Electron. Commun. Probab. **0** (2025), article no. 0, 1–9.
https://doi.org/10.1214/YY-TN
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Some central limit theorems for critical beta-splitting trees*

Alexander Iksanov† Anatolii Nikitin‡ Roman Yakymiv§

Abstract

We further explore a connection initially unveiled in Iksanov (2025) between critical beta-splitting trees and infinite 'balls-in-boxes' schemes. Using the connection, we derive a new joint central limit theorem for components of the height of a leaf chosen uniformly at random in the discrete version of a critical beta-splitting tree. Also, we obtain a joint central limit theorem for the heights in the discrete and continuous versions of a critical beta-splitting tree.

Keywords: central limit theorem; infinite occupancy; random tree; subordinator.
MSC2020 subject classifications: Primary 60C05; 60F05, Secondary 05C05; 60K05.
Submitted to ECP on October 13, 2025, final version accepted on .

1 Setting and result

For $\beta \in (-2, +\infty)$ and integer $m \geq 2$, define the probability distribution $q^{(\beta)}(m) := (q^{(\beta)}(m,i),\, 1 \leq i \leq m-1)$ by

$$q^{(\beta)}(m,i) := \frac{1}{a^{(\beta)}(m)} \frac{\Gamma(\beta+i+1)\Gamma(\beta+m-i+1)}{\Gamma(i+1)\Gamma(m-i+1)}, \quad i = 1, 2, \ldots, m-1,$$

where Γ is the Euler gamma function and $a^{(\beta)}(m)$ is the normalizing constant. Following [2], we construct a random rooted tree called *beta-splitting tree with n leaves* by recursively splitting the set $\{1, 2, \ldots, n\}$ of 'elements'. There is a left edge at the root leading to a left subtree with G_n elements $\{1, 2, \ldots, G_n\}$ and a right edge leading to a right subtree with $n - G_n$ elements $\{G_n + 1, G_n + 2, \ldots, n\}$. Here, G_n is a random variable with distribution $q^{(\beta)}(n)$. Recursively, each set of m elements with $m \geq 2$ is split into two subsets, with a left subset having a random size with distribution $q^{(\beta)}(m)$. The procedure stops once all sets of elements have become leaves (singletons). Here is a possible realization of the procedure with $n = 5$:

$$[12345] \to [12][345] \to [1][2][34][5] \to [1][2][3][4][5]. \tag{1.1}$$

*The work of A. Iksanov was supported by the National Research Foundation of Ukraine (project 2023.03/0059 'Contribution to modern theory of random series')

†Faculty of Computer Science and Cybernetics, Taras Shevchenko National University of Kyiv, Ukraine E-mail: iksan@univ.kiev.ua http://do.csc.knu.ua/iksan

‡Faculty of Natural Sciences, Jan Kochanowski University of Kielce, Poland and Faculty of Economics, National University of Ostroh Academy, Ukraine. E-mail: anatolii.nikitin@ujk.edu.pl

§Faculty of Computer Science and Cybernetics, Taras Shevchenko National University of Kyiv, Ukraine. E-mail: yakymiv@knu.ua

The beta-splitting tree with $\beta = -1$ is called *critical*. For the critical trees,

$$q^{(-1)}(m,i) = \frac{m}{2h_{m-1}}\frac{1}{i(m-i)}, \quad i = 1,2,\ldots,m-1,$$

where $h_m := \sum_{j=1}^m 1/j$. The term 'critical' stems from the fact the leaf heights are of order $n^{-\beta-1}$ for $\beta \neq -1$, whereas these are of order $\log n$ for $\beta = -1$. A detailed investigation of the critical beta-splitting trees was undertaken very recently in a series of papers [3, 5, 7], see also [6, 9, 14, 15]. The state of the art can be traced via the dynamically changing arXiv article [4].

There are two versions of this construction: discrete and continuous. In the discrete construction, the edges are assumed to have lengths 1, whereas in the continuous construction, the edge emanating from a subset of size $m \geq 2$ has a random length with the exponential distribution of mean $1/h_{m-1}$. Let L_n and D_n denote the heights of a leaf chosen uniformly at random in the discrete and continuous versions of a critical beta-splitting tree with n leaves, respectively. The terms 'edge-height' for L_n and 'time-height' for D_n are also used. As the splitting procedure proceeds, the sizes of subtrees containing a uniformly chosen leaf decrease from n to 1. For $j \in \mathbb{N}$, $j \leq n-1$, denote by $L_{n,j}$ the number of the size j decrements in a sequence of the sizes of those subtrees. Referring back to (1.1) and assuming that a uniformly chosen leaf is 3, the sizes of subtrees are 5, 3, 2 and 1, $L_5 = 3$, $L_{5,1} = 2$, $L_{5,2} = 1$ and $L_{5,3} = L_{5,4} = 0$. Plainly, $L_n = \sum_{j=1}^{n-1} L_{n,j}$ and $\sum_{j=1}^{n-1} jL_{n,j} = n-1$.

