Discrete Mathematics 311 (2011) 463–477

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Strong convergence on weakly logarithmic combinatorial assemblies

Eugenijus Manstavičius *

Department of Mathematics and Informatics, Vilnius University, Naugarduko Str. 24, LT-03225 Vilnius, Lithuania
Institute of Mathematics and Informatics, Akademijos 4, LT-08663 Vilnius, Lithuania

ARTICLE INFO

Article history:
Received 18 January 2009
Received in revised form 9 December 2010
Accepted 10 December 2010

Keywords:
Random combinatorial structure
Total variation distance
Almost sure convergence
Iterated logarithm law
Additive function

ABSTRACT

We deal with the random combinatorial structures called assemblies. Instead of the traditional logarithmic condition which assures asymptotic regularity of the number of components of a given order, we assume only lower and upper bounds of this number. Using the author's analytic approach, we generalize the independent process approximation in the total variation distance of the component structure of an assembly. To evaluate the influence of strongly dependent large components, we obtain estimates of the appropriate conditional probabilities by unconditioned ones. The estimates are applied to examine additive functions defined on a new class of structures, called weakly logarithmic. Some analogs of Major's and Feller's theorems which concern almost sure behavior of sums of independent random variables are proved.

1. Introduction

We are concerned with the value distribution problems of mappings defined on *Abelian partitional complexes* [6] or *assemblies* [2] and begin with the definitions which can be found in the latter systematic study.

Let σ be a set of $n \geq 1$ points, partitioned into subsets so that there are $k_j(\sigma) > 0$ subsets of size j, $1 \leq j \leq n$. Set $\bar{k}(\sigma) := \big(k_1(\sigma), \dots, k_n(\sigma)\big)$. If we define $\ell(\bar{s}) := 1s_1 + \cdots + ns_n$ for $\bar{s} = (s_1, \dots, s_n) \in \mathbb{Z}_+^n$, then $\ell\big(\bar{k}(\sigma)\big) = n$. Assume that in each subset of size j, $1 \leq j \leq n$, by some rule one of m_j, $0 < m_j < \infty$, possible *structures* can be chosen. For instance, the latter might be a *cycle*, then $m_j = (j-1)!$. A subset with a structure is a *component* of σ, and the set σ itself is called an assembly [2]. In the mentioned instance, σ was just a *permutation*. Using all possible partitions of σ and the same rule to define a structure in a component, we get the class $\mathcal{A}_n$ of assemblies of size n. Let $\mathcal{A}_0$ be comprised of the empty set. The union

$$\mathcal{A}_0 \cup \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_n \cup \cdots$$

forms the whole class of assemblies. Its basic parameters appear in the conditions posed on the sequence m_j, $j \geq 1$.

There are

$$n! \prod_{j=1}^{n} \left(\frac{1}{j!}\right)^{s_j} \frac{1}{s_j!}$$

ways to partition an n-set into subsets, so that $\bar{k}(\sigma) = \bar{s}$ if $\ell(\bar{s}) = n$ and $\bar{s} \in \mathbb{Z}_+^n$. Hence, there are

$$Q_n(\bar{s}) := n! \prod_{j=1}^{n} \left(\frac{m_j}{j!}\right)^{s_j} \frac{1}{s_j!}$$

* Corresponding address: Department of Mathematics and Informatics, Vilnius University, Naugarduko Str. 24, LT-03225 Vilnius, Lithuania.
E-mail address: eugenijus.manstavicius@mif.vu.lt.

0012-365X/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.disc.2010.12.014

assemblies with the component vector $\bar{k}(\sigma)=\bar{s}$, and the total number of them in the class $\mathcal{A}_n$ equals

$$|\mathcal{A}_n|=\sum_{\ell(\bar{s})=n} Q_n(\bar{s}).$$

On the class $\mathcal{A}_n$, one can define the uniform probability measure denoted by

$$\nu_n(\ldots)=|\mathcal{A}_n|^{-1}|\{\sigma\in\mathcal{A}_n:\ \ldots\}|.$$

From now on, $\sigma\in\mathcal{A}_n$ is an elementary event. Following the tradition of probabilistic number theory and in contrast to [2], we prefer to write $f(\sigma)$ instead of f for an arbitrary random variable (r.v.) defined on $\mathcal{A}_n$. The component vector $\bar{k}(\sigma)$ has the following distribution:

$$\nu_n(\bar{k}(\sigma)=\bar{s})=\mathbf{1}\{\ell(\bar{s})=n\}\frac{n!}{|\mathcal{A}_n|}\prod_{j=1}^{n}\frac{1}{s_j!}\left(\frac{m_j}{j!}\right)^{s_j},$$

where $\bar{s}=(s_1,\ldots,s_n)\in\mathbb{Z}_+^n$. This leads to the *Conditioning Relation* (see [2, page 48])

$$\nu_n(\bar{k}(\sigma)=\bar{s})=P\big(\bar{\xi}=\bar{s}|\ell(\bar{\xi})=n\big),\tag{1}$$

where $\bar{\xi}:=(\xi_1,\ldots,\xi_n)$ and ξ_j, $1\le j\le n$, are mutually independent Poisson r.vs defined on some probability space $\{\Omega,\mathcal{F},P\}$ with $\mathbf{E}\xi_j=u^jm_j/j!$. Here $u>0$ is an arbitrary number.

Apart from permutations, very important examples of assemblies are mappings of a finite set into itself interpreted as their functional digraphs. In this case, the components are just the components of the digraph. For mappings, we have $|\mathcal{A}_n|=n^n$ and

$$m_j/j!=(1/j)\sum_{s=0}^{j-1}j^s/s!\sim e^j/2j,\quad j\to\infty.$$

More examples can be found in [2].

Equality (1) is sometimes used in the following alternative definition of random assemblies. Such a structure σ of size n is understood as an unordered collection of $k_j(\sigma)$ components, defined on disjoint sets of j labelled points, such that the probability distribution of $\bar{k}(\sigma)$ is given by (1), where ξ_j, $1\le j\le n$, are arbitrary independent Poisson r.vs. In the particular case of $\mathbf{E}\xi_j=\theta/j$, where $\theta>0$ is a constant, we arrive at random permutations σ taken with the probability

$$\nu_n^{(\theta)}\big(\{\sigma\}\big)=\theta^{k_1(\sigma)+\cdots+k_n(\sigma)}/\theta(\theta+1)\cdots(\theta+n-1).$$

Then the quantity on the right-hand side of (1) is the Ewens Sampling Formula

$$P\big(\{\bar{s}\}\big)=\mathbf{1}\{\ell(\bar{s})=n\}\prod_{j=1}^{n}\left(\frac{\theta}{j}\right)^{s_j}\frac{1}{s_j!},\quad \bar{s}\in\mathbb{Z}_+^n,$$

playing a great role in various statistical applications. In what follows, we also call $\nu_n^{(\theta)}$ the Ewens Probability Measure on the symmetric group.

The so-called *Logarithmic Condition* (see [2]) in the case of assemblies requires that

$$m_j/j!\sim\theta y^j/j$$

for some constants $y>0$ and $\theta>0$ as $j\to\infty$. Under this condition, it is natural and technically convenient to take $u=y^{-1}$, which yields the relation $\mathbf{E}\xi_j\sim\theta/j$ as $j\to\infty$. The above presented examples of assemblies satisfy this condition, we may call them logarithmic.

There are a few ways to induce new classes of assemblies, e.g., refinement and coloring which are mentioned in [2, Section 2.4]. In the second approach, given a class of structures, one colors each of m_j components of size $j\ge 1$ with one of $n_j\ge 1$ colors and gets m_jn_j new components. Collecting them into assemblies one defines a new class. Now, a requirement of the asymptotic regularity for m_jn_j as $j\to\infty$ in the logarithmic condition becomes rather restrictive. This condition is not satisfied in the problems for the assemblies missing some component sizes. For the recent development of a probabilistic theory for such structures, see [16] and the references therein. Another path leading to non-logarithmic assemblies is related to weighted probability measures. For this, we refer to the author's papers [12,14] and Zacharovas' dissertation [26] containing generalizations of the Ewens probability in the symmetric group. Finally, the results (see [25,26,15]) on the value distribution of mappings defined on powers of permutations also concern the non-logarithmic structures. Having all that in mind, we introduce a new class of assemblies.

Definition. Let $n\ge 1$, and let μ_n be a probability measure on $\mathcal{A}_n$. The pair $\big(\mathcal{A}_n,\mu_n\big)$ is called weakly logarithmic if there exists a random vector $\bar{\xi}=(\xi_1,\ldots,\xi_n)$, maybe depending on n, with mutually independent Poissonian coordinates such that

$$\mu_n\big(\bar{k}(\sigma)=\bar{s}\big)=P\big(\bar{\xi}=\bar{s}|\ell(\bar{\xi})=n\big)$$

for each $\bar{s}\in\mathbb{Z}_+^n$ and

$$\frac{\theta'}{j}\leq\lambda_j:=\mathbf{E}\xi_j\leq\frac{\theta''}{j} \tag{2}$$

uniformly in $j\geq 1$ and $n\geq 1$ for some positive constants θ' and θ''.

