Combinatorics, Probability and Computing (2014) 23, 889–913. © Cambridge University Press 2014
doi:10.1017/S0963548314000376

On Statistics of Permutations Chosen From the Ewens Distribution

TATJANA BAKSHAJEVA and EUGENIJUS MANSTAVIČIUS

Faculty of Mathematics and Informatics, Vilnius University, Naugarduko str. 24, LT-03225 Vilnius, Lithuania
(e-mail: tania.kargina@gmail.com)
Institute of Mathematics and Informatics, Vilnius University, Akademijos 4, LT-08663 Vilnius, Lithuania
(e-mail: eugenijus.manstavicius@mif.vu.lt)

Received 23 March 2013; revised 29 August 2013; first published online 22 August 2014

We explore the asymptotic distributions of sequences of integer-valued additive functions defined on the symmetric group endowed with the Ewens probability measure as the order of the group increases. Applying the method of factorial moments, we establish necessary and sufficient conditions for the weak convergence of distributions to discrete laws. More attention is paid to the Poisson limit distribution. The particular case of the number-of-cycles function is analysed in more detail. The results can be applied to statistics defined on random permutation matrices.

2010 *Mathematics subject classification*: Primary 60C05
Secondary 05A16, 20P05

1. Introduction

We deal with asymptotic value distribution problems of mappings defined on the symmetric group $\mathbf{S}_n$ as $n \to \infty$. Let $\sigma \in \mathbf{S}_n$ be an arbitrary permutation and $\sigma = \kappa_1 \cdots \kappa_w$ its representation as the product of independent cycles κ_i, and let $w := w(\sigma)$ be their number. If $k_j(\sigma)$, $1 \leqslant j \leqslant n$, denotes the number of cycles of length j in this decomposition, then

$$\bar{k}(\sigma) := \big(k_1(\sigma), \dots, k_n(\sigma)\big)$$

is called the cycle structure vector. The Ewens probability measure on the subsets $A \subset \mathbf{S}_n$ is defined by

$$\nu_n(A) := \nu_{n,\theta}(A) = \frac{1}{\theta^{(n)}} \sum_{\sigma \in A} \theta^{w(\sigma)},$$

where $\theta > 0$ is a fixed parameter and $\theta^{(n)} := \theta(\theta+1)\cdots(\theta+n-1)$. Set $\ell(\bar{s}) = 1s_1 + \cdots + ns_n$ if $\bar{s} = (s_1, \dots, s_n) \in \mathbb{Z}_+^n$. An easy combinatorial argument gives the distribution of the cycle structure vector

$$\nu_n\big(\bar{k}(\sigma)=\bar{s}\big) = \frac{n!}{\theta^{(n)}}\prod_{j=1}^{n}\left(\frac{\theta}{j}\right)^{s_j}\frac{1}{s_j!} =: P_\theta(\{\bar{s}\}) \tag{1.1}$$

if $\bar{s}\in\ell^{-1}(n)\subset\mathbb{Z}_+^n$. The probability $P_\theta(\{\bar{s}\})$ defines an induced measure on $\ell^{-1}(n)$ and is called the Ewens Sampling Formula (henceforth denoted by ESF). If ξ_j, $j\geqslant 1$, denote independent Poisson random variables given on some probability space $\{\Omega,\mathcal{F},P\}$ with $\mathbb{E}\xi_j=\theta/j$ and $\bar{\xi}:=(\xi_1,\dots,\xi_n)$, then

$$\nu_n\big(\bar{k}(\sigma)=\bar{s}\big) = P\big(\bar{\xi}=\bar{s}|\,\ell(\bar{\xi})=n\big),\quad \bar{s}\in\ell^{-1}(n). \tag{1.2}$$

Under the probability measure ν_n, the quantities $k_j(\sigma)$, $1\leqslant j\leqslant n$, are dependent random variables. We prefer to leave the elementary event σ in its notation, although this is in contrast to other random variables defined as above on a non-specialized space $\{\Omega,\mathcal{F},P\}$. It is known (see [2, Theorem 7.7]) that the total variation distance

$$\frac{1}{2}\sum_{s_1,\dots,s_r\geqslant 0}\big|\nu_n\big(k_1(\sigma)=s_1,\dots,k_r(\sigma)=s_r\big)-P(\xi_1=s_1,\dots,\xi_r=s_r)\big| = o(1) \tag{1.3}$$

if and only if $r=o(n)$. Here and in what follows, we assume that $n\to\infty$. The book [2] is a good reference for the listed and many more properties of the ESF.

In the present paper, we discuss the asymptotic value distribution of an additive (completely additive) function $h:\mathbf{S}_n\to\mathbb{R}$ with respect to ν_n. Such a function is defined via a real array $\{a_j, j\geqslant 1\}$, by setting

$$h(\sigma) := \sum_{j\leqslant n} a_j k_j(\sigma). \tag{1.4}$$

Taking arrays a_{nj}, $1\leqslant j\leqslant n$, $n\geqslant 1$, we obtain sequences of functions $h_n(\sigma)$. So, if $a_{nj}=1$ for $j\in J_n\subset\{1,\dots,n\}$ and $a_{nj}=0$ otherwise, we have the number-of-cycles function which includes the restricted to J_n cycle lengths only. In what follows, we denote it by $w(\sigma,J_n)$. Apart from the latter example, the additive functions are involved in many combinatorial, algebraic and statistical problems. The function $h(\sigma)$ defined via $a_j=\log j$, $j\leqslant n$, approximates the logarithm of the group theoretical order of almost all permutations $\sigma\in\mathbf{S}_n$ well (see [10] and [36] or [37]). Particular additive functions appear in physical models via Hamiltonians in Bose gas theory (see [4]–[6] and the references therein). They are indispensable in treating the random permutation matrix ensemble. Let $M:=M(\sigma):=\big(\mathbf{1}\{i=\sigma(j)\}\big)$, $1\leqslant i,j\leqslant n$ and $\sigma\in\mathbf{S}_n$, be such a matrix taken with the weighted frequency $\nu_n(\{M\})=\nu_n(\{\sigma\})=\theta^{w(\sigma)}/\theta^{(n)}$, let

$$Z_n(x;\sigma) := \det\big(I-xM(\sigma)\big) = \prod_{j\leqslant n}(1-x^j)^{k_j(\sigma)} \tag{1.5}$$

be its characteristic polynomial, and let $e^{2\pi i\varphi_j(\sigma)}$, where $\varphi_j(\sigma)\in[0,1)$ and $j\leqslant n$, be its eigenvalues. The papers [12], [13], [34], [35], and [39], and some arXiv preprints (see, for instance, [1] and the references therein), concern $\log|Z_n(x;\sigma)|$, $\Im\log Z_n(x;\sigma)$ or the

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

trace-related statistics

$$\mathrm{Trf}(\sigma) := \sum_{j\leqslant n} f\big(\varphi_j(\sigma)\big) = \sum_{j\leqslant n} k_j(\sigma) \sum_{0\leqslant s\leqslant j-1} f\left(\frac{s}{j}\right), \tag{1.6}$$

where $f : [0,1] \to \mathbb{R}$ is a function. An indicator function $f = \mathbf{1}_A$ of an interval $A \subset [0,1]$ or other integer-valued functions fall within the scope of the present paper.

So far, the general problem to find necessary and sufficient conditions under which the distribution function

$$T_n(x) := \nu_n\big(h_n(\sigma) < x\big)$$

weakly converges to a limit law is out of reach even for $\theta = 1$. The case when the sequence $a_{nj} = a_j$, $j \geqslant 1$, *i.e.*, does not depend on n, is easier. Then the answer is given by an analogue of the three series theorem of Kolmogorov (see Theorem 8.25 in [2]). If $\theta = 1$ and $h_n(\sigma) = h(\sigma)/\beta(n)$, where the function $h(\sigma)$ is fixed and $\beta(n) > 0$, $\beta(n) \to \infty$ but $\beta(un)/\beta(n) \to 1$ for every fixed $0 < u < 1$ (slowly oscillating at infinity), necessary and sufficient conditions were established in the second author's paper [27], which contains an extensive reference list of earlier papers by other authors. If $\beta(n)$ is regularly varying at infinity, the first results go back to [20]. The problem remains open if no *a fortiori* condition on $\beta(n)$ is taken. On the other hand, for partial sum processes defined by additive functions, convergence of distributions in appropriate function spaces to infinitely divisible measures implies slow oscillation of $\beta(n)$. This further yields necessary and sufficient convergence conditions even for generalized Ewens probability measures. For the latest account in this direction, we refer to [8].

In the present paper, we focus on sequences of additive functions $h_n(\sigma)$ defined via a_{nj}. Henceforth, we will often use the abbreviation $a_j = a_{nj}$ without the index n and take $a_j = 0$ if $j > n$. For $\theta = 1$, the partial sums of such functions have been used to model stochastic processes [22]. Recently, we succeeded in establishing necessary and sufficient conditions for the weak law of large numbers if $\theta \geqslant 1$ (see [17]). Some success in proving general limit theorems has been achieved for the integer-valued functions $h_n(\sigma)$. The case for $\theta = 1$ was explored by the second author in [23]–[25]. For $\theta > 0$, the first author in [14] and [15] obtained an exhaustive result for the sequence $w(\sigma, J_n)$. We now generalize this line of research with the case $a_j \in \mathbb{Z}_+$ for $j \leqslant n$ and supply a few examples shedding more light on the class of possible limit distribution for $w(\sigma, J_n)$. On the other hand, one of the purposes of the present paper is to demonstrate the factorial moment method. The approach proved to be useful in a series of number-theoretical papers by J. Šiaulys [29]–[31]. The idea lies in analysis of the expressions of moments. Although involved, they contain the key information useful in establishing necessary and sufficient conditions for the convergence of distributions.