In Theorem 1.7 of [7], the following central limit theorems were proved. Let $\mathrm{Normal}\,(0,1)$ denote a random variable with the normal distribution of zero mean and unit variance. As usual, $\stackrel{\mathrm{d}}{\to}$ denotes convergence in distribution.

Proposition 1.1. *As $n \to \infty$,*

$$\frac{L_n - (2\zeta(2))^{-1}(\log n)^2}{((2\zeta(3)/3(\zeta(2))^3)(\log n)^3)^{1/2}} \stackrel{\mathrm{d}}{\to} \mathrm{Normal}\,(0,1), \quad n \to \infty,$$

where $\zeta(s) := \sum_{n\geq 1} n^{-s}$ for $s > 1$, and

$$\frac{D_n - (\zeta(2))^{-1}\log n}{((2\zeta(3)/(\zeta(2))^3)\log n)^{1/2}} \stackrel{\mathrm{d}}{\to} \mathrm{Normal}\,(0,1), \quad n \to \infty.$$

The proof in [7] was based on an asymptotic analysis of recurrences satisfied by $u \mapsto \mathbb{E}e^{uL_n}$ and $u \mapsto \mathbb{E}e^{uD_n}$ for $u \geq 0$ and $n \in \mathbb{N}$, respectively. The recurrences are actually consequences of the distributional equalities

$$L_n \stackrel{\mathrm{d}}{=} L_{n-J_n} + 1, \quad n \geq 2, \quad D_n \stackrel{\mathrm{d}}{=} D_{n-J_n} + \tau_n, \quad n \geq 2, \quad D_1 = L_1 = 0, \tag{1.2}$$

where $\stackrel{\mathrm{d}}{=}$ denotes equality of distributions. Here, J_n is independent of $(L_k, D_k)_{k\geq 2}$ on the right-hand sides and has distribution

$$\mathbb{P}\{J_n = k\} = \frac{1}{kh_{n-1}}, \quad k = 1,2,\ldots,n-1,$$

and τ_n is independent of J_n and $(D_k)_{k\geq 2}$ and has the exponential distribution of mean $1/h_{n-1}$. In a more recent article [15], the central limit theorem for L_n was derived from Theorem 2.1 in [16], which provides sufficient conditions ensuring the validity of a central limit theorem for distributional solutions to recurrences like (1.2) and more general ones. Even more recently, the scope of that approach was extended in [9]. Among other things, this enabled the authors of [9] to prove a central limit theorem for the total length of a continuous-time critical beta-splitting tree with n leaves as $n \to \infty$.

The purpose of this note is to prove[1] the joint convergence of $(L_{n,1},\dots,L_{n,j},L_n,D_n)$, properly centered and normalized, as $n\to\infty$ for each $j\in\mathbb{N}$.

Theorem 1.2. *As $n\to\infty$,*

$$\Big(\frac{L_n-(2\zeta(2))^{-1}(\log n)^2}{(2\zeta(3)(\zeta(2))^{-3}(\log n)^3)^{1/2}},\frac{D_n-(\zeta(2))^{-1}\log n}{(2\zeta(3)(\zeta(2))^{-3}\log n)^{1/2}}\Big)\ \overset{\mathrm{d}}{\to}\ \Big(\int_0^1 W^\ast(u)\mathrm{d}u, W^\ast(1)\Big),\quad (1.3)$$

where $W^\ast:=(W^\ast(u))_{u\in[0,1]}$ is a standard Brownian motion.

For any fixed $j\in\mathbb{N}$, as $n\to\infty$,

$$\Big(\frac{L_{n,r}-(\zeta(2)r)^{-1}\log n}{(\log n)^{1/2}}\Big)_{1\le r\le j}\ \overset{\mathrm{d}}{\to}\ \big((2\zeta(3)(\zeta(2))^{-3})^{1/2}r^{-1}W^\ast(1)+(r\zeta(2))^{-1/2}Y_r\big)_{1\le r\le j}\quad (1.4)$$

for the same Brownian motion $W^\ast$ as in (1.3). Here, Y_1, $Y_2,\dots,Y_j$ are independent random variables with the standard normal distribution, which are independent of $W^\ast$. Also, limit relations (1.3) and (1.4) hold jointly.

Remark 1.3. Observe that $\mathbb{E}\big[\big(\int_0^1 W^\ast(u)\mathrm{d}u\big)^2\big]=1/3$. This explains the fact that the denominator of the fraction involving L_n in Theorem 1.2 differs by the factor $3^{-1/2}$ from that in Proposition 1.1.

It was shown in [14] that, for each $n\in\mathbb{N}$, there is a connection between beta-splitting tree with n leaves and certain ‘balls-in-boxes’ scheme with n balls and infinitely many boxes. Using that connection, discussed in some detail in Section 2, we provide in Section 3 a reasonably simple proof of Theorem 1.2.

