

# Limit theorems for conditioned non-generic Galton–Watson trees

Igor Kortchemski

Laboratoire de Mathématiques, UMR 8628 CNRS, Université Paris-Sud, 91405 Orsay Cedex, France. E-mail: [igor.kortchemski@normalesup.org](mailto:igor.kortchemski@normalesup.org)

Received 13 March 2013; revised 30 August 2013; accepted 13 September 2013

---

**Abstract.** We study a particular type of subcritical Galton–Watson trees, which are called non-generic trees in the physics community. In contrast with the critical or supercritical case, it is known that condensation appears in certain large conditioned non-generic trees, meaning that with high probability there exists a unique vertex with macroscopic degree comparable to the total size of the tree. Using recent results concerning subexponential distributions, we investigate this phenomenon by studying scaling limits of such trees and show that the situation is completely different from the critical case. In particular, the height of such trees grows logarithmically in their size. We also study fluctuations around the condensation vertex.

**Résumé.** Nous étudions une classe particulière d'arbres de Galton–Watson sous-critiques, appelés arbres non-génériques en physique. Contrairement au cas critique ou surcritique, il est connu qu'une condensation apparaît dans certains grands arbres non-génériques conditionnés, c'est-à-dire qu'avec grande probabilité il existe un unique sommet de degré macroscopique comparable à la taille totale de l'arbre. En utilisant des résultats récents relatifs à des lois sousexponentielles, nous étudions ce phénomène en étudiant les limites d'échelles de tels arbres et montrons que la situation est complètement différente du cas critique. En particulier, la hauteur de ces arbres croît logarithmiquement en leur taille. Nous étudions aussi les fluctuations autour du sommet de condensation.

*MSC:* Primary 60J80; 60F17; secondary 05C80; 05C05

*Keywords:* Condensation; Subcritical Galton–Watson trees; Scaling limits; Subexponential distributions

---

## Introduction

The behavior of large Galton–Watson trees whose offspring distribution  $\mu = (\mu_i)_{i \geq 0}$  is *critical* (meaning that the mean of  $\mu$  is 1) and has *finite variance* has drawn a lot of attention. If  $t_n$  is a Galton–Watson tree with offspring distribution  $\mu$  (in short a  $\text{GW}_\mu$  tree) conditioned on having total size  $n$ , Kesten [22] proved that  $t_n$  converges locally in distribution as  $n \rightarrow \infty$  to the so-called critical Galton–Watson tree conditioned to survive. Aldous [1] studied the scaled asymptotic behavior of  $t_n$  by showing that the appropriately rescaled contour function of  $t_n$  converges to the Brownian excursion.

These results have been extended in different directions. The “finite second moment” condition on  $\mu$  has been relaxed by Duquesne [11], who showed that when  $\mu$  belongs to the domain of attraction of a stable law of index  $\theta \in (1, 2]$ , the appropriately rescaled contour function of  $t_n$  converges toward the normalized excursion of the  $\theta$ -stable height process, which codes the so-called  $\theta$ -stable tree (see also [24]). In a different direction, several authors have considered trees conditioned by other quantities than the total size, for example by the height [23,28] or the number of leaves [25,30].

### Non-critical Galton–Watson trees

Kennedy [21] noticed that, under certain conditions, the study of non-critical offspring distributions reduces to the study of critical ones. More precisely, if  $\lambda > 0$  is a fixed parameter such that  $Z_\lambda = \sum_{i \geq 0} \mu_i \lambda^i < \infty$ , set  $\mu_i^{(\lambda)} = \mu_i \lambda^i / Z_\lambda$  for  $i \geq 0$ . Then a  $\text{GW}_\mu$  tree conditioned on having total size  $n$  has the same distribution as a  $\text{GW}_{\mu^{(\lambda)}}$  tree conditioned on having total size  $n$ . Thus, if one can find  $\lambda > 0$  such that both  $Z_\lambda < \infty$  and  $\mu^{(\lambda)}$  is critical, then studying a conditioned non-critical Galton–Watson tree boils down to studying a critical one. This explains why the critical case has been extensively studied in the literature.

Let  $\mu$  be a probability distribution such that  $\mu_0 > 0$  and  $\mu_k > 0$  for some  $k \geq 2$ . We are interested in the case where there exist *no*  $\lambda > 0$  such that both  $Z_\lambda < \infty$  and  $\mu^{(\lambda)}$  is critical (see [17], Section 8, for a characterization of such probability distributions). An important example is when  $\mu$  is subcritical (i.e., of mean strictly less than 1) and  $\mu_i \sim c/i^\beta$  as  $i \rightarrow \infty$  for a fixed parameter  $\beta > 2$ . The study of such  $\text{GW}_\mu$  trees conditioned on having a large fixed size was initiated only recently by Jonsson and Stefánsson [20] who called such trees *non-generic* trees. They studied the above-mentioned case where  $\mu_i \sim c/i^\beta$  as  $i \rightarrow \infty$ , with  $\beta > 2$ , and showed that if  $t_n$  is a  $\text{GW}_\mu$  tree conditioned on having total size  $n$ , then with probability tending to 1 as  $n \rightarrow \infty$ , there exists a unique vertex of  $t_n$  with maximal degree, which is asymptotic to  $(1 - m)n$  where  $m < 1$  is the mean of  $\mu$ . This phenomenon is called *condensation* and appears in a variety of statistical mechanics models such as the Bose–Einstein condensation for bosons, the zero-range process [13,19] or the Backgammon model [4] (see Fig. 1).

Jonsson and Stefánsson [20] have also constructed an infinite random tree  $\widehat{\mathcal{T}}$  (with a unique vertex of infinite degree) such that  $t_n$  converges locally in distribution toward  $\widehat{\mathcal{T}}$  (meaning roughly that the degree of every vertex of  $t_n$  converges toward the degree of the corresponding vertex of  $\widehat{\mathcal{T}}$ ). See Section 2.3 below for the description of  $\widehat{\mathcal{T}}$ . In [17], Janson has extended this result to simply generated trees and has in particular given a very precise description of local properties of Galton–Watson trees conditioned on their size.

In this work, we are interested in the existence of scaling limits for the random trees  $t_n$ . When scaling limits exist, one often gets information concerning the global structure of the tree.

### Notation and assumptions

Throughout this work  $\theta > 1$  will be a fixed parameter. We say that a probability distribution  $(\mu_j)_{j \geq 0}$  on the nonnegative integers satisfies Assumption  $(H_\theta)$  if the following two conditions hold:

- (i)  $\mu$  is subcritical, meaning that  $0 < \sum_{j=0}^{\infty} j \mu_j < 1$ .
- (ii) There exists a measurable function  $\mathcal{L}: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $\mathcal{L}(x) > 0$  for  $x$  large enough and  $\lim_{x \rightarrow \infty} \mathcal{L}(tx) / \mathcal{L}(x) = 1$  for all  $t > 0$  (such a function is called slowly varying) and  $\mu_n = \mathcal{L}(n) / n^{1+\theta}$  for every  $n \geq 1$ .

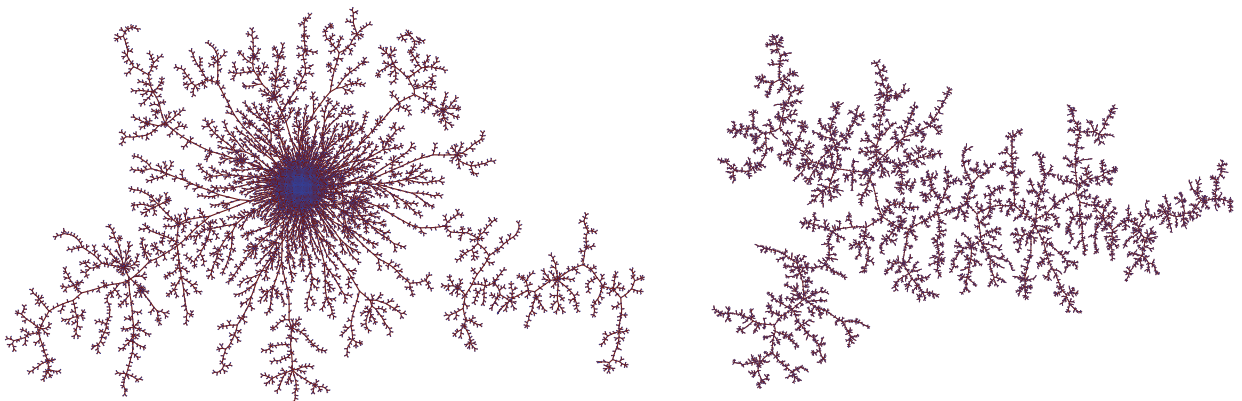


Fig. 1. The first figure shows a large non-generic Galton–Watson tree. The second figure shows a large critical Galton–Watson tree with finite variance.

Let  $|\tau|$  be the total progeny or size of a tree  $\tau$ . Condition (ii) implies that  $\mathbb{P}_\mu(|\tau| = n) > 0$  for sufficiently large  $n$ . Note that (ii) is more general than the analogous assumption in [17,20], where only the case  $\mathcal{L}(x) \rightarrow c$  as  $x \rightarrow \infty$  was studied in detail. Throughout this text,  $\theta > 1$  is a fixed parameter and  $\mu$  is a probability distribution on  $\mathbb{Z}_+$  satisfying the Assumption  $(H_\theta)$ . In addition, for every  $n \geq 1$  such that  $\mathbb{P}_\mu(|\tau| = n) > 0$ ,  $\mathfrak{t}_n$  is a  $\text{GW}_\mu$  tree conditioned on having  $n$  vertices (note that  $\mathfrak{t}_n$  is well defined for  $n$  sufficiently large). The mean of  $\mu$  will be denoted by  $\mathfrak{m}$  and we set  $\gamma = 1 - \mathfrak{m}$ .

We are now ready to state our main results which concern different aspects of non-generic trees. We are first interested in the condensation phenomenon and derive properties of the maximal degree. We then find the location of the vertex of maximal degree. Finally we investigate the global behavior of non-generic trees by studying their height.

### Condensation

If  $\tau$  is a (finite) tree, we denote by  $\Delta(\tau)$  the maximal out-degree of a vertex of  $\tau$  (the out-degree of a vertex is by definition its number of children). If  $\tau$  is a finite tree, let  $u_\star(\tau)$  be the smallest vertex (in the lexicographical order, see Definition 1.3 below) of  $\tau$  with maximal out-degree. The following result states that, with probability tending to 1 as  $n \rightarrow \infty$ , there exists a vertex of  $\mathfrak{t}_n$  with out-degree roughly  $\gamma n$  and that the deviations around this typical value are of order roughly  $O(n^{1/(2 \wedge \theta)})$ , and also that the out-degrees of all the other vertices of  $\mathfrak{t}_n$  are of order roughly  $O(n^{1/(2 \wedge \theta)})$ . In particular the vertex with maximal out-degree is unique with probability tending to 1 as  $n \rightarrow \infty$ .

**Theorem 1.** *There exists a slowly varying function  $L$  such that if  $B_n = L(n)n^{1/(2 \wedge \theta)}$ , the following assertions hold:*

- (i) *We have  $\frac{\Delta(\mathfrak{t}_n)}{\gamma n} \xrightarrow[n \rightarrow \infty]{(\mathbb{P})} 1$ .*
- (ii) *Let  $D_n$  be the maximal out-degree of vertices of  $\mathfrak{t}_n$  except  $u_\star(\mathfrak{t}_n)$ . If  $\theta \geq 2$ , then  $D_n/B_n$  converges in probability to 0 as  $n \rightarrow \infty$ . If  $\theta \in (1, 2)$ , then*

$$\mathbb{P}\left(\frac{D_n}{B_n} \leq u\right) \xrightarrow[n \rightarrow \infty]{} \exp\left(\frac{1}{\Gamma(1 - \theta)} u^{-\theta}\right), \quad u \geq 0,$$

where  $\Gamma$  is Euler’s Gamma function.

- (iii) *Let  $(Y_t)_{t \geq 0}$  be a spectrally positive Lévy process with Laplace exponent  $\mathbb{E}[\exp(-\lambda Y_t)] = \exp(t\lambda^{2 \wedge \theta})$ . Then*

$$\frac{\Delta(\mathfrak{t}_n) - \gamma n}{B_n} \xrightarrow[n \rightarrow \infty]{(d)} -Y_1.$$

When  $\mu$  has finite variance  $\sigma^2 \in (0, \infty)$ , one may take  $B_n = \sigma\sqrt{n/2}$ . Theorem 1 has already been proved when  $\mu_n \sim c/n^{1+\theta}$  as  $n \rightarrow \infty$  (i.e., when  $\mathcal{L} = c + o(1)$ , in which case one may choose  $L$  to be a constant function) by Jonsson and Stefánsson [20] for (i) and Janson [17] for (iii). However, our techniques are different and are based on a coding of  $\mathfrak{t}_n$  by a conditioned random walk combined with recent results of Armendáriz and Loulakis [2] concerning random walks whose jump distribution is subexponential, which imply that, informally, the tree  $\mathfrak{t}_n$  looks like a finite spine of geometric length decorated with independent  $\text{GW}_\mu$  trees, and on top of which are grafted  $\Delta(\mathfrak{t}_n)$  independent  $\text{GW}_\mu$  trees (see Proposition 2.6 and Corollary 2.7 below for precise statements). The main advantage of this approach is that it enables us to obtain new results concerning the structure of  $\mathfrak{t}_n$ .

### Localization of the vertex of maximal degree

We are also interested in the location of the vertex of maximal degree  $u_\star(\mathfrak{t}_n)$ . Before stating our results, we need to introduce some notation. If  $\tau$  is a tree, let  $U(\tau)$  be the index in the lexicographical order of the first vertex of  $\mathfrak{t}_n$  with maximal out-degree (when the vertices of  $\tau$  are ordered starting from index 0). Note that the number of children of  $u_\star(\tau)$  is  $\Delta(\tau)$ . Denote the generation of  $u_\star(\tau)$  by  $|u_\star(\tau)|$ .