In what follows, let $\Rightarrow$ stand for the weak convergence and let $F_Y(x)$ be a distribution function of a random variable Y concentrated on $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$. The mean value with respect to ν_n of a function $g(\sigma)$ defined on $\mathbf{S}_n$ will be denoted by $\mathbb{E}_n g(\sigma)$. Set $x_{(r)} = x(x-1)\cdots(x-r+1)$, $r \geqslant 1$ for the falling factorial and $x_{(0)} = 1$. Henceforth, unless indicated otherwise, we take $i, r, r_i, j, j_i \in \mathbb{N}$ and $j, j_i \leqslant n$. The first two theorems involve

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

the quantity

$$\Upsilon_n(l,m;h) := \sum_{u=1}^{l} \theta^u \sum_{\substack{r_1+\cdots+r_u=l \\ 1\leqslant r_i\leqslant m, i\leqslant u}} \binom{l-1}{r_1-1} \cdots \binom{l-r_1-\cdots-r_{u-1}-1}{r_u-1}$$
$$\times \sum_{j_1+\cdots+j_u<n} \frac{a_{j_1(r_1)}\cdots a_{j_u(r_u)}}{j_1\cdots j_u}\left(1-\frac{j_1+\cdots+j_u}{n}\right)^{\theta-1},$$

which is an approximation of the lth factorial moment of an appropriately truncated additive function obtained from $h(\sigma)$.

Let $\rho_n(m)$ be a generic error term, whose exact definition will vary but will satisfy the relation

$$\lim_{m\to\infty} \limsup_{n\to\infty} |\rho_n(m)| = 0.$$

Henceforth, the estimates $O(\cdot)$ and $o(\cdot)$ as $n\to\infty$ will depend on θ and some other quantities. To stress possible dependence on a parameter, say K, we will write $O_K(\cdot)$ and $o_K(\cdot)$. The first result concerns necessary conditions for convergence.

Theorem 1.1. *For $\theta \geqslant 1$, let $h_n(\sigma)$ be a sequence of integer-valued additive functions defined via $a_j = a_{nj}$, and let Y be a random variable taking values in $\mathbb{Z}_+$ and such that $\mathbb{E}Y^\alpha < \infty$ for $\alpha \geqslant 2+\varepsilon > 2$. If $T_n(x) \Rightarrow F_Y(x)$, then*

$$\lim_{n\to\infty} \sum_{j<n} \frac{\mathbf{1}\{a_j \leqslant -1\}}{j}\left(1-\frac{j}{n}\right)^{\theta-1} = 0 \tag{1.7}$$

and

$$\Upsilon_n(l,m;h) - \mathbb{E}Y_{(l)} = \rho_n(m) \tag{1.8}$$

for each fixed natural number $l \leqslant \alpha - 1 - \varepsilon$.

We have to confess that the technical condition $\theta \geqslant 1$ in Theorem 1.1 and in some subsequent results concerning the necessity is undesirable. Sufficient convergence conditions are given by the following result.

Theorem 1.2. *Let $\theta > 0$ and let $h_n(\sigma)$ be a sequence of integer-valued additive functions. Assume that condition* (1.7) *is satisfied. If there exists a sequence $\Upsilon(l)$ such that*

$$\Upsilon_n(l,m;h) - \Upsilon(l) = \rho_n(m) \tag{1.9}$$

for every $l \in \mathbb{N}$ and

$$\sum_{l=0}^{\infty} \frac{\Upsilon(l)2^l}{l!} < \infty,$$

then $T_n(x) \Rightarrow F_Y(x)$ and $\mathbb{E}Y_{(l)} = \Upsilon(l)$ for $l \geqslant 1$.

The following corollary gives necessary and sufficient convergence conditions for more specialized cases. Let $\Pi_\mu(x)$ be the distribution function of the Poisson law with parameter $\mu > 0$.

Corollary 1.3. *For $\theta \geqslant 1$, let $h_n(\sigma)$ be a sequence of integer-valued additive functions. The convergence $T_n(x) \Rightarrow \Pi_\mu(x)$ holds if and only if condition* (1.7) *holds and*

$$\Upsilon_n(l, m; h) - \mu^l = \rho_n(m) \tag{1.10}$$

for every $l \in \mathbb{N}$.

As the total variation estimate approximation (1.3) shows, the Poisson distribution appears as a limit if the cycles of lengths up to $r = o(n)$ are involved. By the next corollary, we demonstrate that one can find $a_j = a_{nj}$, $n/2 < j \leqslant n$, which defines a sequence of additive functions obeying a Poisson limit law with a sufficiently small μ. Such a phenomenon has been observed in [22] if $\theta = 1$. The construction involves the following strictly increasing in $x \in [1/2, 1]$ function

$$t_\theta(x) := \theta \int_{1/2}^{x} (1-u)^{\theta-1} \frac{du}{u}.$$

We will prove that $t_\theta(1) < 1$ if $\theta \geqslant 1$.

Corollary 1.4. *Let $\theta \geqslant 1$, $\mu \leqslant -\log(1 - t_\theta(1))$, and $a_j = a_{nj} \in \mathbb{Z}_+$ be such that*

$$\lim_{n\to\infty} \sum_{j \leqslant n/2} \frac{\mathbf{1}\{a_j \neq 0\}}{j} = 0. \tag{1.11}$$

The convergence $T_n(x) \Rightarrow \Pi_\mu(x)$ holds if and only if

$$\theta \sum_{n/2<j<n} \frac{\mathbf{1}\{a_j = k\}}{j} \left(1 - \frac{j}{n}\right)^{\theta-1} = \mathrm{e}^{-\mu} \frac{\mu^k}{k!} + o_k(1) \tag{1.12}$$

for every $k \in \mathbb{N}$.

If $a_j = a_{nj}$ remain bounded for the overwhelming proportion of $j \leqslant n$, we obtain some results for $\theta < 1$. Introduce the quantity $\Upsilon_n(l; h) := \Upsilon_n(l, \infty; h)$. As we will prove in Lemma 3.4 below, it represents the main asymptotical term of the lth factorial moment of $h(\sigma)$, namely,

$$\mathbb{E}_n h(\sigma)_{(l)} = \Upsilon_n(l; h) + O_{K,l}\left(\frac{1 + \log^l n}{n^{1\wedge\theta}}\right),$$

where $a \wedge b := \min\{a, b\}$, provided that $a_j \in \mathbb{Z}_+$ and $a_j \leqslant K$ if $j \leqslant n$.

Theorem 1.5. *Let $\theta > 0$, $a_j = a_{nj} \in \mathbb{Z}_+$, $j \leqslant n$, and, for some $K \in \mathbb{N}$,*

$$\sum_{j<n} \frac{\mathbf{1}\{a_j \geqslant K\}}{j} \left(1 - \frac{j}{n}\right)^{\theta-1} = o_K(1). \tag{1.13}$$

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

The convergence $T_n(x) \Rightarrow F_Y(x)$ *holds if and only if there exists a sequence* $\Upsilon(l)$ *such that*

$$\lim_{n\to\infty} \Upsilon_n(l;h) = \Upsilon(l) \tag{1.14}$$

for every $l \in \mathbb{N}$. *If this condition holds, then* $\mathbb{E}Y_{(l)} = \Upsilon(l)$ *for* $l \geqslant 1$.

For the number-of-cycles function $w(\sigma, J_n)$, condition (1.14) attains the simplest form. Let the asterisk $*$ over a sum mean the conditions $a_j = a_{nj} = 1$ and $a_{j_i} = 1$, $i \geqslant 1$, or, equivalently, for $j, j_i \in J_n$. Set

$$V_n(x) := \nu_n\big(w(\sigma, J_n) < x\big)$$

and

$$\upsilon_n(l) := \Upsilon_n(l, w) = \theta^l \sum_{j_1+\cdots+j_l<n}^{*} \frac{1}{j_1\cdots j_l}\left(1 - \frac{j_1+\cdots+j_l}{n}\right)^{\theta-1}, \quad l \geqslant 1.$$

The next corollary of Theorem 1.5 has been proved in [14] and [15].

Corollary 1.6. *Let* $\theta > 0$ *and* $J_n \subset \{1, \ldots, n\}$ *be arbitrary. The distribution function* $V_n(x) \Rightarrow F_Y(x)$ *if and only if there exists a sequence* $\upsilon(l)$, $l \geqslant 1$, *such that*

$$\lim_{n\to\infty} \upsilon_n(l) = \upsilon(l) \tag{1.15}$$

for every $l \in \mathbb{N}$. *If the latter condition is satisfied, then* $\mathbb{E}Y_{(l)} = \upsilon(l)$ *for* $l \geqslant 1$.

Here is a particular case.

Corollary 1.7. *Let* $\theta \geqslant 1$ *and* $J_n \subset \{1, \ldots, n\}$ *be arbitrary. The convergence* $V_n(x) \Rightarrow \Pi_\mu(x)$ *holds if and only if there exists a sequence* $r = r(n) = o(n)$ *such that conditions*

$$\sum_{j\leqslant r}^{*} \frac{\theta}{j} = \mu + o(1), \quad \sum_{r<j<n}^{*} \frac{1}{j}\left(1 - \frac{j}{n}\right)^{\theta-1} = o(1) \tag{1.16}$$

are satisfied.

Corollary 1.7 demonstrates that a Poisson law can appear in the limit for $w(\sigma, J_n)$ only if the proportion of $j \in J_n \cap [\varepsilon n, n]$ is negligible for each $0 < \varepsilon < 1$. If $\theta = 1$, this fact has been observed in [25] even for bounded $a_j = a_{nj}$ where $j \leqslant n$. Non-degenerate limit distributions concentrated on the finite set $\{0, 1, \ldots, L-1\}$ where $L \geqslant 2$ are of particular interest.

Corollary 1.8. *Let* $\theta > 0$, $J_n \subset \{1, \ldots, n\}$, *and* $L \in \mathbb{N} \setminus \{1\}$ *be arbitrary. Assume that* Y *is a random variable taking values in* $\{0, 1, \ldots, L-1\}$ *and such that* $\mathbb{E}Y_{(l)} = \upsilon(l)$ *if* $l \leqslant L-1$. *If*

$$\lim_{n\to\infty} \theta^l \sum_{n/L<j_1,\ldots,j_l<n}^{*} \frac{\mathbf{1}\{j_1+\cdots+j_l<n\}}{j_1\cdots j_l}\left(1 - \frac{j_1+\cdots+j_l}{n}\right)^{\theta-1} = \upsilon(l) \tag{1.17}$$

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

for each $l \leqslant L-1$ *and*

$$\sum_{j\leqslant n/L}^{*} \frac{1}{j} = o_L(1), \tag{1.18}$$

then $V_n(x) \Rightarrow F_Y(x)$.

