Lithuanian Mathematical Journal, Vol.X, No. x, 20xx, ***Author's Version*** *March 9, 2021*

A sharp inequality for the variance with respect to the Ewens Sampling Formula

Ž. Baronėnas, E. Manstavičius and P. Šapokaitė

(e-mail: eugenijus.manstavicius@mif.vu.lt)
Vilnius University, Institute of Mathematics, Naugarduko 24, LT-03225

Dedicated to Memory of Professor Jonas Kubilius

Abstract. The variance of a linear statistic defined on the symmetric group endowed with the Ewens probability is examined. Despite the dependence of the summands, it can be bounded from above by a constant multiple of the sum of variances of the summands. We find the exact value of this constant. The analysis of the appearing quadratic forms and eigenvalue search is built upon the exponential matrices and discrete Hahn's polynomials.

MSC: 60C05, 05A16, 20P05
Keywords: Random permutation, quadratic form, eigenvalue, discrete Hahn's polynomials, Turán-Kubilius inequality.

1 Introduction and results

The variance of a linear statistic defined on the symmetric group endowed with the Ewens probability is examined in the paper. The main obstacle to overcome in this seemingly simple problem is the dependence of the summands. We propose an approach built upon exponential matrices and special functions.

Let $\mathbb{S}_n$ denote the symmetric group of permutations σ acting on $n \in \mathbb{N}$ letters. Each $\sigma \in \mathbb{S}_n$ has a unique representation (up to the order) by the product of independentcycles κ_i:

$$\sigma = \kappa_1 \cdots \kappa_w \tag{1.1}$$

where $w = w(\sigma)$ denotes the number of cycles. Denote by $k_j(\sigma) \geq 0$ the number of cycles in (1.1) of length j for $1 \leq j \leq n$ and introduce the *cycle vector* $\bar{k}(\sigma) = (k_1(\sigma), \ldots, k_n(\sigma))$.

As usual, set $(x)_m = x(x+1)\cdots(x+m-1)$, where $m \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, for the increasing factorial. Denote also

$$\Theta(m) = \binom{\theta+m-1}{m} = \frac{(\theta)_m}{m!} = [x^m](1-x)^{-\theta}, \tag{1.2}$$

where $[x^m]f(x)$ stands for the mth coefficient of a power series $f(x)$ and $\theta > 0$ is a parameter. The *Ewens Probability Measure* $\nu_{n,\theta}$ on $\mathbb{S}_n$ is defined by

$$\nu_{n,\theta}\big(\{\sigma\}\big) = \theta^{w(\sigma)}/(\theta)_n, \quad \sigma \in \mathbb{S}_n.$$

This gives the probability space $(\mathbb{S}_n, 2^{\mathbb{S}_n}, \nu_{n,\theta})$ and every mapping $f : \mathbb{S}_n \to \mathbb{R}$ becomes a random variable (r.v.) defined on it. Throughout the paper, we write it as $f(\sigma)$ leaving the elementary event σ, in contrast to r.vs defined on some unspecified probability space $\{\Omega, \mathcal{F}, \mathbf{P}\}$.

Set $\ell(\bar{s}) = 1s_1 + \cdots + ns_n$ for a vector $\bar{s} = (s_1, \ldots, s_n) \in {\mathbb{N}_0}^n$. The equality $\ell(\bar{k}(\sigma)) = n$, valid for each $\sigma \in \mathbb{S}_n$, shows the dependence of the r.vs $k_j(\sigma)$ with respect to $\nu_{n,\theta}$. It is well known (see, for example, [1, Sect. 2.3]) that the distribution of $\bar{k}(\sigma)$ can be written as the conditional distribution of $\bar{\xi} = (\xi_1, \ldots, \xi_n)$, where ξ_j, $1 \le j \le n$, are mutually independent Poisson r.vs with parameter $\mathbf{E}\xi_j = \theta/j$. Indeed, direct calculation shows that

$$\nu_{n,\theta}\big(\bar{k}(\sigma) = \bar{s}\big) = \mathbf{1}\big\{\ell(\bar{s}) = n\big\}\Theta(n)^{-1} \prod_{j=1}^{n} \Big(\frac{\theta}{j}\Big)^{s_j} \frac{1}{s_j!} = \mathbf{P}\big(\bar{\xi} = \bar{s} \,|\, \ell(\bar{\xi}) = n\big). \tag{1.3}$$

Here $\mathbf{1}\{\cdot\}$ stands for the indicator function. The probability in (1.3), ascribed to the vector $\bar{s} \in \mathbb{N}_0^n$, is called the *Ewens Sampling Formula*. It has been introduced by W. J. Ewens [5] to model the mutation of genes. For a comprehensive account of the recent applications of this ubiquitous distribution in combinatorics and statistics, see [1], [7], [6], or survey [4] and the subsequent comments on it.