2 Our strategy

It is important for what follows that the distributional equalities in (1.2) hold true jointly, also together with the corresponding distributional recurrences for $(L_{n,1},\dots,L_{n,j})$, $j\in\mathbb{N}$, $j\le n-1$. Fix such a j. For $n\ge1$, put $X_n:=(L_{n,1},\dots,L_{n,j},L_n,D_n)$ and $Y_n:=(\mathbb{1}_{\{J_n=1\}},\dots,\mathbb{1}_{\{J_n=j\}},1,\tau_n)$. Then

$$X_n\overset{\mathrm{d}}{=}X_{n-J_n}+Y_n,\quad n\ge2,\quad X_1=\underbrace{(0,0,0,\dots,0)}_{j+2\text{ zeros}}.\quad (2.1)$$

Let $S:=(S(t))_{t\ge0}$ be a subordinator (an almost surely increasing Lévy process) with $S(0)=0$, zero drift, no killing and the Lévy measure ν defined by

$$\nu(\mathrm{d}x)=\frac{\mathrm{e}^{-x}}{1-\mathrm{e}^{-x}}\,\mathbb{1}_{(0,\infty)}(x)\,\mathrm{d}x.\quad (2.2)$$

Let $E_1,\dots,E_n$ be independent random variables with the exponential distribution of unit mean, which are independent of S. The closed range of S has zero Lebesgue measure and splits the positive halfline into infinitely many disjoint intervals that we call gaps. To obtain the aforementioned ‘balls-in-boxes’ scheme, take the gaps in the role of boxes and the points of the exponential sample in the role of balls. Denote by K_n the number of gaps occupied by at least one E_j, $j=1,2,\dots,n$. Also, for $j\in\mathbb{N}$, $j\le n$, denote by $K_{n,j}$ the number of gaps occupied by j elements of the exponential sample $E_1,\dots,E_n$. Note that $K_n=\sum_{j=1}^n K_{n,j}$ and $\sum_{j=1}^n jK_{n,j}=n$.

Put $S^{\leftarrow}(t):=\inf\{u\ge0: S(u)>t\}$ for $t\ge0$ and $E_{n,n}:=\max(E_1,\dots,E_n)$ for $n\in\mathbb{N}$. Here is the basic observation which enables us to connect critical beta-splitting trees with the ‘balls-in-boxes’ scheme.

[1] This article is partly based on the unpublished manuscript ‘Another proof of CLT for critical beta-splitting tree’ by A. Iksanov written in May, 2024. That manuscript contained a new proof of the joint distributional convergence of (L_n,D_n). New results on the distributional convergence of $(L_{n,1},\dots,L_{n,j},L_n,D_n)$ were obtained by A. Nikitin and R. Yakymiv.

Proposition 2.1. *For each $n \geq 1$ and each $j \in \mathbb{N}$, $j \leq n$,*

$$(L_{n+1,1}, \ldots, L_{n+1,j}, L_{n+1}, D_{n+1}) \overset{\mathrm{d}}{=} (K_{n,1}, \ldots, K_{n,j}, K_n, S^{\leftarrow}(E_{n,n})).$$

With this at hand, limit relation (1.3) in Theorem 1.2 is a consequence of the following result, in which ν is a nonzero Lévy measure, not necessarily the same as given in (2.2). Put

$$\Phi(t) := \int_{(0,\,\infty)} \big(1 - \exp(-t(1-\mathrm{e}^{-x}))\big)\nu(\mathrm{d}x), \quad t > 0,$$

$$\mathtt{m}_2 := \mathrm{Var}\,[S(1)] = \int_{(0,\,\infty)} x^2\nu(\mathrm{d}x), \quad \mathtt{m}_1 := \mathbb{E}[S(1)] = \int_{(0,\,\infty)} x\nu(\mathrm{d}x).$$

Theorem 2.2. *Assume that $t \mapsto \Phi(\mathrm{e}^t)$ is regularly varying at ∞ of positive index β and that $\mathtt{m}_2 \in (0, \infty)$. Then, as $n \to \infty$,*

$$\Big(\frac{K_n - \mathtt{m}_1^{-1}\int_1^n y^{-1}\Phi(y)\mathrm{d}y}{(\mathtt{m}_2\mathtt{m}_1^{-3}\log n)^{1/2}\Phi(n)}, \frac{S^{\leftarrow}(E_{n,n}) - \mathtt{m}_1^{-1}\log n}{(\mathtt{m}_2\mathtt{m}_1^{-3}\log n)^{1/2}}\Big) \overset{\mathrm{d}}{\to} \Big(\beta\int_0^1 W(1-u)u^{\beta-1}\mathrm{d}u, W(1)\Big),$$

where $W := \big(W(u)\big)_{u\in[0,1]}$ is a standard Brownian motion.

