Bernoulli **24**(2), 2018, 1033–1052
DOI: 10.3150/16-BEJ847

Functional central limit theorems in $L^2(0, 1)$ for logarithmic combinatorial assemblies

KOJI TSUKUDA

Biostatistics Center, Kurume University, 67 Asahi-machi, Kurume-city, Fukuoka, 830-0011, Japan.
E-mail: k.tsukuda@gmail.com

Functional central limit theorems in $L^2(0, 1)$ for logarithmic combinatorial assemblies are presented. The random elements argued in this paper are viewed as elements taking values in $L^2(0, 1)$ whereas the Skorokhod space is argued as a framework of weak convergences in functional central limit theorems for random combinatorial structures in the literature. It enables us to treat other standardized random processes which converge weakly to a corresponding Gaussian process with additional assumptions.

Keywords: functional central limit theorem; logarithmic assembly; Poisson approximation; random mappings; the Ewens sampling formula

1. Introduction

The aim of this paper is proving functional central limit theorems (FCLTs) in $L^2(0, 1)$ for logarithmic assemblies via Poisson approximations. Let $\{Z_j\}_{j=1}^{\infty}$ be a sequence of independent Poisson random variables with $\mathsf{E}[Z_j] = \lambda_j$ for all $j = 1, 2, \dots$, and consider a sequence of random variables $\{C_j^n\}_{j=1}^{\infty}$ for a positive integer n whose law is determined by the conditioning relation

$$\mathsf{P}\left[C_1^n = c_1, \dots, C_n^n = c_n\right] = \mathsf{P}\left[Z_1 = c_1, \dots, Z_n = c_n \,\middle|\, \sum_{j=1}^{n} jZ_j = n\right] \tag{1.1}$$

for $1 \le j \le n$ and $C_j^n = 0$ for $j \ge n+1$. It means that, considering assemblies, C_j^n denotes the number of components whose sizes are j for $1 \le j \le n$, so they are called component counts of a partition (see, e.g., Arratia *et al.* [3]). For these component counts, under some conditions for $\{\lambda_j\}_{j=1}^{\infty}$, the following FCLT holds (Arratia *et al.* [3], page 1354): the random process

$$\left(\frac{\sum_{j=1}^{[n^u]} C_j^n - \theta u \log n}{\sqrt{\theta \log n}}\right)_{0 \le u \le 1} \tag{1.2}$$

converges weakly to the standard Brownian motion $(B(u))_{0 \le u \le 1}$ in the Skorokhod space $D[0, 1]$ as $n \to \infty$, where θ is some positive constant. In the present paper, we shall establish an alternative FCLT. That is, we shall show that the random process

$$X_n(\cdot) = \left(\frac{\sum_{j=1}^{[n^u]} C_j^n - \sum_{j=1}^{[n^u]} \lambda_j}{\sqrt{\sum_{j=1}^{[n^u]} \lambda_j}}\right)_{0 < u < 1} \tag{1.3}$$

converges weakly to

$$G(\cdot) = \left(\frac{B(u)}{\sqrt{u}}\right)_{0<u<1}$$

in $L^2(0,1)$ as $n \to \infty$. The difference between (1.2) and (1.3) is the standardization. We need to note that our theory requires a set of stronger conditions that stems from the difference in the denominator of (1.3). Moreover, we shall show that the random process

$$X'_n(\cdot) = \left(\left(\frac{\sum_{j=1}^{[n^u]} C_j^n - \theta u \log n}{\sqrt{\theta u \log n}}\right) 1\left\{\frac{\varepsilon}{\log n} < u < 1\right\}\right)_{0<u<1} \tag{1.4}$$

converges weakly to $G(\cdot)$ in $L^2(0,1)$ as $n \to \infty$, where the notation $1\{\cdot\}$ denotes the indicator function and ε is a positive constant.

It is fruitful to show the convergence to 0 of the total variation distance between the laws of component counts and of independent random variables; see, for example, Arratia and Tavaré [4] and Arratia *et al.* [2]. One of existing results is the weak convergence of the random process (1.2) in $D[0,1]$ stated above. FCLTs for logarithmic combinatorial structures in $D[0,1]$ are originally proved for specific structures: for random permutations by DeLaurentis and Pittel [6], for random mappings by Hansen [10] and for the Ewens sampling formula by Hansen [11]. The proof of Hansen [10,11] are by a direct way to check the tightnesses and convergences of finite dimensional marginal distributions. Arratia and Tavaré [4] proved the functional central limit theorems for the Ewens sampling formula and random mappings through elegant ways via Poisson approximations. Arratia *et al.* [2] proved that it is possible to apply such strategy to general logarithmic combinatorial structures. On the other hand, this paper supplies a new result to total variation distance results for random combinatorial structures. The proof strategy is to approximate $\sum_{j=1}^{[n^u]} C_j^n$ by a Poisson process and use a FCLT in $L^2(0,1)$ for Poisson processes. The proofs of the approximations are based on the asymptotic result about the total variation distance, which has been already established.

Let us show some notations. We shall mainly argue asymptotic behaviors when n tends to infinity and denote a convergence in probability and a weak convergence by $\to^p$ and $\Rightarrow$, respectively. Let us denote by $A =^d B$ that the laws of random elements A and B are the same. The notations $a \wedge b$ and $a \vee b$ for real numbers a, b mean $\min(a,b)$ and $\max(a,b)$, respectively. Consider the inner product

$$\langle z_1, z_2\rangle_{L^2} = \int_0^1 z_1(u) z_2(u)\, du,$$

where $z_1(\cdot)$ and $z_2(\cdot)$ are real valued functions on $(0,1)$. Introduce $L^2(0,1)$ as equivalence classes of square integrable real valued functions on $(0,1)$, that is, the set of all measurable functions $z : (0,1) \to \mathbf{R}$ which satisfy $\|z\|_{L^2} = \sqrt{\langle z, z\rangle_{L^2}} < \infty$. This space is a separable Hilbert space with respect to L^2 distance $\|z_1 - z_2\|_{L^2}$.

In Section 2, a FCLT in $L^2(0,1)$ for Poisson processes is prepared. This result shall be used in Section 5 and Section 7. In Section 3, Poisson approximations in $L^2(0,1)$ for $X_n(\cdot)$ are presented.

The sufficient condition appearing in the precedent sections are discussed in Section 4. Section 5 is devoted to showing a FCLT in $L^2(0,1)$ which asserts the weak convergence of $X_n(\cdot)$ based on the results given in the previous two sections. This result is applied to two important examples: the Ewens sampling formula and random mappings in Section 6. Another FCLT in $L^2(0,1)$, which asserts the weak convergence of $X_n'(\cdot)$, is presented in Section 7.

2. A FCLT in $L^2(0,1)$ for Poisson processes

There is a FCLT in $D[0,1]$ for Poisson processes, which is used to prove the weak convergence of (1.2), see the proof of Theorem 5 in Arratia and Tavaré [4]. First, let us prove a FCLT in $L^2(0,1)$ for Poisson processes with another standardization, by which the limit becomes $(B(u)/\sqrt{u})_{0<u<1}$.