**Theorem 2.** *The following three convergences hold:*

- (i) *For  $i \geq 0$ ,  $\mathbb{P}(U(\mathfrak{t}_n) = i) \xrightarrow[n \rightarrow \infty]{} \gamma \cdot \mathbb{P}_\mu(|\tau| \geq i + 1)$ .*

(ii) As  $n \rightarrow \infty$ ,  $|u_\star(t_n)|$  converges in distribution toward a geometric random variable of parameter  $1 - m$ , i.e.,

$$\mathbb{P}\left(|u_\star(t_n)| = i\right) \xrightarrow{n \rightarrow \infty} (1 - m)m^i, \quad i \geq 0.$$

See Section 2.3 below for the description of  $\widehat{\mathcal{T}}$ . Using different methods, a result similar to assertion (i) (as well as Propositions 2.2 and 2.4 below) has been proved by Durrett [12], Theorem 3.2, in the context of random walks when  $t_n$  is a  $\text{GW}_\mu$  tree conditioned on having at least  $n$  vertices and in addition  $\mu$  has finite variance. However, the so-called local conditioning by having a fixed number of vertices is often much more difficult to analyze (see, e.g., [11,28]). Note that  $\sum_{i \geq 1} \mathbb{P}_\mu(|\tau| \geq i) = \mathbb{E}_\mu[|\tau|] = 1/\gamma$ , so that the limit in (i) is a probability distribution. The proof of (i) combines the coding of  $t_n$  by a conditioned random walk with the previously mentioned results of Armendáriz and Loulakis. The proof of the second assertion uses (i) together with the local convergence of  $t_n$  toward the infinite random tree  $\widehat{\mathcal{T}}$ , which has been obtained by Jonsson and Stefánsson [20] in a particular case and then generalized by Janson [17], and was already mentioned above.

We are also interested in the sizes of the subtrees grafted on  $u_\star(t_n)$ . If  $\tau$  is a tree, for  $1 \leq j \leq \Delta(\tau)$ , let  $\xi_j(\tau)$  be the number of descendants of the  $j$ th child of  $u_\star(\tau)$  and set  $Z_j(\tau) = \xi_1(\tau) + \xi_2(\tau) + \dots + \xi_j(\tau)$ . If  $I$  is an interval, we let  $\mathbb{D}(I, \mathbb{R})$  denote the space of all right-continuous with left limits (càdlàg) functions  $I \rightarrow \mathbb{R}$ , endowed with the Skorokhod  $J_1$ -topology (see [5], Chapter 3, and [15], Chapter VI, for background concerning the Skorokhod topology). If  $x \in \mathbb{R}$ , let  $\lfloor x \rfloor$  denote the greatest integer smaller than or equal to  $x$ . Recall that  $(Y_t)_{t \geq 0}$  is the spectrally positive Lévy process with Laplace exponent  $\mathbb{E}[\exp(-\lambda Y_t)] = \exp(t\lambda^{2 \wedge \theta})$ . Recall the sequence  $(B_n)_{n \geq 1}$  from Theorem 1.

**Theorem 3.** *The following convergence holds in distribution in  $\mathbb{D}([0, 1], \mathbb{R})$ :*

$$\left( \frac{Z_{\lfloor \Delta(t_n)t \rfloor}(t_n) - \Delta(t_n)t/\gamma}{B_n}, 0 \leq t \leq 1 \right) \xrightarrow{n \rightarrow \infty} \left( \frac{1}{\gamma} Y_t, 0 \leq t \leq 1 \right).$$

Note that in the case when  $\mu$  has finite variance, we have  $\theta \geq 2$  and  $Y$  is just a constant times standard Brownian motion. Let us mention that Theorem 3 is used in [18] to study scaling limits of random planar maps with a unique large face (see [18], Proposition 3.1) and is also used in [8] to study the shape of large supercritical site-percolation clusters on random triangulations.

**Corollary 1.** *If  $\theta \geq 2$ ,  $\max_{1 \leq i \leq \Delta(t_n)} \xi_i(t_n)/B_n$  converges in probability toward 0 as  $n \rightarrow \infty$ . If  $\theta < 2$ , for every  $u > 0$  we have:*

$$\mathbb{P}\left(\frac{1}{B_n} \max_{1 \leq i \leq \Delta(t_n)} \xi_i(t_n) \leq u\right) \xrightarrow{n \rightarrow \infty} \exp\left(\frac{1}{\gamma^\theta \Gamma(1 - \theta)} u^{-\theta}\right).$$

The dichotomy between the cases  $\theta < 2$  and  $\theta \geq 2$  arises from the fact that  $Y$  is continuous if and only if  $\theta \geq 2$ .

### Height of non-generic trees

One of the main contributions of this work is to understand the growth of the height  $\mathcal{H}(t_n)$  of  $t_n$ , which is by definition the maximal generation in  $t_n$ . We establish the key fact that  $\mathcal{H}(t_n)$  grows logarithmically in  $n$ :

**Theorem 4.** *For every sequence  $(\lambda_n)_{n \geq 1}$  of positive real numbers tending to infinity:*

$$\mathbb{P}\left(\left|\mathcal{H}(t_n) - \frac{\ln(n)}{\ln(1/m)}\right| \leq \lambda_n\right) \xrightarrow{n \rightarrow \infty} 1.$$

Note that the situation is completely different from the critical case, where  $\mathcal{H}(t_n)$  grows like a power of  $n$ . Theorem 4 implies that  $\mathcal{H}(t_n)/\ln(n) \rightarrow \ln(1/m)$  in probability as  $n \rightarrow \infty$ , thus partially answering Problem 20.7 in [17]. Proposition 2.11 below also shows that this convergence holds in all the spaces  $\mathbb{L}_p$  for  $p \geq 1$ . Theorem 4 can be intuitively explained by the fact that the height of  $t_n$  should be close to the maximum of the height of  $\gamma n$  independent subcritical  $\text{GW}_\mu$  trees, which is indeed of order  $\ln(n)$ .

Since  $t_n$  grows roughly as  $\ln(n)$  as  $n \rightarrow \infty$ , it is natural to wonder if one could hope to obtain a scaling limit after rescaling the distances in  $t_n$  by  $\ln(n)$ . We show that the answer is negative and that we cannot hope to obtain a nontrivial scaling limit for  $t_n$  for the Gromov–Hausdorff topology, in sharp contrast with the critical case (see [11]). This partially answers a question of Janson [17], Problem 20.11.

**Theorem 5.** *The sequence  $(\ln(n)^{-1} \cdot t_n)_{n \geq 1}$  is not tight for the Gromov–Hausdorff topology, where  $\ln(n)^{-1} \cdot t_n$  stands for the metric space obtained from  $t_n$  by multiplying all distances by  $\ln(n)^{-1}$ .*

The Gromov–Hausdorff topology is the topology on compact metric spaces (up to isometries) defined by the Gromov–Hausdorff distance, and is often used in the study of scaling limits of different classes of random graphs (see [7], Chapter 7, for background concerning the Gromov–Hausdorff topology).

However, we establish the convergence of the finite-dimensional marginal distributions of the height function coding  $t_n$ . If  $\tau$  is a tree, for  $0 \leq i \leq |\tau| - 1$ , denote by  $H_i(\tau)$  the generation of the  $i$ th vertex of  $\tau$  in the lexicographical order.

**Theorem 6.** *Let  $k \geq 1$  be an integer and fix  $0 < t_1 < \dots < t_k < 1$ . Then*

$$(H_{\lfloor nt_1 \rfloor}(t_n), H_{\lfloor nt_2 \rfloor}(t_n), \dots, H_{\lfloor nt_k \rfloor}(t_n)) \xrightarrow[n \rightarrow \infty]{(d)} (1 + \mathbf{e}_0 + \mathbf{e}_1, 1 + \mathbf{e}_0 + \mathbf{e}_2, \dots, 1 + \mathbf{e}_0 + \mathbf{e}_k),$$

where  $(\mathbf{e}_i)_{i \geq 0}$  is a sequence of i.i.d. geometric random variables of parameter  $1 - m$  (i.e.,  $\mathbb{P}(\mathbf{e}_0 = i) = (1 - m)m^i$  for  $i \geq 0$ ).

Informally, the random variable  $\mathbf{e}_0$  describes the length of the spine, and the random variables  $(\mathbf{e}_1, \dots, \mathbf{e}_k)$  describe the height of vertices chosen in a forest of independent subcritical  $\text{GW}_\mu$  trees. Note that these finite-dimensional marginal distributions converge without scaling, even though the height of  $t_n$  is of order  $\ln(n)$ .

This text is organized as follows. We first recall the definition and basic properties of Galton–Watson trees. In Section 2, we establish limit theorems for large conditioned non-generic Galton–Watson trees. We conclude by giving possible extensions and formulating some open problems.

## 1. Galton–Watson trees

### 1.1. Basic definitions

We briefly recall the formalism of plane trees (also known in the literature as rooted ordered trees) which can be found in [27] for example.

**Definition 1.1.** *Let  $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$  be the set of all nonnegative integers and let  $\mathbb{N}$  be the set of all positive integers. Let also  $U$  be the set of all labels defined by:*

$$U = \bigcup_{n=0}^{\infty} (\mathbb{N})^n,$$

where by convention  $(\mathbb{N})^0 = \{\emptyset\}$ . An element of  $U$  is a sequence  $u = u_1 \cdots u_k$  of positive integers and we set  $|u| = k$ , which represents the “generation” of  $u$ . If  $u = u_1 \cdots u_i$  and  $v = v_1 \cdots v_j$  belong to  $U$ , we write  $uv = u_1 \cdots u_i v_1 \cdots v_j$  for the concatenation of  $u$  and  $v$ . In particular, we have  $u\emptyset = \emptyset u = u$ . Finally, a plane tree  $\tau$  is a finite or infinite subset of  $U$  such that:

- (i)  $\emptyset \in \tau$ ,
- (ii) if  $v \in \tau$  and  $v = ui$  for some  $i \in \mathbb{N}$ , then  $u \in \tau$ ,
- (iii) for every  $u \in \tau$ , there exists  $k_u(\tau) \in \{0, 1, 2, \dots\} \cup \{\infty\}$  (the number of children of  $u$ ) such that, for every  $j \in \mathbb{N}$ ,  $uj \in \tau$  if and only if  $1 \leq j \leq k_u(\tau)$ .

Note that in contrast with [26,27] we allow the possibility  $k_u(\tau) = \infty$  in (iii). In the following, by *tree* we will always mean plane tree, and we denote the set of all trees by  $\mathbb{T}$  and the set of all finite trees by  $\mathbb{T}_f$ . We will often view each vertex of a tree  $\tau$  as an individual of a population whose  $\tau$  is the genealogical tree. The total progeny or size of  $\tau$  will be denoted by  $|\tau| = \text{Card}(\tau)$ . If  $\tau$  is a tree and  $u \in \tau$ , we define the shift of  $\tau$  at  $u$  by  $T_u\tau = \{v \in U; uv \in \tau\}$ , which is itself a tree.

**Definition 1.2.** Let  $\rho$  be a probability measure on  $\mathbb{Z}_+$ . The law of the Galton–Watson tree with offspring distribution  $\rho$  is the unique probability measure  $\mathbb{P}_\rho$  on  $\mathbb{T}$  such that:

- (i)  $\mathbb{P}_\rho(k_\emptyset = j) = \rho(j)$  for  $j \geq 0$ ,
- (ii) for every  $j \geq 1$  with  $\rho(j) > 0$ , conditionally on  $\{k_\emptyset = j\}$ , the subtrees  $T_1\tau, \dots, T_j\tau$  are independent and identically distributed with distribution  $\mathbb{P}_\rho$ .

A random tree whose distribution is  $\mathbb{P}_\rho$  will be called a Galton–Watson tree with offspring distribution  $\rho$ , or in short a  $\text{GW}_\rho$  tree.

In the sequel, for every integer  $j \geq 1$ ,  $\mathbb{P}_{\rho,j}$  will denote the probability measure on  $\mathbb{T}^j$  which is the distribution of  $j$  independent  $\text{GW}_\rho$  trees. The canonical element of  $\mathbb{T}^j$  is denoted by  $\mathbf{f}$ . For  $\mathbf{f} = (\tau_1, \dots, \tau_j) \in \mathbb{T}^j$ , let  $|\mathbf{f}| = |\tau_1| + \dots + |\tau_j|$  be the total progeny of  $\mathbf{f}$ .

### 1.2. Coding Galton–Watson trees

We now explain how trees can be coded by three different functions. These codings are important in the understanding of large Galton–Watson trees (see Fig. 2).

**Definition 1.3.** We write  $u < v$  for the lexicographical order on the labels  $U$  (for example  $\emptyset < 1 < 21 < 22$ ). Let  $\tau$  be a finite tree and order the individuals of  $\tau$  in lexicographical order:  $\emptyset = u(0) < u(1) < \dots < u(|\tau| - 1)$ . The height process  $H(\tau) = (H_n(\tau), 0 \leq n \leq |\tau|)$  is defined, for  $0 \leq n < |\tau|$ , by:

$$H_n(\tau) = |u(n)|.$$

We set  $H_{|\tau|}(\tau) = 0$  for technical reasons. The height  $\mathcal{H}(\tau)$  of  $\tau$  is by definition  $\max_{0 \leq n < |\tau|} H_n(\tau)$ .

Consider a particle that starts from the root and visits continuously all the edges of  $\tau$  at unit speed, assuming that every edge has unit length. When the particle leaves a vertex, it moves toward the first non visited child of this vertex if there is such a child, or returns to the parent of this vertex. Since all the edges are crossed twice, the total time needed to explore the tree is  $2(|\tau| - 1)$ . For  $0 \leq t \leq 2(|\tau| - 1)$ ,  $C_\tau(t)$  is defined as the distance to the root of the position of

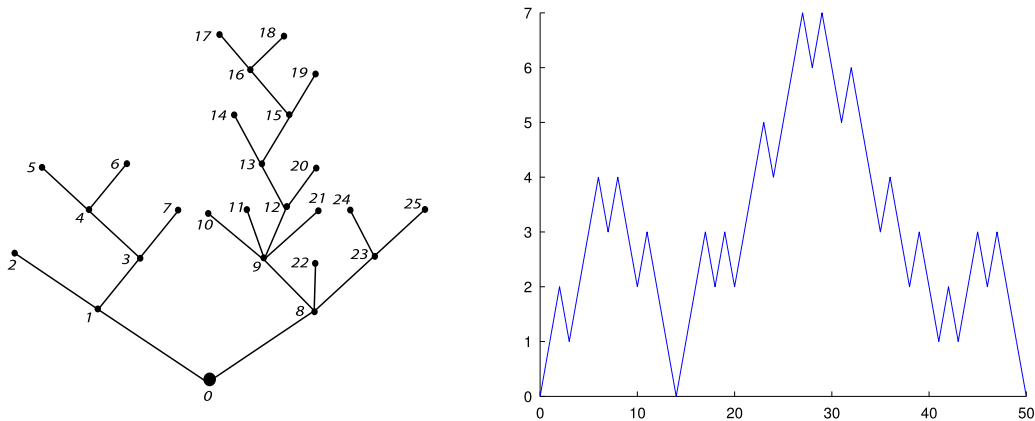


Fig. 2. A tree  $\tau$  with its vertices indexed in lexicographical order and its contour function ( $C_u(\tau); 0 \leq u \leq 2(|\tau| - 1)$ ). Here,  $|\tau| = 26$ .

