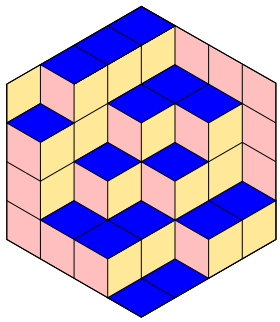


ALGEBRAIC COMBINATORICS

Alperen Özdemir

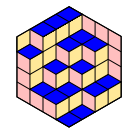
Random walks generated by the Ewens distribution on the symmetric group

Volume 6, issue 4 (2023), p. 907-927.

https://doi.org/10.5802/alco.290

Algebraic Combinatorics is published by The Combinatorics Consortium and is a member of the Centre Mersenne for Open Scientific Publishing
www.tccpublishing.org www.centre-mersenne.org
e-ISSN: 2589-5486

Algebraic Combinatorics
Volume 6, issue 4 (2023), p. 907–927
https://doi.org/10.5802/alco.290



Random walks generated by the Ewens distribution on the symmetric group

Alperen Özdemir

Abstract This paper studies Markov chains on the symmetric group S_n where the transition probabilities are given by the Ewens distribution with parameter $\theta > 1$. The eigenvalues are identified to be proportional to the content polynomials of partitions. We show that the mixing time is bounded above by a constant depending only on the parameter if θ is fixed. However, if it agrees with the number of permuted elements ($\theta = n$), the sequence of chains has a total variation cutoff at $\frac{\log n}{\log 2}$.

1. Introduction

The Ewens distribution on the symmetric group is defined by probabilities that are exponentially proportional to the number of cycles of permutations, with a base parameter $\theta > 0$. It originated in population genetics [12], then has gained a broader reach. It is initially defined on partitions of integers, which is related to the symmetric group by cycle decompositions. It arises from well-known cycle generating processes, Feller Coupling and the Chinese restaurant process. (See [3].) Crane [6] has an extensive survey on Ewens sampling formula.

The Ewens distribution can also be considered a deformation of the uniform distribution on the symmetric group. Hanlon [14] studied a Metropolis Markov chain driven by random transpositions with stationary distribution as the Ewens measure. The spectral analysis of this chain is further carried out in [10], and the mixing time is studied. More recently, Jiang [15] showed a cutoff result for the random transposition Metropolis chain.

The chains we study in this paper have transition probabilities given by Ewens distributions. At each stage, a permutation is randomly chosen and multiplied by the permutation of the current state. We are interested in finding the mixing time of the chains depending on the base parameter θ of the distribution. A more precise description is as follows:

Let σ and τ be two permutations in S_n. Define $\gamma(\sigma)$ to be the number of cycles in σ. For a given $\theta > 1$, possibly a function of n, we raise it to the number of cycles

Manuscript received 19th September 2022, accepted 31st December 2022.

Keywords. Ewens distribution, mixing time, random walks on groups, YJM elements.

Acknowledgements. The author would like to thank Jason Fulman for the suggestion of the problem.

of permutations to obtain the probabilities. So, if the chain is at permutation σ at a certain stage, the probability of moving to τ at the next stage is

$$P_\theta(\tau\sigma^{-1}) = \frac{\theta^{\gamma(\tau\sigma^{-1})}}{\theta^{(n)}}, \tag{1}$$

where $\frac{1}{\theta^{(n)}}$ is the normalizing constant. Explicitly,

$$\theta^{(n)} = \sum_{i=0}^{n} \begin{bmatrix} n \\ i \end{bmatrix} \theta^i = \theta(\theta+1)\cdots(\theta+n-1),$$

where $\begin{bmatrix} n \\ i \end{bmatrix}$ is a Stirling number of the first kind, which counts the number of permutations of cycle length i in S_n (see [2, Section 1.4]).

Let us briefly mention some properties of our Markov chain. The chain is irreducible and aperiodic, since the transition probability is non-zero for any permutation. It is time-homogenous, simply because we have the same transition probabilities regardless of the stage. It is also reversible, following from the fact that any permutation and its inverse, which have the same cycle type, have the same probability of being selected at any stage. From the very same fact, the transition matrix is also symmetric, which implies it is doubly stochastic. Therefore, the stationary distribution is the uniform distribution over S_n.

Next we define the distance concept used throughout the paper. Let the probability assigned to $\sigma \in S_n$ be $P_\theta^t(\sigma)$ after running the chain for t steps. The total variation distance between P_θ^t and the uniform distribution π is

$$d_\theta(t) := \|P_\theta^t - \pi\|_{TV} = \frac{1}{2}\sum_{\sigma\in S_n} |P_\theta^t(\sigma) - \pi(\sigma)| = \max_{S\subseteq S_n} |P_\theta^t(S) - \pi(S)|. \tag{2}$$

Our goal is to investigate the convergence time with respect to the total variation distance. It is a common phenomenon in this setting that the convergence to the stationary distribution shows a *cutoff*, a sharp transition. The following definition of cutoff is found in [20]. A sequence of Markov chains exhibits a cutoff at the *mixing time* $\{t_n\}$ with a *window* of size $\{w_n\}$ if

(i) $\lim_{n\to\infty} \frac{w_n}{t_n} = 0$,
(ii) $\lim_{\gamma\to-\infty} \liminf_{n\to\infty} d_{TV}(t_n + \gamma w_n) = 1$,
(iii) $\lim_{\gamma\to\infty} \limsup_{n\to\infty} d_{TV}(t_n + \gamma w_n) = 0$.

Also, see [8] for an insightful treatment of the concept and [24] for another definition along with a long list of examples. Our result reads:

Theorem 1.1. *Let P_θ be the Markov chain on S_n defined above in* (1)*, and $c > 1$ be a constant.*

(i) *If $t \geqslant c\frac{\theta^2+1}{2}$, then $d_\theta(t) < C(\theta)e^{-c}$ for some constant $C(\theta)$ depending only on θ.*
(ii) *If $\theta = n$, then the chains have a total variation cutoff at $t_n = \frac{\log n}{\log 2}$ with $w_n = 1$.*

For the rest of the paper, we first present the representation theory tools in Section 2. Then we identify the eigenvalues of the Markov chain using Young–Jucys–Murphy elements. They turn out to be functions of contents of Young diagrams. In the following section, we prove the first part of Theorem 1.1 by using Plancherel growth process. We also provide a simpler proof for only the integer values of θ. In Section 4, we prove the second part of the theorem.

2. Representation theory techniques

We summarize the techniques that are used in the following sections. We start with the method developed in [11] to study the random transposition walk. For the connection to the group representation theory, first consider the *Fourier transform* of the measure μ defined on a finite group G, evaluated at any representation λ of G,

$$\widehat{\mu}(\lambda) = \sum_{g \in G} \mu(g)\lambda(g).$$

Then its inverse transform is

$$\mu(g) = \sum_{\lambda \text{ irred.}} d_\lambda tr(\lambda(g^{-1})\widehat{\mu}(\lambda)),$$

where the sum is over all irreducible representations of G and d_λ stands for the dimension of the representation λ. The inverse transform leads to Plancherel's formula stated below:

$$\sum_{g \in G} |\mu(g)|^2 = \frac{1}{|G|} \sum_{\lambda \text{ irred.}} d_\lambda tr(\widehat{\mu}(\lambda)\widehat{\mu}(\lambda)^T). \tag{3}$$

It establishes the connection with the total variation distance if we take μ on the left-hand side to be the difference of two measures.

We also note a fact about the Fourier transform of convolutions of measures, as a Markov chain at step t can be viewed as t-fold convolution measure of the transition probabilities. The convolution of two measures μ and ν is defined to be

$$\mu * \nu(g) = \sum_{h \in G} \mu(gh^{-1})\nu(h).$$

The Fourier transform of a convolution satisfies

$$\widehat{\mu * \nu}(g) = \widehat{\mu}(g)\widehat{\nu}(g). \tag{4}$$

2.1. Young–Jucys–Murphy elements and eigenvalues.

We will focus on the symmetric group S_n. The irreducible representations of S_n are indexed by partitions of integers. We use the notation $\lambda = (\lambda_1, \lambda_2, \ldots) \vdash n$ for partitions of n. λ' refers to the conjugate partition of λ. The dimension of an irreducible representation λ, d_λ, is equal to the number of standard Young tableaux of shape λ. See, for instance, [23, Chapter 2].

First we state an upper bound theorem, which was proved in [7, Chapter 3B] using Plancherel formula (3). As we identify the eigenvalues, we construct the upper bound at the same time. We first note that the irreducible representations of the symmetric group S_n are indexed by partitions of n; see [23, Chapter 2].

Lemma 2.1 ([7]). *Let μ be a probability distribution over S_n and π be the uniform distribution. Then,*

$$\|\mu - \pi\|_{TV}^2 \leqslant \frac{1}{4} \sum_{\substack{\lambda \vdash n \\ \lambda \neq (n)}} d_\lambda \, tr(\widehat{\mu}(\lambda)\widehat{\mu}(\lambda)^T).$$

In our case, we replace μ by P_θ^t. The first of the two observations to be made on P_θ^t, in order to simplify the right-hand side, is P_θ^t is t-fold convolution of P_θ. Therefore, by (4), $\widehat{P_\theta^t}(\lambda) = \widehat{P_\theta}(\lambda)^t$.