Lemma 2.1. *Let $(N_t)_{t\geq 0}$ be a homogeneous Poisson process which satisfies $N_0 = 0$ and its intensity is $\lambda > 0$. Define the non-decreasing function $u \mapsto s_n(u)$ for $0 \leq u \leq 1$ and $n = 1, 2, \ldots$ which satisfies*

$$\inf_{u\in(\tau,1)} s_n(u) > 0$$

for all $0 < \tau < 1$,

$$\lim_{n\to\infty}\left(\frac{\sup_{u\in(0,1)}|s_n(u) - Ku\log n|}{\log n}\right) = 0 \tag{2.1}$$

for some positive constant K and

$$\lim_{n\to\infty}\left(\int_0^1 \frac{du}{(s_n(u))^\delta}\right) = 0 \tag{2.2}$$

for some $\delta > 0$. Then, the random process

$$\mathcal{M}_n(\cdot) = \left(\frac{N_{s_n(u)} - \lambda s_n(u)}{\sqrt{\lambda s_n(u)}}\right)_{0<u<1}$$

converges weakly to the Gaussian process $G(\cdot) = (B(u)/\sqrt{u})_{0<u<1}$ in $L^2(0,1)$ as $n \to \infty$, where $B(\cdot)$ is the standard Brownian motion.

Proof. First of all, it holds that

$$\mathcal{M}_n(u) = \frac{1}{\sqrt{\lambda}}\int_0^{s_n(1)} \frac{1\{t \leq s_n(u)\}}{\sqrt{s_n(u)}}(dN_t - \lambda\, dt)$$

for any $u \in (0,1)$. Since it holds that

$$\mathsf{E}\big[\|\mathcal{M}_n\|_{L^2}^2\big] = 1,$$

$\mathcal{M}_n(\cdot)$ almost surely takes its value in $L^2(0,1)$. For the proof, we use Theorem 1.8.4 in the book of van der Vaart and Wellner [15] which is based on the tightness criterion by Prokhorov [12].

(A) Convergence of the inner product. Fix an arbitrary $h \in L^2(0,1)$. The Fubini theorem yields that

$$\langle \mathcal{M}_n, h\rangle_{L^2} = \frac{1}{\sqrt{\lambda}} \int_0^{s_n(1)} \left(\int_0^1 \frac{1\{t \le s_n(u)\}}{\sqrt{s_n(u)}} h(u)\, du \right) (dN_t - \lambda\, dt).$$

The process

$$\left(\frac{1}{\sqrt{\lambda}} \int_0^{s} \left(\int_0^1 \frac{1\{t \le s_n(u)\}}{\sqrt{s_n(u)}} h(u)\, du \right) (dN_t - \lambda\, dt) \right)_{0 \le s \le s_n(1)}$$

is a martingale relative to the filtration $\mathcal{F}_s = \{\mathcal{F}_0 \vee \sigma(N_t; t \in (0,s))\}$, where $\mathcal{F}_0$ is a σ-field which is independent of $\sigma(N_t; t \in (0,s))$, with the predictable quadratic variation process

$$\begin{aligned}
&\left(\frac{1}{\lambda} \int_0^{s} \left(\int_0^1 \frac{1\{t \le s_n(u)\}}{\sqrt{s_n(u)}} h(u)\, du \right)^2 \lambda\, dt \right)_{0 \le s \le s_n(1)} \\
&= \left(\int_0^{s} \left(\int_0^1 \int_0^1 \frac{1\{t \le s_n(u)\} 1\{t \le s_n(v)\}}{\sqrt{s_n(u) s_n(v)}} h(u) h(v)\, du\, dv \right) dt \right)_{0 \le s \le s_n(1)}.
\end{aligned}$$

So, the predictable quadratic variation of $\langle \mathcal{M}_n, h\rangle_{L^2}$, denoted by $\langle\langle \mathcal{M}_n, h\rangle_{L^2}\rangle$, is

$$\int_0^1 \int_0^1 \left(\frac{s_n(u) \wedge s_n(v)}{\sqrt{s_n(u) s_n(v)}} \right) h(u) h(v)\, du\, dv.$$

As for the integrand, it holds that

$$\left| \frac{s_n(u) \wedge s_n(v)}{\sqrt{s_n(u) s_n(v)}} h(u) h(v) \right| \le \left| h(u) h(v) \right|,$$

$$\int_0^1 \int_0^1 \left| h(u) h(v) \right| du\, dv \le \int_0^1 h(u)^2\, du < \infty$$

and

$$\lim_{n\to\infty} \left(\frac{s_n(u) \wedge s_n(v)}{\sqrt{s_n(u) s_n(v)}} \right) = \lim_{n\to\infty} \left(\frac{K((u \log n) \wedge (v \log n))}{\sqrt{K^2 \log n^u \log n^v}} \right) = \frac{u \wedge v}{\sqrt{uv}}.$$

Thence, the dominated convergence theorem yields that

$$\lim_{n\to\infty} \langle\langle \mathcal{M}_n, h\rangle_{L^2}\rangle = \int_0^1 \int_0^1 \frac{(u \wedge v) h(u) h(v)}{\sqrt{uv}}\, du\, dv$$

and it is equal to

$$\int_0^1 \int_0^1 \frac{\mathsf{E}[B(u)B(v)]h(u)h(v)}{\sqrt{uv}}\,du\,dv = \mathsf{E}\big[\langle G, h\rangle_{L^2}^2\big].$$

Let us check the Lyapunov type condition. The Schwarz inequality yields that

$$\begin{aligned}
&\int_0^{s_n(1)} \frac{1}{\lambda^{1+\delta}} \left(\int_0^1 \frac{1\{t \le s_n(u)\}}{\sqrt{s_n(u)}} h(u)\,du \right)^{2+2\delta} \lambda\,dt \\
&\quad \le \int_0^{s_n(1)} \frac{1}{\lambda^{\delta}} \left(\int_0^1 \frac{1\{t \le s_n(u)\}}{s_n(u)}\,du \right)^{1+\delta} \left(\int_0^1 h(u)^2\,du \right)^{1+\delta} dt \\
&\quad \le \int_0^{s_n(1)} \frac{1}{\lambda^{\delta}} \int_0^1 \frac{1\{t \le s_n(u)\}}{(s_n(u))^{1+\delta}}\,du \left(\int_0^1 h(u)^2\,du \right)^{1+\delta} dt.
\end{aligned}$$

The right-hand side is equal to

$$\begin{aligned}
&\frac{1}{\lambda^{\delta}} \int_0^1 \int_0^{s_n(1)} 1\big\{t \le s_n(u)\big\}\,dt \frac{1}{(s_n(u))^{1+\delta}}\,du \left(\int_0^1 h(u)^2\,du \right)^{1+\delta} \\
&\quad = \frac{1}{\lambda^{\delta}} \int_0^1 \frac{1}{(s_n(u))^{\delta}}\,du \left(\int_0^1 h(u)^2\,du \right)^{1+\delta}.
\end{aligned}$$

It converges to 0 as $n \to \infty$ by the assumption. Therefore, the convergence of the inner product, $\langle \mathcal{M}_n, h\rangle_{L^2} \Rightarrow \langle G, h\rangle_{L^2}$ as $n \to \infty$, is proved by the martingale CLT.