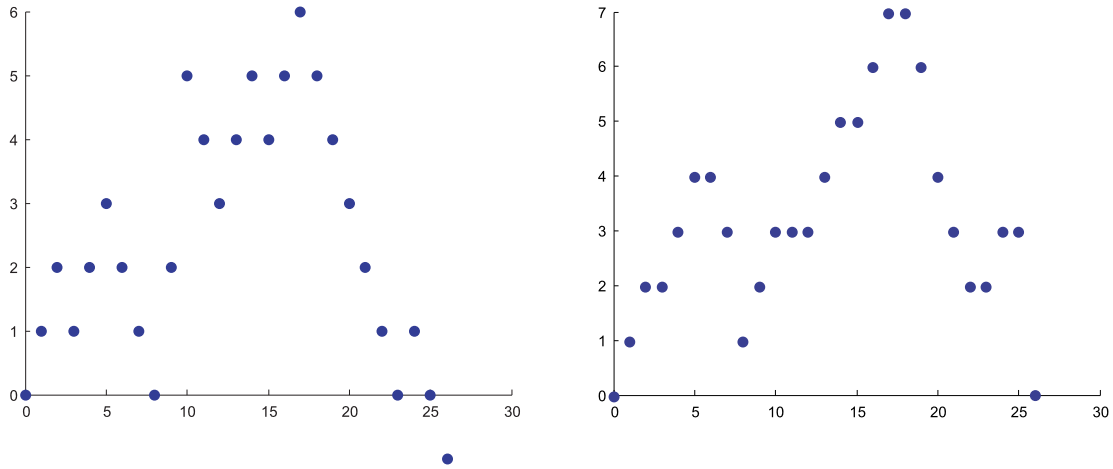


Fig. 3. The Lukasiewicz path  $\mathcal{W}(\tau)$  and the height function  $H(\tau)$  of  $\tau$ .

the particle at time  $t$ . For technical reasons, we set  $C_\tau(t) = 0$  for  $t \in [2(|\tau| - 1), 2|\tau|]$ . The function  $C(\tau)$  is called the contour function of the tree  $\tau$ . See Fig. 3 for an example, and [11], Section 2, for a rigorous definition.

Finally, the Lukasiewicz path  $\mathcal{W}(\tau) = (\mathcal{W}_n(\tau), 0 \leq n \leq |\tau|)$  of  $\tau$  is defined by  $\mathcal{W}_0(\tau) = 0$  and for  $0 \leq n \leq |\tau| - 1$ :

$$\mathcal{W}_{n+1}(\tau) = \mathcal{W}_n(\tau) + k_{u(n)}(\tau) - 1.$$

Note that necessarily  $\mathcal{W}_{|\tau|}(\tau) = -1$  and that  $U(t_n) = \min\{j \geq 0; \mathcal{W}_{j+1}(t_n) - \mathcal{W}_j(t_n) = \Delta(t_n) - 1\}$ , where we recall that  $U(t_n)$  is the index in the lexicographical order of the first vertex of  $t_n$  with maximal out-degree.

The following proposition explains the importance of the Lukasiewicz path. Let  $\rho$  be a critical or subcritical probability distribution on  $\mathbb{N}$  with  $\rho(1) < 1$ .

**Proposition 1.4.** *Let  $(W_n)_{n \geq 0}$  be a random walk with starting point  $W_0 = 0$  and jump distribution  $v(k) = \rho(k + 1)$  for  $k \geq -1$ . Set  $\zeta = \inf\{n \geq 0; W_n = -1\}$ . Then  $(W_0, W_1, \dots, W_\zeta)$  has the same distribution as the Lukasiewicz path of a  $\text{GW}_\rho$  tree. In particular, the total progeny of a  $\text{GW}_\rho$  tree has the same law as  $\zeta$ .*

**Proof.** See [26], Proposition 1.5. □

We next extend the definition of the Lukasiewicz path to a forest. If  $\mathbf{f} = (\tau_i)_{1 \leq i \leq j}$  is a forest, set  $n_0 = 0$  and  $n_p = |\tau_1| + |\tau_2| + \dots + |\tau_p|$  for  $1 \leq p \leq j$ . Then, for every  $0 \leq i \leq p - 1$  and  $0 \leq k < |\tau_{i+1}|$ , set

$$\mathcal{W}_{n_i+k}(\mathbf{f}) = \mathcal{W}_k(\tau_{i+1}) - i.$$

Note that  $(\mathcal{W}_{n_i+k}(\mathbf{f}) + i; 0 \leq k \leq |\tau_{i+1}|)$  is the Lukasiewicz path of  $\tau_{i+1}$  and that  $\min\{0 \leq i \leq n_j; \mathcal{W}_i(\mathbf{f}) = -k\} = n_k$  for  $1 \leq k \leq j$ .

Finally, the following result will be useful.

**Proposition 1.5.** *Let  $(W_n)_{n \geq 0}$  be the random walk introduced in Proposition 1.4 with  $\rho = \mu$ . Then*

- (i)  $\mathbb{P}(\forall i \geq 1, W_i \leq -1) = \gamma$ .
- (ii) For every  $i \geq 0$ ,  $\mathbb{P}(\forall m \leq i, W_m \geq 0) = \mathbb{P}_\mu(|\tau| \geq i + 1)$ .

**Proof.** By [31], Theorem 1 in Chapter 2, we have  $\mathbb{P}(\forall i \geq 1, W_i \leq -1) = -\mathbb{E}[W_1] = 1 - m$ . The second assertion is an immediate consequence of Proposition 1.4. □

### 1.3. The Vervaat transformation

For  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{Z}^n$  and  $i \in \mathbb{Z}/n\mathbb{Z}$ , denote by  $\mathbf{x}^{(i)}$  the  $i$ th cyclic shift of  $\mathbf{x}$  defined by  $x_k^{(i)} = x_{i+k \bmod n}$  for  $1 \leq k \leq n$ .

**Definition 1.6.** Let  $n \geq 1$  be an integer and let  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{Z}^n$ . Set  $w_j = x_1 + \dots + x_j$  for  $1 \leq j \leq n$  and let the integer  $i_*(\mathbf{x})$  be defined by  $i_*(\mathbf{x}) = \inf\{j \geq 1; w_j = \min_{1 \leq i \leq n} w_i\}$ . The Vervaat transform of  $\mathbf{x}$ , denoted by  $\mathbf{V}(\mathbf{x})$ , is defined to be  $\mathbf{x}^{(i_*(\mathbf{x}))}$ .

The following fact is well known (see, e.g., [29], Section 5):

**Proposition 1.7.** Let  $(W_n, n \geq 0)$  be as in Proposition 1.4 and  $X_k = W_k - W_{k-1}$  for  $k \geq 1$ . Fix an integer  $n \geq 1$  such that  $\mathbb{P}(W_n = -1) > 0$ . The law of  $\mathbf{V}(X_1, \dots, X_n)$  under  $\mathbb{P}(\cdot | W_n = -1)$  coincides with the law of  $(X_1, \dots, X_n)$  under  $\mathbb{P}(\cdot | \zeta = n)$ .

From Proposition 1.4, it follows that the law of  $\mathbf{V}(X_1, \dots, X_n)$  under  $\mathbb{P}(\cdot | W_n = -1)$  coincides with the law of  $(\mathcal{W}_1(t_n), \mathcal{W}_2(t_n) - \mathcal{W}_1(t_n), \dots, \mathcal{W}_n(t_n) - \mathcal{W}_{n-1}(t_n))$  where  $t_n$  is a  $\text{GW}_\rho$  tree conditioned on having total progeny equal to  $n$ .

We now introduce the Vervaat transformation in continuous time.

**Definition 1.8.** Set  $\mathbb{D}_0([0, 1], \mathbb{R}) = \{\omega \in \mathbb{D}([0, 1], \mathbb{R}); \omega(0) = 0\}$ . The Vervaat transformation in continuous time, denoted by  $\mathcal{V} : \mathbb{D}_0([0, 1], \mathbb{R}) \rightarrow \mathbb{D}([0, 1], \mathbb{R})$ , is defined as follows. For  $\omega \in \mathbb{D}_0([0, 1], \mathbb{R})$ , set  $g_1(\omega) = \inf\{t \in [0, 1]; \omega(t-) \wedge \omega(t) = \inf_{[0,1]} \omega\}$ . Then define:

$$\mathcal{V}(\omega)(t) = \begin{cases} \omega(g_1(\omega) + t) - \inf_{[0,1]} \omega, & \text{if } g_1(\omega) + t \leq 1, \\ \omega(g_1(\omega) + t - 1) + \omega(1) - \inf_{[0,1]} \omega, & \text{if } g_1(\omega) + t \geq 1. \end{cases}$$

By combining the Vervaat transformation with limit theorems under the conditional probability distribution  $\mathbb{P}(\cdot | W_n = -1)$  and using Proposition 1.4 we will obtain information about conditioned Galton–Watson trees. The advantage of dealing with  $\mathbb{P}(\cdot | W_n = -1)$  is to avoid any positivity constraint.

### 1.4. Slowly varying functions

Recall that a measurable function  $L : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is said to be slowly varying if  $L(x) > 0$  for  $x$  large enough and  $\lim_{x \rightarrow \infty} L(tx)/L(x) = 1$  for all  $t > 0$ . Let  $L : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a slowly varying function. Without further notice, we will use the following standard facts:

- (i) The convergence  $\lim_{x \rightarrow \infty} L(tx)/L(x) = 1$  holds uniformly for  $t$  in a compact subset of  $(0, \infty)$ .
- (ii) Fix  $\varepsilon > 0$ . There exists a constant  $C > 1$  such that  $\frac{1}{C}x^{-\varepsilon} \leq L(nx)/L(n) \leq Cx^\varepsilon$  for every integer  $n$  sufficiently large and  $x \geq 1$ .

These results are immediate consequences of the so-called representation theorem for slowly varying functions (see, e.g., [6], Theorem 1.3.1).

## 2. Limit theorems for conditioned non-generic Galton–Watson trees

In the sequel,  $(W_n; n \geq 0)$  denotes the random walk introduced in Proposition 1.4 with  $\rho = \mu$ . Note that  $\mathbb{E}[W_1] = -\gamma < 0$ . Set  $X_0 = 0$  and  $X_k = W_k - W_{k-1}$  for  $k \geq 1$ . It will be convenient to work with centered random walks, so we also set  $\bar{W}_n = W_n + \gamma n$  and  $\bar{X}_n = X_n + \gamma$  for  $n \geq 0$  so that  $\bar{W}_n = \bar{X}_1 + \dots + \bar{X}_n$ . Obviously,  $W_n = -1$  if and only if  $\bar{W}_n = \gamma n - 1$ .



### 2.1. Invariance principles for conditioned random walks

In this section, our goal is to prove Theorem 1. We first introduce some notation. Denote by  $T : \bigcup_{n \geq 1} \mathbb{R}^n \rightarrow \bigcup_{n \geq 1} \mathbb{R}^n$  the operator that interchanges the last and the (first) maximal component of a finite sequence of real numbers:

$$T(x_1, \dots, x_n)_k = \begin{cases} \max_{1 \leq i \leq n} x_i, & \text{if } k = n, \\ x_n, & \text{if } x_k > \max_{1 \leq i < k} x_i \text{ and } x_k = \max_{k \leq i \leq n} x_i, \\ x_k, & \text{otherwise.} \end{cases}$$

Since  $\mu$  satisfies Assumption  $(H_\theta)$ , we have  $\mathbb{P}(\overline{W}_1 \in (x, x + 1]) \sim \mathcal{L}(x)/x^{1+\theta}$  as  $x \rightarrow \infty$ . Then, by [10], Theorem 9.1, we have:

$$\mathbb{P}(\overline{W}_n \in (x, x + 1]) \underset{n \rightarrow \infty}{\sim} n \cdot \mathbb{P}(\overline{W}_1 \in (x, x + 1]), \quad (1)$$

uniformly in  $x \geq \varepsilon n$  for every fixed  $\varepsilon > 0$ . In other words, the distribution of  $\overline{W}_1$  is  $(0, 1]$ -subexponential, so that we can apply a recent result of Armendáriz and Loulakis [2] concerning conditioned random walks with subexponential jump distribution. In our particular case, this result can be stated as follows:

**Theorem 2.1 (Armendáriz and Loulakis, Theorem 1 in [2]).** *For  $n \geq 1$  and  $x > 0$ , let  $\mu_{n,x}$  be the probability measure on  $\mathbb{R}^n$  which is the distribution of  $(\overline{X}_1, \dots, \overline{X}_n)$  under the conditional probability distribution  $\mathbb{P}(\cdot | \overline{W}_n \in (x, x + 1])$ .*

*Then for every  $\varepsilon > 0$ , we have:*

$$\lim_{n \rightarrow \infty} \sup_{x \geq \varepsilon n} \sup_{A \in \mathcal{B}(\mathbb{R}^{n-1})} |\mu_{n,x} \circ T^{-1}[A \times \mathbb{R}] - \mu^{\otimes(n-1)}[A]| = 0.$$

As explained in [2], this means that under  $\mathbb{P}(\cdot | \overline{W}_n \in (x, x + 1])$ , asymptotically one gets  $n - 1$  independent random variables after forgetting the largest jump.

The proof of Theorem 1 is based on the following invariance principle concerning a conditioned random walk with negative drift, which is a simple consequence of Theorem 2.1.