We prefer to stay within the theory of random permutations. Apart from $w(\sigma)$, other linear statistics (or *completely additive functions*)

$$h(\sigma) :\,= a_1 k_1(\sigma) + \cdots + a_n k_n(\sigma), \tag{1.4}$$

where $\bar{a} := (a_1, \ldots, a_n) \in \mathbb{R}^n$ is a non-zero vector, continue to raise an interest. For example, $h(\sigma)$ with $a_j = \log j$, $j \le n$, is a good approximation for the logarithm of the group-theoretical order of $\sigma \in \mathbb{S}_n$ (see [1] or [22]). The case with $a_j = \{xj\}$, where $\{u\}$ stands for the fractional part of $u \in \mathbb{R}$, is met in the theory of random permutation matrices (see [21]).

For an arbitrary $h(\sigma)$, the problem of finding necessary and sufficient conditions, assuring the weak convergence of distributions

$$\nu_{n,\theta}\big(h(\sigma) - \alpha(n) \le x\beta(n)\big), \tag{1.5}$$

where $\alpha(n) \in \mathbb{R}$ and $\beta(n) \to \infty$ as $n \to \infty$, is still open (see [1, Sect. 8.5] or [13] and the references therein). Obstacles in the necessity part arise because of the dependence of the summands as shown by (1.3). This also happens in the analysis of power moments carried out by the second author [11] and [12] even in the case $\theta = 1$. Let us now focus on the variance.

By $\mathbf{E}_{n,\theta} f(\sigma)$ and $\mathbf{V}ar_{n,\theta} f(\sigma)$ we denote the mean value and the variance of a r.v. $f(\sigma)$ defined on $\mathbb{S}_n$ with respect to $\nu_{n,\theta}$. For the particular function $h(\sigma)$ in (1.4), we also set $A_{n,\theta}(\bar{a}) = \mathbf{E}_{n,\theta} h(\sigma)$ and $D_{n,\theta}(\bar{a}) = \mathbf{V}ar_{n,\theta} h(\sigma)$. Applying Watterson's [20] formulas (see [1, (5.6), p. 96]) for the factorial moments of $k_j(\sigma)$, one easily finds (see [15] for the details) the expressions

$$A_{n,\theta}(\bar{a}) = \theta \sum_{j \le n} \frac{a_j}{j} \frac{\Theta(n-j)}{\Theta(n)}$$

and

$$\begin{aligned} D_n(\bar{a}) &= \theta \sum_{j \le n} \frac{a_j^2}{j} \frac{\Theta(n-j)}{\Theta(n)} + \theta^2 \sum_{i+j \le n} \frac{a_i a_j}{ij} \frac{\Theta(n-i-j)}{\Theta(n)} \\ &\quad - \theta^2 \bigg(\sum_{j \le n} \frac{a_j}{j} \frac{\Theta(n-j)}{\Theta(n)} \bigg)^2 \\ &=: \theta B_n(\bar{a}) + \theta^2 \Delta_n(\bar{a}), \end{aligned} \tag{1.6}$$

if $n \ge 2$ and

$$B_n(\bar{a}) = \sum_{j \le n} \frac{a_j^2}{j} \frac{\Theta(n-j)}{\Theta(n)}.$$

The latter quantity is close to the sum of variances of the summands in the definition of $h(\sigma)$. In fact, formula (4) from [15] shows that

$$\sum_{j\le n} a_j^2 \mathbf{V}ar_{n,\theta} k_j(\sigma) - \theta B_n(\bar a) = O\Big(n^{-(1\wedge\theta)} B_n(\bar a)\Big),$$

where $a\wedge b = \min\{a,b\}$ if $a,b\in\mathbb{R}$, with an absolute constant in the symbol $O(\cdot)$. We also have (see [15])

$$D_n(\bar a) \le C\theta B_n(\bar a) \tag{1.7}$$

uniformly in $n\ge 2$ with an absolute constant C which can be specified. If $\theta\ge 1$, one can take $C=2$. The purpose of the present paper is to find the exact value of C in (1.7).

Theorem 1. *Let $\theta>0$ be arbitrary and $n\ge 2$. Then*

$$\tau_n(\theta) := \sup\left\{ \frac{D_n(\bar a)}{\theta B_n(\bar a)} : \quad \bar a\in\mathbb{R}^n\setminus\{0\}\right\} = \frac{\theta+2}{\theta+1}.$$

The supremum is achieved taking $a_j=(\theta+2)j^2-(2n+\theta)j$ where $1\le j\le n$.

The pioneering results obtained in [11] and [16] showed that $\tau_n(1)=3/2+O(n^{-1})$ and $\tau_n(2)=4/3+O(n^{-1})$. The approach originated in Kubilius' paper [10] was based upon the extremal properties of the Jacobi polynomials. It unavoidably added a vanishing error term to the result. Recently J. Klimavičius and the second author [8] established that $\tau_n(1)=3/2$ for all $n\ge 2$. Theorem 1 resumes the research for an arbitrary $\theta>0$. It is directly related to the above mentioned problem concerning distributions (1.5). Applying Theorem 1, we obtain that the weak convergence of (1.5) with $\beta(n)=\sqrt{\theta B_n(a)}$, which is natural to use, can take place only to the limit laws having variance not exceeding $(\theta+2)/(\theta+1)$.

