 + \frac{e^c}{2}\right)\right) \geq H^2\left(\mathrm{Poiss}(1), \mathrm{Poiss}\left(1 + \frac{e^c}{2}\right)\right), \tag{3.17}$$

where H^2 is the **(squared) Hellinger distance**. The right hand side of the lower bound of Theorem 1.1 is an explicit computation of the Hellinger distance between two Poisson random variables. $\square$

4 Upper Bound

In this section, we make a few comments on the upper bound of Theorem 1.1. From [21, Lemma 12.18(ii)] along with Theorem 1.2, we get the following upper bound.

Lemma 4.1. *Let $|A| = |B| = n$ and $N = 2n$. Let U be the uniform distribution on $\mathfrak{S}_N$. Then, any time $t \geq 0$, we have the bound*

$$d_{\mathsf{TV}}(P^t(x, \cdot), U)^2 \leq \frac{1}{4} \sum_{\substack{\lambda \vdash N \\ \lambda \neq (N)}} \sum_{\lambda : c^{\lambda}_{\mu,\nu} > 0} c^{\lambda}_{\mu,\nu} f_\lambda f_\mu f_\nu \cdot |\operatorname{Eig}^{\lambda}_{\mu,\nu}|^{2t}. \tag{4.2}$$

We prove the upper bound in Theorem 1.1 by splitting up the sum in Equation 4.2 into several zones based on the indexing partition. Approximately, we split up the space of partitions of N into approximately three zones. These zones are as follows:

Definition 4.3. Let $a^* := 2 - a^{-1}$. We define the **Red**, **Blue**, and **Yellow** zones explicitly by

$$\begin{aligned}
\text{Zone}_R &:= \{\lambda : \lambda_1, \lambda_1^* \le n\},\\
\text{Zone}_B &:= \{\lambda : n \le \lambda_1 \le a^* n, \lambda_1^* \le n\} \cup \{\lambda : n \le \lambda_1^* \le a^* n, \lambda_1 \le n\},\\
\text{Zone}_Y &:= \{\lambda : 2n - 1 \ge \lambda_1 \ge a^* n\} \cup \{\lambda : \lambda_1^* \ge a^* n\}.
\end{aligned}$$

Diagrammatically, these zones are depicted in Figure 2 where the space of partitions is represented as a two-dimensional plane where the x-axis corresponds to the length of the first row and the y-axis corresponds to the length of the first column. In the analysis, we also partition each zone into several subzones.

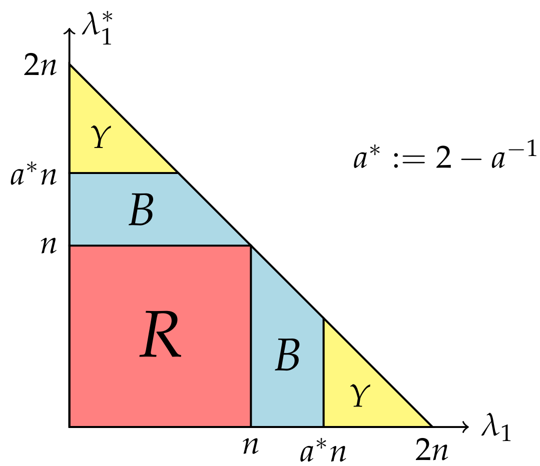

Figure 2: Cartesian plane with x-axis λ_1 and y-axis λ_1^*

Over each sub-zone, we bound the maximum value of $\text{Eig}_{\mu,\nu}^{\lambda}$ and the dimension of $c_{\mu,\nu}^{\lambda} f_\lambda f_\mu f_\nu$ depending on the length of the first row and first column. The main terms in the sum come from partitions with long first row. To show that the contribution coming from the other partitions is negligible, we need a good understanding of when $c_{\mu,\nu}^{\lambda} > 0$. Equivalently, we need to understand the triples (λ, μ, ν) which appear non-trivially in Equation 4.2. To do this, we rely on the theory of hives [18] and more generally the representation theory of semisimple Lie algebras.

Remark 4.4. For random transpositions, Diaconis and Shahshahani use zones similar to Figure 2. However, in the biased case, the analysis is considerably more delicate and we need to divide the red, blue, and yellow zones into *additional* sub-zones. Some of these sub-zones then need to be divided into even smaller sub-zones, where these sub-zones are determined *dynamically*. For more details of this analysis, we direct the reader to the preprint [24].

Acknowledgements

The authors thank Ezra Edelman for introducing us to the Hellinger distance. The authors thank Lauren Williams for helpful comments on the extended abstract. The authors also thank Johnny Gao, Laura Colmenarejo, and Grant Barkley for enlightening conversations about Littlewood–Richardson coefficients and the hive model. Finally, the authors thank the anonymous reviewers for their helpful comments.