Conversely, if $\theta \leqslant 1$ *and* $V_n(x) \Rightarrow F_Y(x)$, *then conditions* (1.17) *and* (1.18) *are satisfied.*

The rest of the paper is organized as follows. Section 2 contains examples and some observations on the results formulated above. Using the opportunity, we include an addendum to the first author's papers [14] and [15] concerning the class of possible limit laws for $w(\sigma, J_n)$. This was stimulated by Lugo's Conjecture stated in [19] and disproved below. Section 3 contains formulae of factorial moments and required lower estimates of some frequencies. The proofs of Theorems 1.1 and 1.2 are presented in Sections 4 and 5. The last section deals with the case of bounded a_j, including the function $w(\sigma, J_n)$.

2. Examples and a discussion of limit laws

Example for Corollary 1.4. Let $t_\theta(x)$ be the function defined above, $\theta \geqslant 1$, and

$$0 < \mu \leqslant -\log(1 - t_\theta(1)).$$

Introduce the sequence $1/2 = d_0 < d_1 < \cdots$ by

$$t_\theta(d_k) = \mathrm{e}^{-\mu} \sum_{l=1}^{k} \frac{\mu^l}{l!}, \quad k = 1, 2, \ldots,$$

and set $a_j = k$ if $nd_{k-1} < j \leqslant nd_k$ and $a_j = 0$ otherwise. We claim that the additive function $h_n(\sigma)$ defined via such a_j satisfies condition (1.12).

First, check that the function $t_\theta(x)$ is strictly increasing in x and $t_\theta(1) < 1$. Indeed, the latter is evident if $1 \leqslant \theta \leqslant 1/\log 2$. Otherwise, $t_\theta(1) \leqslant t_{1/\log 2}(1)$ for $\theta \geqslant 1/\log 2$. The properties observed ensure that the sequence d_k is correctly defined. Approximating the sum by the Riemann integral, we have

$$\begin{aligned} &\sum_{n/2<j<n} \frac{\theta \mathbf{1}\{a_j = k\}}{j} \left(1 - \frac{j}{n}\right)^{\theta-1} \\ &\quad = \sum_{nd_{k-1}<j<nd_k} \frac{\theta \mathbf{1}\{a_j = k\}}{j} \left(1 - \frac{j}{n}\right)^{\theta-1} \\ &\quad = t_\theta(d_k) - t_\theta(d_{k-1}) + o_k(1) = \mathrm{e}^{-\mu} \frac{\mu^k}{k!} + o_k(1). \end{aligned}$$

The claim is established.

Bounded $a_j = a_{nj}$. Let $K, k \geqslant 1$ be integers, $0 \leqslant a_j \leqslant K$ if $n/(k+1) < j \leqslant n/k$ and $a_j = 0$ elsewhere. For an additive function h_n defined via such a_j, we have $\Upsilon_n(l, K; h) = 0$ if $l > kK$. Indeed, by analysing the expression of $\Upsilon_n(l, K; h)$, we see that the non-zero summands correspond to the indices $r_i \leqslant K$, $i \leqslant u \leqslant k$, satisfying $r_1 + \cdots + r_u = l$. The

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

latter is impossible if $l > kK$. Consequently, if a limit law exists for such a function h_n, it should be concentrated on a finite set.

On the other hand, we have failed to answer the following question.

Problem 1. Is it possible to construct an additive function obeying a Poisson limit law using unbounded $a_j = a_{nj}$ such that $a_j = 0$ for all but $j \in (n/(k+1), n/k]$ where $k \geqslant 2$?

M. Lugo's Conjecture [19]. By definition, a random variable $q\Pi$ is called (k,λ) quasi-Poisson if it has the distribution

$$P(q\Pi = i) = \sum_{j=i}^{k} \binom{j}{i} (-1)^{j-1} \lambda^j, \quad i = 0, 1, \ldots k,$$

where $0 < \lambda \leqslant 1$. The factorial moment $\mathbb{E}q\Pi_{(l)} = \lambda^l$ if $l \leqslant k$ and $\mathbb{E}q\Pi_{(l)} = 0$ if $l > k$. If $\theta = 1$, it is easy to define a subset J_n so that $w(\sigma, J_n)$ obeys the quasi-Poisson limit law. In fact this and some other results from [18] were included in Theorem 1.3 of the second author's paper [24]. Lugo wrote on page 13 of [19]: 'in the case of the Ewens distribution, the following conjecture seems reasonable'.

Conjecture 2.1 ([19]). *The expected number of cycles of length in $[\gamma n, \delta n]$ of a permutation of $\{1, \ldots, n\}$ chosen from the Ewens distribution approaches*

$$\lambda = \int_{\gamma}^{\delta} (1-x)^{\theta-1} \frac{dx}{x}$$

as $n \to \infty$. Furthermore, in the case where $1/(k+1) \leqslant \gamma < \delta < 1/k$ for some positive integer k, the distribution of the number of cycles converges in distribution to quasi-Poisson (k, λ).

The factor θ is missing in the formula for λ and the second claim of the conjecture is false if $\theta \neq 1$. Here is a counter-example.

Examine $w(\sigma, J_n)$, where $J_n = \{j : \ n/3 < j \leqslant n/2\}$. A routine approximation of sums by the Riemann integrals yields the asymptotic formulas for the first two factorial moments via the quantities

$$\begin{aligned} \upsilon_n(1) &= \sum_{n/3 < j \leqslant n/2} \frac{\theta}{j} \left(1 - \frac{j}{n}\right)^{\theta-1} + o(1) \\ &= \theta \int_{1/3}^{1/2} (1-u)^{\theta-1} \frac{du}{u} + o(1) =: \lambda + o(1) \end{aligned}$$

and

$$\begin{aligned} \upsilon_n(2) &= \theta^2 \sum_{n/3 < i,j \leqslant n/2} \frac{1}{ij} \left(1 - \frac{i+j}{n}\right)^{\theta-1} + o(1) \\ &= \theta^2 \int_{1/3}^{1/2} \int_{1/3}^{1/2} (1-u-v)^{\theta-1} \frac{du\,dv}{uv} + o(1). \end{aligned}$$

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

Hence

$$
\begin{aligned}
&\upsilon_n(1)^2 - \upsilon_n(2) \\
&\quad = \theta^2 \int_{1/3}^{1/2} \int_{1/3}^{1/2} \big((1-u)^{\theta-1}(1-v)^{\theta-1} - (1-u-v)^{\theta-1}\big) \frac{du\,dv}{uv} \\
&\qquad + o(1) \geqslant c_0 > 0
\end{aligned}
$$

if $\theta > 1$, c_0 is sufficiently small, and n is sufficiently large. If $\theta < 1$, the difference $\upsilon_n(1)^2 - \upsilon_n(2) < -c_0 < 0$ for sufficiently large n. Observing also that $\upsilon_n(l) = 0$ if $l \geqslant 3$, by Corollary 1.6, we see that the function $w(\sigma, J_n)$ obeys a limit distribution but it is not the $(2, \lambda)$ quasi-Poisson.

Problem 2. Let $\mathcal{L}$ be the class of possible limit distributions for $w(\sigma, J_n)$, for arbitrary $J_n \subset \{1, \ldots, n\}$, under the Ewens probability measure ν_n. Find a description of $\mathcal{L}$.

We now present some related examples based on Corollary 1.6.

Bernoulli distribution $\mathrm{Be}(p)$, where p is the parameter $p \in (0,1)$. We claim that $\mathrm{Be}(p) \in \mathcal{L}$ if $p \leqslant t_\theta(1)$, where $t_\theta(x)$ is the previously defined function on $[1/2, 1]$ and $\theta \geqslant 1$. The construction is based on the factorial moments. For $\mathrm{Be}(p)$, they are $\upsilon(1) = p$ and $\upsilon(l) = 0$ if $l \geqslant 2$. It suffices, therefore, to find an α such that $t_\theta(\alpha) = p$ and to define $J_n = \{j \leqslant n : n/2 < j \leqslant \alpha n\}$. By a simple approximation of the sum by the integral, we verify condition (1.15) and find that $\upsilon_n(1) = p + o(1)$.

Binomial distribution $\mathrm{Bi}(M, p)$, where $M \in \mathbb{N}$ and $p \in (0,1)$. Now, the factorial moments are equal to $M_{(l)}p^l$ if $l = 1, 2, \ldots, M$ and to zero if $l = M+1, M+2, \ldots$. For simplicity, we confine ourselves to a particular case of $\theta = 1$ and $M = 2$. We claim that $\mathrm{Bi}(2, p) \in \mathcal{L}$ if

$$0 < p \leqslant (1/2)\log 3 = 0.405\ldots.$$

The idea to construct such an example has been shown in [33]. Let us take two temporary parameters $0 < \alpha \leqslant \log 2$ and $0 < \beta \leqslant \log(3/2)$. Define the sequence of sets of natural numbers

$$J_n = \mathbb{N} \cap \big((n/3, (n/3)\mathrm{e}^\alpha] \cup (2n/3, (2n/3)\mathrm{e}^\beta]\big).$$

The factorial moments of the additive function $w(\sigma, J_n)$ are approximated by

$$
\begin{aligned}
\upsilon_n(1) &= \sum_{j \in J_n} \frac{1}{j} = \alpha + \beta + o(1) \leqslant \log 3 + o(1), \\
\upsilon_n(2) &= \sum_{i,j \in J_n} \mathbf{1}\{i + j \leqslant n\} \frac{1}{ij} = \Bigg(\sum_{n/3 < j \leqslant (n/3)\mathrm{e}^\alpha} \frac{1}{j} \Bigg)^2 = \alpha^2 + o(1),
\end{aligned}
$$

and $\upsilon_n(l) = 0$ if $l \geqslant 3$. To get the binomial distribution, we have to require that

$$2\alpha^2 = (\alpha + \beta)^2.$$

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

Hence

$$\alpha = (\sqrt{2}+1)\beta.$$

Given $p \leqslant (1/2)\log 3$, we can choose β and, consequently, α so that

$$2p = \alpha + \beta = (\sqrt{2}+2)\beta \leqslant \log 3.$$

Now, taking $\beta = (2-\sqrt{2})p$ and $\alpha = p\sqrt{2}$, due to the condition on p, we are done.