(B) Asymptotic tightness. It is sufficient to prove

$$\lim_{J\to\infty} \limsup_{n\to\infty} \mathsf{E}\left[\sum_{j>J} \langle \mathcal{M}_n, e_j\rangle_{L^2}^2 \right] = 0,$$

where $\{e_j(\cdot)\}_{j=1}^{\infty}$ is a complete orthonormal system of $L^2(0,1)$. The Fubini theorem yields that

$$\langle \mathcal{M}_n, e_j\rangle_{L^2} = \frac{1}{\sqrt{\lambda}} \int_0^{s_n(1)} \left(\int_0^1 \frac{1\{t \le s_n(u)\}}{\sqrt{s_n(u)}} e_j(u)\,du \right) (dN_t - \lambda\,dt).$$

It holds that

$$\limsup_{n\to\infty} \mathsf{E}\left[\sum_{j>J} \langle \mathcal{M}_n, e_j\rangle_{L^2}^2 \right] = \limsup_{n\to\infty} \mathsf{E}\left[\|\mathcal{M}_n\|_{L^2}^2 - \sum_{j=1}^{J} \langle \mathcal{M}_n, e_j\rangle_{L^2}^2 \right]. \tag{2.3}$$

For the first term of the integrand in the right-hand side of (2.3), it holds that

$$\limsup_{n\to\infty} \mathsf{E}\big[\|\mathcal{M}_n\|_{L^2}^2\big] = 1,$$

and, for the second term, it holds that

$$
\begin{aligned}
0 &\geq \limsup_{n\to\infty} \mathsf{E}\left[-\sum_{j=1}^{J}\langle \mathcal{M}_n, e_j\rangle_{L^2}^2\right] \\
&= -\liminf_{n\to\infty} \mathsf{E}\left[\sum_{j=1}^{J}\langle \mathcal{M}_n, e_j\rangle_{L^2}^2\right] \\
&= -\liminf_{n\to\infty} \sum_{j=1}^{J} \frac{1}{\lambda}\int_0^{s_n(1)} \left(\int_0^1 \frac{1\{t \leq s_n(u)\}}{\sqrt{s_n(u)}} e_j(u)\, du\right)^2 \lambda\, dt \qquad (2.4) \\
&= -\liminf_{n\to\infty} \sum_{j=1}^{J} \int_0^1 \int_0^1 \int_0^{s_n(1)} \frac{1\{t \leq (s_n(u)\wedge s_n(v))\}}{\sqrt{s_n(u)s_n(v)}}\, dt\, e_j(u)e_j(v)\, du\, dv \\
&= -\liminf_{n\to\infty} \int_0^1 \int_0^1 \frac{s_n(u)\wedge s_n(v)}{\sqrt{s_n(u)s_n(v)}} \left(\sum_{j=1}^{J} e_j(u)e_j(v)\right) du\, dv.
\end{aligned}
$$

Applying the Fatou–Lebesgue theorem, the right-hand side of (2.4) is bounded above by

$$
-\int_0^1 \int_0^1 \liminf_{n\to\infty} \frac{s_n(u)\wedge s_n(v)}{\sqrt{s_n(u)s_n(v)}} \left(\sum_{j=1}^{J} e_j(u)e_j(v)\right) du\, dv, \qquad (2.5)
$$

since

$$
\left|\frac{s_n(u)\wedge s_n(v)}{\sqrt{s_n(u)s_n(v)}} \left(\sum_{j=1}^{J} e_j(u)e_j(v)\right)\right| \leq \sum_{j=1}^{J} \left|e_j(u)e_j(v)\right|
$$

and

$$
\int_0^1 \int_0^1 \sum_{j=1}^{J} \left|e_j(u)e_j(v)\right| du\, dv \leq \sum_{j=1}^{J} \int_0^1 \left(e_j(u)\right)^2 du = J < \infty.
$$

By the condition (2.1), (2.5) is equal to

$$
-\int_0^1 \int_0^1 \frac{u\wedge v}{\sqrt{uv}} \left(\sum_{j=1}^{J} e_j(u)e_j(v)\right) du\, dv = -\mathsf{E}\left[\sum_{j=1}^{J}\langle G, e_j\rangle_{L^2}^2\right].
$$

The Bessel inequality yields that

$$
\sum_{j=1}^{J}\langle G, e_j\rangle_{L^2}^2 \leq \|G\|_{L^2}^2,
$$

so the dominated convergence theorem yields that

$$\lim_{J\to\infty} \mathsf{E}\left[\sum_{j=1}^{J} \langle G, e_j\rangle_{L^2}^2\right] = \mathsf{E}\left[\sum_{j=1}^{\infty} \langle G, e_j\rangle_{L^2}^2\right] = \mathsf{E}\big[\|G\|_{L^2}^2\big] = 1.$$

Hence, (2.3) converges to 0 as $J \to \infty$.

Because (A) and (B) hold, the conclusion follows from the Theorem 1.8.4 of van der Vaart and Wellner [15]. This completes the proof. □

Consider the case

$$s_n(u) = \sum_{j=1}^{[n^u]} \lambda_j,$$

for $0 \le u \le 1$. The condition (2.1) shall be discussed in Section 4. The condition (2.2) for $\delta = 1/2$ can be written by

$$\int_0^1 \frac{du}{\sqrt{\sum_{j=1}^{[n^u]} \lambda_j}} = \sum_{k=1}^{n-1} \int_{\frac{\log k}{\log n}}^{\frac{\log k+1}{\log n}} \frac{du}{\sqrt{\sum_{j=1}^{k} \lambda_j}} = \frac{1}{\log n} \sum_{k=1}^{n-1} \frac{\log(k+1) - \log k}{\sqrt{\sum_{j=1}^{k} \lambda_j}} \to 0,$$

where the equalities hold for $n \ge 2$. Nevertheless, it may be difficult to check this condition. However, for example, by the following proposition, we easily verify the condition (2.2) if $j\lambda_j$ is a positive constant (the Ewens sampling formula) or if $j\lambda_j$ is non-decreasing with respect to j and $\lambda_1 > 0$ (e.g., random mappings), see Section 6.