In our notation, the *logarithmic assemblies* are characterized by the condition $\lambda_j\sim\theta/j$ as $j\to\infty$, where $\theta>0$ is a constant, and no dependence on n is allowed.

The main result of this paper is the following total variation (TV) approximation. Let $\mathcal{L}(X)$ be the distribution of a random variable X. Afterwards the index r, $1\leq r\leq n$, added to the vectors $\bar{k}(\sigma)$ and $\bar{\xi}$ will denote that only the first r coordinates are taken. Let $x_+=\max\{x,0\}$ for $x\in\mathbb{R}$ and let $f(x)\ll g(x)$ be the abbreviation of $f(x)=O\big(g(x)\big)$.

Theorem (*TV Approximation*)**.** *Let $(\mathcal{A}_n,\mu_n)$ be weakly logarithmic. There exists a positive constant $c_1=c_1(\theta',\theta'')$ such that*

$$\rho_{TV}\Big(\mathcal{L}\big(\bar{k}_r(\sigma)\big),\mathcal{L}(\bar{\xi}_r)\Big):=\sum_{\bar{s}\in\mathbb{Z}_+^r}\Big(\mu_n\big(\bar{k}_r(\sigma)=\bar{s}\big)-P(\bar{\xi}_r=\bar{s})\Big)_+\ll\Big(\frac{r}{n}\Big)^{c_1} \tag{3}$$

uniformly in $1\leq r\leq n$. The constant in $\ll$ depends on θ' and θ'' only.

In fact, the estimate (3) is nontrivial if $r\leq c_2n$, where $c_2>0$ is a sufficiently small constant depending on θ' and θ''. It is known (see [2] and the references therein) that, for permutations taken with equal probabilities, the remainder term in (3) is decreasing exponentially in the ratio n/r. On the other hand, for the permutations under the Ewens distribution with $\theta\neq 1$, the error is of order r/n. The same bound holds for logarithmic assemblies satisfying the conditions

$$\epsilon_j:=\theta^{-1}j\lambda_j-1\ll j^{-\alpha},\qquad \epsilon_j-\epsilon_{j+1}\ll j^{1+\alpha},\quad \alpha>0,$$

and some mild condition on $P(\xi_j=s)$ for $j\geq 1$ and $s\geq 2$. This and more involved results can be found in [2]. Considering the non-logarithmic case, we have failed to obtain $c_1=1$.

The main ingredient in the proof of the estimate (3) is Proposition 1. It concerns the asymptotic behavior of a ratio of the coefficients of two power series having rather particular shapes.

Our next task concerns the estimates of conditional discrete probabilities by appropriate unconditional ones. The very idea goes back to Ruzsa's paper [21] in probabilistic number theory. In [10], we adopted it and obtained upper estimates of the tail probabilities involving the cycle vector $\bar{k}(\sigma)$ of a random permutation σ. In the joint paper with Babu [3], the idea was extended to permutations taken with the Ewens probability and later, jointly with Norkūnienė [17], we adopted it for logarithmic assemblies. We now develop the same principle for the weakly logarithmic assemblies.

The problem is to estimate $P\big(\bar{\xi}\in A|\ell(\bar{\xi})=n\big)$ in terms of $P(\bar{\xi}\in B)$, where $A\subset B\subset\mathbb{Z}_+^n$. As it has been shown in [13] by examples, such universal estimates are impossible if $A=B$. Evidently, the problem is trivial if $B=\mathbb{Z}_+^n$ is allowed. Seeking a convenient and non-trivial extension of A we exploit geometric properties of the semi-lattice $\mathbb{Z}_+^n$. For two vectors $\bar{s}=(s_1,\ldots,s_n)$ and $\bar{t}=(t_1,\ldots,t_n)$, we set $\bar{s}\perp\bar{t}$ if $s_1t_1+\cdots+s_nt_n=0$ and write $\bar{s}\leq\bar{t}$ if $s_j\leq t_j$ for each $j\leq n$. Further, we adopt the notation $\bar{s}\parallel\bar{t}$ for the expression "$\bar{s}$ exactly enters $\bar{t}$" which means that $\bar{s}\leq\bar{t}$ and $\bar{s}\perp\bar{t}-\bar{s}$. For arbitrary subset $U\subset\mathbb{Z}_+^n$, we define its extension

$$V=V(U)=\big\{\bar{s}=\bar{t^1}+\bar{t^2}-\bar{t^3}\colon \bar{t^1},\bar{t^2},\bar{t^3}\in U,\bar{t^1}\perp(\bar{t^2}-\bar{t^3}),\bar{t^3}\parallel\bar{t^2}\big\}. \tag{4}$$

The use of triples of vectors from U is rather natural because, in general, $U\not\subset U\pm U$ if $\bar{0}\notin U$. The additional conditions in (4) assure that $V\subset\mathbb{Z}_+^n$ and are motivated by the subsequent applications. Set also $\bar{A}=\mathbb{Z}_+^n\setminus A$ and $\theta=\min\{1,\theta'\}$.

Theorem (*Upper Estimate*)**.** *Let $(\mathcal{A}_n,\mu_n)$ be weakly logarithmic and $\bar{\xi}$ be the Poissonian random vector introduced in the Definition. Then, for an arbitrary $U\subset\mathbb{Z}_+^n$,*

$$\mu_n\big(\bar{k}(\sigma)\in\overline{V}\big)=P\big(\bar{\xi}\in\overline{V}|\ell(\bar{\xi})=n\big)\ll P^{\theta}(\bar{\xi}\in\overline{U})+\mathbf{1}\{\theta<1\}n^{-\theta},$$

where the implicit constant depends on θ' and θ'' only.

The Upper Estimate becomes more transparent when applied to the value distributions of additive functions. We demonstrate this in a fairly general context. Let $(\mathbb{G},+)$ be an Abelian group and $h_j(s)$, $j\in\mathbb{N}$, $s\in\mathbb{Z}_+$, be a two-dimensional sequence in $\mathbb{G}$ satisfying the condition $h_j(0)=0$ for each $j\geq 1$. Then we can define an *additive function* $h\colon\mathcal{A}_n\to\mathbb{G}$ by

$$h(\sigma)=\sum_{j\leq n}h_j\big(k_j(\sigma)\big). \tag{5}$$

Corollary 1. *Let $(\mathbb{G}, +)$ be an Abelian group and $h: \mathcal{A}_n \to \mathbb{G}$ be an additive function. Uniformly in $A \subset \mathbb{G}$,*

$$\mu_n\big(h(\sigma) \notin A + A - A\big) \ll P^\theta\left(\sum_{j\le n} h_j(\xi_j) \notin A\right) + \mathbf{1}\{\theta < 1\} n^{-\theta}.$$

Corollary 2. *Let $h: \mathcal{A}_n \to \mathbb{R}$ be an additive function. Uniformly in $a \in \mathbb{R}$ and $u \ge 0$,*

$$\mu_n\big(|h(\sigma) - a| \ge u\big) \ll P^\theta\left(\left|\sum_{j\le n} h_j(\xi_j) - a\right| \ge u/3\right) + \mathbf{1}\{\theta < 1\} n^{-\theta}.$$

The example presented in [13] shows that even for permutations the constant $1/3$ in the last estimate cannot be substituted by $1 - \varepsilon$ or by $5/7$.

As in the case of logarithmic assemblies, the TV Approximation and the Upper Estimate can be used to prove general limit theorems for additive functions defined on $\mathcal{A}_n$. One can deal with the one-dimensional case (see, for instance, [2, Section 8.5]) or examine weak convergence of random combinatorial processes (see [3,4,11], and [2, Section 8.1]). This approach can be applied to examine strong convergence of sums of r.vs defined on $\mathcal{A}_n$. Extending [19], we now obtain an analog of the functional law of iterated logarithm. It can be compared with Major's [9] result for independent r.vs, generalizing the celebrated Strassen's theorem.

It is worth stressing that we deal with random variables which are defined on a sequence of probability spaces, not on a fixed space. This raises the first obstacle to be overcome; therefore, we adopt some basic definitions.

