Adv. Appl. Probab. **56,** 1064–1105 (2024)
*doi:*10.1017/apr.2024.3

CONVERGENCE OF THE HEIGHT PROCESS OF SUPERCRITICAL GALTON–WATSON FORESTS WITH AN APPLICATION TO THE CONFIGURATION MODEL IN THE CRITICAL WINDOW

SERTE DONDERWINKEL ,* *McGill University*

Abstract

We show joint convergence of the Łukasiewicz path and height process for slightly supercritical Galton–Watson forests. This shows that the height processes for supercritical continuous-state branching processes as constructed by Lambert (2002) are the limit under rescaling of their discrete counterparts. Unlike for (sub-)critical Galton–Watson forests, the height process does not encode the entire metric structure of a supercritical Galton–Watson forest. We demonstrate that this result is nonetheless useful, by applying it to the configuration model with an independent and identically distributed power-law degree sequence in the critical window, of which we obtain the metric space scaling limit in the product Gromov–Hausdorff–Prokhorov topology, which is of independent interest.

Keywords: Random trees; branching processes; Galton–Watson trees; Bienaymé trees; metric space convergence

2020 Mathematics Subject Classification: Primary 60J80
Secondary 60C05

1. Introduction

In this work, we study the scaling limit of the genealogical structure of a slightly supercritical Galton–Watson forest by showing convergence of its height process under rescaling. The height process is defined using a depth-first exploration of the forest. Consequently, it encodes only a part of a supercritical Galton–Watson forest: a walker that executes a depth-first exploration will never cross the first path to infinity that it encounters. We show that our convergence result can nevertheless be applied, by using it to establish metric space convergence of the configuration model in the critical window. Moreover, our result shows that the height process for supercritical continuous-state branching processes (CSBPs) that was defined by Lambert in [41] is in fact the limit under rescaling of its discrete counterpart.

We will briefly introduce the encoding of a (random) forest by a *Łukasiewicz path* and a *height process* in the discrete setting. Furthermore, we will introduce the Łukasiewicz path and height process in the continuous setting. We then state our results and methods, after which we provide an overview of related work.

Received 12 November 2021; accepted 19 December 2023.
* Postal address: Burnside Hall, 805 Sherbrooke Street West, Montréal, Quebec H3A 0B9, Canada. Email address: serte.donderwinkel@mcgill.ca

1.1. Encoding forests by processes

The encoding of random trees and forests in the discrete setting and the continuum by (excursions of) random processes has been around for a long time; see e.g. [5, 24, 25, 28, 43, 48, 50].

Let T be an ordered rooted finite tree, say $|T| = n$. Let $v_0, \dots, v_{n-1}$ denote the vertices of the tree visited in depth-first order, so that v_0 is the root of the tree. We will define the height function and Łukasiewicz path of T. Both of these functions uniquely characterize T. The height process of T, referred to as h, is defined as

$$h(i) = d_T(v_i, v_0);$$

i.e. for all i, $h(i)$ equals the distance from v_i to the root. Moreover, for all $i \in \{1, \dots, n\}$, let y_i be the number of children of v_{i-1}, and set $y_0 = 1$. Then the Łukasiewicz path of T is defined by

$$x(i) = \sum_{0 \le j \le i} (y_j - 1)$$

for $i = 0, \dots, n$. Thus, $x(i)$ is the total number of younger siblings of v_i and its ancestors, where *younger* means that they are explored later in the depth-first search. Note that

$$h(i) = \#\{j < i : x(j) = \min\{x(k) : j \le k \le i\}\}, \tag{1}$$

which is proved by Le Gall in [42]. These encodings of trees by walks can easily be extended to ordered forests by concatenating the Łukasiewicz paths and height processes.

We can use the correspondence between forests and their Łukasiewicz paths to construct Galton–Watson forests from random walks. Suppose $(D_1, D_2, \dots)$ are independent and identically distributed (i.i.d.) random variables with law π on $\mathbb{N} = \{0, 1, 2 \dots\}$. Then

$$(S(k), k \ge 0) := \left(\sum_{i=1}^{k} (D_i - 1), k \ge 0 \right)$$

is the Łukasiewicz path of a random forest in which all vertices have independent offspring with law π. We refer to such a forest as a π-Galton–Watson forest (GW (π)). We will write $(H(k), k \ge 0)$ to denote the height process corresponding to $(S(k), k \ge 0)$.

The continuous counterpart of the Galton–Watson process is a *continuous-state branching process* (CSBP) (see e.g. [30]). The rôle of the Łukasiewicz path is played by a Lévy process $(L_t, t \ge 0)$ without negative jumps (i.e. a *spectrally positive* Lévy process). The law of L is completely characterized by its Laplace exponent ϕ, defined via

$$\mathbb{E}\left[\exp(-\theta L_t)\right] = \exp(t\phi(\theta)).$$

Then, for $\zeta_L = \inf\{t > 0 : L_t \le 0\}$,

$$\varphi(t) = \int_0^t \frac{1}{L_s} ds, \quad t \in (0, \zeta_L],$$

and $\varphi^{-1} : (0, \varphi(\zeta_L)) \to \mathbb{R}$, the inverse of φ, the CSBP with branching mechanism ϕ, which we refer to as CSBP (ϕ), can be defined as

$$\left(L_{\varphi^{-1}(t)}, 0 \le t \le \zeta_L\right).$$

We wish to define the height process corresponding to $(L_t, t \geq 0)$ (which encodes the genealogy of the CSBP corresponding to the consecutive excursions of $(L_t, t \geq 0)$ above its infimum) analogously to (1), so we should define a functional $H(L) = (H_t, t \geq 0)$ of L such that H_t in some sense measures the 'size' of the set $\{s \leq t : L_{s-} = \inf\{L_r : r \in [s, t]\}\}$. In [43], it was established that if L almost surely does not drift to ∞ and satisfies

$$\int_1^\infty \frac{du}{\phi(u)} < \infty,$$

then there exists a continuous process $(H_t, t \geq 0)$ such that

$$H_t = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^t \mathbb{1}_{\{L_s - \inf\{L_r : r \in [s,t]\} \leq \epsilon\}},$$

with the limit in probability for all $t \in [0, \infty)$. Further results were proved in [24]. The excursions of $(H_t, t \geq 0)$ above 0 encode metric spaces called *Lévy trees*, which are defined in [23]. In [41], the definition of $H(L)$ was extended to spectrally positive Lévy processes L that drift to ∞ almost surely.

1.2. Results and methods

For each n, let $D_1^n, D_1^2, \ldots$ be an i.i.d. sequence of random variables with law π_n, and set

$$(S^n(k), k \geq 0) = \left(\sum_{i=1}^k (D_i^n - 1), k \geq 0 \right).$$

Let $(H^n(k), k \geq 0)$ be the corresponding height process as defined in Section 1.1. We impose the following conditions:

(C1) There exist a nondecreasing sequence of positive integers $(\gamma_n, n \geq 1)$ that converges to ∞ and a Lévy process $(L_t, t \geq 0)$ on $\mathbb{R}$ that does not have downward jumps and is of infinite variation, such that

$$\left(n^{-1} S^n_{\lfloor n\gamma_n t \rfloor}, t \geq 0 \right) \xrightarrow{d} (L_t, t \geq 0) \tag{2}$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$.

(C2) For Y_m^n, the number of vertices at height m in the first n trees of the forest encoded by $(S^n(k), k \geq 0)$, for every $\delta > 0$, we have

$$\liminf_{n \to \infty} \mathbb{P}\left[Y^n_{\lfloor \delta \gamma_n \rfloor} = 0 \right] > 0. \tag{3}$$

The following result is proved in [24].

Theorem 1. ([Theorem 1.4.3 and Theorem 2.2.1].) *Suppose Conditions (C1) and (C2) hold, and in addition the following hold:*

(C3*) *The Laplace exponent $\phi(\theta)$ of $(L_t, t \geq 0)$ satisfies*

$$\int_1^\infty \frac{du}{\phi(u)} < \infty.$$

(C4*) *Almost surely, L_t does not drift to ∞ as $t \to \infty$.*

(C5*) *We have $\mathbb{E}[D_1^n] \leq 1$ for all n.*

Then there exists a continuous modification of the height process of $(L_t, t \geq 0)$, say $(H_t, t \geq 0)$, such that

$$\left(\gamma_n^{-1} H^n_{\lfloor n\gamma_n t \rfloor}, t \geq 0\right) \xrightarrow{d} (H_t, t \geq 0)$$

as $n \to \infty$, jointly with (2).

Our main result is the analogue of Theorem 1 for supercritical Galton–Watson processes.

Theorem 2. *Suppose Conditions (C1) and (C2) hold, and in addition the following hold:*

(C3) *For $\phi(\theta)$ the Laplace exponent of $(L_t, t \geq 0)$ and $\xi > 0$ the unique value such that $\phi(\xi) = 0$, we have*

$$\int_1^\infty \frac{du}{\phi(u + \xi)} < \infty.$$

(C4) *We have $L_t \to \infty$ almost surely as $t \to \infty$.*

(C5) *We have $\mathbb{E}[D_1^n] > 1$ for all n.*

Then there exists a continuous modification of the height process of $(L_t, t \geq 0)$, say $(H_t, t \geq 0)$, such that

$$\left(\gamma_n^{-1} H^n_{\lfloor n\gamma_n t \rfloor}, t \geq 0\right) \xrightarrow{d} (H_t, t \geq 0)$$

as $n \to \infty$, jointly with (2).

Roughly speaking, Condition (C2) ensures that the maximal height in a forest of a number n of GW (π_n) trees, rescaled by γ_n^{-1}, may not exceed δ. An analogous fact holds for CSBP (ϕ). Informally, Condition (C3) ensures that a Lévy tree encoded by an excursion of L_t above its infimum, conditional on the excursion being finite, has finite height almost surely. This is a necessary condition for a continuous modification of the height process to exist. Finally, Conditions (C4) and (C5) ensure that GW (π_n) and CSBP (ϕ) respectively are supercritical.

Our method is based on the following pathwise construction of $(S^n(k), H^n(k), k \geq 0)$:

1. Up to the first time that it hits its overall infimum, $(S^n(k), k \geq 0)$ encodes a random number of finite trees, which are independent and each have the law of a tree in GW (π_n) conditioned to be finite. This yields a pathwise construction of $(S^n(k), H^n(k), k \geq 0)$ up to the first time that $(S^n(k), k \geq 0)$ hits its overall infimum.

2. After the first time that it hits its overall infimum, $(S^n(k), k \geq 0)$ encodes an infinite spine with trees attached to the left of the infinite spine, and vertices attached to the right of the infinite spine. To be precise, for every vertex v on the infinite spine, a random number of independent copies of a tree in GW (π_n) conditioned to be finite are attached to v left of the infinite spine, and a random number of vertices (that will never be visited) are attached to the right of the infinite spine. This yields a pathwise construction of $(S^n(k), H^n(k), k \geq 0)$ after the first time that $(S^n(k), k \geq 0)$ hits its overall infimum. This

construction is similar to the pathwise construction for Galton–Watson processes with immigration defined in [25].

(See e.g. [36], [44, Section 5.7], and [6, Chapter I.12] for the laws of the trees encoded by $(S^n(k), k \geq 0)$ before and after it hits its overall infimum, although height processes are not considered in these works.)

We define a similar pathwise construction of $(L_t, H_t, t \geq 0)$, which is standard for the pre-infimum process (see e.g. [10, Chapter VII]), and based on [41, Section 5] for the post-infimum process. We then show convergence in distribution under rescaling of the pathwise construction of $(S^n(k), H^n(k), k \geq 0)$ to the pathwise construction of $(L_t, H_t, t \geq 0)$.

Finally, in Section 4 we use Theorem 2 to show metric space convergence of the largest components of a uniform graph with an i.i.d. heavy-tailed degree sequence in the critical window, which extends the main result of [17]. For a precise statement of this result, and an overview of earlier work on related graph models, we refer the reader to Section 4.

1.3. Related work

While there is an extensive literature concerning the convergence of height processes for critical branching processes, we are only aware of two works that consider the supercritical case: [14, 22]. In [14, Theorem B2], Broutin, Duquesne, and Wang discuss convergence of the height process of a model similar to the model we consider here. However, in that work, the authors only consider the forest before the first infinite line of descent. In [22], Duquesne studies a class of supercritical branching processes with a single infinite line of descent (*sin-trees*). The author encodes the resulting trees with two processes, one that encodes the genealogical structure on the left, and one that encodes the genealogical structure on the right side of the infinite line of descent. For this model, convergence under rescaling of the discrete contour processes to the continuum height process is proved. The pathwise construction that we use resembles the pathwise construction used in [22].

In [26], convergence of supercritical Galton–Watson trees under rescaling is studied through a different lens. In Theorem 4.15 of that work, Duquesne and Winkel show the convergence of a class of supercritical Galton–Watson forests to a Lévy forest in the sense of Gromov–Hausdorff convergence on locally compact rooted real trees. Although their theorem applies to the entire forest, and not just to the trees and pendant subtrees to the left of the first infinite line of descent, convergence in the Gromov–Hausdorff topology on locally compact rooted real trees does not imply convergence of 'depth-first ordering' of the vertices in the tree, as convergence of the height process does.

An alternative approach to viewing the height process of a supercritical branching process is introduced in [18] for branching processes with a quadratic branching mechanism, and extended to general CSBPs in [1]. The approach is to build the super-critical tree up to a given level a, such that the tree can be encoded by a height process. Then, the law of the resulting tree is absolutely continuous with respect to the law of a (sub-)critical Lévy tree whose branches above level a are removed, which is referred to as *pruning*. Furthermore, this family of processes indexed by a satisfies a consistency property, and hence there exists a projective limit. It is established in [2] that the limit object has the same law as the supercritical Lévy tree defined in [25]. In [32], He and Luan define an analogous pruning operation for supercritical Galton–Watson trees and prove that the contour functions of these truncated Galton–Watson trees converge weakly to the processes constructed by Abraham and Delmas in [1].

2. The construction of the height process

In this section, we will combine results from the literature in order to give a construction of the height process of a supercritical Galton–Watson process in the discrete case, and of a supercritical CSBP in the continuum.

2.1. The height process in the discrete case

In this section, we describe a pathwise construction of the height process corresponding to S^n. We do this by considering separately the process before and after it hits its overall infimum. This corresponds to considering the laws of finite and infinite trees in a supercritical branching process separately; this idea can be traced back to Harris [31]. In the descriptions of the two parts, we will make use of a process that is locally absolutely continuous to S^n, which we will denote by $\hat{S}^n$. We will refer to to $\hat{S}^n$ as 'S^n conditioned to die out'. This formalizes, in the random walk framework, the well-known fact that a supercritical Galton–Watson process conditioned to die out is a subcritical Galton–Watson process.

For $t \geq 0$, define

$$\mathcal{L}_{S^n(1)}(\theta) = \mathbb{E}\left[\exp(-\theta S^n(1))\right],$$

and set $\phi_n(\theta) = \log \mathcal{L}_{S^n(1)}(\theta)$. Since $\mathbb{E}[S^n(1)] > 0$, we have $\mathcal{L}_{S^n(1)}{}'(0) < 0$, so by the convexity of $\mathcal{L}_{S^n(1)}(\theta)$ and the fact that $\mathcal{L}_{S^n(1)}(0) = 1$, there is a unique $\xi_n > 0$ such that $\mathcal{L}_{S^n(1)}(\xi_n) = 1$ and $\phi_n(\xi_n) = 0$. Let $\mathbb{P}_n$ be the law of S^n, and let

$$\mathcal{F}_k^n := \sigma(S^n(m),\, m \leq k)$$

be the natural filtration of S^n. Let $\mathbb{P}_n^{\#}$ be locally absolutely continuous with respect to $\mathbb{P}_n$, with

$$\mathrm{d}\mathbb{P}_n^{\#}|_{\mathcal{F}_k^n} = \exp\big(-\xi_n S^n(k)\big)\,\mathrm{d}\mathbb{P}_n|_{\mathcal{F}_k^n},$$

and let $\hat{S}^n$ be a process which under $\mathbb{P}_n$ has the law of S^n under $\mathbb{P}_n^{\#}$. Let

$$\tau^n(m) = \inf\left\{k : S^n(k) = -m\right\},$$

so that $\{\tau^n(m) < \infty\}$ may be interpreted as the event that at least m trees in the Galton–Watson forest are finite. The following properties of ξ_n and $\hat{S}^n$ are standard (see e.g. [6, Chapter I.12]).

Theorem 3. *The following hold:*

1. *For any $m \geq 0$,*

$$\exp(-\xi_n m) = \mathbb{P}_n\left[\tau^n(m) < \infty\right].$$

2. *If $\Lambda \in \mathcal{F}^n_{\tau^n(m)}$ for some $m > 0$, then*

$$\mathbb{P}_n^{\#}[\Lambda] = \mathbb{P}_n\left[\Lambda \,|\, \tau^n(m) < \infty\right].$$

3. *The process $\hat{S}^n$ is a downward skip-free random walk on the integers.*

The following lemma is a first reason why $\hat{S}^n$ plays an important rôle in the pathwise construction of S^n.

Lemma 1. *Let G^n be a geometric random variable with success probability equal to $1 - \exp(-\xi_n)$, i.e. $\mathbb{P}_n(G^n = k) = \exp(-k\xi_n)(1 - \exp(-\xi_n))$, independent of $\hat{S}^n$. Then the pre-infimum process of S^n has the law of $\hat{S}^n$, stopped at the first time it reaches level G^n.*

Proof. Note that the negative of the overall infimum of S^n, which we denote by $-I^n_\infty$, equals the number of finite trees in the forest defined by S^n viewed as a Łukasiewicz path. Hence, it is distributed as a geometric random variable with parameter $\exp(-\xi_n)$. Let ρ^n denote the time when S^n first reaches I^n_∞. We have that

$$(S^n(j) : 0 \le j \le \rho^n) \text{ under } \mathbb{P}_n(\ \cdot\ | I^n_\infty = -m)$$

has the same distribution as

$$(S^n(j) : 0 \le j \le \tau^n(m)) \text{ under } \mathbb{P}_n\left(\ \cdot\ | \tau^n(m) < \infty\right),$$

which, by Theorem 3, is equal in distribution to

$$(\hat{S}^n(j) : 0 \le j \le \tau^n(m)) \text{ under } \mathbb{P}^{\#}_n.$$

Combining this with the distribution of I^n_∞, we find that for G_n a random variable with the geometric distribution with success probability $\exp(-\xi_n)$ independent of S^n, $(S^n(j) : 0 \le j \le \rho^n)$ under $\mathbb{P}_n$ has the same distribution as $(\hat{S}^n(j) : 0 \le j \le \tau^n(G^n))$ under $\mathbb{P}^{\#}_n$. □

By Theorem 3, $\hat{S}^n$ is the Łukasiewicz path of a subcritical Galton–Watson forest. We let $\hat{H}^n$ be its height process as defined in (1), i.e.

$$\hat{H}^n(i) = \#\left\{j < i : \hat{S}^n(j) = \min\{\hat{S}^n(k) : j \le k \le i\}\right\}.$$

Then Lemma 1 has the following corollary.