Proposition 2.1. *Consider a sequence* $\{\lambda_j\}_{j=1}^{\infty}$ *of positive real numbers. If*

$$L = \inf_{j=1,2,\ldots} j\lambda_j > 0$$

holds, then it holds that

$$\lim_{n\to\infty} \left(\int_0^1 \frac{du}{\sqrt{\ell_n'(u)}} \right) = 0, \tag{2.6}$$

where

$$\ell_n'(u) = \sum_{j=1}^{[n^u]} \lambda_j \qquad (0 \le u \le 1).$$

Proof. It holds that

$$\ell_n'(u) = \sum_{j=1}^{[n^u]} \frac{1}{j} j\lambda_j > L \sum_{j=1}^{[n^u]} \frac{1}{j} > L \log\big(\big[n^u\big] + 1\big) > Lu \log n,$$

for any $0 \le u \le 1$. Therefore, it follows that

$$0 < \int_0^1 \frac{du}{\sqrt{\ell_n'(u)}} < \frac{1}{\sqrt{L \log n}} \int_0^1 \frac{du}{\sqrt{u}} = \frac{2}{\sqrt{L \log n}} \to 0$$

as $n \to \infty$. This completes the proof. □

3. Poisson approximations

Define the total variation distance $d_{\mathrm{TV}}(X, Y)$ between the laws of random variables X and Y which take their values in finite or countably infinite space S by

$$d_{\mathrm{TV}}(X, Y) = \sup_{A \subset S} \big| \mathsf{P}(X \in A) - \mathsf{P}(Y \in A) \big|.$$

It holds that

$$d_{\mathrm{TV}}(X, Y) = \frac{1}{2} \sum_{s \in S} \big| \mathsf{P}(X = s) - \mathsf{P}(Y = s) \big|.$$

Define

$$T_{mn} = \sum_{j=m+1}^{n} j Z_j \qquad (0 \le m < n),$$

and

$$d_b(n) = d_{\mathrm{TV}}\big((C_1^n, \dots, C_b^n), (Z_1, \dots, Z_b) \big)$$

for $1 \le b \le n$, where $\{C_j^n\}_{j=1}^n$ and $\{Z_j\}_{j=1}^n$ are sequences of random variables introduced in Section 1. Then,

$$d_b(n) = \frac{1}{2} \sum_{r=0}^{\infty} \mathsf{P}[T_{0b} = r] \left| 1 - \frac{\mathsf{P}[T_{bn} = n - r]}{\mathsf{P}[T_{0n} = n]} \right|$$

holds, which is the equation (33) of Arratia and Tavaré [5]. Asymptotic properties of $d_b(n)$ have been already established, see, for example, Arratia *et al.* [2,3]. The following lemma, which guarantees a Poisson approximation in $L^2(0, 1)$

$$\left\| X_n(\cdot) - \frac{\sum_{j=1}^{[n \cdot]} Z_j - \sum_{j=1}^{[n \cdot]} \lambda_j}{\sqrt{\sum_{j=1}^{[n \cdot]} \lambda_j}} \right\|_{L^2}^2 \to^p 0,$$

holds by the convergence of $d_b(n)$ to 0 with some additional conditions, where $X_n(\cdot)$ is defined in (1.3).

Lemma 3.1. *Consider an assembly. Let*

$$g_\theta(n) = \sup_{u\in(0,1)} \left| \sum_{j=1}^{[n^u]} \left(\frac{\theta}{j} - \lambda_j \right) \right| = \sup_{u\in(0,1)} \left| \ell_n(u) - \ell_n'(u) \right|, \tag{3.1}$$

where

$$\ell_n(u) = \sum_{j=1}^{[n^u]} \frac{\theta}{j}, \qquad \ell_n'(u) = \sum_{j=1}^{[n^u]} \lambda_j,$$

for $0 \le u \le 1$. *Assume* $\lambda_1 > 0$ *and*

$$\lim_{n\to\infty} \left(\frac{g_\theta(n) \log\log n}{\log n} \right) = 0 \tag{3.2}$$

for some $\theta > 0$. *If*

$$\sup_{i\ge1} i\lambda_i < \infty, \qquad \liminf_{i\to\infty} i\lambda_i > 0$$

and if

$$\lim_{n\to\infty} d_b(n) = 0 \tag{3.3}$$

for some $b = b(n)$ *such that* $b/n = o(1)$ *and*

$$\lim_{n\to\infty} \left(\log\left(\frac{n}{b}\right) \sqrt{\frac{(g_\theta(n)+1)\log\log n}{\log n}} \right) = 0,$$

then it holds that

$$\int_0^1 \left| \frac{\sum_{j=1}^{[n^u]} (C_j^n - Z_j)}{\sqrt{\ell_n'(u)}} \right|^2 du \to^p 0.$$

Proof. We have

$$\begin{aligned}
0 &\le \int_0^1 \left| \frac{\sum_{j=1}^{[n^u]} (C_j^n - Z_j)}{\sqrt{\ell_n'(u)}} \right|^2 du \\
&= \int_0^1 \left| \frac{\sum_{j=1}^{[n^u]} (C_j^n - Z_j)}{\sqrt{\ell_n(u)}} \right|^2 \left(1 + \frac{\ell_n(u) - \ell_n'(u)}{\ell_n'(u)} \right) du \\
&\le \int_0^1 \left| \frac{\sum_{j=1}^{[n^u]} (C_j^n - Z_j)}{\sqrt{\ell_n(u)}} \right|^2 \left(1 + \frac{\sup_{u\in(0,1)} (\ell_n(u) - \ell_n'(u))}{\inf_{u\in(0,1)} \ell_n'(u)} \right) du \\
&\le \left(1 + \frac{g_\theta(n)}{\lambda_1} \right) \int_0^1 \left| \frac{\sum_{j=1}^{[n^u]} (C_j^n - Z_j)}{\sqrt{\ell_n(u)}} \right|^2 du.
\end{aligned} \tag{3.4}$$

The right factor of the right-hand side of (3.4) is evaluated by

$$
\begin{aligned}
&\int_0^1 \left| \frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\ell_n(u)}} \right|^2 du \\
&\quad = \int_0^1 \left| \frac{\sum_{j=1}^{n} 1\{j \le n^u\}(C_j^n - Z_j)}{\sqrt{\ell_n(u)}} \right|^2 du \\
&\quad = \int_0^1 \sum_{j=1}^{n}\sum_{k=1}^{n} \frac{1\{j \le n^u\}(C_j^n - Z_j)1\{k \le n^u\}(C_k^n - Z_k)}{\ell_n(u)}\, du \qquad (3.5)\\
&\quad = \sum_{j=1}^{n}\sum_{k=1}^{n} \int_0^1 \frac{1\{(j \vee k) \le n^u\}(C_j^n - Z_j)(C_k^n - Z_k)}{\ell_n(u)}\, du \\
&\quad \le \sum_{j=1}^{n}\sum_{k=1}^{n} \int_0^1 \frac{1\{(j \vee k) \le n^u\}}{\ell_n(u)}\, du \big|(C_j^n - Z_j)(C_k^n - Z_k)\big|.
\end{aligned}
$$

For $j = k = 1$, it holds that

$$
\begin{aligned}
\int_0^1 \frac{1\{(j \vee k) \le n^u\}}{\ell_n(u)}\, du &= \int_0^1 \frac{1}{\ell_n(u)}\, du \\
&< \int_0^{\frac{1}{\log n}} \frac{1}{\theta \log(1 + [n^u])}\, du + \int_{\frac{1}{\log n}}^1 \frac{1}{u\theta \log n}\, du \\
&< \frac{1}{\theta \log n}\left(\frac{1}{\log 2} + \log\log n\right)
\end{aligned}
$$

since $\ell_n(u) > \theta \log([n^u] + 1) > \theta u \log n$ for any $u \in (0, 1)$. For other (j, k), it holds that

$$
\begin{aligned}
\int_0^1 \frac{1\{(j \vee k) \le n^u\}}{\ell_n(u)}\, du &= \int_{\frac{\log(j\vee k)}{\log n}}^1 \frac{1}{\ell_n(u)}\, du \\
&< \int_{\frac{\log(j\vee k)}{\log n}}^1 \frac{1}{u\theta \log n}\, du \\
&= \frac{1}{\theta \log n}\big(\log\log n - \log\log(j \vee k)\big).
\end{aligned}
$$