**Proposition 2.2.** *Let  $R$  be a uniformly distributed random variable on  $[0, 1]$ . Then the following convergence holds in  $\mathbb{D}([0, 1], \mathbb{R})$ :*

$$\left( \frac{W_{\lfloor nt \rfloor}}{n}, 0 \leq t \leq 1 \mid W_n = -1 \right) \xrightarrow[n \rightarrow \infty]{(d)} (-\gamma t + \gamma \mathbb{1}_{R \leq t}, 0 \leq t \leq 1). \quad (2)$$

**Proof.** By the definition of  $\overline{W}$ , it is sufficient to check that the following convergence holds in  $\mathbb{D}([0, 1], \mathbb{R})$ :

$$\left( \frac{\overline{W}_{\lfloor nt \rfloor}}{n}, 0 \leq t \leq 1 \mid \overline{W}_n = \gamma n - 1 \right) \xrightarrow[n \rightarrow \infty]{(d)} (\gamma \mathbb{1}_{R \leq t}, 0 \leq t \leq 1), \quad (3)$$

where  $R$  is a uniformly distributed random variable on  $[0, 1]$ . Denote by  $V_n$  the coordinate of the first maximal component of  $(\overline{X}_1, \dots, \overline{X}_n)$ . Set  $\widetilde{W}_0 = 0$  and for  $1 \leq i \leq n - 1$  set:

$$\widetilde{W}_i = \begin{cases} \overline{X}_1 + \overline{X}_2 + \dots + \overline{X}_i, & \text{if } i < V_n, \\ \overline{X}_1 + \overline{X}_2 + \dots + \overline{X}_{V_n-1} + \overline{X}_{V_n+1} + \dots + \overline{X}_{i+1}, & \text{otherwise.} \end{cases}$$

By Theorem 2.1, for every  $\varepsilon > 0$ :

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \forall t \in [0, 1], \left| \frac{\widetilde{W}_{\lfloor (n-1)t \rfloor}}{n-1} \right| \leq \varepsilon \mid \overline{W}_n = \gamma n - 1 \right) - \mathbb{P} \left( \forall t \in [0, 1], \left| \frac{\overline{W}_{\lfloor (n-1)t \rfloor}}{n-1} \right| \leq \varepsilon \right) = 0. \quad (4)$$

Next, by the functional strong law of large numbers,

$$\mathbb{P}\left(\forall t \in [0, 1], \left| \frac{\widetilde{W}_{\lfloor (n-1)t \rfloor}}{n-1} \right| \leq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 1. \tag{5}$$

Combining (5) with (4), we get the following convergence in  $\mathbb{D}([0, 1], \mathbb{R})$ :

$$\left( \frac{\widetilde{W}_{\lfloor (n-1)t \rfloor}}{n-1}; 0 \leq t \leq 1 \mid \overline{W}_n = \gamma n - 1 \right) \xrightarrow[n \rightarrow \infty]{(\mathbb{P})} \mathbf{0}, \tag{6}$$

where  $\mathbf{0}$  stands for the constant function equal to 0 on  $[0, 1]$ . In addition, note that on the event  $\{\overline{W}_n = \gamma n - 1\}$ , we have  $\overline{X}_{V_n} = \gamma n - 1 - \widetilde{W}_{\lfloor n-1 \rfloor}$ . The following joint convergence in distribution thus holds in  $\mathbb{D}([0, 1], \mathbb{R}) \times \mathbb{R}$ :

$$\left( \left( \frac{\widetilde{W}_{\lfloor (n-1)t \rfloor}}{n-1}; 0 \leq t \leq 1 \right), \frac{\overline{X}_{V_n}}{n} \mid \overline{W}_n = \gamma n - 1 \right) \xrightarrow[n \rightarrow \infty]{(\mathbb{P})} (\mathbf{0}, \gamma). \tag{7}$$

Standard properties of the Skorokhod topology then show that the following convergence holds in  $\mathbb{D}([0, 1], \mathbb{R})$ :

$$\left( \frac{\overline{W}_{\lfloor nt \rfloor}}{n} - \frac{\overline{X}_{V_n}}{n} \mathbb{1}_{\{t \geq V_n/n\}}; 0 \leq t \leq 1 \mid \overline{W}_n = \gamma n - 1 \right) \xrightarrow[n \rightarrow \infty]{(\mathbb{P})} \mathbf{0}. \tag{8}$$

Next, note that the convergence (7) implies that under  $\mathbb{P}(\cdot \mid \overline{W}_n = \gamma n - 1)$ ,  $(\overline{X}_1, \dots, \overline{X}_n)$  has a unique maximal component with probability tending to one as  $n \rightarrow \infty$ . Since the distribution of  $(\overline{X}_1, \dots, \overline{X}_n)$  under  $\mathbb{P}(\cdot \mid \overline{W}_n = \gamma n - 1)$  is cyclically exchangeable, one easily gets that the law of  $V_n/n$  under  $\mathbb{P}(\cdot \mid \overline{W}_n = \gamma n - 1)$  converges to the uniform distribution on  $[0, 1]$ . Also from (7) we know that  $\overline{X}_{V_n}/n$  under  $\mathbb{P}(\cdot \mid \overline{W}_n = \gamma n - 1)$  converges in probability to  $\gamma$ . It follows that

$$\left( \frac{\overline{X}_{V_n}}{n} \mathbb{1}_{\{V_n/n \leq t\}}, 0 \leq t \leq 1 \mid \overline{W}_n = \gamma n - 1 \right) \xrightarrow[n \rightarrow \infty]{(d)} (\gamma \mathbb{1}_{R \leq t}, 0 \leq t \leq 1), \tag{9}$$

where  $R$  is uniformly distributed over  $[0, 1]$ . Since (8) holds in probability, we can combine (8) and (9) to get (3). This completes the proof.  $\square$

Before proving Theorem 1, we need to introduce some notation. For  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{Z}^n$ , set  $\mathcal{M}(\mathbf{x}) = \max_{1 \leq i \leq n} x_i$ . Recall the notation  $\mathbf{V}(\mathbf{x})$  for the Vervaat transform of  $\mathbf{x}$ . Note that  $\mathcal{M}(\mathbf{x}) = \mathcal{M}(\mathbf{V}(\mathbf{x}))$ . Let  $F : \mathbb{R} \rightarrow \mathbb{R}$  be a bounded continuous function. Recall that  $\Delta(t_n)$  denotes the maximal out-degree of a vertex of  $t_n$ . Since the maximal jump of  $\mathcal{W}(t_n)$  is equal to  $\Delta(t_n) - 1$ , it follows from the remark following Proposition 1.7 that:

$$\begin{aligned} \mathbb{E}[F(\Delta(t_n))] &= \mathbb{E}[F(\mathcal{M}(\mathbf{V}(X_1, X_2, \dots, X_n)) + 1) \mid W_n = -1] \\ &= \mathbb{E}[F(\mathcal{M}(X_1, X_2, \dots, X_n) + 1) \mid W_n = -1]. \end{aligned} \tag{10}$$

Recall that since  $\mu$  satisfies Assumption  $(H_\theta)$ ,  $\overline{W}_1$  belongs to the domain of attraction of a spectrally positive strictly stable law of index  $2 \wedge \theta$ . Hence there exists a slowly varying function  $L$  such that  $\overline{W}_n / (L(n)n^{1/(2 \wedge \theta)})$  converges in distribution toward  $Y_1$ . We set  $B_n = L(n)n^{1/(2 \wedge \theta)}$  and prove that Theorem 1 holds with this choice of  $B_n$ . The function  $L$  is not unique, but if  $\tilde{L}$  is another slowly function with the same property we have  $L(n)/\tilde{L}(n) \rightarrow 1$  as  $n \rightarrow \infty$ . So our results do not depend on the choice of  $L$ . Note that when  $\mu$  has finite variance  $\sigma^2$ , one may take  $B_n = \sigma \sqrt{n/2}$ , and when  $\mathcal{L} = c + o(1)$  one may choose  $L$  to be a constant function.

We are now ready to prove Theorem 1.

**Proof of Theorem 1.** If  $Z \in \mathbb{D}([0, 1], \mathbb{R})$ , denote by  $\overline{\Delta}(Z) = \sup_{0 < s < 1} (Z_s - Z_{s-})$  the largest jump of  $Z$ . Since  $\overline{\Delta} : \mathbb{D}([0, 1], \mathbb{R}) \rightarrow \mathbb{R}$  is continuous, from Proposition 2.2 we get that, under the conditioned probability measure  $\mathbb{P}(\cdot \mid W_n = -1)$ ,  $\mathcal{M}(X_1, X_2, \dots, X_n)/n$  converges in probability toward  $\gamma$  as  $n \rightarrow \infty$ . Assertion (i) immediately follows from (10).

For the second assertion, keeping the notation of the proof of Proposition 2.2, we get by Theorem 2.1 that for every bounded continuous function  $F : \mathbb{D}([0, 1], \mathbb{R}) \rightarrow \mathbb{R}$

$$\lim_{n \rightarrow \infty} \left| \mathbb{E} \left[ F \left( \frac{\tilde{W}_{\lfloor (n-1)t \rfloor}}{B_n}; 0 \leq t \leq 1 \right) \middle| \bar{W}_n = \gamma n - 1 \right] - \mathbb{E} [F(Y_t, 0 \leq t \leq 1)] \right| = 0.$$

Since the jumps of  $(\tilde{W}_{\lfloor (n-1)t \rfloor}; 0 \leq t \leq 1)$  have the same distribution as the out-degrees, minus one, of all the vertices of  $\mathfrak{t}_n$ , except  $u_*(\mathfrak{t}_n)$ , and by continuity of the map  $Z \mapsto \bar{\Delta}(Z)$  on  $\mathbb{D}([0, 1], \mathbb{R})$ , it follows that

$$\frac{D_n}{B_n} \xrightarrow[n \rightarrow \infty]{(d)} \bar{\Delta}(Y_t, 0 \leq t \leq 1).$$

If  $\theta \geq 2$ ,  $Y$  is continuous and  $\bar{\Delta}(Y_t, 0 \leq t \leq 1) = 0$ . If  $\theta < 2$ , the result easily follows from the fact that the Lévy measure of  $Y$  is  $\nu(dx) = \mathbb{1}_{\{x>0\}} dx / (\Gamma(-\theta)x^{1+\theta})$ .

For (iii), if  $V_n$  is as in the proof of Proposition 2.2, note that we have

$$\mathcal{M}(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n) = \bar{X}_{V_n} = \gamma n - 1 - \sum_{i \neq V_n} \bar{X}_i = \gamma n - 1 - \tilde{W}_{n-1}$$

on the event  $\{\bar{W}_n = \gamma n - 1\}$ . As noted in [2], formula (2.7), it follows from (6) that

$$\left( \frac{\mathcal{M}(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n) - \gamma n}{B_n} \middle| \bar{W}_n = \gamma n - 1 \right) \xrightarrow[n \rightarrow \infty]{(d)} -Y_1.$$

Since  $\mathcal{M}(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n) = \mathcal{M}(X_1, X_2, \dots, X_n) + \gamma$ , we thus get that

$$\left( \frac{\mathcal{M}(X_1, X_2, \dots, X_n) + 1 - \gamma n}{B_n} \middle| W_n = -1 \right) \xrightarrow[n \rightarrow \infty]{(d)} -Y_1.$$

Assertion (iii) then immediately follows from (10). This completes the proof. □

**Remark 2.3.** *The preceding proof shows that assertion (i) in Theorem 1 remain true when  $\mu$  is subcritical and both (1) and Theorem 2.1 hold. These conditions are more general than those of Assumption  $(H_\theta)$ : see, e.g., [10], Section 9, for examples of probability distributions that do not satisfy Assumption  $(H_\theta)$  but such that (1) holds. Note that assertion (ii) in Theorem 1 relies on the fact that  $\mu$  belongs to the domain of attraction of a stable law. Note also that there exist subcritical probability distributions such that none of the assertions of Theorem 1 hold (see [17], Example 19.37, for an example).*

By applying the Vervaat transformation in continuous time to the convergence of Proposition 2.2, standard properties of the Skorokhod topology imply the following invariance principle for the Lukasiewicz path coding  $\mathfrak{t}_n$  (we leave details to the reader since we will not need this result later). See Fig. 4 for a simulation.

**Proposition 2.4.** *The following assertions hold.*

(i) *We have:*

$$\sup_{0 \leq i \leq U(\mathfrak{t}_n)} \frac{W_i(\mathfrak{t}_n)}{n} \xrightarrow[n \rightarrow \infty]{(\mathbb{P})} 0.$$

(ii) *The following convergence holds in distribution in  $\mathbb{D}([0, 1], \mathbb{R})$ :*

$$\left( \frac{W_{\lfloor nt \rfloor \vee (U(\mathfrak{t}_n)+1)}(\mathfrak{t}_n)}{n}, 0 \leq t \leq 1 \right) \xrightarrow[n \rightarrow \infty]{(d)} (\gamma(1-t), 0 \leq t \leq 1).$$

Property (i) shows that  $(W_{\lfloor nt \rfloor}(\mathfrak{t}_n)/n, 0 \leq t \leq 1)$  does not converge in distribution in  $\mathbb{D}([0, 1], \mathbb{R})$  toward  $(\gamma(1-t), 0 \leq t \leq 1)$  and this explains why we look at the Lukasiewicz path only after time  $U(\mathfrak{t}_n)$  in (ii).

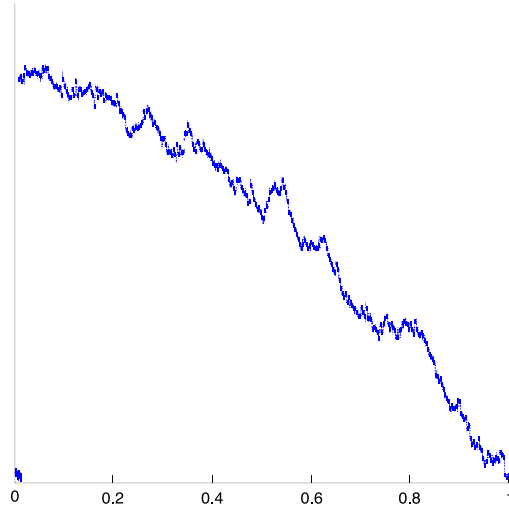


Fig. 4. A simulation of a Lukasiewicz path of a large non-generic tree.

