Article:

Bate, Michael orcid.org/0000-0002-6513-2405, Connor, Stephen orcid.org/0000-0002-9785-2159 and Matheau-Raven, Oliver (2021) Cutoff for a One-sided Transposition Shuffle. The Annals of Applied Probability. pp. 1746-1773. ISSN 1050-5164

Cutoff for a One-sided Transposition Shuffle

Michael E. Bate, Stephen B. Connor, and Oliver Matheau-Raven*
University of York, UK

October 20, 2020

Abstract

We introduce a new type of card shuffle called a *one-sided transposition shuffle*. At each step a card is chosen uniformly from the pack and then transposed with another card chosen uniformly from *below* it. This defines a random walk on the symmetric group generated by a distribution which is non-constant on the conjugacy class of transpositions. Nevertheless, we provide an explicit formula for all eigenvalues of the shuffle by demonstrating a useful correspondence between eigenvalues and standard Young tableaux. This allows us to prove the existence of a total-variation cutoff for the one-sided transposition shuffle at time $n \log n$. We also study a weighted generalisation of the shuffle which, in particular, allows us to recover the well known mixing time of the classical random transposition shuffle.

Keywords and phrases:
Mixing time; Cutoff phenomenon; Coupon collecting; Representation theory; Young tableaux

2010 Mathematics Subject Classification:
Primary: 60J10
Secondary: 20C30

1 Introduction

Consider a stacked deck of n distinct cards, whose positions are labelled by elements of the set $[n] \coloneqq \{1, \dots, n\}$ from bottom to top. Any shuffle which involves choosing two positions and switching the cards found there is called a transposition shuffle, and may be viewed as a random walk on the symmetric group S_n. (If the two positions coincide then no cards are moved.) Diaconis and Shahshahani [6] were the first to study transposition shuffles using the representation theory of S_n; they famously showed that the *random transposition shuffle*, in which the two positions are chosen independently and uniformly on $[n]$, takes $(n/2) \log n$ steps to randomise the order of the deck. (This is known as the *mixing time* of the shuffle.) The use of representation theory, both in [6] and successive works (e.g. [1, 3]), relies heavily on the fact that the distribution generating the shuffle is constant on conjugacy classes of S_n.

Since then a variety of algebraic and probabilistic techniques have been employed to study different types of transposition shuffle. Notable examples include the "transpose top and random" shuffle [16], the "adjacent transpositions" shuffle [9], and a generalisation of the latter in which the two cards are constrained to lie within a certain (cyclical) distance of one another [2]. All of these shuffles have the property that, at each step, the transposition to be applied is chosen uniformly from within some set which generates the entire group S_n.

Semi-random transposition shuffles form an interesting class of Markov chains (see e.g. [12–14]). In this class the right hand chooses a card uniformly at random, the left hand chooses a card via some (possibly time-inhomogeneous) independent stochastic process, and then the two chosen cards are transposed. A universal upper bound of $O(n \log n)$ on the mixing time of any semi-random transposition shuffle was established by Mossel, Peres, and Sinclair [12].

*PhD funded by the EPSRC grant: EP/N509802/1

In the present work we introduce a new class of shuffles called *one-sided transposition shuffles*: these have the defining property that at step i the right hand's position (R^i) is chosen according to some arbitrary distribution on $[n]$, and given the value of R^i the distribution of the left hand's position (L^i) is supported on the set $\{1,\ldots,R^i\}$. For the majority of this paper we shall focus on the case when the right and left hands are both chosen uniformly from their possible ranges, but in Section 4 we shall extend our main results to the case when R^i is chosen using a particular type of weighted distribution.

Our setup differs significantly from the previously studied shuffles mentioned above. The dependence between the left and right hands means that this does not fall into the class of semi-random transpositions. Furthermore, although the generating set for our shuffle is the entire conjugacy class of transpositions, the distribution that we impose upon this set is in general far from uniform. (E.g. when right and left hands are both uniform on their permitted ranges, the probabilities attached to different transpositions range from $1/n^2$ to $1/2n$; see Definition 1 below.) We note that there can clearly be no universal upper bound on the mixing time of these shuffles without imposing further constraints on the distribution of the left hand, since the shuffle can be slowed arbitrarily by increasing the probability that the two hands choose the same position.

In order to state our results we briefly introduce some notation and terminology.

Definition 1. *The (unbiased) one-sided transposition shuffle P_n is the ergodic random walk on S_n generated by the following distribution on the conjugacy class of transpositions:*

$$P_n(\tau) = \begin{cases} \frac{1}{n}\cdot\frac{1}{j} & \text{if } \tau = (i\,j) \text{ for some } 1 \le i \le j \le n \\ 0 & \text{otherwise.} \end{cases}$$

We use the convention that all 'transpositions' $(i\,i)$ are equal to the identity element id, and therefore $P_n(id) = \frac{1}{n}(1+\frac{1}{2}+\cdots+\frac{1}{n}) = \frac{H_n}{n}$, where H_k denotes the k^{th} harmonic number. We write P_n^t for the t–fold convolution of P_n with itself.

This shuffle is clearly both reversible and transitive, and has stationary distribution equal to the uniform distribution on S_n, denoted π_n; that is, $P_n^t(\sigma) \to 1/n!$ as $t \to \infty$ for all $\sigma \in S_n$. In order to study the rate of this convergence, we begin by recalling that the total variation distance between two probability distributions P and Q on S_n is defined by

$$\|P - Q\|_{\text{TV}} = \sup_{A\subseteq S_n} |P(A) - Q(A)| = \frac{1}{2}\sum_{\sigma\in S_n} |P(\sigma) - Q(\sigma)|.$$

Now consider a sequence of distributions $\{Q_n\}_{n\in\mathbb{N}}$ on state spaces $\{S_n\}_{n\in\mathbb{N}}$ with corresponding stationary distributions $\{\pi_n\}_{n\in\mathbb{N}}$. We may define the total variation mixing time $t_n^{\text{mix}}(\varepsilon)$ for the distribution Q_n as follows:

$$t_n^{\text{mix}}(\varepsilon) = \min\{t : \|Q_n^t - \pi_n\|_{\text{TV}} < \varepsilon\}.$$

It is well known that many natural sequences of this kind exhibit behaviour known as a *cutoff phenomenon*, whereby the convergence to equilibrium occurs more and more sharply as $n \to \infty$.

Definition 2. *A sequence of distributions $\{Q_n\}$ exhibits a (total variation)* cutoff *at time $\{t_n\}$ with a* window *of size $\{w_n\}$ if $w_n = o(t_n)$ and the following limits hold:*

$$\lim_{c\to\infty}\limsup_{n\to\infty}\|Q_n^{t_n+cw_n} - \pi_n\|_{\text{TV}} = 0$$
$$\lim_{c\to\infty}\liminf_{n\to\infty}\|Q_n^{t_n-cw_n} - \pi_n\|_{\text{TV}} = 1\,.$$

Existence of a cutoff implies that $t_n^{\text{mix}}(\varepsilon) \sim t_n$ for all $\varepsilon \in (0,1)$. The main conclusion of this work is that the one-sided transposition shuffle exhibits a cutoff at time $t_n = n\log n$.

Theorem 3. *The one-sided transposition shuffle P_n exhibits a cutoff at time $n\log n$. Specifically, for any $c_1 > 0$ and $c_2 > 2$*

$$\limsup_{n\to\infty}\|P_n^{n\log n + c_1 n} - \pi_n\|_{\text{TV}} \le \sqrt{2}e^{-c_1}\,, \tag{1}$$

and $$\liminf_{n\to\infty}\|P_n^{n\log n - n\log\log n - c_2 n} - \pi_n\|_{\text{TV}} \ge 1 - \frac{\pi^2}{6(c_2-2)^2}\,. \tag{2}$$

The lower bound on the mixing time in (2) will be obtained via a coupling argument which allows us to compare the one-sided transposition shuffle to a variation of a coupon-collecting problem. To establish the upper bound for the cutoff in Theorem 3 we shall make use of a classical ℓ^2 bound on total variation distance.

Lemma 4 (Lemma 12.16 [11])**.** *Let Q be the transition matrix for a reversible, transitive, aperiodic Markov chain on finite state space $\mathcal{X}$, with stationary distribution π. Let the eigenvalues of Q be denoted β_i, with $1 = \beta_1 > \beta_2 \geq \cdots \geq \beta_{|\mathcal{X}|} > -1$. Then*

$$4\|Q^t - \pi\|_{\mathrm{TV}}^2 \leq \sum_{i\neq 1} \beta_i^{2t}.$$

The spectrum of the one-sided transposition shuffle will be analysed using the representation theory of the symmetric group, and our method exhibits some algebraic features which are of independent interest. The key result is describing an explicit method for obtaining the eigenvectors of P_{n+1} from those of P_n; this is called *lifting* eigenvectors. It is interesting that the technique of lifting eigenvectors allows the use of representation theory to analyse shuffles which are not constant on conjugacy classes. In order to work, this technique requires a particularly close relationship between the shuffle on n and $n+1$ cards – in our examples of one-sided transposition shuffles on $n+1$ cards, the $(n+1)$-th card only moves when $R^i = n+1$, and for all other choices of R^i the shuffle on $n+1$ cards behaves like the shuffle on n cards (see equation (29) for more details). An analysis involving the lifting of eigenvectors was first used in the recent work of Dieker and Saliola [7] which studied the eigenspaces of the *random-to-random shuffle*; this analysis was used by Bernstein and Nestoridi [4] to prove the existence of a cutoff for this shuffle at time $(3/4)n \log n$. Lafrenière [10] subsequently showed that similar lifting techniques could be applied to more general "symmetrized shuffling operators". In this paper we make several non-trivial changes to the technique developed by Dieker and Saliola in order to employ it in the analysis of the one-sided transposition shuffle: we believe that this is the first time such a technique has been shown to be applicable to either a non-symmetrized shuffle or a transposition shuffle. We suspect that with suitable modifications the technique of lifting eigenvectors may be used to analyse a whole variety of shuffles for which the standard technique of discrete Fourier transforms fails.

Using this method we prove that each eigenvalue of P_n corresponds to a *standard Young tableau*, and may be computed explicitly from the entries in the tableau. That we are able to find such explicit results is remarkable given that the distribution generating the shuffle is not constant on conjugacy classes. Definitions, notation and proofs will be carefully laid out below, but for ease of reference we state here the main two results which will underpin our analysis, proofs of which can be found in the Appendix:

Theorem 5. *The eigenvalues of P_n are labelled by standard Young tableaux of size n, and the eigenvalue represented by a tableau of shape λ appears d_λ times, where d_λ is the dimension of λ.*

Lemma 6. *For a tableau T of shape λ the eigenvalue corresponding to T is given by*

$$\operatorname{eig}(T) = \frac{1}{n} \sum_{\substack{\text{boxes} \\ (i,j)}} \frac{j-i+1}{T(i,j)}, \tag{3}$$

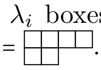
where the sum is performed over all boxes (i,j) in T.

The organisation of the remainder of the paper is as follows. In Section 2 we first explore some important properties of the eigenvalues for P_n, and then use these to prove the upper bound on the mixing time given by Theorem 3. The corresponding lower bound will be proved in Section 3, using entirely probabilistic arguments. Finally, in Section 4 we will consider a generalisation of the one-sided transposition shuffle, in which R^i is chosen according to a non-uniform distribution: we show that the algebraic technique developed for P_n holds in this more general setting, and that the well-known mixing time for the (standard) random transposition shuffle may be recovered in this way.

2 Upper Bound

2.1 Eigenvalue Analysis

Before we establish relations of the eigenvalues for P_n we first will recall some standard definitions about partitions, Young diagrams and Young tableaux [8]. A *partition* of n is a tuple of positive integers $\lambda = (\lambda_1, \ldots, \lambda_r)$ such that, $\sum \lambda_i = n$, and $\lambda_1 \geq \cdots \geq \lambda_r$. We write $\lambda \vdash n$. We call r the *length* of λ and denote it $l(\lambda)$. We denote the partition $(1, \ldots, 1) \vdash n$ as (1^n).

Every partition λ has an associated *Young diagram*, made by forming a left adjusted stack of boxes with rows labelled downwards and row i having λ_i boxes. We often blur the distinction between a partition and its Young diagram, e.g. $(4,2) =$ . We may refer to the boxes of a diagram λ by using coordinates (i,j) to mean the box in the i^{th} row and j^{th} column. Given a partition $\lambda \vdash n$, we may form the *transpose* of λ, denoted λ', by turning rows into columns in the Young diagram. We have $\lambda \vdash n$ if and only if $\lambda' \vdash n$.

We have a partial order on partitions of n called the *dominance order*: in terms of Young diagrams, for two partitions $\mu, \lambda \vdash n$, we write $\lambda \trianglerighteq \mu$ if we can form μ by moving boxes of λ down and to the left. We have $\lambda \trianglerighteq \mu$ if and only if $\mu' \trianglerighteq \lambda'$ [8, Lemma 1.4.11].

Given two partitions μ, λ of different sizes, we say $\mu \subset \lambda$ if μ is fully contained in λ when we align the Young diagrams of μ and λ at the top left corners; equivalently, if we write $\lambda = (\lambda_1, \ldots, \lambda_r)$ and $\mu = (\mu_1, \ldots, \mu_s)$, this means that $s \leq r$ and $\mu_i \leq \lambda_i$ for each $1 \leq i \leq s$.