By virtue of (1.3), the result can be reformulated for the conditional variance of the linear statistics

$$Y_n := a_1\xi_1+\cdots+a_n\xi_n.$$

We obtain the following optimal inequality.

Corollary. Let $n\ge 2$ and $a_j\in\mathbf{R}$, $1\le j\le n$, be arbitrary. Then

$$\mathbf{V}ar\big(Y_n \,\big|\, \ell(\bar\xi)=n\big) \le \frac{\theta(\theta+2)}{\theta+1}\sum_{j=1}^n \frac{a_j^2}{j}\,\frac{\Theta(n-j)}{\Theta(n)}.$$

The problem concerns the quadratic forms $\Delta_n(\bar a)$ and $B_n(\bar a)$. The substitution

$$a_j = \Big(\frac{j\Theta(n)}{\Theta(n-j)}\Big)^{1/2} x_j, \quad 1\le j\le n,$$

reduces $B_n(\bar a)$ to the square of Euclidean norm $||\bar x||^2$ of the vector $\bar x=(x_1,\dots,x_n)\in\mathbb{R}^n$. Then $\Delta_n(\bar a)$ becomes a quadratic form, denoted afterwards by $\mathcal{M}_n(\bar x) := \bar x M_n \bar x'$, where $\bar x'$ is the transpose of $\bar x$ and $M_n=((m_{ij}))$, $1\le i,j\le n$, is the matrix with entries

$$m_{ij} = \frac{\Theta(n-i-j)}{\big(ij\Theta(n-i)\Theta(n-j)\big)^{1/2}} - \Big(\frac{\Theta(n-i)}{i\Theta(n)}\Big)^{1/2}\Big(\frac{\Theta(n-j)}{j\Theta(n)}\Big)^{1/2}. \tag{1.8}$$

Here we assume that $\Theta(-k) = 0$ if $k \in \mathbb{N}$. Now, by virtue of (1.6),

$$\begin{aligned}\tau_{n,\theta} &= 1 + \theta \sup_{\bar{x} \neq \bar{0}} \Big(||\bar{x}||^{-2} \mathcal{M}_n(\bar{x}) \Big) = 1 + \theta \sup_{\bar{x} \neq \bar{0}} \Big(||\bar{x}||^{-2} \sum_{r=1}^{n} \mu_r x_r^2 \Big) \\ &= 1 + \theta \max_{1 \le r \le n} \mu_r, \end{aligned} \tag{1.9}$$

where $\{\mu_1, \dots, \mu_n\}$ is the spectrum of matrix M_n. So, Theorem 1 follows from the following proposition.

Theorem 2. *The spectrum of the matrix M_n comprises*

$$\mu_r = \frac{(-1)^r (r-1)!}{(\theta)_r}, \quad 1 \le r \le n.$$

For the eigenvector corresponding to the maximal μ_2, one may take the vector with coordinates

$$\Big((\theta+2)j - (2n+\theta) \Big) \Big(j \Theta(n-j) \Big)^{1/2}, \quad 1 \le j \le n.$$

The proof of Theorem 2 presented in the next section is built upon exponential matrices.

The problem of finding the remaining eigenvectors of matrix M_n also raises an interest. Actually, they were already used in the continuation of present paper [14] dealing with the lower estimates of $D_n(\bar{a})$. Solving is based upon particular cases of the generalized hypergeometric series which are exposed, for example, in [3] or [9, Chapter 9]. The hint to exploit them stems from [8]. We may confine ourselves to the case of polynomials which, in the traditional notation, can be written as

$${}_{p+2}F_q\big(-m, -x, (a_p); (b_q); z \big) = \sum_{k=0}^{m} \frac{(-m)_k (-x)_k (a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k k!} z^k,$$

where $p, q, m \in \mathbb{N}_0$, $a_1, \dots, a_p; b_1, \dots, b_q \in \mathbb{R}$ are parameters. Moreover, it suffices to reckon the discrete Hahn's polynomials

$$Q_r(x; \alpha, \beta, n) = {}_3F_2\big(-r, -x, r+\alpha+\beta+1; \alpha+1, -n+1; 1 \big)$$

by specifying the parameters to $\alpha = 1$ and $\beta = \theta - 1$. This agrees with the notation in [9, Section 9.5] with $N = n - 1$. The values of polynomials

$$q_r(x) = Q_r(x-1; 1, \theta-1, n), \quad 0 \le r \le n-1,$$

at $x = j \in \{1, \dots, n\}$ can be calculated using the recurrence formula [9, (9.5.3)]. Namely, $q_0(j) = 1$,