Now let ν be as in (2.2). According to a Hurwitz identity (see, for instance, formula (23.2.7) in [1]),

$$\mathtt{m}_\mathtt{r} := \int_{(0,\infty)} x^r\nu(\mathrm{d}x) = \int_0^1 x^{-1}(-\log(1-x))^r\mathrm{d}x = \Gamma(r+1)\zeta(r+1), \quad r > 0,$$

where Γ is the Euler gamma function. In particular, $\mathtt{m}_1 := \zeta(2)$ and $\mathtt{m}_2 = 2\zeta(3)$. Further,

$$\begin{aligned}\Phi(t) = \int_{(0,\,\infty)} \big(1-\exp\{-t(1-\mathrm{e}^{-x})\}\big)\nu(\mathrm{d}x) &= \int_0^\infty \big(1-\exp\{-t(1-\mathrm{e}^{-x})\}\big)\frac{\mathrm{e}^{-x}}{1-\mathrm{e}^{-x}}\mathrm{d}x \\ &= \int_0^t y^{-1}(1-\mathrm{e}^{-y})\mathrm{d}y = \log t + \gamma + O(\mathrm{e}^{-t}), \quad t \to \infty, \end{aligned} \tag{2.3}$$

where $\gamma = \int_0^1 y^{-1}(1-\mathrm{e}^{-y})\mathrm{d}y - \int_1^\infty y^{-1}\mathrm{e}^{-y}\mathrm{d}y$ is the Euler-Mascheroni constant. This shows that β in Theorem 2.2 is equal to 1 and

$$\int_1^n y^{-1}\Phi(y)\mathrm{d}y = (\log n)^2/2 + \gamma\log n + o(\log n) = (\log n)^2/2 + o((\log n)^{3/2}), \quad n \to \infty.$$

Thus, limit relation (1.3) in Theorem 1.2 does indeed follow from Proposition 2.1 and Theorem 2.2.

In view of Proposition 2.1, limit relation (1.4) and the joint convergence in Theorem 1.2 follow from Proposition 2.3 given next. We write $\Phi^{(r)}$ to denote the rth derivative of Φ.

Proposition 2.3. *Let $j \in \mathbb{N}$. Assume that*

$$\Phi(t) = \log t + c + O(t^{-\varepsilon}), \quad t \to \infty \tag{2.4}$$

for some constants $c \in \mathbb{R}$ and $\varepsilon > 0$, and that ν has a continuous density on $(0,\infty)$ and satisfies $\nu([y,\infty]) = O(\mathrm{e}^{-\delta y})$ as $y \to \infty$ for some $\delta > 0$. Assume also that, for each

$r = 1, \dots, j$, the function $t \mapsto (-1)^{r+1}\Phi^{(r)}(t)t^r$ is nondecreasing on $(1,\infty)$. Then, as $n \to \infty$,

$$\Big(\frac{K_{n,r} - (\mathtt{m}_1 r)^{-1}\log n}{(\log n)^{1/2}}\Big)_{1\le r\le j} \overset{\mathrm{d}}{\to} \big((\mathtt{m}_2\mathtt{m}_1^{-3})^{1/2} r^{-1} W(1) + (r\mathtt{m}_1)^{-1/2} Z_r\big)_{1\le r\le j} \tag{2.5}$$

for the same Brownian motion W as in Theorem 2.2. Here, $Z_1, Z_2, \dots, Z_j$ are independent random variables with the standard normal distribution, which are independent of W. Also, limit relations of Theorem 2.2 and Proposition 2.3 hold jointly, namely, as $n \to \infty$,

$$\begin{aligned}&\Big(\Big(\frac{K_{n,r} - (\mathtt{m}_1 r)^{-1}\log n}{(\log n)^{1/2}}\Big)_{1\le r\le j}, \frac{K_n - (2\mathtt{m}_1)^{-1}(\log n)^2}{(\log n)^{3/2}}, \frac{S^{\leftarrow}(E_{n,n}) - \mathtt{m}_1^{-1}\log n}{(\log n)^{1/2}}\Big)\\ &\overset{\mathrm{d}}{\to} \Big(\big((\mathtt{m}_2\mathtt{m}_1^{-3})^{1/2} r^{-1} W(1) + (r\mathtt{m}_1)^{-1/2} Z_r\big)_{1\le r\le j}, (\mathtt{m}_2\mathtt{m}_1^{-3})^{1/2}\int_0^1 W(u)\mathrm{d}u, (\mathtt{m}_2\mathtt{m}_1^{-3})^{1/2} W(1)\Big).\end{aligned} \tag{2.6}$$

We shall derive relation (2.5) from Theorem 15 in [12]. The distributional convergence of K_n as stated in Proposition 2.3 was earlier obtained in Theorem 12 of [12]. The proof of this result given in [12] is different from our proof of Theorem 2.2.

The measure ν as in (2.2) has a continuous density on $(0,\infty)$ and satisfies $\nu([y,\infty]) \sim \mathrm{e}^{-y}$ as $y \to \infty$. According to (2.3), the corresponding Φ satisfies (2.4) with $c = \gamma$ and any $\varepsilon > 0$. Further, for $r \in \mathbb{N}$, put $F_r(t) := 1 - \mathrm{e}^{-t}(1 + t + t^2/2 + \ldots + t^{r-1}/(r-1)!)$ for $t \ge 0$. Thus, F_r is the distribution function of a gamma distributed random variable with parameters r and 1. It can be checked with the help of mathematical induction that

$$(-1)^{r+1} t^r \Phi^{(r)}(t) = (r-1)! F_r(t), \quad t \ge 0.$$

In particular, the function $t \mapsto (-1)^{r+1} t^r \Phi^{(r)}(t)$ is increasing on $[0,\infty)$. Thus, Proposition 2.3 applies to the 'balls-in-boxes' scheme with ν as in (2.2).