Given a partition $\lambda \vdash n$, we may form a *Young tableau* T by putting numbers into the boxes of (the Young diagram of) λ. A *standard Young tableau* T is one in which the numbers $1, \ldots, n$ occur once each and such that the values are increasing across rows and down columns. The set of standard Young tableaux of shape λ is denoted by $\mathrm{SYT}(\lambda)$. The size of the set $\mathrm{SYT}(\lambda)$ is called the *dimension* of λ, denoted d_λ. For a tableau T, form the transpose of T denoted T' by turning rows into columns and preserving the value in each box. If $T \in \mathrm{SYT}(\lambda)$, then $T' \in \mathrm{SYT}(\lambda')$. Given a tableau T, let $T(i,j)$ be the value of box (i,j) in T if it exists and undefined otherwise.

For any $\lambda \vdash n$, define the tableau $T_\lambda^{\rightarrow}$ by inserting the numbers $1, \ldots, n$ from left to right. Define the tableau $T_\lambda^{\downarrow}$ by inserting the numbers $1, \ldots, n$ from top to bottom. From the introduction we know the eigenvalues for P_n are labelled by Young tableaux of size n, and Lemma 6 gives an explicit formula for the eigenvalue associated to a given tableau. Before applying the bound in Lemma 4, we first investigate relationships between the eigenvalues. We show that the eigenvalue corresponding to $T \in \mathrm{SYT}(\lambda)$ is bounded by the eigenvalues for $T_\lambda^{\rightarrow}$ and $T_\lambda^{\downarrow}$. To simplify our upper bound calculation, we prove that we only need to consider the partitions for which $T_\lambda^{\rightarrow}$ gives a positive eigenvalue. Lastly, we prove that the eigenvalues corresponding to $T_\lambda^{\rightarrow}, T_\lambda^{\downarrow}$ decrease as one moves down the dominance order of partitions. We first illustrate some of the preceding definitions and discussion with an example.

Example 7. *Let $\lambda = (3,2) \vdash 5$. Then $\mathrm{SYT}(\lambda)$ has five elements, so $d_\lambda = 5$. These 5 tableaux, together with the associated eigenvalues calculated using Lemma 6, are given in Table 1 below. In this table, $T_\lambda^{\rightarrow}$ is the first tableau listed and $T_\lambda^{\downarrow}$ is the last one; we can see that the corresponding eigenvalues bound all the others.*

$T \in \mathrm{SYT}((3,2))$	1 2 3 / 4 5	1 2 4 / 3 5	1 2 5 / 3 4	1 3 4 / 2 5	1 3 5 / 2 4
$\mathrm{eig}(T)$	0.64	0.59	0.57	$0.52\overline{3}$	$0.50\overline{3}$

Table 1: Eigenvalues corresponding to $T \in \mathrm{SYT}((3,2))$.

We can now begin our analysis of the eigenvalues by showing how swapping numbers in a tableau affects the corresponding eigenvalue.

Lemma 8. *Let T be a general Young tableau. Suppose we form a new tableau S by swapping two values in T which have coordinates $(i_1, j_1), (i_2, j_2)$ in T. WLOG assume $T(i_1, j_1) < T(i_2, j_2)$.*

Then the change in corresponding eigenvalues satisfies the following inequality:

$$\mathrm{eig}(S) - \mathrm{eig}(T) \begin{cases} \geq 0 & \text{if } (i_1 - i_2) + (j_2 - j_1) \geq 0 \\ < 0 & \text{if } (i_1 - i_2) + (j_2 - j_1) < 0\,. \end{cases}$$

Importantly, if we move the larger entry down and to the left the change in eigenvalue is non-negative; if it moves up and to the right then the change is negative.

Proof. Since S and T agree in all but two entries, Lemma 6 tells us that the difference in eigenvalues is given by

$$\begin{aligned} \mathrm{eig}(S) - \mathrm{eig}(T) &= \frac{1}{n}\left(\frac{j_1 - i_1 + 1}{T(i_2, j_2)} + \frac{j_2 - i_2 + 1}{T(i_1, j_1)}\right) - \frac{1}{n}\left(\frac{j_1 - i_1 + 1}{T(i_1, j_1)} + \frac{j_2 - i_2 + 1}{T(i_2, j_2)}\right) \\ &= \frac{(i_1 - i_2) + (j_2 - j_1)}{n}\left(\frac{1}{T(i_1, j_1)} - \frac{1}{T(i_2, j_2)}\right). \end{aligned}$$

□

This result allows us to prove that the eigenvalues for $T^{\downarrow}_{\lambda}$ and $T^{\rightarrow}_{\lambda}$ bound all others for $\mathrm{SYT}(\lambda)$.

Lemma 9. *Let $\lambda \vdash n$. For any $T \in \mathrm{SYT}(\lambda)$ we have the following inequality:*

$$\mathrm{eig}(T^{\downarrow}_{\lambda}) \leq \mathrm{eig}(T) \leq \mathrm{eig}(T^{\rightarrow}_{\lambda}). \tag{4}$$

Proof. Reading across the rows of T, beginning with the first row, identify the first box in which T and $T^{\rightarrow}_{\lambda}$ have different entries; write (i,j) for the coordinates of this box. Due to the way in which $T^{\rightarrow}_{\lambda}$ is constructed, $T(i,j) > T^{\rightarrow}_{\lambda}(i,j)$. Furthermore, the number $T(i,j) - 1$ must occur strictly below and to the left of $T(i,j)$, since T is a standard Young tableau. Swapping entries $T(i,j) - 1$ and $T(i,j)$ produces a new element of $\mathrm{SYT}(\lambda)$ whose corresponding eigenvalue is no smaller than $\mathrm{eig}(T)$, thanks to Lemma 8.

We can therefore iterate this procedure, swapping $T(i,j) - 1$ with $T(i,j) - 2$ etc, until $T(i,j) = T^{\rightarrow}_{\lambda}(i,j)$. Note that at this point the entries in the first $T(i,j)$ boxes of T and $T^{\rightarrow}_{\lambda}$ must agree, moreover these entries are now fixed in place. We now proceed to the next box in which T and $T^{\rightarrow}_{\lambda}$ differ, and repeat: this results in a sequence of swaps which make the entries of T agree with those in $T^{\rightarrow}_{\lambda}$, and which can only ever cause the corresponding eigenvalue to increase. This proves the second inequality in Lemma 9, and the first one follows via an analogous argument on the columns of T.

□

The next result and its corollary establish that when bounding eigenvalues, we only need to consider those given by $T^{\rightarrow}_{\lambda}$.

Lemma 10. *Let $\lambda \vdash n$. For any $T \in \mathrm{SYT}(\lambda)$ we have*

$$\mathrm{eig}(T) + \mathrm{eig}(T') = \frac{2H_n}{n}.$$

Proof. Let $T \in SYT(\lambda)$. Then

$$\begin{aligned} \mathrm{eig}(T) + \mathrm{eig}(T') &= \frac{1}{n}\sum_{\substack{\text{boxes}\\(i,j)\in T}} \frac{j - i + 1}{T(i,j)} + \frac{1}{n}\sum_{\substack{\text{boxes}\\(j,i)\in T'}} \frac{i - j + 1}{T'(j,i)} \\ &= \frac{1}{n}\sum_{\substack{\text{boxes}\\(i,j)\in T}} \frac{j - i + 1}{T(i,j)} + \frac{1}{n}\sum_{\substack{\text{boxes}\\(i,j)\in T}} \frac{-(j - i) + 1}{T(i,j)} = \frac{2H_n}{n}. \end{aligned}$$

□

Corollary 11. *Let $\lambda = (\lambda_1, \ldots, \lambda_r) \vdash n$, and suppose we have $\mathrm{eig}(T^{\downarrow}_{\lambda}) \leq 0$, then we have*

$$\mathrm{eig}(T^{\rightarrow}_{\lambda'}) \geq |\mathrm{eig}(T^{\downarrow}_{\lambda})| \geq 0. \tag{5}$$

Proof. It follows from Lemma 10 that $\mathrm{eig}(T^{\rightarrow}_{\lambda'}) + \mathrm{eig}(T^{\downarrow}_{\lambda}) = 2H_n/n$. Thus if $\mathrm{eig}(T^{\downarrow}_{\lambda}) \le 0$ then

$$\mathrm{eig}(T^{\rightarrow}_{\lambda'}) = \frac{2H_n}{n} - \mathrm{eig}(T^{\downarrow}_{\lambda}) \ge -\mathrm{eig}(T^{\downarrow}_{\lambda}) = |\mathrm{eig}(T^{\downarrow}_{\lambda})| \ge 0\,.$$

□

We end this section by establishing a relationship between eigenvalues and the dominance ordering on partitions.

Lemma 12. *Let $\lambda, \mu \vdash n$. If $\lambda \trianglerighteq \mu$ then*

$$\begin{aligned} \mathrm{eig}(T^{\rightarrow}_{\lambda}) &\ \ge\ \mathrm{eig}(T^{\rightarrow}_{\mu}) & (6)\\ \textit{and}\quad \mathrm{eig}(T^{\downarrow}_{\lambda}) &\ \ge\ \mathrm{eig}(T^{\downarrow}_{\mu}). & (7)\end{aligned}$$

Proof. If we can show the statements hold for any partition μ which is formed from λ by moving only one box then inductively it will hold for all $\lambda \trianglerighteq \mu$. Suppose μ is formed from λ by moving a box from row a to row b, with $a < b \le l(\lambda) + 1$ (if $b = l(\lambda) + 1$ then a new row is created by placing the removed box on the very bottom of the diagram). The box we move goes from coordinates (a, λ_a) of λ to $(b, \lambda_b + 1)$ in μ.

We shall prove first that $\mathrm{eig}(T^{\rightarrow}_{\lambda}) \ge \mathrm{eig}(T^{\rightarrow}_{\mu})$. Since $T^{\rightarrow}_{\lambda}$ and $T^{\rightarrow}_{\mu}$ are both numbered from left to right, the effect of moving a box from row a to row b is that $T^{\rightarrow}_{\mu}(i,j) = T^{\rightarrow}_{\lambda}(i,j) - 1$ for any box (i,j) with $a < i \le b$; boxes in all other rows contain the same values in both tableaux. Using equation (3), and remembering to include a term to account for the box being moved, we find that:

$$\begin{aligned} n(\mathrm{eig}(T^{\rightarrow}_{\lambda}) - \mathrm{eig}(T^{\rightarrow}_{\mu})) &= \left(\frac{\lambda_a - a + 1}{T^{\rightarrow}_{\lambda}(a,\lambda_a)} - \frac{(\lambda_b + 1) - b + 1}{T^{\rightarrow}_{\mu}(b,\lambda_b+1)}\right) + \sum_{\substack{(i,j)\in T_\lambda \cap T_\mu \\ \text{with } a<i\le b}} \left[\frac{1}{T^{\rightarrow}_{\lambda}(i,j)} - \frac{1}{T^{\rightarrow}_{\mu}(i,j)}\right](j - i + 1)\\ &\ge \left(\frac{\lambda_a - a + 1}{T^{\rightarrow}_{\lambda}(a,\lambda_a)} - \frac{(\lambda_b + 1) - b + 1}{T^{\rightarrow}_{\mu}(b,\lambda_b+1)}\right) + (\lambda_a - a + 1)\left(\frac{1}{T^{\rightarrow}_{\mu}(b,\lambda_b+1)} - \frac{1}{T^{\rightarrow}_{\lambda}(a,\lambda_a)}\right)\\ &= \frac{(\lambda_a - \lambda_b) + (b-a) - 1}{T^{\rightarrow}_{\mu}(b,\lambda_b+1)} \ge 0. \end{aligned}$$

The first inequality holds because all the square-bracketed terms in the sum are negative; we upper bound $j - i + 1 \le \lambda_a - a + 1$, and the resulting sum telescopes. The final inequality holds because $(\lambda_a - \lambda_b) \ge 1$ and $(b - a) \ge 1$.

For the second inequality, recall that $\lambda \trianglerighteq \mu$ if and only if $\mu' \trianglerighteq \lambda'$. Therefore, using the first established inequality we find that $\mathrm{eig}(T^{\rightarrow}_{\mu'}) \ge \mathrm{eig}(T^{\rightarrow}_{\lambda'})$. Now Lemma 10 gives $-\mathrm{eig}(T^{\downarrow}_{\mu}) \ge -\mathrm{eig}(T^{\downarrow}_{\lambda})$ and thus we recover the desired inequality. □

2.2 Upper Bound Analysis

In this section we complete the proof of the upper bound of Theorem 3, making use of the results of Section 2.1. The analysis splits into two parts, dealing separately with those partitions λ having either "large" or "small" first row.