$$q_1(j) = -\frac{(\theta+2)j - (\theta+2n)}{2(n-1)}$$

and

$$A_r q_{r+1}(j) = (A_r + C_r - j + 1) q_r(j) - C_r q_{r-1}(j),$$

where

$$A_r = \frac{(r+\theta+1)(r+2)(n-r-1)}{(2r+\theta+1)(2r+\theta+2)}, \qquad C_r = \frac{r(r+\theta+n)(r+\theta-1)}{(2r+\theta)(2r+\theta+1)}$$

and $r = 1, \ldots, n-1$. Moreover, we have the following orthogonality property [9, (9.5.2)]:

$$< q_l, q_r >:= \sum_{j=1}^{n} j q_l(j) q_r(j) \Theta(n-j) = \delta_{lr} \pi_r^2, \tag{1.10}$$

where δ_{lr} is the Kronecker symbol and

$$\pi_r^2 = \frac{(r+\theta+1)_n (\theta)_r}{(2r+\theta+1)(r+1)(n-r)_r (n-1)!} > 0.$$

Note that, $q_r(x)$, $1 \le r \le n-1$, can be also obtained uniquely by the Gram–Schmidt orthogonalization procedure starting with $q_0(x) = 1$ and applying the inner product (1.10). They form an orthogonal basis in the vector space of polynomials whose degrees do not exceed $n-1$. Exploiting this, in the isomorphic Euclidean space $\mathbb{R}^n$, we easily find the needed canonical basis for the matrix M_n.

Theorem 3. *The system of the vectors*

$$\bar{e}_r = (e_{r1}, \ldots, e_{rn}), \quad 1 \le r \le n,$$

where

$$e_{rj} = \pi_{r-1}^{-1} q_{r-1}(j) \sqrt{j \Theta(n-j)}, \quad 1 \le j \le n,$$

is an orthonormal basis in $\mathbb{R}^n$. Moreover, the vector $\bar{e}_r$ is the eigenvector of matrix M_n corresponding to μ_r for each $1 \le r \le n$.

The proof will be presented in the last section of the paper.

Finally, the distributions of mappings defined on random permutations taken according to the Ewens probability are close to that defined on logarithmic decomposable combinatorial structures (see [1]); therefore, we hope that our method is applicable when estimating the variances of similar statistics defined in such classes.

2 Proof of Theorem 2

The idea is to find a matrix L_n such that the product

$$\mathrm{e}^{L_n} M_n \mathrm{e}^{-L_n} =: \big((w_{ij}) \big)$$

is the triangle matrix with $w_{ij} = 0$ if $1 \le j < i \le n$ and $w_{jj} = \mu_j$ if $1 \le j \le n$. This implies that the eigenvalues of M_n are listed on the main diagonal of the product, as desired.

At first, we recall two identities from [17].

Lemma 1. *Let $M, m \in \mathbb{N}_0$ and $a, b \in \mathbb{R}$. Then*

$$\sum_{k=0}^{M} \binom{a+k}{k} \binom{b-k}{M-k} = \sum_{k=0}^{M} \binom{a+b-k}{M-k} \tag{2.1}$$

and

$$\sum_{k=0}^{M} (-1)^k \binom{M}{k} \binom{a-k}{m} = \binom{a-M}{m-M}. \tag{2.2}$$

*Proof S*ee formulas (43) on page 618 and (56) on page 619 of [17].

Proof of Theorem 2. Let us introduce the matrix $L_n(\theta) = \big((l_{ij})\big)$ with the entries $l_{ij} = 0$ for all $1 \le i, j \le n$ but for $i = j+1$, where

$$l_{j+1,j} = -\bigg(\frac{(j+1)j\Theta(n-j-1)}{\Theta(n-j)}\bigg)^{1/2}, \quad 1 \le j \le n-1.$$

Consider the powers $L_n^k(\theta) =: \big((l_{ij}^{(k)})\big)$, $0 \le k \le n-1$. The nonzero entries of $L_n^k(\theta)$ fill up the kth, $1 \le k \le n-1$, diagonal under the main one. By induction, we observe that

$$\begin{aligned} l_{j+k,j}^{(k)} &= l_{j+k,j+1}^{(k-1)} l_{j+1,j} \\ &= l_{j+k,j+k-1} l_{j+k-1,j+k-2} \cdots l_{j+1,j} \\ &= (-1)^k \prod_{r=0}^{k-1} \big((j+r+1)(j+r)\big)^{1/2} \bigg(\prod_{r=0}^{k-1} \frac{\Theta(n-j-r-1)}{\Theta(n-j-r)}\bigg)^{1/2} \\ &= (-1)^k (j)_k \Big(\frac{j+k}{j}\Big)^{1/2} \bigg(\frac{\Theta(n-j-k)}{\Theta(n-j)}\bigg)^{1/2} \end{aligned}$$

if $1 \le k \le n-j$. Hence the matrix $V := \mathrm{e}^{L_n(\theta)} =: \big((v_{ij})\big)$ has $v_{ij} = 0$ if $1 \le i < j \le n$ and

$$\begin{aligned} v_{ij} = \frac{l_{ij}^{(i-j)}}{(i-j)!} &= (-1)^{i-j} \binom{i-1}{j-1} \Big(\frac{i}{j}\Big)^{1/2} \bigg(\frac{\Theta(n-i)}{\Theta(n-j)}\bigg)^{1/2} \\ &= (-1)^{i-j} \binom{i}{j} \Big(\frac{j}{i}\Big)^{1/2} \bigg(\frac{\Theta(n-i)}{\Theta(n-j)}\bigg)^{1/2} \end{aligned}$$

if $i \ge j$. Moreover, $V^{-1} = \mathrm{e}^{-L_n(\theta)} = \big((|v_{ij}|)\big)$ if $1 \le i, j \le n$.