Let (S, d) be a separable metric space. Assume that $X, X_1, X_2, \dots, X_n$ are S-valued random variables all defined on the probability space $\{\Omega_n, \mathcal{F}_n, P_n\}$. Denote by $d(Y, A) := \inf\{d(Y, Z): Z \in A\}$, $A \subset S$, $Y \in S$, the distance from Y to A. We say that X_m converges to $X\{P_n\}$-almost surely ($\{P_n\}$-a.s.), if for each $\varepsilon > 0$

$$\lim_{n_1\to\infty} \limsup_{n\to\infty} P_n\Big(\max_{n_1\le m\le n} d(X_m, X) \ge \varepsilon\Big) = 0.$$

If $P_n = P$ does not depend on n, our definition agrees with that of the classical almost sure convergence (see [20, Chapter X]). A compact set $A \subset S$ is called a *cluster* for the sequence X_m if, for each $\varepsilon > 0$ and each $Y \in A$,

$$\lim_{n_1\to\infty} \limsup_{n\to\infty} P_n\Big(\max_{n_1\le m\le n} d(X_m, A) \ge \varepsilon\Big) = 0$$

and

$$\lim_{n_1\to\infty} \liminf_{n\to\infty} P_n\Big(\min_{n_1\le m\le n} d(X_m, Y) < \varepsilon\Big) = 1.$$

We denote the last two relations, by

$$X_m \Rightarrow A \quad (\{P_n\}\text{-a.s.}).$$

Let $C[0, 1]$ be the Banach space of continuous functions on the interval $[0, 1]$ with the supremum distance $\rho(\cdot, \cdot)$. The set of absolutely continuous functions g such that $g(0) = 0$ and

$$\int_0^1 \big(g'(t)\big)^2 \mathrm{d}t \le 1$$

is called the *Strassen set* $\mathcal{K}$. We shall show that it is the cluster set of some combinatorial processes constructed using partial sums

$$h(\sigma, m) := \sum_{j\le m} h_j\big(k_j(\sigma)\big),$$

where $h_j(s) \in \mathbb{R}$ and $1 \le m \le n$. Set $a_j = h_j(1)$,

$$A(m) := \sum_{j=1}^m a_j(1 - \mathrm{e}^{-\lambda_j}), \qquad B^2(m) := \sum_{j=1}^m a_j^2 \mathrm{e}^{-\lambda_j}\big(1 - \mathrm{e}^{-\lambda_j}\big),$$

and $\beta(m) = B(m)\sqrt{2LLB(m)}$, where $Lx := \log\max\{x, \mathrm{e}\}$, $x \in \mathbb{R}$, and $1 \le m \le n$. We denote by $u_m(\sigma, t)$ the polygonal line joining the points

$$(0, 0), \qquad \big(B^2(i), h(\sigma, i) - A(i)\big), \quad 1 \le i \le m,$$

and set

$$U_m(\sigma, t) = \beta(m)^{-1} u_m\big(\sigma, B^2(m)t\big), \quad \sigma \in \mathcal{A}_n,\ 0 \le t \le 1,$$

for $1 \le m \le n$.

In the next two theorems and their corollaries, applying the definition of weakly logarithmic structures (see Definition), we assume that the involved independent Poisson r.vs $\xi_j, j \geq 1$, do not depend on n. The first result generalizes the cases examined in [10,18,19].

Theorem 1. *Let* $(\mathcal{A}_n, \mu_n)$ *be weakly logarithmic. If* $B(n) \to \infty$ *and*

$$a_j = \mathrm{o}\left(\frac{B(j)}{\sqrt{LLB(j)}}\right), \quad j \to \infty, \tag{6}$$

then

$$U_m(\sigma, \cdot) \Rightarrow \mathcal{K} \quad (\{\mu_n\}\text{-a.s.}). \tag{7}$$

Applying continuous functionals defined on the space $C[0, 1]$, we derive partial cases of the last theorem.

Corollary 3. *Let the conditions of Theorem 1 be satisfied. The following relations hold* $\{\mu_n\}$*-a.s.:*

(i) $U_m(\sigma, 1) \Rightarrow [-1, 1]$;
(ii) $\big(U_m(\sigma, 1/2), U_m(\sigma, 1)\big) \Rightarrow \{(u, v) \in \mathbb{R}^2 : u^2 + (v-u)^2 \leq 1/2\}$;
(iii) *if* $U_{m'}(\sigma, 1/2) \Rightarrow \sqrt{2}/2$ *for some subsequence* $m' \to \infty$, *then* $U_{m'}(\sigma, \cdot) \Rightarrow g_1$, *where*

$$g_1(t) = \begin{cases} t\sqrt{2} & \text{if } 0 \leq t \leq 1/2, \\ \sqrt{2}/2 & \text{if } 1/2 \leq t \leq 1; \end{cases}$$

(iv) *if* $U_{m'}(\sigma, 1/2) \Rightarrow 1/2$ *and* $U_{m'}(\sigma, 1) \Rightarrow 0$ *for some subsequence* $m' \to \infty$, *then* $U_{m'}(\sigma, \cdot) \Rightarrow g_2$, *where*

$$g_2(t) = \begin{cases} t & \text{if } 0 \leq t \leq 1/2, \\ 1-t & \text{if } 1/2 \leq t \leq 1. \end{cases}$$

Using other more sophisticated functionals (see, e.g., [7, Chapter I]), one can proceed in a similar manner. Claim (i) includes the assertion that

$$|h(\sigma, m) - A(m)| \leq (1+\varepsilon)\beta_m$$

holds uniformly in m, $n_1 \leq m \leq n$, for asymptotically almost all $\sigma \in \mathcal{A}_n$ as n and n_1 tend to infinity. Moreover, it shows that the upper bound is sharp apart from the term $\varepsilon\beta(m)$. An idea how to improve this error goes back to Feller's paper [5]. It has been exploited by the author in the case of a special additive function defined on permutations. Recently, this was extended for the logarithmic assemblies [17]. We now formulate a more general result.

We say that an increasing sequence ψ_m, $m \geq 1$, belongs to the *upper class* Ψ^+ (respectively, *the lower class* Ψ^-) if

$$\lim_{n_1\to\infty} \limsup_{n\to\infty} \mu_n\Big(\max_{n_1 \leq m \leq n} \psi_m^{-1}|h(\sigma, m) - A(m)| \geq 1\Big) = 0, \tag{8}$$

$$\left(\lim_{n_1\to\infty} \liminf_{n\to\infty} \mu_n\Big(\max_{n_1 \leq m \leq n} \psi_m^{-1}|h(\sigma, m) - A(m)| \geq\Big) = 1\right).$$

Theorem 2. *Let* $(\mathcal{A}_n, \mu_n)$ *be weakly logarithmic and* $B(n) \to \infty$. *Assume that a positive sequence* $\phi_n \to \infty$ *is such that*

$$a_j = \mathrm{O}\left(\frac{B(j)}{\phi_j^3}\right), \quad j \geq 1. \tag{9}$$

If the series

$$\sum_{j=1}^{\infty} \frac{a_j^2 \phi_j}{jB^2(j)} \mathrm{e}^{-\phi_j^2/2} \tag{10}$$

converges, then $B(m)\phi_m \in \Psi^+$. *If the series (10) diverges, then* $B(m)\phi_m \in \Psi^-$.

Since the series

$$\sum_{j=1}^{\infty} \frac{a_j^2}{j} \frac{(LLB(j))^{1/2}}{B^2(j)(LB(j))^{1+x}}$$

converges for $x = \varepsilon$ and diverges for $x = -\varepsilon$, the last theorem implies (i) in Corollary 3 under a bit stronger condition.

Let us illustrate Theorem 2. For brevity, set $L_1 u = Lu$ and $L_k u = L(L_{k-1}u)$ if $k \geq 2$. Denote $\gamma_{2m}^2(\pm\varepsilon) := 2(1 \pm \varepsilon)L_2 B(m)$,

$$\gamma_{3m}^2(\pm\varepsilon)/2 := L_2 B(m) + \frac{3}{2}(1 \pm \varepsilon)L_3 B(m),$$

and

$$\gamma_{sm}^2(\pm\varepsilon)/2 := L_2B(m) + \frac{3}{2}L_3B(m) + L_4B(m) + \cdots + (1\pm\varepsilon)L_sB(m)$$

for $s \geq 4$.

Corollary 4. *Under the conditions of Theorem 2, we have*

$$B(m)\gamma_{sm}(\varepsilon) \in \Psi^+$$

and

$$B(m)\gamma_{sm}(-\varepsilon) \in \Psi^-$$

for each $s \geq 2$.

What do we obtain for the sequence of additive functions

$$s(\sigma, m) := \sum_{j\leq m} \mathbf{1}\{k_j(\sigma) \geq 1\}$$

counting the number of components with different sizes $1 \leq j \leq m \leq n$? Assume that $(\mathcal{A}_n, \mu_n)$ is weakly logarithmic and, in addition,

$$\sum_{j\leq m} \lambda_j = \varkappa \log m + O(1),$$

where $\theta' \leq \varkappa \leq \theta''$ is a constant. For our example, $a_j \equiv 1$ and $A(m) = B^2(m) + O(1) = \varkappa \log m + O(1)$ as $m \to \infty$. Consequently, we may assert by Theorem 1 that, "for almost all $\sigma \in \mathcal{A}_n$", the Strassen set $\mathcal{K}$ is the cluster of the sequence of functions

$$\Big(s\big(\sigma, (\varkappa/2)tLm\big) - (\varkappa/2)tLm\Big)/\sqrt{2\varkappa LmL_3m}, \quad n_1 \leq m \leq n,\ 0 \leq t \leq 1,$$

as $n \to \infty$ and $n_1 \to \infty$. Similar claims hold for the projections or other functionals as indicated in Corollary 3.