It yields the bound for the right-hand side of (3.5)

$$
\begin{aligned}
&\frac{1}{\theta \log n}\left(\frac{1}{\log 2} + \log\log n\right)\big|(C_1^n - Z_1)(C_1^n - Z_1)\big| \\
&\quad - \frac{\log\log 2}{\theta \log n}\big|(C_2^n - Z_2)\big(2(C_1^n - Z_1) + (C_2^n - Z_2)\big)\big|
\end{aligned}
$$

$$+\sum_{j=1}^{n}\sum_{k=2}^{n}\frac{\log\log n}{\theta\log n}\left|\left(C_j^n - Z_j\right)\left(C_k^n - Z_k\right)\right|$$
$$< \frac{2}{\theta\log n}\left(\frac{1}{\log 2} - \log\log 2 + \log\log n\right)\left(\sum_{j=1}^{n}\left|C_j^n - Z_j\right|\right)^2.$$

So, it is sufficient to prove that

$$\sqrt{\left(1+\frac{g_\theta(n)}{\lambda_1}\right)\frac{\log\log n}{\log n}}\sum_{j=1}^{n}\left|C_j^n - Z_j\right| \to^p 0.$$

It follows from the next lemma. This completes the proof. □

The following lemma was used in the proof of the previous lemma. It is slightly different from Lemma 2 of Arratia *et al.* [3], though the proof is essentially the same.

Lemma 3.2. *Consider an assembly which satisfies*

$$\sup_{i\geq 1} i\lambda_i < \infty, \qquad \liminf_{i\to\infty} i\lambda_i > 0, \tag{3.6}$$

$$\lim_{n\to\infty} d_b(n) = 0 \tag{3.7}$$

for some $b = b(n)$ *such that* $b/n = o(1)$ *and*

$$\lim_{n\to\infty}\left(\log\left(\frac{n}{b}\right)\sqrt{\frac{f(n)}{\log n}}\right) = 0,$$

where $f(n)/\log n = o(1)$. *Then, there is a coupling satisfying*

$$\sqrt{\frac{f(n)}{\log n}}\sum_{j=1}^{n}\left|C_j^n - Z_j\right| \to^p 0$$

as $n \to \infty$.

Proof. For any $1 \leq b = b(n) \leq n$, the triangle inequality yields that

$$\sum_{j=1}^{n}\left|C_j^n - Z_j\right| \leq \sum_{j=1}^{b}\left|C_j^n - Z_j\right| + \sum_{j=b+1}^{n} C_j^n + \sum_{j=b+1}^{n} Z_j.$$

As for the first term in the right-hand side, for any $\epsilon > 0$, it holds that

$$\mathsf{P}\left[\sum_{j=1}^{b}\left|C_j^n - Z_j\right| > \epsilon\right] \leq \mathsf{P}\left[\left(C_1^n, \ldots, C_b^n\right) \neq (Z_1, \ldots, Z_b)\right] = d_b(n) \to 0$$

by the assumption (3.7). Both of the expectation of the second term and third term are $O(\log(n/b))$ by the assumptions (3.6), see the proof of Lemma 2 in Arratia *et al.* [3]. This completes the proof. $\square$

4. On logarithmic conditions

When $\lambda_j = \theta/j$ for all $j = 1, 2, \ldots$ with some $\theta > 0$, the conditioning relation (1.1) yields that

$$\mathsf{P}\big[\big(C_1^n, \ldots, C_n^n\big) = (c_1, \ldots, c_n)\big] = \frac{n!}{(\theta)_n} \prod_{j=1}^{n} \left(\frac{\theta}{j}\right)^{c_j} \frac{1}{c_j!} 1\left\{\sum_{j=1}^{n} j c_j = n\right\}, \tag{4.1}$$

where $(\theta)_n$ denotes $\theta \times (\theta+1) \times \cdots \times (\theta+n-1)$. This law is the Ewens sampling formula (ESF) which appears firstly in Ewens [8]. Some random structures can be regarded as "perturbations" of ESF in some sense. The meaning of "perturbations" is, for example, the logarithmic condition

$$\lim_{j\to\infty} j\mathsf{P}[Z_j = 1] = \lim_{j\to\infty} j\mathsf{E}[Z_j] = \theta \tag{4.2}$$

see Arratia and Tavaré [5] and Arratia *et al.* [2], or an approximation of a generating function of the sequence $\{\lambda_j\}_{j=1}^{\infty}$ to ESF near the singularity, see Flajolet and Soria [9] and Arratia *et al.* [3]. Note that the conditions (3.6) and (3.7) in Lemma 3.2 follow from the logarithmic condition (4.2), see Theorem 3.1 of Arratia *et al.* [2] and Section 4.1 of Arratia *et al.* [3]. There exists a unified approach by Arratia *et al.* [2] for structures satisfying the uniform logarithmic condition (ULC)

$$|\varepsilon_{il}| \le e(i) c_l \qquad (l = 1, 2, \ldots),$$

$$\lim_{i\to\infty} e(i) = 0, \qquad D_1 = \sum_{l=1}^{\infty} l c_l < \infty,$$

where $\varepsilon_{i1} = i\mathsf{P}[Z_i = 1] - \theta$ and $\varepsilon_{il} = i\mathsf{P}[Z_i = l]$ for $l = 2, 3, \ldots$. In particular, assemblies, multisets, and selections satisfying the logarithmic condition also satisfy the uniform logarithmic condition (Arratia *et al.* [2], Proposition 1.1). In this paper, the magnitude of the "perturbation" is measured by $g_\theta(n)$ which is defined in (3.1) and it also relates to the condition (2.1) of Lemma 2.1.

It follows from $|\lambda_j - \theta/j| = |\sum_{l=1}^{\infty} l\varepsilon_{jl}/j| \le D_1 e(j)/j$ for any $j = 1, \ldots, n$ that

$$\sum_{j=1}^{n} \left|\lambda_j - \frac{\theta}{j}\right| \le D_1 \sum_{j=1}^{n} \frac{e(j)}{j}.$$

If $\sum_{j=1}^{\infty} e(j)/j < \infty$ holds, then the series $\sum_{j=1}^{\infty} (\theta/j - \lambda_j)$ is absolutely convergent, so $g_\theta(n) = O(1)$ as $n \to \infty$. Especially, when an assembly is considered, $e(j)$ can be taken as

$$\max\left(\frac{1}{j}, \sup_{i\ge j} \big|i\mathsf{P}[Z_i = 1] - \theta\big|\right),$$

see the proof of Proposition 1.1 in Arratia *et al.* [2]. Thence, if

$$\sup_{i\geq j}\left|i\mathsf{P}[Z_i=1]-\theta\right| = \sup_{i\geq j}\left|i\lambda_i e^{-\lambda_i}-\theta\right| = O\left(\frac{1}{\log j}\right) \tag{4.3}$$

holds as $j\to\infty$, $e(j)=O((\log j)^{-1})$ follows for enough large j, so we have

$$g_\theta(n)\leq D_1\sum_{j=1}^{n}\frac{e(j)}{j}=O(\log\log n)$$

as $n\to\infty$. It shows that (3.2) is met.