Corollary 1. *We have that*

$$(S^n(m), H^n(m), m \le \rho^n) \overset{d}{=} (\hat{S}^n(m), \hat{H}^n(m), m \le \tau^n(G^n)).$$

We will now characterize the post-infimum process and its height process. Let $\mathbb{P}^{\uparrow}_n$ be the law of $(S^n(k), k \ge 0)$ conditioned to be non-negative for all k. Note that we are conditioning on an event of non-zero probability because S^n is supercritical, so this is well-defined.

The following lemma suggests an adaptation of S^n that has the same height process, and that turns out to be easier to work with.

Lemma 2. *For a discrete Łukasiewicz path $(X_k, k \ge 0)$, set $\underline{\underline{X}}_n = \min\{X_m, m \ge n\}$. We will refer to $\underline{\underline{X}}_n$ as the future infimum process. Then, the height processes of X and $X - \underline{\underline{X}}$ are the same.*

Proof. Let H^* denote the height process of $X - \underline{\underline{X}}$. Then, by definition,

$$H^*_n = \#\left\{m < n : X_m - \underline{\underline{X}}_m = \min\{X_j - \underline{\underline{X}}_j, m \le j \le n\}\right\}.$$

Set $\underline{\underline{g}}(n) = \max\{m < n : X_m = \underline{\underline{X}}_m\}$. Then, for $m \le \underline{\underline{g}}(n)$, we have that $\underline{\underline{X}}_m = \min\{X_j : m \le j \le n\}$ and $\min\{X_j - \underline{\underline{X}}_j, m \le j \le n\} = 0$. Also, for $m > \underline{\underline{g}}(n)$, we have $\underline{\underline{X}}_m = \underline{\underline{X}}_n$. Hence, indeed,

$$\begin{aligned} H^*_n =& \#\left\{m \le \underline{\underline{g}}(n) : X_m - \underline{\underline{X}}_m = \min\{X_j - \underline{\underline{X}}_j, m \le j \le n\}\right\} \\ &+ \#\left\{\underline{\underline{g}}(n) < m < n : X_m - \underline{\underline{X}}_m = \min\{X_j - \underline{\underline{X}}_j, m \le j \le n\}\right\} \end{aligned}$$

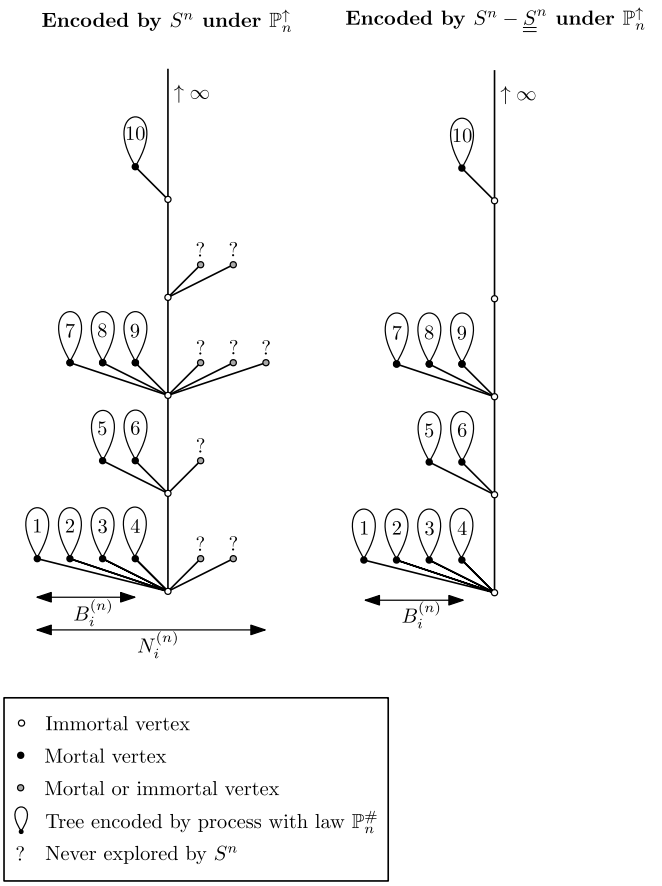

FIGURE 1. The information captured in S^n and $S^n - \underline{\underline{S}}^n$ under $\mathbb{P}_n^{\uparrow}$.

$$
\begin{aligned}
&= \#\left\{m \le \underline{\underline{g}}(n) : X_m - \min\{X_j : m \le j \le n\} = 0\right\} \\
&\quad + \#\left\{\underline{\underline{g}}(n) < m < n : X_m - \underline{\underline{X}}_n = \min\{X_j - \underline{\underline{X}}_n,\, m \le j \le n\}\right\} \\
&= \#\left\{m \le n : X_m = \min\{X_j,\, m \le j \le n\}\right\} = H_n,
\end{aligned}
$$

where H_n is the height process of X_n. □

Note that since S^n drifts to $+\infty$ almost surely, $S^n - \underline{\underline{S}}^n$ under $\mathbb{P}_n^{\uparrow}$ is recurrent in zero. Indeed, if S^n drifts to $+\infty$, we have $\underline{\underline{S}}^n(k) > -\infty$ for all $k > 0$. Then, by the definition of $\underline{\underline{S}}^n$, there is a finite $l > k$ such that $S^n(l) = \underline{\underline{S}}^n(l)$. Moreover, $S^n - \underline{\underline{S}}^n$ is non-negative. Combining these facts implies that $S^n - \underline{\underline{S}}^n$ is the Łukasiewicz path of an infinite spine with finite trees attached to it only on the left-hand side. S^n describes the same infinite spine with trees on its left-hand side, but also contains information on the total number of children (including children to the right of the spine) of vertices on the spine. See Figure 1. We will use this description to study the distributions of these paths.

Using terminology introduced by Janson in [36], we will call a vertex *immortal* if it is the root of an infinite tree. Otherwise we will call it *mortal*. Note that since S^n drifts to $+\infty$

almost surely, every vertex has a positive probability of being immortal. Hence there are almost surely countably infinitely many infinite spines. Note that in every generation, we only visit the vertices to the left of the leftmost immortal vertex. This means that, under $\mathbb{P}_n^{\uparrow}$, an excursion of $S^n - \underline{\underline{S}}^n$ above 0 consists of an increment (of size, say, $B^{(n)}$) corresponding to the mortal older brothers of the first immortal vertex, and then an excursion with the law of S^n starting at $B^{(n)}$ conditioned to hit 0 in finite time. This corresponds to sampling first the number of trees to the left of the infinite spine and then the shapes of those trees.

The sample paths of S^n can be constructed from the sample paths of $S^n - \underline{\underline{S}}^n$, by adding the randomness that encodes the number of younger brothers of vertices on the leftmost path to infinity. This corresponds to replacing the jumps of size $B^{(n)}$ by jumps of size $N^{(n)} - 1$, with $N^{(n)}$ being distributed as the total generation size given $B^{(n)}$. This is illustrated in Figure 1. (A similar decomposition of Galton–Watson trees conditioned on non-extinction is discussed by Lyons and Peres in Section 5.7 of [44], where they do not introduce the encoding by a Łukasiewicz path and height process.)

Note that the height at time k in the tree encoded by S^n under $\mathbb{P}_n^{\uparrow}$ is given by the height of the position along the infinite spine where the finite subtree containing the vertex visited at time k is attached plus the height of the vertex in the finite subtree.

We will now investigate the joint distribution of $N^{(n)}$ and $B^{(n)}$.

Lemma 3. *Let $N^{(n)}$ be the number of children in a set of offspring conditioned to contain at least one immortal vertex, let $B^{(n)}$ be the number of older brothers of the oldest immortal vertex in such a set of offspring, and let $\exp(-\xi_n)$ be the probability that, under $\mathbb{P}_n$, a tree dies out. Then for $l > k$ we have*

$$\mathbb{P}(B^{(n)} = k, N^{(n)} = l) = \exp(-k\xi_n)\mathbb{P}(S_1^n = l - 1). \tag{4}$$

Proof. Let ψ_n denote the probability generating function of the offspring distribution under $\mathbb{P}_n$ (i.e. the law of Z_1^n under $\mathbb{P}_n$). Let M_n be the random variable representing the number of mortal children of an immortal parent, and let J_n be the random variable representing the number of immortal children of an immortal parent. Then, in [36], Janson gives the generating function of the joint law of M_n and J_n as

$$\mathbb{E}\left[x^{M_n} y^{J_n}\right] = \frac{\psi_n\left(\exp(-\xi_n)x + (1 - \exp(-\xi_n))y\right) - \psi_n\left(\exp(-\xi_n)x\right)}{1 - \exp(-\xi_n)}. \tag{5}$$

Also, given the number of mortal and immortal children, they appear in a uniformly random order. It is then straightforward that the generating function of the total number of children of an immortal parent is given by

$$\mathbb{E}\left[x^{N^{(n)}}\right] = \frac{\psi_n(x) - \psi_n(\exp(-\xi_n)x)}{1 - \exp(-\xi_n)},$$

so we obtain that for $k = 1, 2, \ldots,$

$$\mathbb{P}(N^{(n)} = k) = \frac{1 - \exp(-k\xi_n)}{1 - \exp(-\xi_n)}\mathbb{P}(S_1^n = k - 1).$$

Using (5) to analyze the generating function of the joint law of $N^{(n)}$ and J_n, we see that, conditional on the value of $N^{(n)}$, J_n is distributed as a binomial random variable with parameters $N^{(n)}$ and $1 - \exp(-\xi_n)$ conditioned to be at least 1. Since the mortal and immortal children

appear in a uniform order, conditional on the generation size $N^{(n)}$, the number of mortal older brothers of the first immortal vertex, $B^{(n)}$, is distributed as a geometric random variable with success parameter $\exp(-\xi_n)$ conditioned to be at most $N^{(n)}-1$. We obtain that

$$\mathbb{P}(B^{(n)}=k|N^{(n)}=l)=\frac{\exp(-k\xi_n)(1-\exp(-\xi_n))}{1-\exp(-l\xi_n)},$$

which proves the statement. □

These findings on the post-infimum process can be summarized as follows.

Proposition 1. *The sample paths of S^n and $S^n-\underline{\underline{S}}^n$ under $\mathbb{P}_n^{\uparrow}$ can be constructed by concatenating excursions above the future infimum as follows. Sample a countably infinite number of independent copies of $B^{(n)}$ and $N^{(n)}$ according to*

$$\mathbb{P}(B^{(n)}=k,\,N^{(n)}=l)=\begin{cases}\exp(-k\xi_n)\mathbb{P}(S_1^n=l-1), & 0\le k<l,\\ 0 & \text{otherwise.}\end{cases}$$

Start an excursion with an increment of size $N^{(n)}-1$ above the previous future infimum, and continue from there as a process with law $\mathbb{P}_n^{\#}$. Kill the excursion at $N^{(n)}-B^{(n)}-1$ above the previous future infimum, which will be the new future infimum.

By replacing the jumps of size $N^{(n)}-1$ by jumps of size $B^{(n)}$ we obtain $S^n-\underline{\underline{S}}^n$.

2.1.1. *A pathwise construction.* The characterization of the pre- and post-infimum processes justifies the following pathwise construction of $(S^n, S^n-\underline{\underline{S}}^n, H^n)$ under $\mathbb{P}_n$. This is similar to the pathwise construction of the encoding processes of a Galton–Watson process with immigration given in Section 2.2 of [22].

See Figure 2 for a graphical representation of the construction.

- Let $\hat{S}^n(k)$ be a random walk with law $\mathbb{P}_n^{\#}$, and let $\hat{H}^n(k)$ be the corresponding height process. Set $\hat{I}^n(k)=\min_{m\le k}\hat{S}^n(k)$. The trees encoded by $\hat{S}^n(k)$ will be used as the finite trees that are explored before the first infinite tree is encountered, and as the finite subtrees to the left of the leftmost infinite spine which are rooted at the vertices on the infinite spine.
- Let G^n be distributed as a geometric random variable with mean $\exp(-\xi_n)$, which gives the number of finite trees explored by the process. We then let $(B_1^{(n)}, N_1^{(n)}),(B_2^{(n)}, N_2^{(n)}),\dots$ be i.i.d. pairs of random variables, independent of $\hat{S}^n$, distributed as $(B^{(n)}, N^{(n)})$. Set $Q^n(0)=G_n$, $Q^n(k)=G_n+\sum_{i=1}^{k}N_i^{(n)}$ for $k\ge 1$. Then $N_m^{(n)}$ will be the number of children attached to the mth vertex on the infinite spine. Also, set $R^n(0)=G_n$ and $R^n(k)=G_n+\sum_{i=1}^{k}B_i^{(n)}$ for $k\ge 1$. Then $B_m^{(n)}$ will be the number of finite subtrees to the left of the infinite spine rooted at the mth vertex on the infinite spine.
- Define $F^n(k)=\inf\{l\le k-1\colon -\hat{I}^n(k-1-l)<R^n(l)\}$. Then $F^n(k)$ will be the number of vertices located on the leftmost infinite spine among the first k vertices that we visit in our depth-first exploration.
- Then define

$$S^n(k)=-G^{(n)}+\hat{S}^n\left(k-F^{(n)}(k)\right)+Q^n\left(F^{(n)}(k)\right)-F^n(k),$$

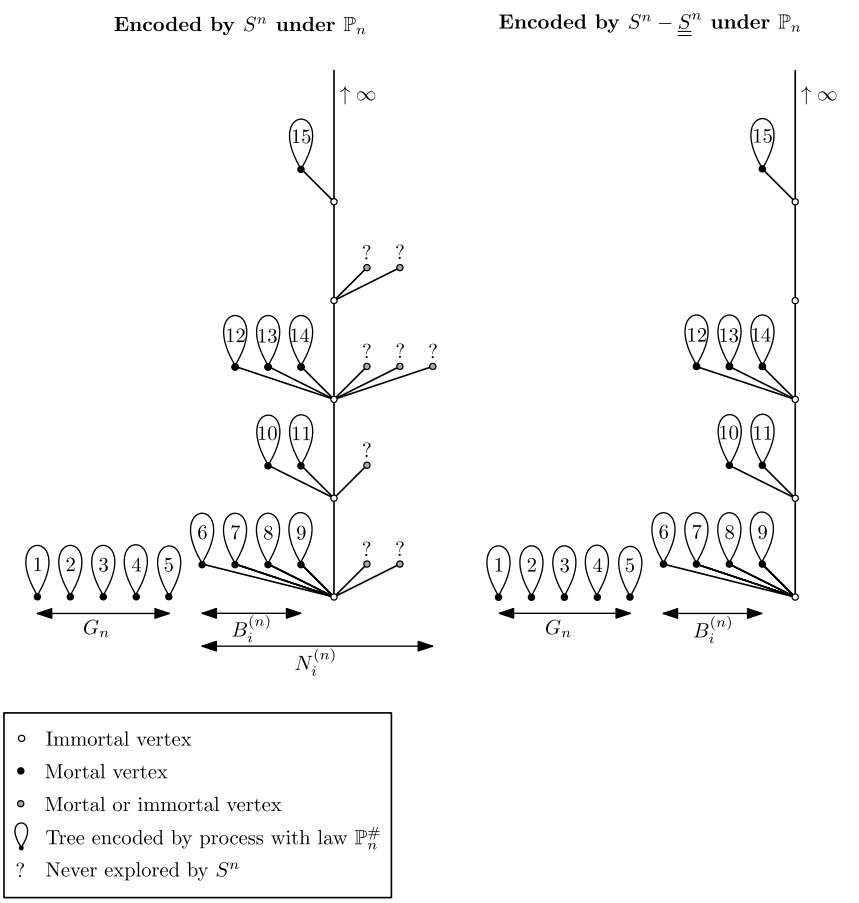

FIGURE 2. The information captured in S^n and $S^n - \underline{\underline{S}}^n$ under $\mathbb{P}_n$. The finite trees are encoded by the pre-infimum process, and the infinite spine and its pendant subtrees are encoded by the post-infimum process.

so that

$$(S^n - \underline{\underline{S}}^n)(k) = \hat{S}^n(k - F^{(n)}(k)) + R^n\left(F^{(n)}(k)\right)$$

and

$$\begin{aligned} H^n(k) = &\left(\hat{H}^n(k - F^{(n)}(k)) + F^{(n)}(k)\right) \mathbb{1}_{\{F^{(n)}(k-1)=F^{(n)}(k)\}} \\ &+ \left(\hat{H}^n\left(k - (F^{(n)}(k) - 1)\right) + F^{(n)}(k) - 1\right) \mathbb{1}_{\{F^{(n)}(k-1)<F^{(n)}(k)\}}. \end{aligned} \tag{6}$$

By Corollary 1, Proposition 1, and Lemma 2, the constructed process has the desired distribution.

In the next section, we will introduce the continuous counterpart of this construction. In Section 3, we will use Equation (6) to show the joint convergence under rescaling of the discrete processes to the continuous processes.

2.2. The height process of a supercritical Lévy process

Just as in the discrete case, we will obtain a pathwise construction of a supercritical spectrally positive Lévy process and its height process by considering its pre- and post-infimum processes separately. As before, let L be such a Lévy process with Laplace exponent ϕ. Define $I_\infty = \inf\{L_t, t \geq 0\}$, and set $\rho = \sup\{t > 0 : L_t = I_\infty\}$. The process on $[0, \rho]$ will be referred to as the *pre-infimum process*, and the process on $[\rho, \infty)$ (shifted up by $-I_\infty$ so that it starts at 0) will be referred to as the *post-infimum process*. Informally, as in the discrete case, the pre-infimum process encodes the $\mathbb{R}$-trees without a path of infinite length, and the post-infimum process encodes the metric structure to the left of the leftmost path of infinite length in the first $\mathbb{R}$-tree that contains such a path.