For the second one, first consider $\widehat{P_\theta}$ as an element of the group algebra, $\mathbb{C}(S_n)$. A well-known fact about the center of the group algebra, $\mathbf{Z}(\mathbb{C}(S_n))$, is that it has basis consisting of the sum of the all elements in the same conjugacy class. Since the

density function defined by P_θ has an equal weight for the permutations in the same conjugacy class, $\widehat{P_\theta}(\lambda)$ is in $\mathbf{Z}(\mathbb{C}(S_n))$. But then $\widehat{P_\theta}$ is a constant times the identity matrix of the same dimension with λ, since the representation of any element in the center is in that form as a result of Schur's lemma. We have more to infer from this fact, but for now we use below the implication $\widehat{P_\theta}(\lambda) = \widehat{P_\theta}(\lambda)^T$.

Putting two observations together, the upper bound can be expressed as

$$\|P_\theta^t - U\|_{TV} \leqslant \frac{1}{4} \sum_{\substack{\lambda \vdash n \\ \lambda \neq (n)}} d_\lambda tr(\widehat{P_\theta}(\lambda)^{2t}). \tag{5}$$

Next, we evaluate $\widehat{P_\theta}(\lambda)$. We already stated that $\widehat{P_\theta}(\lambda)$ is a constant multiple of the identity matrix. In order to identify the constant we first introduce Young–Jucys–Murphy elements. The detailed treatment of the subject can be found in [5, 26]. For $i = 2, ..., n$, consider the following elements $R_i \in \mathbb{C}(S_n)$:

$$R_i = (1, i) + (2, i) + \cdots + (i - 1, i), \tag{6}$$

which were independently introduced in [16] and [22]. These elements form a basis for the Gelfand–Tsetlin subalgebra of $\mathbb{C}(S_n)$, which is generated by the centers of the symmetric groups $S_1, .., S_n$. It is a maximal commutative subalgebra, and is the algebra of all operators diagonal in the Gelfand–Tsetlin basis. The diagonal entries of the irreducible representations of R_i are the *contents* of boxes that contain i in the standard Young tableaux. The content of a box in the Young diagram of is defined by its position in the diagram, it is the column number of the box minus the row number of it.

To be more specific, let us take a partition λ of n. There are d_λ standard Young tableaux of λ, each of which is associated with a fixed diagonal entry of $\lambda(R_i)$ for all $i = 2, ..., n$. In fact, the order of the diagonal entries does not matter for our purpose (see Corollary 2.4), so we only consider the first entry. We let the upper-left diagonal of $\lambda(R_i)$ be determined by the position of i in the standard Young tableau which has entries according to the lexicographic order, i.e., 1 through λ_1 in the first part, $\lambda_1 + 1$ through $\lambda_1 + \lambda_2 + 1$ in the second part and so on. Denote the upper-left corner element of $\lambda(R_i)$ by $c_\lambda(i)$. So the $d_\lambda \times d_\lambda$ diagonal matrix $\lambda(R_i)$ is in the form

$$\lambda(R_i) = \begin{pmatrix} c_\lambda(i) & & & \\ & * & & \\ & & \ddots & \\ & & & * \end{pmatrix}.$$

We give an example below: We give an example below:

Example 2.2. Let $\lambda = (3, 1) \vdash 4$. The Young tableaux for this partition are

$$T_1 = \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 \\ \cline{1-1} \end{array} \qquad T_2 = \begin{array}{|c|c|c|} \hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1} \end{array} \qquad T_3 = \begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 \\ \cline{1-1} \end{array}$$

The ith diagonal entry of $\lambda(R_j)$ is the content of the box containing j in T_i. So we have

$$\lambda(R_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \ \lambda(R_3) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \lambda(R_4) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Although no R_i for $i \in \{2, ..., n\}$ is in the center of the group algebra, any symmetric polynomial with variables $R_2, ..., R_n$ is in $\mathbf{Z}(\mathbb{C}(S_n))$. See [13, Lecture 2] for the details. We are particularly interested in the elementary symmetric functions of YJM elements, the reason is the following theorem:

THEOREM 2.3 ([9]). *Let R_i be defined as in* (6)*. The kth elementary symmetric function in $R_2, ..., R_n$ is the sum over all permutations in conjugacy classes of S_n with $n-k$ cycles in the group algebra, i.e.,*

$$e_k(R_2, ..., R_n) = \sum_{2 \leqslant i_1 < \cdots < i_k \leqslant n} R_{i_1} \cdots R_{i_k} = \sum_{\substack{\sigma \in S_n \\ \gamma(\sigma) = n-k}} \sigma.$$

A useful corollary, which is related to our observations on the representations of R_i, is as follows:

COROLLARY 2.4. *Let λ be an irreducible representation of S_n. Then*

$$\lambda\left(e_k(R_2, ..., R_n)\right) = e_k(c_\lambda(2), ..., c_\lambda(n))\, I_{d_\lambda}.$$

Proof. Since $e_k(R_2, ..., R_n)$ is the sum of elements in the same conjugacy classes, it is in the center of the group algebra. Therefore, its representation is a constant multiple of the identity matrix by Schur's lemma. So it suffices to consider the upper-left corner entries of each $\lambda(R_i)$ for $i \in \{2, ..., n\}$, which are $c_\lambda(i)$'s. □

Now we are ready to calculate $\widehat{P_\theta}(\lambda)$. By Theorem 2.3 and its corollary,

$$\begin{aligned}
\widehat{P_\theta}(\lambda) &= \sum_{\sigma \in S_n} P_\theta(\sigma)\lambda(\sigma) \\
&= \sum_{k=0}^{n-1} \frac{\theta^{n-k}}{\theta^{(n)}} \sum_{\substack{\sigma \in S_n \\ \gamma(\sigma) = n-k}} \lambda(\sigma) \\
&= \sum_{k=0}^{n-1} \frac{\theta^{n-k}}{\theta^{(n)}}\, \lambda\left(e_k(R_2, ..., R_n)\right) \\
&= \left(\sum_{k=0}^{n-1} \frac{\theta^{n-k}}{\theta^{(n)}}\, e_k(c_\lambda(2), ..., c_\lambda(n)) \right) I_{d_\lambda} \\
&= \left(\frac{\theta^n}{\theta^{(n)}} \sum_{k=0}^{n-1} e_k\left(\frac{c_\lambda(2)}{\theta}, ..., \frac{c_\lambda(n))}{\theta} \right) \right) I_{d_\lambda}.
\end{aligned}$$

Next, we use the identity

$$\prod_{i=1}^{n}(1 + x_i) = \sum_{k=1}^{n} e_k(x_1, ..., x_n)$$

to obtain above,

$$\begin{aligned}
\widehat{P_\theta}(\lambda) &= \left(\frac{\theta^n}{\theta^{(n)}} \prod_{i=2}^{n} \left(1 + \frac{c_\lambda(i)}{\theta} \right) \right) I_{d_\lambda} \\
&= \left(\frac{\theta}{\theta^{(n)}} \prod_{i=2}^{n} (\theta + c_\lambda(i)) \right) I_{d_\lambda}
\end{aligned}$$

Since $c_\lambda(1) = 0$ for all $\lambda \vdash n$, the expression eventually simplifies to

$$\widehat{P_\theta}(\lambda) = \left(\prod_{i=1}^{n} \left(\frac{\theta + c_\lambda(i)}{\theta + i - 1} \right) \right) I_{d_\lambda}. \tag{7}$$

Since P_θ is a symmetric matrix, the eigenvalues of irreducible representations give complete list of eigenvalues of P_θ. See [11, Section 4] for the details. Hence, the eigenvalues of P_θ with their multiplicities are

$$\beta_{\lambda,\theta} = \prod_{i=1}^{n} \left(\frac{\theta + c_\lambda(i)}{\theta + i - 1} \right) \text{ with multiplicity } d_\lambda^2 \text{ for } \lambda \vdash n. \tag{8}$$

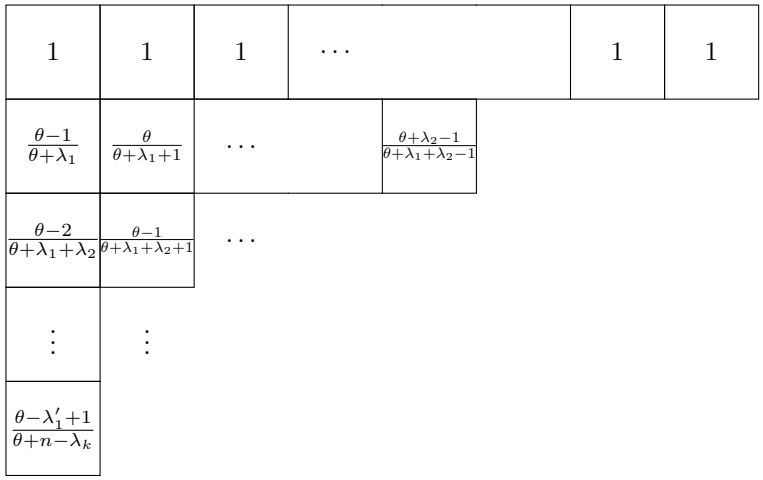


FIGURE 1. The factors of β_λ, the eigenvalue associated with $\lambda = (\lambda_1, \ldots, \lambda_k) \vdash n$.