Let us pay more attention to Corollary 4. Define

$$\delta_{sm}^2(\pm\varepsilon)/2 := L_3m + \frac{3}{2}L_4m + L_5m + \cdots + (1\pm\varepsilon)L_sm, \quad s \geq 3.$$

Returning to the expanded notation, we obtain

$$\lim_{n_1\to\infty} \limsup_{n\to\infty} \mu_n\left(\max_{n_1\leq m\leq n} \frac{|s(\sigma, m) - \varkappa Lm|}{\delta_{sm}(+\varepsilon)\sqrt{\varkappa Lm}} \geq 1\right) = 0 \tag{11}$$

and

$$\lim_{n_1\to\infty} \liminf_{n\to\infty} \mu_n\left(\max_{n_1\leq m\leq n} \frac{|s(\sigma, m) - \varkappa Lm|}{\delta_{sm}(-\varepsilon)\sqrt{\varkappa Lm}} \geq 1\right) = 1 \tag{12}$$

for every $0 < \varepsilon < 1$. This can be reversed.

Introduce the ordered statistics

$$1 \leq j_1(\sigma) < \cdots < j_s(\sigma), \quad s := s(\sigma, n),$$

of different component sizes appearing in a random $\sigma \in \mathcal{A}_n$. Denote

$$\eta_{sm}^2(\pm\varepsilon)/2 := L_2m + \frac{3}{2}L_3m + L_4m + \cdots + (1\pm\varepsilon)L_sm$$

for $s \geq 2$. Using $s\big(\sigma, j_m(\sigma)\big) = m$, we obtain from (11) and (12)

$$\lim_{n_1\to\infty} \limsup_{n\to\infty} \mu_n\left(\max_{n_1\leq m\leq s} \frac{|\varkappa Lj_m(\sigma) - m|}{\eta_{sm}(+\varepsilon)\sqrt{m}} \geq 1\right) = 0$$

and

$$\lim_{n_1\to\infty} \liminf_{n\to\infty} \mu_n\left(\max_{n_1\leq m\leq s} \frac{|\varkappa Lj_m(\sigma) - m|}{\eta_{sm}(-\varepsilon)\sqrt{m}} \geq 1\right) = 1$$

for every $0 < \varepsilon < 1$. Consequently, asymptotically for almost all $\sigma \in \mathcal{A}_n$,

$$|\varkappa Lj_m(\sigma) - m| \leq \eta_{sm}(+\varepsilon)\sqrt{m}$$

uniformly in $m \in [n_1, s(\sigma, n)]$ if $n \to \infty$ and $n_1 \to \infty$ arbitrarily slowly. This is sharp in the described sense.

The main argument in deriving Theorems 1 and 2 is the same; therefore, we will omit the proofs of the second result and its corollaries. The technical details in the case of logarithmic assemblies can be found in [17]. Finally, we observe that by substituting r.vs ξ_j, $1 \leq j \leq n$, by appropriate independent geometrically distributed and negative binomial r.vs, one can similarly extend the logarithmic classes of additive arithmetical semigroups and weighted multisets (see [2]).

2. Proof of the TV Approximation

We apply the analytic method developed in our papers (see [12,14], and the references therein). Initially, the approach was used to obtain asymptotic formulas for some Fourier transforms of distributions. That led to general one-dimensional limit theorems, including the optimal remainder term estimates. In this regard, apart from the above mentioned, the works by Zacharovas [24–26] were noticeable. On the other hand, the investigations carried out by Arratia et al. (see [1,2], and the references therein) have shown that the deeper TV approximation can be obtained by other analytic means.

Both approaches are based on the expression of the TV distance in terms of the local probabilities including large deviations for very special r.vs (see Lemma 1). In contrast to the previous investigations, we are not seeking the asymptotic formulae for each of these probabilities but examine their ratios. This way, we succeed in avoiding the regularity assumptions which used to be posed on the generating series of assemblies (see [22,23,8]). The comparative analysis already proved to be useful in the author's recent papers (see [15,16]).

The first lemma reduces the problem to a one-dimensional case. For $\bar{s} = (s_1, \ldots, s_n)$, set $\ell_{ij}(\bar{s}) = (i+1)s_{i+1} + \cdots + js_j$ if $0 \le i < j \le n$. Moreover, let $\ell_r(\bar{s}) := \ell_{0r}(\bar{s})$, where $1 \le r \le n$. Then $\ell_n(\bar{s}) = \ell(\bar{s})$.

Lemma 1. *We have*

$$\rho_{TV}\big(\mathcal{L}(\bar{k}_r(\sigma)), \mathcal{L}(\bar{\xi}_r)\big) = \rho_{TV}\big(\mathcal{L}(\bar{\xi}_r | \ell(\bar{\xi}) = n), \mathcal{L}(\bar{\xi}_r)\big)$$
$$= \sum_{m \in \mathbb{Z}_+} P\big(\ell_r(\bar{\xi}) = m\big)\left(1 - \frac{P\big(\ell_{rn}(\bar{\xi}) = n-m\big)}{P\big(\ell(\bar{\xi}) = n\big)}\right)_+. \tag{13}$$

Proof. See [2, p. 69]. □

Consequently, the ratio of probabilities on the right-hand side of (13) is now the main objective. The limiting behavior of the probabilities can be rather complicated for weakly logarithmic assemblies but, as we will show in the sequel, the ratio of probabilities in (13) is regular. Since

$$P\big(\ell_{rn}(\bar{\xi}) = m\big) = \frac{1}{2\pi \mathrm{i}} \int_{|z|=1} \frac{1}{z^{m+1}} \exp\left\{\sum_{r<j\le n} \lambda_j (z^j - 1)\right\} \mathrm{d}z, \tag{14}$$

one can apply the analytic technique (see [12] or [14]) which has been elaborated to compare the Taylor coefficients of two power series. Namely, if $d_j \in \mathbb{R}_+$ and $f_j \in \mathbb{C}$, $1 \le j \le n$, are two sequences (either of them may depend on n or on other parameters) and

$$D(z) := \exp\left\{\sum_{j\le n} \frac{d_j}{j} z^j\right\} =: \sum_{s=0}^{\infty} D_s z^s,$$
$$F(z) := \exp\left\{\sum_{j\le n} \frac{f_j}{j} z^j\right\} =: \sum_{s=0}^{\infty} F_s z^s,$$

then, under certain conditions, we have obtained asymptotic formulas for F_n/D_n as $n \to \infty$. As in [12], we now also assume the inequalities

$$d' \le d_j \le d'' \tag{15}$$

for all $1 \le j \le n$ and some positive constants $d' \le d''$. In our case, f_j are very special; therefore, we can simplify the previous argument and get rid of (2.4) in [12]. The goal now is to find the ratio F_m/D_n preserving some uniformity.

Set, for brevity,

$$e_r = \exp\left\{-\sum_{j\le r} \frac{d_j}{j}\right\}.$$

Proposition 1. *Assume that the sequence d_j, $1 \le j \le n$, satisfies Condition (15). Given r, $0 \le r \le n$, set $f_j = d_j$ if $r < j \le n$ and $f_j = 0$ if $j \le r$. Let $0 \le \eta \le 1/2$ and $1/n \le \delta \le 1/2$ be arbitrary. Then there exist positive constants $c = c(d')$ and $c_3 = c_3(d', d'')$ such that*

$$F_m/(e_r D_n) - 1 \ll \eta\delta^{-1} + \delta^c + \big((r/n)\mathbf{1}\{r \ge 1\}\big)\delta^{-1-c_3}$$

uniformly in (r, m) such that

$$0 \le r \le \delta n, \qquad n(1-\eta) \le m \le n. \tag{16}$$

Here and in the proof of this claim, the constant in $\ll$ depends on d' and d'' only.