Lemma 4 allows us to upper bound the total variation distance in terms of the non-trivial eigenvalues of the transition matrix. Using Lemma 6 we see that the trivial eigenvalue corresponds to the one-dimensional partition $\lambda = (n)$, and so Theorem 5 implies that

$$4\|P^t_n - \pi_n\|^2_{\mathrm{TV}} \le \sum_{\substack{\lambda\vdash n\\ \lambda\neq(n)}} d_\lambda \sum_{T\in SYT(\lambda)} \mathrm{eig}(T)^{2t}\,.$$

Recall from Lemma 9 that for any $T \in \mathrm{SYT}(\lambda)$ the eigenvalue corresponding to T may be bounded by those corresponding to $T^{\downarrow}_{\lambda}$ and $T^{\rightarrow}_{\lambda}$. With this in mind, we let $\Lambda^{\rightarrow}_n = \{\lambda \vdash n : |\mathrm{eig}(T^{\downarrow}_{\lambda})| \le \mathrm{eig}(T^{\rightarrow}_{\lambda})\}$ and $\Lambda^{\downarrow}_n = \{\lambda \vdash n : |\mathrm{eig}(T^{\downarrow}_{\lambda})| > |\mathrm{eig}(T^{\rightarrow}_{\lambda})|\}$; note that these are disjoint sets, with

$\Lambda_n^{\rightarrow} \subseteq \{\lambda \vdash n : \mathrm{eig}(T_\lambda^{\rightarrow}) \geq 0\}$ and $\Lambda_n^{\downarrow} \subseteq \{\lambda \vdash n : \mathrm{eig}(T_\lambda^{\downarrow}) < 0\}$. Using Lemma 9 and then Corollary 11 we relax the upper bound as follows:

$$\begin{aligned}
4\|P_n^t - \pi_n\|_{\mathrm{TV}}^2 &\leq \mathrm{eig}(T_{(1^n)})^{2t} + \sum_{\substack{\lambda\in\Lambda_n^{\rightarrow}\\ \lambda\neq(n)}} d_\lambda \sum_{T\in SYT(\lambda)} \mathrm{eig}(T)^{2t} + \sum_{\substack{\lambda\in\Lambda_n^{\downarrow}\\ \lambda\neq(1^n)}} d_\lambda \sum_{T\in SYT(\lambda)} \mathrm{eig}(T)^{2t} \\
&\leq \mathrm{eig}(T_{(1^n)})^{2t} + \sum_{\substack{\lambda\in\Lambda_n^{\rightarrow}\\ \lambda\neq(n)}} d_\lambda^2 \,\mathrm{eig}(T_\lambda^{\rightarrow})^{2t} + \sum_{\substack{\lambda\in\Lambda_n^{\downarrow}\\ \lambda\neq(1^n)}} d_\lambda^2 \,\mathrm{eig}(T_\lambda^{\downarrow})^{2t} \\
&\leq \mathrm{eig}(T_{(1^n)})^{2t} + \sum_{\substack{\lambda:\mathrm{eig}(T_\lambda^{\rightarrow})\geq 0\\ \lambda\neq(n)}} d_\lambda^2 \,\mathrm{eig}(T_\lambda^{\rightarrow})^{2t} + \sum_{\substack{\lambda:\mathrm{eig}(T_\lambda^{\downarrow})<0\\ \lambda\neq(1^n)}} d_\lambda^2 \,\mathrm{eig}(T_\lambda^{\downarrow})^{2t} \\
&\leq \mathrm{eig}(T_{(1^n)})^{2t} + \sum_{\substack{\lambda:\mathrm{eig}(T_\lambda^{\rightarrow})\geq 0\\ \lambda\neq(n)}} d_\lambda^2 \,\mathrm{eig}(T_\lambda^{\rightarrow})^{2t} + \sum_{\substack{\lambda:\mathrm{eig}(T_\lambda^{\downarrow})<0\\ \lambda'\neq(n)}} d_{\lambda'}^2 \,\mathrm{eig}(T_{\lambda'}^{\rightarrow})^{2t} \\
&\leq \mathrm{eig}(T_{(1^n)})^{2t} + 2\sum_{\substack{\lambda:\mathrm{eig}(T_\lambda^{\rightarrow})\geq 0\\ \lambda\neq(n)}} d_\lambda^2 \,\mathrm{eig}(T_\lambda^{\rightarrow})^{2t} . \qquad (8)
\end{aligned}$$

(In the penultimate line we have used Corollary 11 and the fact that $d_{\lambda'} = d_\lambda$. The final inequality follows by a second application of Corollary 11: if λ satisfies $\mathrm{eig}(T_\lambda^{\downarrow}) < 0$ then $\mathrm{eig}(T_{\lambda'}^{\rightarrow})$ must be non-negative.)

The first term in (8) is simple to deal with at time $t = n\log(n) + cn$. We have already observed that $\mathrm{eig}\left(T_{(n)}\right) = 1$, and so Lemma 10 implies that $\mathrm{eig}\left(T_{(1^n)}\right) = 2H_n/n - 1$. This means that

$$\mathrm{eig}\left(T_{(1^n)}\right)^{2t} = \left(1 - \frac{2H_n}{n}\right)^{2(n\log n + cn)} \qquad (9)$$

and, using the bound $1 - x \leq e^{-x}$, we see that this tends to zero for any fixed c as $n \to \infty$.

It therefore remains to bound the sum in (8). The partitions with the biggest eigenvalues will be those with large first rows λ_1, and so we split the analysis into two parts according to this value; by *large* partitions we mean those with $\lambda_1 \geq 3n/4$, and *small* partitions are those with $\lambda_1 < 3n/4$. Large partitions give the biggest eigenvalues for P_n and must be dealt with carefully; it is these which will determine the mixing time of the shuffle. Small partitions have correspondingly large dimensions, but eigenvalues which are small enough to give control around any time of order $n\log(n)$. We begin by identifying the partition at the top of the dominance ordering for any fixed value of λ_1, which allows us to employ Lemma 12. If $\lambda \vdash n$ has first row equal to $\lambda_1 = n - k$, then by moving boxes up and to the right it follows trivially that

$$\lambda \trianglelefteq \begin{cases} (n-k, k) & \text{if } k \leq \frac{n}{2} \\ (n-k, n-k, \dots) = (n-k, \star) & \text{if } \frac{n}{2} < k \leq n-1 , \end{cases}$$

where we write $(n-k, \star)$ for the partition which has as many rows of $n-k$ boxes as possible, with the last row being formed from whatever is left over; it will transpire that in this case only the size of the first two rows will be important for our bounds.

For each k we also need a bound on sum of the squared dimensions of all partitions with $\lambda_1 = n - k$, and for this we use:

Lemma 13 (Corollary 2 of [6])**.**

$$\sum_{\substack{\lambda\vdash n\\ \lambda_1 = n-k}} d_\lambda^2 \leq \binom{n}{k}^2 k! \leq \frac{n^{2k}}{k!} .$$

Large Partitions

Let λ be a partition satisfying $\mathrm{eig}(T_\lambda^{\rightarrow}) \geq 0$, and for which $\lambda_1 = n - k$ for some $k \leq n/4$. We have observed above that $\lambda \trianglelefteq (n-k, k)$, and so Lemma 12 suggests that we look at the eigenvalue of $T^{\rightarrow}_{(n-k,k)}$. Using our eigenvalue formula from Lemma 6 we calculate this as follows, with the first/second sum corresponding to the first/second row of $T^{\rightarrow}_{(n-k,k)}$:

$$\begin{aligned}\operatorname{eig}(T^{\rightarrow}_{(n-k,k)}) &= \frac{1}{n}\sum_{j=1}^{n-k}\frac{j}{T^{\rightarrow}_{(n-k,k)}(1,j)} + \frac{1}{n}\sum_{j=1}^{k}\frac{j-1}{T^{\rightarrow}_{(n-k,k)}(2,j)} = \frac{n-k}{n} + \frac{1}{n}\sum_{j=1}^{k}\frac{j-1}{n-k+j} \\ &= 1 - \frac{(n-k+1)}{n}(H_n - H_{n-k+1}) - \frac{1}{n}.\end{aligned} \tag{10}$$

We now use this, along with the inequality $1-x \le e^{-x}$, to bound the contribution of large partitions to the sum in (8):

$$\begin{aligned}\sum_{k=1}^{n/4}\sum_{\substack{\lambda:\operatorname{eig}(T^{\rightarrow}_{\lambda})\ge 0\\ \lambda_1=n-k}} d_\lambda^2 \operatorname{eig}(T^{\rightarrow}_{\lambda})^{2t} &\le \sum_{k=1}^{n/4}\frac{n^{2k}}{k!}\operatorname{eig}\left(T^{\rightarrow}_{(n-k,k)}\right)^{2t} \qquad \text{(by Lemma 13)} \\ &\le \sum_{k=1}^{n/4}\frac{n^{2k}}{k!}\left(1-\frac{(n-k+1)}{n}(H_n-H_{n-k+1})-\frac{1}{n}\right)^{2t} \\ &\le \sum_{k=1}^{n/4}\frac{n^{2k}}{k!}e^{-2t\left(\frac{(n-k+1)}{n}(H_n-H_{n-k+1})+\frac{1}{n}\right)} \\ &\le e^{-2c}\sum_{k=1}^{n/4}\frac{n^{2k-2(n-k+1)(H_n-H_{n-k+1})-2}}{k!},\end{aligned} \tag{11}$$

where in the last step we have substituted $t = n\log n + cn$. The ratio of $(k+1)^{th}$ to k^{th} terms in this sum is given by

$$\frac{n^{2(H_n-H_{n-k})}}{k+1}. \tag{12}$$

For large n this ratio is less than one for all $k = 1, \ldots, n/4$. Indeed, a little analysis shows that for large n the largest value of the ratio over this range of k is achieved when $k = 1$, at which point it equals $n^{2/n}/2$. For sufficiently large n this ratio is thus bounded above by $3/4$, say, which permits us to bound the sum in (11) by a geometric series with initial term 1:

$$e^{-2c}\sum_{k=1}^{n/4}\frac{n^{2k-2(n-k+1)(H_n-H_{n-k+1})-2}}{k!} \le e^{-2c}\sum_{k=1}^{n/4}(3/4)^{k-1} \le 4e^{-2c}. \tag{13}$$

Small partitions

Now consider a partition λ satisfying $\operatorname{eig}(T^{\rightarrow}_{\lambda}) \ge 0$ and for which $\lambda_1 = n-k$ with $n/4 < k \le n-2$. Suppose first of all that $n/4 < k \le n/2$; as in the large partition case, any such partition is dominated by $(n-k,k)$, and the same calculation as in equation (10) shows that

$$\operatorname{eig}\left(T^{\rightarrow}_{(n-k,k)}\right) = \frac{n-k}{n} + \frac{1}{n}\sum_{j=1}^{k}\frac{j-1}{n-k+j}. \tag{14}$$

Now consider the case when $k > n/2$. Here we have already identified that $\lambda \trianglelefteq (n-k,\star)$, and so we proceed by calculating the eigenvalue of $T^{\rightarrow}_{(n-k,\star)}$. Note first that for any box (i,j) with $i \ge 3$,

$$\frac{j-i+1}{T^{\rightarrow}_{(n-k,\star)}(i,j)} = \frac{j-i+1}{(i-1)(n-k)+j} \le \frac{(n-k)}{(i-1)(n-k)+(n-k)} \le \frac{1}{3}.$$

Using this inequality in conjunction with Lemma 6 we bound $\operatorname{eig}(T^{\rightarrow}_{(n-k,\star)})$ as follows:

$$\begin{aligned}\operatorname{eig}(T^{\rightarrow}_{(n-k,\star)}) &= \frac{1}{n}\sum_{j=1}^{n-k}\frac{j}{T^{\rightarrow}_{(n-k,\star)}(1,j)} + \frac{1}{n}\sum_{j=1}^{n-k}\frac{j-1}{T^{\rightarrow}_{(n-k,\star)}(2,j)} + \frac{1}{n}\sum_{\substack{(i,j)\\ i\ge 3}}\frac{j-i+1}{T^{\rightarrow}_{(n-k,\star)}(i,j)} \\ &\le \frac{n-k}{n} + \frac{1}{n}\sum_{j=1}^{n-k}\frac{j-1}{n-k+j} + \frac{n-2(n-k)}{3n}.\end{aligned} \tag{15}$$

We now observe that (15) provides an upper bound for the expression in (14). Indeed, for $n/4 < k \le n/2$ we may write

$$\frac{n-k}{n} + \frac{1}{n}\sum_{j=1}^{n-k}\frac{j-1}{n-k+j} + \frac{n-2(n-k)}{3n} - \mathrm{eig}(T^{\rightarrow}_{(n-k,k)}) = \frac{1}{n}\sum_{j=k+1}^{n-k}\left(\frac{j-1}{n-k+j} - \frac{1}{3}\right)$$
$$= \frac{2(n-2k)}{3n} - \frac{(n-k+1)}{n}(H_{2(n-k)} - H_n)\,.$$

Substituting $k = \gamma n$, this final expression is bounded below for any $n \ge 15$ by the function $f(\gamma)$, where $f : [1/4, 1/2] \to \mathbb{R}$ is defined by

$$f(\gamma) = \frac{2(1-2\gamma)}{3} - \left(1 - \gamma + \frac{1}{15}\right)\log(2(1-\gamma))\,.$$

This function is non-negative for all $\gamma \in [1/4, 1/2]$, thus completing our claim.