Therefore the upper bound on the total variation distance (5) is reduced to

$$\begin{aligned} d_\theta(t) &\leqslant \frac{1}{4} \sum_{\substack{\lambda \vdash n \\ \lambda \neq (n)}} d_\lambda \, tr\left(\left(\prod_{i=1}^{n} \left(\frac{\theta + c_\lambda(i)}{\theta + i - 1} \right)^{2t} \right) I_{d_\lambda} \right) \\ &\leqslant \frac{1}{4} \sum_{\substack{\lambda \vdash n \\ \lambda \neq (n)}} d_\lambda^2 \prod_{i=1}^{n} \left(\frac{\theta + c_\lambda(i)}{\theta + i - 1} \right)^{2t}. \end{aligned} \tag{9}$$

3. Proof of Theorem 1.1 part (i)

In this section, we prove the first part of the theorem. It can be explicitly stated as saying that the convergence (to their stationary distributions) times of the Markov chains defined on S_n by the Ewens distributions with a fixed parameter θ are bounded above by a constant, which depends on only θ, for all n.

We first study a sum

$$Z(\theta) := \sum_{\lambda \vdash n} d_\lambda^2 \beta_{\lambda,\theta}^2, \tag{10}$$

which is none other than the sum of the eigenvalues of P_θ^2. We will relate this expression to upper bound on the mixing time. Observe that unlike (9) we included the trivial partition (n). We will calculate it explicitly. First we need some definitions.

3.1. Plancherel growth process. Let $\mathbb{Y}$ denote the *Young's lattice*, which is the lattice of all integer partitions. A partial order on $\mathbb{Y}$ is defined by the inclusion of Young diagrams, i.e., for $\lambda, \mu \in \mathbb{Y}$ we say μ is covered by λ, and denote it by $\mu \nearrow \lambda$, if λ can be obtained from μ by adding a box to an admissible position in its Young diagram. This box adding process gives a Markov chain on $\mathbb{Y}$, known as the Plancherel growth process; see [17]. Its transition probabilities are given by

$$p(\lambda, \Lambda) = \frac{d_\Lambda}{|\Lambda| d_\lambda} \text{ for } \lambda \nearrow \Lambda.$$

The law of the chain at stage n is the Plancherel measure on S_n, which is defined as

$$\mathcal{P}(\lambda) = \frac{d_\lambda^2}{n!} \text{ for } \lambda \vdash n.$$

We use this process to compute (10) recursively. Let us list some properties that the transition measure satisfies. Firstly,

$$d_\Lambda = \sum_{\lambda : \lambda \nearrow \Lambda} d_\lambda \quad \text{and} \quad d_\lambda = \frac{1}{|\Lambda|} \sum_{\Lambda : \lambda \nearrow \Lambda} d_\Lambda. \tag{11}$$

Secondly, we state the relation of this measure to contents, similar to the way it is presented in [18, Section 10]. Let X be the content of the random box added to the diagram of λ according to the transition measure. We have

$$\begin{aligned} \mathbf{E}_\lambda(X) &= \sum_{\Lambda : \lambda \nearrow \Lambda} c(\Lambda \smallsetminus \lambda) \frac{d_\Lambda}{|\Lambda| d_\lambda} = 0, \\ \mathbf{E}_\lambda(X^2) &= \sum_{\Lambda : \lambda \nearrow \Lambda} c(\Lambda \smallsetminus \lambda)^2 \frac{d_\Lambda}{|\Lambda| d_\lambda} = |\lambda|. \end{aligned} \tag{12}$$

Lemma 3.1. *If $Z_n(\theta)$ be defined as in* (10)*, then*

$$Z_n(\theta) = \binom{n+\theta^2-1}{n} \Big/ \binom{n+\theta-1}{n}^2.$$

Proof. By the first equation in (11), we have

$$Z_n(\theta) = \sum_{\lambda \vdash n} d_\lambda^2 \beta_{\lambda,\theta}^2 = \sum_{\mu \vdash n-1} d_\mu^2 \beta_{\mu,\theta}^2 \sum_{\lambda : \mu \nearrow \lambda} \frac{d_\lambda}{d_\mu} \left(\frac{\beta_{\lambda,\theta}}{\beta_{\mu,\theta}} \right)^2.$$

Next, using the formula (8) and the equations in (12),

$$\begin{aligned} Z_n(\theta) &= \sum_{\mu \vdash n-1} d_\mu^2 \beta_{\mu,\theta}^2 \sum_{\lambda : \mu \nearrow \lambda} \frac{d_\lambda}{d_\mu} \left(\frac{\theta + c(\lambda \smallsetminus \mu)}{n+\theta-1} \right)^2 \\ &= \frac{n}{(n+\theta-1)^2} \sum_{\mu \vdash n-1} d_\mu^2 \beta_{\mu,\theta}^2 \left(\mathbf{E}(\theta^2) + \mathbf{E}(2\theta X) + \mathbf{E}(X^2) \right) \\ &= n \frac{n+\theta^2-1}{(n+\theta-1)^2} Z_{n-1}(\theta). \end{aligned}$$

Since $Z_1(\theta) = 1$, the recursive relation above gives the result. □

We also note the following combinatorial identity. Since $Z(\theta)$ is in fact the sum of the eigenvalues of P_θ^2, it will agree with the trace of P_θ^2. Each element along the diagonal of P_θ^2 is obtained by the sum of squares of elements in any row, since P_θ is symmetric. Putting these observations together, we obtain:

$$\frac{1}{n!}\sum_{\sigma\in S_n}\theta^{2\gamma(\sigma)} = \binom{n+\theta^2-1}{n}.$$

If we let $Y(\pi)$ be the number of cycles in a uniformly chosen permutation $\pi \in S_n$, and define $\Gamma_n(\pi) = \theta^{Y(\pi)}$, we have

$$\mathbf{E}(\Gamma_n) = \binom{n+\theta-1}{n} \quad \text{and} \quad \mathbf{E}(\Gamma_n^2) = \binom{n+\theta^2-1}{n}.$$

The first identity above is obtained by computing the trace of P_θ.

3.2. Upper bound. Before evaluating the upper bound on the total variation distance, we give an upper bound on the eigenvalues. In particular we show that the eigenvalue associated with $\lambda = (n-1,1)$ is the largest among all but the trivial partition (n).

Lemma 3.2. *Let $\lambda \neq (n)$ be a partition of n. Then $\beta^2_{(n-1,1),\theta} \geqslant \beta^2_{\lambda,\theta}$ for all $\theta > 1$.*

Proof. We will compare the factors of the eigenvalues $\beta_{\lambda,\theta}$ and $\beta_{(n-1,1),\theta}$ going through all boxes in their diagrams according to lexicographic order. We ignore the product of denominators in (8), $\theta^{(n)}$, as its common to all partitions. Every partition has a box in the upper left corner of its diagram, so the first factors agree. Since λ is not the trivial partition, its diagram also includes a box in the leftmost position of its second row, which we cross out with the box of $(n-1,1)$ in the same position. Finally, we distinguish two cases, the positive (1) and the negative (2) factors of $\beta_{\lambda,\theta}$. Since $\theta > 1$, $\beta_{(n-1,1),\theta}$ does not have a negative factor.

Skipping the aforementioned two boxes, we start with the box in the first row and the second column of $(n-1,1)$. It has the factor $\theta+2-1 = \theta+1$. While the partition λ can have a box at the same position or a box at the second row and the second column or at the third row and the first column. In the last two cases, the factor is $\theta-1$, which is positive and smaller than of $(n-1,1)$. We repeat this argument for all boxes with positive factors.

Suppose the kth box of λ has a negative factor, i.e., $\theta+c_\lambda(k+2) < 0$. Since $c_\lambda(k+2)$ is minimized if the first k boxes of λ are located in th first column, we have $c_\lambda(k+2) > -(k+1)$. On the other hand, the kth factor for the partition $(n-1,1)$ is $\theta+k$. Since $\theta > 1$, we have $|\theta+k| \geqslant |\theta + c_\lambda(k+2)|$, which proves the second case and the lemma. □

We are ready to take on (5),

$$\begin{aligned} d_\theta(t) &\leqslant \sum_{\substack{\lambda\vdash n\\ \lambda\neq(n)}} d_\lambda^2 \beta_\lambda^{2(t+1)} \\ &\leqslant \left(\max_{\substack{\lambda\vdash n\\ \lambda\neq(n)}} \beta_\lambda^2\right)^t (Z(\theta)-1) \\ &\leqslant \left(\frac{\theta-1}{\theta+n-1}\right)^{2t} \binom{n+\theta^2-1}{n}\binom{n+\theta-1}{n}^{-2}. \end{aligned} \tag{13}$$

Next, we bound the binomial terms in the sum. Let us take $k \leqslant \theta < k+1$ for $k \in \mathbb{Z}^+$. For the first term, we have

$$\binom{n+\theta^2-1}{n} < \frac{n^{(k+1)^2}}{k^2!}.$$

We need a lower bound for the latter,

$$\binom{n+\theta-1}{n} > \frac{n^k}{(k+1)!}.$$

So, if we take $t = c\frac{k^2+1}{2}$ for some $c > 1$,

$$\begin{aligned} d_\theta(t) &\leqslant \frac{1}{4}\frac{n^{(k+1)^2}}{k^2!}\frac{(k+1)!^2}{n^{2k}}\left(\frac{k-1}{k+n-1}\right)^{c(k^2+1)} \\ &= C(k)n^{1-c} \end{aligned}$$

where $C(k)$ is a some constant depending only on $k = \lfloor\theta\rfloor$. Since c is also independent of n, we can take any fixed value for n, say $n \geqslant 3$, and conclude that $d_\theta(t) < C(k)e^{-c}$. This concludes the proof of Theorem 1.1 part (i).