When applying Lemma 2.1 in Section 5, the case $s_n(u)=\sum_{j=1}^{[n^u]}\lambda_j$ is considered. The condition (2.1) is met if $g_\theta(n)=o(\log n)$ for some $\theta>0$, because

$$\begin{aligned}\sum_{j=1}^{[n^u]}\lambda_j &= \sum_{j=1}^{[n^u]}\left(\lambda_j-\frac{\theta}{j}\right)+\theta\left(\sum_{j=1}^{[n^u]}\frac{1}{j}-u\log n\right)+\theta u\log n\\ &\leq g_\theta(n)+\theta+\theta u\log n\end{aligned}$$

holds for any $u\in(0,1)$. Moreover, we can choose $b=b(n)$ in Lemma 3.1 such as $n/((\log n)^2)$ if $g_\theta(n)=O(\log\log n)$. The condition $\sum_{j=1}^{\infty}e(j)/j<\infty$ appeared in Arratia *et al.* [2] as an additional sufficient condition for the results in their paper, and the slightly weaker condition (4.3) is new.

5. A FCLT in $L^2(0,1)$ for logarithmic combinatorial assemblies

Let us show the main assertion of this paper: a functional central limit theorem in $L^2(0,1)$ for logarithmic assemblies.

Theorem 5.1. *Consider an assembly. Assume* (2.6) *and the conditions in Lemma* 3.1 *for some* $\theta>0$. *Then, the random process* $(X_n(u))_{0<u<1}$ *defined in* (1.3) *converges weakly to* $(G(u))_{0<u<1}=(B(u)/\sqrt{u})_{0<u<1}$ *in* $L^2(0,1)$ *as* $n\to\infty$.

Proof. Lemma 3.1 yields that

$$\left\|\frac{\sum_{j=1}^{[n^\cdot]}C_j^n-\ell_n'(\cdot)}{\sqrt{\ell_n'(\cdot)}}-\frac{\sum_{j=1}^{[n^\cdot]}Z_j-\ell_n'(\cdot)}{\sqrt{\ell_n'(\cdot)}}\right\|_{L^2}\to^p 0.$$

Let $(N_t^1)_{t\geq 0}$ be the homogeneous Poisson process with unit intensity, then it holds that

$$\sum_{j=1}^{[n^u]}Z_j=^d\sum_{j=1}^{[n^u]}\left(N^1_{\ell_j'(1)}-N^1_{\ell_{j-1}'(1)}\right)=N^1_{\ell'_{[n^u]}(1)}=N^1_{\ell_n'(u)},$$

for any $u \in (0,1)$. Lemma 2.1 yields that

$$\left(\frac{\sum_{j=1}^{[n^u]} Z_j - \ell'_n(u)}{\sqrt{\ell'_n(u)}}\right)_{0<u<1} \Rightarrow \big(G(u)\big)_{0<u<1}$$

in $L^2(0,1)$. Theorem 2.7(iv) in van der Vaart [14] yields the conclusion. This completes the proof. □

Using a sufficient condition stated before, the following corollary holds.

Corollary 5.1. *Consider an assembly. Assume the conditions* (3.3) *for* $b(n)$ *satisfying the conditions in Lemma* 3.1, (4.3),

$$\lim_{j\to\infty} j\lambda_j = \theta, \tag{5.1}$$

and

$$\inf_{j=1,2,\ldots} j\lambda_j > 0. \tag{5.2}$$

Then, the random process $(X_n(u))_{0<u<1}$ *defined in* (1.3) *converges weakly to* $(G(u))_{0<u<1} = (B(u)/\sqrt{u})_{0<u<1}$ *in* $L^2(0,1)$ *as* $n \to \infty$.

6. Examples

As examples for Theorem 5.1 and Corollary 5.1, let us apply them to the component counts for the Ewens sampling formula and $\{n^n\}$ uniform random mappings. Concerning FCLTs, Poisson approximations for "small components" are important. As it is already seen in Section 3, a convergence $d_b(n) \to 0$ for some $b = b(n)$ guarantees the approximation. For special cases of the Ewens sampling formula (ESF) and random mappings, $d_b(n) \to 0$ if, and only if, $b = o(n)$, see Theorem 2 and Theorem 10 of Arratia and Tavaré [4]. For general logarithmic assemblies, there are some options to prove the convergence, and here let us introduce two of them.

The Δ-domain with parameters $\eta > 0$ and $0 < \phi < \pi/2$ is defined by

$$\Delta(\eta,\phi) = \big\{z \in \mathbf{C};\ |z| \le 1+\eta,\ \big|\arg(z-1)\big| \ge \phi\big\}.$$

Consider the exponential generating function of $\{\lambda_j\}_{j=1}^{\infty}$

$$f(z) = \exp\left(\sum_{j=1}^{\infty} \lambda_j z^j\right).$$

For concrete examples, see Section 4.3 in Arratia *et al.* [3]. The following theorems, by Arratia *et al.* [3] or Arratia *et al.* [2] respectively, can be used to see the conditions of Lemma 3.1.

Theorem 6.1 (Arratia *et al.* [3], Theorem 3). *For some* $\eta_0 > 0$ *and* $0 < \phi < \pi/2$, *define* $\Delta = \Delta(\eta_0, \phi)$. *Assume* f *is analytic on* $\Delta \setminus \{1\}$ *and there exist positive constants* δ, θ, K *such that* $f(z) = K(1-z)^{-\theta}(1 + O((1-z)^{\delta}))$ *as* $z \to 1$ *in* Δ. *Consider an assembly such that* $\sup_{j=1,2,\dots} j\lambda_j < \infty$. *If* $(b(n)\log n)/n \to 0$ *holds, then it holds that*

$$d_b(n) = \frac{|1-\theta|}{2n}\mathsf{E}\big[\big|T_{0b} - \mathsf{E}[T_{0b}]\big|\big] + o\left(\frac{b}{n}\right) = O\left(\frac{b}{n}\right)$$

as $n \to \infty$.

Theorem 6.2 (Arratia *et al.* [2], Theorem 3.1). *Consider a random structure which satisfies uniform logarithmic condition. If* $b(n)/n \to 0$ *holds, then it holds that*

$$\lim_{n\to\infty} d_b(n) = 0.$$

6.1. The Ewens sampling formula

Consider the component counts of ESF. The probability mass function (pmf) of the component counts is determined by the conditioning relation (1.1): the pmf is (4.1). In the case of ESF, since $\lambda_j = \theta/j$ for $\theta > 0$ and $j = 1, 2, \dots,$ $j\lambda_j$ is a constant $\theta > 0$. Note that $\ell_n(\cdot) = \ell_n'(\cdot)$, so $g_\theta(n) = 0$.