The laws outside $\mathcal{L}$. It is easy to see that $\upsilon_n(l) \leqslant \upsilon_n(1)^l$ if $\theta \geqslant 1$. Consequently, the inequality should be preserved by the laws in $\mathcal{L}$. In fact this observation is due to J. Šiaulys and G. Stepanauskas [32]. Distributions such as the geometric with a parameter $p \in (0,1)$ or a mixed Poisson distribution $F_Y(x) = \Pi(x;\beta,\lambda,\tau)$ defined by the factorial moments

$$\mathbb{E}Y_{(l)} = \beta\lambda^l + (1-\beta)\tau^l, \quad l = 1,2,\ldots,$$

where $0<\beta<1$, $\lambda,\tau>0$, and $\lambda \neq \tau$, do not belong to $\mathcal{L}$ if $\theta \geqslant 1$.

3. Lemmas

In this section we present exact expressions of the factorial moments of a completely additive function $h(\sigma)$ defined via $a_j \in \mathbb{R}$. Particular attention is devoted to the case of bounded a_j and approximations. Denote

$$\psi_n(m) = \frac{n!}{\theta^{(n)}}\,\frac{\theta^{(m)}}{m!} = \prod_{k=m+1}^{n}\left(1+\frac{\theta-1}{k}\right)^{-1},$$

where $0 \leqslant m \leqslant n$. It is well known that

$$\frac{\theta^{(m)}}{m!} = \frac{m^{\theta-1}}{\Gamma(\theta)}\left(1+O\left(\frac{1}{m}\right)\right), \quad m \geqslant 1,$$

where $\Gamma(u)$ is Euler's gamma function. Hence

$$\psi_n(m) = \left(\frac{m}{n}\right)^{\theta-1}\left(1+O\left(\frac{1}{m}\right)\right), \quad 1 \leqslant m \leqslant n. \tag{3.1}$$

Henceforth, we will use the inequalities

$$\psi_n(n-i-j) \geqslant \psi_n(n-i)\psi_n(n-j) \quad \text{if } \theta \leqslant 1$$

and

$$\psi_n(n-i-j) \leqslant \psi_n(n-i)\psi_n(n-j) \quad \text{if } \theta \geqslant 1,$$

valid for $0 \leqslant i,j \leqslant n$.

It is worth recalling Watterson's formula.

Lemma 3.1. *For $(j_1,\ldots,j_r) \in \mathbb{Z}_+^r$, $l = 1j_1+\cdots+rj_r$ and $1 \leqslant r \leqslant n$,*

$$\mathbb{E}_n\left\{\prod_{i=1}^{r} k_{i(j_i)}(\sigma)\right\} = \psi_n(n-l)\mathbf{1}\{l \leqslant n\}\prod_{i=1}^{r}\left(\frac{\theta}{i}\right)^{j_i}. \tag{3.2}$$

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

Proof. See (5.6) on page 96 of [2]. □

The next lemma somewhat extends the previous formula.

Lemma 3.2. *Let $\theta > 0$. For a completely additive function $h(\sigma)$ and every $k \in \mathbb{N}$, we have*

$$\begin{aligned}\mathbb{E}_n h(\sigma)_{(k)} &= \gamma_n(k)\\ &:= \sum_{u=1}^{k} \theta^u \sum_{r_1+\cdots+r_u=k} \binom{k-1}{r_1-1} \cdots \binom{k-r_1-\cdots-r_{u-1}-1}{r_u-1}\\ &\quad\times \sum_{j_1+\cdots+j_u \leqslant n} \frac{a_{j_1(r_1)}\cdots a_{j_u(r_u)}}{j_1\cdots j_u} \varphi_n\big(n-(j_1+\cdots+j_u)\big). \qquad (3.3)\end{aligned}$$

Proof. We first derive a recurrence relation for

$$\beta_n(k) := \big(\theta^{(n)}/n!\big)\gamma_n(k),$$

where $\beta_n(0) = \theta^{(n)}/n!$ if $n \geqslant 0$ and $\beta_0(k) = 0$ if $k \geqslant 1$. At this stage, it suffices to apply the formal power series algebra (see [11], Chapter III). Let z and w be formal variables, $\varphi_0(z) := 1$, and

$$\varphi_n(z) := \frac{\theta^{(n)}}{n!}\mathbb{E}_n z^{h(\sigma)} = \sum_{k\geqslant 0} \frac{\beta_n(k)}{k!} z^k.$$

Thus, $\beta_n(k)/k! = [z^k]\varphi_n(z)$, the kth power series coefficient (in fact, of a polynomial). The application of formal derivatives simplifies the notation and allows us to use the classical rules of differentiation. For this reason, we also write $\beta_n(k) = \varphi_n^{(k)}(z)|_{z=1}$.

Grouping over the classes of σ with the common cycle vector and using Cauchy's formula for the cardinality of a class, we have

$$\varphi_n(z) = \sum_{\ell(\bar{k})=n} \prod_{j=1}^{n} \left(\frac{\theta z^{a_j}}{j}\right)^{k_j} \frac{1}{k_j!}.$$

This leads to the formal power series equations

$$\sum_{n\geqslant 0} \varphi_n(z) w^n = \exp\left\{\theta \sum_{j\geqslant 1} \frac{z^{a_j}}{j} w^j\right\}$$

and

$$\begin{aligned}\sum_{n\geqslant 0} \varphi_n'(z) w^n &= \theta \sum_{m\geqslant 0} \varphi_m(z) w^m \cdot \sum_{j\geqslant 1} \frac{a_j z^{a_j-1}}{j} w^j\\ &= \theta \sum_{n\geqslant 0} \left(\sum_{j\leqslant n} \varphi_{n-j}(z) \frac{a_j z^{a_j-1}}{j}\right) w^n.\end{aligned}$$

Hence

$$\varphi_n'(z) = \theta \sum_{j\leqslant n} \varphi_{n-j}(z) \frac{a_j z^{a_j-1}}{j}. \qquad (3.4)$$

Taking the $(k-1)$th-order derivatives with respect to z, we arrive at

$$\varphi_n^{(k)}(z) = \theta \sum_{j\leqslant n} \sum_{l=0}^{k-1} \binom{k-1}{l} \frac{a_{j(l+1)} z^{a_j-l-1}}{j} \varphi_{n-j}^{(k-1-l)}(z).$$

This is a relation between two polynomials, thus we may take $z=1$ to obtain

$$\beta_n(k) = \theta \sum_{r=1}^{k-1} \binom{k-1}{r-1} \sum_{j\leqslant n} \frac{a_{j(r)}}{j} \beta_{n-j}(k-r) + \theta \sum_{j\leqslant n} \frac{a_{j(k)}}{j} \frac{\theta^{(n-j)}}{(n-j)!}. \tag{3.5}$$

We now apply mathematical induction to prove that

$$\beta_n(k) = \sum_{u=1}^{k} \theta^u \sum_{r_1+\cdots+r_u=k} \binom{k-1}{r_1-1} \cdots \binom{k-r_1-\cdots-r_{u-1}-1}{r_u-1} \\ \times \sum_{j_1+\cdots+j_u\leqslant n} \frac{a_{j_1(r_1)}\cdots a_{j_u(r_u)}}{j_1\cdots j_u} \frac{\theta^{(n-j_1-\cdots-j_u)}}{(n-j_1-\cdots-j_u)!}. \tag{3.6}$$

A direct application of (3.4) yields

$$\beta_n(1) = \theta \sum_{j\leqslant n} \frac{a_j}{j} \frac{\theta^{(n-j)}}{(n-j)!}.$$

Assume that the induction hypothesis (3.6) holds for $\beta_{n-j}(k-r)$ if $k-r\geqslant 1$. Applying this formula, we use summation indices $r_2,\ldots$ and $j_2,\ldots$ leaving r_1 and j_1 for the summation in (3.5) with respect to r and j. So, using the assumption in (3.5), we obtain

$$\beta_n(k) = \theta \sum_{r_1=1}^{k-1} \binom{k-1}{r_1-1} \sum_{j_1\leqslant n} \frac{a_{j_1(r_1)}}{j_1} \\ \times \sum_{u=2}^{k-r_1+1} \theta^{u-1} \sum_{r_2+\cdots+r_u=k-r_1} \binom{k-r_1-1}{r_2-1} \cdots \binom{k-r_1-\cdots-r_{u-1}-1}{r_u-1} \\ \times \sum_{j_2+\cdots+j_u\leqslant n-j_1} \frac{a_{j_2(r_2)}\cdots a_{j(r_u)}}{j_2\cdots j_u} \frac{\theta^{(n-j_1-\cdots-j_u)}}{(n-j_1-\cdots-j_u)!} + \theta \sum_{j=1}^{n} \frac{a_{j(k)}}{j} \frac{\theta^{(n-j)}}{(n-j)!}.$$

Interchanging the order of summation, we arrive at

$$\beta_n(k) = \sum_{u=2}^{k} \theta^u \sum_{r_1+\cdots+r_u=k} \binom{k-1}{r_1-1} \cdots \binom{k-r_1-\cdots-r_{u-1}-1}{r_u-1} \\ \times \sum_{j_1+\cdots+j_u\leqslant n} \frac{a_{j_1(r_1)}\cdots a_{j_u(r_u)}}{j_1\cdots j_u} \frac{\theta^{(n-j_1-\cdots-j_u)}}{(n-j_1-\cdots-j_u)!} + \theta \sum_{j=1}^{n} \frac{a_{j(k)}}{j} \frac{\theta^{(n-j)}}{(n-j)!}.$$

The last sum equals the summand corresponding to $u=1$ in the previous sum over u. Joining them together, we obtain (3.6). Further, dividing it by $\theta^{(n)}/n!$, we complete the proof of this lemma. □

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

Corollary 3.3. *Assume that $a_j \in \{0,1\}$ if $j \leqslant n$ and let the asterisk $*$ over a sum stand for the condition $a_j = 1$. Then*

$$\mathbb{E}_n h(\sigma)_{(k)} = \gamma_n(k) = \theta^k \sum_{j_1 \leqslant n}^{*} \frac{1}{j_1} \cdots \sum_{j_k \leqslant n}^{*} \frac{\mathbf{1}\{j_1 + \cdots + j_k \leqslant n\}}{j_k} \psi_n\big(n - (j_1 + \cdots + j_k)\big).$$

Henceforth, the symbol $\ll$ is used as an analogue of $O(\cdot)$ and $a \asymp b$ means that $a \ll b$ and $b \ll a$.