We have just shown that for any $\lambda \vdash n$ satisfying $\mathrm{eig}(T^{\rightarrow}_{\lambda}) \ge 0$ and for which $\lambda_1 = n-k$ with $n/4 < k \le n-2$,

$$\begin{aligned}\mathrm{eig}\,(T^{\rightarrow}_{\lambda}) &\le \frac{n-k}{n} + \frac{1}{n}\sum_{j=1}^{n-k}\frac{j-1}{n-k+j} + \frac{n-2(n-k)}{3n} \\ &= \frac{n-k}{n} + \frac{n-k-1-(n-k+1)(H_{2(n-k)} - H_{n-k+1})}{n} + \frac{2k-n}{3n} \\ &= 1 - \frac{(4k-2n+3)}{3n} - \frac{(n-k+1)}{n}(H_{2(n-k)} - H_{n-k+1})\,.\end{aligned}$$

Using the inequalities $1 - x \le e^{-x}$ for all x, and $(x+1)(H_{2x} - H_{x+1}) > (x-1)\log 2$ for all integers $x \ge 2$, we are able to bound the contribution from small partitions to the sum in (8) at time $t = n\log n + cn$ as follows:

$$\begin{aligned}\sum_{k=n/4}^{n-2}\sum_{\substack{\lambda:\,\mathrm{eig}(T^{\rightarrow}_{\lambda})\ge 0\\ \lambda_1 = n-k}} d_{\lambda}^2\mathrm{eig}\,(T^{\rightarrow}_{\lambda})^{2t} &\le \sum_{k=n/4}^{n-2}\frac{n^{2k}}{k!}\,e^{-\frac{2t}{n}\left(\frac{4k-2n+3}{3} + (n-k+1)(H_{2(n-k)} - H_{n-k+1})\right)} \\ &\le e^{-2c}\sum_{k=n/4}^{n-2}\frac{n^{\frac{4n-2k-6}{3} - 2(n-k-1)\log 2}}{k!}\,. \qquad (16)\end{aligned}$$

Once again writing $k = \gamma n$, now for $\gamma \in [1/4, 1]$, a straightforward application of Stirling's formula shows that for large n the dominant term in the summand of (16) takes the form $n^{g(\gamma)n/3}$, where $g(\gamma) = 4 - 5\gamma - 6(1-\gamma)\log 2 < 0$ for all $\gamma \in [1/4, 1]$. It follows that, for any fixed c, the sum in (16) tends to zero as $n \to \infty$. Combining this result with the bounds in (8), (9) and (13) completes the proof of the upper bound on the mixing time in Theorem 3.

3 Lower Bound

To complete Theorem 3 we need to prove the lower bound on the mixing time. To do this we employ the usual trick of finding a set of permutations $B_n \subset S_n$ which has significantly different probability under the equilibrium distribution π_n and the one-sided transposition measure P_n^t. The definition of total variation distance then immediately yields a simple lower bound:

$$\|P_n^t - \pi_n\|_{\mathrm{TV}} \ge P_n^t(B_n) - \pi_n(B_n)\,.$$

In particular, we follow in the steps of [6] and find a suitable set B_n by considering the number of fixed points within (a certain part of) the deck. Estimation of $P_n^t(B_n)$ then reduces to a novel variant of the classical coupon collector's problem.

Recall that the deck of n cards are labelled $\{1, \ldots, n\}$ from bottom to top. One step of the one-sided transposition shuffle on n cards may be modelled by firstly choosing a position

$R^i \sim U\{1,\dots n\}$ with our right hand, and then choosing a position (below our right hand) $L^i \sim U\{1,\dots,R^i\}$ with our left hand and transposing the cards in the chosen positions.

Since the left hand always chooses a position below that of the right hand, it is intuitively clear that our shuffle is relatively unlikely to transpose two cards near to the top of the deck. This leads us to focus attention on a set of positions at the top of the deck: write V_n for the top part of the deck, where

$$V_n = \{n - n/m + 1, \dots, n-1, n\}\,,$$

and where $m = m(n)$ is to be chosen later. We shall be keeping track of fixed points within this part of the deck. Let

$$B_n = \{\rho \in S_n \,|\, \rho \text{ has at least 1 fixed point in } V_n\}\,.$$

Note that V_n contains n/m positions, and so we may upper bound the size of B_n by choosing one position in V_n to fix and considering all permutations of the other $n-1$ positions. This shows that $|B_n| \le (n/m)(n-1)!$, and hence $\pi_n(B_n) \le 1/m$.

To bound the value of $P_n^t(B_n)$ we reduce the problem to studying a simpler Markov chain linked to coupon collecting. When either of our hands (R^i, L^i) picks a new (previously untouched) card we shall say that this card gets *collected*. The uncollected cards in V_n at time t are those which have not yet been picked by either hand, and thus the size of this set gives us a lower bound on the number of fixed points in V_n. Writing U_n^t for the set of uncollected cards in V_n after t one-sided transposition shuffles, it follows that

$$P_n^t(B_n) \ge \mathbb{P}(|U_n^t| \ge 1)\,. \tag{17}$$

We wish to show that at time $t = n\log n - n\log\log n$ the probability on the right hand side of (17) is large. The difficulty with the analysis here is that in the one-sided transposition shuffle the value of L^i is clearly not independent of R^i. This means that a standard coupon-collecting argument for the time taken to collect all of the cards/positions in V_n cannot be applied in our setting, and a little more work is therefore required.

Note that at each step there are four possibilities: both hands collect new cards, only one hand does (left or right) or neither does. This permits us to bound the change in the number of *collected* cards as follows:

$$\begin{aligned}
|V_n \smallsetminus U_n^{t+1}| &= |V_n \smallsetminus U_n^t| + |\{L^{t+1}, R^{t+1}\} \cap U_n^t| \\
&\le |V_n \smallsetminus U_n^t| + 2\cdot \mathbb{1}[L^{t+1} \in U_n^t] + \mathbb{1}[L^{t+1} \notin U_n^t, R^{t+1} \in U_n^t]\,,
\end{aligned} \tag{18}$$

where $\mathbb{1}[\cdot]$ is an indicator function. Now, since the left hand is more likely to choose positions towards the bottom of the pack,

$$\mathbb{P}(L^{t+1} \in U_n^t) \le \mathbb{P}(L^{t+1} \in \hat{U}_n^t)\,,$$

where $\hat{U}_n^t = \{n - n/m + 1, \dots, n - n/m + |U_n^t|\}$, i.e. the $|U_n^t|$ lowest numbered positions in V_n. Furthermore,

$$\begin{aligned}
\mathbb{P}(L^{t+1} \in \hat{U}_n^t) &= \frac{1}{n}\sum_{k\in\hat{U}_n^t} \mathbb{P}\left(L^{t+1} \in \hat{U}_n^t \,|\, R^{t+1} = k\right) + \frac{1}{n}\sum_{k\in V_n\smallsetminus\hat{U}_n^t} \mathbb{P}\left(L^{t+1} \in \hat{U}_n^t \,|\, R^{t+1} = k\right) \\
&= \frac{1}{n}\sum_{k\in\hat{U}_n^t} \frac{k - (n - n/m)}{k} + \frac{1}{n}\sum_{k\in V_n\smallsetminus\hat{U}_n^t} \frac{|U_n^t|}{k} \\
&\le \frac{\sum_{k=1}^{|U_n^t|} k + (n/m - |U_n^t|)|U_n^t|}{n(n - n/m)} \le \frac{|U_n^t|}{(m-1)n}\,.
\end{aligned} \tag{19}$$

The probability of the final event in (18) is simple to bound:

$$\mathbb{P}(L^{t+1} \notin U_n^t, R^{t+1} \in U_n^t) \le \mathbb{P}(R^{t+1} \in U_n^t) \le \frac{|U_n^t|}{n}\,. \tag{20}$$

Using (18), (19) and (20) together, we now define a counting process M_n^t which stochastically dominates the number of collected cards $|V_n \setminus U_n^t|$ at all times:

$$M_n^0 = 0\,;$$
$$\mathbb{P}(M_n^{t+1} = M_n^t + k) = \begin{cases} \frac{1}{(m-1)n}\left(\frac{n}{m} - M_n^t\right) & \text{if } k = 2 \\ \frac{1}{n}\left(\frac{n}{m} - M_n^t\right) & \text{if } k = 1 \\ 1 - \frac{m}{(m-1)n}\left(\frac{n}{m} - M_n^t\right) & \text{if } k = 0\,. \end{cases} \tag{21}$$

Combining this with (17) we obtain the following bound on $P_n^t(B_n)$:

$$P_n^t(B_n) \geq \mathbb{P}\left(M_n^t < n/m\right)\,. \tag{22}$$

We are interested in the time at which the process M_n^t first reaches level n/m, where we now take $m = m(n) = \log n$.

Lemma 14. *Let $\mathcal{T} = \min\{t : M_n^t \geq n/\log n\}$. Then for any $c > 2$,*

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{T} < n\log n - n\log\log n - cn) \leq \frac{\pi^2}{6(c-2)^2}\,.$$

Before proving Lemma 14 we first show how this result quickly leads to a proof of the lower bound in Theorem 3. Writing $t = n\log n - n\log\log n$ we obtain

$$\begin{aligned} \|P_n^{t-cn} - \pi_n\|_{\mathrm{TV}} &\geq P_n^{t-cn}(B_n) - \pi_n(B_n) \geq \mathbb{P}\left(M_n^{t-cn} < n/m(n)\right) - 1/m(n) \\ &= \mathbb{P}(\mathcal{T} > t - cn) - \frac{1}{\log n} \geq 1 - \frac{\pi^2}{6(c-2)^2} - \frac{1}{\log n}\,, \end{aligned}$$

as required.

Proof of Lemma 14. Let $\mathcal{T}_i$ be the time spent by the process M_n^t in each state $i \geq 0$. We want to find $\mathcal{T} = \mathcal{T}_0 + \mathcal{T}_1 + \cdots + \mathcal{T}_{(n/m)-1}$.

From (21) we see that

$$p_i := \mathbb{P}(M_n^{t+1} > M_n^t \,|\, M_n^t = i) = \frac{m}{(m-1)n}\left(\frac{n}{m} - i\right)\,. \tag{23}$$

In the standard coupon collecting problem each of the random variables $\mathcal{T}_i$ has a geometric distribution with success probability p_i. Here, however, we have to take into account the chance that our counting process M_n increments by two, leading it to spend zero time at some state. Note first that

$$\mathbb{P}(M_n^{t+1} = M_n^t + 2 \,|\, M_n^{t+1} > M_n^t) = \frac{1}{m}\,,$$

independently of the value of M_n^t. Prior to spending any time in state i, the process M_n^t must visit (at least) one of the states $i-1$ or $i-2$. A simple argument shows that

$$\mathbb{P}(\mathcal{T}_i > 0 \,|\, \mathcal{T}_{i-1} > 0) = 1 - \frac{1}{m}\,, \quad \text{and} \quad \mathbb{P}(\mathcal{T}_i > 0 \,|\, \mathcal{T}_{i-2} > 0) = 1 - \frac{1}{m}\left(1 - \frac{1}{m}\right) \geq 1 - \frac{1}{m}\,.$$

Therefore $\mathbb{P}(\mathcal{T}_i > 0) \geq 1 - \frac{1}{m}$ for all states i, and so $\mathcal{T}_i$ stochastically dominates the random variable $\mathcal{T}_i'$ with mass function

$$\mathbb{P}(\mathcal{T}_i' = k) = \begin{cases} 1/m & k = 0 \\ (1 - 1/m)p_i(1-p_i)^{k-1} & k \geq 1\,. \end{cases} \tag{24}$$

It follows that $\mathbb{P}(\mathcal{T} < t) \leq \mathbb{P}(\mathcal{T}' < t)$ for any t, where $\mathcal{T}' = \mathcal{T}_0' + \mathcal{T}_1' + \cdots + \mathcal{T}_{(n/m)-1}'$. Setting $m = m(n) = \log n$ we may bound the expectation and variance of $\mathcal{T}'$:

$$\begin{aligned} \mathrm{E}[\mathcal{T}'] &= \sum_{i=0}^{n/m-1} \frac{m-1}{mp_i} = \left(\frac{m-1}{m}\right)^2 n\log(n/m) \geq n\log n - n\log\log n - 2n\,; \\ \mathrm{Var}[\mathcal{T}'] &\leq \sum_{i=0}^{n/m-1} \frac{1}{p_i^2} \leq \sum_{i=1}^{n/m} \frac{n^2}{i^2} \leq \frac{\pi^2}{6}n^2\,. \end{aligned}$$

Finally, applying Chebyshev's inequality yields the following for any $c > 2$:

$$\mathbb{P}\left(\mathcal{T}' < n\log(n) - n\log\log n - cn\right) \le \mathbb{P}\left(|\mathcal{T}' - \mathrm{E}[\mathcal{T}']| > (c-2)n\right) \le \frac{\pi^2}{6(c-2)^2}.$$

□

4 Biased One-sided Transpositions

In this section we generalise the result of Theorem 3 by allowing the right hand to choose from a more general distribution on $[n]$.

Definition 15. *Given a* weight function $w : \mathbb{N} \to (0, \infty)$, *let* $N_w(n) = \sum_{i=1}^{n} w(i)$ *denote the cumulative weight up to* n. *Then the* biased one-sided transposition shuffle $P_{n,w}$ *is the random walk on* S_n *generated by the following distribution on transpositions:*

$$P_{n,w}(\tau) = \begin{cases} \frac{w(j)}{N_w(n)} \cdot \frac{1}{j} & \text{if } \tau = (i\, j) \text{ for some } 1 \le i \le j \le n \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

This shuffle allows a general distribution for the position chosen by the right hand, R^i, with

$$\mathbb{P}(R^i = j) = \frac{w(j)}{N_w(n)}, \text{ for } 1 \le j \le n, \tag{26}$$

but maintains the property that the left hand L^i chooses a position uniformly on the set $\{1, \ldots, R^i\}$. Importantly the weight of each position j may only depend on j and not the size of the deck n. This setup implies that the biased shuffle preserves the recursive algebraic structure identified in the Appendix, and that the results of Theorems 32 and 41 still hold (up to minor changes in constants). It follows that the eigenvalues of the biased one-sided transposition shuffles are still represented by standard Young tableaux, and that the eigenvalue associated to a tableau T may be computed in a similar way.