The conditions (5.1) and (5.2) are obvious. The condition (4.3) follows from

$$\sup_{i\geq j}\theta\left|e^{-\theta/i} - 1\right| = \theta\left(1 - e^{-\theta/j}\right) = O\left(\frac{1}{j}\right)$$

as $j \to \infty$. By Theorem 6.2, (3.3) is met for $b(n) = n/\log n$. Therefore, Corollary 5.1 yields a FCLT

$$\left(\frac{\sum_{j=1}^{[n^u]} C_j^n - \ell_n(u)}{\sqrt{\ell_n(u)}}\right)_{0<u<1} \Rightarrow \left(\frac{B(u)}{\sqrt{u}}\right)_{0<u<1} \tag{6.1}$$

in $L^2(0,1)$ as $n \to \infty$ for the component counts whose law is given by (4.1).

Remark 6.1. Theorem 1 of Arratia *et al.* [1] yields that

$$\mathsf{E}\left[\sup_{u\in(0,1)}\left|\frac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\log n}}\right|\right] \leq \frac{\sum_{j=1}^{n}\mathsf{E}[|C_j^n - Z_j|]}{\sqrt{\log n}} = O\left(\frac{1}{\sqrt{\log n}}\right)$$

as $n \to \infty$, see also Theorem 1 of Arratia and Tavaré [4]. As an analog of the result, it may be of interest to see the order of

$$\mathsf{E}\left[\left\|\frac{\sum_{j=1}^{[n^{\cdot}]}(C_j^n - Z_j)}{\sqrt{\ell_n(\cdot)}}\right\|_{L^2}^2\right]. \tag{6.2}$$

The thesis of the author considered it in Lemma 9.2.1, but unfortunately the proof, specifically the inequality in (9.2.5) page 118, was wrong. Note that, as it is stated before, a FCLT (6.1) in $L^2(0,1)$ holds without the convergence of (6.2) to 0.

6.2. Random mappings

Consider the component counts of $\{n^n\}$ uniform random mappings, in which case,

$$\lambda_j = \frac{e^{-j}}{j}\sum_{i=0}^{j-1}\frac{j^i}{i!}$$

for $j = 1, 2, \ldots$ and the conditioning relation (1.1) gives the pmf

$$\mathsf{P}\big[\big(C_1^n, \ldots, C_n^n\big) = (c_1, \ldots, c_n)\big] = \frac{n!e^n}{n^n}\prod_{j=1}^{n}\frac{\lambda_j^{c_j}}{c_j!}1\left\{\sum_{j=1}^{n} jc_j = n\right\}. \tag{6.3}$$

It holds that $\theta = 1/2$.

First, let us verify the conditions (3.2) and (2.6). Consider a sequence of independent Poisson random variables $\{P_j\}_{j=1}^{\infty}$ with $\mathsf{E}[P_j] = j$ for all $j = 1, 2, \ldots$. It holds that

$$\mathsf{P}(P_j < j) = \sum_{i=1}^{j-1}\mathsf{P}(P_j = i) = e^{-j}\sum_{i=0}^{j-1}\frac{j^i}{i!} = j\lambda_j.$$

Teicher [13] proves in their second inequality of (8) that $j\lambda_j$, which is denoted by $A_{j-1,j}$ with their notation, is increasing with respect to j. It holds that

$$\inf_{j=1,2,\ldots} j\lambda_j = \lambda_1 = \frac{1}{e} > 0.$$

Moreover, the convergence

$$\lim_{j\to\infty} j\lambda_j = \lim_{j\to\infty}\mathsf{P}\left(\frac{P_j - j}{\sqrt{j}} < 0\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} e^{-x^2/2}\,dx = \frac{1}{2}$$

holds, which is the essentially same as the equation (9) of Teicher [13]. It yields that $1/(2j) > \lambda_j$ for all $j = 1, 2, \ldots$. By this inequality and

$$\sum_{j=1}^{\infty}\left(\frac{1}{2j} - \lambda_j\right) = \frac{\log 2}{2},$$

which is the equation (31) in Donnelly *et al.* [7], it holds that

$$g_{\frac{1}{2}}(n) = \sum_{j=1}^{n}\left(\frac{1}{2j} - \lambda_j\right) = O(1).$$

Moreover, Theorem 6.1 yields $d_b(n) \to 0$ for $b = b(n)$ such that $b(n) = n/((\log n)^2)$ e.g. Therefore, Theorem 5.1 yields a FCLT in $L^2(0,1)$ for the component counts whose law is given by (6.3).

7. Another FCLT in $L^2(0,1)$ for logarithmic combinatorial assemblies

In this section, let us prove the weak convergence in $L^2(0,1)$ to $(B(u)/\sqrt{u})_{0<u<1}$ of the random process $(X_n'(u))_{0<u<1}$ defined in (1.4). In the case of the Ewens sampling formula and random mappings, the assumptions in Lemma 3.2 and (7.1), which are the assumptions of the weak convergence, are met.

Theorem 7.1. *Consider an assembly satisfying the assumptions in Lemma 3.2 with $f(n) = \log\log n$ and*

$$\lim_{n\to\infty}\left(g_\theta(n)\sqrt{\frac{\log\log n}{\log n}}\right) = 0 \tag{7.1}$$

for some $\theta > 0$, where $g_\theta(n)$ is defined in (3.1). Then, $(X_n'(u))_{0<u<1}$ defined in (1.4) converges weakly to $(G(u))_{0<u<1} = (B(u)/\sqrt{u})_{0<u<1}$ in $L^2(0,1)$ as $n \to \infty$.

Proof. Denote

$$\mathcal{P}_n(\cdot) = \left(\frac{N^1_{\theta u \log n} - \theta u \log n}{\sqrt{\theta u \log n}}\right)_{0<u<1},$$

where $(N_t^1)_{t\geq 0}$ is the Poisson process with unit intensity. It holds that

$$X_n'(u) - \mathcal{P}_n(u) = \begin{cases} -\mathcal{P}_n(u) & \left(0 < u < \dfrac{\varepsilon}{\log n}\right), \\ \dfrac{\sum_{j=1}^{[n^u]}(C_j^n - Z_j)}{\sqrt{\theta u \log n}} + \dfrac{\sum_{j=1}^{[n^u]} Z_j - N^1_{\theta u \log n}}{\sqrt{\theta u \log n}} & \left(\dfrac{\varepsilon}{\log n} < u < 1\right). \end{cases}$$

In order to prove a Poisson approximation in $L^2(0,1)$:

$$\left\| X_n'(\cdot) - \mathcal{P}_n(\cdot) \right\|^2_{L^2} \to^p 0, \tag{7.2}$$

it is sufficient to prove the followings:

$$\int_0^{\frac{\varepsilon}{\log n}} \frac{(N^1_{\theta u \log n} - \theta u \log n)^2}{\theta u \log n}\, du \to^p 0, \tag{7.3}$$

$$\int_{\frac{\varepsilon}{\log n}}^1 \frac{(\sum_{j=1}^{[n^u]}(C_j^n - Z_j))^2}{\theta u \log n}\, du \to^p 0, \tag{7.4}$$

$$\int_{\frac{\varepsilon}{\log n}}^{1} \frac{(\sum_{j=1}^{[n^u]} Z_j - N^1_{\theta u \log n})^2}{\theta u \log n} du \to^p 0. \tag{7.5}$$