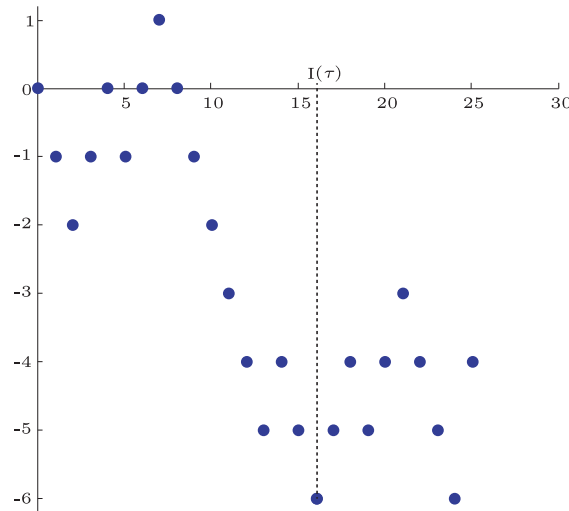


Fig. 5. The modified Lukasiewicz path  $\tilde{\mathcal{W}}(\tau)$  of the tree  $\tau$  appearing in Fig. 2. Here  $n = 26$ ,  $\Delta(\tau) = 4$ ,  $I(\tau) = 16$  and  $U(\tau) = 9$ .

2.2. Description of the Lukasiewicz path after removing the vertex of maximal degree

Recall that  $U(\tau)$  is the index in the lexicographical order of the first vertex of  $\tau$  with maximal out-degree. We first define a modified version  $\tilde{\mathcal{W}}(\tau)$  of the Lukasiewicz path as follows. Set  $n = |\tau|$ , and for  $1 \leq i \leq n - U(\tau) - 1$ , set  $\tilde{\mathcal{X}}_i(\tau) = \mathcal{W}_{U(\tau)+i+1}(\tau) - \mathcal{W}_{U(\tau)+i}(\tau)$  and for  $n - U(\tau) \leq i \leq n - 1$  set  $\tilde{\mathcal{X}}_i(\tau) = \mathcal{W}_{i+1-(n-U(\tau))}(\tau) - \mathcal{W}_{i-(n-U(\tau))}(\tau)$ . In other words,  $\tilde{\mathcal{X}}_1(\tau), \dots, \tilde{\mathcal{X}}_{n-1}(\tau)$  are the increments of  $\mathcal{W}(t_n)$ , shifted cyclically to start just after the maximum jump (which is not included). Then set  $\tilde{\mathcal{W}}_i(\tau) = \tilde{\mathcal{X}}_1(\tau) + \tilde{\mathcal{X}}_2(\tau) + \dots + \tilde{\mathcal{X}}_i(\tau)$  for  $0 \leq i \leq n - 1$  (see Fig. 5 for an example). Note that  $\Delta(\tau) = -\tilde{\mathcal{W}}_{n-1}(\tau)$ . Finally, set

$$I(\tau) = \min \left\{ i \in \{0, 1, \dots, n - 1\}; \tilde{\mathcal{W}}_i(\tau) = \min_{0 \leq j \leq n-1} \tilde{\mathcal{W}}_j(\tau) \right\}.$$

We now introduce some notation. For every  $1 \leq i \leq \Delta(\tau)$ , let  $\mathcal{T}_i(\tau)$  be the tree of descendants of the  $i$ th child of  $u_*(\tau)$ . Set also  $\mathcal{F}_{i,j}(\tau) = (\mathcal{T}_i(\tau), \dots, \mathcal{T}_j(\tau))$ . Finally, for  $1 \leq k \leq \Delta(\tau)$ , let  $\zeta_k(\tau) = \inf\{i \geq 0; \tilde{\mathcal{W}}_i(\tau) = -k\}$ . The following result explains the reason why we introduce  $\tilde{\mathcal{W}}(\tau)$ .

**Proposition 2.5.** *The following assertions hold.*

- (i) *We have  $U(\tau) = n - 1 - I(\tau)$ .*
- (ii) *For  $1 \leq k \leq \Delta(\tau)$ ,  $(\tilde{\mathcal{W}}_0, \tilde{\mathcal{W}}_1, \dots, \tilde{\mathcal{W}}_{\zeta_k(\tau)})$  is the Lukasiewicz path of the forest  $\mathcal{F}_{1,k}(\tau)$ .*
- (iii) *The vectors  $(\mathcal{W}_0(\tau), \mathcal{W}_1(\tau), \dots, \mathcal{W}_{U(\tau)}(\tau))$  and*

$$(\tilde{\mathcal{W}}_{I(\tau)}(\tau) - \tilde{\mathcal{W}}_{I(\tau)}(\tau), \tilde{\mathcal{W}}_{I(\tau)+1}(\tau) - \tilde{\mathcal{W}}_{I(\tau)}(\tau), \dots, \tilde{\mathcal{W}}_{n-1}(\tau) - \tilde{\mathcal{W}}_{I(\tau)}(\tau))$$

*are equal.*

This should be clear from the relation between  $\mathcal{W}$  and  $\tilde{\mathcal{W}}$  (see Fig. 5), and a formal proof would not be enlightening. We now prove that the random variables  $\tilde{\mathcal{X}}_i(t_n)$  are asymptotically independent.

**Proposition 2.6.** *We have:*

$$\sup_{A \in \mathcal{B}(\mathbb{R}^{n-1})} |\mathbb{P}((\tilde{\mathcal{X}}_1(t_n), \dots, \tilde{\mathcal{X}}_{n-1}) \in A) - \mathbb{P}((X_1, \dots, X_{n-1}) \in A)| \xrightarrow{n \rightarrow \infty} 0.$$

**Proof.** We keep the notation introduced in the proof of Proposition 2.2. By Proposition 1.7 and by the definition of  $\tilde{\mathcal{W}}(t_n)$ , we have

$$(\tilde{\mathcal{X}}_i(t_n); 0 \leq i \leq n - 1) \stackrel{(d)}{=} (X_{V_n+1}, X_{V_n+2}, \dots, X_n, X_1, \dots, X_{V_n-1}) \quad \text{under } \mathbb{P}(\cdot | W_n = -1).$$

On the event that  $(X_1, \dots, X_n)$  has a unique maximal component, we have

$$(X_{V_n+1}, X_{V_n+2}, \dots, X_n, X_1, \dots, X_{V_n-1}) \stackrel{(d)}{=} (X_1, X_2, \dots, X_{V_n-1}, X_{V_n+1}, \dots, X_n).$$

Indeed, since the distribution of  $(X_1, \dots, X_n)$  under  $\mathbb{P}(\cdot | W_n = -1)$  is cyclically exchangeable,  $V_n$  is uniformly distributed on the latter event. But we have already seen that under  $\mathbb{P}(\cdot | W_n = -1)$ ,  $(X_1, \dots, X_n)$  has a unique maximal component with probability tending to one as  $n \rightarrow \infty$ . The conclusion follows from Theorem 2.1.  $\square$

The following corollary will be useful.

**Corollary 2.7.** *Fix  $\eta \in (0, 1)$ . We have*

$$\sup_{A \in \mathcal{B}(\mathbb{T}_f^{[\lfloor \gamma \eta n \rfloor]})} |\mathbb{P}((\mathcal{T}_1(t_n), \dots, \mathcal{T}_{[\lfloor \gamma \eta n \rfloor]}(t_n)) \in A) - \mathbb{P}_\mu^{\otimes [\lfloor \gamma \eta n \rfloor]}(A)| \xrightarrow{n \rightarrow \infty} 0.$$

This means that, as  $n \rightarrow \infty$ , the random variables  $\mathcal{T}_1(t_n), \dots, \mathcal{T}_{[\lfloor \gamma \eta n \rfloor]}(t_n)$  are asymptotically independent  $\text{GW}_\mu$  trees.

**Proof.** Since  $\inf\{i \geq 0, W_i = -\lfloor \gamma \eta n \rfloor\}/n$  converges in probability towards  $\eta$  as  $n \rightarrow \infty$ , Proposition 2.6 entails that  $\tilde{\zeta}_{[\lfloor \gamma \eta n \rfloor]}(t_n)/n \rightarrow \eta$  in probability as  $n \rightarrow \infty$ . This implies that for every  $\varepsilon \in (0, 1 - \eta)$ , with probability tending to 1 as  $n \rightarrow \infty$ ,  $\mathcal{T}_1(t_n), \dots, \mathcal{T}_{[\lfloor \gamma \eta n \rfloor]}(t_n)$  only depends on  $\tilde{\mathcal{X}}_1(t_n), \dots, \tilde{\mathcal{X}}_{[\lfloor \gamma(\eta+\varepsilon)n \rfloor]}(t_n)$ . The conclusion immediately follows from Proposition 2.6.  $\square$

2.3. Location of the vertex with maximal out-degree

The main tool for studying the modified Lukasiewicz path  $\tilde{\mathcal{W}}$  is a time-reversal procedure which we now describe. For a sequence  $(a_i)_{i \geq 0}$  and for every integer  $n \geq 1$ , we let  $(a_i^{(n)})_{0 \leq i \leq n-1}$  be the sequence defined by  $a_i^{(n)} = a_{n-1} - a_{n-1-i}$ .

**Proof of Theorem 2(i).** For every tree  $\tau$ , writing  $\tilde{\mathcal{W}}$  instead of  $\tilde{\mathcal{W}}(\tau)$  to simplify notation, note that by Proposition 2.5(i) we have

$$U(\tau) = \max \left\{ 0 \leq k \leq n - 1; \tilde{\mathcal{W}}_k^{(n)} = \sup_{0 \leq j \leq n-1} \tilde{\mathcal{W}}_j^{(n)} \right\}.$$

It follows from Proposition 2.5(i) and Proposition 2.6 that, for every  $i \geq 0$ ,

$$\mathbb{P}(U(t_n) = i) - \mathbb{P} \left( i = \max \left\{ 0 \leq k \leq n - 1; W_k^{(n)} = \sup_{0 \leq j \leq n-1} W_j^{(n)} \right\} \right) \xrightarrow{n \rightarrow \infty} 0.$$

Since  $(W_i^{(n)})_{0 \leq i \leq n-1}$  and  $(W_i)_{0 \leq i \leq n-1}$  have the same distribution, we have

$$\max \left\{ 0 \leq k \leq n - 1; W_k^{(n)} = \sup_{0 \leq j \leq n-1} W_j^{(n)} \right\} \stackrel{(d)}{=} \max \left\{ 0 \leq k \leq n - 1; W_k = \sup_{0 \leq j \leq n-1} W_j \right\}.$$

In addition,  $W$  has a negative drift and tends almost surely to  $-\infty$ , hence

$$\mathbb{P} \left( i = \max \left\{ 0 \leq k \leq n - 1; W_k = \sup_{0 \leq j \leq n-1} W_j \right\} \right) \xrightarrow{n \rightarrow \infty} \mathbb{P} \left( i = \max \left\{ k \geq 0; W_k = \sup_{j \geq 0} W_j \right\} \right).$$

A simple argument based once again on time-reversal shows that the probability appearing in the right-hand side of the previous expression is equal to

$$\mathbb{P}(\forall m \leq i, W_m \geq 0) \cdot \mathbb{P}(\forall j \geq 1, W_j \leq -1).$$

Assertion (i) in Theorem 2 then follows from Proposition 1.5. □

**Remark 2.8.** One similarly shows that

$$n - 1 - \tilde{\zeta}_{\Delta(t_n)}(t_n) \xrightarrow[n \rightarrow \infty]{(d)} \sup \{i \geq 0; W_i = 0\}. \tag{11}$$

Indeed, since  $\Delta(t_n) = -\tilde{\mathcal{W}}_{n-1}(t_n)$ , we have  $n - 1 - \tilde{\zeta}_{\Delta(t_n)}(t_n) = \max \{0 \leq i \leq n - 1; \tilde{\mathcal{W}}_i^{(n)} = 0\}$ , and (11) follows by the same arguments as those in the proof of Theorem 2(i).

To prove the other assertions of Theorem 2, we will need the size-biased distribution associated with  $\mu$ , which is the distribution of the random variable  $\zeta^*$  such that:

$$\mathbb{P}(\zeta^* = k) := \frac{k\mu_k}{\mathfrak{m}}, \quad k = 0, 1, \dots$$

The following result concerning the local convergence of  $t_n$  as  $n \rightarrow \infty$  will be useful. We refer the reader to [17], Section 6, for definitions and background concerning local convergence of trees (note that we need to consider trees that are not locally finite, so that this is slightly different from the usual setting).

Let  $\hat{\mathcal{T}}$  be the infinite random tree constructed as follows. Start with a spine composed of a random number  $S$  of vertices, where  $S$  is defined by:

$$\mathbb{P}(S = i) = (1 - \mathfrak{m})\mathfrak{m}^{i-1}, \quad i = 1, 2, \dots \tag{12}$$

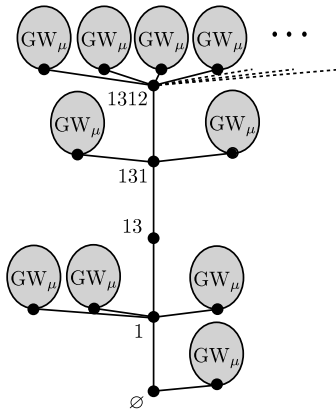


Fig. 6. An illustration of  $\widehat{\mathcal{T}}$ . Here, the spine is composed of the vertices  $\emptyset, 1, 13, 13, 131, 1312$ .

Then attach further branches as follows (see also Fig. 6). At the top of the spine, attach an infinite number of branches, each branch being a  $\text{GW}_\mu$  tree. At all the other vertices of the spine, a random number of branches distributed as  $\zeta^* - 1$  is attached to either to the left or to the right of the spine, each branch being a  $\text{GW}_\mu$  tree. At a vertex of the spine where  $k$  new branches are attached, the number of new branches attached to the left of the spine is uniformly distributed on  $\{0, \dots, k\}$ . Moreover all random choices are independent.