Let us start with a sketch of the proof and the following notation. Let $K,\ 1 \le \delta n < K \le n,\ K > r$, be a parameter to be chosen later. For a fixed $0 < \alpha < 1$, we introduce the functions

$$q(z) := \sum_{r<j\le n} d_j z^{j-1}, \qquad G_1(z) = \exp\left\{\alpha \sum_{r<j\le K} \frac{d_j}{j} z^j\right\},$$
$$G_2(z) = \exp\left\{-\alpha \sum_{K<j\le n} \frac{d_j}{j} z^j\right\}, \qquad G_3(z) = F^\alpha(z) - G_1(z).$$

We denote by $[z^k]U(z)$ the kth Taylor coefficient of an analytic at zero function $U(z)$. Observe that

$$[z^k]G_3(z) \le [z^k]F^\alpha(z), \quad k \ge 0. \tag{17}$$

Set further $T = (\delta n)^{-1}$,

$$\Delta = \{z = e^{it} : T < |t| \le \pi\}, \qquad \Delta_0 = \{z = e^{it} : |t| \le T\}.$$

Seeking F_m, we use the following identity

$$\begin{aligned} F_m &= \frac{1}{2\pi i m} \int_{|z|=1} \frac{F'(z)}{z^m}\, dz \\ &= \frac{1}{2\pi i m}\left(\int_{\Delta_0} + \int_\Delta\right) \frac{F'(z)\big(1 - G_2(z)\big)}{z^m}\, dz + \frac{1}{2\pi i m}\int_{|z|=1} \frac{F'(z)G_2(z)}{z^m}\, dz =: J_0 + J_1 + J_2. \end{aligned} \tag{18}$$

In what follows, we estimate the integrals J_1 and J_2 and, changing the integrand, reduce J_0 to the main term of an asymptotic formula for D_n. This yields the desired expression for the ratio F_m/D_n in the needed region for m. Lemmas 6 and 7 formalizes all this.

The estimates of J_1 and J_2 are rather involved. The first two lemmas below just hide some part of the technical calculations. The integral J_2 is estimated in Lemma 4. Apart from Lemma 2, here we use an elementary argument. Estimating J_1 in Lemma 5, we apply Lemmas 2 and 3 together with Parseval's equality for power series. This is a rather fruitful idea going back to the earlier author's papers (see references in [14]). It allows obtaining an estimate on a large arc without any extension of the series outside the convergence disk. We hope that the following diagram of inter-relations in our proofs will enhance readability. If we use the notation $a \propto \{b, c\}$ to indicate that the proof of Lemma a requires Lemmas b and c, then

$$7 \propto \{2, 6\}, \qquad 6 \propto \{4, 5\}, \qquad 5 \propto \{2, 3\}, \qquad 4 \propto \{2\}.$$

It remains to prove the listed lemmas.

Lemma 2. *We have*

$$D(1)n^{-1} \ll D_n \ll D(1)n^{-1}$$

for all $n \ge 1$.

Proof. This is Lemma 3.1 from [12]. □

Lemma 3. *Let* $\delta n \ge 1$. *Then*

$$\max_{T\le|t|\le\pi} |F(e^{it})| \ll e_r D(1)\delta^{d'}$$

uniformly in $0 \le r \le \delta n$.

Proof. By definition,

$$\begin{aligned} \frac{|F(e^{it})|}{D(1)} &= e_r \frac{|F(e^{it})|}{F(1)} = e_r \exp\left\{\sum_{r<j\le n} \frac{d_j(\cos tj - 1)}{j}\right\} \\ &\le e_r \exp\left\{d' \sum_{\delta n<j\le n} \frac{\cos tj - 1}{j}\right\} \end{aligned} \tag{19}$$

uniformly in $0 \le r \le \delta n$. We now use the relation

$$S(x,t) := \sum_{j\le x} \frac{\cos tj - 1}{j} = \log\min\left\{1, \frac{2\pi}{x|t|}\right\} + O(1),$$

valid for all $x \geq 1$ and $|t| \leq \pi$. It shows that $S(\delta n, t) \ll 1$ for $T = (\delta n)^{-1} \leq |t| \leq \pi$. Hence, for such t,

$$S(n, t) - S(\delta n, t) \leq S(n, T) + O(1) = \log \delta + O(1).$$

This yields the desired claim. □

Lemma 4. *If* $0 < \alpha < 1$ *and* $\delta n \geq 1$, *then*

$$J_2 \ll D_n e_r (K/n)^{\alpha d'}$$

uniformly in $n/2 \leq m \leq n$.

Proof. For brevity, let

$$u_s := [z^s] G_1(z), \qquad v_l := [z^l] F^{1-\alpha}(z), \quad s, l \geq 0.$$

Since

$$F'(z) G_2(z) = q(z) G_1(z) F^{1-\alpha}(z),$$

by Cauchy's formula, we have

$$J_2 = \frac{1}{2\pi \mathrm{i} m} \int_{|z|=1} q(z) G_1(z) F^{1-\alpha}(z) \frac{\mathrm{d}z}{z^m} = \frac{1}{m} \sum_{r<j\leq m} d_j \sum_{s+l=m-j} u_s v_l.$$

Hence, by condition (15),

$$\begin{aligned} J_2 &\leq \frac{2d''}{n} \sum_{s\leq n} u_s \sum_{l\leq n} v_l \leq \frac{2d''}{n} F^{1-\alpha}(1) G_1(1) \\ &= \frac{2d'' F(1)}{n} \exp\left\{ -\alpha \sum_{K<j\leq n} \frac{d_j}{j} \right\} \ll D_n e_r (K/n)^{\alpha d'}. \end{aligned}$$

In the last step, we used Lemma 2.

The lemma is proved. □

Lemma 5. *Let* $0 < \alpha < 1$ *be arbitrary and* $\delta n \geq 1$. *Then*

$$J_1 \ll \frac{e_r n D_n}{K} \delta^{d'(1-\alpha)}$$

uniformly in $n/2 \leq m \leq n$ *and* $0 \leq r \leq \delta n$.

Proof. Recalling the previous notation, we can rewrite

$$J_1 = \frac{1}{2\pi \mathrm{i} m} \int_{\Delta} q(z) F^{1-\alpha}(z) G_3(z) \frac{\mathrm{d}z}{z^m}.$$

Hence, by Lemma 3,

$$\begin{aligned} J_1 &\ll n^{-1} \max_{z\in\Delta} |F(z)|^{1-\alpha} \int_{|z|=1} |q(z)|\, |G_3(z)|\, |\mathrm{d}z| \\ &\ll n^{-1} \left(e_r D(1) \delta^{d'} \right)^{1-\alpha} \left(\int_{|z|=1} |q(z)|^2 |\mathrm{d}z| \right)^{1/2} \left(\int_{|z|=1} |G_3(z)|^2 |\mathrm{d}z| \right)^{1/2}. \end{aligned}$$

By Parseval's equality,

$$\int_{|z|=1} |q(z)|^2 |\mathrm{d}z| = 2\pi \sum_{r<j\leq n} d_j^2 \leq 2\pi (d'')^2 n$$

and, recalling (17),

$$\begin{aligned} \int_{|z|=1} |G_3(z)|^2 |\mathrm{d}z| &\leq 2\pi \sum_{l>K} \left([z^l] G_3(z) \right)^2 \leq \frac{2\pi}{K^2} \sum_{l=1}^{\infty} l^2 \left([z^l] F^{\alpha}(z) \right)^2 \\ &= \frac{1}{K^2} \int_{|z|=1} |(F^{\alpha}(z))'|^2 |\mathrm{d}z| \end{aligned}$$

$$= \frac{\alpha^2}{K^2}\int_{|z|=1}|F^\alpha(z)|^2|q(z)|^2|\mathrm{d}z|$$
$$\ll \frac{(e_rD(1))^{2\alpha}}{K^2}\int_{|z|=1}|q(z)|^2|\mathrm{d}z| \ll \frac{(e_rD(1))^{2\alpha}n}{K^2}.$$

Collecting the last three estimates, by Lemma 2, we obtain the desired estimate of J_1.

Lemma 5 is proved. □

At this stage, we have the following estimate.

Lemma 6. *If condition (15) is satisfied and $\delta n \geq 1$, then there exists a positive constant $c = c(d')$ such that*

$$F_m = J_0 + \mathrm{O}\big(e_rD_n\delta^c\big) \tag{20}$$

uniformly in $0 \leq r \leq \delta n$ and $n/2 \leq m \leq n$. Moreover,

$$D_n = \frac{1}{2\pi \mathrm{i}n}\int_{\Delta_0} D'(z)\big(1-G_2(z)\big)\frac{\mathrm{d}z}{z^n} + \mathrm{O}\big(D_n\delta^c\big). \tag{21}$$

Proof. It suffices to apply Lemmas 4 and 5 with $K = \delta^{c(\alpha)}n$, where

$$c(\alpha) = \min\big\{1, d'(1-\alpha)/(\alpha d'+1)\big\},$$

and optimize the function $d'\alpha c(\alpha)$ with respect to $\alpha \in (0, 1)$. Its maximum is attained at $\alpha = \alpha_0 := \big(\sqrt{1+d'}-1\big)/d'$. This yields

$$c(\alpha_0) = \min\{1, \sqrt{1+d'}-1\}.$$

Hence, if $d' \leq 3$, then (20) holds with $c = d'\alpha_0c(\alpha_0) = (\sqrt{1+d'}-1)^2$. If $d' > 3$, then the choice $\alpha = \alpha_1 := (d'-1)/2d'$ gives $c(\alpha_1) = 1$. Consequently, (20) holds with $c = d'\alpha_1 = (d'-1)/2$.