3 Proofs

Proof of Proposition 2.1. For fixed $n \in \mathbb{N}$ and each $t \ge 0$, put

$$N_n(t) := \#\{1 \le k \le n : E_k > S(t)\}.$$

Then $N_n := (N_n(t))_{t\ge 0}$ is a decreasing integer-valued continuous-time Markov chain which starts at n and gets absorbed at 0. The variable $S^{\leftarrow}(E_{n,n})$ is the absorption time at 0. According to Theorem 5.2(i) in [11], N_n goes from m to $m-k$ with probability

$$\begin{aligned}\frac{\binom{m}{k}\int_{(0,\infty)}(1-\mathrm{e}^{-x})^k \mathrm{e}^{-x(m-k)}\nu(\mathrm{d}x)}{\sum_{k=1}^m \binom{m}{k}\int_{(0,\infty)}(1-\mathrm{e}^{-x})^k \mathrm{e}^{-x(m-k)}\nu(\mathrm{d}x)} &= \frac{\binom{m}{k}\int_0^1 x^{k-1}(1-x)^{m-k}\mathrm{d}x}{\sum_{k=1}^m \binom{m}{k}\int_0^1 x^{k-1}(1-x)^{m-k}\mathrm{d}x}\\ &= \frac{1}{kh_m}, \quad k = 1, 2, \dots, m.\end{aligned}$$

Thus, denoting by J_n^* the number of exponential points (out of n) falling into the left-most occupied gap we infer

$$\mathbb{P}\{J_n^* = k\} = \frac{1}{kh_n} = \mathbb{P}\{J_{n+1} = k\}, \quad k = 1, 2, \dots, n.$$

Let τ_n^* be the time of the first decrement of N_n. This random variable has the exponential distribution with mean

$$\frac{1}{\sum_{k=1}^n \binom{n}{k}\int_{(0,\infty)}(1-\mathrm{e}^{-x})^k \mathrm{e}^{-x(n-k)}\nu(\mathrm{d}x)} = \frac{1}{h_n}.$$

Fix any $j \in \mathbb{N}$, $j \leq n$. Now, with $X_n^* := (K_{n,1}, \dots, K_{n,j}, K_n, S^{\leftarrow}(E_{n,n}))$ and $Y_n^* := (\mathbb{1}_{\{J_n^*=1\}}, \dots, \mathbb{1}_{\{J_n^*=j\}}, 1, \tau_n^*)$ for $n \geq 1$, we infer

$$X_n^* \overset{\mathrm{d}}{=} X_{n-J_n^*}^* + Y_n^*, \quad n \geq 1, \quad X_0^* = \underbrace{(0, 0, 0, \dots, 0)}_{j+2 \text{ zeros}},$$

where J_n^* is independent of Y_n^* and $(X_k^*)_{k \geq 1}$ on the right-hand side. Comparing this distributional equality with (2.1) completes the proof of the proposition. □

Proof of Theorem 2.2. Let $D([0,1])$ be the Skorohod space of càdlàg functions defined on $[0,1]$ equipped with the J_1-topology. In what follows, $\Rightarrow$ denotes weak convergence on this space. It is a standard fact that

$$\Big(\frac{S(ut) - \mathtt{m}_1 ut}{(\mathtt{m}_2 t)^{1/2}}\Big)_{u \in [0,1]} \Rightarrow (-W(u))_{u \in [0,1]}, \quad t \to \infty$$

(it is more convenient to write here $-W(u)$ rather than $W(u)$). By inversion (see, for instance, Lemma 1 in [17])

$$\Big(\frac{S^{\leftarrow}(ut) - \mathtt{m}_1^{-1} ut}{(\mathtt{m}_2 \mathtt{m}_1^{-3} t)^{1/2}}\Big)_{u \in [0,1]} \Rightarrow (W(u))_{u \in [0,1]}, \quad t \to \infty. \tag{3.1}$$

Specializing this to one-dimensional convergence and recalling that $E_{n,n} - \log n$ converges in distribution to a random variable with the Gumbel distribution we conclude that

$$\frac{S^{\leftarrow}(E_{n,n}) - \mathtt{m}_1^{-1} \log n}{(\mathtt{m}_2 \mathtt{m}_1^{-3} \log n)^{1/2}} \overset{\mathrm{d}}{\to} W(1), \quad n \to \infty.$$

This demonstrates the distributional convergence of the second coordinate in Theorem 2.2. The distributional convergence of the first coordinate was proved in Theorem 3.1(a) of [10]. Thus, we only need to combine these into the joint convergence.