Lemma 16. *The eigenvalues for the biased one-sided transposition shuffle* $P_{n,w}$ *on* n *cards are indexed by standard Young tableau of size* n. *For a standard Young tableau* T *of size* n *and* $m \in [n]$ *define a function* $T(m)$ *by setting* $T(m) = j - i + 1$ *if and only if* $T(i,j) = m$. *The eigenvalue corresponding to a tableau* T *is given by*

$$\mathrm{eig}(T) = \frac{1}{N_w(n)} \sum_{\substack{\text{boxes} \\ (i,j)}} \frac{j-i+1}{T(i,j)} \cdot w(T(i,j)) = \frac{1}{N_w(n)} \sum_{m=1}^{n} T(m) \frac{w(m)}{m}.$$

We now focus on a natural choice of weight function of the form $w(j) = j^\alpha$; we shall denote the resulting shuffle as $P_{n,\alpha}$, and write $N_\alpha(n)$ in place of $N_w(n)$. For $\alpha = 0$ we recover the unbiased one-sided transposition shuffle P_n, while if $\alpha > 0$ ($\alpha < 0$) the right hand is biased towards the top (respectively, bottom) of the deck.

Theorem 17. *Define the time* $t_{n,\alpha}$ *as,*

$$t_{n,\alpha} = \begin{cases} N_\alpha(n)/n^\alpha & \text{if } \alpha \le 1 \\ N_\alpha(n)/N_{\alpha-1}(n) & \text{if } \alpha \ge 1. \end{cases}$$

The biased one-sided transposition shuffle $P_{n,\alpha}$ *exhibits a total variation cutoff at time* $t_{n,\alpha} \log n$ *for all* $\alpha \in \mathbb{R}$. *Specifically for any* $c_1 > 5/2, c_2 > \max\{2, 3-\alpha\}$ *we have:*

$$\limsup_{n\to\infty} \|P_{n,\alpha}^{t_{n,\alpha}(\log n + c_1)} - \pi_n\| \le Ae^{-2c_1} \text{ for a universal constant } A, \text{ for all } \alpha$$

and

$$\liminf_{n\to\infty} \|P_{n,\alpha}^{t_{n,\alpha}(\log n - \log\log n - c_2)} - \pi_n\| \ge \begin{cases} 1 - \frac{\pi^2}{6(c_2-3+\alpha)^2} & \text{for } \alpha \le 1 \\ 1 - \frac{\pi^2}{6(c_2-2)^2} & \text{for } \alpha \ge 1. \end{cases}$$

The asymptotics of the cutoff times for the biased one-sided transposition shuffle are summarised in Table 2.

	$\alpha \in (-\infty,-1)$	$\alpha = -1$	$\alpha \in (-1,1]$	$\alpha \in (1,\infty)$
$t_{n,\alpha}\log n$	$\zeta(-\alpha)n^{-\alpha}\log n$	$n(\log n)^2$	$\frac{1}{1+\alpha}n\log n$	$\frac{\alpha}{1+\alpha}n\log n$

Table 2: Asymptotics of the cutoff time $t_{n,\alpha}\log n$ as $n \to \infty$, for different values of α.

Note that the fastest mixing time is obtained when $\alpha = 1$; using this weight function the shuffle is constant on the conjugacy class of transpositions, with transition probabilities similar to those of the classical random transpositions: $P_{n,1}((i\,j)) = 2/(n(n+1))$. In this case we obtain a mixing time of $t_{n,1} \sim (n/2)\log n$ which agrees with that of the random transposition shuffle. The mixing time is bounded above by $n\log n$ for all $\alpha > 1$, but as $\alpha \to -\infty$ the mixing time is unbounded; in particular, when $\alpha < -1$ the mixing time is of order $O(n^{-\alpha}\log n)$.

Theorem 17 is proved by generalising the results of Sections 2 and 3. Many of the arguments are almost identical to those already presented, so in what follows we shall simply sketch the main differences. We note that bounding the mixing time when using more general monotonic weight functions (but still satisfying Definition 15) is relatively straightforward: Lemma 16 indicates that the upper bound on the mixing time is determined by whether $w(n)/n$ is increasing or decreasing in n. (See [?] for further details.) Here we restrict attention to the case when $w(j) = j^\alpha$ since for this family of shuffles we are able to show the existence of a cutoff.

4.1 Upper Bound for Biased One-sided Transpositions

We first of all note that the shuffles $P_{n,\alpha}$ are still aperiodic, transitive, and reversible, meaning we may once again use Lemma 4 to upper bound the mixing time. Furthermore, for $\alpha \le 1$ the main results of Section 2.1 still hold, meaning that it again makes sense to bound the eigenvalues of large and small partitions separately. For $\alpha \ge 1$ we will introduce a new tableau which will help us bound the eigenvalues for $P_{n,\alpha}$. After establishing bounds on our eigenvalues for all α we present a combined argument for the upper bound in Theorem 17.

Lemma 18. *Let $\lambda \vdash n$ with $\lambda_1 = n-k$. Then the eigenvalue $\mathrm{eig}(T^{\rightarrow}_\lambda)$ for the shuffle $P_{n,\alpha}$ with $\alpha \le 1$ may be bounded as follows:*

$$\mathrm{eig}(T^{\rightarrow}_\lambda) \le \begin{cases} 1 - \frac{(n-k+1)kn^\alpha}{nN_\alpha(n)} & \text{if } k \le n/4 \\ 1 - \frac{kn^\alpha}{2N_\alpha(n)} & \text{if } n/4 < k. \end{cases}$$

Proof. For $k \le n/2$, the maximum partition in the dominance order for the class of partitions with $\lambda_1 = n-k$ is $(n-k,k)$, and so $\mathrm{eig}(T^{\rightarrow}_\lambda) \le \mathrm{eig}\left(T^{\rightarrow}_{(n-k,k)}\right)$. The eigenvalue of $T^{\rightarrow}_{(n-k,k)}$ may be calculated by summing over the two rows of the partition $(n-k,k)$ and using Lemma 16, as follows:

$$\begin{aligned} N_\alpha(n)\,\mathrm{eig}\left(T^{\rightarrow}_{(n-k,k)}\right) &= \sum_{m=1}^{n-k} m^\alpha + \sum_{m=n-k+1}^{n} (m-n+k-1)m^{\alpha-1} \\ &= \sum_{m=1}^{n} m^\alpha - (n-k+1)\sum_{m=n-k+1}^{n} m^{\alpha-1} \\ &\le N_\alpha(n) - \frac{(n-k+1)kn^\alpha}{n}. \end{aligned}$$

This immediately proves the desired inequality for $k \le n/4$, and also yields the stated bound for $n/4 < k \le n/2$.

For $k > n/2$ we once again need to bound $\mathrm{eig}\left(T^{\rightarrow}_{(n-k,\star)}\right)$. Letting $\nu = (n-k,\star)$ for ease of notation we calculate as follows:

$$\begin{aligned}
N_\alpha(n)\,\mathrm{eig}(T_\nu^{\rightarrow}) &= \sum_{j=1}^{n-k} j^\alpha + \sum_{i=2}^{l(\nu)}\sum_{j=1}^{\nu_i}(j-i+1)((i-1)(n-k)+j)^{\alpha-1} \\
&= N_\alpha(n) - \sum_{i=2}^{l(\nu)}\sum_{j=1}^{\nu_i}(i-1)(n-k+1)\frac{((i-1)(n-k)+j)^\alpha}{(i-1)(n-k)+j} \\
&\le N_\alpha(n) - \frac{(n-k)n^\alpha}{n}\sum_{i=2}^{l(\nu)}(i-1)\nu_i\,.
\end{aligned}$$

By definition of the partition ν, each row but the last has size $n-k$, and the final row has size $\nu_{l(\nu)} = n-(l(\nu)-1)(n-k)$. In addition, since $l(\nu) = \lceil n/(n-k)\rceil$ we may write $l(\nu) = n/(n-k)+x$ for some $0 \le x < 1$. Substituting these values we obtain:

$$\begin{aligned}
N_\alpha(n)\,\mathrm{eig}(T_\nu^{\rightarrow}) &\le N_\alpha(n) - \frac{(n-k)n^\alpha}{n}\frac{(l(\nu)-1)(2n-(n-k)l(\nu))}{2} \\
&= N_\alpha(n) - \frac{n^\alpha}{2n}(n-(1-x)(n-k))(n-x(n-k)) \\
&= N_\alpha(n) - \frac{n^\alpha}{2n}(nk + x(1-x)(n-k)^2) \\
&\le N_\alpha(n) - \frac{kn^\alpha}{2}\,.
\end{aligned}$$

□

With $\alpha \ge 1$ the main results (all but Lemma 12) of Section 2.1 hold with the roles of $T_\lambda^{\downarrow}$ and $T_\lambda^{\rightarrow}$ interchanged, and so the bound on total variation in equation (8) now involves $\mathrm{eig}(T_\lambda^{\downarrow})$. Therefore, we look to bound the eigenvalue of $T_\lambda^{\downarrow}$ when $\lambda_1 = n-k$, and to do so we need to introduce a new tableau, $T_\lambda^{\searrow}$.

Definition 19. *Let $T_\lambda^{\searrow}$ define the Young tableau formed by filling in the diagonals of λ from left to right, with each diagonal filled from top to bottom. For example,*

$$T^{\searrow}_{(3,2,1)} = \begin{array}{|c|c|c|}\hline 3 & 5 & 6 \\ \hline 2 & 4 \\ \cline{1-2} 1 \\ \cline{1-1}\end{array}\,, \qquad T^{\searrow}_{(4,2)} = \begin{array}{|c|c|c|c|}\hline 2 & 4 & 5 & 6 \\ \hline 1 & 3 \\ \cline{1-2}\end{array}\,.$$

Lemma 20. *For $\alpha \ge 1$, and $\lambda \vdash n$ with $\lambda_1 = n-k$ we have,*

$$\mathrm{eig}(T_\lambda^{\downarrow}) \le \mathrm{eig}\left(T^{\searrow}_{(n-k,\star)}\right).$$

Proof. Recall that, for any tableau T_λ, $T_\lambda(m) = j-i+1$ if and only if m appears in box (i,j) of T_λ. Note that this function is constant on integers appearing in the same diagonal of T_λ, and that $T_\lambda^{\searrow}(m)$ is non-decreasing in m.

Now consider all the values of $T_\lambda^{\downarrow}(m)$ for $m \in [n]$, including repeats, and order them from smallest to largest as $c_1 \le c_2 \le \cdots \le c_n$. Using Lemma 16, we may upper bound $\mathrm{eig}(T_\lambda^{\downarrow})$ as follows:

$$\begin{aligned}
N_\alpha(n)\cdot\mathrm{eig}(T_\lambda^{\downarrow}) &= \sum_{m=1}^{n} T_\lambda^{\downarrow}(m)\cdot m^{\alpha-1} \le \sum_{m=1}^{n} c_m \cdot m^{\alpha-1} \\
&= \sum_{m=1}^{n} T_\lambda^{\searrow}(m)\cdot m^{\alpha-1} \le \sum_{m=1}^{n} T^{\searrow}_{(n-k,\star)}(m)\cdot m^{\alpha-1} = N_\alpha(n)\cdot\mathrm{eig}(T^{\searrow}_{(n-k,\star)})\,.
\end{aligned}$$

The first inequality follows from the fact that $m^{\alpha-1}$ is increasing in m for $\alpha \ge 1$, and so pairing up the constants c_m and $m^{\alpha-1}$ cannot decrease the value of the sum. For the second inequality, notice that $(n-k,\star)$ is obtained from λ by moving boxes up and to the right. Thus the diagonal containing m in $T^{\searrow}_{(n-k,\star)}$ is (weakly) to the right of the corresponding diagonal in $T_\lambda^{\searrow}$, and so $T_\lambda^{\searrow}(m) \le T^{\searrow}_{(n-k,\star)}(m)$ for all m.

□

Lemma 21. *For all $k \le n-2$ and all $m \in [n]$,*

$$T^{\searrow}_{(n-k,\star)}(m) \le \frac{n-k}{n} \cdot m. \tag{27}$$

Proof. Let us write $l(n-k,\star) = l^+ = \lceil \frac{n}{n-k} \rceil$ and $l^- = \lfloor \frac{n}{n-k} \rfloor$. If m_1, m_2 belong to the same diagonal then $T^{\searrow}_{(n-k,\star)}(m_1) = T^{\searrow}_{(n-k,\star)}(m_2)$, and hence if (27) holds for the smallest m on a diagonal it holds for every entry of that diagonal. Furthermore, the bound trivially holds for any m whose box (i,j) satisfies $j-i+1 \le 0$ (for which the left hand side of (27) is non-positive). Combining these two observations, we see that it suffices to prove the bound for those values of m which appear in the first row of $T^{\searrow}_{(n-k,\star)}$.

Note that no diagonal can contain more than l^+ boxes: call diagonals with l^- or fewer boxes *short* diagonals, and all others *long* diagonals. Note that long diagonals can only exist when $l^+ = l^- + 1$. Any long diagonals clearly occur strictly before the short ones, when working left to right along the first row. If the box $(1,j)$ lies on a long diagonal, then the numbering pattern for $T^{\searrow}$ implies that this box will contain the integer $m = \binom{l^+}{2} + 1 + (j-1)l^+$. For this value of m, the right hand side of (27) becomes

$$\begin{aligned} \frac{n-k}{n}\left(\frac{l^+(l^+-1)}{2} + 1 + (j-1)l^+\right) &= \frac{(n-k)l^+}{n}\left(\frac{l^+-1}{2} + \frac{1}{l^+} - 1 + j\right) \\ &\ge j = T^{\searrow}_{(n-k,\star)}(m), \end{aligned}$$

thanks to the definition of l^+ and the fact that $(x-1)/2 + 1/x \ge 1$ if $x \ge 2$.