To obtain (21), use (20) with $r = 0$ and $m = n$.

The lemma is proved. □

Lemma 7. *If $0 \leq \eta \leq 1/2$ and $1/n \leq \delta \leq 1/2$ are arbitrary, then*

$$J_0/e_rD_n - 1 \ll \eta\delta^{-1} + \delta^c + (r/n)\mathbf{1}\{r \geq 1\}\delta^{-1-c_3}, \quad c_3 := cd''/d',$$

uniformly in $n(1-\eta) \leq m \leq n$ and $0 \leq r \leq \delta n$. The constant c has been defined in the proof of Lemma 6.

Proof. If $z \in \Delta_0$ and $r \geq 1$, then

$$\begin{aligned}
F'(z) &= e_rD(z)\exp\Big\{-\sum_{j\leq r}\frac{d_j}{j}(z^j-1)\Big\}q(z)\\
&= e_rD(z)\Big(1+\mathrm{O}\Big(\frac{r}{\delta n}\Big)\Big)\Big(\sum_{j\leq n}-\sum_{j\leq r}\Big)d_jz^{j-1}\\
&= e_rD'(z)\big(1+\mathrm{O}\big(r/\delta n\big)\big) + \mathrm{O}\big(re_rD(1)\big)
\end{aligned}$$

and

$$z^{-m} = z^{-n}\big(1+\mathrm{O}(\eta\delta^{-1})\big).$$

Moreover, recalling the choice $K = \delta^{c(\alpha_0)}n$ if $d' \leq 3$, we have

$$1-G_2(z) \ll \exp\big\{d''\alpha_0c(\alpha_0)\log(1/\delta)\big\} = \exp\big\{cd''/d'\log(1/\delta)\big\} = \delta^{-c_3}.$$

If $d' > 3$, then $K = \delta n$ and $c = d'\alpha_1$, therefore the last estimate holds again.

Consequently, by virtue of $m^{-1} = n^{-1}\big(1+\mathrm{O}(\eta)\big)$, from Lemma 2 and Eq. (21), we obtain

$$\begin{aligned}
J_0 &= \frac{e_r}{2\pi \mathrm{i}n}\Big(1+\mathrm{O}\Big(\Big(\frac{r}{n}+\eta\Big)\frac{1}{\delta}\Big)\Big)\int_{\Delta_0}D'(z)\big(1-G_2(z)\big)\frac{\mathrm{d}z}{z^n} + \mathrm{O}\big(e_rD_n(r/n)\delta^{-1-c_3}\big)\\
&= e_rD_n\Big(1+\mathrm{O}\big((r/n+\eta)\delta^{-1}\big)\Big)\big(1+\mathrm{O}(\delta^c)\big) + \mathrm{O}\big(e_rD_n(r/n)\delta^{-1-c_3}\big).
\end{aligned}$$

This implies the desired estimate for $r \geq 1$. If $r < 1$, the terms having the fraction r/n do not appear.

The lemma is proved. □

Proof of Proposition 1. Apply (20) and the last lemma. □

Proof of TV Approximation. We now apply Lemma 1 and Proposition 1 with $d_j = \lambda_j$. condition (15) for weakly logarithmic assemblies is satisfied with $d' = \theta'$ and $d'' = \theta''$. We may leave the previous meaning of the constants c and c_3. From (14) and Proposition 1 with $\eta = (r/n)^{1/2}$, and $\delta = (r/n)^x$, where $0 < x < \min\{1/2,\, 1/(1+c_3)\}$ is a fixed number, we obtain

$$\frac{P\big(\ell_{rn}(\bar\xi) = n-m\big)}{P\big(\ell(\bar\xi) = n\big)} - 1 \ll (r/n)^{1/2-x} + (r/n)^{cx} + (r/n)^{1-(1+c_3)x}$$
$$\ll (r/n)^{c_1}, \quad c_1 = c_1(\theta', \theta'') > 0,$$

uniformly in $0 \le m \le \sqrt{rn}$ provided that $1 \le r \le 2^{-1/x} n =: c_2 n$.

The summands over $m > \sqrt{rn}$ in (13) contribute not more than

$$(rn)^{-1/2} \mathbf{E} \ell_r(\bar\xi) = (rn)^{-1/2} \sum_{j \le r} j\lambda_j \le \theta'' (r/n)^{1/2}.$$

Hence, by (13), we obtain

$$\rho_{TV}\big(\mathcal{L}\big(\bar k_r(\sigma)\big), \mathcal{L}(\bar\xi_r)\big) \ll (r/n)^{c_1},$$

where $1 \le r \le c_2 n$. Since the claim of the TV Approximation is trivial for $c_2 n < r \le n$, we have finished its proof. □

3. Proof of Theorem 1 and its corollaries

Set $\mathbb{Z}_+^n(m) = \{\bar s \in \mathbb{Z}_+^n : \ell(\bar s) = m\}$ where $0 \le m \le n$. For arbitrary distributions $p_j(k)$, $1 \le j \le n$, on $\mathbb{Z}_+$ we define the product measure on $\mathbb{Z}_+^n$ by

$$P(\{\bar k\}) = \prod_{j \le n} p_j(k_j), \quad \bar k = (k_1, \ldots, k_n) \in \mathbb{Z}_+^n.$$

Denote for brevity $P_n = P(\mathbb{Z}_+^n(n))$. Let $V = V(U)$ be the extension of an arbitrary subset $U \subset \mathbb{Z}_+^n$ defined in (4).

Lemma 8. *Suppose $n \ge 1$ and assume that there exist positive constants c_4, c_5, C_1, C_2 such that*

(i) *$p_j(0) \ge c_4$ for all $1 \le j \le n$;*

(ii) *$P\big(\mathbb{Z}_+^n(m)\big) \le C_1 \left(\frac{n}{m+1}\right)^{1-\theta} P_n$ for $0 \le m \le n-1$ and for some $0 < \theta \le 1$;*

(iii) *$P_n \ge c_5 n^{-1}$;*

(iv) *for $1 \le m \le n$,*

$$\sum_{\substack{k \ge 1,\, j \le n \\ kj = m}} \frac{p_j(k)}{p_j(0)} \le \frac{C_2}{m}.$$

Then

$$P\big(\overline{V} \mid \mathbb{Z}_+^n(m)\big) \le C P^\theta(\overline{U}) + C_1 C_2 \theta^{-1} n^{-\theta} \mathbf{1}\{\theta < 1\},$$

where

$$C := \max\left\{\frac{32}{c_4^2}, \frac{C_2}{c_5} + \frac{4C_1}{c_4} + \frac{C_1 C_2}{\theta}\right\}.$$

Proof. See [3], Appendix. □

Proof of Upper Estimate. It suffices to check conditions (i)–(iv) of the last lemma for the Poissonian probabilities $p_j(k)$ with parameters λ_j. By virtue of condition (2), (i) and (iv) are trivial. Further, we find

$$P\big(\mathbb{Z}_+^n(m)\big) = P\left(\sum_{j=1}^m j\xi_j = m,\, \xi_{m+1} = 0, \ldots, \xi_n = 0\right)$$
$$= \exp\left\{-\sum_{j=1}^n \lambda_j\right\} \sum_{\ell_m(\bar k) = m} \prod_{j=1}^m \frac{\lambda_j^{k_j}}{k_j!}$$
$$= \exp\left\{-\sum_{j=1}^n \lambda_j\right\} [z^m] \exp\left\{\sum_{j \le m} \lambda_j z^j\right\}, \quad 0 \le m \le n.$$

Hence, applying Lemma 2, we obtain

$$P\big(\mathbb{Z}_+^n(m)\big) \asymp \frac{1}{m+1}\exp\left\{-\sum_{j=m+1}^{n}\lambda_j\right\}$$

for $0 \le m \le n$, where $a \asymp b$ means $a \ll b \ll a$. This and condition (2) implies (ii) and (iii).

The theorem is proved. □

Proof of Corollary 1. Apply the Upper Estimate to

$$U = \left\{\bar{t} \in \mathbb{Z}_+^n : H(\bar{t}) \in A\right\},$$

where $H(\bar{t}) := \sum_{j\le n} h_j(t_j)$, and check that

$$V(U) \subset \left\{\bar{s} \in \mathbb{Z}_+^n : H(\bar{s}) \in A + A - A\right\}.$$

Now

$$\begin{aligned}\mu_n\big(h(\sigma) \notin A + A - A\big) &= P\big(H(\bar{\xi}) \notin A + A - A | \ell(\bar{\xi}) = n\big)\\ &\le P\big(\bar{\xi} \notin V(U) | \ell(\bar{\xi}) = n\big)\\ &\ll P^{\theta}\big(\bar{\xi} \notin U\big) + \mathbf{1}\{\theta' < 1\}n^{-\theta'}\\ &= P^{\theta}\big(H(\bar{\xi}) \notin A\big) + \mathbf{1}\{\theta' < 1\}n^{-\theta'}.\end{aligned}$$

Corollary 1 is proved. □

Proof of Corollary 2. Apply the previous corollary to $\mathbb{G} = \mathbb{R}$ and $A = \{t : |t - a| \le u/3\}$. □

4. Proof of Theorem 1

We adopt the argument used in the case of permutations [10] and for the logarithmic assemblies [19].