For each $t > 0$, put

$$W_t(u) := \frac{S^{\leftarrow}(t(1-u)) - \mathtt{m}_1^{-1} t(1-u)}{(\mathtt{m}_2 \mathtt{m}_1^{-3} t)^{1/2}}, \quad u \in [0,1].$$

An inspection of the proof of Theorem 3.1(a) of [10] reveals that the distributional convergence of the first coordinate in Theorem 2.2 is driven by that of $\int_{[0,\,1]} W_t(u) \mathrm{d}_u(\Phi(\mathrm{e}^{tu})/\Phi(\mathrm{e}^t))$. Thus, we are left with showing that, for all real α_1 and α_2,

$$\alpha_1 W_t(0) + \alpha_2 \int_{[0,\,1]} W_t(u) \mathrm{d}_u(\Phi(\mathrm{e}^{tu})/\Phi(\mathrm{e}^t)) = \int_{[0,\,1]} (\alpha_1 W_t(0) + \alpha_2 W_t(u)) \mathrm{d}_u(\Phi(\mathrm{e}^{tu})/\Phi(\mathrm{e}^t))$$
$$\overset{\mathrm{d}}{\to} \beta \int_0^1 (\alpha_1 W(1) + \alpha_2 W(1-u)) u^{\beta-1} \mathrm{d}u = \alpha_1 W(1) + \alpha_2 \beta \int_0^1 W(1-u) u^{\beta-1} \mathrm{d}u, \quad t \to \infty.$$

Since, for each $t > 0$, $u \mapsto \Phi(\mathrm{e}^{tu})/\Phi(\mathrm{e}^t)$, $u \in [0,1]$ is the distribution function of a nonnegative random variable, $\lim_{t \to \infty} (\Phi(\mathrm{e}^{tu})/\Phi(\mathrm{e}^t)) = u^\beta$ for each $u \geq 0$ and, in view of (3.1), $(\alpha_1 W_t(0) + \alpha_2 W_t(u))_{u \in [0,1]} \Rightarrow (\alpha_1 W(1) + \alpha_2 W(1-u))_{u \in [0,1]}$ as $t \to \infty$, this convergence is secured by Lemma 4.2 (a). The proof of Theorem 2.2 is complete. □

Proof of Proposition 2.3. Let $\pi := (\pi(t))_{t \geq 0}$ denote a Poisson process on $[0, +\infty)$ of unit intensity. For $r \in \mathbb{N}$ and $t \geq 0$, denote by $\mathcal{K}(t, r)$ the number of gaps that contain r elements of the Poissonized sample $E_1, \dots, E_{\pi(t)}$.

By Theorem 15 in [12], under the assumptions of Proposition 2.3, excluding the monotonicity assumption, for any fixed $j \in \mathbb{N}$,

$$\Big(\frac{K_{n,r} - (\mathtt{m}_1 r)^{-1}\log n}{(\log n)^{1/2}}\Big)_{1\leq r\leq j} \overset{\mathrm{d}}{\to} (X_r)_{1\leq r\leq j}, \quad n\to\infty,$$

where $(X_r)_{1\leq r\leq j}$ is a centered Gaussian vector with $\mathbb{E}[X_r X_s] = \mathtt{m}_2\mathtt{m}_1^{-3}(rs)^{-1} + \mathbb{1}_{\{r=s\}}(\mathtt{m}_1 r)^{-1}$ for $r, s = 1, 2, \ldots, j$. Also,

$$\Big(\frac{\mathcal{K}(t,r) - (\mathtt{m}_1 r)^{-1}\log t}{(\log t)^{1/2}}\Big)_{1\leq r\leq j} \overset{\mathrm{d}}{\to} (X_r)_{1\leq r\leq j}, \quad t\to\infty, \tag{3.2}$$

see p. 596 in [12]. The formula for the covariance entails that

$$X_r = (\mathtt{m}_2\mathtt{m}_1^{-3})^{1/2} r^{-1} Z + (\mathtt{m}_1 r)^{-1/2} Z_r, \quad r = 1, 2, \ldots, j,$$

where Z, $Z_1, \ldots, Z_j$ are independent random variables with the standard normal distribution. It remains to identify Z with $W(1)$, where W is a Brownian motion as in Theorem 2.2. To this end, we use a decomposition, for $r = 1, \ldots, j$,

$$\mathcal{K}(\mathrm{e}^t, r) - \mathtt{m}_1^{-1}\int_0^t h_r(y)\mathrm{d}y = \Big(\mathcal{K}(\mathrm{e}^t, r) - \int_{[0,\,t]} h_r(t-x)\mathrm{d}S^{\leftarrow}(x)\Big)$$
$$+ \Big(\int_{[0,\,t]} h_r(t-x)\mathrm{d}S^{\leftarrow}(x) - \mathtt{m}_1^{-1}\int_0^t h_r(y)\mathrm{d}y\Big) =: A(t,r) + B(t,r),$$

where $h_r(t) := (-1)^{r+1}(r!)^{-1}\Phi^{(r)}(\mathrm{e}^t)\mathrm{e}^{rt}$ for $t \in \mathbb{R}$. We shall show that

$$\Big(\frac{B(t,r)}{t^{1/2}}\Big)_{1\leq r\leq j} \overset{\mathrm{d}}{\to} (\mathtt{m}_2\mathtt{m}_1^{-3})^{1/2} W(1)(r^{-1})_{1\leq r\leq j}, \quad t\to\infty, \tag{3.3}$$

where $(W(u))_{u\in[0,1]}$ is a standard Brownian motion as in (3.1), and that

$$\int_0^t h_r(y)\mathrm{d}y = r^{-1}\log t + O(1), \quad t\to\infty. \tag{3.4}$$

In view of (3.2), we then have

$$\Big(\frac{A(t,r)}{t^{1/2}}\Big)_{1\leq r\leq j} \overset{\mathrm{d}}{\to} ((\mathtt{m}_1 r)^{-1/2} Z_r)_{1\leq r\leq j}, \quad t\to\infty.$$

Finally, these together with the argument given in the proof of Theorem 2.2 will justify formula (2.6).