3.3. An elementary proof for integer parameters. Let $k \in \mathbb{Z}^+$. By the hook-length formula, we may write $Z(k)$ explicitly as

$$Z(k) = \sum_{\lambda\vdash n} \frac{(n!)^2}{(\prod_{i=1}^n h(i))^2} \prod_{i=0}^{n-1}\left(\frac{k+c_\lambda(i+1)}{k+i}\right)^{2t}, \tag{14}$$

Then we can rearrange it as the following:

$$\begin{aligned} Z(k) &= \sum_{\lambda\vdash n}\left(\frac{n!}{\prod_{i=1}^n k+i-1}\right)^2\left(\prod_{i=1}^n \frac{k+c_\lambda(i)}{h(i)}\right)^2 \\ &= \binom{n+k-1}{n}^{-2}\sum_{\lambda\vdash n} s_\lambda^2(1^k) \end{aligned} \tag{15}$$

where $s_\lambda(1^k) = \prod_{i=0}^n \frac{k+c(i)}{h(i)}$ is the Schur function with k non-zero variables, each being equal to 1. Note that $s_\lambda(1^k)$ is zero unless $\lambda_1' \leqslant k$. See [21, Section I-3] for details.

Consider the Cauchy identity,

$$\prod_{i,j}(1-x_iy_j)^{-1} = \sum_{\lambda \text{ part.}} s_\lambda(\mathbf{x})s_\lambda(\mathbf{y})$$

where $\mathbf{x} = (x_1, x_2, ...)$, $\mathbf{y} = (y_1, y_2, ...)$ and the sum is over all partitions λ of any size. The proof is found in [21, Section I-4]. We are interested in partitions λ such that $|\lambda| = n$, so we consider the coefficient of t^n for the generating function

$$F(\mathbf{x}, \mathbf{y}; t) = \prod_{i,j}(1-x_iy_jt)^{-1}.$$

Then we take $x_1 = \cdots = x_k = 1$ and $x_l = 0$ for $l > k$. We do the same for $\mathbf{y}$ to obtain

$$\sum_{\lambda \text{ part.}} s_\lambda^2(1^k) = \prod_{i=1}^k\prod_{j=1}^k(1-t)^{-1} = (1-t)^{-k^2}.$$

The right-hand side is the generating function for complete symmetric functions, and the coefficient of t^n is given by $h_n = \binom{n+k^2-1}{n}$. Therefore,

$$\sum_{\lambda\vdash n} s_\lambda^2(1^k) = \binom{n+k^2-1}{n}.$$

We then can proceed as before and show the same upper bound. The reason that we cannot simply extend the result from integers to continuum is the failure of eigenvalue

monotonicity in the non-integer case. Observe that the eigenvalues are always non-negative for integer values of θ, which is not true in general.

4. Proof of Theorem 1.1 part (ii)

4.1. Defining representation. It follows from the definition of the total variation distance (2) that

$$d_\theta(t) \geqslant |P_\theta^t(A) - \pi(A)| \tag{16}$$

for any subset A of S_n. A statistic commonly used to choose an expedient subset in (16) is the number of fixed points. Define B_k to be the set of permutations with less than or equal to k fixed points. This particular choice of subsets establishes an important connection with the symmetric group representations.

The defining representation of S_n is the n-dimensional representation ρ where

$$\rho(\sigma)_{ij} = \begin{cases} 1 \ \sigma(j) = i \\ 0 \text{ otherwise} \end{cases}$$

for $\sigma \in S_n$. Denote the character of the representation, in other words, the trace of the matrices given above, by $\chi_\rho(\sigma)$. Observe that $\chi_\rho(\sigma)$ counts the number of fixed points of $\sigma \in S_n$. Therefore $\sigma \in B_k$ if and only if $\chi_\rho(\sigma) \leqslant k$.

Next, we consider χ_ρ as a random variable under the law of the Markov chain at step t. The defining representation is reducible and decomposed as

$$\rho = S^{(n)} \oplus S^{(n-1,1)} \tag{17}$$

where S^λ is the Specht module associated with partition λ of n, noting that the partitions of n are in one to one correspondence with the irreducible representations of S_n. See [23, Chapter 2] for details. For the second moment of χ_ρ, we consider the decomposition of $\rho \otimes \rho$ as

$$\rho \otimes \rho = 2S^{(n)} \oplus 3S^{(n-1,1)} \oplus S^{(n-2,2)} \oplus S^{(n-2,1^2)}. \tag{18}$$

Then we use the facts below,

$$\chi_{\rho_1 \oplus \rho_2} = \chi_{\rho_1} + \chi_{\rho_2},$$
$$\chi_{\rho_1 \otimes \rho_2} = \chi_{\rho_1} \cdot \chi_{\rho_2},$$

and invoke some Fourier analytic results, which can be found in [4, Chapter 16] with detailed proofs and in [7, Section 2C] in most relevance to our case, to calculate the first two moments of χ_ρ. We have the following expressions for the first two moments of a distribution μ over S_n.

$$\begin{aligned} &\mathbf{E}_\mu(\chi_\rho) = \operatorname{tr}(\hat{\mu}(n)) + \operatorname{tr}(\hat{\mu}(n-1,1)), \\ &\mathbf{E}_\mu(\chi_\rho^2) = 2\operatorname{tr}(\hat{\mu}(n)) + 3\operatorname{tr}(\hat{\mu}(n-1,1)) + \operatorname{tr}(\hat{\mu}(n-2,2)) + \operatorname{tr}(\hat{\mu}(n-2,1,1)). \end{aligned} \tag{19}$$

Then using the first two moments, the probabilities can be bounded by Chebyshev's inequality,

$$\mathbf{Pr}(\chi_\rho \leqslant \mathbf{E}(\chi_\rho) - a) \leqslant \frac{\operatorname{Var}(\chi_\rho)}{\operatorname{Var}(\chi_\rho) + a^2}. \tag{20}$$

In particular, the distribution of χ_ρ under the uniform distribution over S_n is well-known from the classical matching problem. It is given by

$$\pi(B_k) = \mathbf{Pr}_\pi(\chi_\rho \geqslant k) = \frac{1}{(k-1)!} \sum_{l=k}^{n} \frac{(-1)^{l-k}}{l(l-k)!} \tag{21}$$

for $k = 1, 2, \ldots,$ whose limiting distribution is Poisson with parameter one. See [25] for various derivations of this result.

4.2. Lower bound. The decompositions (17) and (18) of the defining representation allow us to evaluate the expression in (19) with respect to the law of P_θ^t. We first identify the traces of the Fourier transform by our previous derivation (7) for $\widehat{P_\theta^t}(\lambda)$ for $\lambda \vdash n$.

$$\begin{aligned}
\operatorname{tr}(\widehat{P_\theta^t}(n)) &= 1,\\
\operatorname{tr}(\widehat{P_\theta^t}(n-1,1)) &= (n-1)\left(\frac{\theta-1}{\theta+n-1}\right)^t,\\
\operatorname{tr}(\widehat{P_\theta^t}(n-2,2)) &= \frac{n(n-3)}{2}\left(\frac{\theta-1}{\theta+n-1}\right)^t\left(\frac{\theta-2}{\theta+n-2}\right)^t,\\
\operatorname{tr}(\widehat{P_\theta^t}(n-2,2)) &= \frac{(n-1)(n-2)}{2}\left(\frac{\theta-1}{\theta+n-1}\right)^t\left(\frac{\theta}{\theta+n-2}\right)^t.
\end{aligned}$$

If we take $t = \frac{\log n}{\log(\theta+n)-\log(\theta)} - \frac{\gamma}{\log(\theta+n)-\log(\theta)}$ where $\gamma > 0$, then

$$\begin{aligned}
\lim_{n\to\infty} \mathbf{E}_{P_\theta^t}(\chi_\rho) &= 1 + e^\gamma\\
\lim_{n\to\infty} \mathbf{E}_{P_\theta^t}(\chi_\rho^2) &= 2 + 3(1+e^\gamma) + e^{2\gamma}.
\end{aligned}$$

Therefore, $\operatorname{Var}_{P_\theta^t}(\chi_\rho)$ is asymptotically $4 + e^\gamma$.

We take n large enough and define A to be the set of permutations with less than or equal to $e^{\gamma-1}+1$ fixed points, i.e.,

$$A = \{\sigma \in S_n : \chi_\rho(\sigma) \leqslant e^{\gamma-1}+1\}.$$

Then by Chebyshev's inequality (20),

$$P_\theta^t(A) = \mathbf{Pr}_{P_\theta^t}(\chi_\rho \leqslant 1 + e^{\gamma-1}) \leqslant \frac{4+e^\gamma}{4+e^\gamma+e^{2\gamma-2}} < \frac{1}{e^{\gamma-2}}.$$

For the uniform distribution, we have

$$\pi(A) = \mathbf{Pr}_\pi(\chi_\rho \leqslant 1 + e^{\gamma-1}) \geqslant 1 - \frac{1}{e^\gamma!}$$

by (21). Putting them together,

$$\begin{aligned}
d_\theta(t) &= \max_{S\subseteq S_n} |P_\theta^t(S) - \pi(S)|\\
&\geqslant |P_\theta^t(A) - \pi(A)|\\
&\geqslant 1 - \frac{1}{e^\gamma!} - \frac{1}{e^{\gamma-2}} > 1 - \frac{1}{e^\gamma}.
\end{aligned}$$

Therefore, we have the lower bound

$$d_n\left(\frac{\log n}{\log 2} - \gamma\right) = \left\| P_\theta^{\frac{\log n}{\log 2}-\gamma} - \pi \right\|_{TV} \geqslant 1 - \frac{1}{e^\gamma}. \tag{22}$$

4.3. Dimension bounds and a partial order on partitions. We present bounds on the sum of d_λ^2's for certain sets of partitions to be used in the proof of the theorem. A bound for a fixed length of the first row is given by

Lemma 4.1 ([7]). *Let $|\lambda| = n$ and λ_1 be the length of the first row of λ. Then,*

$$\sum_{\lambda=(\lambda_1,\dots)} d_\lambda^2 \leqslant \binom{n}{\lambda_1}^2 (n-\lambda_1)!. \tag{23}$$

The proof in [7] is by bounding the number of standard Young tableaux. First choose λ_1 elements for the first row, then count all possible standard Young tableaux that can be formed from the remaining cells.