Lemma 3.4. *If $a_j \in \mathbb{Z}_+ \cap [0,m]$ for $j \leqslant n$, then*

$$\begin{aligned}\gamma_n(k) = \sum_{u=1}^{k} \theta^u &\sum_{r_1+\cdots+r_u=k} \binom{k-1}{r_1-1} \cdots \binom{k-r_1-\cdots-r_{u-1}-1}{r_u-1} \\ &\times \sum_{j_1+\cdots+j_u<n} \frac{a_{j_1(r_1)} \cdots a_{j_u(r_u)}}{j_1 \cdots j_u} \left(1 - \frac{j_1+\cdots+j_u}{n}\right)^{\theta-1} \\ &\quad + O_{m,k}\left(\frac{1+\log^k n}{n^{1\wedge\theta}}\right). \qquad (3.7)\end{aligned}$$

Proof. It suffices to deal with the case of $\theta \neq 1$ and n sufficiently large. Set $\Delta_n(k)$ for the difference of $\gamma_n(k)$ in (3.3) and the main term in its approximation (3.7). Using (3.3) and the given bound of a_j, we have

$$\begin{aligned}\Delta_n(k) \ll \sum_{u=1}^{k} C_u(k,m) &\sum_{j_1,\ldots,j_u<n} \frac{\mathbf{1}\{j_1+\cdots+j_u<n\}}{j_1\cdots j_u} \frac{1}{n-(j_1+\cdots+j_u)} \\ &\times \left(1 - \frac{j_1+\cdots+j_u}{n}\right)^{\theta-1} \\ &\quad + n^{1-\theta} \sum_{u=1}^{k} C_u(k,m) \sum_{j_1,\ldots,j_u\leqslant n} \frac{\mathbf{1}\{j_1+\cdots+j_u=n\}}{j_1\cdots j_u}.\end{aligned}$$

Here

$$C_u(k,m) := \sum_{\substack{r_1+\cdots+r_u=k \\ 1\leqslant r_i\leqslant m, i\leqslant u}} \binom{k-1}{r_1-1} \cdots \binom{k-r_1-\cdots-r_{u-1}-1}{r_u-1} \ll_{m,k} 1$$

if $1 \leqslant u \leqslant k$. Using the latter, we see that a typical sum to be estimated is

$$\begin{aligned}&\sum_{j_1<n} \frac{1}{j_1} \cdots \sum_{j_u<n} \frac{\mathbf{1}\{j_1+\cdots+j_u<n\}}{j_u\big(n-(j_1+\cdots+j_u)\big)} \left(1 - \frac{j_1+\cdots+j_u}{n}\right)^{\theta-1} \\ &\quad + n^{1-\theta} \sum_{j_1\leqslant n} \frac{1}{j_1} \cdots \sum_{j_u\leqslant n} \frac{\mathbf{1}\{j_1+\cdots+j_u=n\}}{j_u} =: R_{nu} + r_{nu}, \qquad (3.8)\end{aligned}$$

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

where $1 \leqslant u \leqslant k$. Now, in the sums of the second remainder term, at least one $j_i \geqslant n/u$, $1 \leqslant i \leqslant u$. Hence, estimating one of the u sums trivially, we obtain

$$r_{nu} \leqslant \frac{u\log(u+1)}{n^\theta} \sum_{j_1 \leqslant n} \frac{1}{j_1} \cdots \sum_{j_{u-1} \leqslant n} \frac{\mathbf{1}\{j_1 + \cdots + j_{u-1} \leqslant n - n/u\}}{j_{u-1}}$$
$$\leqslant \frac{u\log(u+1)}{n^\theta} \biggl(\sum_{j \leqslant n} \frac{1}{j}\biggr)^{u-1} \ll_k \frac{\log^{u-1} n}{n^\theta}$$

for $1 \leqslant u \leqslant k$.

For brevity, introduce temporarily the notation $J = j_1 + \cdots + j_u$ and $j = j_{u+1}$. We will apply the mathematical induction for either of the sums in the splitting

$$R_{n,u+1} \ll \sum_{j_1<n} \frac{1}{j_1} \cdots \sum_{j_u<n} \frac{\mathbf{1}\{J<n\}}{j_u} \sum_{j \leqslant (n-J)/2} \frac{1}{j} \frac{1}{(n-J)-j}$$
$$+ \sum_{j_1<n} \frac{1}{j_1} \cdots \sum_{j_u<n} \frac{\mathbf{1}\{J<n\}}{j_u} \sum_{(n-J)/2<j<n-J} \frac{1}{j} \frac{1}{(n-J)-j} \biggl(1 - \frac{J+j}{n}\biggr)^{\theta-1}$$
$$=: R'_{n,u+1} + R''_{n,u+1}.$$

Now,

$$R'_{n1} + R''_{n1} = \sum_{j \leqslant n/2} \frac{1}{j} \frac{1}{n-j} + \sum_{n/2<j<n} \frac{1}{j} \frac{1}{n-j} \biggl(1 - \frac{j}{n}\biggr)^{\theta-1}$$
$$\ll \frac{\log n}{n} + \frac{1}{n^\theta} \sum_{n/2<j<n} (n-j)^{\theta-2} \ll \frac{\log n}{n} + \frac{1}{n^{1\wedge\theta}} \ll \frac{\log n}{n^{1\wedge\theta}}.$$

Assuming that $R'_{nu} \ll_u (\log^u n)/n$, we have

$$R'_{n,u+1} \ll R'_{nu} \log n \ll_u (\log^{u+1} n)/n$$

in either of the cases $\theta < 1$ or $\theta > 1$. Further, if $\theta > 1$, then $(1-(J+j)/n)^{\theta-1} \leqslant 1$. An easy estimation of the innermost sum now implies

$$R''_{n,u+1} \ll R'_{nu} \log n \ll_u (\log^{u+1} n)/n.$$

If $\theta < 1$, then

$$R''_{n,u+1} \ll \sum_{j_1<n} \frac{1}{j_1} \cdots \sum_{j_u<n} \frac{\mathbf{1}\{J<n\}}{j_u(n-J)}$$
$$\times \sum_{(n-J)/2<j<n-J} \biggl(1 - \frac{J+j}{n}\biggr)^{\theta-1} \frac{1}{(n-J)-j}$$
$$\ll \frac{1}{n^{\theta-1}} \sum_{j_1<n} \frac{1}{j_1} \cdots \sum_{j_u<n} \frac{\mathbf{1}\{J<n\}}{j_k(n-J)} \sum_{1 \leqslant s<n} s^{\theta-2}$$
$$\ll_u \frac{1}{n^{\theta-1}} \frac{\log^u n}{n} = \frac{\log^u n}{n^\theta}$$

since the last inner sum is bounded and the remaining iterated sum was estimated before.

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

Collecting all the estimates, we return to (3.8) and conclude that

$$R_{nu} + r_{nu} \ll_u (\log^u n) n^{-(\theta \wedge 1)}$$

for sufficiently large n. Inserting this into the expression $\Delta_n(k)$, we finish the proof of the lemma. □

In a similar way, we can handle the growth of the factorial moments $\upsilon_n(l)$ corresponding to the function $w(\sigma, J_n)$ as $l \to \infty$.

Lemma 3.5. *Let $J_n \subset \{1, \ldots, n\}$ be arbitrary. If $\theta \geqslant 1$, then $\upsilon_n(l) \leqslant \upsilon_n(1)^l$ for every $n, l \in \mathbb{N}$. If $\theta < 1$, then there exists a positive constant C depending on θ only such that*

$$\upsilon_n(l) \leqslant C^l \big(\upsilon_n(1) + 1\big)^l$$

for every $l \in \mathbb{N}$.

Proof. The proof of the first assertion is straightforward. In the case $\theta < 1$, we apply the induction. Examine the innermost sum on the right-hand side of the inequality

$$\upsilon_n(l+1) \leqslant \theta^l \sum_{j_1, \ldots, j_l \in J_n} \frac{\mathbf{1}\{S < n\}}{j_1 \cdots j_l} \sum_{j \in J_n} \frac{\mathbf{1}\{j < n - S\}}{j} \left(1 - \frac{S+j}{n}\right)^{\theta - 1}, \tag{3.9}$$

where *pro tem* $S := j_1 + \cdots + j_l$ and $j := j_{l+1}$. The summands over $j \leqslant (n-S)/2$ contribute no more than

$$2^{1-\theta}(1 - S/n)^{\theta-1}\upsilon_n(1)$$

and

$$\begin{aligned}
&\sum_{j \in J_n} \frac{\mathbf{1}\{(n-S)/2 < j < n - S\}}{j} \left(1 - \frac{S+j}{n}\right)^{\theta-1} \\
&\qquad \leqslant \frac{2}{n^{\theta-1}(n-S)} \sum_{j \in J_n} \mathbf{1}\{(n-S)/2 < j < n-S\}(n-S-j)^{\theta-1} \\
&\qquad \leqslant \frac{2}{n^{\theta-1}(n-S)} \sum_{k < (n-S)/2} k^{\theta-1} \leqslant C_1 \left(1 - \frac{S}{n}\right)^{\theta-1}.
\end{aligned}$$

The last two estimates and (3.9) yield

$$\upsilon_n(l+1) \leqslant (2 \vee C_1)\big(\upsilon_n(1) + 1\big)\upsilon_n(l).$$

Consequently, the desired assertion holds with $C = 2 \vee C_1 := \max\{2, C_1\}$. □

Let us introduce the concentration function

$$Q_n(u) = \sup_{x \in \mathbb{R}} \nu_n(|h(\sigma) - x| < u), \quad u \geqslant 0,\ x \in \mathbb{R},$$

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

and

$$D_n(u;\lambda) = \sum_{j\leqslant n} \frac{u^2 \wedge (a_j - \lambda j)^2}{j}, \quad D_n(u) = \min_{\lambda\in\mathbb{R}} D_n(u;\lambda).$$

Lemma 3.6. *We have*

$$Q_n(u) \ll u D_n(u)^{-1/2} \tag{3.10}$$

for every $\theta > 0$.