Proof of (3.3). We intend to check that, for any $\alpha_1, \alpha_2, \ldots, \alpha_j \in \mathbb{R}$,

$$\frac{\sum_{r=1}^j \alpha_r \int_{[0,\,t]} h_r(t-x)\mathrm{d}(S^{\leftarrow}(x) - \mathtt{m}_1^{-1}x)}{(\mathtt{m}_2\mathtt{m}_1^{-3}t)^{1/2}} \overset{\mathrm{d}}{\to} W(1)\sum_{r=1}^j \alpha_r r^{-1}, \quad t\to\infty.$$

Integrating by parts we conclude that the left-hand side is equal to

$$\sum_{r=1}^j \alpha_r\Big(h_r(0)\frac{S^{\leftarrow}(t) - \mathtt{m}_1^{-1}t}{(\mathtt{m}_2\mathtt{m}_1^{-3}t)^{1/2}} + \int_{(0,\,1]}\frac{S^{\leftarrow}(t(1-x)) - \mathtt{m}_1^{-1}t(1-x)}{(\mathtt{m}_2\mathtt{m}_1^{-3}t)^{1/2}}\mathrm{d}_x h_r(tx)\Big).$$

For $r = 1, \ldots, j$, the functions h_r are nondecreasing by assumption and satisfy $\lim_{t\to\infty} h_r(t) = r^{-1}$ by Lemma 4.1. Also, relation (3.1) entails

$$\Big(\frac{S^{\leftarrow}(t(1-x)) - \mathtt{m}_1^{-1}t(1-x)}{(\mathtt{m}_2\mathtt{m}_1^{-3}t)^{1/2}}\Big)_{x\in[0,1]} \Rightarrow (W(1-x))_{x\in[0,1]}, \quad t\to\infty$$

in the J_1-topology on $D([0,1])$. With these at hand, an application of Lemma 4.2(b) ensures that

$$\sum_{r=1}^{j}\alpha_r\Big(h_r(0)\frac{S^{\leftarrow}(t)-\mathtt{m}_1^{-1}t}{(\mathtt{m}_2\mathtt{m}_1^{-3}t)^{1/2}}+\int_{(0,\,1]}\frac{S^{\leftarrow}(t(1-x))-\mathtt{m}_1^{-1}t(1-x)}{(\mathtt{m}_2\mathtt{m}_1^{-3}t)^{1/2}}\mathrm{d}_x h_r(tx)\Big)$$

$$\overset{\mathrm{d}}{\to}\ W(1)\sum_{r=1}^{j}\alpha_r h_r(0)+W(1)\sum_{r=1}^{j}\alpha_r(r^{-1}-h_r(0))=W(1)\sum_{r=1}^{j}\alpha_r r^{-1}.$$

Proof of (3.4). We use mathematical induction. If $r=1$, then

$$\int_0^t h_1(y)\mathrm{d}y=\Phi(\mathrm{e}^t)-\Phi(1)=\log t+O(1),\quad t\to\infty.$$

Assume that (3.4) holds with $r=s$. Then using this together with $h_s(t)=O(1)$ as $t\to\infty$, we infer

$$\int_0^t h_{s+1}(y)\mathrm{d}y=h_s(t)-h_s(0)+s(s+1)^{-1}\int_0^t h_s(y)\mathrm{d}y=(s+1)^{-1}\log t+O(1),\quad t\to\infty.$$

The proof of Proposition 2.3 is complete. □

4 Appendix

Here, we collect a couple of auxiliary results. Recall that

$$\Phi(t)=\int_{(0,\,\infty)}(1-\exp(-t(1-\mathrm{e}^{-x})))\nu(\mathrm{d}x),\quad t>0.$$

Lemma 4.1. *Let $c>0$. The relation*

$$\lim_{t\to\infty}(\Phi(\lambda t)-\Phi(t))=c\log\lambda \tag{4.1}$$

for all $\lambda>0$ implies that, for each $r\in\mathbb{N}$,

$$\lim_{t\to\infty}(-1)^{r+1}t^r\Phi^{(r)}(t)=c(r-1)!. \tag{4.2}$$

Conversely, if relation (4.2) *holds for some $r\in\mathbb{N}$, then* (4.1) *holds.*

Proof. Since Φ' is a nondecreasing function, relation (4.1) is equivalent to $\lim_{t\to\infty}t\Phi'(t)=c$ by Theorem 3.6.8 in [8]. For the converse part of the lemma, assume that (4.2) holds for some $r\geq 2$. Applying Karamata's theorem (Proposition 1.5.10 in [8]) $r-1$ times we obtain (4.2) with $r=1$.