Next, by Lemma 4.1, we bound the sum of dimension squares for the partitions with their first row larger than a linear order of n. We use the notation $f(n) = \mathcal{O}(g(n))$ to mean $\limsup_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| < \infty$, and $f(n) = o(g(n))$ for $\limsup_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| = 0$.

Lemma 4.2. *Let $|\lambda| = n$ and $\alpha \in (0,1)$. Define $Q_1 = \{\lambda \vdash n : \alpha n \leqslant \lambda_1\}$ and $Q_2 = \left\{\lambda \vdash n : \alpha n \leqslant \lambda_1 \text{ and } \frac{\lambda_1}{2} \leqslant \lambda_2 \leqslant \lambda_1\right\}$. As $n \to \infty$, for all $\epsilon > 0$,*

(i) $\sum_{\lambda\in Q_1} d_\lambda^2 = \mathcal{O}(n^{(1-\alpha+\epsilon)n})$,

(ii) $\sum_{\lambda\in Q_2} d_\lambda^2 = \mathcal{O}(n^{(1-\frac{3}{2}\alpha+\epsilon)n})$.

Proof. The binomial term in (23) is simply bounded by 4^n considering

$$\binom{n}{\lambda_1} < \sum_{i=0}^{n} \binom{n}{k} = 2^n.$$

The second term in (23) is bounded by the inequality, $(n-\lambda_1)! \leqslant n^{n-\lambda_1}$. Combining the two observations, we have

$$\begin{aligned}\sum_{\alpha n\leqslant\lambda_1} d_\lambda^2 &\leqslant \sum_i \binom{n}{\alpha n + i}^2 (n-\alpha n-i)! \\ &\leqslant 4^n \sum_i (n-\alpha n-i)! \\ &\leqslant n 4^n n^{(1-\alpha)n} = \mathcal{O}(n^{(1-\alpha+\epsilon)n})\end{aligned}$$

for all $\epsilon > 0$, which proves part (i).

To prove part (ii), first note that the proof of Lemma 4.1 can easily be extended to yield the fact

$$\sum_{\lambda=(\lambda_1,\lambda_2,\dots)} d_\lambda^2 \leqslant \binom{n}{\lambda_1}^2 \binom{n-\lambda_1}{\lambda_2}^2 (n-\lambda_1-\lambda_2)!.$$

Therefore, similar to part (i),

$$\begin{aligned}\sum_{\substack{\alpha n\leqslant\lambda_1 \\ \frac{\lambda_1}{2}\leqslant\lambda_2\leqslant\lambda_1}} d_\lambda^2 &\leqslant \sum_{\alpha n\leqslant\lambda_1} \sum_{\frac{\lambda_1}{2}\leqslant\lambda_2} \binom{n}{\lambda_1}^2 \binom{n}{\lambda_2}^2 (n-\lambda_1-\lambda_2)! \\ &\leqslant \sum_i \sum_j \binom{n}{\alpha n + i}^2 \binom{n}{\frac{1}{2}\alpha n + j}^2 (n - \frac{3}{2}\alpha n - i - j)! \\ &\leqslant n^2 4^{2n} n^{(1-\frac{3}{2}\alpha)n} = \mathcal{O}(n^{(1-\frac{3}{2}\alpha+\epsilon)n}).\end{aligned}$$

□

Next, we define a partial order on the partitions of a fixed natural number n, known as the *dominance order*, which is to be used in the proof.

Definition 4.3 ([21]). *Let $\lambda = (\lambda_1, \lambda_2, ...)$ and $\mu = (\mu_1, \mu_2, ...)$ be partitions of n. We say that λ dominates μ, and denote by $\lambda \succeq \mu$, if*

$$\sum_{i=1}^{j} \lambda_i \geqslant \sum_{i=1}^{j} \mu_i \text{ for all } j \geqslant 1.$$

If $\lambda \succeq \mu$ and $\lambda \neq \mu$, then we write $\lambda \succ \mu$.

4.4. Upper bound. Let us take $\beta_\lambda = \beta_{\lambda,n}$. The idea is to show that the large eigenvalues have small multiplicities and vice versa. In summary, we first divide $\{\lambda : \lambda \vdash n\}$ into regions with respect to the length of the first two rows, and identify the partitions which dominate every other partition in those regions. Then the bounds on the dimensions given above are employed.

We start by defining two sets of partitions, of which the eigenvalues are computationally manageable. The first set consists of partitions where the cells in the Young diagram are stacked up and right as much as possible for a fixed ratio of the length of the first row to the total numbers of boxes. Formally, for a fixed $\alpha \in (0,1)$,

$$\lambda^\alpha_{(1)} := (\alpha n, \dots, \alpha n, rn) \tag{24}$$

where $\alpha q + r = 1$ for some $q \in \mathbb{N}$ and $0 \leqslant r < \alpha$. Although the definition makes sense only if αn and rn are integers, one can choose the closest integers as n tends to infinity and the proof below works properly.

The set of partitions defined above respects the dominance order (see Definition 4.3) in the sense that $\lambda^\eta_{(1)} \succeq \lambda^\zeta_{(1)}$ if and only if $\eta \geqslant \zeta$. In fact, more is true. It is easy to check that

$$\lambda^\alpha_{(1)} \succeq \lambda \quad \text{ if and only if } \quad \lambda_1 \leqslant \alpha n. \tag{25}$$

In the second set of partitions, the length of the second row is fixed to be the half of the first row's length if the length of the first row is less than $\frac{n}{2}$, otherwise it is half of the number of remaining cells. More precisely, the definition is as follows:

$$\lambda^\alpha_{(2)} := \begin{cases} \left(\alpha n, \frac{\alpha}{2}n \dots, \frac{\alpha}{2}n, \tilde{r}n\right) & \alpha \in \left(0, \frac{1}{2}\right] \\ \left(\alpha n, \frac{1-\alpha}{2}n, \frac{1-\alpha}{2}n\right) & \alpha \in \left(\frac{1}{2}, 1\right). \end{cases} \tag{26}$$

where $0 \leqslant \tilde{r} < \frac{\alpha}{2}$. Similarly,

$$\lambda^\alpha_{(2)} \succeq \lambda \quad \text{ if } \quad \lambda_1 \leqslant \alpha n,\ \lambda_2 \leqslant \min\left\{\frac{\alpha}{2}n, \frac{1-\alpha}{2}n\right\}. \tag{27}$$

$\{\lambda^\alpha_{(1)}\}_\alpha$ lies through the upper sides of the triangle in Figure 2, while $\{\lambda^\alpha_{(2)}\}_\alpha$ lies strictly below it through the inscribed triangle in the figure. Then the dominance order can be interpreted as $\lambda^\alpha_{(i)}$ dominates every partition to the right under the triangle it lies on.

Next, we estimate the eigenvalues corresponding to those partitions defined above.

Lemma 4.4. *Let $|\lambda| = n$ and $\alpha q + r = 1$ for some fixed α as in definition* (24)*. Then we have,*

$$\beta_{\lambda^\alpha_{(1)}} \leqslant C \frac{(1+\alpha)^{q(1+\alpha)n}(1+r)^{(1+r)n}}{2^{2n}},$$

where C is a constant depending on α.

Proof. We use the formula for the eigenvalues (8) to have

$$\begin{aligned}
\beta_{\lambda^{\alpha}_{(1)}} &= \prod_{i=1}^{q}\prod_{j=1}^{\alpha n}\frac{n-i+j}{n+(i-1)\alpha n+j-1}\times\prod_{j=1}^{rn}\frac{n-q+j}{2n-rn+j-1}\\
&\leqslant \prod_{i=1}^{q}\prod_{j=1}^{\alpha n}\frac{n+j}{n+(i-1)\alpha n+j}\times\prod_{j=1}^{rn}\frac{n+j}{2n-rn+j}\\
&\leqslant \prod_{i=1}^{q}\frac{[(1+\alpha)n]!}{n!}\frac{[(1+(i-1)\alpha)n]!}{[(1+i\alpha)n]!}\times\frac{[(1+r)n]!}{n!}\frac{[(1+q\alpha)n]!}{(2n)!}\\
&= \frac{([(1+\alpha)n]!)^{q}\,[(1+r)n]!}{(n!)^{q}\,(2n)!}
\end{aligned}$$

For n large enough, applying the formula (see [1, 6.1.38])

$$x! = \sqrt{2\pi}x^{x+\frac{1}{2}}e^{-x+\frac{\xi}{12x}}, \qquad \text{for some } 0<\xi<1, \tag{28}$$

to the factorials above, we obtain

$$\begin{aligned}
\beta_{\lambda^{\alpha}_{(1)}} &\leqslant e^{\frac{q}{12n}}\sqrt{\frac{(1+\alpha)^{q}(1+r)}{2}}\,\frac{(1+\alpha)^{q(1+\alpha)n}\,(1+r)^{(1+r)n}}{2^{2n}}\\
&\leqslant 2^{2/\alpha}\,\frac{(1+\alpha)^{q(1+\alpha)n}\,(1+r)^{(1+r)n}}{2^{2n}}.
\end{aligned}$$

□

In fact, for our purpose, only the cases where either $r=0$ or $r=\frac{\alpha}{2}$ are relevant. For the second set of partitions, we have the following lemma.