It remains to deal with the short diagonals which contain a box in the first row. For these diagonals we now work from right to left, and consider boxes $(1, n-k+1-j)$ for $j = 1, 2, \ldots$. Let $m(j)$ denote the integer appearing in box $(1, n-k+1-j)$. We know that $m(1) = n$, and it is clear that (27) holds for box $(1, n-k)$. It is then straightforward to see that

$$m(j) = m(j-1) - \min\{j, l^-\}$$

for any $j \ge 2$ for which box $(1, n-k+1-j)$ is part of a short diagonal. The result for all short diagonals now follows quickly by induction (using $l^-(n-k)/n \le 1$). □

Lemma 22. *Let $\lambda \vdash n$ with $\lambda_1 = n-k$. Then the eigenvalue $\mathrm{eig}(T^{\downarrow}_{\lambda})$ for the shuffle $P_{n,\alpha}$ with $\alpha \ge 1$ may be bounded as follows:*

$$\mathrm{eig}(T^{\downarrow}_{\lambda}) \le \begin{cases} 1 - \frac{k(n-k)N_{\alpha-1}(n)}{nN_\alpha(n)} & \text{if } k \le n/4 \\ 1 - \frac{k}{n} & \text{for all } k. \end{cases}$$

Proof. We begin by quickly proving the second bound for all k using Lemmas 20 and 21:

$$\mathrm{eig}(T^{\downarrow}_{\lambda}) \le \mathrm{eig}\left(T^{\searrow}_{(n-k,\star)}\right) = \frac{1}{N_\alpha(n)} \sum_{m=1}^{n} T^{\searrow}_{(n-k,\star)}(m) \cdot m^{\alpha-1} \le \frac{1}{N_\alpha(n)} \sum_{m=1}^{n} \frac{(n-k)}{n} \cdot m^{\alpha} = 1 - \frac{k}{n}.$$

Now we prove the bound for $k \le n/4$, by finding a tighter bound on $T^{\searrow}_{(n-k,\star)}$ and once again using Lemma 20. For $k \le n/4$, we know $(n-k,\star) = (n-k,k)$, and thanks to the order in which the boxes are filled, we see that $T^{\searrow}_{(n-k,\star)}(m) \le m/2$ for $1 \le m \le 2k$, and $T^{\searrow}_{(n-k,\star)}(m) = m-k$ for $2k+1 \le m \le n$. This gives us the following simple bound:

$$\begin{aligned} N_\alpha(n) \cdot \mathrm{eig}(T^{\searrow}_{(n-k,\star)}) &\le \sum_{m=1}^{2k} \frac{m}{2} m^{\alpha-1} + \sum_{m=2k+1}^{n} (m-k)m^{\alpha-1} \\ &= N_\alpha(n) - kN_{\alpha-1}(n) + \left(kN_{\alpha-1}(2k) - \frac{1}{2}N_\alpha(2k)\right). \end{aligned} \tag{28}$$

Now,

$$\begin{aligned}
kN_{\alpha-1}(2k) - \frac{1}{2}N_\alpha(2k) &= k\sum_{m=1}^{2k} m^{\alpha-1}\left(1-\frac{m}{2k}\right) \le k\int_0^{2k} x^{\alpha-1}\left(1-\frac{x}{2k}\right)\mathrm{d}x \\
&= \frac{k(2k)^\alpha}{\alpha(1+\alpha)} \\
&\le \frac{2k^2(n/2)^{\alpha-1}}{\alpha(1+\alpha)} \qquad \text{(since } k \le n/4\text{)} \\
&\le k^2(n/2)^{\alpha-1},
\end{aligned}$$

since $\alpha \ge 1$. Finally, an application of Jensen's inequality shows that $(n/2)^{\alpha-1} \le N_{\alpha-1}(n)/n$, and combining this with (28) yields the desired result.

□

Returning to Theorem 17, recall that $t_{n,\alpha} = N_\alpha(n)/n^\alpha$ if $\alpha \le 1$ and $t_{n,\alpha} = N_\alpha(n)/N_{\alpha-1}(n)$ if $\alpha \ge 1$. Following the argument of Section 2.2 and using Lemmas 18 and 22 for the appropriate values of α, we may show that for any $\alpha \in \mathbb{R}$:

$$4\|P^t_{n,\alpha} - \pi_n\|^2_{\mathrm{TV}} \le (\mathrm{eig}(1^n))^{2t} + 2\sum_{k=1}^{n/4}\binom{n}{k}^2 k!\left(1-\frac{(n-k)k}{n}t^{-1}_{n,\alpha}\right)^{2t} + 2\sum_{k=n/4}^{n-2}\binom{n}{k}^2 k!\left(1-\frac{k}{2}t^{-1}_{n,\alpha}\right)^{2t}.$$

Substituting $t = t_{n,\alpha}(\log n + c)$ and once again using the inequality $1 - x \le e^{-x}$, we are left with two sums to control; both have previously been shown to be bounded above by Ae^{-4c}, for some universal constant A, when $c > 5/2$ and n is sufficiently large [5, page 42].

4.2 Lower Bound for Biased One-sided Transpositions

4.2.1 Case: $\alpha \le 1$

We use a coupon-collecting argument as in Section 3, once again letting $V_n = \{n - n/m + 1, \dots, n-1, n\}$ with $m = \log n$, and considering the set $B_n = \{\rho \in S_n \mid \rho \text{ has at least 1 fixed point in } V_n\}$. Equation (18) still holds for the biased version of the shuffle, but we now modify the bounds in (19) and (20) as follows, using the inequality $k^\alpha \le (\frac{m}{m-1})^{1-\alpha} n^\alpha$ for all $k \in V_n$ (which holds for all $\alpha \le 1$):

$$\begin{aligned}
\mathbb{P}(L^{t+1} \in \hat{U}^t_n) &= \sum_{k\in\hat{U}^t_n}\frac{w(k)}{N_\alpha(n)}\frac{k-(n-n/m)}{k} + \sum_{k\in V_n\setminus\hat{U}^t_n}\frac{w(k)}{N_\alpha(n)}\frac{|U^t_n|}{k} \\
&\le \frac{(\frac{m}{m-1})^{1-\alpha}n^\alpha}{N_\alpha(n)}\frac{\sum_{k=1}^{|U^t_n|}k + (n/m - |U^t_n|)|U^t_n|}{(n-n/m)} \le \frac{(\frac{m}{m-1})^{1-\alpha}n^\alpha|U^t_n|}{N_\alpha(n)(m-1)}; \\
\mathbb{P}(R^{t+1} \in U^t_n) &\le \frac{(\frac{m}{m-1})^{1-\alpha}n^\alpha|U^t_n|}{N_\alpha(n)}.
\end{aligned}$$

Using these as before we construct a counting process M^t_n which dominates the number of collected cards; the expression for p_i in (23) becomes

$$p_i = \left(\frac{m}{m-1}\right)^{2-\alpha}\frac{n^\alpha}{N_\alpha(n)}\left(\frac{n}{m} - i\right),$$

and this is easily checked to be strictly less than one for n sufficiently large, whatever the value of $\alpha \le 1$. The remainder of the analysis mirrors the unbiased case: using the new expression for p_i the distribution of the random variable $\mathcal{T}'_i$ is exactly as given in (24), and we arrive at

$$\begin{aligned}
\mathrm{E}[\mathcal{T}'] &= \sum_{i=0}^{n/m-1}\frac{m-1}{mp_i} \ge \frac{N_\alpha(n)}{n^\alpha}(\log n - \log\log n - (3-\alpha)); \\
\mathrm{Var}[\mathcal{T}'] &\le \sum_{i=0}^{n/m-1}\frac{1}{p_i^2} \le \frac{\pi^2}{6}\frac{N_\alpha(n)^2}{n^{2\alpha}}.
\end{aligned}$$

The proof of the lower bound is completed by using these new bounds in Chebychev's inequality.

4.2.2 Case: $\alpha \geq 1$

With $\alpha \geq 1$, the right hand is now more likely to choose cards near the top of the deck, and so it makes sense to swap the roles of the right and left hands in our coupon-collecting argument. To that end, let $V_n = \{1, \ldots, n/m\}$ and let

$$B_n = \{\rho \in S_n \,|\, \rho \text{ has at least 1 fixed point in } V_n\}.$$

As in Section 3 we let U_n^t denote the set of uncollected cards in V_n after t steps of the biased one-sided transposition shuffle. The change in the number of collected cards in (18) then holds with L^{t+1} and R^{t+1} interchanged, and we may replace the inequalities in (19) and (20) with the following bounds (which hold for sufficiently large n):

$$\mathbb{P}(R^{t+1} \in U_n^t) = \frac{1}{N_\alpha(n)} \sum_{i \in U_n^t} i^\alpha \leq \frac{|U_n^t|}{N_\alpha(n)} (n/m)^\alpha \leq \frac{|U_n^t|}{N_\alpha(n)} \frac{n^\alpha}{\alpha(m-1)} \leq \frac{|U_n^t|}{N_\alpha(n)} \frac{N_{\alpha-1}(n)}{(m-1)};$$

$$\mathbb{P}(R^{t+1} \notin U_n^t, L^{t+1} \in U_n^t) \leq \frac{|U_n^t|}{N_\alpha(n)} \sum_{i=|U_n^t|+1}^{n} \frac{i^\alpha}{i} \leq \frac{|U_n^t| N_{\alpha-1}(n)}{N_\alpha(n)}.$$

Analysis of the resulting counting process M_n^t shows that for $\alpha \geq 1$ the biased one-sided transposition shuffle satisfies the lowed bound in Theorem 17.

Appendix: Lifting Eigenvectors Analysis

We work with the group algebra $\mathfrak{S}_n = \mathbb{C}[S_n]$ and its representations. We begin with some of the background and basic constructions.

For each $n \in \mathbb{N}$ let $[n] = \{1, \ldots, n\}$ denote the set consisting of the first n natural numbers. Given $n \in \mathbb{N}$ we denote by W^n the set of words of length n in the elements of $[n]$, where by a word of length n we simply mean a string $w = w_1 \cdot w_2 \cdot \ldots \cdot w_n$ of n elements from $[n]$, allowing repeats. Note that in forming words we simply regard the elements of $[n]$ as distinct symbols; we separate symbols by a dot $\cdot$. (It is notationally convenient later on that these symbols are positive integers.) There is a natural action of the symmetric group S_n on W^n: given a word $w = w_1 \cdot w_2 \cdot \ldots \cdot w_n \in W^n$ and an element $\sigma \in S_n$, we let $\sigma(w) := w_{\sigma^{-1}(1)} \cdot w_{\sigma^{-1}(2)} \cdot \ldots \cdot w_{\sigma^{-1}(n)} \in W^n$. We emphasise that this is the action of S_n on words by place permutations, it is NOT the action of S_n induced by its action on $[n]$, e.g. $(123)(2 \cdot 3 \cdot 2) = 2 \cdot 2 \cdot 3 \neq 3 \cdot 1 \cdot 3$. We denote by M^n the complex space with basis W^n, so M^n is an n^n-dimensional vector space over $\mathbb{C}$. The S_n-action by place permutations on W^n extends to give M^n the structure of an $\mathfrak{S}_n$-module.

To each word $w \in W^n$ we can associate an n-tuple of non-negative integers, which we call its *evaluation*, as follows. For $1 \leq i \leq n$, let $\text{eval}_i(w)$ count the number of occurrences of the symbol i in the word w, and then let $\text{eval}(w) := (\text{eval}_1(w), \ldots, \text{eval}_n(w))$. Note that $\sum_{i=1}^n \text{eval}_i(w) = n$ for any word w in M^n. If in addition $\text{eval}(w)$ is a non-increasing sequence of integers, then we identify $\text{eval}(w)$ with the corresponding partition of n (this is a matter of "forgetting" any zeroes at the end of the tuple).

For a partition $\lambda \vdash n$, a Young tableau T of shape λ naturally corresponds to a word in W^n. Write the word $w(T) = w_1 \cdot w_2 \cdot \ldots \cdot w_n$ by setting $w_{T(i,j)} = i$ for each box (i, j) in T. Equivalently, the numerical entries in the i^{th} row of T tell us in which positions to put the symbol i in the word $w(T)$. Note that two tableaux give the same word if and only if they have the same shape and the same set of entries in each row (i.e., they correspond to the same *tabloid*).

Definition 23. *To every partition $\lambda \vdash n$ we may associate a simple module S^λ of $\mathfrak{S}_n$ called the* Specht module *for λ. The Specht module S^λ has dimension d_λ.*

Definition 24. *For an n-tuple $\lambda = (\lambda_1, \ldots, \lambda_n)$ of non-negative integers summing to n, define M^λ to be the span of the words $w \in W^n$ with* $\text{eval}(w) = \lambda$. *Since S_n acts by place permutations, this is clearly an $\mathfrak{S}_n$-submodule of M^n.*

If $\lambda \vdash n$ is a partition of n, then we allow ourselves an abuse of notation and also consider λ as an n-tuple by adding some zeroes on the end (if necessary). We can then attach the module M^λ to a partition λ. There is a unique copy of the Specht module S^λ as a submodule of M^λ

Let us explicitly identify the unique copy of $S^\lambda \subset M^\lambda$ as mentioned in the above definition. Let T be a standard Young tableau of shape λ. Define $C_T \subseteq S_n$ to be the column stabilizer for T – that is C_T is the subgroup of S_n consisting of permutations which permute the elements in each column of T. Corresponding to T we have the word $w(T) \in M^\lambda$ as above. Form a new element $s(T)$ of M^λ, where

$$s(T) := \sum_{\sigma \in \mathrm{ColStab}(T)} \mathrm{sign}(\sigma) w(\sigma(T)).$$

The Specht module S^λ is the subspace of M^λ with basis the elements $s(T)$, where T runs over all standard Young tableaux of shape λ.