(In [9], Bertoin uses a different strategy, and splits the process at the infimum attained on a compact time interval. We will not use this approach, and so we will not discuss his results here.)

On an infinite time horizon, the following results on spectrally positive Lévy processes are available:

1. By Théorème 2 in Bertoin [8], the pre-infimum process of a spectrally positive Lévy process that drifts to $+\infty$ has the same law as the process 'conditioned to drift to $-\infty$' stopped at an independent exponential level. Informally, this result says that the pre-infimum process encodes $\mathbb{R}$-trees conditioned to die out.

2. By Lemma 4.1 in Millar [45] (which is rephrased as 'Théorème (Millar)' in [8]; we will use the latter version of the result), the post-infimum process of a spectrally positive Lévy process that drifts to $+\infty$ has the same law as the process conditioned to stay positive, and is independent of the pre-infimum process. Informally, this result says that the post-infimum process encodes (part of) a single $\mathbb{R}$-tree conditioned to be infinite.

3. In [41], Lambert describes the height process corresponding to a class of spectrally positive Lévy processes conditioned to stay positive.

We start with an overview of the results by Bertoin [8] that we use. Firstly, since L is supercritical, there is a unique $\xi > 0$ such that $\phi(\xi) = 0$. Let $\mathbb{P}$ be the law of L, and set

$$\mathcal{F}_t := \sigma(L_s, s \leq t).$$

Then $(\exp(-\xi L_t), t \geq 0)$ is a $\mathbb{P}$-martingale. Let $\mathbb{P}^\#$ be locally absolutely continuous with respect to $\mathbb{P}$, with

$$\mathrm{d}\mathbb{P}^\#|_{\mathcal{F}_t} = \exp(-\xi L_t)\mathrm{d}\mathbb{P}|_{\mathcal{F}_t}. \tag{7}$$

Let $\hat{L}$ be a process which under $\mathbb{P}$ has the law of L under $\mathbb{P}^\#$. The following analogue of Theorem 3 is a straightforward consequence of the fact that $\exp(-\xi L_t)$ is a martingale. See [10, Chapter VII].

Theorem 4. *The following hold:*

1. *For* $\tau(x) = \inf\{t > 0 : L_t = -x\}$,

$$\mathbb{P}[\tau(x) < \infty] = \exp(-\xi x).$$

2. *If* $\Lambda \in \mathcal{F}_{\tau(x)}$ *for some* $x > 0$, *then*

$$\mathbb{P}^\#[\Lambda] = \mathbb{P}[\Lambda | \tau(x) < \infty].$$

3. *We have that $\mathbb{P}^{\#}$ is the law of a spectrally positive subcritical Lévy process with Laplace exponent $\phi^{\#}(\cdot) = \phi(\cdot + \xi)$.*

The following theorem is then proved in [8] as Théorème 2.

Theorem 5. ([8, Thàreme 2].) *Let E be an exponential random variable with rate ξ. The pre-infimum process of L has the law of $\hat{L}$, independent of E, stopped at the first time it reaches level E.*

These observations, together with Proposition 1.4.3 of Duquesne and Le Gall [24], imply the following proposition.

Proposition 2. *There exists a continuous modification of the height process corresponding to $\hat{L}$, which we will denote by $\hat{H}$. Moreover, for*

$$\rho = \sup\{t : L_t = \inf_{s \geq 0} L_s\},$$

we have $\rho < \infty$ almost surely; furthermore, there exists a continuous modification of the height process of L up to ρ, which we will refer to as H, and for E an exponential random variable with rate ξ,

$$(L_t, H_t, t \leq \rho) \overset{d}{=} (\hat{L}_t, \hat{H}_t, t \leq \tau(E)).$$

Proof. Since L is supercritical, ρ is finite almost surely. Theorem 4.3 and Condition (C3) ensure that the conditions of Theorem 1.4.3 in [24] are satisfied, which implies that the height process corresponding to $\hat{L}$ exists and has a continuous modification. Theorem 5 then yields the proposition. □

We will now focus on the post-infimum process. We will use two important results from the literature. Firstly, by Lemma 4.1 in Millar [45], the post-infimum process $(L_{\rho+t} - I_\infty, t \geq 0)$ has the same law as L conditioned to stay positive and is independent of the pre-infimum process. Call the law of L conditioned to stay positive $\mathbb{P}^{\uparrow}$. For the definition of this process, see [8, 15, 16]. The height process of L under $\mathbb{P}^{\uparrow}$ is characterized by Lambert in [41], and is obtained via the continuous counterpart of the construction in Section 2.1. Indeed, in [41], Lambert defines

$$\underline{\underline{L}}_t := \inf\{L_s, s \geq t\},$$

and in [41, Lemma 8ii] he shows that the height processes of $L - \underline{\underline{L}}$ and L are equal. Then, in [41, Theorem 3], he gives a pathwise construction of $L - \underline{\underline{L}}$ and its local time at 0 under $\mathbb{P}^{\uparrow}$, by viewing $L - \underline{\underline{L}}$ as a continuous-time branching process with immigration, an object introduced in [40]. Lemma 8ii of [41] also illustrates how to construct the height process corresponding to $L - \underline{\underline{L}}$. Translating these results to our setting yields the following proposition, which is the continuous counterpart of the pathwise construction of $S^n - \underline{\underline{S}}^n$ under $\mathbb{P}_n^{\uparrow}$ described in Proposition 1.

Theorem 6. ([41, Theorem 3, Lemma 8ii].) *Recall the definition of $\mathbb{P}^{\#}$ from (7). Let $\hat{L}$ be a process which under $\mathbb{P}$ has the law of L under $\mathbb{P}^{\#}$, and let $\phi^{\#}$ be its Laplace exponent. Let $\widetilde{R}$ be a subordinator with Laplace exponent $\frac{\phi(\cdot)}{\cdot + \xi}$ independent of $\hat{L}$. Then $\widetilde{R}$ is a subordinator. For $t \geq 0$, define the right inverse of $\widetilde{R}$ by*

$$\widetilde{R}_t^{-1} := \inf\{s : \widetilde{R}_s > t\}$$

and set

$$\hat{I}_t = \inf_{s \le t} \hat{L}_s.$$

Then, for L^* defined by

$$L^*_t = \hat{L}_t + \widetilde{R}\left(\widetilde{R}^{-1}(-\hat{I}_t)\right),$$

$L - \underline{\underline{L}}$ under $\mathbb{P}^{\uparrow}$ has the same law as L^*. Moreover, the local time of L^* at 0 may be defined by

$$\ell^*_t = \widetilde{R}^{-1}(-\hat{I}_t).$$

Finally, suppose that $\hat{H}$ is a continuous modification of the height process corresponding to $\hat{L}$. Then

$$H^*_t := \ell^*_t + \hat{H}_t$$

is a continuous modification of the height process corresponding to L^*_t.

Combining the above proposition with the characterization of the pre-infimum process in Proposition 2, we obtain the following result.

Proposition 3. *Let $\hat{L}$ be a process which under $\mathbb{P}$ has the law of L under $\mathbb{P}^{\#}$, satisfying Condition (C3), so that its height process is well-defined and has a continuous modification $\hat{H}$. Define $\hat{I}_t = \inf\{\hat{L}_s : s \le t\}$. Let E be an exponential random variable with rate ξ. Let $\widetilde{D}$ be distributed as in the statement of Theorem 6, independent of $\hat{L}$ and E, and set*

$$R_t = \widetilde{R}_t + E$$

and

$$R^{-1}_t := \inf\{s : R_s > t\}.$$

Then the height process of L is well-defined and has a continuous modification H. Moreover, H is also the height process of $L - \underline{\underline{L}}$, and

$$(L_t - \underline{\underline{L}}_t, H_t, t \ge 0) \overset{d}{=} (\hat{L}_t + R(R^{-1}(-\hat{I}_t)), \hat{H}_t + R^{-1}(-\hat{I}_t), t \ge 0).$$

Proof. The existence of $\hat{H}$ follows from Proposition 2. The construction of $L - \underline{\underline{L}}$ follows from Proposition 2 for the pre-infimum process, and from [45, Lemma 4.1] and Theorem 6 for the post-infimum process.

We claim that $(\hat{H}_t + R^{-1}(-\hat{I}_t), t \ge 0)$ is the height process corresponding to the process

$$(\hat{L}_t + R(R^{-1}(-\hat{I}_t)), t \ge 0).$$

Firstly, note that for $\rho^* = \sup\{t : \hat{I}_t = -E\}$,

$$(\hat{L}_t + R(R^{-1}(-\hat{I}_t)), \hat{H}_t + R^{-1}(-\hat{I}_t), t \in [0, \rho^*]) = (\hat{L}_t + E, \hat{H}_t, t \in [0, \rho^*]).$$

The height process is, by definition, not affected by translating the Łukasiewicz path by a constant, so on $[0, \rho^*]$, the claim follows. On $[\rho^*, \infty)$, the claim follows from Theorem 6.

Finally, we claim that the height processes of L and $L - \underline{\underline{L}}$ agree. Set $I_\infty = \inf\{L_t, t \ge 0\}$ and $\rho = \sup\{t > 0 : L_t = I_\infty\}$, and observe that

$$(L_t - \underline{\underline{L}}_t, t \in [0, \rho]) = (L_t - I_\infty, t \in [0, \rho]).$$

Again because the height process is not affected by adding a constant to the Łukasiewicz path, the claim follows on $[0, \rho]$. On $[\rho, \infty)$, the claim follows from Lemma 8ii in [41]. □

3. Joint convergence of the height process and Łukasiewicz path

In this section, we will show the convergence of the discrete Łukasiewicz path S^n and height process H^n to their continuous counterparts L and H under rescaling. The convergence result relies on the construction of the discrete and continuous processes introduced in (6) and Proposition 3 respectively.

We will start by proving the joint convergence of $\hat{S}^n$ and its height process $\hat{H}^n$ under rescaling. Recall that by Condition (C1) there exists a nondecreasing sequence of positive integers $(\gamma_n, n \geq 1)$ that converges to ∞ such that

$$\left(n^{-1} S^n_{\lfloor n\gamma_n t\rfloor}, t \geq 0\right) \xrightarrow{d} (L_t, t \geq 0)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$.

Theorem 7. *We have that*

$$\left(n^{-1} \hat{S}^n_{\lfloor n\gamma_n t\rfloor}, \gamma_n^{-1} \hat{H}^n_{\lfloor n\gamma_n t\rfloor}, t \geq 0\right) \xrightarrow{d} \left(\hat{L}_t, \hat{H}_t, t \geq 0\right)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})^2$ *as* $n \to \infty$.

Proof. We will first show convergence under rescaling of the Łukasiewicz path, i.e.

$$\left(n^{-1} \hat{S}^n_{\lfloor n\gamma_n t\rfloor}, t \geq 0\right) \xrightarrow{d} \left(\hat{L}_t, t \geq 0\right)$$

in the Skorokhod topology as $n \to \infty$, after which we will use Theorem 2.3.1 of Duquesne and Le Gall [24] to show the joint convergence with the height process. Since $\hat{S}^n$ is a downward skip-free random walk on the integers, it will be sufficient to show that for all $\theta > 0$,

$$\mathbb{E}\left[\exp\left(-\theta n^{-1} \hat{S}^n_{\lfloor n\gamma_n t\rfloor}\right)\right] \to \mathbb{E}\left[\exp(-\theta \hat{L}_t)\right] = \exp(t\phi(\theta + \xi)). \tag{8}$$

(See e.g. [14, Lemma A3].) Note that by definition,

$$\begin{aligned}
\mathbb{E}\left[\exp\left(-\theta n^{-1} \hat{S}^n_{\lfloor n\gamma_n t\rfloor}\right)\right] &= \mathbb{E}\left[\exp\left(-\xi_n S^n_{\lfloor n\gamma_n t\rfloor}\right) \exp\left(-\theta n^{-1} S^n_{\lfloor n\gamma_n t\rfloor}\right)\right] \\
&= \mathbb{E}\left[\exp\left(-\left(n\xi_n + \theta\right) n^{-1} S^n_{\lfloor n\gamma_n t\rfloor}\right)\right] \\
&= \exp\left(\lfloor n\gamma_n t\rfloor \phi_n\left(n^{-1}\left(n\xi_n + \theta\right)\right)\right).
\end{aligned}$$

For sake of brevity, we will denote $n\gamma_n \phi_n(n^{-1}\,\cdot)$ by $\bar{\phi}_n(\,\cdot\,)$. By the convergence under rescaling of S^n to L, we know that $\bar{\phi}_n$ converges to ϕ pointwise. We will first show that

$$n\xi_n \to \xi \tag{9}$$

as $n \to \infty$; then we will show that there is a $b < \xi$ such that for any $B > b$,

$$\bar{\phi}_n(\,\cdot\,) \to \phi(\,\cdot\,) \tag{10}$$

uniformly on $[b,B]$. Together, (9) and (10) imply (8).

Note that by definition, $n\xi_n$ is the unique non-trivial zero of $\bar{\phi}_n$. By convexity of ϕ and by $\phi(0) = \phi(\xi) = 0$, we see that for all $\epsilon > 0$,

$$\phi(\xi - \epsilon) < 0 < \phi(\xi + \epsilon).$$

So by the pointwise convergence of $\bar{\phi}_n$ to ϕ, we have that, for all n large enough,

$$\bar{\phi}_n(\xi - \epsilon) < 0 < \bar{\phi}_n(\xi + \epsilon).$$

But then $n\xi_n \in (\xi - \epsilon, \xi + \epsilon)$, which implies (9). To prove (10), given the pointwise convergence of $\bar{\phi}_n$ to ϕ, it will be sufficient to find a $y < \xi$ such that, for all n large enough, $\bar{\phi}_n$ is monotone on $[y, \infty)$. By convexity of $\bar{\phi}_n$ for all n, it is enough to show that there is an $x < y < \xi$ such that

$$\bar{\phi}_n(x) < \bar{\phi}_n(y) < 0 \tag{11}$$

for all n large enough. However, by convexity of ϕ and by $\phi(0) = \phi(\xi)$, there exist $x < y < 0$ such that

$$\phi(x) < \phi(y) < 0.$$

The pointwise convergence of $\bar{\phi}_n$ to ϕ then implies (11), and hence (10) and (8).

We now wish to apply [24, Theorem 2.3.1] to obtain joint convergence under rescaling of the Łukasiewicz path and height process. Considering Theorem 3.3, Condition (C3), and (8), the only condition that is left to check is that, for Y^n_m the number of individuals in generation m in the Galton–Watson branching process given by the first n trees in the forest encoded by S^n, we have for all $\delta > 0$

$$\liminf_{n\to\infty} \mathbb{P}^{\#}\left[Y^n_{\lfloor \delta\gamma_n \rfloor} = 0\right] > 0.$$

We claim that this is equivalent to the statement above under $\mathbb{P}$. Indeed,

$$\mathbb{P}\left[Y^n_{\lfloor \delta\gamma_n \rfloor} = 0\right] = \mathbb{P}\left[\text{the first } n \text{ trees are finite}\right] \times \mathbb{P}^{\#}\left[Y^n_{\lfloor \delta\gamma_n \rfloor} = 0\right],$$

so the equivalence follows from the fact that

$$\mathbb{P}\left[\text{the first } n \text{ trees are finite}\right] = \exp(-n\xi_n) \to \exp(-\xi).$$

The statement then follows from

$$\liminf_{n\to\infty} \mathbb{P}\left[Y^n_{\lfloor \delta\gamma_n \rfloor} = 0\right] > 0,$$

which is the assumption (3). □

We will now show joint convergence under rescaling of Q^n and R^n, which is the content of the next theorem. Suppose, without loss of generality, that the Laplace exponent of L is given by

$$\phi(\theta) = a\theta + \frac{1}{2}b\theta^2 + \int_0^\infty \nu(dx)(\exp(-\theta x) - 1 + \theta(x \wedge 1))$$

for $a > 0$, $b \geq 0$, and ν a measure on $[0, \infty)$ that satisfies $\int_0^\infty (1 \wedge x)\nu(x) < \infty$, so that (a, b, ν) is the characteristic triple of L.

Define

$$\tilde{\nu}(dx, dy) = \exp(-\xi x)\nu(dy)dx\mathbb{1}_{x<y}$$

and let $(R_t, Q_t)_{t\geq 0}$ be a Lévy process with Lévy measure $\tilde{\nu}$, drift vector $(b,2b)$, and no Gaussian component, so that $(R_t)_{t\geq 0}$ and $(Q_t)_{t\geq 0}$ are both subordinators. Let E be an exponential random variable with rate ξ.

Theorem 8. *It holds that*

$$\left(n^{-1}R^n_{\lfloor \gamma_n t\rfloor}, n^{-1}Q^n_{\lfloor \gamma_n t\rfloor}, t\geq 0\right)\xrightarrow{d}(E+R_t, E+Q_t, t\geq 0)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^2_+)$ *as* $n\to\infty$.