Lemma 4.5. *Let $|\lambda|=n$ and $\tilde{r}=0$ in the definition (26) for $\alpha\in\left(0,\frac{1}{2}\right]$. Then there exists $\tilde{q}\in\mathbb{N}$ such that $\alpha\tilde{q}=2$. The eigenvalues are bounded as*

$$\beta_{\lambda^{\alpha}_{(2)}} \leqslant C\left(1+\frac{\alpha}{2}\right)^{\alpha n/2}\beta_{\lambda^{\alpha/2}_{(1)}},$$

where C is a constant depending on α.

Proof. The eigenvalue formula (8) gives

$$\begin{aligned}
\beta_{\lambda^{\alpha}_{(2)}} &= \prod_{i=1}^{\tilde{q}-2}\prod_{j=1}^{\alpha n/2}\frac{n-i+j}{n+(i+1)\frac{\alpha n}{2}+j-1}\\
&= \prod_{j=1}^{\alpha n/2}\frac{n+\frac{\alpha n}{2}+j-1}{n-(\tilde{q}-1)+j}\prod_{i=1}^{\tilde{q}-1}\prod_{j=1}^{\alpha n/2}\frac{n-i+j}{n+i\frac{\alpha n}{2}+j-1}\\
&\leqslant \left(\frac{n+\frac{\alpha n}{2}-1}{n-(\tilde{q}-1)}\right)^{\alpha n/2}\beta_{\lambda^{\alpha/2}_{(1)}}\\
&= C\left(1+\frac{\alpha}{2}\right)^{\alpha n/2}\beta_{\lambda^{\alpha/2}_{(1)}}
\end{aligned}$$

for some constant C. □

Having established bounds on set of eigenvalues, we then bound the total variation distance for $\theta=n$ using the upper bound formula (9). We separate the range of the

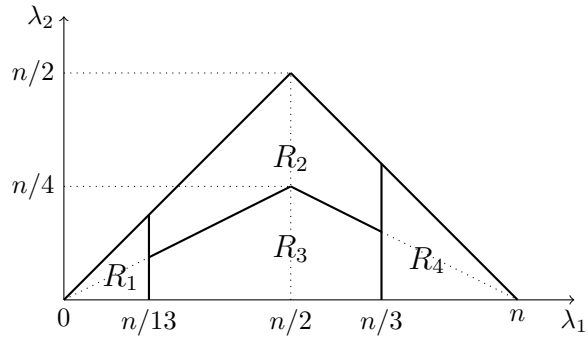


FIGURE 2. Planar representation of partitions with respect to the length of the first two rows of Young diagram

sum into four regions as depicted in Figure 2. In set notation, they are

$$R_1 = \left\{\lambda \vdash n : \lambda_1 \leqslant \frac{n}{13}\right\},$$
$$R_2 = \left\{\lambda \vdash n : \frac{n}{13} < \lambda_1 \leqslant \frac{n}{3},\ \min\left\{\frac{\lambda_1}{2}, \frac{1-\lambda_1}{2}\right\} \leqslant \lambda_2\right\},$$
$$R_3 = \left\{\lambda \vdash n : \frac{n}{13} < \lambda_1 \leqslant \frac{n}{3},\ \lambda_2 < \min\left\{\frac{\lambda_1}{2}, \frac{1-\lambda_1}{2}\right\}\right\},$$
$$R_4 = \left\{\lambda \vdash n : \frac{n}{3} < \lambda_1 < n\right\}.$$

We take $t^* = \frac{\log n}{\log 2}$ throughout the proof. Also note that

$$\beta_\lambda^{2t^*} = \beta_\lambda^{\frac{2}{\log 2}\log n} = n^{\frac{2}{\log 2}\log \beta_\lambda}. \tag{29}$$

4.4.1. *Region 1.* We show that the eigenvalues are small enough after a linear order of n, so that when multiplied with d_λ^2 the sum is of smaller order than a constant.

We take $r = 0$ in the statement of Lemma 4.4, therefore $\alpha = \frac{1}{q}$ for some $q \in \mathbb{N}$. By Lemma 4.4,

$$\begin{aligned}\log \beta_{\lambda^\alpha_{(1)}} &= \left(q\left(1+\frac{1}{q}\right)\log\left(1+\frac{1}{q}\right) - 2\log 2\right) n \\ &= \left((1+q)\left(\frac{1}{q} - \frac{1}{2q^2} + \frac{1}{3q^3} - \cdots\right) - 2\log 2\right) n \\ &\leqslant \left(1 + \frac{1}{2q} - 2\log 2\right) n \end{aligned} \tag{30}$$

Then we bound (29).

$$\begin{aligned}\log_n \left(\beta_{\lambda^\alpha_{(1)}}\right)^{2t^*} &= \frac{2}{\log 2}\log \beta_{\lambda^\alpha_{(1)}} \\ &= \left(\frac{2}{\log 2} + \frac{1}{q\log 2} - 4\right) n \\ &\approx \left(\frac{1.44}{q} - 1.11\right) n\end{aligned}$$

So if we choose q large enough, or equivalently α small enough, $\left(\beta_{\lambda^\alpha_{(1)}}\right)^{2t^*} \leqslant n^{-n}$. It turns out that the smallest integer q that satisfies the inequality is 13. We have,

$$\begin{aligned}\sum_{\lambda\in R_1} d_\lambda^2\beta_\lambda^{2t^*} &\leqslant \left(\beta_{\lambda_{(1)}^{1/13}}\right)^{2t^*} \sum_{\substack{\lambda\vdash n\\ \lambda_1\leqslant n/13}} d_\lambda^2\\ &\leqslant n^{-1.01n}\sum_{\lambda\vdash n} d_\lambda^2\\ &\leqslant n^{-1.01n} n!\\ &= o(n^{-0.01n}).\end{aligned}$$

4.4.2. *Region 2.* First we briefly justify the need for the restriction on the length of the second row of Young diagrams. Consider the partition $\lambda = (\frac{n}{2}, \frac{n}{2})$, and the corresponding eigenvalue

$$\beta_{(\frac{n}{2},\frac{n}{2})} = \prod_{i=1}^{n/2} \frac{n+i-1}{n+\frac{n}{2}-1} = \frac{\binom{\frac{3}{2}n-1}{\frac{n}{2}}}{\binom{2n-1}{\frac{n}{2}}} = C\frac{\left(\frac{3}{2}\right)^{3n/2}\left(\frac{3}{2}\right)^{3n/2}}{2^{2n}} = C\frac{3^{3n}}{2^{5n}}$$

for some constant C. Then,

$$\beta_{(\frac{n}{2},\frac{n}{2})}^{2t^*} = Cn^{\frac{2}{\log 2}n[3\log 3 - 5\log 2]} \approx Cn^{-0.49n}.$$

On the other hand, $\sum_{\lambda=(\frac{n}{2},\dots)} d_\lambda^2$ is of order at least $\frac{n}{2}!$. To see this, first fill out the first row by $1, 2, \dots, \frac{n}{2}$, then there are exactly $\frac{n}{2}!$ ways to obtain a standard Young diagram, which gives a lower bound. But since $n^{(1/2-\epsilon)n} = o(\frac{n}{2}!)$, $\sum_{\lambda=(\frac{n}{2},\dots)} d_\lambda^2\,\beta_{(\frac{n}{2},\frac{n}{2})}^{2t^*}$ is of order larger than a constant.