For example, if $\lambda = (3,1)$ then we have 3 standard Young tableaux,

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1}\end{array} \qquad \begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 \\ \cline{1-1}\end{array}$$

and we only have the first column to permute in each case. It is easy to check that

$$S^{(3,1)} = \langle 1\cdot 1\cdot 1\cdot 2 - 2\cdot 1\cdot 1\cdot 1,\ 1\cdot 1\cdot 2\cdot 1 - 2\cdot 1\cdot 1\cdot 1,\ 1\cdot 2\cdot 1\cdot 1 - 2\cdot 1\cdot 1\cdot 1\rangle.$$

It is a standard result in the theory that S^λ has multiplicity 1 as a summand of the permutation module M^λ.

Lemma 25 (Theorem 2.11.2 of [15]). *For $\lambda \vdash n$ we have,*

$$M^\lambda \cong \bigoplus_{\mu \trianglerighteq \lambda} K_{\lambda,\mu} S^\mu.$$

where $K_{\lambda,\mu}S^\mu$ denotes a direct sum of $K_{\lambda,\mu}$ copies of S^μ. The coefficients $K_{\lambda,\mu}$ are called Kostka numbers, and for all $\lambda \vdash n$ we know $K_{\lambda,\lambda} = 1$.

The submodule $M^{(1^n)}$ is of particular importance because it is spanned by the $n!$ permutations of the word $1\cdot 2\cdot \ldots \cdot n \in W^n$, and therefore we can model our shuffle on n cards by considering a linear operator on this space. In representation-theoretic terms, $M^{(1^n)}$ can be identified as the regular module for $\mathfrak{S}_n$; we may state a classical result about this regular module in the notation we have now set up.

Corollary 26 (Example 2.11.6 of [15]).

$$M^{(1^n)} \cong \bigoplus_{\lambda \vdash n} d_\lambda S^\lambda \ \text{ as } \mathfrak{S}_n\text{-modules.}$$

For modelling our shuffle P_n acting on the space $\mathfrak{S}_n$ we need to turn it into a linear operator. In fact we may turn it into an element Q_n of our group algebra $\mathfrak{S}_n$.

Definition 27. *Let $n \in \mathbb{N}$. The one-sided transposition shuffle on n cards may be viewed as the following element of the group algebra $\mathfrak{S}_n$.*

$$\sum_{1\leq i\leq j\leq n} P_n((i\,j))(i\,j) = \sum_{1\leq i\leq j\leq n} \frac{1}{nj}(i\,j).$$

To simplify our calculations it is convenient to scale this operator by n, so we introduce a new operator

$$Q_n := \sum_{1\leq i\leq j\leq n} \frac{1}{j}(i\,j).$$

By realising the shuffling operator as an element of the group algebra we can concentrate on finding the eigenvalues of Q_n on $M^{(1^n)}$. Furthermore, applying Corollary 26, we can reduce the problem of finding eigenvalues for the shuffle on $M^{(1^n)}$ to the problem of finding eigenvalues on the Specht modules S^λ. Moreover, since the operator is acting as an element of the group algebra, we are then free to study its action on the natural copy of S^λ inside M^λ to solve this problem, rather than having to stick to the copies of S^λ inside $M^{(1^n)}$. This turns out to be very useful, because there is a natural way of relating eigenvectors and eigenvalues corresponding to different partitions according to the branching rules in Figure 1. Note that we allow the special case $n = 0$ with partition (0) and corresponding Young diagram $\varnothing$.

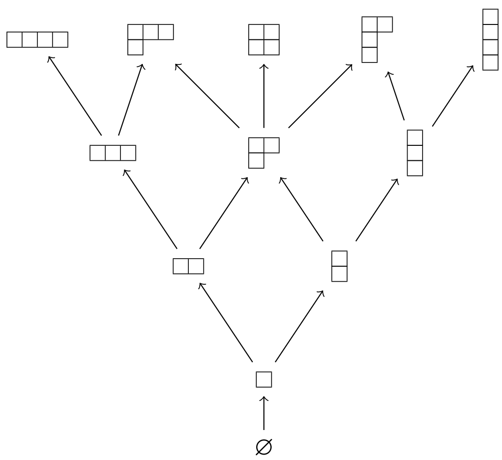


Figure 1: Young's lattice for partitions of size $n \in \{0, 1, 2, 3, 4\}$.

The one-sided transposition shuffle admits a recursive structure which is seen when we focus on the difference of Q_{n+1} and Q_n,

$$Q_{n+1} - Q_n = \frac{1}{(n+1)} \sum_{1 \leq i \leq (n+1)} (i \;\; n+1). \tag{29}$$

This signifies that the only difference between the one-sided transposition shuffle on $n+1$ and n cards is the movement of the new card in position $n+1$. We now define some important linear operators which will allow us to study the Q_n inductively using equation (29).

Definition 28. *We define two linear operators on the spaces spanned by words. To do so, it is enough to define the effect on any given word.*

1. *Let* $a \in [n+1]$. *Define the* adding operator $\Phi_a : M^n \to M^{n+1}$ *as follows: given a word* $w \in W^n$, *define*
$$\Phi_a(w) := w \cdot a,$$
(i.e., append the symbol a *to the end of the word* w*).*

2. *Let* $a, b \in [n]$. *Define the* switching operator $\Theta_{b,a} : M^n \to M^n$ *as follows: given a word* $w = w_1 \cdot w_2 \cdot \ldots \cdot w_n \in W^n$, *define*
$$\Theta_{b,a}(w) := \sum_{\substack{1 \leq i \leq n \\ w_i = b}} w_1 \cdot \ldots \cdot w_{i-1} \cdot a \cdot w_{i+1} \cdot \ldots \cdot w_n,$$
(i.e., for each occurrence of b *in the word* w*, replace that occurrence with* a *and sum the resulting words).*

Remark 29. *There is some ambiguity in the definitions just given – since* $1 \in [n]$ *for all* n*, for example, strictly speaking we should define* Φ_1 *separately for each* n*. However, this would burden us with even more notation, and it should always be clear from the context which domain and codomain we are considering.*

The adding operator on words defined above is an analogue of the process of adding boxes to Young diagrams. The process of adding boxes to Young diagrams may be seen in Figure 1. We now set up some notation to describe this more precisely.

Definition 30. *Given an n-tuple $\lambda = (\lambda_1, \dots, \lambda_r)$ of non-negative integers summing to n and an element $a \in [n+1]$, we form an $(n+1)$-tuple denoted $\lambda + e_a$ by first adding a zero to the end of λ and then adding 1 to this $(n+1)$-tuple in position a. Then $\lambda + e_a$ is an $(n+1)$-tuple of non-negative integers summing to $n+1$. E.g. $(2,1,0) + e_3 = (2,1,0,0) + (0,0,1,0) = (2,1,1,0)$.*

With this notation in hand, we have the following easy lemma:

Lemma 31. *Given $a \in [n+1]$ and an n-tuple λ of non-negative integers summing to n, we have*

$$\Phi_a : M^\lambda \to M^{\lambda + e_a},$$

i.e., the restriction of Φ_a to M^λ has image in $M^{\lambda+e_a}$.

Given $a, b \in [n]$ and n-tuples λ, μ of non-negative integers summing to n with $\lambda + e_a = \mu + e_b$, we have

$$\Theta_{b,a} : M^\lambda \to M^\mu,$$

i.e., the restriction of $\Theta_{b,a}$ to M^λ has image in M^μ.

Our next result establishes the crucial equation upon which all the subsequent results in this section rely. It relates the shuffle on n cards to that on $n+1$ cards, and allows us to lift the eigenvalues.

Theorem 32. *Given $n \in \mathbb{N}$, we have*

$$Q_{n+1} \circ \Phi_a - \Phi_a \circ Q_n = \frac{1}{n+1}\Phi_a + \frac{1}{n+1}\sum_{1 \le b \le n} \Phi_b \circ \Theta_{b,a}\,. \tag{30}$$

Proof. It suffices to prove the result on words. Let $w = w_1 \cdot \ldots \cdot w_n$ be a word of length n and let $a \in [n+1]$. Consider the two terms on the left hand side applied to w:

$$(Q_{n+1} \circ \Phi_a)(w) = \sum_{1 \le i \le j \le n+1} \frac{1}{j}(i\,j)(w \cdot a) = \frac{1}{n+1}\sum_{\substack{j=n+1 \\ 1 \le i \le n+1}} (i\,j)(w \cdot a) + \sum_{1 \le i \le j \le n} \frac{1}{j}(i\,j)(w \cdot a) \tag{31}$$

$$(\Phi_a \circ Q_n)(w) = \left(\sum_{1 \le i \le j \le n} \frac{1}{j}(i\,j)(w)\right) \cdot a. \tag{32}$$

The second summation in (31) cancels with (32) because the adjoined a is in the $(n+1)$-th place, therefore it never moves and may be brought outside. This leaves us with the following:

$$(Q_{n+1} \circ \Phi_a - \Phi_a \circ Q_n)(w) = \frac{1}{n+1}\sum_{1 \le i \le n+1} (i \;\; n+1)(w \cdot a). \tag{33}$$

If $i = n+1$ we move nothing, giving the term $w \cdot a = \Phi_a(w)$. Otherwise we apply the transposition $(i \;\; n+1)$ to $w \cdot a$. This has the same effect as replacing the i^{th} symbol w_i in w with a and then appending w_i on the end of the new word. Since we do this for all symbols in w, the net effect is the same as $\sum_{1 \le b \le n} \Phi_b \circ \Theta_{b,a}$ applied to w. (The operator $\Theta_{b,a}$ systematically finds all occurrences of the letter b in w and replaces with an a, and then Φ_b puts the b back on the end. Since $w \in W^n$, all possibilities are exhausted by letting b range over $1 \le b \le n$.) This completes the proof. □

In terms of shuffling cards, we can interpret (30) as taking into account the difference between shuffling a deck and then adding a card versus adding a card and then shuffling. If we can understand how the operators Φ_a and $\Theta_{b,a}$ behave, this then inductively tells us how the shuffle on $n+1$ cards behaves using information about the shuffle on n cards. Equation (30) is our analogue of the similar equation found in [7, Theorem 38]. We now record a key property of the switching operators $\Theta_{a,b}$.

Lemma 33 (See Section 2.9 of [15])**.** *The maps $\Theta_{b,a}$ are $\mathfrak{S}_n$-module morphisms.*

Proof. This is clear from the definitions: since S_n is acting by place permutations, it amounts to the same thing to replace an occurrence of the symbol b with a symbol a and then permute the word as to first permute the word and then replace the same symbol b in its new position with an a. □

The above result is used to prove our next important lemma. Recall that given a partition λ we may add boxes to it in certain places to form a new partition. By our blurring of the distinction between partitions of n and n-tuples, if we add a box on row i the new partition formed is $\lambda + e_i$. Our next lemma shows how our switching operators behave when restricted to Specht modules.

Lemma 34 (Lemma 44 of [7])**.** *Let $\lambda, \alpha \vdash n$ be such that $\lambda + e_a = \alpha + e_b$ for some $a, b \in [n]$. Then $\Theta_{b,a}$ is non-zero on S^λ if and only if λ dominates α. In particular, if $a < b$ then $\Theta_{b,a}(S^\lambda) = 0$.*

Proof. Since S^λ is simple and $\Theta_{b,a}$ is a module homomorphism, the image $\Theta_{b,a}(S^\lambda)$ is 0 or isomorphic to S^λ, by Schur's lemma. But $\Theta_{b,a}(S^\lambda)$ lies in M^α because of the relationship $\lambda + e_a = \alpha + e_b$, and M^α has a submodule isomorphic to S^λ if and only if λ dominates α (Lemma 25). This gives the first assertion of the lemma.

To finish, note that in terms of diagrams the fact that $\lambda + e_a = \alpha + e_b$ corresponds to the fact that we can get from the diagram for λ to that for α by moving a box from row b to row a. Hence, under the given hypothesis, we have that λ dominates α if and only if $a \geq b$. □

The preceding result shows that when we restrict equation (30) to a Specht module S^λ we can change the index of the summation in the final term on the right hand side, as follows.

Corollary 35 (Similar to Corollary 45 of [7])**.**

$$(Q_{n+1} \circ \Phi_a - \Phi_a \circ Q_n)|_{S^\lambda} = \frac{1}{n+1} \Phi_a|_{S^\lambda} + \frac{1}{n+1} \sum_{1 \leq b \leq a} \Phi_b \circ \Theta_{b,a}|_{S^\lambda} \tag{34}$$

Having restricted equation (30) to the Specht module S^λ, we now analyse the image in the module $M^{\lambda+e_a}$ (note that it is clear from the left hand side of (30) that we do land in $M^{\lambda+e_a}$).