**Theorem 2.9 (Jonsson and Stefánsson [20], Janson [17]).** *The trees  $t_n$  converges locally in distribution toward  $\widehat{\mathcal{T}}$  as  $n \rightarrow \infty$ .*

**Proof of Theorem 2(ii) and (iii).** By Skorokhod’s representation theorem (see, e.g., [5], Theorem 6.7) we can suppose that the convergence  $t_n \rightarrow \widehat{\mathcal{T}}$  as  $n \rightarrow \infty$  holds almost surely for the local topology. Let  $u_\star \in \widehat{\mathcal{T}}$  be the vertex of the spine with largest generation. By (12), we have for  $i \geq 0$ :

$$\mathbb{P}(|u_\star| = i) = (1 - m)m^i. \tag{13}$$

Recall the notation  $U(t_n)$  for the index of  $u_\star(t_n)$ . Let  $\varepsilon > 0$ . By assertion (i) of Theorem 2, which was proved at the beginning of this section, we can fix an integer  $K$  such that, for every  $n$ ,  $\mathbb{P}(U(t_n) \leq K) > 1 - \varepsilon$ . From the local convergence of  $t_n$  to  $\widehat{\mathcal{T}}$  (and the properties of local convergence, see in particular Lemma 6.3 in [17]) we can easily verify that

$$\mathbb{P}(\{u_\star(t_n) \neq u_\star\} \cap \{U(t_n) \leq K\}) \xrightarrow{n \rightarrow \infty} 0.$$

We conclude that  $\mathbb{P}(u_\star(t_n) \neq u_\star) \rightarrow 0$  as  $n \rightarrow \infty$ . Assertion (ii) of Theorem 2 now follows from (13). □

Note that assertion (i) in Theorem 2 was needed to prove assertion (ii). Indeed, the local convergence of  $t_n$  toward  $\widehat{\mathcal{T}}$  would not have been sufficient to get that  $\mathbb{P}(u_\star(t_n) \neq u_\star) \rightarrow 0$ .

#### 2.4. Subtrees branching off the vertex with maximum out-degree

Before proving Theorem 3, we gather a few useful ingredients. It is well known that the mean number of vertices of a  $\text{GW}_\mu$  tree at generation  $n$  is  $m^n$ . As a consequence, we have  $\mathbb{E}_\mu[|\tau|] = 1 + m + m^2 + \dots = 1/(1 - m) = 1/\gamma$ . Moreover, for  $n \geq 1$ , by Kemperman’s formula (see, e.g., [29], Section 5):

$$\mathbb{P}_\mu(|\tau| = n) = \frac{1}{n} \mathbb{P}(W_n = -1) = \frac{1}{n} \mathbb{P}(\overline{W}_n = \gamma n - 1) \underset{n \rightarrow \infty}{\sim} \frac{\mathcal{L}(n)}{(\gamma n)^{1+\theta}}, \tag{14}$$

where we have used (1) for the last estimate. It follows that the total progeny of a  $\text{GW}_\mu$  tree belongs to the domain of attraction of a spectrally positive strictly stable law of index  $2 \wedge \theta$ . Hence we can find a slowly varying function  $L'$

such that the law of  $(\zeta(\mathbf{f}) - n/\gamma)/(L'(n)n^{1/(2\wedge\theta)})$  under  $\mathbb{P}_{\mu,n}$  converges as  $n \rightarrow \infty$  to the law of  $Y_1$ , where we recall that  $\mathbb{P}_{\mu,j}$  is the law of a forest of  $j$  independent  $\text{GW}_\mu$  trees. We set  $B'_n = L'(n)n^{1/(2\wedge\theta)}$ .

Let  $(Z_i)_{i \geq 0}$  be the random walk which starts at 0 and whose jump distribution has the same law as the total progeny of a  $\text{GW}_\mu$  tree. Note that  $\mathbb{P}(Z_j = k) = \mathbb{P}_{\mu,j}(\zeta(\mathbf{f}) = k)$ . Hence the distribution of  $Z_1$  belongs to the domain of attraction of a spectrally positive strictly stable law of index  $2 \wedge \theta$ . In particular, the following convergence holds in distribution in the space  $\mathbb{D}([0, 1], \mathbb{R})$ :

$$\left( \frac{Z_{\lfloor nt \rfloor} - nt/\gamma}{B'_n}, 0 \leq t \leq 1 \right) \xrightarrow[n \rightarrow \infty]{(d)} (Y_t, 0 \leq t \leq 1). \tag{15}$$

Finally, the following technical result establishes a useful link between  $B_n$  and  $B'_n$ .

**Lemma 2.10.** *We have  $B'_n/B_n \rightarrow 1/\gamma^{1+1/(2\wedge\theta)}$  as  $n \rightarrow \infty$ .*

The proof of Lemma 2.10 is postponed to the end of this section.

We are now ready to prove Theorem 3.

**Proof of Theorem 3.** We shall show that for every fixed  $\eta \in (0, 1)$ :

$$\left( \frac{Z_{\lfloor \Delta(t_n)t \rfloor}(t_n) - \Delta(t_n)t/\gamma}{B_n}, 0 \leq t \leq \eta \right) \xrightarrow[n \rightarrow \infty]{(d)} \left( \frac{1}{\gamma} Y_t, 0 \leq t \leq \eta \right). \tag{16}$$

Since  $\Delta(t_n)/(\gamma n) \rightarrow 1$  in probability as  $n \rightarrow \infty$  (Theorem 1(i)), recalling Lemma 2.10, the desired result will follow from a time-reversal argument since the vectors  $(\xi_1(t_n), \xi_2(t_n), \dots, \xi_{\Delta(t_n)}(t_n))$  and  $(\xi_{\Delta(t_n)}(t_n), \xi_{\Delta(t_n)-1}(t_n), \dots, \xi_1(t_n))$  have the same distribution. Tightness follows from the time-reversal argument and also continuity at  $t = 1$ .

Since  $\Delta(t_n)/(\gamma n) \rightarrow 1$  in probability as  $n \rightarrow \infty$ , by [25], Lemma 5.7, it is sufficient to establish that

$$\left( \frac{Z_{\lfloor \gamma n t \rfloor}(t_n) - nt}{B_n}, 0 \leq t \leq \eta \right) \xrightarrow[n \rightarrow \infty]{(d)} \left( \frac{1}{\gamma} Y_t, 0 \leq t \leq \eta \right). \tag{17}$$

To this end, note that  $B'_{\lfloor \gamma n \rfloor}/B_n \rightarrow 1/\gamma$  as  $n \rightarrow \infty$  and that  $(Z_1(t_n), Z_2(t_n), \dots, Z_{\lfloor \gamma n \eta \rfloor}(t_n))$  are asymptotically independent by Corollary 2.7. The conclusion immediately follows from (15) applied with  $\lfloor \gamma n \rfloor$  instead of  $n$ .  $\square$

We conclude this section by proving Lemma 2.10.

**Proof of Lemma 2.10.** Let  $\sigma^2$  be the variance of  $\mu$ . Note that  $\sigma^2 = \infty$  if  $\theta \in (1, 2)$ ,  $\sigma^2 < \infty$  if  $\theta > 2$  and that we can have either  $\sigma^2 = \infty$  or  $\sigma^2 < \infty$  for  $\theta = 2$ . When  $\sigma^2 = \infty$ , the desired result follows from classical results expressing  $B_n$  in terms of  $\mu$ . Indeed, in the case  $\theta < 2$ , we may choose  $B_n$  and  $B'_n$  such that (see, e.g., [25], Theorem 1.10):

$$\frac{B'_n}{B_n} = \frac{\inf\{x \geq 0; \mathbb{P}_\mu(|\tau| \geq x) \leq 1/n\}}{\inf\{x \geq 0; \mu([x, \infty)) \leq 1/n\}}.$$

Property (ii) in Assumption  $(H_\theta)$  and (14) entail that  $\mathbb{P}_\mu(|\tau| \geq x)/\mu([x, \infty)) \rightarrow 1/\gamma^{1+\theta}$  as  $x \rightarrow \infty$ . The result easily follows. The case when  $\sigma^2 = \infty$  and  $\theta = 2$  is treated by using similar arguments. We leave details to the reader.

We now concentrate on the case  $\sigma^2 < \infty$ . Note that necessarily  $\theta \geq 2$ . Let  $\sigma'^2$  be the variance of  $|\tau|$  under  $\mathbb{P}_\mu$  (from (14) this variance is finite when  $\sigma^2 < \infty$ ). We shall show that  $\sigma' = \sigma/\gamma^{3/2}$ . The desired result will then follow since we may take  $B_n = \sigma\sqrt{n/2}$  and  $B'_n = \sigma'\sqrt{n/2}$  by the classical central limit theorem. In order to calculate  $\sigma'^2$ , we introduce the Galton–Watson process  $(\mathfrak{Z}_i)_{i \geq 0}$  with offspring distribution  $\mu$  such that  $\mathfrak{Z}_0 = 1$ . Recall that  $\mathbb{E}[\mathfrak{Z}_i] = m^i$ . Then note that:

$$\sigma'^2 = \mathbb{E}_\mu[|\tau|^2] - \mathbb{E}_\mu[|\tau|]^2 = \mathbb{E} \left[ \left( \sum_{i=0}^{\infty} \mathfrak{Z}_i \right)^2 \right] - \frac{1}{\gamma^2}.$$



Since  $(\mathfrak{Z}_i/m^i)_{i \geq 0}$  is a martingale with respect to the filtration generated by  $(\mathfrak{Z}_i)_{i \geq 0}$ , we have  $\mathbb{E}[Z_i Z_j] = m^{j-i} \mathbb{E}[Z_i^2]$ . Also using the well-known fact that for  $i \geq 1$  the variance of  $\mathfrak{Z}_i$  is  $\sigma^2 m^{i-1} (m^i - 1)/(m - 1)$  (see, e.g., [3], Section 1.2), write:

$$\begin{aligned} \mathbb{E} \left[ \left( \sum_{i=0}^{\infty} \mathfrak{Z}_i \right)^2 \right] &= \sum_{i=0}^{\infty} \mathbb{E}[\mathfrak{Z}_i^2] + 2 \sum_{0 \leq i < j} m^{j-i} \mathbb{E}[\mathfrak{Z}_i^2] = \sum_{i=0}^{\infty} \mathbb{E}[\mathfrak{Z}_i^2] \left( 1 + \frac{2m}{1-m} \right) \\ &= \left( 1 + \sum_{i=1}^{\infty} \left( \frac{\sigma^2 m^{i-1} (m^i - 1)}{m - 1} + m^{2i} \right) \right) \left( 1 + \frac{2m}{1-m} \right) \\ &= \frac{\sigma^2}{\gamma^3} + \frac{1}{\gamma^2}. \end{aligned}$$

This entails  $\sigma' = \sigma/\gamma^{3/2}$  and the conclusion follows.  $\square$

**Proof of Corollary 1.** Recall that  $\overline{\Delta}(Z) = \sup_{0 < s < 1} (Z_s - Z_{s-})$  denotes the largest jump of  $Z \in \mathbb{D}([0, 1], \mathbb{R})$ . It follows from the continuity of  $Z \rightarrow \overline{\Delta}(Z)$  and Theorem 3 that:

$$\frac{1}{B_n} \max_{1 \leq i \leq \Delta(t_n)} \xi_i(t_n) \xrightarrow[n \rightarrow \infty]{(d)} \frac{1}{\gamma} \sup_{s \in (0, 1]} (Y_s - Y_{s-}).$$

If  $\theta \geq 2$ ,  $Y$  is continuous and the first assumption of Corollary 1 follows. If  $\theta < 2$ , the result easily follows from the fact that the Lévy measure of  $Y$  is  $\nu(dx) = \mathbb{1}_{\{x > 0\}} dx / (\Gamma(-\theta)x^{1+\theta})$ .  $\square$

## 2.5. Height of large conditioned non-generic trees

We now prove Theorem 4. If  $\mathbf{f} = (\tau_1, \dots, \tau_k)$  is a forest, its height  $\mathcal{H}(\mathbf{f})$  is by definition  $\max(\mathcal{H}(\tau_1), \dots, \mathcal{H}(\tau_k))$ . Recall that for  $1 \leq i \leq \Delta(\tau)$ , let  $\mathcal{T}_i(\tau)$  is the tree of descendants of the  $i$ th child of  $u_*(\tau)$  and that  $\mathcal{F}_{i,j}(\tau) = (\mathcal{T}_i(\tau), \dots, \mathcal{T}_j(\tau))$ .

**Proof of Theorem 4.** If  $\tau$  is a tree, let  $\mathcal{H}_*(\tau) = 1 + \mathcal{H}(\mathcal{F}_{1, \Delta(\tau)}(\tau))$  be the height of the subtree of descendants of  $u_*(\tau)$  in  $\tau$ . By Theorem 2(ii), the generation of  $u_*(t_n)$  converges in distribution. It is thus sufficient to establish that, if  $(\lambda_n)_{n \geq 1}$  of positive real numbers tending to infinity:

$$\mathbb{P} \left( \left| \mathcal{H}_*(t_n) - \frac{\ln(n)}{\ln(1/m)} \right| \leq \lambda_n \right) \xrightarrow[n \rightarrow \infty]{} 1. \quad (18)$$

To simplify notation, set  $\mathcal{H}_{i,j}^{(n)} = \mathcal{H}(\mathcal{F}_{i,j}(t_n))$  and  $p_n = \ln(n)/\ln(1/m) - \lambda_n$ . Let us first prove the lower bound, i.e.,  $\mathbb{P}(\mathcal{H}_*(t_n) \leq p_n) \rightarrow 0$  as  $n \rightarrow \infty$ . It is plain that  $\mathbb{P}(\mathcal{H}_*(t_n) \leq p_n) \leq \mathbb{P}(\mathcal{H}_{1, \lfloor \gamma n/2 \rfloor}^{(n)} \leq p_n)$ . In addition, by Corollary 2.7,

$$\mathbb{P}(\mathcal{H}_{1, \lfloor \gamma n/2 \rfloor}^{(n)} \leq p_n) - \mathbb{P}_{\mu, \lfloor \gamma n/2 \rfloor}(\mathcal{H}(\mathbf{f}) \leq p_n) \xrightarrow[n \rightarrow \infty]{} 0.$$

But

$$\mathbb{P}_{\mu, \lfloor \gamma n/2 \rfloor}(\mathcal{H}(\mathbf{f}) \leq p_n) = (1 - \mathbb{P}_{\mu}(\mathcal{H}(\tau) > p_n))^{\lfloor \gamma n/2 \rfloor}.$$

Since  $\mu$  satisfies Assumption  $(H_\theta)$ , we have  $\sum_{i \geq 1} i \ln(i) \mu_i < \infty$ . It follows from [14], Theorem 2, that there exists a constant  $c > 0$  such that:

$$\mathbb{P}_{\mu}(\mathcal{H}(\tau) > k) \underset{k \rightarrow \infty}{\sim} c \cdot m^k. \quad (19)$$

Hence  $\mathbb{P}_{\mu}(\mathcal{H}(\tau) > p_n) \sim c \cdot 1/(n \cdot m^{\lambda_n})$  as  $n \rightarrow \infty$ . Consequently  $\mathbb{P}_{\mu, \lfloor \gamma n/2 \rfloor}(\mathcal{H}(\mathbf{f}) \leq p_n)$  tends to 0 as  $n \rightarrow \infty$ , and the proof of the lower bound is complete.