Proof. See [15]. □

The last lemma is used to obtain lower estimates of the required frequencies below. Let $J \subset \{j : j \leqslant n\}$ be an arbitrary non-empty set, possibly depending on n, and $\overline{J} = \{j : j \leqslant n\} \setminus J$.

Lemma 3.7. *Let* $\theta \geqslant 1$, $K > 0$, *and* J *be such that*

$$\sum_{j\in J} \frac{1}{j} \leqslant K < \infty. \tag{3.11}$$

Denote

$$\mu_n(K) = \inf_J \nu_n\big(k_j(\sigma) = 0 \;\forall\, j \in J\big),$$

where the infimum is taken over J *satisfying* (3.11). *For a sufficiently large* $n_0(K)$, *there exists a positive constant* $c(K)$, *depending at most on* θ *and* K, *such that* $\mu_n(K) \geqslant c(K)$ *if* $n \geqslant n_0(K)$.

Moreover, for any $I \subset J \cap [1, n - n_0(K)]$ *and*

$$\widetilde{S}_n := \bigcup_{j\in I} S_n^j := \bigcup_{j\in I} \{\sigma \in \mathbf{S}_n : k_j(\sigma) = 1,\; k_i(\sigma) = 0 \;\forall i \in J \setminus \{j\}\},$$

we have that

$$\nu_n(\widetilde{S}_n) \geqslant c(K) \sum_{j\in I} \frac{1}{j} \psi_n(n-j) \gg \sum_{j\in I} \frac{1}{j} \left(1 - \frac{j}{n}\right)^{\theta-1} \tag{3.12}$$

provided that $n \geqslant 2n_0(K)$.

Proof. The first claim is Corollary 1.3 of Theorem 1.2 (see [16]). The second claim is proved in [17]. □

The next observation presents the possibility of applying the previous lemma of the sieve type.

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

Lemma 3.8. *Assume that $h_n(\sigma)$ is defined via $a_j = a_{nj} \in \mathbb{Z}$ and $T_n(x) \Rightarrow F(x)$. Then*

$$\sum_{j\leqslant n} \frac{\mathbf{1}\{a_j \neq 0\}}{j} \ll 1 \tag{3.13}$$

provided that $\theta \geqslant 1$ or $\theta < 1$ and

$$\sum_{j\leqslant n} \frac{\mathbf{1}\{|a_j| \geqslant K\}}{j} \leqslant K_1 \tag{3.14}$$

for some positive constants K and K_1. Now, the constant in (3.13) *depends also on F, K, and K_1.*

Proof. Since the limit law has an atom, we obtain a lower estimate of the concentration function $Q_n(u) \geqslant c > 0$ for every $u > 0$ if n is sufficiently large. Now applying Lemma 3.6, we have $D_n(u,\lambda) \ll c^{-2}u^2$ for some $\lambda = \lambda_n \in \mathbb{R}$. This, if $u \to 0$, yields the estimate

$$\sum_{j\leqslant n} \frac{\mathbf{1}\{a_j \neq \lambda j\}}{j} \ll 1. \tag{3.15}$$

In fact, $\lambda \in \mathbb{Z}$. Indeed, if $\|\cdot\|$ denotes the distance to the nearest integer, we have

$$1 \gg D_n(1,\lambda) \geqslant \sum_{j\leqslant n} \frac{\|\lambda j\|^2}{j}$$

and, further, $\lambda =: \tilde{\lambda} + \delta$, where $\tilde{\lambda} = \tilde{\lambda}_n \in \mathbb{Z}$ and $\delta = \delta_n = O(n^{-1})$. Now the inequality $(x+y)^2 \leqslant 2x^2 + 2y^2$, $x, y \in \mathbb{R}$, implies

$$D_n(u,\tilde{\lambda}) \leqslant 2D_n(u,\lambda) + 2\sum_{j\leqslant n} \frac{u^2 \wedge (\delta j)^2}{j} \ll u^2.$$

Consequently, we may proceed with $\lambda \in \mathbb{Z}$.

If $\theta \geqslant 1$ then, denoting $J := \{j \leqslant n : a_j \neq \lambda j\}$ and

$$h_n(\sigma) = \lambda\ell(\bar{k}(\sigma)) + \sum_{j\in J}(a_j - \lambda j)k_j(\sigma) =: \lambda n + \tilde{h}_n(\sigma),$$

by Lemma 3.7, for sufficiently large n,

$$\nu_n(h_n(\sigma) = \lambda n) = \nu_n(\tilde{h}_n(\sigma) = 0) \geqslant \nu_n\big(k_j(\sigma) = 0 \;\forall\, j \in J\big) \geqslant c_1 > 0.$$

Hence, if $\lambda n \to \infty$ for some subsequence of $n \to \infty$, at least c_1 of the probability distribution mass of $h_n(\sigma)$ disappears at infinity. This contradicts the assumption of the theorem. Hence $\lambda = \lambda_n \ll n^{-1}$, and thus $\lambda = 0$ eventually. Now, the estimate $D_n(1,0) \ll 1$ contains (3.13).

Assume that $\theta < 1$ is arbitrary and $|a_j| \leqslant K$ for the most part of $j \leqslant n$ in the sense of (3.14). Now, manipulating the latter and the estimate (3.15), we obtain the bound $\lambda \ll K/n$, which implies that $\lambda = 0$ eventually.

The lemma is proved. □

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

Remark. If all $a_j \in \mathbb{Z}$, $j \geqslant 1$, do not depend on n, then the additive function possesses a limit distribution, *i.e.*, $\nu_n(h(\sigma) < x) \Rightarrow F(x)$, if and only if the series

$$\sum_{j\geqslant 1} \frac{\mathbf{1}\{a_j \neq 0\}}{j}$$

converges. The fact is well known [2, Theorem 8.25], since the three series in an analogue of Kolmogorov's theorem for the integer-valued functions reduce to this one. The last lemma gives a very short proof of the necessity.

4. Proof of Theorem 1.1

First, we observe that, if the limit random variable is concentrated on $\mathbb{Z}_+$, we may confine ourselves to non-negative additive functions. Set a^+ for the non-negative part of $a \in \mathbb{R}$. Let $h_n^{(+)}(\sigma)$ be the additive function defined as in (1.4) via a_j^+, where $j \leqslant n$, and

$$T_n^{(+)}(x) := \nu_n\big(h_n^{(+)}(\sigma) < x\big).$$

Proposition 4.1. *If $\theta \geqslant 1$, then the convergence $T_n(x) \Rightarrow F(x)$ is equivalent to $T_n^{(+)}(x) \Rightarrow F(x)$ together with condition* (1.7). □

Proof. Assume that $T_n(x) \Rightarrow F_Y(x)$. Then, by Lemma 3.8, condition (3.13) holds. Set

$$I := \{j \leqslant n - n_0 : \ a_j \leqslant -1\} \subset J := \{j \leqslant n, \ a_j \neq 0\},$$

where $n \geqslant n_0$ and $n_0 \in \mathbb{N}$, depending on F, is sufficiently large. Define, as in Lemma 3.7,

$$S_n^j = \{\sigma \in \mathbf{S}_n : \ k_j(\sigma) = 1, k_i(\sigma) = 0 \quad \forall i \in J \setminus \{j\}\}$$

and observe that $h_n(\sigma) = a_j \leqslant -1$ for all $\sigma \in S_n^j$ with $j \in I$. From Lemma 3.7 and the above assumption we have

$$o(1) = \nu_n\big(h_n(\sigma) \leqslant -1\big) \geqslant \nu_n\bigg(\bigcup_{j\in I} S_n^j\bigg) \gg \sum_{j\in I} \frac{1}{j}\bigg(1 - \frac{j}{n}\bigg)^{\theta-1}.$$

The sum can also be extended over $n - n_0 \leqslant j \leqslant n$. This proves the necessity of (1.7).

Further, having (1.7), we claim that $T_n^{(+)}(x) \Rightarrow F_Y(x)$ is equivalent to $T_n(x) \Rightarrow F_Y(x)$. Indeed,

$$\begin{aligned} \nu_n\big(h_n(\sigma) \neq h_n^{(+)}(\sigma)\big) &\leqslant \sum_{\substack{j\leqslant n\\ a_j\leqslant -1}} \nu_n\big(k_j(\sigma) \geqslant 1\big) \\ &\leqslant \sum_{\substack{j\leqslant n\\ a_j\leqslant -1}} \mathbb{E}_n k_j(\sigma) = \sum_{\substack{j\leqslant n\\ a_j\leqslant -1}} \frac{\theta}{j}\psi_n(n-j) = o(1). \end{aligned}$$

We have used a particular case of formula (3.2).

Proposition 4.1 is proved. □

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

Proof of Theorem 1.1. We have from Lemma 3.8 that the sum (3.13) is bounded by a constant C_F. Further, by Proposition 4.1, we see that condition (1.7) is satisfied, and we may assume that $h_n(\sigma) \in \mathbb{Z}_+$. In what follows, the constants involved in the estimates can depend on F.

For an integer $m \geqslant 1$, we set $a_j(m) = a_j$ if $a_j \leqslant m$ and $a_j(m) = m$ otherwise, and introduce the truncated functions

$$h_n(\sigma; m) := \sum_{j \leqslant n} a_j(m) k_j(\sigma).$$

By Lemmas 3.2 and 3.4, we have

$$\mathbb{E}_n h_n(\sigma, m)_{(l)} = \Upsilon_n(l, m; h) + o_{m,l}(1).$$

So, the goal lies in proving that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{E}_n h_n(\sigma, m)_{(l)} = \lim_{m\to\infty} \liminf_{n\to\infty} \mathbb{E}_n h_n(\sigma, m)_{(l)} = \mathbb{E} Y_{(l)} =: \Upsilon(l) \tag{4.1}$$

for each natural number $l \leqslant \alpha - 1 - \varepsilon$.