Proof. First we recall that $R^n_0 = Q^n_0 \sim \text{Geom}\,(\exp(-\xi_n))$. Recalling that $n\xi_n\to\xi$, we get that for any $x>0$,

$$\begin{aligned}\mathbb{E}\left[\exp(-xn^{-1}R^n_0)\right] &= \frac{1-\exp(-\xi_n)}{1-\exp(-\xi_n - xn^{-1})}\\ &\to \frac{\xi}{\xi+x},\end{aligned}$$

so we can conclude that

$$(n^{-1}R^n_0, n^{-1}Q^n_0)\xrightarrow{d}(E,E). \tag{12}$$

We aim to use [33, Theorem VII.2.9] to prove the convergence under rescaling of

$$\left(\sum_{i=1}^k B^{(n)}_i, \sum_{i=1}^k (N^{(n)}_i - 1), k\geq 1\right).$$

We will approximate this random walk by a Poissonized version. To that end, define

$$\nu_n(dx) = n\gamma_n\sum_{i=0}^{\infty}\delta_{(i-1)/n}(dx)\mathbb{P}(D=i).$$

Let $(\bar{S}^n_t, t\geq 0)$ denote the compound Poisson process defined by

$$\mathbb{E}\left[\exp(iu\bar{S}^n_t)\right] = \exp\left(t\int_{-1/n}^{\infty}\nu_n(dx)(e^{iux}-1)\right).$$

We claim that

$$\bar{S}^n_1\xrightarrow{d}L_1. \tag{13}$$

Indeed, $\bar{S}^n_1$ is a Poissonized version of $n^{-1}S^n_{\lfloor n\gamma_n\rfloor}$, so (13) is implied by (2). Now, let $h:\mathbb{R}\to\mathbb{R}$ be a bounded function such that $h(x)=-x$ on a neighborhood of 0. Then Theorem VII.2.9 in [33] implies the following facts:

1. We have $\int_{-1/n}^{\infty}\nu_n(dx)h(x)\to a_h$ for some a_h as $n\to\infty$.
2. As $n\to\infty$,
$$\int_{-1/n}^{\infty}\nu_n(dx)(h(x))^2\to b+\int_0^{\infty}\nu(dx)(h(x))^2. \tag{14}$$
3. For bounded continuous $f:\mathbb{R}^+\to\mathbb{R}$ such that $f(x)=o(x^2)$ when $x\downarrow 0$, it holds that
$$\int_{-1/n}^{\infty}\nu_n(dx)f(x)\to\int_0^{\infty}\nu(dx)f(x). \tag{15}$$

Recalling that

$$\mathbb{P}(B^{(n)} = k, N^{(n)} = l) = \exp(-k\xi_n)\mathbb{P}(D = l)\mathbb{1}_{k<l},$$

we now define

$$\tilde{\nu}_n(dx, dy) = \gamma_n \sum_{j=0}^{\infty} \sum_{i=0}^{j-1} \delta_{i/n, j/n}(dx, dy) \exp(-\xi_n i)\mathbb{P}(D = j).$$

Then, by an argument similar to the one used before, for $(\bar{R}_t^n, \bar{Q}_t^n, t \geq 0)$ denoting the process with stationary increments defined by

$$\mathbb{E}\left[\exp(iu_1\bar{R}_t^n + iu_2\bar{Q}_t^n)\right] = \exp\left(t \int_0^{\infty} \int_0^{\infty} \tilde{\nu}_n(dx, dy)(e^{iu_1 x + iu_2 y} - 1)\right),$$

we get that for all t, $(\bar{R}_t^n, \bar{Q}_t^n)$ is a Poissonized version of

$$\left(n^{-1} \sum_{i=1}^{\lfloor \gamma_n t \rfloor} B_i^n, n^{-1} \sum_{i=1}^{\lfloor \gamma_n t \rfloor} N_i^n\right).$$

We will show that for all t,

$$(\bar{R}_t^n, \bar{Q}_t^n) \xrightarrow{d} (R_t, Q_t), \tag{16}$$

which implies the functional convergence by a result in Kallenberg's book [39, Theorem 16.14]. By Theorem VII.2.9 in [33], using the truncation function $H(x, y) = (x \wedge 1, y \wedge 1)$, the following properties imply (16):

1. We have
   (a) $\int_0^{\infty} \int_0^{\infty} \tilde{\nu}_n(dx, dy)(x \wedge 1) \to \int_0^{\infty} \int_0^{\infty} \tilde{\nu}(dx, dy)(x \wedge 1) + b$, and
   (b) $\int_0^{\infty} \int_0^{\infty} \tilde{\nu}_n(dx, dy)(y \wedge 1) \to \int_0^{\infty} \int_0^{\infty} \tilde{\nu}(dx, dy)(y \wedge 1) + 2b$.
2. We have
   (a) $\int_0^{\infty} \int_0^{\infty} \tilde{\nu}_n(dx, dy)(x \wedge 1)^2 \to \int_0^{\infty} \int_0^{\infty} \tilde{\nu}(dx, dy)(x \wedge 1)^2$,
   (b) $\int_0^{\infty} \int_0^{\infty} \tilde{\nu}_n(dx, dy)(y \wedge 1)^2 \to \int_0^{\infty} \int_0^{\infty} \tilde{\nu}(dx, dy)(y \wedge 1)^2$, and
   (c) $\int_0^{\infty} \int_0^{\infty} \tilde{\nu}_n(dx, dy)(x \wedge 1)(y \wedge 1) \to \int_0^{\infty} \int_0^{\infty} \tilde{\nu}(dx, dy)(y \wedge 1)^2$.
3. For all continuous, bounded $F : \mathbb{R}^2 \to \mathbb{R}$ that are 0 on a neighborhood of 0,
$$\int_0^{\infty} \int_0^{\infty} \tilde{\nu}_n(dx, dy)F(x, y) \to \int_0^{\infty} \int_0^{\infty} \tilde{\nu}(dx, dy)F(x, y).$$

We will prove the conditions one by one, starting with 1a. We note that

$$\begin{aligned}\int_0^{\infty} \int_0^{\infty} \tilde{\nu}_n(dx, dy)(x \wedge 1) &= \gamma_n \sum_{i=0}^{\infty} \sum_{j=0}^{i-1} \mathbb{P}(D = i) \exp(-\xi_n j)(j/n \wedge 1) \\ &= \frac{1}{n} \int_{-1/n}^{\infty} \nu_n(dx) \sum_{j=0}^{\lfloor nx \rfloor} \exp(-\xi_n j)(j/n \wedge 1).\end{aligned}$$

We will first argue that

$$\frac{1}{n(x\wedge 1)}\sum_{j=0}^{\lfloor nx\rfloor}\exp(-\xi_n j)(j/n\wedge 1)\to\frac{1}{1\wedge x}\int_0^x dy\exp(-\xi y)(y\wedge 1) \tag{17}$$

uniformly in all $x>0$ as $n\to\infty$. Recall that $n\xi_n=\xi+o(1)$, which, together with the Riemann integrability of $\exp(-\xi y)(y\wedge 1)$ on compact intervals, implies pointwise convergence in (17). The facts that $\int_0^\infty dy\exp(-\xi y)(y\wedge 1)<\infty$ and that $\int_0^x dy\exp(-\xi y)(y\wedge 1)=O(x^2)$ as $x\to 0$, together with the uniform continuity of $f(x,y):=\frac{1}{x\wedge 1}\exp(-\xi y)(y\wedge 1)$ on $(\delta,\infty)\times(0,R)$ for any $\delta>0$ and $R>0$, then imply that convergence in (17) is in fact uniform, as claimed.

Since $\int_{-1/n}^\infty \nu_n(dx)(x\wedge 1)$ is convergent as $n\to\infty$, and in particular bounded, we have that

$$\left|\int_{-1/n}^\infty \nu_n(dx)\frac{1}{n}\sum_{j=0}^{\lfloor nx\rfloor}\exp(-\xi_n j)(j/n\wedge 1)dy-\int_{-1/n}^\infty \nu_n(dx)\int_0^x dy\exp(-\xi y)(y\wedge 1)\right| \tag{18}$$

converges to 0 as $n\to\infty$. So in order to prove 1a, it is sufficient to show that

$$\int_{-1/n}^\infty \nu_n(dx)\int_0^x dy\exp(-\xi y)(y\wedge 1)\to\int_0^\infty\int_0^\infty \tilde{\nu}(dx,dy)(x\wedge 1)+b.$$

To see this, consider a truncation function h, and define

$$g_u(x):=e^{iux}-1+iuh(x)+u^2\int_0^x dy\exp(-\xi y)(y\wedge 1).$$

Then, note that $\int_0^x dy\exp(-\xi y)(y\wedge 1)=\frac{x^2}{2}+o(x^2)$ as $x\to 0$, so $g_u(x)=o(x^2)$ as $x\to 0$. Furthermore, g_u is bounded, because $\int_0^x dy\exp(-\xi y)(y\wedge 1)$ is bounded. Then, by (15),

$$\int_{-1/n}^\infty \nu_n(dx)g_u(x)\to\int_0^\infty \nu(dx)g_u(x). \tag{19}$$

Moreover, since

$$\begin{aligned}&\mathbb{E}\left[\exp(iu\bar{S}^n)\right]=\exp\left(-iu\int_{-1/n}^\infty \nu_n(dx)h(x)+\int_{-1/n}^\infty \nu_n(dx)(e^{iux}-1+uih(x))\right)\\&\to\mathbb{E}\left[\exp(iuL_1)\right]=\exp\left(-iua_h-bu^2+\int_0^\infty \nu(dx)(e^{iux}-1+uih(x))\right)\end{aligned}$$

and $\int_{-1/n}^\infty \nu_n(dx)h(x)\to a_h$, we find that

$$\int_{-1/n}^\infty \nu_n(dx)(e^{iux}-1+uih(x))\to -bu^2+\int_0^\infty \nu(dx)(e^{iux}-1+uih(x)),$$

which, combined with (19), implies that

$$\int_{-1/n}^\infty \nu_n(dx)\int_0^x dy\exp(-\xi y)(y\wedge 1)\to\int_0^\infty\int_0^\infty \tilde{\nu}(dx,dy)(x\wedge 1)+b,$$

proving 1a.

A similar proof shows that also 1b holds.

To prove 2a, note that

$$\int_0^\infty \int_0^\infty \tilde{\nu}_n(dx, dy)(x \wedge 1)^2 = \gamma_n \sum_{i=0}^{\infty} \sum_{j=0}^{i-1} \mathbb{P}(D^n = i) \exp(-\xi_n j)(j/n \wedge 1)^2$$
$$= \frac{1}{n} \int_{-1/n}^{\infty} \nu_n(dx) \sum_{j=0}^{\lfloor nx \rfloor} \exp(-\xi_n j)(j/n \wedge 1)^2 dy.$$

The first fact we need is that

$$\left| \frac{1}{n} \int_{-1/n}^{\infty} \nu_n(dx) \sum_{j=0}^{\lfloor nx \rfloor} \exp(-\xi_n j)(j/n \wedge 1)^2 dy - \int_{-1/n}^{\infty} \nu_n(dx) \int_0^x dy \exp(-\xi y)(y \wedge 1)^2 \right|$$

converges to 0 as $n \to \infty$, which is proved in a similar manner to (18). Then, since $\int_0^x dy \exp(-\xi y)(y \wedge 1)^2 = o(x^2)$ as $x \to 0$, and it is a bounded function of x, we can use (15) to obtain 2a. The proofs of 2b, 2c, and 3 are similar.

By Theorem VII.2.9 in [33], for all $t > 0$,

$$(\bar{R}_t^n, \bar{Q}_t^n) \xrightarrow{d} (R_t, Q_t),$$

which proves the statement. □

The final steps in the proof of Theorem 2 are similar to the proof of Theorem 1.5 in [22].

Recall that $F^n(k) = \inf\{l : -\hat{I}^n(k-1-l) < R^n(l)\}$. We also define its continuous counterpart

$$(F_t, t \geq 0) = \left(\inf\{s : -\hat{I}_t < R_s\}, t \geq 0\right).$$

We will now use the convergence of (R^n, Q^n) under rescaling to prove the convergence under rescaling of F^n to F. For this we need the following technical lemma, whose proof can be found in the appendix.

Lemma 4. *Suppose $f_n \to f$ in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$, with f_n increasing for all n, and f increasing and continuous. Furthermore, suppose $g_n \to g$ in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$, with g_n increasing for all n, g strictly increasing, and $g_n(t) \to \infty$ and $g(t) \to \infty$ as $t \to \infty$. Then*

$$(\inf\{s \geq 0 : f(t) < g(s)\}, t \geq 0)$$

is continuous, and

$$(\inf\{s \geq 0 : f_n(t) < g_n(s)\}, t \geq 0) \to (\inf\{s \geq 0 : f(t) < g(s)\}, t \geq 0)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$.

Moreover, for $(\epsilon_n)_{n \geq 0}$ such that $\epsilon_n \downarrow 0$,

$$(\inf\{s \geq 0 : f_n(t - \epsilon_n s) < g_n(s)\}, t \geq 0) \to (\inf\{s \geq 0 : f(t) < g(s)\}, t \geq 0)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$.

Lemma 5. *We have that*

$$\left(\gamma_n^{-1} F^n(\lfloor n\gamma_n t\rfloor), t \geq 0\right) \xrightarrow{d} (F_t, t \geq 0)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ *as* $n \to \infty$*, jointly with the joint convergence of* R^n *and* Q^n *under rescaling (Theorem 8) and the joint convergence of* $\hat{S}^n$ *and* $\hat{H}^n$ *under rescaling (Theorem 7), where* (R^n, Q^n) *is independent of* $(\hat{S}^n, \hat{H}^n)$*. Moreover,* $(F_t, t \geq 0)$ *is continuous almost surely.*

Proof. We need to show that

$$\left(\gamma_n^{-1} \inf\{l : -\hat{I}^n(\lfloor n\gamma_n t\rfloor - 1 - l) < R^n_l\}, t \geq 0\right) \xrightarrow{d} \left(\inf\{s : -\hat{I}_t < R_s\}, t \geq 0\right)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$ and that $\left(\inf\{s : -\hat{I}_t < R_s\}, t \geq 0\right)$ is continuous, for which we will use Lemma 4. Firstly, recall that L is of infinite variation, so that $b > 0$, or $\nu((0, \infty)) = \infty$. In the first case, it is obvious that R is strictly increasing. In the second case, note that since R has Lévy measure

$$\int_{y\in[0,\infty)} \tilde{\nu}(dx, dy) = \exp(-\xi x)\nu((x, \infty))dx,$$

the intensity of jumps of size ϵ goes to ∞ as ϵ goes to 0, which implies that the jumps of R are dense, and its sample paths are strictly increasing with probability 1. By Skorokhod's representation theorem, we may work on a probability space where the joint convergence of R^n and Q^n under rescaling (Theorem 8) and the joint convergence of $\hat{S}^n$ (and $\hat{I}^n$) and $\hat{H}^n$ under rescaling (Theorem 7) all hold almost surely. Then, by Lemma 4, since $\hat{I}$ is continuous ($\hat{L}$ is spectrally positive), we have that

$$\left(\gamma_n^{-1} \inf\{l : -\hat{I}^n(\lfloor \gamma_n nt\rfloor - 1 - l) < R^n_l\}, t \geq 0\right) \xrightarrow{\text{a.s.}} \left(\inf\{s : -\hat{I}_t < R_s\}, t \geq 0\right)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$ and F_t is continuous almost surely. The result follows. □

From Theorem 7, Theorem 8, and Lemma 5 we know that, as $n \to \infty$,

$$\begin{gathered}\left(n^{-1}\hat{S}^n(\lfloor n\gamma_n t\rfloor), \gamma_n^{-1}\hat{H}^n(\lfloor n\gamma_n t\rfloor), t \geq 0\right) \xrightarrow{d} \left(\hat{L}_t, \hat{H}_t, t \geq 0\right),\\ \left(n^{-1}R^n(\lfloor \gamma_n t\rfloor), n^{-1}Q^n(\lfloor \gamma_n t\rfloor), t \geq 0\right) \xrightarrow{d} (R_t, Q_t, t \geq 0)\text{, and}\\ \left(\gamma_n^{-1}F^n(\lfloor n\gamma_n t\rfloor), t \geq 0\right) \xrightarrow{d} (F_t, t \geq 0),\end{gathered} \tag{20}$$

jointly, in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})^2$, $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^2)$, and $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ respectively. We would like to show the convergence under rescaling of $Q^n(F^n(k))$ and of $R^n(F^n(k))$ jointly with the convergence in (20), since these quantities appear in the pathwise construction of S^n and $S^n - \underline{S}^n$ in Equation (6). For this we need a technical lemma, which follows immediately from the characterization of convergence in the Skorokhod topology given in the book of Ethier and Kurtz [29, Proposition 3.6.5].

Lemma 6. *Let* E *be a metric space, and suppose* $h_n \to h$ *in* $\mathbb{D}(\mathbb{R}_+, E)$ *and* $f_n \to f$ *in* $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$ *as* $n \to \infty$*. If* f_n *and* f *are non-decreasing and* f *is continuous, then*

$$h_n \circ f_n \to h \circ f$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^2)$ *as* $n \to \infty$*.*

Lemma 7. *We have*

$$\left(n^{-1}D^n\left(F^n\left(\lfloor n\gamma_n t\rfloor\right)\right), n^{-1}Q^n\left(F^n\left(\left(\lfloor n\gamma_n t\rfloor\right)\right)\right), t\geq 0\right)\xrightarrow{d}\left(D_{F_t}, Q_{F_t}, t\geq 0\right)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^2)$ *as* $n\to\infty$*, jointly with the convergence in* (20).

Proof. By Skorokhod's representation theorem we may work on a space where the convergence of (20) holds almost surely. The result then follows from Lemma 6. □

Lemma 8. *We have*

$$\begin{aligned}&\left(n^{-1}\hat{S}^n\left(\lfloor n\gamma_n t\rfloor - F^n\left(\lfloor n\gamma_n t\rfloor\right)\right), \gamma_n^{-1}\hat{H}^n\left(\lfloor n\gamma_n t\rfloor - F^n\left(\lfloor n\gamma_n t\rfloor\right)\right), t\geq 0\right)\\ &\xrightarrow{d}(\hat{L}_t, \hat{H}_t\geq 0)\end{aligned}$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})^2$ *as* $n\to\infty$*, jointly with the convergence in* (20).

Proof. Firstly, note that $k - F^n(k)$ equals the number of steps not spent on the spine up to time k and so is a non-decreasing function of k. Then, note that

$$\lfloor n\gamma_n t\rfloor - F^n\left(\lfloor n\gamma_n t\rfloor\right) = \left\lfloor n\gamma_n\left(t-\gamma_n^{-1}n^{-1}F^n\left(\lfloor n\gamma_n t\rfloor\right)\right)\right\rfloor,$$

and

$$\left(t-\gamma_n^{-1}n^{-1}F^n\left(\lfloor n\gamma_n t\rfloor\right), t\geq 0\right)\to(t, t\geq 0)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ almost surely as $n\to\infty$. We may use Skorokhod's representation theorem to work on a space where the convergence in (20) holds almost surely, and then Lemma 6 gives the result. □

We will now prove Theorem 2.