More technical obstacles arise calculating

$$\begin{aligned} w_{ij} &= \sum_{\substack{1 \le r \le i \\ j \le s \le n}} v_{ir} m_{rs} |v_{sj}| \\ &= \sum_{\substack{1 \le r \le i \\ j \le s \le n}} (-1)^{i-r} \binom{i}{r} \Big(\frac{r}{i}\Big)^{1/2} \bigg(\frac{\Theta(n-i)}{\Theta(n-r)}\bigg)^{1/2} \cdot \frac{\Theta(n-r-s)}{\big(rs\Theta(n-r)\Theta(n-s)\big)^{1/2}} \\ &\qquad\qquad \times \binom{s-1}{j-1} \Big(\frac{s}{j}\Big)^{1/2} \bigg(\frac{\Theta(n-s)}{\Theta(n-j)}\bigg)^{1/2} \\ &\quad - \sum_{\substack{1 \le r \le i \\ j \le s \le n}} (-1)^{i-r} \binom{i}{r} \Big(\frac{r}{i}\Big)^{1/2} \bigg(\frac{\Theta(n-i)}{\Theta(n-r)}\bigg)^{1/2} \cdot \bigg(\frac{\Theta(n-r)}{r\Theta(n)}\bigg)^{1/2} \bigg(\frac{\Theta(n-s)}{s\Theta(n)}\bigg)^{1/2} \\ &\qquad\qquad \times \binom{s-1}{j-1} \Big(\frac{s}{j}\Big)^{1/2} \bigg(\frac{\Theta(n-s)}{\Theta(n-j)}\bigg)^{1/2} \\ &=: \Sigma_1 - \Sigma_2. \end{aligned}$$

Here

$$\varSigma_1 = (-1)^i \Big(\frac{\Theta(n-i)}{ij\Theta(n-j)}\Big)^{1/2} \sum_{1\le r\le i\wedge n-j} \frac{(-1)^r}{\Theta(n-r)}\binom{i}{r} \sum_{j\le s\le n-r} \Theta(n-r-s)\binom{s-1}{j-1}.$$

After the change $s = n-r-k$ and application of (1.2), the inner sum reduces to that given in (2.1). It equals

$$\begin{aligned}\sum_{k=0}^{n-r-j} \binom{\theta-1+k}{k}\binom{n-r-1-k}{n-r-j-k} &= \sum_{k=0}^{n-r-j} \binom{\theta+n-r-2-k}{n-r-j-k} \\ &= \sum_{l=0}^{n-r-j} \binom{\theta-2+j+l}{l} = [x^{n-r-j}]\frac{1}{(1-x)^{\theta+j}} = \binom{\theta+n-r-1}{n-r-j}.\end{aligned}$$

Hence

$$\varSigma_1 = (-1)^i \Big(\frac{\Theta(n-i)}{ij\Theta(n-j)}\Big)^{1/2} \sum_{1\le r\le i\wedge n-j} \frac{(-1)^r}{\Theta(n-r)}\binom{i}{r}\binom{\theta+n-r-1}{n-r-j}.$$

Similarly,

$$\varSigma_2 = \frac{(-1)^i}{\Theta(n)}\Big(\frac{\Theta(n-i)}{ij\Theta(n-j)}\Big)^{1/2} \sum_{1\le r\le i} (-1)^r\binom{i}{r} \sum_{j\le s\le n} \Theta(n-s)\binom{s-1}{j-1}.$$

Since

$$\sum_{j\le s\le n} \binom{s-1}{j-1}\Theta(n-s) = [x^n]\Big(\frac{x^j}{(1-x)^j}\cdot\frac{1}{(1-x)^\theta}\Big) = \binom{\theta+n-1}{n-j},$$

we obtain

$$\varSigma_2 = \frac{(-1)^{i+1}}{\Theta(n)}\Big(\frac{\Theta(n-i)}{ij\Theta(n-j)}\Big)^{1/2}\binom{\theta+n-1}{n-j}.$$

Consequently,

$$\begin{aligned} w_{ij} &= \varSigma_1 - \varSigma_2 \\ &= (-1)^i \Big(\frac{\Theta(n-i)}{ij\Theta(n-j)}\Big)^{1/2} \sum_{0\le r\le i\wedge n-j} \frac{(-1)^r}{\Theta(n-r)}\binom{i}{r}\binom{\theta+n-r-1}{n-r-j} \\ &=: (-1)^i \Big(\frac{\Theta(n-i)}{ij\Theta(n-j)}\Big)^{1/2}\cdot\varSigma. \end{aligned} \tag{2.3}$$