Now set  $q_n = \ln(n)/\ln(1/m) + \lambda_n$ . The proof of the fact that  $\mathbb{P}(\mathcal{H}_\star(t_n) \geq q_n) \rightarrow 0$  as  $n \rightarrow \infty$  is similar and we only sketch the argument. Write:

$$\mathbb{P}(\mathcal{H}_\star(t_n) \geq q_n) \leq \mathbb{P}_\mu(\mathcal{H}_{1, \lfloor \Delta(t_n)/2 \rfloor}^{(n)} \geq q_n) + \mathbb{P}_\mu(\mathcal{H}_{\lfloor \Delta(t_n)/2 \rfloor + 1, \Delta(t_n)}^{(n)} \geq q_n).$$

Since  $\mathcal{H}_{\Delta(t_n)+1, \Delta(t_n)}^{(n)}$  has the same distribution as  $\mathcal{H}_{1, \Delta(t_n) - \lfloor \Delta(t_n)/2 \rfloor}^{(n)}$ , it suffices to show that the first term of the last sum tends to 0 as  $n \rightarrow \infty$ . By Theorem 1(i), we have  $\Delta(t_n)/2 \leq \lfloor 2\gamma n/3 \rfloor$  with probability tending to 1 as  $n \rightarrow \infty$ . It is thus sufficient to establish that  $\mathbb{P}_\mu(\mathcal{H}_{1, \lfloor 2\gamma n/3 \rfloor}^{(n)} \geq q_n) \rightarrow 0$  as  $n \rightarrow \infty$ . By arguments similar to those of the proof of the lower bound, it is enough to check that

$$\mathbb{P}_{\mu, \lfloor 2\gamma n/3 \rfloor}(\mathcal{H}(\mathbf{f}) \geq q_n) \xrightarrow[n \rightarrow \infty]{} 0.$$

This follows from (19), the fact that  $\mathbb{P}_{\mu, \lfloor 2\gamma n/3 \rfloor}(\mathcal{H}(\mathbf{f}) \geq q_n) = 1 - (1 - \mathbb{P}_\mu(\mathcal{H}(\tau) \geq q_n))^{\lfloor 2\gamma n/3 \rfloor}$  combined with the asymptotic behavior  $\mathbb{P}_\mu(\mathcal{H}(\tau) \geq q_n) \sim c \cdot m^{\lambda_n}/n$  as  $n \rightarrow \infty$ . This completes the proof of the upper bound and establishes (18).  $\square$

Theorem 4 implies that  $\mathcal{H}(t_n)/\ln(n) \rightarrow \ln(1/m)$  in probability as  $n \rightarrow \infty$ . We next show that this convergence holds in  $\mathbb{L}^p$  for every  $p \geq 1$ .

**Proposition 2.11.** *For every  $p \geq 1$ , we have*

$$\mathbb{E}[\mathcal{H}(t_n)^p] \underset{n \rightarrow \infty}{\sim} \frac{\ln(n)^p}{\ln(1/m)^p}.$$

**Proof.** The following proof is due to an anonymous referee. It is sufficient to show there exists  $K > 0$  such that

$$\mathbb{E}[\mathcal{H}(t_n)^p \mathbb{1}_{\{\mathcal{H}(t_n) > K \ln(n)\}}] \xrightarrow[n \rightarrow \infty]{} 0.$$

By (14), we have  $\mathbb{P}(|\tau| = n) \geq n^{-2-\theta}$  for  $n$  sufficiently large, and in addition by (19) we have  $\mathbb{P}(\mathcal{H}(\tau) > K \ln(n)) \leq 2c \cdot n^{-K \ln(1/m)}$  for  $n$  sufficiently large. Hence, bounding the height of  $\mathcal{H}(t_n)$  by  $n$ , we get

$$\begin{aligned} \mathbb{E}[\mathcal{H}(t_n)^p \mathbb{1}_{\{\mathcal{H}(t_n) > K \ln(n)\}}] &\leq n^p \cdot \mathbb{P}(\mathcal{H}(t_n) > K \ln(n)) \\ &\leq n^p \cdot \mathbb{P}(\mathcal{H}(\tau) > K \ln(n)) / \mathbb{P}(|\tau| = n) \\ &\leq 2c \cdot n^{p+2+\theta-K \ln(1/m)}. \end{aligned}$$

It thus suffices to choose  $K > (p + 2 + \theta)/\ln(1/m)$ . This completes the proof.  $\square$

### 2.6. Scaling limits of non-generic trees

We turn to the proof of Theorem 5.

**Proof of Theorem 5.** Fix  $\eta \in (0, 1/\ln(1/m))$ . We shall show that, with probability tending to one as  $n \rightarrow \infty$ , at least  $\ln(n)$  trees among the  $\lfloor \gamma n/2 \rfloor$  trees  $\mathcal{T}_1(t_n), \dots, \mathcal{T}_{\lfloor \gamma n/2 \rfloor}(t_n)$  have height at least  $\eta \ln(n)$ . This will indeed show that, with probability tending to one as  $n \rightarrow \infty$ , the number of balls of radius less than  $\eta$  needed to cover  $\ln(n)^{-1} \cdot t_n$  tends to infinity. By standard properties of the Gromov–Hausdorff topology (see [7], Proposition 7.4.12) this implies that the sequence of random metric spaces  $(\ln(n)^{-1} \cdot t_n)_{n \geq 1}$  is not tight.

If  $\mathbf{f} = (\tau_1, \dots, \tau_j)$  is a forest, let  $E_n(\mathbf{f})$  be the event defined by

$$E_n(\mathbf{f}) = \{\text{at most } \ln(n) \text{ trees among } \tau_1, \dots, \tau_j \text{ have height at least } \eta \ln(n)\}.$$

It is thus sufficient to prove that  $\mathbb{P}(\mathbf{E}_n(\mathcal{T}_1(t_n), \dots, \mathcal{T}_{\lfloor \gamma n/2 \rfloor}(t_n)))$  converges toward 0 as  $n \rightarrow \infty$ . As previously, by Corollary 2.7, it is sufficient to establish that

$$\mathbb{P}_{\mu, \lfloor \gamma n/2 \rfloor}(\mathbf{E}_n(\mathbf{f})) \xrightarrow{n \rightarrow \infty} 0. \quad (20)$$

Now denote by  $N_n$  the number of trees among a forest of  $\lfloor \gamma n/2 \rfloor$  independent  $\text{GW}_\mu$  trees of height at least  $\eta \ln(n)$ . Using (19) and setting  $\eta' = \eta \ln(1/m)$ , we get that for a certain constant  $C > 0$ ,  $N_n$  dominates a binomial random variable  $\text{Bin}(\lfloor \gamma n/2 \rfloor, Cn^{\eta'})$ , which easily implies that  $\mathbb{P}(N_n \leq \ln(n)) \rightarrow 0$  as  $n \rightarrow \infty$ . This shows (20) and completes the proof.  $\square$

Note that Theorem 5 implies that there is no nontrivial scaling limit for the contour function coding  $t_n$ , since convergence of scaled contour functions imply convergence in the Gromov–Hausdorff topology (see, e.g., [26], Lemma 2.3).

### 2.7. Finite dimensional marginals of the height function

We first extend the definition of the height function to a forest. If  $\mathbf{f} = (\tau_i)_{1 \leq i \leq j}$  is a forest, set  $n_0 = 0$  and  $n_p = |\tau_1| + |\tau_2| + \dots + |\tau_p|$  for  $1 \leq p \leq j$ . Then, for every  $0 \leq i \leq p-1$  and  $0 \leq k \leq |\tau_{i+1}|$ , set

$$\mathbf{H}_{n_i+k}(\mathbf{f}) = H_k(\tau_{i+1}).$$

Note that the excursions of  $\mathbf{H}(\mathbf{f})$  above 0 are the  $(\mathbf{H}_{n_i+k}(\mathbf{f}); 0 \leq k \leq |\tau_{i+1}|)$ . The Lukasiewicz path  $\mathcal{W}(\mathbf{f})$  and height function  $\mathbf{H}(\mathbf{f})$  satisfy the following relation (see, e.g., [26], Proposition 1.7, for a proof): For every  $0 \leq n \leq |\mathbf{f}|$ ,

$$\mathbf{H}_n(\mathbf{f}) = \text{Card}\left(\left\{k \in \{0, 1, \dots, n-1\}; \mathcal{W}_k(\mathbf{f}) = \inf_{k \leq j \leq n} \mathcal{W}_j(\mathbf{f})\right\}\right). \quad (21)$$

Recall that  $(W_n)_{n \geq 0}$  stands for the random walk introduced in Proposition 1.4 with  $\rho = \mu$  and that  $X_k = W_k - W_{k-1}$  for  $k \geq 1$ . For every  $n \geq 0$ , set

$$H_n = \text{Card}\left(\left\{0 \leq k \leq n-1; W_k = \inf_{k \leq j \leq n} W_j\right\}\right), \quad J_n = n - \min\left\{0 \leq i \leq n; W_i = \min_{0 \leq j \leq n} W_j\right\}.$$

Finally, for  $k \geq 0$ , set

$$M_k = \text{Card}\left(\left\{1 \leq i \leq k; W_i = \max_{0 \leq j \leq i} W_j\right\}\right), \quad T = \sup\left\{i \geq 0, W_i = \sup_{j \geq 0} W_j\right\}.$$

Since  $W$  drifts almost surely to  $-\infty$ ,  $T$  is almost surely finite, and  $M_T$  is distributed according to a geometric random variable of parameter  $\mathbb{P}(\forall i \geq 1, W_i \leq -1) = \gamma$ , by Proposition 1.5(i).

The following result, which is an unconditioned version of Theorem 6, will be useful.

**Lemma 2.12.** *For every  $0 < s < 1$ , the following convergence holds in distribution:*

$$(H_{\lfloor ns \rfloor}, H_n, J_n) \xrightarrow[n \rightarrow \infty]{(d)} (\mathbf{e}_1, M_T, T), \quad (22)$$

where  $\mathbf{e}_1$  is geometric random variable of parameter  $\gamma$ , independent of  $(X_n)_{n \geq 1}$ .

**Proof.** Set  $W_i^n = W_{n-\lfloor ns \rfloor+i} - W_{n-\lfloor ns \rfloor}$  for  $i \geq 0$ , and  $M_k^n = \text{Card}(\{1 \leq i \leq k; W_i^n = \max_{0 \leq j \leq i} W_j^n\})$ . Set also

$$T_n = \max\left\{i \in \{0, 1, \dots, n\}, W_i = \sup_{0 \leq j \leq n} W_j\right\}.$$

Notice that  $(W_i^n, i \geq 0)$  has the same distribution as  $(W_i, i \geq 0)$ . Using the fact that  $(W_i, 0 \leq i \leq n)$  and  $(W_n - W_{n-i}, 0 \leq i \leq n)$  have the same distribution, we get that

$$(H_{\lfloor ns \rfloor}, H_n, J_n) \stackrel{(d)}{=} (M_{\lfloor ns \rfloor}^n, M_n, T_n).$$

Let  $F_1 : \mathbb{Z} \rightarrow \mathbb{R}_+$ ,  $F_2 : \mathbb{Z}^2 \rightarrow \mathbb{R}_+$  be bounded functions and fix  $\varepsilon > 0$ . Choose  $N_0 > 0$  such that  $\mathbb{P}(T > N_0) < \varepsilon$ . For  $n \geq N_0$ , note that  $M_n = M_{N_0}$  and  $T_n = T_{N_0}$  on the event  $T \leq N_0$ . Hence for  $n \geq N_0$ :

$$|\mathbb{E}[F_1(M_{[ns]}^n)F_2(M_n, T_n)] - \mathbb{E}[F_1(M_{[ns]}^n)F_2(M_{N_0}, T_{N_0})]| \leq C\varepsilon,$$

where  $C > 0$  is a constant depending only on  $F_1, F_2$  (and which may change from line to line). Next, using the fact that  $M_{[ns]}^n$  is independent of  $(M_{N_0}, T_{N_0})$  for  $n > N_0/(1-s)$ , we get that for  $n > N_0/(1-s)$ ,

$$|\mathbb{E}[F_1(M_{[ns]}^n)F_2(M_n, T_n)] - \mathbb{E}[F_1(M_{[ns]}^n)]\mathbb{E}[F_2(M_{N_0}, T_{N_0})]| \leq C\varepsilon.$$

The conclusion immediately follows since  $M_{[ns]}^n$  has the same distribution as  $M_{[ns]}$  and since  $(M_n, T_n)$  converges in distribution toward  $(M_T, T)$  as  $n \rightarrow \infty$ .  $\square$

**Remark 2.13.** *It is straightforward to adapt the proof of Lemma 2.12 to get that for every  $0 < t_1 < t_2 < \dots < t_k < 1$  and  $b > 0$ , the following convergence holds in distribution:*

$$(H_{[nt_1-b]}, H_{[nt_2-b]}, \dots, H_{[nt_k-b]}, H_{n-1}, J_{n-1}) \xrightarrow[n \rightarrow \infty]{(d)} (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k, M_T, T),$$

where  $(\mathbf{e}_i)_{1 \leq i \leq k}$  are i.i.d. geometric random variables of parameter  $\gamma$ , independent of  $(M_T, T)$ .

Recall that for  $1 \leq i \leq \tilde{j} \leq \Delta(\tau)$ ,  $\mathcal{T}_i(\tau)$  is the tree of descendants of the  $i$ th child of  $u_*(\tau)$ , that  $\mathcal{F}_{i,j}(\tau) = (\mathcal{T}_i(\tau), \dots, \mathcal{T}_j(\tau))$  and that  $\zeta_k(\tau) = \inf\{i \geq 0; \tilde{\mathcal{W}}_i(\tau) = -k\}$  for  $1 \leq k \leq \Delta(\tau)$ . We are now ready to prove Theorem 6.