Proof of Theorem 2. Let $(\hat{L}, \hat{H})$, $\hat{I}$, (R,Q), and $F_t = \inf\{s : -\hat{I}_t < R_s\}$ be as defined in Section 2.2. Then, for $\underline{\underline{S}}^n$ the future infimum of S^n, we know by the pathwise construction of S^n, $S^n - \underline{\underline{S}}^n$, and H^n given in (6), and by Lemmas 8, 7, and 5, that

$$\begin{aligned}&\left(n^{-1}S^n\left(\lfloor tn\gamma_n\rfloor\right), n^{-1}(S^n-\underline{\underline{S}}^n)\left(\lfloor tn\gamma_n\rfloor\right), \gamma_n^{-1}H^n\left(\lfloor tn\gamma_n\rfloor\right), t\geq 0\right)\\ &\xrightarrow{d}\left(\hat{L}_t+Q_{F_t}, \hat{L}_t+R_{F_t}, \hat{H}_t+F_t, t\geq 0\right)\end{aligned}\tag{21}$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^3)$ as $n\to\infty$. By assumption, we have that

$$\left(n^{-1}S^n\left(\lfloor n\gamma_n t\rfloor\right), t\geq 0\right)\xrightarrow{d}(L_t, t\geq 0),$$

so

$$L_t\overset{d}{=}\hat{L}_t+Q_{F_t},$$

and by construction, $\hat{L}_t+R_{F_t}$ equals $\hat{L}_t+Q_{F_t}$ minus its future infimum. Then, by Proposition 3, we know that $\hat{H}_t+F_t$ is the height process corresponding to $\hat{L}_t+R_{F_t}$ and hence to $\hat{L}_t+Q_{F_t}$. The result follows. □

4. Application to the configuration model in the critical window

This section contains new results on the scaling limit of the configuration model with i.i.d. power-law degrees in the critical window. We use Theorem 2 to extend the methods in Conchon-Kerjan and Goldschmidt [17] from the critical point to the critical window.

The configuration model is a method of constructing a multigraph with a given degree sequence that was introduced by Bollobás in [13]:

> Consider n vertices labeled by $[n]$ and a sequence $\mathbf{d} = (d_i)_{i\in[n]} \in \mathbb{N}^n$ such that $\sum_{i\in[n]} d_i$ is even. We will sample a multigraph such that the degree of vertex i is equal to d_i for every $i \in [n]$. The configuration model on n vertices having degree sequence $\mathbf{d}$ is constructed as follows. Equip vertex j with d_j half-edges. Two half-edges create an edge once they are paired. Pick any half-edge and pair it with a uniformly chosen half-edge from the remaining unpaired half-edges and keep repeating the above procedure until all half-edges are paired.

Note that the graph constructed by the above procedure may contain self-loops or multiple edges. It can be shown that, conditionally on the constructed multigraph being simple, the law of such graphs is uniform over all possible simple graphs with degree sequence $\mathbf{d}$. Furthermore, as shown in [35], under very general assumptions, the asymptotic probability of the graph being simple is positive. For a discussion of the configuration model and standard results, see [51, Chapter 7].

4.1. Model and result

We use the configuration model to construct a uniform graph with a random degree sequence. The model we consider is as follows.

Fix $\lambda \in \mathbb{R}$. Most quantities that will be defined depend on λ. To avoid overcomplicating the notation, this will not be made explicit unless necessary to avoid confusion. For each $n \in \mathbb{N}$, let $D_1^n, D_2^n, \ldots, D_n^n \geq 1$ be an i.i.d. degree sequence satisfying the following properties (which are labeled by 'CM' for 'configuration model'):

(CM1) For $\mu_n := \mathbb{E}[D_1^n]$, we have $\mu_n \to \mu$ as $n \to \infty$, with μ not depending on λ.

(CM2) We have $\mathbb{E}\left[\left(D_1^n\right)^2\right] = \left(2 + \lambda n^{-\frac{\alpha-1}{\alpha+1}}\right)\mathbb{E}[D_1^n]$ for some $\alpha \in (1, 2)$.

(CM3) We have $\mathbb{P}(D_1^n = k) \sim c_n k^{-(\alpha+2)}$ as $k \to \infty$, with $c_n > 0$ for all n and $c_n \to c$ as $n \to \infty$.

(CM4) For Z^n a random variable such that $\mathbb{P}(Z^n = k) = k\mathbb{P}(D_1^n = k)/\mu_n$, and $(S^n(k), k \geq 0)$ a random walk with steps distributed as $Z^n - 2$, let Y_m^n be the number of vertices at height m in the first $\lfloor n^{1/(\alpha+1)} \rfloor$ trees of the forest encoded by $(S^n(k), k \geq 0)$. Then, for every $\delta > 0$,

$$\liminf_{n\to\infty} \mathbb{P}\left[Y^n_{\lfloor \delta n^{\frac{\alpha-1}{\alpha+1}} \rfloor} = 0\right] > 0.$$

Let $G_1^n, G_2^n, \ldots$ be the components of a uniformly random graph with i.i.d. degrees that are distributed as D_1^n, with the components listed in decreasing order of size. View each G_i^n as a compact measured metric spaces by equipping it with the graph distance d_i^n, and the counting measure on its vertices, μ_i^n. More generally, a compact measured metric space will be denoted by a triple (G, d, μ), for (G, d) a compact metric space and μ a finite Borel measure on (G, d). Formally, each G_i^n is an element of the Polish space of isometry-equivalence classes

of measured metric spaces, endowed with the Gromov–Hausdorff–Prokhorov distance. For a discussion of the topology, we refer the reader to [3, Section 2]. We will prove the following theorem.

Theorem 9. *There exists a sequence of random compact measured metric spaces*

$$((\mathcal{G}_1, d_1, \mu_1), (\mathcal{G}_2, d_2, \mu_2), \dots)$$

such that, as $n \to \infty$,

$$\left(\left(G_i^n, n^{-\frac{\alpha-1}{\alpha+1}} d_i^n, n^{-\alpha/(\alpha-1)} \mu_i^n\right), i \geq 1\right) \xrightarrow{d} ((\mathcal{G}_i, d_i, \mu_i), i \geq 1)$$

in the sense of the product Gromov–Hausdorff–Prokhorov topology.

In the case where $\lambda = 0$, and the degree distribution does not depend on n, this result was already known from [17, Theorem 1.1]; this is known as the critical case. Intuitively, criticality entails that for large n, and for (V_n, W_n) an edge chosen uniformly at random from the graph, the expected degree of V_n is roughly 2. Our contribution is then to prove the theorem for all $\lambda \in \mathbb{R}$ and for degree distributions depending on n. This is known as the *critical window*, in which, for large n, and for (V_n, W_n) an edge chosen uniformly at random from the graph, the expected degree of V_n is roughly $2 + \lambda n^{-\frac{\alpha-1}{\alpha+1}}$.

4.2. Related work

Most results on the configuration model have been obtained for models with a deterministic degree sequence. The phase transition for the undirected setting was shown in [37, 46, 47]. The limiting law of the rescaled component sizes at criticality and in the critical window were obtained by Riordan [49] under the assumption that the degrees are bounded. Dhara, van der Hofstad, van Leeuwaarden, and Sen showed convergence under rescaling of the component sizes and surpluses in the critical window in the finite-third-moment setting [19] and in the heavy-tailed regime [20]. Bhamidi, Dhara, van der Hofstad, and Sen obtained metric space convergence in the critical window in [11], a result that the authors later improved to a stronger topology in [12]; both of these results also hold conditional on the constructed multigraph being simple. We will discuss the results of Bhamidi, Dhara, van der Hofstad, and Sen further at the end of this section.

Configuration models with a *random* degree sequence are considered in [17, 38]. Joseph [38] showed convergence of the component sizes and surpluses of the large components under rescaling at criticality, both for degree distributions with finite third moments and for the heavy-tailed regime. Conchon-Kerjan and Goldschmidt [17] showed product Gromov–Hausdorff–Prokhorov convergence of the large components at criticality in these two regimes, and showed that the results also hold conditioned on the resulting graph being simple.

In recent work [21], the author and Xie showed metric space convergence under rescaling of the strongly connected components of the directed configuration model at criticality.

4.2.1. *Results in [11, 12].* We now describe the model that is considered in [11, 12] to provide a comparison with our result. Our description of the results in [11] closely follows the presentation in [17].

Let $\mathfrak{d}_1^n \geq \cdots \geq \mathfrak{d}_n^n$ be a family of deterministic degree sequences, such that $\sum_{i=1}^n \mathfrak{d}_i^n$ is even, and if $\mathfrak{D}_n$ denotes the degree of a vertex chosen uniformly at random, the following conditions hold (they are labeled by 'DD' for 'deterministic degrees'):

(DD1) We have $n^{-1/(\alpha+1)}\mathfrak{d}_i^n \to \vartheta_i$ as $n \to \infty$ for each $i \geq 1$, where $\vartheta_1 \geq \vartheta_2 \geq \cdots \geq 0$ is such that $\sum_{i\geq 1} \vartheta^3 < \infty$, but $\sum_{i\geq 1} \vartheta_i^2 = \infty$.

(DD2) We have $\mathfrak{D}_n \xrightarrow{d} \mathfrak{D}$, along with the convergence of its first two moments, for some random variable $\mathfrak{D}$ with $\mathbb{P}(\mathfrak{D} = 1) > 0$, $\mathbb{E}[\mathfrak{D}] = \mu$, and $\mathbb{E}[\mathfrak{D}(\mathfrak{D}-1)]/\mathbb{E}[\mathfrak{D}] = \theta > 1$, and

$$\lim_{K\to\infty} \limsup_{n\to\infty} n^{-3/(\alpha+1)} \sum_{i\geq K+1} \left(\mathfrak{d}_i^n\right)^3 = 0.$$

Let $\theta_n = \mathbb{E}\left[\mathfrak{D}_n(\mathfrak{D}_n - 1)\right]/\mathbb{E}\left[\mathfrak{D}_n\right]$. The authors of [11, 12] sample a uniform graph with this degree sequence and perform percolation at parameter

$$p_n(\lambda) = \frac{1}{\theta_n} + \tilde{\lambda} n^{-\frac{\alpha-1}{\alpha+1}},$$

for some $\tilde{\lambda} \in \mathbb{R}$, which yields a graph in the critical window. Call the resulting degree sequence $(\mathfrak{D}_1^{n,\tilde{\lambda}}, \ldots, \mathfrak{D}_n^{n,\tilde{\lambda}})$. In this setting, [11, Theorem 2.2] is the precise analogue of our Theorem 9 in the Gromov-weak topology. In [12], this result is strengthened to convergence in the Gromov–Hausdorff–Prokhorov topology, under the following additional assumptions:

(DD3) For $\mathfrak{D}_n^*$ the degree of a size-biased pick from $\mathfrak{d}_1^n, \ldots, \mathfrak{d}_n^n$, there exists $c_0 > 0$ such that for all $1 \leq l \leq \mathfrak{d}_1^n$ and $n \geq 1$,

$$\mathbb{P}\left(\mathfrak{D}_n^* \geq l\right) \geq \frac{c_0}{l^\alpha}.$$

(DD4) For $\beta_i^n = n^{-2/(\alpha+1)} \sum_{j=1}^{i-1} \left(\mathfrak{d}_j^n\right)^2$, there exists a sequence $(k_n)_{n\geq 1}$ with $k_n \to \infty$ and $k_n = o\left(n^{1/(\alpha+1)}\right)$ such that $\beta_{k_n}^n / \log n \to \infty$ as $n \to \infty$.

(DD5) For all $\epsilon > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} n^{-1/(\alpha+1)} \sum_{i\in(k,kn)} e^{-\epsilon\beta_i^n} \mathfrak{d}_i^n = 0.$$

These extra assumptions allow the authors of [12] to show that the components in their graph model satisfy the *global lower mass-bound property* [12, Theorems 1.2 and 1.3], which allows them to extend the results in [11] to convergence in the Gromov–Hausdorff–Prokhorov topology using results from [7].

The limit object in [11, 12] is constructed by making vertex identifications in *tilted inhomogeneous continuum random trees*. The scaling limit of the depth-first walk that Bhamidi, Dhara, van der Hofstad, and Sen consider is a *thinned Lévy process*, whereas we show convergence to a measure-changed Lévy process. The connection between our results and theirs will become clear in the following section.

4.2.2. *Relationship to percolation.* We will illustrate that the law of a degree sequence that is obtained by bond percolation on the *half-edges* of a sequence of vertices with i.i.d. degrees in the supercritical regime satisfies the conditions of Theorem 9. This is approximately the degree distribution after bond percolation on the *edges* of a uniform random graph with such a supercritical random degree sequence, although we ignore dependence between the degrees of different vertices. Using results of Janson [34], such mild dependence can be shown to have a negligible effect on the properties of the graph, but we omit the straightforward details. We also show how our results are related to the results in [11, 12].

Let D be a random variable in $\mathbb{N}$ that satisfies $\mathbb{E}[D] = \mu$, $\mathbb{E}[D^2] = \rho > 2\mu$ and $\mathbb{P}(D = k) \sim ck^{-(\alpha+2)}$ for $\alpha \in (1, 2)$ as $k \to \infty$. Define $C_\alpha = \frac{c\Gamma(2-\alpha)}{\alpha(\alpha-1)}$, so that the Laplace transform $\mathcal{L}_D$ of D satisfies

$$\mathcal{L}_D(\theta) = \mathbb{E}[\exp(-\theta D)] = 1 - \mu\theta + \frac{\rho}{2}\theta^2 - \frac{C_\alpha}{\alpha+1}\theta^{\alpha+1} + o(\theta^{\alpha+1})$$

for $\theta \to 0$. View D as the degree distribution. Then, keep every half-edge with probability

$$p(\lambda, n) = \frac{1 + \lambda n^{-\frac{\alpha-1}{\alpha+1}}}{\frac{\rho}{\mu} - 1},$$

and call the resulting degree distribution $B^{(\lambda,n)}$.

In the next paragraph, we will show that the conditions of Theorem 9 are satisfied for $B^{(\lambda,n)}$. Then, for $\mathfrak{d}_1^n, \dots, \mathfrak{d}_n^n$ denoting a sample of i.i.d. random variables $\mathfrak{D}_1, \dots, \mathfrak{D}_n$ with the same distribution as D, Conditions (DD1) and (DD2) are satisfied almost surely for some sequence of random variables ϑ_i [20, Section 2.2]. Moreover, the order statistics of the percolated degree sequence with $\tilde{\lambda} = \lambda/(\rho/\mu - 1)$ closely resemble an ordered sample of i.i.d. random variables distributed as $B^{(\lambda,n)}$. Therefore, it should be the case that, in this particular set-up, the limit of Bhamidi, Dhara, van der Hofstad, and Sen corresponds to the limit in Theorem 9.

We will now verify the conditions of Theorem 9 for $B^{(\lambda,n)}$. Note that the Laplace transform $\mathcal{L}_{B^{(\lambda,n)}}$ of $B^{(\lambda,n)}$ satisfies

$$\mathcal{L}_{B^{(\lambda,n)}}(\theta) = \mathbb{E}[\exp(D \log(1 - p(\lambda, n) + p(\lambda, n)e^{-\theta}))], \tag{22}$$

so that

$$\begin{aligned}\mathcal{L}_{B^{(\lambda,n)}}(\theta) =& 1 - p(\lambda, n)\mu\theta + p(\lambda, n)\mu\left(2 + \lambda n^{-\frac{\alpha-1}{\alpha+1}}\right)\frac{\theta^2}{2} - \frac{C_\alpha}{\alpha+1}p(\lambda, n)^{\alpha+1}\theta^{\alpha+1} \\ &+ o(\theta^{\alpha+1})\end{aligned}$$

as $\theta \to 0$. This implies that Conditions (CM1), (CM2), and (CM3) are satisfied with $\mu_n = p(\lambda, n)\mu \to \mu/(\frac{\rho}{\mu} - 1)$ and $c_n = cp(\lambda, n)^{\alpha+1} \to c/(\frac{\rho}{\mu} - 1)^\alpha$ as $n \to \infty$.

We now check Condition (CM4). We will drop the dependency on λ from the notation, unless it is necessary to avoid confusion. Let $\tilde{D}$ be a random variable with the size-biased distribution of D, i.e. for all $k \in \mathbb{N}$,

$$\mathbb{P}(\tilde{D} = k) = \frac{k\mathbb{P}(D = k)}{\mathbb{E}[D]},$$

and similarly, let $\tilde{B}_n$ be a random variable with the size-biased distribution of $B^{(n,\lambda)}$, i.e.

$$\mathbb{P}(\tilde{B}_n = k) = \frac{k\mathbb{P}(B^{(n,\lambda)} = k)}{\mathbb{E}\left[B^{(n,\lambda)}\right]}.$$

Let $g_{\tilde{D}-1}(x)$ and $g_{\tilde{B}_n-1}(x)$ be the probability generating functions of $\tilde{D} - 1$ and $\tilde{B}_n - 1$ respectively. Then (22) implies that

$$g_{\tilde{B}_n-1}(x) = g_{\tilde{D}-1}(1 - p(\lambda, n) + p(\lambda, n)x).$$

Letting $g_{\tilde{B}_n-1}^{\circ k}$ denote the kth iterate of $g_{\tilde{B}_n-1}$, note that Condition (CM4) is equivalent to

$$\liminf_{n\to\infty}\left(g_{\tilde{B}_n-1}^{\circ\left\lfloor \delta n^{\frac{\alpha-1}{\alpha+1}}\right\rfloor}(0)\right)^{\lfloor n^{1/(\alpha+1)}\rfloor} > 0. \tag{23}$$

(See for instance the discussion below Theorem 2.3.1 in [24].) As in the proof of [14, Proposition 5.25], it is sufficient to show that, for $t \in (0, \infty)$ and $r_n(t)$ the value such that

$$\int_{r_n(t)}^{1} \frac{dr}{g_{\tilde{B}_n-1}(1-r)-1+r} = n^{\frac{\alpha-1}{\alpha+1}}t,$$

we have

$$\limsup_{n\to\infty} n^{1/(\alpha+1)} r_n(t) < \infty.$$

Note that

$$\begin{aligned}\int_{r_n(t)}^{1} \frac{dr}{g_{\tilde{B}_n-1}(1-r)-1+r} &= \int_{r_n(t)}^{1} \frac{dr}{g_{\tilde{D}-1}\,(1-p(\lambda,n)r)-1+r} \\ &= \int_{p(\lambda,n)r_n(t)}^{p(\lambda,n)} \frac{ds}{p(\lambda,n)g_{\tilde{D}-1}\,(1-s)-p(\lambda,n)+s},\end{aligned}$$

where we change variable to $s = p(\lambda, n)r$. An elementary calculation yields that

$$g_{\tilde{D}-1}\,(1-s) = 1-(\rho/\mu-1)\,s + \frac{C_\alpha}{\mu}s^\alpha + o(s^\alpha)$$

as $s \to 0$, which implies that

$$p(\lambda,n)g_{\tilde{D}-1}\,(1-s) - p(\lambda,n) + s = -\lambda n^{-\frac{\alpha-1}{\alpha+1}}s + p(\lambda,n)\frac{C_\alpha}{\mu}s^\alpha + p(\lambda,n)o(s^\alpha)$$

as $s \to 0$. Then, setting $v = n^{1/(\alpha+1)}s$ implies that

$$\int_{r_n(t)}^{1} \frac{dr}{g_{\tilde{B}_n-1}(1-r)-1+r} = n^{\frac{\alpha-1}{\alpha+1}} \int_{n^{1/(\alpha+1)}p(\lambda,n)r_n(t)}^{n^{1/(\alpha+1)}p(\lambda,n)} \frac{dv}{-\lambda v + \frac{C_\alpha}{\mu}v^\alpha + o(v^\alpha)},$$

so that, since $p(n, \lambda) = \Theta(1)$ as $n \to \infty$, we obtain

$$\limsup_{n\to\infty} n^{1/(\alpha+1)} r_n(t) < \infty$$

as required. This implies that Condition (CM4) is satisfied and so, indeed, performing bond percolation on the half-edges of a sequence of vertices with i.i.d. degrees in the supercritical regime yields a degree distribution that satisfies the conditions of Theorem 9.