Let $Z_1, Z_2, \ldots, Z_n$ be independent random variables defined on some probability space $\{\Omega, \mathcal{F}, P\}$, with $\mathbf{E}Z_j = 0$, $\mathbf{E}Z_j^2 < \infty$, $j = 1, 2, \ldots,$ and

$$S_m = \sum_{j=1}^{m} Z_j, \qquad D_m^2 = \sum_{j=1}^{m}\mathbf{E}Z_j^2.$$

We define the polygonal lines $s_n(\cdot)$: $[0, D_n^2] \to \mathbb{R}$ by

$$s_n(t) = S_m\frac{D_{m+1}^2 - t}{D_{m+1}^2 - D_m^2} + S_{m+1}\frac{t - D_m^2}{D_{m+1}^2 - D_m^2}$$

if $D_m^2 \le t < D_{m+1}^2$ and $0 \le m \le n - 1$. Set also

$$S_n(t) = \frac{s_n(D_n^2 t)}{\sqrt{2D_n^2 LLD_n^2}}$$

for $0 \le t \le 1$ and $n \in \mathbb{N}$.

Lemma 9. *Let $D(n) \to \infty$ as $n \to \infty$. Assume that there exists a sequence*

$$M_n = \mathrm{o}\left(\frac{D_n}{\sqrt{LLD_n}}\right)$$

such that

$$P\big(|Z_n| \le M_n\big) = 1$$

for each $n \ge 1$. Then

$$S_n(\cdot) \Rightarrow \mathcal{K} \quad (P\text{-a.s.}).$$

Proof. This is Major's Theorem [9]. □

We will apply Lemma 9 to $Z_j = a_j\big(\eta_j - (1-\mathrm{e}^{-\lambda_j})\big)$, where $\eta_j := \mathbf{1}\{\xi_j \ge 1\}$ and $1 \le j \le n$. Then $D_n^2 = B^2(n)$ and condition (6) will be at our disposal. To simplify the calculations, we introduce another sequence of additive functions

$$\tilde{h}(\sigma, m) := \sum_{j=1}^{m} a_j \mathbf{1}\{k_j(\sigma) \ge 1\}, \quad m \le n.$$

Let $\tilde{u}_m(\sigma, t)$ and $\widetilde{U}_m(\sigma, t)$ be the combinatorial processes defined as $u_m(\sigma, t)$ and $U_m(\sigma, t)$ using $\tilde{h}(\sigma, m)$ instead of $h(\sigma, m)$. Set also $Y_m = a_1\eta_1 + \cdots + a_m\eta_m$ for $1 \le m \le n$.

Lemma 10. *For an arbitrary* $\varepsilon > 0$,

$$\lim_{n_1\to\infty} \limsup_{n\to\infty} \mu_n\Big(\max_{n_1\le m\le n} \rho\big(\widetilde{U}_m(\sigma, \cdot), U_m(\sigma, \cdot)\big) \ge \varepsilon\Big) = 0. \tag{22}$$

Proof. If j and j' are the consecutive numbers from the set $I := \{j \le m: a_j \ne 0\}$, then, by virtue of the definition of $u_m(\sigma, t)$,

$$\max\left\{|\widetilde{U}_m(\sigma, t) - U_m(\sigma, t)| : \frac{B^2(j)}{B^2(m)} \le t \le \frac{B^2(j')}{B^2(m)}\right\} \le \beta^{-1}(m) \max\Big\{|\tilde{h}(\sigma, j) - h(\sigma, j)|, |\tilde{h}(\sigma, j') - h(\sigma, j')|\Big\}.$$

Hence

$$\begin{aligned}
\mu_n\Big(\max_{n_1\le m\le n} \rho\big(\widetilde{U}_m(\sigma, \cdot), U_m(\sigma, \cdot)\big) \ge \varepsilon\Big) &\le \mu_n\Big(\max_{n_1\le m\le n} \max_{j\in I} |\tilde{h}(\sigma, j) - h(\sigma, j)| \ge \varepsilon\beta(n_1)\Big) \\
&\le \mu_n\left(\sum_{j=1}^{n} |h_j(k_j(\sigma)) - a_j \cdot \mathbf{1}\{k_j(\sigma) \ge 1\}| \ge \varepsilon\beta(n_1)\right) \\
&\ll P^\theta\left(\sum_{j=1}^{n} |h_j(\xi_j) - a_j \cdot \mathbf{1}\{\eta_j \ge 1\}| \ge (\varepsilon/3)\beta(n_1)\right) + \mathrm{o}(1).
\end{aligned}$$

In the last step, we applied Corollary 2. In its turn, if $K > 2$ is arbitrary, the probability appearing on the right-hand side can be majorized by

$$P\big(\exists j \le K : \xi_j \ge K\big) + P\big(\exists j > K : \xi_j \ge 2\big) + P\left(\sum_{j\le K}\big(|h_j(\xi_j)| + |a_j|\eta_j\big) \ge (\varepsilon/3)\beta(n_1),\ 2 \le \xi_j \le K,\ \forall j \le K\right).$$

Since $\beta(n_1) \to \infty$ as $n_1 \to \infty$, the last probability is negligible. The first two of them do not exceed

$$\sum_{j\le K}\sum_{k\ge K} \frac{\mathrm{e}^{-\lambda_j}\lambda_j^k}{k!} + \sum_{j\ge K}\sum_{k\ge 2} \frac{\mathrm{e}^{-\lambda_j}\lambda_j^k}{k!} \ll K^{-1}.$$

Collecting the estimates, since K is arbitrary, we obtain the desired claim of Lemma 10.

In the sequel, we use only the functions $\tilde{h}(\sigma, m)$ and the processes $\widetilde{U}_m(\sigma, t)$ writing them without the "tilde". □

Lemma 11. *Let* $1 \le k \le n$, $0 < b_n \le b_{n-1} \le \cdots \le b_1$, *and* $\varepsilon > 0$ *be arbitrary. For* $h = \tilde{h}$, *if* $n \to \infty$, *we have*

$$\begin{aligned}
\mu_n\Big(\max_{k\le m\le n} b_m|h(\sigma, m) - A(m)| \ge \varepsilon\Big) &\ll P^\theta\Big(\max_{k\le m\le n} b_m|Y_m - A(m)| \ge \varepsilon/3\Big) + \mathrm{o}(1) \\
&\le 3^{2\theta}\varepsilon^{-2\theta}\left(b_k^2 B^2(k) + \sum_{k\le j\le n} b_j^2 a_j^2 \mathrm{e}^{-\lambda_j}(1 - \mathrm{e}^{-\lambda_j})\right)^\theta + \mathrm{o}(1).
\end{aligned}$$

Proof. The first estimate follows from Corollary 1 applied to $\mathbb{G} = \mathbb{R}^{n-r+1}$,

$$A = \Big\{(s_r, \ldots, s_n) \in \mathbb{R}^{n-r+1}: \max_{r\le m\le n} |s_m - A(m)| < \varepsilon/3\Big\},$$

and

$$h(\sigma) = \big(h(\sigma, r), \ldots, h(\sigma, n)\big).$$

The second inequality in Lemma 11 is just a partial case of Theorem 13 in Chapter III of [20].

The lemma is proved. □

Let r, $n_1 \le r \le n$, be a parameter, $q := \max\{j \in I : j \le r\}$, and

$$u_m^{(r)}(\sigma, t) = \begin{cases} u_m(\sigma, t) & \text{if } t \le B^2(q), \\ u_m(\sigma, B^2(q)) & \text{if } t > B^2(q). \end{cases}$$

Denote $U_m^{(r)}(\sigma, t) := u_m^{(r)}\big(\sigma, B^2(m)t\big)/\beta(m)$. Similarly, let

$$s_m^{(r)}(t) = \begin{cases} s_m(t) & \text{if } t \le B^2(q), \\ s_m\big(B^2(q)\big) & \text{if } t > B^2(q) \end{cases}$$

and $S_m^{(r)}(t) = s_m^{(r)}\big(tB(m)\big)/\beta(m)$.