We restrict our attention to a smaller region, and use part (ii) of Lemma 4.2 to bound the dimensions. First, we choose a sequence of partitions $\{\lambda_{(1)}^{\alpha_i}\}_i$ in this region, for which $\alpha_0 = \frac{2}{3}, \alpha_1 = \frac{1}{2}, \alpha_2 = \frac{2}{5}$ and $\alpha_i = \frac{1}{i}$ for $i \geqslant 3$. By the dominance order relations (25), we have

$$\sum_{\lambda\in R_2} d_\lambda^2\,(\beta_\lambda)^{2t^*} \leqslant \sum_i \left(\beta_{\lambda_{(1)}^{\alpha_i}}\right)^{2t^*} \sum_{\substack{\lambda\in R_2\\ \alpha_{i+1}n\leqslant\lambda_1\leqslant\alpha_i n}} d_\lambda^2.$$

Then by part (ii) of Lemma 4.2,

$$\sum_{\lambda\in R_2} d_\lambda^2\,(\beta_\lambda)^{2t^*} \leqslant \sum_i \left(\beta_{\lambda_{(1)}^{\alpha_i}}\right)^{2t^*} n^{(1-\frac{3}{2}\alpha_{i+1})n}.$$

Next, we bound the eigenvalues. For $i = 0, 1$ and 2, we have the following calculations by Lemma 4.4:

$$\begin{aligned}\log\beta_{(1)}^{\alpha_0} &= \log\beta_{(\frac{2n}{3},\frac{n}{3})} \leqslant \left(\frac{5}{3}\log\frac{5}{3} + \frac{4}{3}\log\frac{4}{3} - 2\log 2\right)n \leqslant -0.15n,\\ \log\beta_{(1)}^{\alpha_1} &= \log\beta_{(\frac{n}{2},\frac{n}{2})} \leqslant \left(3\log\frac{3}{2} - 2\log 2\right)n \leqslant -0.17n,\\ \log\beta_{(1)}^{\alpha_2} &= \log\beta_{(\frac{2n}{5},\frac{2n}{5},\frac{n}{5})} \leqslant \left(\frac{14}{5}\log\frac{7}{5} + \frac{6}{5}\log\frac{6}{5} - 2\log 2\right)n \leqslant -0.22n.\end{aligned}$$

Further computations using (29) yield

$$\sum_{i=0}^{2} \left(\beta_{\lambda_{(1)}^{\alpha_i}}\right)^{2t^*} n^{(1-\frac{3}{2}\alpha_{i+1})n} = o(n^{-0.09n}). \tag{31}$$

For $i \geqslant 3$, since $\alpha_i = \frac{1}{i}$, we can use the bound (30) to obtain

$$\log \beta_{(1)}^{\alpha_i} \leqslant \left(1 + \frac{1}{2i} - 2\log 2\right) n.$$

Then by (29),

$$\begin{aligned}&\sum_{i\geqslant 3} \left(\beta_{\lambda_{(1)}^{\alpha_i}}\right)^{2t^*} n^{(1-\frac{3}{2}\alpha_{i+1})n} \\ =&\sum_{i\geqslant 3} \exp\left\{\left(\frac{2}{\log 2} - 4 + \frac{1}{i\log 2} + 1 - \frac{3}{2(i+1)}\right) n\log n\right\}.\end{aligned}$$

It is easy to see that the right-hand side is a decreasing function of i for $i \geqslant 3$. Therefore, if we plug in $i = 3$ and carry out the calculations,

$$\begin{aligned}\sum_{i\geqslant 3} \left(\beta_{\lambda_{(1)}^{\alpha_i}}\right)^{2t^*} n^{(1-\frac{3}{2}\alpha_{i+1})n} &\leqslant \sum_{i\geqslant 3} \exp\left\{\left(\frac{2}{\log 2} - 4 + \frac{1}{3\log 2} + 1 - \frac{3}{8}\right) n\log n\right\} \\ &= o(n^{-0.008n}).\end{aligned}$$

Combining with (31), we eventually have

$$\sum_{\lambda\in R_2} d_\lambda^2 \left(\beta_\lambda\right)^{2t^*} = o(n^{-0.008n}).$$

4.4.3. *Region 3.* This region is treated very similar to Region 2. We consider the second set of partitions (26) this time with the same choice of α_i's as in the previous case.

By the dominance order relations (27) mentioned above, we have

$$\sum_{\lambda\in R_3} d_\lambda^2 \left(\beta_\lambda\right)^{2t^*} \leqslant \sum_i \left(\beta_{\lambda_{(2)}^{\alpha_i}}\right)^{2t^*} \sum_{\substack{\lambda\in R_3 \\ \alpha_{i+1}n\leqslant\lambda_1\leqslant\alpha_i n}} d_\lambda^2.$$

By part (i) of Lemma 4.2,

$$\sum_{\lambda\in R_3} d_\lambda^2 \left(\beta_\lambda\right)^{2t^*} \leqslant \sum_i \left(\beta_{\lambda_{(2)}^{\alpha_i}}\right)^{2t^*} n^{(1-\alpha_{i+1})n}.$$

For $i = 0$, we can bound the eigenvalue by direct calculations,

$$\begin{aligned}\beta_{(2)}^{\alpha_0} = \beta_{\left(\frac{2n}{3},\frac{n}{6},\frac{n}{6}\right)} &= \prod_{i=1}^{2}\prod_{j=1}^{n/6} \frac{n-i+j}{\frac{5n}{3} + (i-1)\frac{n}{6} + j - 1} \\ &\leqslant C\frac{\left(\frac{7}{6}\right)^{14n/6}\left(\frac{5}{3}\right)^{5n/3}}{2^{2n}},\end{aligned}$$

which follows from the proof of Lemma 4.4. Similarly,

$$\begin{aligned}\beta_{(2)}^{\alpha_1} =\beta_{\left(\frac{n}{2},\frac{n}{4},\frac{n}{4}\right)} &\leqslant C\frac{\left(\frac{5}{4}\right)^{5n/2}\left(\frac{3}{2}\right)^{3n/2}}{2^{2n}}, \\ \beta_{(2)}^{\alpha_2} =\beta_{\left(\frac{2n}{5},\frac{n}{5},\frac{n}{5},\frac{n}{5}\right)} &\leqslant C\frac{\left(\frac{6}{5}\right)^{18n/5}\left(\frac{7}{5}\right)^{7n/5}}{2^{2n}}.\end{aligned}$$

Therefore,

$$\log \beta_{(2)}^{\alpha_0} \leqslant \left(\frac{14}{6}\log\frac{7}{6} + \frac{5}{3}\log\frac{5}{2} - 2\log 2\right) n \leqslant -0.175n,$$
$$\log \beta_{(2)}^{\alpha_1} \leqslant \left(\frac{5}{2}\log\frac{5}{4} + \frac{3}{2}\log\frac{3}{2} - 2\log 2\right) n \leqslant -0.22n,$$
$$\log \beta_{(2)}^{\alpha_2} \leqslant \left(\frac{18}{5}\log\frac{6}{5} + \frac{7}{5}\log\frac{7}{5} - 2\log 2\right) n \leqslant -0.25n.$$

Then we have

$$\sum_{i=0}^{2} \left(\beta_{\lambda_{(2)}^{\alpha_i}}\right)^{2t^*} n^{(1-\alpha_{i+1})n} = o(n^{-0.005n}). \tag{32}$$

For $i \geqslant 3$, we apply Lemma 4.4 and 4.5 to obtain

$$\begin{aligned}\log \beta_{(2)}^{\alpha_i} &\leqslant \left(\frac{1}{2i}\log\left(1+\frac{1}{2i}\right) + 2i\left(1+\frac{1}{2i}\right)\log\left(1+\frac{1}{2i}\right) - 2\log 2\right) n \\ &= \left(\left(2i+1+\frac{1}{2i}\right)\left(\frac{1}{2i} - \frac{1}{8i^2} + \frac{1}{24i^3} - \frac{1}{64i^4} + \cdots\right) - 2\log 2\right) n \\ &\leqslant \left(1 + \frac{1}{4i} + \frac{5}{24i^2} - 2\log 2\right) n.\end{aligned}$$

Then by (29),

$$\begin{aligned}&\sum_{i=3}^{12} \left(\beta_{\lambda_{(2)}^{\alpha_i}}\right)^{2t^*} n^{(1-\alpha_{i+1})n} \\ =&\sum_{i=3}^{12} \exp\left\{\left(\frac{2}{\log 2} - 4 + \frac{1}{2i\log 2} + \frac{5}{12i^2\log 2} + 1 - \frac{1}{(i+1)}\right) n\log n\right\}.\end{aligned}$$

Now we observe that

$$\frac{2}{\log 2} - 3 + \frac{5}{12i^2\log 2} \leqslant \frac{2}{\log 2} - 3 + \frac{5}{108\log 2} < 0,$$

for $i \geqslant 3$. One can also verify that $\frac{1}{2i\log 2} - \frac{1}{(i+1)}$ is an increasing function of i for $i \geqslant 3$. Therefore, we can plug in $i = 12$ to obtain an upper bound.

$$\sum_{i=3}^{12} \left(\beta_{\lambda_{(2)}^{\alpha_i}}\right)^{2t^*} n^{(1-\alpha_{i+1})n} \leqslant \sum_{i=3}^{12} \exp\left\{\left(\frac{1}{144\log 2} - \frac{1}{13}\right) n\log n\right\} = o(n^{-0.06n}).$$

Putting together with (32), we have

$$\sum_{\lambda\in R_3} d_\lambda^2 \left(\beta_\lambda\right)^{2t^*} = o(n^{-0.005n}).$$

4.4.4. *Region 4.* The cases covered above allows us to conclude that if $t \geqslant \frac{2}{\log 2}\log n$, Region 4 determines the convergence rate. Take $t = \frac{\log n}{\log 2} + \gamma$ for $\gamma > 0$.
Let $m = n - \lambda_1$. Then we have

$$\begin{aligned}\beta_{(\lambda_1, n-\lambda_1)} &= \prod_{j=1}^{m} \frac{n+j-1}{2n-m+j-1} \\ &\leqslant \left(\frac{n+m}{2n}\right)^m\end{aligned}$$

Using the fact that $\log(1+x) < x$,

$$\log \beta_{(n-m,m)} \leqslant m\left[-\log 2 + \log\left(1+\frac{m}{n}\right)\right] \leqslant -m\log 2 + \frac{m^2}{n}, \tag{33}$$

which implies

$$\begin{aligned}\beta_{(n-m,m)}^{2t} = \beta_{(n-m,m)}^{\frac{2\log n}{\log 2}}\,\beta_{(n-m,m)}^{2\gamma} &\leqslant n^{\frac{2m^2}{n\log 2}-2m}\,\beta_{(n-1,1)}^{2\gamma} \\ &\leqslant n^{\frac{2m^2}{n\log 2}-2m}\left(\frac{1}{2}\right)^{2\gamma}.\end{aligned}$$