4.2.3. *The methods in [17].* In this section, we will further discuss the results and methods of Conchon-Kerjan and Goldschmidt [17]. As mentioned previously, Conchon-Kerjan and Goldschmidt study a specific case of the model defined in Section 4.1, namely the case where the degree sequence does not depend on n and $\lambda = 0$. They prove Theorem 9 for that family of models, which is the content of [17, Theorem 1.1]. (Their result includes the case $\alpha = 2$, which we do not consider here.)

The limit object is referred to as the *α-stable graph* for $\alpha \in (1, 2)$ (and the *Brownian graph* for $\alpha = 2$). They obtain an additional result that identifies the components of the limit object as $\mathbb{R}$-trees encoded by tilted excursions of an α-stable spectrally positive Lévy process for $\alpha \in (1, 2)$ (and tilted excursions of a Brownian motion for $\alpha = 2$) with additional vertex identifications. We cannot obtain such a description of the limit components in Theorem 9 because of their lack of self-similarity.

We now give an informal overview of the proof of [17, Theorem 1.1]. Much of the proof transfers over to our setting without change, so after this overview we will focus on the parts of the proof that are different in our setting. The method in [17] relies on using the configuration model in a depth-first manner, which we describe below.

From the description of the configuration model, it is clear that we can pick an order of connecting half-edges to our convenience. Hence we will choose an order that makes it similar to a depth-first exploration process. First, sample an i.i.d. degree sequence $D_1, \dots, D_n$ with $D_1 \geq 1$ almost surely. Start from a vertex v chosen with probability proportional to D_v and label its half-edges in an arbitrary way. We maintain a *stack*, which is an ordered list of the half-edges that we have seen but have not yet explored. Add all the half-edges of v to the stack, ordered according to their labels, with the lowest label being on top of the stack. From now on, at each step, if the stack is non-empty, take the half-edge from the top of the stack, and sample its pair uniformly among the unpaired half-edges, i.e. the remaining half-edges on the stack, and the unexplored half-edges not on the stack. If the paired half-edge was not on the stack, say it was linked to vertex w, then remove the paired half-edges from the system and place the remaining $D_w - 1$ half-edges of w on the top of the stack, arbitrarily labeled and in decreasing order of label, so that the lowest label of a half-edge of w is now on top of the stack (unless the degree of w is 1). If the paired half-edge was on the stack, remove both paired half-edges from the system. If the stack is empty, we start a new connected component by selecting an unexplored vertex with probability proportional to its degree, and putting its half-edges on top of the stack.

The argument then proceeds as follows:

1. Conditionally on $D_1, \dots, D_n$, if we order the vertices by the time their first half-edge is paired in the configuration model, the ordered degree sequence $\hat{D}_1^n, \dots, \hat{D}_n^n$ is a size-biased random ordering of $D_1, \dots, D_n$, and the forest encoded by the Łukasiewicz path $\tilde{S}_m^n = \sum_{i=1}^m \left(\hat{D}_i^n - 2\right)$ is closely related to the components of the multigraph given by the configuration model.

2. For $i \neq j$, in general, $\hat{D}_i^n$ is not equal to $\hat{D}_j^n$ in distribution, and $\hat{D}_i^n$ and $\hat{D}_j^n$ are dependent. These facts make $\tilde{S}^n$ hard to study. However, for $Z_1, \dots, Z_n$ i.i.d. with $\mathbb{P}(Z_1 = k) = \frac{k\mathbb{P}(D_1=k)}{\mathbb{E}[D_1]}$, for $m \leq n$ there exists a function ϕ_m^n such that for $g : \mathbb{N}^m \to \mathbb{R}$,

$$\mathbb{E}[g(\hat{D}_1^n, \hat{D}_2^n, \dots, \hat{D}_m^n)] = \mathbb{E}[\phi_m^n(Z_1, Z_2, \dots, Z_m) g(Z_1, Z_2, \dots, Z_m)].$$

Moreover, ϕ_m^n behaves well under rescaling, which allows the authors of [17] to study $S_m := \sum_{i=1}^m (Z_i - 2)$, and then use the measure change to translate results on this simpler process to results on $\tilde{S}^n$.

3. Indeed, under rescaling, S converges to an α-stable spectrally positive Lévy process L, jointly with its height process, and this result is used to show that $\tilde{S}$ (up to time $O(n^{\alpha/(\alpha+1)})$) converges to a process that is locally absolutely continuous to L, jointly with its height process.

4. The excursions of $\tilde{S}^n$ above its running infimum and the corresponding excursions above 0 of its height process encode individual trees in the forest. It is shown that the longest excursions explored up to time $\Theta(n^{\alpha/(\alpha+1)})$ converge under rescaling. The theory of size-biased point processes, developed by Aldous in [4], is then used to show that, in fact, with high probability, all large excursions are observed in the first $\Theta(n^{\alpha/(\alpha+1)})$ steps, and that the excursions listed in decreasing order of length converge as well.

5. By adding extra randomness and making some vertex identifications, one can modify the forest encoded by $\tilde{S}^n$ to yield a multigraph that is equal in law to the graph created by the configuration model, and these modifications behave well under rescaling the graph and taking limits.

6. Finally, the authors show that conditioning on the graph not containing multiple edges and loops does not affect the distribution of the largest components. This follows from an argument adapted from Joseph [38], which shows that the first loops and multiple edges are sampled far beyond the time scale $\Theta(n^{\alpha/(\alpha+1)})$, and so their presence or absence cannot affect the scaling limit.

4.3. Adapting the methods in [17] to the critical window

The largest barrier to generalizing the methods in [17] is showing the convergence under rescaling of S, jointly with its height process. The results proved in Section 3 allow us to do this. After that, it is trivial to extend most of the arguments in [17] to the critical window.

The convergence under rescaling of S is the content of Proposition 4. We then discuss the results in [17] that are not trivially extended to the critical window and need some further justification.

Let $D_1^n, \dots, D_n^n$ be i.i.d. with a degree distribution as specified in 4.1. Recall that the degree distribution depends on both λ and n, but that we have dropped the dependency on λ in the notation.

We consider the configuration model executed in depth-first order on vertices with degrees $D_1^n, \dots, D_n^n$. Let $(\hat{D}_1^n, \dots, \hat{D}_n^n)$ denote the degrees in order of discovery, so that $(\hat{D}_1^n, \dots, \hat{D}_n^n)$ is distributed as a size-biased random ordering of $D_1^n, \dots, D_n^n$. This is defined as follows.

Let Σ be a random permutation of $\{1, \dots, n\}$ such that

$$\mathbb{P}(\Sigma = \sigma \,|\, D_1^n, \dots, D_n^n) = \frac{D_{\sigma(1)}^n}{\sum_{j=1}^n D_{\sigma(j)}^n} \frac{D_{\sigma(2)}^n}{\sum_{j=2}^n D_{\sigma(j)}^n} \cdots \frac{D_{\sigma(n)}^n}{D_{\sigma(n)}^n}.$$

Then, by Proposition 3.2 of [17],

$$\begin{aligned} &\mathbb{P}(\hat{D}_1^n = k_1, \hat{D}_2^n = k_2, \dots, \hat{D}_n^n = k_n) \\ &= n! k_1 \mathbb{P}(D_1^n = k_1) k_2 \mathbb{P}(D_1^n = k_2) \cdots k_n \mathbb{P}(D_1^n = k_n) \prod_{i=1}^n \frac{1}{\sum_{j=i}^n k_j}. \end{aligned}$$

Now let $0 \le m \le n$ and $k_1, k_2, \dots, k_m \ge 1$, and define Ξ^n_{n-m} to be a random variable having the same law as $D^n_{m+1} + \cdots + D^n_n$. Then, for

$$\phi^n_m(k_1, k_2, \dots, k_m) := \frac{n!\mu^m}{(n-m)!}\mathbb{E}\left[\prod_{i=1}^m \frac{1}{\sum_{j=i}^m k_j + \Xi^n_{n-m}}\right],$$

Proposition 3.2 in [17] yields that for $Z^n_1, \dots, Z^n_n$ i.i.d. random variables with the size-biased degree of D^n_1, i.e.

$$\mathbb{P}(Z^n_1 = k) = \frac{k\mathbb{P}(D^n_1 = k)}{\mathbb{E}[D^n_1]},$$

for any test-function $g : \mathbb{N}^m \to \mathbb{R}$, we have

$$\mathbb{E}[g(\hat{D}^n_1, \hat{D}^n_2, \dots, \hat{D}^n_m)] = \mathbb{E}[\phi^n_m(Z^n_1, Z^n_2, \dots, Z^n_m)g(Z^n_1, Z^n_2, \dots, Z^n_m)];$$

that is, ϕ^n_m defines a measure change to get from a vector of size-biased distributed random variables to a vector of size-biased ordered random variables. We note that

$$\tilde{S}^n(k) := \sum_{i=1}^k \left(\hat{D}^n_i - 2\right)$$

is the Łukasiewicz path of a forest that is closely related to the depth-first spanning forest of our graph of interest, because it encodes the degrees in order of discovery. Therefore, the existence of the measure change motivates the study of the limit under rescaling of

$$S^n(k) := \sum_{i=1}^k \left(Z^n_i - 2\right)$$

and its corresponding height process. This is the content of the following proposition.

Proposition 4. *Let L be a spectrally positive α-stable Lévy process with Lévy measure $\pi(dx) = \frac{c}{\mu}x^{-(\alpha+1)}dx$. Then, for any $\lambda \in \mathbb{R}$, there exists a continuous modification of the height process of $(L_t + \lambda t, t \ge 0)$, which we will denote by H^λ. Moreover, for H^n the height process corresponding to S^n, we have that, as $n \to \infty$,*

$$\left(n^{-1/(\alpha+1)}S^n\left(\lfloor tn^{\alpha/(\alpha+1)}\rfloor\right), n^{-\frac{\alpha-1}{\alpha+1}}H^n\left(\lfloor tn^{\alpha/(\alpha+1)}\rfloor\right), t \ge 0\right)$$
$$\xrightarrow{d} (L_t + \lambda t, H^\lambda_t, t \ge 0)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})^2$.

We will use Theorem 2 to prove Proposition 4. We will first study the Laplace transform of Z^n_1, which is the content of the following lemma.

Lemma 9. *Define $\mathcal{L}_{Z^n_1}(s) = \mathbb{E}[\exp(-sZ^n_1)]$. Then*

$$\mathcal{L}_{Z^n_1}(s) = 1 - \left(2 + \lambda n^{-\frac{\alpha-1}{\alpha+1}}\right)s + \frac{C^{(n)}}{\mu_n}s^\alpha + o(s^\alpha)$$

as $s \to 0$.

Proof. Note that

$$\begin{aligned}\mathcal{L}_{Z_1^n}{}''(s) &= \mathbb{E}\left[(Z_1^n)^2 \exp(-sZ_1^n)\right] \\ &= \sum_{k=1}^{\infty} \frac{\mathbb{P}(D_1^n = k)k^3 e^{-sk}}{\mu_n} \\ &= \frac{c_n}{\mu_n} s^{\alpha-2}\Gamma(2-\alpha) + o(s^{\alpha-2})\end{aligned} \tag{24}$$

for $s \to 0$, where the last equality follows from the Euler–Maclaurin formula, using that $\mathbb{P}(Z_1^n = k) \sim \frac{c_n}{\mu_n} k^{-(\alpha+1)}$ as $k \to \infty$. Then, because $\mathbb{E}[Z_1^n] = 2 + \lambda n^{-\frac{\alpha-1}{\alpha+1}}$, integrating twice gives the result. □

Proof of Proposition 4. We will first prove that S^n converges in distribution under rescaling. Set

$$M^n(k) = \sum_{i=1}^{k} \left(Z_i^n - 2 - \lambda n^{-\frac{\alpha-1}{\alpha+1}}\right)$$

and

$$A^n(k) = k\lambda n^{-\frac{\alpha-1}{\alpha+1}},$$

so that

$$S^n(k) = M^n(k) + A^n(k)$$

is the Doob–Meyer decomposition of S^n. Observe that

$$\left(n^{-1/(\alpha+1)} A^n\left(\lfloor tn^{\alpha/(\alpha+1)} \rfloor\right), t \geq 0\right) \to (\lambda t, t \geq 0)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$.

We claim that for every $t \geq 0$,

$$n^{-1/(\alpha+1)} M^n\left(\lfloor tn^{\alpha/(\alpha+1)} \rfloor\right) \xrightarrow{d} L_t. \tag{25}$$

Firstly, observe that for all $\theta > 0$,

$$\mathbb{E}\left[\exp(-\theta L_t)\right] = \exp\left(t \frac{c\Gamma(2-\alpha)}{\mu\alpha(\alpha-1)} \theta^{\alpha}\right).$$

Recall that $C_\alpha = \frac{c\Gamma(2-\alpha)}{\alpha(\alpha-1)}$, and set $C^{(n)} = \frac{c_n\Gamma(2-\alpha)}{\alpha(\alpha-1)}$, so that $C^{(n)} \to C_\alpha$ as $n \to \infty$. We will show that for every $\theta > 0$,

$$\mathbb{E}\left[\exp\left(-\theta n^{-1/(\alpha+1)} M^n\left(\lfloor tn^{\alpha/(\alpha+1)} \rfloor\right)\right)\right] \to \exp\left(t \frac{C_\alpha}{\mu} \theta^{\alpha}\right)$$

as $n \to \infty$, which will prove the claim.

Note that

$$\begin{aligned}&\mathbb{E}\left[\exp(-\theta n^{-1/(\alpha+1)} M^n(\lfloor tn^{\alpha/(\alpha+1)} \rfloor)\right] \\ &= \left[\mathcal{L}_{Z_1^n}(\theta n^{-1/(\alpha+1)}) \exp\left(\theta n^{-1/(\alpha+1)}\left(2 + \lambda n^{-\frac{\alpha-1}{\alpha+1}}\right)\right)\right]^{\lfloor tn^{\alpha/(\alpha+1)} \rfloor}.\end{aligned} \tag{26}$$

Then, Lemma 9 implies that

$$\mathcal{L}_{Z_1^n}(s) = \exp\left(- \left(2 + \lambda n^{-\frac{\alpha-1}{\alpha+1}}\right) s + \frac{C^{(n)}}{\mu_n} s^\alpha + o(s^\alpha) \right)$$

as $s \to 0$. Plugging this into (26), we find that as $n \to \infty$,

$$\mathbb{E}\left[\exp\left(-\theta n^{-1/(\alpha+1)} M^n(\lfloor t n^{\alpha/(\alpha+1)} \rfloor)\right)\right] \to \exp\left(t \frac{C_\alpha}{\mu} \theta^\alpha\right),$$

which proves (25).

Now, using [39, Theorem 16.14], we may deduce that as $n \to \infty$,

$$\left(n^{-1/(\alpha+1)} S^n\left(\lfloor t n^{\alpha/(\alpha+1)} \rfloor\right), t \geq 0\right) \xrightarrow{d} (L_t + \lambda t, t \geq 0) \tag{27}$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$.

In order to also obtain the convergence of the height process under rescaling, we note that for $\lambda \leq 0$, we can directly apply [24, Theorems 1.4.3 and 2.2.1], stated in this work as Theorem 1. In the case of $\lambda > 0$, we apply Theorem 2. In both results, we use the scaling parameters $p = n^{1/\alpha+1}$ and $\gamma_p = n^{\frac{\alpha-1}{\alpha+1}}$. The conditions of the theorems then follow directly from our assumptions on the degree distributions and (27). This implies the existence of a continuous modification of the height process and the claimed convergence. □

We will now prove that in the cases we consider, ϕ_m^n behaves well under rescaling. This is the content of the following proposition, which is a generalization of the proof of [17, Proposition 4.3].

Proposition 5. *Set*

$$\phi(t) = \exp\left(\frac{1}{\mu} \int_0^t s dL_s - \frac{C_\alpha}{(\alpha+1)\mu^{\alpha+1}} t^{\alpha+1}\right).$$

Then

$$\phi(n, t) \xrightarrow{d} \phi(t)$$

as $n \to \infty$, and the sequence $\{\phi(n, t), n \in \mathbb{N}\}$ is uniformly integrable.

The proof will follow the structure of the proof of [17, Proposition 4.3], but we will need to adapt the technical lemmas presented there to our more general setting.