As for (7.3), it follows from

$$\mathsf{E}\left[\int_0^{\frac{\varepsilon}{\log n}} \frac{(N^1_{\theta u \log n} - \theta u \log n)^2}{\theta u \log n} du\right] = \frac{\varepsilon}{\log n} \to 0.$$

As for (7.4), it follows from

$$\begin{aligned}\int_{\frac{\varepsilon}{\log n}}^{1} \frac{(\sum_{j=1}^{[n^u]}(C_j^n - Z_j))^2}{\theta u \log n} du &\le \int_{\frac{\varepsilon}{\log n}}^{1} \frac{(\sum_{j=1}^{n} |C_j^n - Z_j|)^2}{\theta u \log n} du \\ &= \frac{\log\log n - \log \varepsilon}{\theta \log n}\left(\sum_{j=1}^{n}\left|C_j^n - Z_j\right|\right)^2 \\ &\to^p 0\end{aligned}$$

by Lemma 3.2. As for (7.5), it holds that

$$\begin{aligned}&\sup_{u\in(0,1)} \mathsf{E}\left[\left(\sum_{j=1}^{[n^u]} Z_j - N^1_{\theta u \log n}\right)^2\right] \\ &\quad = \sup_{u\in(0,1)} \left(\left|\ell_n'(u) - \theta u \log n\right|^2 + \left|\ell_n'(u) - \theta u \log n\right|\right),\end{aligned} \tag{7.6}$$

since

$$\left|\sum_{j=1}^{[n^u]} Z_j - N^1_{\theta u \log n}\right| =^d \left|N^1_{\ell_n'(u)} - N^1_{\theta u \log n}\right| =^d N^1_{|\ell_n'(u) - \theta u \log n|}$$

holds for any $u \in (0, 1)$. The triangle inequality yields that

$$\sup_{u\in(0,1)} \left|\ell_n'(u) - \theta u \log n\right| \le \sup_{u\in(0,1)} \left|\ell_n'(u) - \ell_n(u)\right| + \sup_{u\in(0,1)} \left|\ell_n(u) - \theta u \log n\right|, \tag{7.7}$$

so (7.6) is $o(\log n / \log\log n)$. That is because the first term of (7.7) is $o(\sqrt{\log n / \log\log n})$ by the assumption (7.1) and the second term is bounded above by θ. Thence, (7.5) holds, because

$$\begin{aligned}&\mathsf{E}\left[\int_{\frac{\varepsilon}{\log n}}^{1} \frac{(\sum_{j=1}^{[n^u]} Z_j - N^1_{\theta u \log n})^2}{\theta u \log n} du\right] \\ &\quad \le \sup_{u\in(0,1)} \left(\mathsf{E}\left[\left(\sum_{j=1}^{[n^u]} Z_j - N^1_{\theta u \log n}\right)^2\right]\right) \int_{\frac{\varepsilon}{\log n}}^{1} \frac{1}{\theta u \log n} du\end{aligned}$$

$$= \sup_{u\in(0,1)} \left(\left| \ell_n'(u) - \theta u \log n \right|^2 + \left| \ell_n'(u) - \theta u \log n \right| \right) \frac{\log\log n - \log \varepsilon}{\theta \log n}.$$

Therefore, (7.2) follows. Setting $s_n(u) = \theta u \log n$ $(0 \le u \le 1)$, $K = \theta$ and $\delta < 1$, the conditions of Lemma 2.1 are met, so Lemma 2.1 yields the weak convergence of $\mathcal{P}_n(\cdot)$ to $G(\cdot)$. This completes the proof. □

Remark 7.1. Comparing with Theorem 5.1, (7.1) is stronger than (3.2), but note that if (4.3) holds, then the condition (7.1) also holds.

Acknowledgements

The contents of this paper are based partly on the thesis of the author. The major part of this work was done when the author was a Research Fellow of Japan Society for the Promotion of Science at SOKENDAI (The Graduate University for Advanced Studies) and Waseda University. This work was partly supported by JSPS KAKENHI Grant Number 26-1487 (Grant-in-Aid for JSPS Fellows). The author wishes to express his sincere thanks to Professors Satoshi Hattori, Shuhei Mano and Yoichi Nishiyama for their insightful comments.

References

[1] Arratia, R., Barbour, A.D. and Tavaré, S. (1992). Poisson process approximations for the Ewens sampling formula. *Ann. Appl. Probab.* **2** 519–535. MR1177897
[2] Arratia, R., Barbour, A.D. and Tavaré, S. (2000). Limits of logarithmic combinatorial structures. *Ann. Probab.* **28** 1620–1644. MR1813836
[3] Arratia, R., Stark, D. and Tavaré, S. (1995). Total variation asymptotics for Poisson process approximations of logarithmic combinatorial assemblies. *Ann. Probab.* **23** 1347–1388. MR1349176
[4] Arratia, R. and Tavaré, S. (1992). Limit theorems for combinatorial structures via discrete process approximations. *Random Structures Algorithms* **3** 321–345. MR1164844
[5] Arratia, R. and Tavaré, S. (1994). Independent process approximations for random combinatorial structures. *Adv. Math.* **104** 90–154. MR1272071
[6] DeLaurentis, J.M. and Pittel, B.G. (1985). Random permutations and Brownian motion. *Pacific J. Math.* **119** 287–301. MR0803120
[7] Donnelly, P.J., Ewens, W.J. and Padmadisastra, S. (1991). Functionals of random mappings: Exact and asymptotic results. *Adv. in Appl. Probab.* **23** 437–455. MR1122869
[8] Ewens, W.J. (1972). The sampling theory of selectively neutral alleles. *Theoret. Population Biology* **3** 87–112; erratum, ibid. **3** (1972), 240; erratum, ibid. **3** (1972), 376. MR0325177
[9] Flajolet, P. and Soria, M. (1990). Gaussian limiting distributions for the number of components in combinatorial structures. *J. Combin. Theory Ser. A* **53** 165–182. MR1041444
[10] Hansen, J.C. (1989). A functional central limit theorem for random mappings. *Ann. Probab.* **17** 317–332. Correction: *Ann. Probab.* **19** (1991) 1393–1396. MR0972788
[11] Hansen, J.C. (1990). A functional central limit theorem for the Ewens sampling formula. *J. Appl. Probab.* **27** 28–43. MR1039182
[12] Prohorov, Yu.V. (1956). Convergence of random processes and limit theorems in probability theory. *Teor. Veroyatn. Primen.* **1** 177–238. MR0084896

[13] Teicher, H. (1955). An inequality on Poisson probabilities. *Ann. Math. Stat.* **26** 147–149. MR0067384
[14] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247
[15] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*: *With Applications to Statistics*. *Springer Series in Statistics*. New York: Springer. MR1385671

Received October 2015