**Proof of Theorem 6.** To simplify, we establish Theorem 6 for  $k = 2$ , the general case being similar. To this end, we fix  $0 < s < t < 1$  and shall show that

$$(H_{[ns]}(t_n), H_{[nt]}(t_n)) \xrightarrow[n \rightarrow \infty]{(d)} (1 + \mathbf{e}_0 + \mathbf{e}_1, 1 + \mathbf{e}_0 + \mathbf{e}_2). \quad (23)$$

We first express  $H_{[ns]}(t_n)$  in terms of the modified Lukasiewicz path  $\tilde{\mathcal{W}}$  which was defined in Section 2.2. To this end we need to introduce some notation. For every tree  $\tau$  and  $0 \leq p \leq |\tau| - 1$ , set

$$\tilde{H}_p(\tau) = \text{Card}\left(\left\{k \in \{0, 1, \dots, p-1\}; \tilde{W}_k(\tau) = \inf_{k \leq j \leq p} \tilde{W}_j(\tau)\right\}\right).$$

Note that by Proposition 2.5(ii),  $(\tilde{H}_1(t_n), \dots, \tilde{H}_{\tilde{\zeta}_{\Delta(t_n)}(t_n)}(t_n))$  is the height function of the forest  $\mathcal{F}_{1, \Delta(t_n)}(t_n)$ . For every  $n \geq 1$  and  $r \in (0, 1)$  such that  $U(t_n) < [nr] < \tilde{\zeta}_{\Delta(t_n)}(t_n)$ , we have  $H_{[nr]}(t_n) = 1 + H_{[nr]-U(t_n)-1}(\mathcal{F}_{1, \Delta(t_n)}(t_n)) + |u_*(\tau)|$ . Hence, using Proposition 2.5(ii) and (21):

$$H_{[nr]}(t_n) = 1 + \tilde{H}_{[nr]-U(t_n)-1}(t_n) + \tilde{H}_{n-1}(t_n).$$

Since  $U(t_n)$  and  $\tilde{\zeta}_{\Delta(t_n)}(t_n)$  converge in distribution (by respectively Theorem 2 and Remark 2.8), we have  $U(t_n) < [ns] < [nt] < \tilde{\zeta}_{\Delta(t_n)}(t_n)$  with probability tending to 1 as  $n \rightarrow \infty$ . By combining Proposition 2.5(i) and Proposition 2.6, we get that:

$$\sup_{A \in \mathcal{B}(\mathbb{R}^2)} |\mathbb{P}((H_{[ns]}(t_n), H_{[nt]}(t_n)) \in A) - \mathbb{P}((1 + H_{[ns]-J_{n-1}-1] + H_{n-1}, 1 + H_{[nt]-J_{n-1}-1] + H_{n-1}) \in A)|$$

converges to 0 as  $n \rightarrow \infty$ . But by Remark 2.13, we have

$$(1 + H_{[ns]-J_{n-1}-1] + H_{n-1}, 1 + H_{[nt]-J_{n-1}-1] + H_{n-1}) \xrightarrow[n \rightarrow \infty]{(d)} (1 + \mathbf{e}_1 + M_T, 1 + \mathbf{e}_2 + M_T)$$

with  $M_T$  independent of  $\mathbf{e}_1, \mathbf{e}_2$ . Since  $M_T$  is distributed according to a geometric random variable of parameter  $\gamma = 1 - m$ , the conclusion immediately follows.  $\square$

### 3. Extensions and comments

We conclude by proposing possible extensions and stating a few open questions.

#### Other types of conditioning

Throughout this text, we have only considered the case of Galton–Watson trees conditioned on having a fixed total progeny. It is natural to consider different types of conditioning. For instance, for  $n \geq 1$ , let  $t_n^h$  be a random tree distributed according to  $\mathbb{P}_\mu(\cdot | \mathcal{H}(\tau) \geq n)$ . In [17], Section 22, Janson has in particular proved that when  $\mu$  is critical or subcritical, as  $n \rightarrow \infty$ ,  $t_n^h$  converges locally to Kesten’s Galton–Watson tree conditioned to survive  $\mathcal{T}^*$ , which a random infinite tree different from  $\widehat{\mathcal{T}}$ . It would be interesting to know whether the theorems of the present work apply in this case.

Another type of conditioning involving the number of leaves has been introduced in [9,25,30]. If  $\tau$  is a tree, denote by  $\lambda(\tau)$  the number of leaves of  $\tau$  (i.e., the number of individuals with no child). For  $n \geq 1$  such that  $\mathbb{P}_\mu(\lambda(\tau) = n) > 0$ , let  $t_n^l$  be a random tree distributed according to  $\mathbb{P}_\mu(\cdot | \lambda(\tau) = n)$ . Do results similar to those we have obtained hold when  $t_n$  is replaced by  $t_n^l$ ? We expect the answer to be positive, since a  $\text{GW}_\mu$  tree with  $n$  leaves is very close to a  $\text{GW}_\mu$  with total progeny  $n/\mu_0$  (see [25] for details), and we believe that the techniques of the present work can be adapted to solve this problem.

#### Concentration of $\mathcal{H}(t_n)$ around $\ln(n)/\ln(1/m)$

By Theorem 4, the sequence of random variable  $(\mathcal{H}(t_n) - \ln(n)/\ln(1/m))_{n \geq 1}$  is tight. It is therefore natural to ask the following question, due to Nicolas Broutin. Does there exist a random variable  $\mathcal{H}$  such that:

$$\mathcal{H}(t_n) - \frac{\ln(n)}{\ln(1/m)} \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{H}?$$

We expect the answer to be negative. Let us give a heuristic argument to support this prediction. In the proof of Theorem 4, we have seen that the height of  $\mathcal{H}(t_n)$  is close to the height of  $\lfloor \gamma n \rfloor$  independent  $\text{GW}_\mu$  trees and the height of each of these trees satisfies the estimate (19). However, if  $(Q_i)_{i \geq 1}$  is an i.i.d. sequence of random variables such that  $\mathbb{P}(Q_1 \geq k) = c \cdot m^k$ , then it is known (see, e.g., [16], Example 4.3) that the random variables

$$\max(Q_1, Q_2, \dots, Q_n) - \frac{\ln(n)}{\ln(1/m)}$$

do not converge in distribution.

#### Other types of trees

Janson [17] gives a very general limit theorem concerning the local asymptotic behavior of simply generated trees conditioned on having a fixed large number of vertices. Let us briefly recall the definition of simply generated trees. Fix a sequence  $\mathbf{w} = (w_k)_{k \geq 0}$  of nonnegative real numbers such that  $w_0 > 0$  and such that there exists  $k > 1$  with  $w_k > 0$  ( $\mathbf{w}$  is called a weight sequence). Let  $\mathbb{T}_f \subset \mathbb{T}$  be the set of all finite plane trees and, for every  $n \geq 1$ , let  $\mathbb{T}_n$  be the set of all plane trees with  $n$  vertices. For every  $\tau \in \mathbb{T}_f$ , define the weight  $w(\tau)$  of  $\tau$  by:

$$w(\tau) = \prod_{u \in \tau} w_{k_u(\tau)}.$$

Then for  $n \geq 1$  set

$$Z_n = \sum_{\tau \in \mathbb{T}_n} w(\tau).$$

For every  $n \geq 1$  such that  $Z_n \neq 0$ , let  $\mathcal{T}_n$  be a random tree taking values in  $\mathbb{T}_n$  such that for every  $\tau \in \mathbb{T}_n$ :

$$\mathbb{P}(\mathcal{T}_n = \tau) = \frac{w(\tau)}{Z_n}.$$

The random tree  $\mathcal{T}_n$  is said to be finitely generated. Galton–Watson trees conditioned on their total progeny are particular instances of simply generated trees. Conversely, if  $\mathcal{T}_n$  is as above, there exists an offspring distribution  $\mu$  such that  $\mathcal{T}_n$  has the same distribution as a  $\text{GW}_\mu$  tree conditioned on having  $n$  vertices if, and only if, the radius of convergence of  $\sum w_i z^i$  is positive (see [17], Section 8).

It would thus be interesting to find out if the theorems obtained in the present work for Galton–Watson trees can be extended to the setting of simply generated trees whose associated radius of convergence is 0. In the latter case, Janson [17] proved that  $\mathcal{T}_n$  converges locally as  $n \rightarrow \infty$  toward a deterministic tree consisting of a root vertex with an infinite number of leaves attached to it. We thus expect that the asymptotic properties derived in the present work will take a different form in this case. We hope to investigate this in future work.

## Acknowledgements

I am grateful to Nicolas Broutin for interesting remarks, to Grégory Miermont for a careful reading of a preliminary version of this work, to Olivier Hénard for Remark 2.3, to Pierre Bertin and Nicolas Curien for useful discussions and to Jean-François Le Gall and an anonymous referee for many useful comments.

## References

- [1] D. Aldous. The continuum random tree. III. *Ann. Probab.* **21** (1993) 248–289. [MR1207226](#)
- [2] I. Armendáriz and M. Loulakis. Conditional distribution of heavy tailed random variables on large deviations of their sum. *Stochastic Process. Appl.* **121** (2011) 1138–1147. [MR2775110](#)
- [3] K. B. Athreya and P. E. Ney. *Branching Processes. Die Grundlehren der Mathematischen Wissenschaften* **196**. Springer, New York, 1972. [MR0373040](#)
- [4] P. Bialas, Z. Burda and D. Johnston. Condensation in the backgammon model. *Nuclear Phys. B* **493** (1997) 505.
- [5] P. Billingsley. *Convergence of Probability Measures*, 2nd edition. *Wiley Series in Probability and Statistics: Probability and Statistics*. Wiley, New York, 1999. [MR1700749](#)
- [6] N. H. Bingham, C. M. Goldie and J. L. Teugels. *Regular Variation. Encyclopedia of Mathematics and its Applications* **27**. Cambridge Univ. Press, Cambridge, 1989. [MR1015093](#)
- [7] D. Burago, Y. Burago and S. Ivanov. *A Course in Metric Geometry. Graduate Studies in Mathematics* **33**. Amer. Math. Soc., Providence, RI, 2001. [MR1835418](#)
- [8] N. Curien and I. Kortchemski. Percolation on random triangulations and stable looptrees. Preprint. Available at [arXiv:1307:6818](#).
- [9] N. Curien and I. Kortchemski. Random non-crossing plane configurations: A conditioned Galton–Watson tree approach. *Random Structures Algorithms*. To appear, 2015.
- [10] D. Denisov, A. B. Dieker and V. Shneer. Large deviations for random walks under subexponentiality: The big-jump domain. *Ann. Probab.* **36** (2008) 1946–1991. [MR2440928](#)
- [11] T. Duquesne. A limit theorem for the contour process of conditioned Galton–Watson trees. *Ann. Probab.* **31** (2003) 996–1027. [MR1964956](#)
- [12] R. Durrett. Conditioned limit theorems for random walks with negative drift. *Z. Wahrsch. Verw. Gebiete* **52** (1980) 277–287. [MR0576888](#)
- [13] S. Großkinsky, G. M. Schütz and H. Spohn. Condensation in the zero range process: Stationary and dynamical properties. *J. Stat. Phys.* **113** (2003) 389–410. [MR2013129](#)
- [14] C. R. Heathcote, E. Seneta and D. Vere-Jones. A refinement of two theorems in the theory of branching processes. *Teor. Veroyatnost. i Primenen.* **12** (1967) 341–346. [MR0217889](#)
- [15] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*, 2nd edition. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **288**. Springer, Berlin, 2003. [MR1943877](#)
- [16] S. Janson. Rounding of continuous random variables and oscillatory asymptotics. *Ann. Probab.* **34** (2006) 1807–1826. [MR2271483](#)
- [17] S. Janson. Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. *Probab. Surv.* **9** (2012) 103–252. [MR2908619](#)
- [18] S. Janson and S. O. Stefánsson. Scaling limits of random planar maps with a unique large face. *Ann. Probab.* To appear, 2015. Available at [arXiv:1212.5072](#).
- [19] I. Jeon, P. March and B. Pittel. Size of the largest cluster under zero-range invariant measures. *Ann. Probab.* **28** (2000) 1162–1194. [MR1797308](#)
- [20] T. Jonsson and S. O. Stefánsson. Condensation in nongeneric trees. *J. Stat. Phys.* **142** (2011) 277–313. [MR2764126](#)
- [21] D. P. Kennedy. The Galton–Watson process conditioned on the total progeny. *J. Appl. Probab.* **12** (1975) 800–806. [MR0386042](#)

- [22] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. Henri Poincaré Probab. Stat.* **22** (1986) 425–487. [MR0871905](#)
- [23] H. Kesten and B. Pittel. A local limit theorem for the number of nodes, the height, and the number of final leaves in a critical branching process tree. *Random Structures Algorithms* **8** (1996) 243–299. [MR1603252](#)
- [24] I. Kortchemski. A simple proof of Duquesne’s theorem on contour processes of conditioned Galton–Watson trees. In *Séminaire de Probabilités XLV* 537–558. *Lecture Notes in Math.* **2078**. Springer, New York, 2013.
- [25] I. Kortchemski. Invariance principles for Galton–Watson trees conditioned on the number of leaves. *Stochastic Process. Appl.* **122** (2012) 3126–3172. [MR2946438](#)
- [26] J.-F. Le Gall. Random trees and applications. *Probab. Surveys* **2** (2005) 245–311. [MR2203728](#)
- [27] J.-F. Le Gall. Random real trees. *Ann. Fac. Sci. Toulouse Math. (6)* **15** (2006) 35–62. [MR2225746](#)
- [28] J.-F. Le Gall. Itô’s excursion theory and random trees. *Stochastic Process. Appl.* **120** (2010) 721–749. [MR2603061](#)
- [29] J. Pitman. *Combinatorial Stochastic Processes. Lectures from the 32nd Summer School on Probability Theory Held in Saint-Flour, July 7–24, 2002.* *Lecture Notes in Math.* **1875**. Springer, Berlin, 2006. [MR2245368](#)
- [30] D. Rizzolo. Scaling limits of Markov branching trees and Galton–Watson trees conditioned on the number of vertices with out-degree in a given set. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** (2) (2015) 512–532.
- [31] L. Takács. *Combinatorial Methods in the Theory of Stochastic Processes*. Robert E. Krieger Publishing Co., Huntington, NY, 1977. Reprint of the 1967 original. [MR0431313](#)