Using the definition of $\Theta(m)$ given in (1.2) and applying identity (2.2), we find that

$$\varSigma = \frac{j!}{(\theta)_j}\sum_{0\le r\le i\wedge n-j} (-1)^r\binom{i}{r}\binom{n-r}{j} = \frac{j!}{(\theta)_j}\binom{n-i}{j-i}.$$

Here $\varSigma = 0$ if $j < i$ and $\varSigma = j!/(\theta)_j$ if $i = j$. Plugging this into (2.3), we obtain

$$w_{ij} = (-1)^i \Big(\frac{\Theta(n-i)}{ij\Theta(n-j)}\Big)^{1/2}\frac{j!}{(\theta)_j}\binom{n-i}{j-i} = \begin{cases} 0 & \text{if } i > j, \\ (-1)^j (j-1)!/(\theta)_j & \text{if } i = j. \end{cases}$$

This proves the main assertion of Theorem 2.

It remains to show that the eigenvector corresponding to μ_2 has the described form. However, this follows immediately from Theorem 3, which we prove in the next section. Theorem 2 is proved.

3 Proof of Theorem 3

We now find all eigenvectors of the matrix M_n. Again, we have to recall a useful identity.

Lemma 2. *Let $p, q, M \in \mathbb{N}_0$, $\alpha, \beta \in \mathbb{R}$, and $a_1, \dots, a_p; b_1 \dots, b_q$ be the parameters such that the hypergeometric series below is correctly defined. Then*

$$\sum_{k=0}^{M} \binom{M}{k} (\alpha)_{M-k} (\beta)_k \cdot {}_{p+1}F_q\big(-k, (a_p); (b_q); 1\big)$$
$$= (\alpha+\beta)_M \cdot {}_{p+2}F_{q+1}\big(-M, \beta, (a_p); \alpha+\beta, (b_q); 1\big).$$

Proof See formula (7) presented on page 388 in [18].

As a corollary, we find the next sum involving the above introduced polynomial

$$Q_r(x; \alpha, \beta, n) = {}_3F_2\big(-r, -x, r+\alpha+\beta+1; \alpha+1, -n+1; 1\big).$$

Lemma 3. *Let $n \geq 2$, $0 \leq M \leq n-1$ and $0 \leq r \leq n-1$. Then*

$$\Sigma_r(M) := \sum_{k=0}^{M} Q_r(k; 1, \theta-1, n)\Theta(M-k)$$
$$= \frac{(\theta+1)_M}{M!} {}_4F_3\big(-M, 1, -r, r+\theta+1; \theta+1, 2, 1-n; 1\big).$$

Proof Apply Lemma 2 for $\alpha = \theta$, $\beta = 1$, and $p = q = 2$.

The obtained expressions of $\Sigma_{r-1}(M), 1 \leq r \leq n$ will be used afterwards. For short, let

$${}_4F_3(-M) = {}_4F_3(-M, 1, 1-r, r+\theta; \theta+1, 2, 1-n; 1).$$

Lemma 4. *Let $\bar{y}_r = (y_{r1}, \dots, y_{rn}) = \pi_{r-1}\bar{e}_r M_n$, $1 \leq r \leq n$, then*

$$y_{ri} = -\left(\frac{\Theta(n-i)}{i}\right)^{1/2} \frac{n}{r(r+\theta-1)} \cdot {}_3F_2\big(-r, -n+i, r+\theta-1; \theta, -n; 1\big) \tag{3.1}$$

if $1 \leq i \leq n$.

*Proof I*n the notation above, applying Lemma 3, we obtain

$$
\begin{aligned}
y_{ri} &= \sum_{j=1}^{n} q_{r-1}(j)\Big(j\Theta(n-j)\Big)^{1/2} m_{ji} \\
&= \frac{1}{(i\Theta(n-i))^{1/2}} \sum_{j=1}^{n-i} Q_{r-1}(j-1;1,\theta-1,n)\Theta(n-j-i) \\
&\quad - \left(\frac{\Theta(n-i)}{i}\right)^{1/2} \frac{1}{\Theta(n)} \sum_{j=1}^{n} Q_{r-1}(j-1;1,\theta-1,n)\Theta(n-j) \\
&= \frac{1}{(i\Theta(n-i))^{1/2}} \Sigma_{r-1}(n-i-1) - \left(\frac{\Theta(n-i)}{i}\right)^{1/2} \frac{1}{\Theta(n)} \Sigma_{r-1}(n-1) \\
&= \frac{1}{\theta}\left(\frac{\Theta(n-i)}{i}\right)^{1/2} \Big[(n-i)\cdot {}_4F_3(-n+i+1) - n\cdot {}_4F_3(-n+1)\Big]
\end{aligned}
$$

if $1 \le i < n$.