Lemma 36 (Lemma 41 of [7])**.** *Suppose $\lambda \vdash n$ and $\lambda + e_a \vdash n+1$. Then the subspace $\Phi_a(S^\lambda)$ is contained in an $\mathfrak{S}_{n+1}$-submodule of $M^{\lambda+e_a}$ that is isomorphic to $\oplus_\mu S^\mu$, where the sum ranges over the partitions μ obtained from λ by adding a box in row i for $i \leq a$.*

Proof. Let w be a word of length n, so that $\Phi_a(w) = w \cdot a$. If the symbol b does not occur in w then

$$\Phi_a(w) = \Theta_{b,a}(\Phi_b(w)).$$

Let $b = l(\lambda) + 1$, so b does not appear in any $w \in M^\lambda$, and consider the $\mathfrak{S}_{n+1}$-submodule N of $M^{\lambda+e_b}$ generated by the elements $x \cdot b$ with $x \in S^\lambda$,

$$N = \langle x \cdot b : x \in S^\lambda \rangle.$$

The submodule N is isomorphic to $\mathrm{Ind}_{\mathfrak{S}_n \times \mathfrak{S}_1}^{\mathfrak{S}_{n+1}}(S^\lambda \otimes S^1)$ (this is essentially the definition of how to induce), and using the branching rules on S_n this decomposes as a multiplicity free direct sum of Specht modules S^μ, where $\mu \vdash n+1$ and $\lambda \subset \mu$ [15, Theorem 2.8.3]. Using the observation at the start of the proof, we obtain

$$\Phi_a(S^\lambda) = \Theta_{b,a}(\Phi_b(S^\lambda)) \subseteq \Theta_{b,a}(\langle \Phi_b(S^\lambda) \rangle) = \Theta_{b,a}(N) \cong \bigoplus_{\mu \supseteq \lambda} \Theta_{b,a}(S^\mu).$$

Now note that $\Theta_{b,a}$ sends any word with evaluation $\lambda + e_b$ to a word with evaluation $\lambda + e_a$, and hence $\Theta_{b,a}(M^{\lambda+e_b}) \subseteq M^{\lambda+e_a}$. It follows that all nonzero summands S^μ appearing on the right hand side occur for $\mu \vdash n+1$ dominating $\lambda + e_a$, and then by Lemma 34 we can conclude that μ is obtained from λ by adding a cell in row i with $i \leq a$, as required. □

Recall that for any $\mathfrak{S}_n$-module V and a partition $\lambda \vdash n$, we have the *isotypic projection* $\pi^\lambda : V \to V$ which projects onto the S^λ-isotypic component of V. Using these projections, we can now define our lifting operators, which will be proven to map eigenvectors of Q_n to those of Q_{n+1}.

Definition 37. *Suppose* $\lambda \vdash n$ *and* $\lambda + e_a = \mu \vdash n+1$ *are two partitions. Define the* lifting operator

$$\kappa_a^{\lambda,\mu} := \pi^\mu \circ \Phi_a : S^\lambda \to S^\mu \subseteq M^\mu.$$

Note that since $\Phi_a(S^\lambda) \subseteq M^\mu$ *and* M^μ *contains a unique copy of* S^μ*, the image of* S^λ *under* $\kappa_a^{\lambda,\mu}$ *is actually contained in* S^μ*.*

The next results give some properties of the lifting operators; these properties depend in an essential way on the choice of Φ_a above. Our choice for Φ_a gives the eigenspaces a different structure to those in [7], in particular we are able to find *all* the eigenvectors for a module S^μ by looking at lifted eigenvectors from partitions $\lambda \subset \mu$.

Corollary 38. *For any* $\lambda \vdash n$ *and* $\lambda + e_a \vdash n+1$*, there exists some* $v \in S^\lambda$ *with*

$$\kappa_a^{\lambda,\lambda+e_a}(v) \neq 0.$$

Proof. If $\kappa_a^{\lambda,\lambda+e_a}(S^\lambda) = 0$, then the image $\Phi_a(S^\lambda)$ lies in the kernel of the projection $\pi^{\lambda+e_a} : M^{\lambda+e_a} \to S^{\lambda+e_a}$, which is an $\mathfrak{S}_{n+1}$-submodule with no component equal to $S^{\lambda+e_a}$. Hence the submodule generated by $\Phi_a(S^\lambda)$ has no component equal to $S^{\lambda+e_a}$. But we previously observed that (with notation as in the proof of Lemma 36)

$$\langle \Phi_a(S^\lambda)\rangle = \langle \Theta_{b,a}(\Phi_b(S^\lambda))\rangle = \Theta_{b,a}(\langle \Phi_b(S^\lambda)\rangle) \cong \Theta_{b,a}(N) \cong \bigoplus_{1\le i\le a} S^{\lambda+e_i}.$$

Since the right hand side contains $S^{\lambda+e_a}$ as a summand, we have a contradiction. □

We already know the map π^μ is an $\mathfrak{S}_{n+1}$-module morphism. Let us realise $\mathfrak{S}_n$ inside $\mathfrak{S}_{n+1}$ as the stabilizer of the $(n+1)^{\text{th}}$ position. Then any $\mathfrak{S}_{n+1}$-module gives rise to an $\mathfrak{S}_n$-module by restriction.

Lemma 39. *The linear operator* $\kappa_a^{\lambda,\lambda+e_a}$ *is a* $\mathfrak{S}_n$*-module morphism with trivial kernel.*

Proof. The key observation is that $\Phi_a(\sigma(v)) = \sigma(\Phi_a(v))$ for all $v \in S^\lambda$ and $\sigma \in \mathfrak{S}_n \subset \mathfrak{S}_{n+1}$. This is obvious, since Φ_a adds an element in the final position which is not affected by σ. Hence $\kappa_a^{\lambda,\lambda+e_a}$ is the composition of two $\mathfrak{S}_n$-module morphisms. The final observation follows from Corollary 38 – since $\kappa_a^{\lambda,\lambda+e_a}$ is a nonzero module morphism with a simple module as its domain, it must be injective by Schur's lemma. □

That the maps $\kappa_a^{\lambda,\lambda+e_a}$ are injective is a key point which simplifies the analysis in this paper compared to that in [7], where the lifting operators can kill eigenvectors. The next results show that $\kappa_a^{\lambda,\lambda+e_a}$ does indeed lift eigenvectors of Q_n into those of Q_{n+1}. To establish this we apply our projection $\pi^{\lambda+e_a}$ to equation (34). We can now state our versions of [7, Lemma 48, Theorem 49]; the proofs follow *mutatis mutandis* from the ones given there (the changes needed are to the constants in equation (30)).

Lemma 40 (Lemma 48 of [7])**.** *For* $\lambda = (\lambda_1, \dots, \lambda_r) \vdash n$*,* $a \in \{1, 2, \dots r+1\}$ *and* $\mu = \lambda + e_i$ *for some* $1 \le i \le a$*,*

$$Q_{n+1} \circ \kappa_a^{\lambda,\mu} - \kappa_a^{\lambda,\mu} \circ Q_n = \frac{2+\lambda_a - a}{n+1}\kappa_a^{\lambda,\mu} + \frac{1}{n+1}\sum_{i\le b\le a} \Theta_{b,a} \circ \kappa_b^{\lambda,\mu}.$$

Proof. This follows from the work in [7]: because we have not changed the switching operators $\Theta_{b,a}$ the proof still holds. The values on the right hand side change to reflect the difference in our equation (30). □

Theorem 41 (Theorem 49 of [7])**.** *For* $\lambda = (\lambda_1, \dots, \lambda_r) \vdash n$*,* $a \in \{1, 2, \dots r+1\}$ *and* $\mu = \lambda + e_i$ *with* $1 \le i \le a$*,*

$$Q_{n+1} \circ \kappa_a^{\lambda,\mu} - \kappa_a^{\lambda,\mu} \circ Q_n = \frac{(2+\lambda_i - i)}{n+1}\kappa_a^{\lambda,\mu}.$$

In particular if $v \in S^\lambda$ *is an eigenvector of* Q_n *with eigenvalue* ε*, then either* $\kappa_a^{\lambda,\mu}(v) = 0$ *or* $\kappa_a^{\lambda,\mu}(v)$ *is an eigenvector of* Q_{n+1} *with eigenvalue*

$$\varepsilon + \frac{2+\lambda_i - i}{n+1}.$$

Proof. This proof also follows from the work in [7]. □

The above theorem tells us exactly how to turn eigenvectors of Q_n into those of Q_{n+1} and, crucially, it also shows how the eigenvalues change in value. The final part of the analysis rests on showing that *all* of the eigenvectors in a Specht module S^μ can be retrieved by lifting from Specht modules S^λ with $\mu = \lambda + e_a$. In fact, we show that these lifted eigenvectors form a basis of S^μ.

Theorem 42. *For any $\mu \vdash n+1$ we may find a basis of eigenvectors of Q_{n+1} for the module S^μ, by lifting the eigenvectors of Q_n belonging in the modules S^λ with $\lambda \vdash n$ and $\lambda \subset \mu$.*

Proof. We proceed by induction, for $n = 1$ we know that the simple modules $S^{(2)}, S^{(1,1)}$ of $\mathfrak{S}_{n+1}$ are both one dimensional. Therefore, the eigenvector $a \in S^{(1)}$ when lifted indeed forms a basis for each simple module.

Consider the simple module of S^μ with $\mu \vdash n+1$. We know classically from the branching rules of S_n [15, Theorem 2.8.3] that the restriction of this module to $\mathfrak{S}_n$ is given by the

$$\mathrm{Res}_{\mathfrak{S}_n}^{\mathfrak{S}_{n+1}}(S^\mu) \cong \bigoplus_{\substack{\lambda \vdash n \\ \lambda \subset \mu}} S^\lambda.$$

Now suppose we have a basis of eigenvectors for every S^λ. By Lemma 39 the map $\kappa^{\lambda,\mu}(S^\lambda)$ gives a basis for S^λ inside of the vector space of $\mathrm{Res}_{\mathfrak{S}_n}^{\mathfrak{S}_{n+1}}(S^\mu)$ which is the same vector space as S^μ. Hence, considering all of the lifted eigenvectors from every S^λ together we find a basis for S^μ. By Theorem 41 the lifted eigenvectors form a basis of eigenvectors for S^μ. □

Inductively, for any $\lambda \vdash n$, Theorem 42 gives us the way to find all the eigenvectors for Q_n belonging to the Specht module S^λ: starting at $\varnothing$ and recursively applying lifting operators until we reach S^λ will give us an eigenvector, and all eigenvectors arise in this way. Note that $S^\varnothing$ has no eigenvectors attached to it, but we allow $\varnothing$ to be an eigenvector with eigenvalue 0, and $\Phi_a(\varnothing) = a$. This agrees with the formula in Theorem 41 because a is the only eigenvector of P_1 with eigenvalue $1 = 0 + (2+0-1)/(1)$. The inductive process of lifting naturally forms one path up Young's lattice which starts at $\varnothing$ and ends at λ. Furthermore, by Theorem 42 each unique path we take $\varnothing \to \lambda$ will result in a distinct eigenvector for S^λ, and all these eigenvectors together form a basis. We now are in a position to prove Theorem 5.

Proof of Theorem 5. Every eigenvector in our constructed basis gives a distinct eigenvalue of S^λ, hence there are d_λ distinct eigenvalues. These are eigenvalues for the shuffle Q_n, and each one appears d_λ times due to the isomorphism in Corollary 26. Therefore overall we have found $\sum_{\lambda \vdash n} d_\lambda^2 = n!$ eigenvalues and thus have a complete set. □

Proof of Lemma 6. Given a standard tableau T of size n, we build it up following its labelling and keeping track of the changes in eigenvalue given by Theorem 41. When box (i,j) is added to T we get a change in eigenvalue of $\frac{2+\lambda_i - i}{n+1} = \frac{2+(j-1)-i}{T(i,j)} = \frac{j-i+1}{T(i,j)}$. After summing these changes for all boxes (i,j) in T we divide by the size of T to normalise the eigenvalue. □

References

[1] Nathanaël Berestycki and Batı Şengül. Cutoff for conjugacy-invariant random walks on the permutation group. *Probability Theory and Related Fields*, 173(3-4):1197–1241, 2019.

[2] Nathanaël Berestycki and Richard Pymar. Effect of scale on long-range random graphs and chromosomal inversions. *The Annals of Applied Probability*, 22(4):1328–1361, 2012.

[3] Nathanaël Berestycki, Oded Schramm, and Ofer Zeitouni. Mixing times for random k -cycles and coalescence-fragmentation chains. *The Annals of Probability*, 39(5):1815–1843, 2011.

[4] Megan Bernstein and Evita Nestoridi.

[5] Persi Diaconis. Group representations in probability and statistics. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA, 1988.

[6] Persi Diaconis and Mehrdad Shahshahani. Generating a random permutation with random transpositions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 57(2):159–179, 1981.

[7] Antonius B. Dieker and Franco V. Saliola. Spectral analysis of random-to-random markov chains. *Advances in Mathematics*, 323:427–485, 2018.

[8] Gordon James and Adalbert Kerber. The Representation Theory of the Symmetric Group. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1984.

[9] Hubert Lacoin. Mixing Time and Cutoff for the Adjacent Transposition Shuffle and the Simple Exclusion. *The Annals of Probability*, 44(2):1426–1487, 2016.

[10] Nadia Lafrenière. Eigenvalues of symmetrized shuffling operators. *Preprint, arXiv:1811.07196*, 2018.

[11] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. Markov Chains and Mixing Times. American Mathematical Society, Providence, RI, 2008.

[12] Elchanan Mossel, Yuval Peres, and Alistair Sinclair. Shuffling by semi-random transpositions. In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 572–581. IEEE, 2004.

[13] Ross G. Pinsky. Cyclic to random transposition shuffles. *Preprint, arXiv:1204.2081*, 2012.

[14] Richard Pymar. Partial mixing of semi-random transposition shuffles. *Preprint, arXiv:1302.2601*, 2013.

[15] Bruce E. Sagan. The symmetric group - representations, combinatorial algorithms, and symmetric functions. Wadsworth & Brooks / Cole mathematics series. Wadsworth, 1991.

[16] Laurent Saloff-Coste. Random walks on finite groups. In *Probability on discrete structures*, pages 263–346. Springer, 2004.