References

[1] D. Aldous and P. Diaconis. "Shuffling cards and stopping times". *Amer. Math. Monthly* **93**.5 (1986), pp. 333–348. DOI.

[2] M. E. Bate, S. B. Connor, and O. Matheau-Raven. "Cutoff for a one-sided transposition shuffle". *Ann. Appl. Probab.* **31**.4 (2021), pp. 1746–1773. DOI.

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

[4] G. Benkart, T. Halverson, and N. Harman. "Dimensions of irreducible modules for partition algebras and tensor power multiplicities for symmetric and alternating groups". *J. Algebraic Combin.* **46**.1 (2017), pp. 77–108. DOI.

[5] N. Berestycki and R. Durrett. "A phase transition in the random transposition random walk". *Probab. Theory Related Fields* **136**.2 (2006), pp. 203–233. DOI.

[6] N. Berestycki and B. ¸Sengül. "Cutoff for conjugacy-invariant random walks on the permutation group". *Probab. Theory Related Fields* **173**.3-4 (2019), pp. 1197–1241. DOI.

[7] M. Bernstein, N. Bhatnagar, and I. Pak. "Cutoff for biased transpositions". 2017. arXiv: 1709.03477.

[8] M. Bernstein and E. Nestoridi. "Cutoff for random to random card shuffle". *Ann. Probab.* **47**.5 (2019), pp. 3303–3320. DOI.

[9] P. Bidigare, P. Hanlon, and D. Rockmore. "A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements". *Duke Math. J.* **99**.1 (1999), pp. 135–174. DOI.

[10] K. S. Brown and P. Diaconis. "Random walks and hyperplane arrangements". *Ann. Probab.* **26**.4 (1998), pp. 1813–1854. DOI.

[11] A. DasGupta. *Asymptotic theory of statistics and probability*. Springer Texts in Statistics. Springer, New York, 2008, pp. xxviii+722.

[12] P. Diaconis and M. Shahshahani. "Generating a random permutation with random transpositions". *Z. Wahrsch. Verw. Gebiete* **57**.2 (1981), pp. 159–179. DOI.

[13] A. B. Dieker and F. V. Saliola. "Spectral analysis of random-to-random Markov chains". *Adv. Math.* **323** (2018), pp. 427–485. DOI.

[14] P. Etingof. "Lie groups and Lie algebras". 2024. arXiv:2201.09397.

[15] L. Flatto, A. M. Odlyzko, and D. B. Wales. "Random shuffles and group representations". *Ann. Probab.* **13**.1 (1985), pp. 154–178. Link.

[16] F. A. Haight. *Handbook of the Poisson distribution*. Vol. No. 11. Publications in Operations Research. John Wiley & Sons, Inc., New York-London-Sydney, 1967, pp. xi+168.

[17] B. Hough. "The random k cycle walk on the symmetric group". *Probab. Theory Related Fields* **165**.1-2 (2016), pp. 447–482. DOI.

[18] A. Knutson and T. Tao. "The honeycomb model of $\mathrm{GL}_n(\mathbf{C})$ tensor products. I. Proof of the saturation conjecture". *J. Amer. Math. Soc.* **12**.4 (1999), pp. 1055–1090. DOI.

[19] H. Lacoin. "Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion". *Ann. Probab.* **44**.2 (2016), pp. 1426–1487. DOI.

[20] L. Le Cam and G. L. Yang. *Asymptotics in statistics*. Second. Springer Series in Statistics. Some basic concepts. Springer-Verlag, New York, 2000, pp. xiv+285. DOI.

[21] D. A. Levin and Y. Peres. *Markov chains and mixing times*. Second. With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson. American Mathematical Society, Providence, RI, 2017, pp. xvi+447. DOI.

[22] P. Matthews. "A strong uniform time for random transpositions". *J. Theoret. Probab.* **1**.4 (1988), pp. 411–423. DOI.

[23] E. Nestoridi and K. Peng. "Mixing times of one-sided k-transposition shuffles". 2021. arXiv:2112.05085.

[24] E. Nestoridi and A. Yan. "Cutoff for the Biased Random Transposition Shuffle". 2024. arXiv:2409.16387.

[25] J. Salez. "Cutoff for non-negatively curved Markov chains". *J. Eur. Math. Soc. (JEMS)* **26**.11 (2024), pp. 4375–4392. DOI.

[26] G. R. White. *Combinatorial Methods in Markov Chain Mixing*. Thesis (Ph.D.)–Stanford University. ProQuest LLC, Ann Arbor, MI, 2017, p. 69. Link.

[27] A. V. Zelevinsky. "A generalization of the Littlewood-Richardson rule and the Robinson-Schensted-Knuth correspondence". *J. Algebra* **69**.1 (1981), pp. 82–94. DOI.