Since $(a)_{l-1} = (a-1)_l/(a-1)$ if $a \neq 1$, we have

$$
\begin{aligned}
&(n-i)\cdot {}_4F_3(-n+i+1) = (n-i)\sum_{l=1}^{r} \frac{(-n+i+1)_{l-1}(1-r)_{l-1}(r+\theta)_{l-1}}{(\theta+1)_{l-1}(-n+1)_{l-1}l!} \\
&= -\frac{\theta n}{r(r+\theta-1)}\left[-1 + \sum_{l=0}^{r} \frac{(-n+i)_l(-r)_l(r+\theta-1)_l}{(\theta)_l(-n)_l l!}\right] \\
&= \frac{\theta n}{r(r+\theta-1)}\big[1 - {}_3F_2\big(-r,-n+i,r+\theta-1;\theta,-n;1\big)\big].
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
n\cdot {}_4F_3(-n+1) &= \frac{\theta n}{r(r+\theta-1)}\big[1 - {}_3F_2\big(-r,-n,r+\theta-1;\theta,-n;1\big)\big] \\
&= \frac{\theta n}{r(r+\theta-1)}\big[1 - {}_2F_1\big(-r,r+\theta-1;\theta;1\big)\big] \\
&= \frac{\theta n}{r(r+\theta-1)},
\end{aligned}
\tag{3.2}
$$

by virtue of the Chu-Vandermonde formula (see *e.g.* [2, (7.16)]). Hence

$$
\begin{aligned}
&(n-i)\cdot {}_4F_3(-n+i+1) - n\cdot {}_4F_3(-n+1) \\
&= -\frac{\theta n}{r(r+\theta-1)} \cdot {}_3F_2\big(-r,-n+i,r+\theta-1;\theta,-n;1\big).
\end{aligned}
$$

Plugging this into the previous expression of y_{ri}, we complete the proof in the case $i < n$.

If $i = n$, then, using Lemma 3 and (3.2), we obtain

$$\begin{aligned} y_{rn} &= -\frac{1}{\sqrt{n}\Theta(n)}\Sigma_{r-1}(n-1) = -\frac{1}{\sqrt{n}\Theta(n)} \cdot \frac{(\theta+1)_{n-1}}{(n-1)!} \cdot \frac{\theta}{r(r+\theta-1)} \\ &= -\frac{\sqrt{n}}{r(r+\theta-1)}. \end{aligned}$$

This is consistent with expression (3.1) given in the lemma.

Lemma 4 is proved.

Proof of Theorem 3. Let $1 \le r \le n$ be fixed. Recall that $q_k(x), 0 \le k \le r$, span the subspace of polynomials whose degrees do not exceed r. Analyse the polynomial appearing in Lemma 4, namely,

$$\Phi_r(x) = {}_3F_2\big(-r, -n+x, r+\theta-1; \theta, -n; 1\big).$$

As we have seen proving (3.2), we have $\Phi_r(0) = 0$. Hence $x^{-1}\Phi_r(x)$ is a polynomial of degree $r-1$, thus there exist constants $c_k \in \mathbb{R}$ such that

$$\Phi_r(x) = x\sum_{k=0}^{r-1} c_k q_r(x).$$

The leading coefficients of the polynomials $\Phi_r(x)$ and $q_r(x)$ are, respectively,

$$\frac{(-1)^r(r+\theta-1)_r}{(\theta)_r(-n)_r}, \qquad \frac{(r+\theta)_{r-1}}{r!(-n+1)_{r-1}}.$$

Consequently,

$$c_{r-1} = \frac{(-1)^r(r+\theta-1)_r}{(\theta)_r(-n)_r} \cdot \frac{r!(-n+1)_{r-1}}{(r+\theta)_{r-1}} = \frac{(-1)^{r-1}r!(r+\theta-1)}{(\theta)_r n}.$$

Now, the result of Lemma 4 can be rewritten as follows:

$$\begin{aligned} y_{ri} &= -\left(\frac{\Theta(n-i)}{i}\right)^{1/2} \frac{ni}{r(r+\theta-1)}\left[c_{r-1}q_{r-1}(i) + \sum_{k=0}^{r-2} c_k q_k(i)\right] \\ &= \big(i\Theta(n-i)\big)^{1/2}\left[\frac{(-1)^r(r-1)!}{(\theta)_r}q_{r-1}(i) + \sum_{k=0}^{r-2} d_k q_k(i)\right] \end{aligned}$$

with some coefficients $d_k = d_k(n, r, \theta)$ for each $1 \le i, r \le n$. Note that the fraction in the brackets is just μ_r found in Theorem 2.