Set $J_n := \{j \leqslant n : \ a_j \neq 0\}$, and then

$$\begin{aligned} \mathbb{E}_n h_n(\sigma, m)_{(l)} &\leqslant m^l \mathbb{E}_n w(\sigma, J_n)_{(l)} = m^l(\upsilon_n(l) + o_l(1)) \\ &\leqslant m^l(\upsilon_n(1)^l + o_l(1)) \ll_{l,F} m^l \end{aligned} \tag{4.2}$$

by virtue of bound (3.13), where the hidden constant depends on l and the limit distribution F.

We now split

$$\mathbb{E}_n h_n(\sigma, m)_{(l)} = E_n(l, m)' + E_n''(l, m) + E_n'''(l, m), \tag{4.3}$$

where

$$\begin{aligned} E_n'(l, m) &= \sum_{b=l}^{m-1} b_{(l)} \nu_n\big(h_n(\sigma) = b\big), \\ E_n''(l, m) &= \sum_{b=m}^{M} b_{(l)} \nu_n\big(h_n(\sigma, m) = b\big), \\ E_n'''(l, m) &= \sum_{b>M} b_{(l)} \nu_n\big(h_n(\sigma, m) = b\big), \end{aligned}$$

and $M = M(m) > m$ is a natural number to be chosen later.

If $l \leqslant \alpha$, then

$$\lim_{n\to\infty} E_n'(l, m) = \sum_{b=l}^{m-1} b_{(l)} P(Y = b) = \Upsilon(l) - \sum_{b \geqslant m} b_{(l)} P(Y = b)$$

for each fixed m and, by virtue of $\mathbb{E} Y^\alpha < \infty$,

$$\lim_{m\to\infty} \sum_{b \geqslant m} b_{(l)} P(Y = b) \leqslant \lim_{m\to\infty} \sum_{b \geqslant m} b^l P(Y = b) = 0.$$

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

In other words,

$$\lim_{m\to\infty}\lim_{n\to\infty}\mathbb{E}_n'(l,m)=\Upsilon(l) \tag{4.4}$$

for each $l \leqslant \alpha$.

Similarly, if $l \leqslant \alpha-1-\varepsilon$,

$$\begin{aligned} E_n''(l,m) &\leqslant \sum_{b=m}^{M} b^l \nu_n\big(h_n(\sigma)\geqslant b\big) = \sum_{b=m}^{M} b^l P(Y\geqslant b)+o_m(1)\\ &\leqslant \mathbb{E}Y^{\alpha}\sum_{b=m}^{\infty}\frac{1}{b^{1+\varepsilon}}+o_m(1)=\rho_n(m). \end{aligned} \tag{4.5}$$

Finally, we have from (4.2)

$$\begin{aligned} E_n'''(l,m) &= \frac{1}{\theta^{(n)}}\sum_{\sigma\in\mathbf{S}_n}\theta^{w(\sigma)}\mathbf{1}\{h_n(\sigma,m)>M\}h_n(\sigma,m)_{(l)}\cdot\frac{h_n(\sigma,m)-l}{h_n(\sigma,m)-l}\\ &\leqslant \frac{1}{M-l}\mathbb{E}_n h_n(\sigma,m)_{(l+1)} \ll_l \frac{m^{l+1}}{M-l}\ll_l \frac{1}{m} \end{aligned}$$

for the choice $M=m^{l+2}$ provided that $m>2l$. Collecting (4.4), (4.5), and the last estimate, from the splitting (4.3), we obtain claim (4.1).

The theorem is proved. □

5. Proof of Theorem 1.2

As we have seen in the proof of Proposition 4.1, condition (1.7) allows us to deal with non-negative functions only. By the condition of the theorem and Lemma 3.4,

$$\mathbb{E}_n h_n(\sigma,m)_{(l)}=\Upsilon_n(l,m;h)+o_{m,l}(1)=\Upsilon(l)+\rho_n(m).$$

Let $L\in\mathbb{N}$ be a fixed number, and examine the expansion of the characteristic function

$$\mathbb{E}_n \mathrm{e}^{\mathrm{i}th_n(\sigma,m)}=\sum_{l=0}^{L}\frac{\mathbb{E}_n h_n(\sigma,m)_{(l)}}{l!}(\mathrm{e}^{\mathrm{i}t}-1)^l+O\bigg(\frac{\mathbb{E}_n h_n(\sigma,m)_{(l)}}{(L+1)!}|\mathrm{e}^{\mathrm{i}t}-1|^{L+1}\bigg),$$

where $t\in\mathbb{R}$ and the constant in $O(\cdot)$ is absolute. We further have

$$\mathbb{E}_n \mathrm{e}^{\mathrm{i}th_n(\sigma,m)}=\sum_{l=1}^{L}\frac{\Upsilon(l)}{l!}(\mathrm{e}^{\mathrm{i}t}-1)^l+O\bigg(\frac{2^L\Upsilon(L+1)}{(L+1)!}\bigg)+\rho_m(n)$$

uniformly in $t\in\mathbb{R}$. In other words,

$$\lim_{m\to\infty}\limsup_{n\to\infty}\bigg|\mathbb{E}_n \mathrm{e}^{\mathrm{i}th(\sigma;m)}-\sum_{l=0}^{L}\frac{\Upsilon(l)}{l!}(\mathrm{e}^{\mathrm{i}t}-1)^l\bigg|\ll\frac{2^L\Upsilon(L+1)}{(L+1)!}$$

for every $L\geqslant 1$.

By the given conditions,

$$\limsup_{n\to\infty}\sum_{\substack{j\leqslant n\\ a_j>m}}\frac{\theta}{j}\Big(1-\frac{j}{n}\Big)^{\theta-1}\leqslant\frac{1}{m}\lim_{r\to\infty}\limsup_{n\to\infty}\sum_{j\leqslant n}\frac{\theta a_j(r)}{j}\Big(1-\frac{j}{n}\Big)^{\theta-1}=\frac{\Upsilon(1)}{m}.$$

Hence

$$\mathbb{E}_n|e^{ith(\sigma;m)} - e^{ith_n(\sigma)}| \leqslant \nu_n(h(\sigma) \neq h(\sigma;m)) \leqslant \sum_{\substack{j\leqslant n\\ a_j>m}} \nu_n(k_j(\sigma) \geqslant 1)$$

$$\leqslant \sum_{\substack{j\leqslant n\\ a_j>m}} \mathbb{E}_n k_j(\sigma) = \sum_{\substack{j\leqslant n\\ a_j>m}} \frac{\theta}{j}\psi_n(n-j) = \rho_m(n).$$

The last two approximations imply

$$\lim_{m\to\infty} \limsup_{n\to\infty} \left|\mathbb{E}_n e^{ith_n(\sigma)} - \sum_{l=0}^{L} \frac{\Upsilon(l)(e^{it}-1)^l}{l!}\right| \ll \frac{2^L \Upsilon(L+1)}{(L+1)!}.$$

It remains to take $L \to \infty$.

The theorem is proved. □

Proof of Corollary 1.4. In the sufficiency part, it suffices to rewrite the factorial moments as follows:

$$\Upsilon_n(l,m;h) = \theta \sum_{n/2<j<n} \frac{a_{j(l)}(m)}{j}\left(1-\frac{j}{n}\right)^{\theta-1}$$

$$= \sum_{k=1}^{m} k_{(l)} \sum_{n/2<j<n} \frac{\theta \mathbf{1}\{a_j = k\}}{j}\left(1-\frac{j}{n}\right)^{\theta-1}.$$

Here we have to check if the inner sums can approach the Poisson probabilities. Since their sum over $k \geqslant 0$ tends to $t_\theta(1) \geqslant 1 - e^{-\mu}$, this is possible. We may continue and get

$$\Upsilon_n(l,m;h) = \sum_{k=1}^{m} k_{(l)}\left(e^{-\mu}\frac{\mu^k}{k!} + o_k(1)\right) = e^{-\mu}\sum_{k=1}^{m} k_{(l)}\frac{\mu^k}{k!} + o_m(1).$$

Hence

$$\Upsilon_n(l,m;h) - \mu^l = \rho_n(m),$$

as desired.

To prove the necessity, we demonstrate another path. Recall that the function

$$\varphi_n(z) = (\theta^{(n)}/n!)\mathbb{E}_n z^{h(\sigma)}, \quad |z| \leqslant 1,$$

satisfies (3.4). If $a_j = 0$ for $j \leqslant n/2$, then $\varphi_{n-j}(z) = \theta^{(n-j)}/(n-j)!$ and, consequently, we obtain

$$e^{\mu(z-1)} + o(1) = \mathbb{E}_n z^{h_n(\sigma)} = 1 + \theta \sum_{n/2<j\leqslant n} \frac{z^{a_j}-1}{j}\psi_n(n-j)$$

$$= 1 + \theta \sum_{k\geqslant 1}(z^k - 1) \sum_{n/2<j\leqslant n} \frac{\mathbf{1}\{a_j = k\}}{j}\left(1-\frac{j}{n}\right)^{\theta-1} + o(1)$$

uniformly in $|z| \leqslant 1$. Applying Cauchy's formula on the circle $|z| = 1$, we complete the proof of the corollary. □

https://doi.org/10.1017/S0963548314000376 Published online by Cambridge University Press

6. The cases with bounded a_j

Proof of Theorem 1.5. Condition (1.13) allows us to explore the case with $0 \leqslant a_j \leqslant K$, $j \leqslant n$, only.

Sufficiency. If $J_n := \{j \leqslant n : a_j \neq 0\}$, then

$$\mathbb{E}_n h_n(\sigma)_{(l)} \leqslant K^l \mathbb{E}_n w(\sigma, J_n)_{(l)} \leqslant K^l C^l \left(\upsilon_n(1) + 1\right)^l \leqslant C_2^l$$

by Lemma 3.5 for every $l \geqslant 1$. Further, it suffices to apply Theorem 1.2.

Necessity. By Lemma 3.8, we obtain from the convergence $T_n(x) \Rightarrow F_Y(x)$ the bound (3.13) which, in turn, yields $\Upsilon_n(1) \ll K$. Indeed, to check this, it suffices to observe that the summands over $n/2 < j < n$ contribute only a bounded quantity. As we have seen, in the sufficiency part,

$$\sup_{n \geqslant 1} \mathbb{E}_n h_n(\sigma)_{(l)} \ll K^l$$

for every fixed $l \geqslant 1$. Now, the weak convergence of distributions also implies the convergence of moments. Namely, we have $\Upsilon_n(l) = \mathbb{E} Y_{(l)} + o_l(1)$ where $l \geqslant 1$.