Proof. The following technical lemmas need justification in the more general setting:

- For $\lambda = 0$ and a degree distribution that does not depend on n, for any $t \geq 0$, by [17, Lemma 4.4] it holds that, as $n \to \infty$,

$$\frac{1}{n} \sum_{k=0}^{\lfloor t n^{\alpha/(\alpha+1)} \rfloor - 1} S^n(k) \xrightarrow{d} \int_0^t L_s ds.$$

  Using the continuous mapping theorem (see e.g. [27, Theorem 3.2.4]), we find that for general λ and for the degree distribution depending on n, for any $t \geq 0$,

$$\frac{1}{n} \sum_{k=0}^{\lfloor t n^{\alpha/(\alpha+1)} \rfloor - 1} S^n(k) \xrightarrow{d} \int_0^t L_s ds + \frac{\lambda t^2}{2}. \tag{28}$$

- Define $\mathcal{L}_{D_1^n}(s) := \mathbb{E}[\exp(-sD_1^n)]$. Then [17, Lemma 4.5a] and the argument thereafter state that for $\lambda = 0$ and a degree distribution that does not depend on n,

$$\begin{aligned}\mathcal{L}_{D_1^n}(s) &= 1 - \mu_n s + \mu_n s^2 - \frac{C_\alpha}{\alpha+1} s^{\alpha+1} + o(s^{\alpha+1}) \\ &= \exp\left(- \mu_n s + \frac{\mu_n(2-\mu_n)}{2} s^2 - \frac{C_\alpha}{\alpha+1} s^{\alpha+1} + o(s^{\alpha+1}) \right).\end{aligned}$$

  In our set-up, we find, similarly to how we obtained $\mathcal{K}_n''(s)$, that

$$\mathcal{L}_{D_1^n}'''(s) = -c\Gamma(2-\alpha)s^{\alpha-2} + o(s^{\alpha-2})$$

  as $s \to 0$, and using $\mathbb{E}[D_1^n] = \mu_n$ and $\mathbb{E}\left[(D_1^n)^2\right] = \left(2 + \lambda n^{-\frac{\alpha-1}{\alpha+1}}\right)\mu_n$, we get, by integrating three times, that

$$\begin{aligned}\mathcal{L}_{D_1^n}(s) &= 1 - \mu_n s + \left(1 + \frac{\lambda}{2} n^{-\frac{\alpha-1}{\alpha+1}}\right)\mu_n s^2 - \frac{C^{(n)}}{\alpha+1} s^{\alpha+1} + o(s^{\alpha+1}) \\ &= \exp\left(- \mu_n s + \frac{\mu_n(2 + \lambda n^{-\frac{\alpha-1}{\alpha+1}} - \mu_n)}{2} s^2 - \frac{C^{(n)}}{\alpha+1} s^{\alpha+1} + o(s^{\alpha+1}) \right)\end{aligned} \tag{29}$$

  as $s \to 0$.

- For $\lambda = 0$ and a degree distribution that does not depend on n, for $m = \lfloor t n^{\alpha/(\alpha+1)} \rfloor$, by [17, Lemma 4.6] it holds that, as $n \to \infty$,

$$\exp\left(m - \frac{2+\mu_n}{2\mu_n}\frac{m^2}{n}\right)\left[\mathcal{L}\left(\frac{m}{n\mu_n}\right)\right]^{n-m} \to \exp\left(- \frac{C_\alpha}{(\alpha+1)\mu^{\alpha+1}} t^{\alpha+1}\right).$$

  By (29) we straightforwardly obtain that, in our set-up, as $n \to \infty$,

$$\exp\left(m - \frac{2+\mu_n}{2\mu_n}\frac{m^2}{n}\right)\left[\mathcal{L}\left(\frac{m}{n\mu_n}\right)\right]^{n-m} \to \exp\left(- \frac{C_\alpha}{(\alpha+1)\mu^{\alpha+1}} t^{\alpha+1} + \frac{\lambda t^2}{2\mu}\right). \tag{30}$$

- For $\lambda = 0$ and a degree distribution that does not depend on n, for $s(0) = 0$ and $s(i) = \sum_{j=1}^{i}(k_j - 2)$, for $i \geq 1$, by [17, Lemma 4.7] it holds that

$$\phi_m^n(k_1, \ldots, k_m) \geq \exp\left(\frac{1}{n\mu_n} \sum_{i=0}^{m} (s(i) - s(m)) - \frac{C_\alpha}{(\alpha+1)\mu^{\alpha+1}} t^{\alpha+1}\right)(1 + o(1)).$$

  Adapting the proof to our set-up, using (30), we find that for $s(0) = 0$ and $s(i) = \sum_{j=1}^{i}(k_j - 2)$, for $i \geq 1$,

$$\phi_m^n(k_1, \ldots, k_m) \geq (1 + o(1)) \exp\left(\frac{1}{n\mu_n} \sum_{i=0}^{m} (s(i) - s(m)) - \frac{C_\alpha t^{\alpha+1}}{(\alpha+1)\mu^{\alpha+1}} + \frac{\lambda t^2}{2\mu}\right).$$

  The proof of this can be found in the appendix.

Now we have that

$$\phi(n,t) \geq \underline{\phi}(n,t) := \exp\left(\frac{1}{n\mu_n}\sum_{i=0}^{m}(S^n(i)-S^n(m)) - \frac{C_\alpha}{(\alpha+1)\mu^{\alpha+1}}t^{\alpha+1} + \frac{\lambda t^2}{2\mu}\right).$$

But then, by (27) and (28), since

$$\frac{1}{n}\sum_{k=0}^{\lfloor tn^{\alpha/(\alpha+1)}\rfloor-1}\left(S^n(k) - S^n(\lfloor tn^{\alpha/(\alpha+1)}\rfloor - 1)\right) \xrightarrow{d} \int_0^t (L_s - L_t)ds - \frac{\lambda t^2}{2},$$

we get that

$$\underline{\phi}(n,t) \xrightarrow{d} \exp\left(\frac{1}{\mu}\int_0^t (L_s - L_t)ds - \frac{C_\alpha}{(\alpha+1)\mu^{\alpha+1}}t^{\alpha+1}\right).$$

We then finish the proof in the same way as the proof of [17, Proposition 3.3] to conclude that

$$\phi(n,t) \xrightarrow{d} \phi(t)$$

as $n \to \infty$, and that the sequence $\{\phi(n,t), n \in \mathbb{N}\}$ is uniformly integrable. The details can be found in the appendix. □

Remember that $n^{-1/(\alpha+1)}\left(S^n\left(\lfloor tn^{\alpha/(\alpha+1)}\rfloor\right), t \geq 0\right)$ converges in law to $(L_t + \lambda t, t \geq 0)$ as $n \to \infty$. We can get from the process $\left(S^n\left(\lfloor sn^{\alpha/(\alpha+1)}\rfloor\right), 0 \leq s \leq t\right)$ to the process $\left(\tilde{S}^n\left(\lfloor sn^{\alpha/(\alpha+1)}\rfloor\right), 0 \leq s \leq t\right)$ via the measure change $\phi(n,t)$. The random variable $\phi(n,t)$ converges in law to $\phi(t)$ as $n \to \infty$. Therefore, we will define the process $(\tilde{K}^\lambda, \tilde{H}^\lambda)$ via the following measure change. For $t \geq 0$, for every non-negative integrable functional $F : \mathbb{D}([0,t], \mathbb{R}^2) \to \mathbb{R}$,

$$\mathbb{E}[F(\tilde{K}^\lambda_s, \tilde{H}^\lambda_s, 0 \leq s \leq t)] = \mathbb{E}\left[\phi(t)F(L_s + \lambda s, H^\lambda_s, 0 \leq s \leq t)\right]. \tag{31}$$

Proposition 6. *We have*

$$\left(n^{-1/(\alpha+1)}\tilde{S}^n\left(\lfloor sn^{\alpha/(\alpha+1)}\rfloor\right), n^{-\alpha/(\alpha+1)}\tilde{H}^n\left(\lfloor sn^{\alpha/(\alpha+1)}\rfloor\right), 0 \leq s \leq t\right)$$
$$\xrightarrow{d} (\tilde{K}^\lambda_s, \tilde{H}^\lambda_s, 0 \leq s \leq t)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^2)$ *as* $n \to \infty$.

Proof. We want to show that for any $t \geq 0$ and any bounded continuous test function $f : \mathbb{D}([0,t], \mathbb{R}^2) \to \mathbb{R}$, for $\tilde{H}^n$ the height process corresponding to $\tilde{S}^n$,

$$\mathbb{E}\left[f\left(n^{-1/(\alpha+1)}\tilde{S}^n\left(\lfloor sn^{\alpha/(\alpha+1)}\rfloor\right), n^{-\alpha/(\alpha+1)}\tilde{H}^n\left(\lfloor sn^{\alpha/(\alpha+1)}\rfloor\right), 0 \leq s \leq t\right)\right]$$
$$\to \mathbb{E}\left[f(\tilde{K}^\lambda_s, \tilde{H}^\lambda_s, 0 \leq s \leq t)\right]$$

as $n \to \infty$. Using our measure change, this is equivalent to showing that for any $t \geq 0$ and any bounded continuous test function $f : \mathbb{D}([0,t], \mathbb{R}^2) \to \mathbb{R}$,

$$\mathbb{E}\left[\phi(n,t)f\left(n^{-1/(\alpha+1)}S^n\left(\lfloor sn^{\alpha/(\alpha+1)}\rfloor\right), n^{-\alpha/(\alpha+1)}H^n\left(\lfloor sn^{\alpha/(\alpha+1)}\rfloor\right), 0 \leq s \leq t\right)\right]$$
$$\to \mathbb{E}[\phi(t)f(L_s + \lambda s, H^\lambda_s, 0 \leq s \leq t)].$$

We now finish as in the proof of [17, Theorem 4.1] to obtain the result. □

The following proposition characterizes the law of $\tilde{K}^{\lambda}$.

Proposition 7. *For $\tilde{K}^{\lambda}(t)$ as defined in (31), for $\theta > 0$,*

$$\mathbb{E}\left[\exp\left(-\theta\tilde{K}^{\lambda}(t)\right)\right] = \exp\left(-\theta\lambda t + \theta C_{\alpha}\frac{t^{\alpha}}{\mu^{\alpha}} + \int_0^t ds \int_0^{\infty} \frac{c}{\mu} x^{-(\alpha+1)} e^{-xs/\mu}(e^{-\theta x} - 1 + \theta x)dx\right).$$

The proof can be found in the appendix.

Proof of Theorem 9. Given Proposition 6, the rest of the proof is completely analogous to the proof of Theorem 1.1 in [17]. Proposition 7 characterizes the law of the encoding process of the components of the limit object. □

Appendix A. Proofs of technical results

Proof of Lemma 4. First note that since f and g are increasing, the function given by $(\inf\{s \geq 0 : f(t) < g(s)\}, t \geq 0)$ is also increasing, and so in particular it has limits from the left and from the right at every point of its domain. Fix $t \geq 0$. Suppose that

$$\lim_{q \uparrow t} \inf\{s \geq 0 : f(q) < g(s)\} < \inf\{s \geq 0 : f(t) < g(s)\}.$$

Then we must have that

(1) there is an $\tilde{s}$ such that $f(t) = g(\tilde{s})$, and

(2) $f(q) < f(t)$ for all $q < t$.

It follows that

$$\lim_{q \uparrow t} \inf\{s \geq 0 : f(q) < g(s)\} = \tilde{s}.$$

But then, since g is strictly increasing,

$$\inf\{s \geq 0 : f(t) < g(s)\} = \inf\{s \geq 0 : g(\tilde{s}) < g(s)\} = \tilde{s},$$

so we must have

$$\lim_{q \uparrow t} \inf\{s \geq 0 : f(q) < g(s)\} = \inf\{s \geq 0 : f(t) < g(s)\}.$$

We also need to show that

$$\lim_{q \downarrow t} \inf\{s \geq 0 : f(q) < g(s)\} = \inf\{s \geq 0 : f(t) < g(s)\}.$$

Fix $\epsilon > 0$. Suppose $\inf\{s \geq 0 : f(t) < g(s)\} = \tilde{s}$, which implies that $f(t) < g(\tilde{s} + \epsilon)$. Then, as $q \downarrow t$, we have $f(q) \downarrow f(t)$, so for $q > t$ small enough, $f(q) < g(\tilde{s} + \epsilon)$. Hence,

$$\inf\{s \geq 0 : f(q) < g(s)\} \leq \tilde{s} + \epsilon,$$

which proves that

$$\lim_{q \downarrow t} \inf\{s \geq 0 : f(q) < g(s)\} = \inf\{s \geq 0 : f(t) < g(s)\}.$$

Therefore, $(\inf\{s \geq 0 : f(t) < g(s)\}, t \geq 0)$ is continuous.

By Proposition 3.6.5 in Ethier and Kurtz [29], proving

$$(\inf\{s \geq 0 : f_n(t) < g_n(s)\}, t \geq 0) \to (\inf\{s \geq 0 : f(t) < g(s)\}, t \geq 0)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$ is then equivalent to showing that for all $t \in \mathbb{R}_+$, and for all $(t_n)_{n\geq 0}$ in $\mathbb{R}_+$ such that $t_n \to t$, we have

$$\inf\{s \geq 0 : f_n(t_n) < g_n(s)\} \to \inf\{s \geq 0 : f(t) < g(s)\}.$$

Suppose $\inf\{s \geq 0 : f(t) < g(s)\} = \tilde{s}$. Fix $\epsilon > 0$ and $(t_n)_{n\geq 0}$ in $\mathbb{R}_+$ such that $t_n \to t$. We will first show that for n large enough,

$$\inf\{s \geq 0 : f_n(t_n) < g_n(s)\} \leq \tilde{s} + \epsilon.$$

By the definition of $\tilde{s}$ and monotonicity, we have that $f(t) < g(\tilde{s} + \epsilon/2)$. Since g is strictly increasing, we have that there is a $\delta > 0$ such that $g(\tilde{s} + \epsilon/2) = f(t) + \delta$. Moreover, since $f_n \to f$ in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$ and f is continuous, $f_n(t_n) \to f(t)$, so we may pick n large enough so that

$$|f_n(t_n) - f(t)| < \delta/2.$$

Since $g_n \to g$ in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$, for n large enough there exists a monotone bijection $\lambda_n : [0, \tilde{s} + 2\epsilon] \to [0, \tilde{s} + 2\epsilon]$ such that

$$\sup_{t\in[0,\tilde{s}+2\epsilon]} |\lambda_n(t) - t| < \epsilon/4$$

and

$$\sup_{t\in[0,\tilde{s}+2\epsilon]} |g_n(\lambda_n(t)) - g(t)| < \delta/2.$$

Hence,

$$f_n(t_n) < f(t) + \delta/2 < g_n(\lambda_n(\tilde{s} + \epsilon/2))$$

and

$$f_n(t_n) < g_n(s)$$

for $s = \tilde{s} + 3\epsilon/4$. Hence, for n large enough,

$$\inf\{s \geq 0 : f_n(t_n) < g_n(s)\} < \tilde{s} + \epsilon.$$

We now want to show that for n large enough,

$$\inf\{s \geq 0 : f_n(t_n) < g_n(s)\} \geq \tilde{s} - \epsilon.$$

Fix $s < \tilde{s} - \epsilon$. We will be done if we can show that $f_n(t_n) \geq g_n(s)$. Since g is strictly increasing, we know that there exists a $\delta > 0$ such that $g(\tilde{s} - \epsilon/2) + \delta = g(\tilde{s} - \epsilon/4)$. Also, by the definition of $\tilde{s}$, we have $f(t) \geq g(\tilde{s} - \epsilon/4)$. Pick n large enough that there exists a monotone bijection $\lambda_n : [0, \tilde{s} + 2\epsilon] \to [0, \tilde{s} + 2\epsilon]$ such that

$$\sup_{t\in[0,\tilde{s}+2\epsilon]} |\lambda_n(t) - t| < \epsilon/4$$

and

$$\sup_{t\in[0,\tilde{s}+2\epsilon]} |g_n(\lambda_n(t)) - g(t)| < \delta/2,$$

and such that

$$|f_n(t_n) - f(t)| < \delta/2.$$

Then

$$\begin{aligned} g_n(s) &\le g_n(\tilde{s} - \epsilon) \\ &\le g(\tilde{s} - \epsilon/2) + \delta/2 \\ &= g(\tilde{s} - \epsilon/4) - \delta/2 \\ &\le f(t) - \delta/2 \\ &\le f_n(t_n), \end{aligned}$$

which proves the statement.

Finally we show that for $(\epsilon_n)_{n\ge 0}$ such that $\epsilon_n \downarrow 0$,

$$(\inf\{s \ge 0 : f_n(t - \epsilon_n s) < g_n(s)\}, t \ge 0) \to (\inf\{s \ge 0 : f(t) < g(s)\}, t \ge 0)$$

in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$. Fix $t \in \mathbb{R}_+$, $t_n \to t$, and $\epsilon > 0$. We need to show that

$$\inf\{s \ge 0 : f_n(t_n - \epsilon_n s) < g_n(s)\} \to \inf\{s \ge 0 : f(t) < g(s)\} =: \tilde{s}.$$

First, note that

$$\inf\{s \ge 0 : f_n(t_n - \epsilon_n s) < g_n(s)\} \le \inf\{s \ge 0 : f_n(t_n) < g_n(s)\},$$

which is smaller than $\tilde{s} + \epsilon$ for n large enough by previous results. Moreover, note that for $s \le \tilde{s} - \epsilon$,

$$f_n(t_n - \epsilon_n s) \ge f_n(t_n - \epsilon_n(\tilde{s} - \epsilon)),$$

which is larger than $g_n(s)$ for n large enough by previous results, since $t_n - \epsilon_n(\tilde{s} - \epsilon) \to t$. Hence,

$$\inf\{s \ge 0 : f_n(t_n - \epsilon_n s) < g_n(s)\} \ge \tilde{s} - \epsilon$$

for n large enough, which concludes the proof. □

We will now prove two technical lemmas needed for the proof of Proposition 5.

Lemma 10. *For* $m = \lfloor t n^{\alpha/(\alpha+1)} \rfloor$, *for*

$$\phi_m^n(k_1, \dots, k_m) = \frac{n! \mu_n^m}{(n-m)!} \mathbb{E}\left[\prod_{i=1}^m \frac{1}{\sum_{j=i}^m k_j + \Xi_{n-m}} \right],$$

$s(0) = 0$ *and* $s(i) = \sum_{j=1}^i (k_j - 2)$, *for* $i \ge 1$, *we have that*

$$\phi_m^n(k_1, \dots, k_m) \ge (1 + o(1)) \exp\left(\frac{1}{n\mu_n} \sum_{i=0}^m (s(i) - s(m)) - \frac{C^{(n)} t^{\alpha+1}}{(\alpha+1)\mu^{\alpha+1}} + \frac{\lambda t^2}{2\mu} \right).$$

The arguments presented are an adaption of the proof of [17, Lemma 4.7].