For $1 \le r \le l \le n$, applying the last formula and the definition of the inner product (1.10), we obtain

$$\begin{aligned} \bar{e}_l M_n \bar{e}_r' &= \pi_{r-1}^{-1}\bar{e}_l y_r' = \frac{1}{\pi_{l-1}\pi_{r-1}}\left[\mu_r < q_{l-1}, q_{r-1} > + \sum_{k=0}^{r-2} d_k < q_{l-1}, q_k >\right] \\ &= \mu_r \delta_{rl} \end{aligned}$$

by virtue of orthogonality. This shows that each $\bar{e}_r$ in the basis is the eigenvector for M_n corresponding to μ_r.

Theorem 3 is proved.

Concluding remark. Comparing the expression $\bar{e}_r M_n = \mu_r \bar{e}_r$ with $\pi_{r-1}^{-1}\bar{y}_r$ given by Lemma 4, we arrive to a seemingly new relation of the generalized hypergeometric functions. For $1 \le i, r \le n$ and $\theta > 0$, it holds that

$$(-1)^{r-1} r! i \cdot {}_3F_2\big(-r+1, -i+1, r+\theta; 2, -n+1; 1\big) \\ = (\theta)_{r-1} n \cdot {}_3F_2\big(-r, -n+i, r+\theta-1; \theta, -n; 1\big).$$

Derivation of it using an appropriate sequence of the so-called contiguous relations (see [19]) would not be short.

Acknowledgment. The authors appreciate the benevolent remarks presented by two anonymous reviewers.

References

1. R. Arratia, A.D. Barbour and S. Tavaré, *Logarithmic Combinatorial Structures: a Probabilistic Approach*, EMS Publishing House, Zürich, 2003.
2. R. Askey, *Orthogonal Polynomials and Special Functions*, University of Wisconsin-Madison, Philadelphia, Pennsylvania, 1975.
3. R. Beals and R. Wong, *Special Functions*, Cambridge University Press, Cambridge, 2012.
4. H. Crane, The ubiquitous Ewens sampling formula, *Statist. Sci.*, **31**:1–19, 2016.
5. W. Ewens, The sampling theory of selectively neutral alleles, *Theor. Pop. Biol.*, **3**:87–112, 1972.
6. Sh. Feng, *The Poisson–Dirichlet Distribution and Related Topics*, Springer, Berlin, 2010.
7. N.S. Johnson, S. Kotz and N. Balakrishnan, *Discrete Multivariate Distributions*, Wiley, New York, 1997.
8. J. Klimavičius and E. Manstavičius, The Turán-Kubilius inequality on permutations, *Annales Univ. Sci. Budapest., Sect. Comp.*, **48**:45–51, 2018.
9. R. Koekoek, P.A. Lesky and R.F. Swarttouw, *Hypergeometric Orthogonal Polynomials and their q-Analogues*, Springer, Berlin, 2010.
10. J. Kubilius, Estimating the second central moment for strongly additive arithmetic functions, *Lietuvos matem. rink.*, **23**:110–117, 1983. Russian.
11. E. Manstavičius, Conditional probabilities in combinatorics. the cost of dependence, in M. Huškov'a and M. Janžura (Eds.), *Prague Stochastics. 7th Prague symposium on asymptotic statistics and 15th Prague conference on information theory, statistical decision functions and random processes*, Charles University, Matfyzpress, Prague, 2006, pp. 523–532.
12. E. Manstavičius, Moments of additive functions defined on the symmetric group, *Acta Appl. Math.*, **97**:119–127, 2007.
13. E. Manstavičius, An analytic method in probabilistic combinatorics, *Osaka J. Math.*, **46**:273–290, 2009.
14. E. Manstavičius, A sharp inequality for the variance with respect to the Ewens sampling formula, *Random Struct. Alg.*, **57**(4):1303–1313, 2020.
15. E. Manstavičius and V. Stepanauskas, On variance of an additive function with respect to a generalized Ewens probability, in M. Bousquet-Mélou and M. Soria (Eds.), *Proceedings of the 25th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, Discrete Math. Theor. Comput. Sci. Proc., BA, Association DMTCS, Nancy, 2014, pp. 301–311.

16. E. Manstavičius and Ž. Žilinskas, On a variance related to the Ewens sampling formula, *Nonlinear Anal. Model. Control*, **16**:453–466, 2011.

17. A.P. Prudnikov, Yu.A. Brychkov and O.I. Marichev, *Integrals and Series*, Volume 1, Taylor & Francis, 1998. Fourth printing.

18. A.P. Prudnikov, Yu.A. Brychkov and O.I. Marichev, *Integrals and Series. More Special Functions*, Volume 3, Gordon & Breach Sci. Publ., New York, 1990.

19. E.D. Rainville, *Special Functions*, Macmillan, New York, 1960.

20. G.A. Watterson, The sampling theory of selectively neutral alleles diffusion model, *Adv. Appl. Probab.*, **6**:463–488, 1974.

21. K. Wieand, *Eigenvalue Distributions of Random Matrices in the Permutation Group and Compact Lie Groups*, PhD thesis, Harvard University, 1998.

22. V. Zacharovas, Distribution of the logarithm of the order of a random permutation, *Lietuvos matem. rink.*, **44**:372–406, 2004. Russian.