The theorem is proved. □

Proof of Corollary 1.7. Only *necessity* requires some argument. By Theorem 1.3, convergence of distributions implies the relations $\upsilon_n(l) \to \mu^l$ where $l \geqslant 1$. Omitting non-negative sums in the difference below, we obtain

$$\begin{aligned} o(1) &= \upsilon_n(1)^l - \upsilon_n(l) \\ &\geqslant \theta^l \sum_{j_1, \dots, j_l \leqslant n}^{*} \frac{\mathbf{1}\{j_1 + \cdots + j_l > n\}}{j_1 \cdots j_l} \left(1 - \frac{j_1}{n}\right)^{\theta - 1} \cdots \left(1 - \frac{j_l}{n}\right)^{\theta - 1} \\ &\geqslant \left(\theta^l \sum_{n/l < j \leqslant n}^{*} \frac{1}{j} \left(1 - \frac{j}{n}\right)^{\theta - 1}\right)^l \end{aligned}$$

for every $l \geqslant 1$. This yields the second of the conditions in (1.16). Using the latter and checking that the factor $(1 - j/n)^{\theta - 1} = 1 + o(1)$ uniformly in $j \leqslant r = o(n)$, we can rewrite the relation $\upsilon_n(1) = \mu + o(1)$ as is given in the first of relations in (1.16). □

Proof of Corollary 1.8. *Sufficiency*. Since (1.17) and (1.18) imply the sufficient condition (1.15) in Corollary 1.7, we are done.

Necessity. In the case discussed, the Lth factorial moment $\mathbb{E}_n h_n(\sigma)_{(L)}$ converges to zero. Hence the relevant formula yields

$$\begin{aligned} o(1) = \mathbb{E}_n w(\sigma, J_n)_{(L)} &\geqslant \theta^l \sum_{j_1, \dots j_L \leqslant n/L}^{*} \frac{1}{j_1 \cdots j_L} \psi_n\big(n - (j_1 + \cdots + j_L)\big) \\ &\geqslant \left(\theta \sum_{j \leqslant n/L}^{*} \frac{1}{j} \psi_n(n - j)\right)^L \end{aligned}$$

for $\theta \leqslant 1$. This is equivalent to (1.18). It also allows us to reduce the problem to the sequence of additive functions with $a_j = 0$ if $j \leqslant n/L$. Then the necessary condition (1.15) reduces to (1.17).

The corollary is proved. □

Concluding remark. Most of the results presented above can be obtained for the generalized Ewens probability measure

$$\nu_{n,\Theta}(A) := \frac{1}{\Theta_n} \sum_{\sigma \in A} \theta_j^{w(\sigma)},$$

where $0 < c_3 \leqslant \theta_j q^{-j} \leqslant C_3 < \infty$ if $j \leqslant n$, $q \geqslant 1$ is a fixed constant, and Θ_n is an appropriate normalization. In some cases, unfortunately, we have to assume that $c_3 = 1$. An analytic technique to deal with the value distribution of mappings defined on $\mathbf{S}_n$ with respect to $\nu_{n,\Theta}$ was proposed by the second author [21]. Later it was extended (see, for example, [26] and [38]) for $\theta_j q^{-j}$ satisfying some averaged conditions. The asymptotic distributions under the generalized Ewens measure of $h_n(\sigma)$ were treated in [26]. The second author's paper [28] provides an approximation in the total variation distance of the truncated cycle vector by an appropriate vector with independent coordinates which is a basic tool for a probabilistic approach. The latter was applied in [8] to prove a functional limit theorem.

The recent papers [3], [7], [9] discuss cases with different behaviour of θ_j, *e.g.*, $\theta_j = \mathrm{e}^{j^\gamma}$, $j \leqslant n$, where $0 < \gamma < 1$. Hopefully, the described method of factorial moments will be of use in these cases too.

Acknowledgements

The authors thank the referees for valuable and benevolent comments.

References

[1] Arous, G. B. and Dang, K. (2011) On fluctuations of eigenvalues of random permutation matrices. `arXiv:1106.2108v1`

[2] Arratia, R. Barbour, A. D. and Tavaré, S. (2003) *Logarithmic Combinatorial Structures: A Probabilistic Approach*, EMS, Zürich.

[3] Barbour, A. D. and Granovsky, B. L. (2005) Random combinatorial structures: The convergent case. *J. Combin. Theory Ser. A* **109** 203–220.

[4] Betz, V. and Ueltschi, D. (2009) Spatial random permutations and infinite cycles. *Commun. Math. Phys.* **285** 469–501.

[5] Betz, V. and Ueltschi, D. (2011) Spatial random permutations with small cycle weights. *Probab. Theory. Rel. Fields* **149** 191–222.

[6] Betz, V. and Ueltschi, D. (2011) Spatial random permutations and Poisson–Dirichlet law of cycle lengths. *Electron. J. Probab.* **16** 1173–1192.

[7] Betz, V., Ueltschi, D. and Velenik, Y. (2011) Random permutations with cycle weights. *Ann. Appl. Probab.* **21** 312–331.

[8] Bogdanas, K. and Manstavičius, E. (2012) Stochastic processes on weakly logarithmic assemblies. In *Analytic and Probabilistic Methods in Number Theory* **5** (A. Laurinčikas *et al.*, eds), *Kubilius Memorial Volume*, TEV, Vilnius, pp. 69–80.

[9] Ercolani, N. M. and Ueltschi, D. (2014) Cycle structure of random permutations with cycle weights. *Random Struct. Alg.* **44** 109–133.

[10] Erdős, P. and Turán, P. (1965) On some problems of a statistical group theory I. *Z. Wahrsch. Verw. Gebiete* **4** 175–186.

[11] Flajolet, P. and Sedgewick, R. (2008) *Analytic Combinatorics*, Cambridge University Press.

[12] Hambly, B., Keevash, P., O'Connell, N. and Stark, D. (2000) The characteristic polynomial of a random permutation matrix. *Stoch. Process. Appl.* **90** 335–346.

[13] Hughes, C., Najnudel, J., Nikeghball, A. and Zeindler, D. (2013) Random permutation matrices under the genereralized Ewens measure. *Ann. Appl. Probab.* **23** 987–1024.

[14] Kargina, T. (2007) Additive functions on permutations and the Ewens probability. *Šiauliai Math. Semin.* **10** 33–41.

[15] Kargina, T. (2009) Asymptotic distributions of the number of restricted cycles in a random permutation. *Lietuvos matem. rink. Proc. LMS* **50** 420–425.

[16] Kargina, T. and Manstavičius, E. (2012) Multiplicative functions on Z_+^n and the Ewens Sampling Formula. *RIMS Kôkyûroku Bessatsu* **B34** 137–151.

[17] Kargina, T. and Manstavičius, E. (2013) The law of large numbers with respect to Ewens probability. *Ann. Univ. Sci. Budapest., Sect. Comp.* **39** 227–238.

[18] Lugo, M. (2009) Profiles of permutations. *Electron. J. Combin.* **16** 1–20.

[19] Lugo, M. (2009) The number of cycles of specified normalized length in permutations. `arXiv:0909.2909vI`

[20] Manstavičius, E. (1996) Additive and multiplicative functions on random permutations. *Lith. Math. J.* **36** 400–408.

[21] Manstavičius, E. (2002) Mappings on decomposable combinatorial structures: Analytic approach. *Combin. Probab. Comput.* **11** 61–78.

[22] Manstavičius, E. (2002) Functional limit theorem for sequences of mappings on the symmetric group. In *Analytic and Probabilistic Methods in Number Theory* **3** (A. Laurinčikas *et al.*, eds), TEV, Vilnius, pp. 175–187.

[23] Manstavičius, E. (2005) The Poisson distribution for the linear statistics on random permutations. *Lith. Math. J.* **45** 434–446.

[24] Manstavičius, E. (2005) Discrete limit laws for additive functions on the symmetric group. *Acta Math. Univ. Ostraviensis* **13** 47–55.

[25] Manstavičius, E. (2008) Asymptotic value distribution of additive function defined on the symmetric group. *Ramanujan J.* **17** 259–280.

[26] Manstavičius, E. (2009) An analytic method in probabilistic combinatorics. *Osaka J. Math.* **46** 273–290.

[27] Manstavičius, E. (2011) A limit theorem for additive functions defined on the symmetric group. *Lith. Math. J.* **51** 211–237.

[28] Manstavičius, E. (2012) On total variation approximations for random assemblies. In *23rd International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms: AofA'12*, DMTCS Proc., pp. 97–108.

[29] Šiaulys, J. (1996) Convergence to the Poisson law II: Unbounded strongly additive functions. *Lith. Math. J.* **36** 393–404.

[30] Šiaulys, J. (1998) Convergence to the Poisson law III: Method of moments. *Lith. Math. J.* **38** 374–390.

[31] Šiaulys, J. (2000) Factorial moments of distributions of additive functions. *Lith. Math. J.* **40** 389–508.

[32] Šiaulys, J. and Stepanauskas, G. (2008) Some limit laws for strongly additive prime indicators. *Šiauliai Math. Semin.* **3** 235–246.

[33] Šiaulys, J. and Stepanauskas, G. (2011) Binomial limit law for additive prime indicators. *Lith. Math. J.* **51** 562–572.

[34] Wieand, K. L. (2000) Eigenvalue distributions of random permutation matrices. *Ann. Probab.* **28** 1563–1587.
[35] Wieand, K. L. (2003) Permutation matrices, wreath products, and the distribution of eigenvalues. *J. Theoret. Probab.* **16** 599–623.
[36] Zacharovas, V. (2002) The convergence rate to the normal law of a certain variable defined on random polynomials. *Lith. Math. J.* **42** 88–107.
[37] Zacharovas, V. (2004) Distribution of the logarithm of the order of a random permutation. *Lith. Math. J.* **44** 296–327.
[38] Zacharovas, V. (2011) Voronoi summation formulae and multiplicative functions on permutations. *Ramanujan J.* **24** 289–329.
[39] Zeindler, D. (2010) Permutation matrices and the moments of their characteristic polynomial. *Electron. J. Probab.* **15** 1092–1118.