For the direct part, assume that (4.2) holds with $r=1$. Now we shall prove by mathematical induction that it holds for all $r\in\mathbb{N}$. Indeed, let (4.2) hold with $r=k$. By Bernstein's criterion, the function Φ' is completely monotone. In particular, $(-1)^{k+2}\Phi^{(k+1)}$ is a nonincreasing function. Now (4.2) with $r=k+1$ follows from a version of the monotone density theorem (see the remark following Theorem 1.7.2 in [8]). □

The second result is essentially Lemma 6.4.2 in [13].

Lemma 4.2. *Let $a,b>0$. Assume that $X_t\Rightarrow X$ as $t\to\infty$ on $D([0,a])$ in the J_1-topology.*
(a) If $(\nu_t)_{t\geq 0}$ is a family of finite measures such that ν_t converge weakly to ν as $t\to\infty$, where ν is a finite measure on $[0,a]$ which is continuous with respect to Lebesgue measure, then

$$\int_{[0,\,a]}X_t(y)\nu_t(\mathrm{d}y)\ \overset{\mathrm{d}}{\to}\ \int_{[0,\,a]}X(y)\nu(\mathrm{d}y),\quad t\to\infty.$$

(b) If $f : [0, \infty) \to [0, \infty)$ *is a nondecreasing function satisfying* $\lim_{t\to\infty} f(t) = b$*, then*

$$\int_{(0,\,a]} X_t(y) \mathrm{d}_y f(ty) \stackrel{\mathrm{d}}{\to} X(0+)(b - f(0+)), \quad t \to \infty.$$

References

[1] Abramowitz, M. and Stegun, I.: Handbook of mathematical functions with formulas, graphs, and mathematical tables. *Dover*, New York, 1964. xiv+1046 pp. MR0167642

[2] Aldous, D.: Probability distributions on cladograms. In *Random discrete structures (Minneapolis, MN, 1993)*, vol. **76** of *IMA Vol. Math. Appl.*, 1–18, Springer, New York, 1996. MR1395604

[3] Aldous, D. and Janson, S.: The critical beta-splitting random tree IV: Mellin analysis of leaf height. *Electron. J. Probab.* **30**, (2025), article no. 69, 39 pp. MR4893424

[4] Aldous, D. J. and Janson, S.: The critical beta-splitting random tree II: overview and open problems. Preprint (2025) available at `https://arxiv.org/abs/2303.02529`

[5] Aldous, D. and Janson, S.: The critical beta-splitting random tree III: the exchangeable partition representation and the fringe tree. *Ann. Inst. Henri Poincaré, Probab. Stat.* (2025+), to appear. Preprint available at `https://arxiv.org/abs/2412.09655`

[6] Aldous, D. J., Janson, S. and Li, X.: The harmonic descent chain. *Electron. Commun. Probab.* **29**, (2024), article no. 77, 10 pp. MR4846604

[7] Aldous, D. and Pittel, B.: The critical beta-splitting random tree I: heights and related results. *Ann. Appl. Probab.* **35**, (2025), 158–195. MR4871707

[8] Bingham, N. H., Goldie, C. M. and Teugels, J. L.: Regular variation. *Cambridge University Press*, Cambridge, 1989. xx+494 pp MR1015093

[9] Brandenberger, A., Chin, B. and Mossel, E.: Asymptotics for the harmonic descent chain and applications to critical beta-splitting trees. Preprint (2025) available at `https://arxiv.org/abs/2505.24821`

[10] Gnedin, A. and Iksanov, A.: Regenerative compositions in the case of slow variation: a renewal theory approach. *Electron. J. Probab.* **17**, (2012), article no. 77, 19 pp. MR2981902

[11] Gnedin, A. and Pitman, J.: Regenerative composition structures. *Ann. Probab.* **33**, (2005), 445–479. MR2122798

[12] Gnedin, A., Pitman, J. and Yor, M.: Asymptotic laws for regenerative compositions: gamma subordinators and the like. *Probab. Theory Relat. Fields.* **135**, (2006), 576–602. MR2240701

[13] Iksanov, A.: Renewal theory for perturbed random walks and similar processes. *Birkhäuser*, Cham, 2016. xiv+250 pp MR585464

[14] Iksanov, A.: The harmonic descent chain and regenerative composition structures. *Electron. Commun. Probab.* **30**, (2025), article no. 11, 3 pp. MR865493

[15] Kolesnik, B.: Critical beta-splitting, via contraction. *Electron. Commun. Probab.* **30**, (2025), article no. 10, 14 pp. MR4865492

[16] Neininger, R. and Rüschendorf, L.: On the contraction method with degenerate limit equation. *Ann. Probab.* **32**, (2004), 2838–2856. MR2078559

[17] Vervaat, W.: Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **23**, (1972), 245–253. MR0321164