Lemma 12. *There exists a sequence $r = r(n)$, $n_1 \le r = o(n)$, such that, for every $\varepsilon > 0$,*

$$\lim_{n_1\to\infty} \limsup_{n\to\infty} P\Big(\max_{n_1\le m\le n} \rho\big(S_m(\cdot), S_m^{(r)}(\cdot)\big) \ge \varepsilon\Big) = 0 \tag{23}$$

and

$$\lim_{n_1\to\infty} \limsup_{n\to\infty} \mu_n\Big(\max_{n_1\le m\le n} \rho\big(U_m(\sigma,\cdot), U_m^{(r)}(\sigma,\cdot)\big) \ge \varepsilon\Big) = 0.$$

Proof. If $P_{n_1,n}(\varepsilon)$ denotes the probability in (23) and $n_1 \le r \le n$, then

$$\begin{aligned} P_{n_1,n}(\varepsilon) &= P\Big(\max_{r\le m\le n} \rho\big(S_m(\cdot), S_m^{(r)}(\cdot)\big) \ge \varepsilon\Big) \\ &= P\left(\max_{r\le m\le n} \frac{1}{\beta(m)} \sup\big\{|s_m(t) - s_m(B^2(q))| : B^2(q) \le t \le B^2(m)\big\} \ge \varepsilon\right) \\ &\le P\Big(\max_{r< m\le n} \beta^{-1}(m)|(Y_m - A(m)) - (Y_r - A(r))| \ge \varepsilon\Big) \\ &\le \varepsilon^{-2}\frac{B^2(n) - B^2(r)}{\beta^2(r)} \end{aligned}$$

by the already mentioned Theorem 13 [20, Chapter III].

The same argument and Lemma 11 (applied in the case $a_j \equiv 0$ if $j \le r$) lead to the estimate

$$\begin{aligned} \mu_n\Big(\max_{n_1\le m\le n} \rho\big(U_m(\sigma,\cdot), U_m^{(r)}(\sigma,\cdot)\big) \ge \varepsilon\Big) &\le \mu_n\Big(\max_{r<m\le n} \beta^{-1}(m)|(h(\sigma,m) - A(q)) - (h(\sigma,r) - A(r))| \ge \varepsilon\Big) \\ &\ll P^{\theta}\Big(\max_{r<m\le n} \beta^{-1}(m)|(Y_m - A(m)) - (Y_r - A(r))| \ge (1/3)\varepsilon\Big) + o(1) \\ &\ll \left(\frac{B^2(n) - B^2(r)}{\beta^2(r)}\right)^{\theta} + o(1) \end{aligned}$$

as $n \to \infty$.

By condition (4), if r is sufficiently large, $r \le j \le n$, and δ, $0 < \delta < 1$, is arbitrary, then $|a_j| \le \delta B(n)/\sqrt{LLB(n)}$. Hence, taking $r = \delta n$ and applying condition (15), we obtain

$$B^2(n) - B^2(r) \ll \delta^2 \log\frac{1}{\delta}\frac{B^2(n)}{LLB(n)}.$$

We now choose $\delta = \delta_n = o(1)$ as $n \to \infty$ so that $\delta \ge 1/\sqrt{n}$. This implies $B^2(n) - B^2(r) = o(\beta^2(r))$. Having in mind the above estimates, we see that, with such an r, the probabilities in Lemma 12 vanish as $n \to \infty$ and $n_1 \to \infty$.

The lemma is proved. □

Proof of Theorem 1. By virtue of the definition of strong convergence and Lemma 12, it suffices to prove that

$$\lim_{n_1\to\infty} \limsup_{n\to\infty} \mu_n\Big(\max_{n_1\le m\le n} \rho\big(U_m^r(\sigma,\cdot), \mathcal{K}\big) \ge \varepsilon\Big) = 0$$

and

$$\lim_{n_1\to\infty} \liminf_{n\to\infty} \mu_n\Big(\min_{n_1\le m\le n} \rho\big(U_m^r(\sigma,\cdot), g\big) < \varepsilon\Big) = 1$$

for each function $g \in \mathcal{K}$ and $\varepsilon > 0$. Since here $r = r(n) \to \infty$ and $r = o(n)$, we can apply TV Approximation and substitute the frequencies by the appropriate probabilities for independent r.vs. Consequently, our task reduces to the proof of

$$\lim_{n_1\to\infty} \limsup_{n\to\infty} P\Big(\max_{n_1 \le m \le n} \rho\big(S_m^r(\cdot), \mathcal{K}\big) \ge \varepsilon \Big) = 0$$

and

$$\lim_{n_1\to\infty} \liminf_{n\to\infty} P\Big(\min_{n_1 \le m \le n} \rho\big(S_m^r(\cdot), g\big) < \varepsilon \Big) = 1.$$

Checking that the last relations follow from Lemmas 9 and 12, we complete the proof of Theorem 1. □

Acknowledgements

The author sincerely thanks the referees for their benevolent remarks which have helped to improve the exposition of this paper.

References

[1] R. Arratia, A.D. Barbour, S. Tavaré, Limits of logarithmic combinatorial structures, Ann. Probab. 28 (2000) 1620–1644.
[2] R. Arratia, A.D. Barbour, S. Tavaré, Logarithmic Combinatorial Structures: A Probabilistic Approach, in: EMS Monographs in Mathematics, EMS Publishing House, Zürich, 2003.
[3] G.J. Babu, E. Manstavičius, Brownian motion and random permutations, Sankhyā A 61 (1999) 312–327.
[4] G.J. Babu, E. Manstavičius, Limit processes with independent increments for the Ewens sampling formula, Ann. Inst. Statist. Math. 54 (2002) 607–620.
[5] W. Feller, The general form of the so-called law of the iterated logarithm, Trans. Amer. Math. Soc. 54 (1943) 373–402.
[6] D. Foata, La série génératrice exponentielle dans les problèmes d'énumération, Séminaire de Mathématiques Supérieures, Les Presses de l'Université de Montréal, Québec, 1974.
[7] D. Freedman, Brownian Motion and Diffusion, Holden-day, San Francisco, 1971.
[8] J. Hansen, Order statistics for decomposable combinatorial structures, Random Structures Algorithms 5 (1994) 517–533.
[9] P. Major, A note on Kolmogorov's law of iterated logarithm, Studia Sci. Math. Hungar. 12 (1977) 191–167.
[10] E. Manstavičius, The law of iterated logarithm for random permutations, Lith. Math. J. 38 (1998) 160–171.
[11] E. Manstavičius, Functional limit theorems for sequences of mappings on the symmetric group, in: A. Dubickas, et al. (Eds.), Analytic and Probab. Methods in Number Theory, TEV, Vilnius, 2002, pp. 175–187.
[12] E. Manstavičius, Mappings on decomposable combinatorial structures: analytic approach, Combin. Probab. Comput. 11 (2002) 61–78.
[13] E. Manstavičius, Conditional probabilities in combinatorics. The cost of dependence, in: M. Hušková, M. Janžura (Eds.), Prague Stochstics 2006, Matfyzpress, Charles University, Prague, 2006, pp. 523–532.
[14] E. Manstavičius, An analytic method in probabilistic combinatorics, Osaka J. Math. 46 (2009) 273–290.
[15] E. Manstavičius, A functional limit theorem on powers of random permutations, Lith. Math. J. 49 (2009) 297–308.
[16] E. Manstavičius, Total variation approximation for random assemblies and a functional limit theorem, Monatsh. Math. 161 (2010) 313–334.
[17] E. Manstavičius, J. Norkūnienė, An analogue of Feller's theorem for logarithmic combinatorial assemblies, Lith. Math. J. 48 (2008) 405–417.
[18] J. Norkūnienė, The law of iterated logarithm for combinatorial assemblies, Lith. Math. J. 46 (2006) 432–445.
[19] J. Norkūnienė, The Strassen law of iterated logarithm for combinatorial assemblies, Lith. Math. J. 47 (2007) 176–183.
[20] V.V. Petrov, Sums of independent random variables, in: Ergebnisse der Mathematik und ihrer Grenzgebiete, B, vol. 82, Springer, New York, 1975.
[21] I.Z. Ruzsa, Generalized moments of additive functions, J. Number Theory 18 (1984) 27–33.
[22] D. Stark, Explicit limits of total variation distance in approximations of random logarithmic assemblies by related Poisson processes, Combin. Probab. Comput. 6 (1997) 87–105.
[23] D. Stark, Total variation asymptotics for related Poisson process approximations of random logarithmic assemblies, Combin. Probab. Comput. 8 (1999) 567–598.
[24] V. Zacharovas, The convergence rate to the normal law of a certain variable defined on random polynomials, Lith. Math. J. 42 (2002) 88–107.
[25] V. Zacharovas, Distribution of the logarithm of the order of a random permutation, Lith. Math. J. 44 (2004) 296–327.
[26] V. Zacharovas, Distribution of random variables on the symmetric group, Doctorial Dissertation, Vilnius University 2004. arXiv:0901.1733.