Proof. We can write

$$\begin{aligned} \phi_m^n(k_1, \dots, k_m) &= \prod_{i=1}^{m-1} \left(1 - \frac{i}{n}\right) \mathbb{E}\left[\prod_{i=1}^m \left(\frac{n\mu_n}{\sum_{j=i}^m k_j + \Xi_{n-m}} \right) \right] \\ &= \mathbb{E}\left[\exp\left(\sum_{i=1}^{m-1} \log\left(1 - \frac{i}{n}\right) - \sum_{i=1}^m \log\left(\frac{\Xi_{n-m}}{n\mu_n} + \frac{1}{n\mu_n} \sum_{j=i}^m k_j \right) \right) \right]. \end{aligned}$$

Then, since $\log(1+x) \le x$ for all $x \in (-1, \infty)$, and there is a $C > 0$ such that for all $1 \le i \le m-1$, $\log(1-i/n) \ge -i/n - Cm^2/n^2$, we have that

$$\begin{aligned}\phi_m^n(k_1, \dots, k_m) &\ge \exp\left(\frac{1}{n\mu_n}\sum_{i=0}^{m}(s(i)-s(m)) - \frac{m(m-1)}{2n} - \frac{Cm^3}{n^2} - m - \frac{m(m+1)}{n\mu_n}\right)\\ &\quad\times \mathbb{E}\left[\exp\left(\frac{-m}{n\mu_n}\Xi_{n-m}\right)\right]\\ &= \exp\left(\frac{1}{n\mu_n}\sum_{i=0}^{m}(s(i)-s(m))\right)\exp\left(m - \frac{(2+\mu_n)m^2}{2\mu_n n}\right)\\ &\quad\times\left[\mathcal{L}_{D_1^n}\left(\frac{m}{n\mu_n}\right)\right]^{n-m}\exp\left(\frac{\mu_n-2}{2\mu_n n} - \frac{Cm^3}{n^2}\right).\end{aligned}$$

Note that by the definition of m, $m^3/n^2 = \Theta\left(n^{(\alpha-2)/(\alpha+1)}\right) = o(1)$, so the final exponent tends to 1 as $n \to \infty$. Then by (30), which states that as $n \to \infty$,

$$\exp\left(m - \frac{2+\mu_n}{2\mu_n}\frac{m^2}{n}\right)\left[\mathcal{L}_{D_1^n}\left(\frac{m}{n\mu_n}\right)\right]^{n-m} \to \exp\left(-\frac{C_\alpha}{(\alpha+1)\mu^{\alpha+1}}t^{\alpha+1} + \frac{\lambda t^2}{2\mu}\right),$$

the desired result follows. □

Furthermore, recall that

$$\phi(n,t) = \phi_m^n(Z_1^n, \dots, Z_m^n)$$

and

$$\phi(t) = \exp\left(-\frac{1}{\mu}\int_0^t s dL_s - \frac{C_\alpha}{(\alpha+1)\mu^{\alpha+1}}t^{\alpha+1}\right).$$

Lemma 11. *As $n \to \infty$,*

$$\phi(n,t) \xrightarrow{d} \phi(t),$$

and $\{\phi(n,t), n \in \mathbb{N}\}$ is uniformly integrable.

The arguments presented are an adaptation of the proof of [17, Proposition 4.3].

Proof. Recall that $S^n(k) = \sum_{i=1}^{k}(Z_i^n - 2)$. By the statement proved above, we have that

$$\phi(n,t) \ge \underline{\phi}(n,t) := (1+o(1))\exp\left(\frac{1}{n\mu_n}\sum_{i=0}^{m}(S^n(i)-S^n(m)) - \frac{C^{(n)}t^{\alpha+1}}{(\alpha+1)\mu^{\alpha+1}} + \frac{\lambda t^2}{2\mu}\right).$$

Then, by Theorem 4 and (28), we get that

$$\frac{1}{n\mu_n}\sum_{i=0}^{m}(S^n(i)-S^n(m)) \xrightarrow{d} \frac{1}{\mu}\int_0^t (L_s^\lambda - L_t^\lambda)ds - \frac{\lambda t^2}{2\mu}.$$

Hence, we get that

$$\begin{aligned}\underline{\phi}(n,t) &\xrightarrow{d} \exp\left(\frac{1}{\mu}\int_0^t (L_s - L_t)ds - \frac{C_\alpha}{(\alpha+1)\mu^{\alpha+1}}t^{\alpha+1}\right)\\ &= \exp\left(-\frac{1}{\mu}\int_0^t s dL_s - \frac{C_\alpha}{(\alpha+1)\mu^{\alpha+1}}t^{\alpha+1}\right),\end{aligned}$$

and the right-hand side equals $\phi(t)$, which has mean 1 by [17, Proposition 3.2]. Also, $\mathbb{E}[\phi(n,t)] = 1$. Then, by [17, Lemma 4.8], we must have that

$$\phi(n,t) \xrightarrow{d} \phi(t),$$

and $\{\phi(n,t), n \in \mathbb{N}\}$ is uniformly integrable. □

Proof of Proposition 7. Let $\tilde{K}^{\lambda}_t$ be defined as in (31): for $t \geq 0$, for every non-negative integrable functional $F : \mathbb{D}([0,t], \mathbb{R}^2) \to \mathbb{R}$,

$$\mathbb{E}[F((\tilde{K}^{\lambda}_s), 0 \leq s \leq t)] = \mathbb{E}\left[\phi(t) F((L_s + \lambda s), 0 \leq s \leq t)\right].$$

We want to show that

$$\mathbb{E}\left[\exp\left(-\theta \tilde{K}^{\lambda}_t\right)\right] = \exp\left(-\theta\lambda t + \theta C_\alpha \frac{t^\alpha}{\mu^\alpha} + \int_0^t ds \int_0^\infty \frac{c}{\mu} x^{-(\alpha+1)} e^{-xs/\mu} dx (e^{-\theta x} - 1 + \theta x)\right).$$

The proof will be an adaptation of the proof of [17, Proposition 6.2]. As before, let L be the spectrally positive α-stable Lévy process having Lévy measure $\pi(dx) = \frac{c}{\mu} x^{-(\alpha+1)} dx$ and Laplace transform

$$\Upsilon(\theta) = \frac{C_\alpha}{\mu} \theta^\alpha.$$

Let X_t be the process with Laplace transform

$$\mathbb{E}\left[\exp(-\theta X_t)\right] = \exp\left(\int_0^t ds \int_0^\infty \pi(dx) \left(e^{-\theta x} - 1 + \theta x\right) e^{-\theta x s}\right),$$

and let

$$A_t = -C_\alpha \frac{t^\alpha}{\mu^\alpha} = -\mu \Upsilon\left(\frac{t}{\mu}\right).$$

Set

$$\tilde{L}^{\lambda}_t = X_t + A_t + \lambda t.$$

We claim that

$$(\tilde{L}^{\lambda}_t, t \geq 0) \stackrel{d}{=} (\tilde{K}^{\lambda}_t, t \geq 0).$$

As in [17], decomposing π gives that

$$\mathbb{E}[\exp(-\theta X_t)] = \exp\left(-\mu\theta \Upsilon\left(\frac{t}{\mu}\right) + \int_0^t \Upsilon\left(\theta + \frac{s}{\mu}\right) ds - \int_0^t \Upsilon\left(\frac{s}{\mu}\right) ds\right),$$

so that

$$\begin{aligned}\mathbb{E}[\exp(-\theta \tilde{L}^{\lambda}_t)] &= \exp\left(\int_0^t \Upsilon\left(\theta + \frac{s}{\mu}\right) ds - \int_0^t \Upsilon\left(\frac{s}{\mu}\right) ds - \theta\lambda t\right) \\ &= \mathbb{E}\left[\exp\left(-\int_0^t \left(\theta + \frac{s}{\mu}\right) dL_s - \int_0^t \Upsilon\left(\frac{s}{\mu}\right) ds - \theta\lambda t\right)\right].\end{aligned}$$

Let $0 = t_0 < t_1 < \cdots < t_m = t$ and let $\theta_1, \ldots, \theta_m \in \mathbb{R}_+$. Then, since X has independent increments, by the argument as presented above,

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(-\sum_{i=1}^{m}\theta_i\left(\tilde{L}^{\lambda}_{t_i}-\tilde{L}^{\lambda}_{t_{i-1}}\right)\right)\right]\\
&=\prod_{i=1}^{m}\mathbb{E}\left[\exp\left(-\int_{t_{i-1}}^{t_i}\left(\theta_i+\frac{s}{\mu}\right)dL_s-\int_{t_{i-1}}^{t_i}\Upsilon\left(\frac{s}{\mu}\right)ds-\theta_i\lambda(t_i-t_{i-1})\right)\right]\\
&=\mathbb{E}\left[\exp\left(-\sum_{i=1}^{m}\int_{t_{i-1}}^{t_i}\left(\theta_i+\frac{s}{\mu}\right)dL_s-\int_{0}^{t}\Upsilon\left(\frac{s}{\mu}\right)ds-\sum_{i=1}^{m}\theta_i\lambda(t_i-t_{i-1})\right)\right]\\
&=\mathbb{E}\left[\exp\left(-\sum_{i=1}^{m}\theta_i\left((L_{t_i}+\lambda t_i)-(L_{t_{i-1}}+\lambda t_{i-1})\right)-\frac{1}{\mu}\int_0^t sdL_s-\int_0^t\Upsilon\left(\frac{s}{\mu}\right)\right)\right],
\end{aligned}$$

where we use the fact that L also has independent increments and perform integration by parts. This proves the statement. □

Acknowledgements

The author would like to thank Christina Goldschmidt for many productive meetings. She would also like to thank Matthias Winkel for helpful input in the early stages of this project, Thomas Duquesne for pointing out interesting literature, and Thomas Hughes for careful proofreading. Finally, she would like to thank an anonymous referee for their diligent reading of the paper and for their useful comments and suggestions.

Funding information

The author was funded by a Clarendon Scholarship and a CRM–ISM Postdoctoral fellowship.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] Abraham, R. and Delmas, J.-F. (2012). A continuum-tree-valued Markov process. *Ann. Prob.* **40**, 1167–1211.

[2] Abraham, R., Delmas, J.-F. and He, H. (2015). Pruning of CRT-sub-trees. *Stoch. Process. Appl.* **125**, 1569–1604.

[3] Addario-Berry, L., Broutin, N., Goldschmidt, C. and Miermont, G. (2017). The scaling limit of the minimum spanning tree of the complete graph. *Ann. Prob.* **45**, 3075–3144.

[4] Aldous, D. (1991). The continuum random tree II: an overview. In *Stochastic Analysis*, eds M. T. Barlow and N. H. Bingham, Cambridge University Press, pp. 23–70.

[5] Aldous, D. (1993). The continuum random tree III. *Ann. Prob.* **21**, 248–289.

[6] Athreya, K. B. and Ney, P. (1972). *Branching Processes*. Springer, Berlin.

[7] Athreya, S., Löhr, W. and Winter, A. (2016). The gap between Gromov-vague and Gromov–Hausdorff-vague topology. *Stoch. Process. Appl.* **126**, 2527–2553.

[8] Bertoin, J. (1991). Sur la décomposition de la trajectoire d'un processus de Lévy spectralement positif en son infimum. *Ann. Inst. H. Poincaré Prob. Statist.* **27**, 537–547.

[9] Bertoin, J. (1993). Splitting at the infimum and excursions in half-lines for random walks and Lévy processes. *Stoch. Process. Appl.* **47**, 17–35.

[10] BERTOIN, J. (1996). *Lévy Processes*. Cambridge University Press.
[11] BHAMIDI, S., DHARA, S., VAN DER HOFSTAD, R. AND SEN, S. (2020). Universality for critical heavy-tailed network models: metric structure of maximal components. *Electron. J. Prob.* **25**, article no. 47.
[12] BHAMIDI, S., DHARA, S., VAN DER HOFSTAD, R. AND SEN, S. (2022). Global lower mass-bound for critical configuration models in the heavy-tailed regime. *Electron. J. Prob.* **27**, 1–29.
[13] BOLLOBÁS, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Europ. J. Combinatorics* **1**, 311–316.
[14] BROUTIN, N., DUQUESNE, T. AND WANG, M. (2021). Limits of multiplicative inhomogeneous random graphs and Lévy trees: limit theorems. *Prob. Theory Relat. Fields* **181**, 865–973.
[15] CHAUMONT, L. (1994). Sur certains processus de Lévy conditionnés à rester positifs. *Stoch. Stoch. Reports* **47**, 1–20.
[16] CHAUMONT, L. (1996). Conditionings and path decompositions for Lévy processes. *Stoch. Process. Appl.* **64**, 39–54.
[17] CONCHON-KERJAN, G. AND GOLDSCHMIDT, C. (2023). The stable graph: the metric space scaling limit of a critical random graph with i.i.d. power-law degrees. *Ann. Prob.* **51**, 1–69.
[18] DELMAS, J.-F. (2008). Height process for super-critical continuous state branching process. *Markov Process. Relat. Fields* **14**, 309–326.
[19] DHARA, S., VAN DER HOFSTAD, R., VAN LEEUWAARDEN, J. S. H. AND SEN, S. (2017). Critical window for the configuration model: finite third moment degrees. *Electron. J. Prob.* **22**, article no. 16.
[20] DHARA, S., VAN DER HOFSTAD, R., VAN LEEUWAARDEN, J. S. H. AND SEN, S. (2020). Heavy-tailed configuration models at criticality. *Ann. Inst. H. Poincaré Prob. Statist.* **56**, 1515–1558.
[21] DONDERWINKEL, S. AND XIE, Z. (2021). *Universality for the directed configuration model with random degrees: metric space convergence of the strongly connected components at criticality*. Preprint. Available at https://arxiv.org/abs/2105.11434.
[22] DUQUESNE, T. (2009). Continuum random trees and branching processes with immigration. *Stoch. Process. Appl.* **119**, 99–129.
[23] DUQUESNE, T. AND LE GALL, J.-F. (2005). Probabilistic and fractal aspects of Lévy trees. *Prob. Theory Relat. Fields* **131**, 553–603.
[24] DUQUESNE, T. AND LE GALL, J.-F. (2002). *Random Trees, Lévy Processes and Spatial Branching Processes* (Astérisque **281**). Société Mathématique de France, Paris.
[25] DUQUESNE, T. AND WINKEL, M. (2007). Growth of Lévy trees. *Prob. Theory Relat. Fields* **139**, 313–371.
[26] DUQUESNE, T. AND WINKEL, M. (2019). Hereditary tree growth and Lévy forests. *Stoch. Process. Appl.* **129**, 3690–3747.
[27] DURRETT, R. (2010). *Probability: Theory and Examples*. Cambridge University Press.
[28] DWASS, M. (1975). Branching processes in simple random walk. *Proc. Amer. Math. Soc.* **51**, 270–274.
[29] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley.
[30] GREY, D. R. (1974). Asymptotic behaviour of continuous time, continuous state-space branching processes. *J. Appl. Prob.* **11**, 669–677.
[31] HARRIS, T. E. (1948). Branching processes. *Ann. Math. Statist.* **19**, 474–494.
[32] HE, H. AND LUAN, N. (2013). A note on the scaling limits of contour functions of Galton–Watson trees. *Electron. Commun. Prob.* **18**, 13 pp.
[33] JACOD, J. AND SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd edn. Springer, Berlin.
[34] JANSON, S. (2009). On percolation in random graphs with given vertex degrees. *Electron. J. Prob.* **14**, 86–118.
[35] JANSON, S. (2009). The probability that a random multigraph is simple. *Combinatorics Prob. Comput.* **18**, 205–225.
[36] JANSON, S. (2012). Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. *Prob. Surveys* **9**, 103–252.
[37] JANSON, S. AND ŁUCZAK, M. J. (2009). A new approach to the giant component problem. *Random Structures Algorithms* **34**, 197–216.
[38] JOSEPH, A. (2014). The component sizes of a critical random graph with given degree sequence. *Ann. Appl. Prob.* **24**, 2560–2594.
[39] KALLENBERG, O. (2002). *Foundations of Modern Probability*. Springer, Cham.
[40] KAWAZU, K. AND WATANABE, S. (1971). Branching processes with immigration and related limit theorems. *Theory Prob. Appl.* **16**, 36–54.
[41] LAMBERT, A. (2002). The genealogy of continuous-state branching processes with immigration. *Prob. Theory Relat. Fields* **122**, 42–70.
[42] LE GALL, J.-F. (1991). Brownian excursions, trees and measure-valued branching processes. *Ann. Prob.* **19**, 1399–1439.
[43] LE GALL, J.-F. AND LE JAN, Y. (1998). Branching processes in Lévy processes: the exploration process. *Ann. Prob.* **26**, 213–252.

[44] LYONS, R. AND PERES, Y. (2017). *Probability on Trees and Networks*. Cambridge University Press.
[45] MILLAR, P. W. (1977). Zero–one laws and the minimum of a Markov process. *Trans. Amer. Math. Soc.* **226**, 365–391.
[46] MOLLOY, M. AND REED, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures Algorithms* **6**, 161–180.
[47] MOLLOY, M. AND REED, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combinatorics Prob. Comput.* **7**, 295–305.
[48] NEVEU, J. AND PITMAN, J. W. (1989). The branching process in a Brownian excursion. In *Séminaire de Probabilités XXIII*, eds J. AZÉMA, M. YOR AND P. A. MEYER, SPRINGER, BERLIN, Heidelberg, pp. 248–257.
[49] RIORDAN, O. (2012). The phase transition in the configuration model. *Combinatorics Prob. Comput.* **21**, 265–299.
[50] ROGERS, L. C. G. (1984). Brownian local times and branching processes. In *Séminaire de Probabilités XVIII 1982/83*, eds J. AZÉMA AND M. YOR, SPRINGER, BERLIN, Heidelberg, pp. 42–55.
[51] VAN DER HOFSTAD, R. (2017). *Random Graphs and Complex Networks*. Cambridge University Press.