Therefore, by Lemma 4.1 and (33),

$$\begin{aligned}\sum_{\lambda\in R_4} d_\lambda^2\,\beta_\lambda^{2t} &= \sum_{2n/3\leqslant\lambda_1\leqslant n-1} d_\lambda^2\,\beta_\lambda^{2t} \\ &\leqslant \left(\frac{1}{2}\right)^{2\gamma}\sum_{m=1}^{n/3}\binom{n}{m}^2 m!\,n^{\frac{2m^2}{n\log 2}-2m}.\end{aligned}$$

To bound the factorial term, we employ the fact $\binom{n}{m} \leqslant \frac{n^m}{m!}$,

$$\begin{aligned}\sum_{\lambda\in R_4} d_\lambda^2\,\beta_\lambda^{2t} &\leqslant \frac{1}{4^\gamma}\sum_{m=1}^{n/3}\frac{n^{2m}}{m!}n^{\frac{2m^2}{n\log 2}-2m} \\ &= \frac{1}{4^\gamma}\sum_{m=1}^{n/3}\frac{n^{\frac{2m^2}{n\log 2}}}{m!}.\end{aligned}$$

An application of Stirling's formula gives

$$\sum_{\lambda\in R_4} d_\lambda^2\,\beta_\lambda^{2t} \leqslant \frac{1}{4^\gamma}\left(1+\frac{1}{2}+\frac{e}{\sqrt{2\pi}}\sum_{m=3}^{n/3}\frac{1}{m^{3/2}}\frac{e^{m-1}n^{\frac{2m^2}{n\log 2}}}{m^{m-1}}+\mathcal{O}(n^{1/n})\right).$$

Let

$$f(m) = \log\left(\frac{e^{m-1}n^{\frac{2\,m^2}{n\log 2}}}{m^{m-1}}\right) = (m-1)+\frac{2m^2\log n}{n\log 2}-(m-1)\log m.$$

Next we show that $f(m)$ is less than 0 in the range of the sum. Taking derivatives, we have

$$f''(m) = \frac{4\log n}{n\log 2}-\frac{1}{m}-\frac{1}{m^2} > 0$$

for $m \geqslant 3$. Therefore f is a convex function, so

$$f(m) \leqslant \max\{f(3), f(n/3)\}$$

for all $m \in [3, n/3]$. It is easy to check that both $f(3)$ and $f(n/3)$ are less than zero, hence f is strictly smaller than zero in the range of the sum. As a result of this,

$$\begin{aligned}\sum_{\lambda\in R_4} d_\lambda^2\,\beta_\lambda^{2t} &\leqslant \frac{1}{4^\gamma}\left(1+\frac{1}{2}+\frac{e}{\sqrt{2\pi}}\sum_{m=3}^{n/3}\frac{1}{m^{3/2}}+\mathcal{O}(n^{1/n})\right) \\ &\leqslant \frac{1}{4^\gamma}\left(1+\frac{1}{2}+\frac{e}{\sqrt{2\pi}}(\zeta(3/2)-1)\right) \\ &\leqslant \frac{1}{4^{\gamma-1}}.\end{aligned}$$

Finally, by (9) we have

$$d_n\left(\frac{\log n}{\log 2}+\gamma\right)=\left\|P_\theta^{\frac{\log n}{\log 2}+\gamma}-\pi\right\|_{TV}\leqslant\frac{1}{4^\gamma}.$$

Combining with (22), we conclude that the chain has a total variation cutoff. This concludes the proof of Theorem 1.1 part (ii).

5. Conclusion

The Plancherel growth process does not seem to be fully exploited in this context. Further work could require estimates on the moments of contents with respect to the transition measure. Some explicit formulae for the moments are provided in [19]. The computations in Section 4.2 suggests the following:

Conjecture 5.1. *Let P_θ be the Markov chain on S_n defined in* (1).

(i) *If $\theta(n)=\mathcal{O}(n^\beta)$ for some $\beta<1$, then the mixing time is bounded above by a constant independent of n.*

(ii) *If $n=\mathcal{O}(\theta(n))$, then the chain has a total variation cutoff at $t_n=\frac{\log n}{\log(n+\theta(n))-\log\theta(n)}$ with a window of size $\frac{1}{\log(n+\theta(n))-\log\theta(n)}$.*

To compare it to a well-known example, we expect that P_θ has the same rate of convergence with the random transposition walk on the symmetric group [11] provided that $\lim_{n\to\infty}\frac{\theta(n)}{n^2/2}=1$.

References

[1] Milton Abramowitz and Irene A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, U.S. Government Printing Office, Washington, D.C., 1964.

[2] Martin Aigner, *A course in enumeration*, Graduate Texts in Mathematics, vol. 238, Springer, Berlin, 2007.

[3] Richard Arratia, A. D. Barbour, and Simon Tavaré, *Poisson process approximations for the Ewens sampling formula*, Ann. Appl. Probab. **2** (1992), no. 3, 519–535.

[4] Ehrhard Behrends, *Introduction to Markov chains (with special emphasis on rapid mixing)*, Advanced Lectures in Mathematics, Friedr. Vieweg & Sohn, Braunschweig, 2000.

[5] Tullio Ceccherini-Silberstein, Fabio Scarabotti, and Filippo Tolli, *Representation theory of the symmetric groups (the Okounkov-Vershik approach, character formulas, and partition algebras)*, Cambridge Studies in Advanced Mathematics, vol. 121, Cambridge University Press, Cambridge, 2010.

[6] Harry Crane, *The ubiquitous Ewens sampling formula*, Statist. Sci. **31** (2016), no. 1, 1–19.

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

[8] ———, *The cutoff phenomenon in finite Markov chains*, Proc. Nat. Acad. Sci. U.S.A. **93** (1996), no. 4, 1659–1664.

[9] Persi Diaconis and Curtis Greene, *Applications of Murphy's elements*, Tech. Report 335, Stanford University Technical Reports, 1989.

[10] Persi Diaconis and Phil Hanlon, *Eigen-analysis for some examples of the Metropolis algorithm*, in Hypergeometric functions on domains of positivity, Jack polynomials, and applications (Tampa, FL, 1991), Contemp. Math., vol. 138, Amer. Math. Soc., Providence, RI, 1992, pp. 99–117.

[11] Persi Diaconis and Mehrdad Shahshahani, *Generating a random permutation with random transpositions*, Z. Wahrsch. Verw. Gebiete (1981), no. 2, 159–179.

[12] Warren J. Ewens, *The sampling theory of selectively neutral alleles*, Theor. Popul. Biol. **3** (1972), no. 1, 87–112.

[13] Adriano M. Garsia and Ömer Eğecioğlu, *Lectures in algebraic combinatorics (Young's construction, seminormal representations, $\mathfrak{sl}(2)$ representations, heaps, basics on finite fields)*, Lecture Notes in Mathematics, vol. 2277, Springer, Cham, 2020.

[14] Phil Hanlon, *A Markov chain on the symmetric group and Jack symmetric functions*, Discrete Math. **99** (1992), no. 1-3, 123–140.

[15] Yunjiang Jiang, *Mixing time of Metropolis chain based on random transposition walk converging to multivariate Ewens distribution*, Ann. Appl. Probab. **25** (2015), no. 3, 1581–1615.
[16] A.-A. A. Jucys, *Symmetric polynomials and the center of the symmetric group ring*, Rep. Mathematical Phys. **5** (1974), no. 1, 107–112.
[17] S. V. Kerov, *Transition probabilities of continual Young diagrams and the Markov moment problem*, Funktsional. Anal. i Prilozhen. **27** (1993), no. 2, 32–49, 96.
[18] ———, *The boundary of Young lattice and random Young tableaux*, in Formal power series and algebraic combinatorics (New Brunswick, NJ, 1994), DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 24, Amer. Math. Soc., Providence, RI, 1996, pp. 133–158.
[19] Michel Lassalle, *Jack polynomials and some identities for partitions*, Trans. Amer. Math. Soc. **356** (2004), no. 9, 3455–3476.
[20] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009, With a chapter by James G. Propp and David B. Wilson.
[21] I. G. Macdonald, *Symmetric functions and Hall polynomials*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1979.
[22] G. E. Murphy, *A new construction of Young's seminormal representation of the symmetric groups*, J. Algebra **69** (1981), no. 2, 287–297.
[23] Bruce E. Sagan, *The symmetric group (representations, combinatorial algorithms, and symmetric functions)*, second ed., Graduate Texts in Mathematics, vol. 203, Springer-Verlag, New York, 2001.
[24] Laurent Saloff-Coste, *Random walks on finite groups*, in Probability on discrete structures, Encyclopaedia Math. Sci., vol. 110, Springer, Berlin, 2004, pp. 263–346.
[25] Lajos Takács, *The problem of coincidences*, Arch. Hist. Exact Sci. (1979/80), no. 3, 229–244.
[26] A. M. Vershik and A. Yu. Okun′kov, *A new approach to representation theory of symmetric groups. II*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **307** (2004), 57–98, 281.

Alperen Özdemir, Georgia Institute of Technology, School of Mathematics, 686 Cherry St NW, Atlanta, GA 30332 (USA)
E-mail : aozdemir6@gatech.edu